

ENDBERICHT

Institut für
Technologie- und
Regionalpolitik – InTeReg

RTG.1999.GF.001-02/03

*gefördert durch das Bundesministerium für Wissenschaft und Verkehr
im Rahmen der zweckgebundenen Förderung der Bundesregierung*

EvinA

***Evaluierung von innovativen Aktionen in der Technologie-,
Struktur- und Arbeitsmarktpolitik: Entwicklung von
Methoden, Indikatoren und „Good Practice“***

*Oliver Fritz, Karin Grasenick, Markus Gruber, Christian Hartmann,
Mirjam Novakovic, Wolfgang Polt, Dorothea Sturn,
Birgit Woitech, Klaus Zinöcker*

Wien – Graz, März 2001

INSTITUT FÜR TECHNOLOGIE- UND REGIONALPOLITIK

EvinA

Evaluierung von **innovativen Aktionen**
in der Technologie-, Struktur- und Arbeitsmarktpolitik:
Entwicklung von Methoden, Indikatoren und
"Good Practice"

Endbericht

gefördert durch das Bundesministerium für Wissenschaft und Verkehr
im Rahmen der zweckgebundenen Förderung der Bundesregierung

*Oliver Fritz, Karin Grasenick, Markus Gruber, Christian Hartmann,
Mirjam Novakovic, Wolfgang Polt, Dorothea Sturn,
Birgit Woitech, Klaus Zinöcker*

Inhaltsverzeichnis

A	EINLEITUNG	II
B	ÜBERBLICK ÜBER DAS PROJEKT EVINA	III
B.1	PROBLEMSTELLUNG.....	III
B.2	ZIELSETZUNG.....	IV
B.3	AUFBAU UND STRUKTUR DES BERICHTES.....	IV
1.	KAPITEL EINS Entwicklung und Einsatzpotential quantitativer und qualitativer Methoden KARIN GRASENICK	
2.	KAPITEL ZWEI Kosten Nutzen Analyse und Fiskalanalyse DOROTHEA STURN, MIRJAM NOVAKOVIC	
3.	KAPITEL DREI Evaluierung von Netzwerken: Bewertung von Lernen und Kooperieren im Rahmen einer clusterorientierten Technologiepolitik CHRISTIAN HARTMANN	
4.	KAPITEL VIER Evaluation im Rahmen der regionalen Strukturpolitik der Europäischen Union MARKUS GRUBER	
5.	KAPITEL FÜNF Die Evaluation von Arbeitsmarktpolitik BIRGIT WOITECH, KLAUS ZINÖCKER	

A EINLEITUNG

Das Projekt EvinA hatte zum Ziel, die Evaluierungspraxis in den zentralen Forschungsbereichen des InTeReg, nämlich der **Struktur-, Technologie- und Arbeitsmarktpolitik** kritisch zu durchleuchten und methodisch weiterzuentwickeln. Anknüpfend an internationale Erfahrungen standen dabei die folgenden Fragen im Vordergrund:

- **Welche Methoden** werden primär angewendet, um innovative Maßnahmen zu evaluieren? Wie sind diese in Hinblick auf ihre Aussagefähigkeit bezüglich der Programmwirkungen zu bewerten?
- Welche **Möglichkeiten der Quantifizierung** von Programmwirkungen gibt es und wo sind Grenzen der Quantifizierung? In welchem Zusammenhang sind sinnvoll quantitative Methoden, in welchem qualitative Methoden einsetzbar und wie ergänzen sich diese Methoden?
- Wie müssen **Programmziele und Evaluierungskriterien** aufeinander abgestimmt sein und wie der Dialog zwischen Programm-DesignerInnen, Programm-ManagerInnen und EvaluatorInnen aussehen, damit Programme realistisch eingeschätzt werden können und Lernen ermöglichen? Wie sieht dieser Zusammenhang dann aus, wenn die Zielsysteme komplex sind und mehr als einen Politikbereich betreffen?
- Wie sind Programmwirkungen abschätzbar, wenn **strategische, schwer zu quantifizierende Ziele** verfolgt werden? Wie beispielsweise lässt sich Lernen oder der Nutzen von Kooperationen (zwischen Unternehmen sowie zwischen Unternehmen und Forschungseinrichtungen) im Rahmen einer clusterorientierten Politik eruieren?

Die Ergebnisse aus diesen vergleichenden Analysen sind im vorliegenden Bericht zusammengefasst. Eine Dokumentation der diesen Analysen zugrundeliegenden Diskussionen und Arbeitsschritte findet sich im den Zwischenbericht (EvinA Zwischenbericht Jänner – Dezember 1999)

Die Zielsetzung dieses Endberichtes war es aber nicht nur, den Erkenntnisgewinn zu dokumentieren, sondern soweit als möglich Schlussfolgerungen für die Evaluierungspraxis in leitfadenartiger Form zu ziehen. Zielgruppen einer solchen Aufbereitung sind EvaluatorInnen, die nach methodischen Weiterentwicklungen ihrer Ansätze streben, zum anderen aber auch Projektträger, Programm-ManagerInnen und die öffentliche Hand, welche Evaluationen ausschreibt und Evaluationsergebnisse implementiert.

B ÜBERBLICK ÜBER DAS PROJEKT EVINA

B.1 Problemstellung

Öffentliche Programme sollen zeigen können, dass die Investitionen in sie lohnen. Evaluationen sollen PolitikerInnen, Programm-ManagerInnen und auch einer informierten Öffentlichkeit darüber Auskunft geben, wie Programme wirken.

Überall dort, wo **Programmziele strategischer Art** sind, etwa die Verbesserung der Kooperationsfähigkeit, die Steigerung der Wettbewerbsfähigkeit oder die Weiterentwicklung von Kompetenzen, kann dies eine schwierige Aufgabe sein. Maßnahmen in diesen Programmen unterscheiden sich hinsichtlich der Bewertung ihrer Wirksamkeit deutlich von Maßnahmen, die einen ganz bestimmten Zweck verfolgen, wie etwa einen Menschen auf den Mond zu bringen:

Therefore, the programs always relate to capacity building, to a knowledge base, and a skill base. They are therefore more difficult to evaluate, because the evaluator cannot check whether a man has been put on the moon, or whether the cancer death rate has gone down." (Rip, A: Implementation and Evaluation of Science and Technology Priorities and Programmes, Twente 1990)

Im Kontext von Programmen in der Struktur-, Technologie- und Arbeitsmarktpolitik kommt erschwerend hinzu, dass viele Programmwirkungen

- nicht unmittelbar auftreten,
- nicht nur die direkten Programmteilnehmer affizieren (man denke an Sekundäreffekte wie Spill-overs und Externalitäten, Multiplikatoreffekte, Linkage-Effekte, Crowding-out Effekte),
- und oft schwer von den Wirkungen anderer Programme oder Ereignisse zu trennen sind.

Daraus resultiert, dass **quantitative Evaluationen schwierig und echte Wirkungsanalysen selten** sind. Häufig ist man mit sehr vagen, schwer verallgemeinerbaren Ergebnissen konfrontiert oder mit allen möglichen Variationen von Pseudoquantifizierungen, wie beispielsweise der vielzitierte Forschungsmultiplikator, aber auch die Abschätzung quantitativer Wirkungen (z.B. wie viele Arbeitsplätze werden mit der Maßnahme geschaffen) in vielen Evaluationen von Strukturfondsprogrammen fällt unter diese Kategorie.

Auf der anderen Seite sind seriöse quantitative Methoden wie die Verwendung ökonomischer Modelle oder Kosten-Nutzen-Analysen sehr aufwendig, bei kleinen Interventionen nur begrenzt anwendbar, haben enorme Datenerfordernisse und hängen in ihren Ergebnissen häufig von einer Vielzahl zu treffender Annahmen ab.

Neben diesen methodischen Fragen ist ein weiteres Problem der Zusammenhang zwischen den **Zielen von Programmen** und den Kriterien für die Evaluation. Idealerweise sollten im Zuge der Programmerstellung bereits Vorstellungen darüber entwickelt werden, wie sich Erfolg oder Misserfolg des Programms ausdrückt. Entsprechend sind Programmziele sorgfältig zu formulieren, so dass unter EntscheidungsträgerInnen, Programm-ManagerInnen, EvaluatorenInnen und ProgrammteilnehmerInnen Klarheit und

Konsens herrscht, was tatsächlich mit dem Programm erreicht werden soll. Dies ist meist nicht der Fall, und bei multidimensionalen sowie strategischen Zielsetzungen auch schwierig.

Angesichts dieser vielschichtigen Problematik kann es nicht einen einzigen Weg zur Entwicklung einer "good practice" geben. Vielmehr sind Weiterentwicklungen auf verschiedenen Ebenen notwendig.

B.2 Zielsetzung

Das Projekt setzte sich zum Ziel, die Evaluierungspraxis in den drei Bereichen Struktur-, Technologie- und Arbeitsmarktpolitik kritisch zu durchleuchten und methodisch weiterzuentwickeln. Anknüpfend an internationale Erfahrungen stehen dabei die folgenden Fragen im Vordergrund:

- **Welche Methoden** werden primär angewendet, um innovative Maßnahmen zu evaluieren? Wie sind diese in Hinblick auf ihre Aussagefähigkeit bezüglich der Programmwirkungen zu bewerten? Welche Methoden stellen welche Anforderungen an die verfügbare Datenbasis und an ein begleitendes Monitoring? In welchem Zusammenhang können Indikatoren entwickelt werden, die relativ einfach zu entwickeln und dennoch methodisch sauber und aussagekräftig sind?
- Welche **Möglichkeiten der Quantifizierung** von Programmwirkungen gibt es und wo sind Grenzen der Quantifizierung? In welchem Zusammenhang sind sinnvoll quantitative Methoden, in welchem qualitative Methoden einsetzbar und wie ergänzen sich diese Methoden?
- Wie müssen **Programmziele und Evaluierungskriterien** aufeinander abgestimmt sein und wie der Dialog zwischen Programm-DesignerInnen, Programm-ManagerInnen und EvaluatorInnen aussehen, damit Programme realistisch eingeschätzt werden können und Lernen ermöglichen? Wie sieht dieser Zusammenhang dann aus, wenn die Zielsysteme komplex sind und mehr als einen Politikbereich betreffen?
- Wie sind Programmwirkungen abschätzbar, wenn **strategische, schwer zu quantifizierende Ziele** verfolgt werden? Wie beispielsweise lässt sich Lernen oder der Nutzen von Kooperationen (zwischen Unternehmen sowie zwischen Unternehmen und Forschungseinrichtungen) im Rahmen einer clusterorientierten Politik eruieren?

B.3 Aufbau und Struktur des Berichtes

Diesen Fragen wurden in folgenden Kapitel nachgegangen:

Das **ersten Kapitel** befasst sich mit der grundlegenden Frage nach den jeweiligen Einsatzmöglichkeiten und Grenzen von quantitativen und qualitativen Methoden. Ein klares Bild über diese Möglichkeiten und Grenzen ist essentiell für die Auswahl bzw. den Mix von Evaluationsmethoden und die Bildung realistischer Erwartungen über deren mögliche Resultate.

Im **zweiten Kapitel** werden zwei, für die Evaluierung von Technologiepolitik zunehmend wichtiger werdende Fragen behandelt: Ansätze zur Kosten-Nutzen-Analyse und der

Fiskalanalyse. Bei ersterer wird gefragt, inwieweit die gesamten gesellschaftlichen Kosten und Erträge einer politischen Intervention erfasst und bewertet werden können. Bei letzterer werden die Auswirkungen der Maßnahme auf die öffentlichen Haushalte zu eruieren versucht. Das hier angeführte Beispiel einer Fiskalanalyse aus dem Bereich Arbeitsmarktpolitik zeigt die mögliche wechselseitige Befruchtung der verschiedenen Gebiete.

Im **dritten Kapitel** wird ein weiteres zentrales Thema für die Evaluierung technologiepolitischer Programme aufgerissen: Technologiepolitik hat in den letzten Jahren zunehmend auf Programme und Initiativen gesetzt, die Kooperationen und Netzwerke zwischen verschiedenen Akteuren fördern. Die Evaluierung solcher Programme ist methodisch anspruchsvoll, weil ein wesentlicher Teil der erhofften Effekte ‚intangibel‘ ist und in der Verbesserung der Kooperationsfähigkeit, der Erhöhten Kompetenz zur Absorption externen Wissen etc. liegt.

Im **vierten Kapitel** wird die Evaluation der regionalen Strukturfonds untersucht. Dabei werden vor dem Erfahrungshintergrund der Evaluation dieser Fonds in der EU bzw. in einzelnen Mitgliedsländern Indikatoren zusammengestellt, die die Kernindikatoren solcher Evaluationen ausmachen und versucht, aus den internationalen Beispielen „good practice“ herauszudestillieren.

Im **fünften Kapitel** schließlich, das sich mit der Evaluierung von arbeitsmarktpolitischen Maßnahmen befasst, wird auf der Basis jüngerer Evaluierungen von Programmen des AMS und des WAFF eine ‚idealtypische Evaluierung der Arbeitsmarktpolitik‘ entworfen. Diese nimmt die Fragestellung des ersten Kapitels wieder auf und skizziert einen idealtypischen Methoden-Mix von quantitativen und qualitativen Ansätzen.

Für die Behandlung der leitenden Fragestellungen wurden Querschnittsthemen formuliert, die in den einzelnen Kapiteln anzusprechen waren. Damit sollte erreicht werden, das im Institut breit vorhandene Evaluations-Know-how dort zu vertiefen, wo es über das spezifische Thema hinaus methodische Fragestellungen gibt, die von allgemeinem Interesse sind.

Die folgende Tabelle zeigt einen Überblick und ordnet die Themen inhaltlich ein.

Thema	Schwerpunkt inhaltlich	Politik-Design: Zusammenhang Zielformulierung und Kriterien der Evaluation	Methoden der Evaluation: Quantitative und qualitative Methoden	spezifische Fragen der Bewertung
1. Entwicklung und Einsatzpotential von qualitativen und quantitativen Methoden	Übergreifend		XXX	
2. Spezifische Probleme bei der Evaluation technologiepolitischer Programme und Maßnahmen	Technologiepolitik			
a) Kosten-Nutzen-Analyse / Fiskalanalyse	Technologiepolitik / Arbeitsmarktpolitik		XXX	XXX
b) Bewertung von Netzwerk/Clusterorientierter Technologiepolitik	Technologiepolitik / Regionalpolitik	XX	X	XX
3. Methoden und Indikatoren für die Evaluierung von Strukturfonds-Programmen	Regionalpolitik	X	XXX	XX
4. Methoden der Evaluation arbeitsmarktpolitischer Programme	Arbeitsmarktpolitik	XXX	XXX	X

1 KAPITEL EINS ENTWICKLUNG UND EINSATZ POTENTIAL QUANTITATIVER UND QUALITATIVER METHODEN

Karin Grasenick

1 KAPITEL EINS	
ENTWICKLUNG UND EINSATZ POTENTIAL QUANTITATIVER UND QUALITATIVER	
METHODEN	1
1.1 EVALUATION UND DIE FRAGE DES METHODISCHEN VORGEHENS	1
1.1.1 <i>Die naturwissenschaftliche, analytisch-nomologische Position</i>	3
1.1.2 <i>Die geisteswissenschaftlich hermeneutisch-interpretative Position</i>	5
1.1.3 <i>Natur- und geisteswissenschaftliche Begriffspaare</i>	7
1.2 DER POSITIVISMUSSTREIT UND SEINE FOLGEN	9
1.2.1 <i>die Kritik an der quantitativen Sozialforschung</i>	9
1.2.2 <i>Grundlagen interpretativer Verfahren</i>	11
1.3 QUALITATIVE VERFAHREN IM SPRACHGEBRAUCH.....	12
1.3.1 <i>(quantitative) Analyse qualitativer Daten</i>	12
1.3.2 <i>explorative Voruntersuchung zur Generierung von Hypothesen</i>	12
1.3.3 <i>sinnverstehende Vorgehen</i>	13
1.4 DIE BEDEUTUNG DER INTERPRETATIVEN SOZIALFORSCHUNG FÜR DIE EVALUATION	15
1.5 DIE VERBINDUNG QUANTITATIVER UND QUALITATIVER METHODEN	15
1.6 FRAGENKATALOG ZUM EVALUATIONSDESIGN.....	18
1.6.1 <i>Fragestellungen zu Rahmenbedingungen und Evaluationsgegenstand</i>	18
1.6.2 <i>Prinzipielle Vorgangsweise für die Design-Entwicklung</i>	19
1.6.3 <i>Fragen zur Klärung der Methoden</i>	19
1.6.4 <i>Checkliste für Erhebungsmethoden</i>	20
1.7 LITERATUR	21

1.1 Evaluation und die Frage des methodischen Vorgehens

Evaluation ist ein in unterschiedlichen Bereichen eingesetztes Instrument, welche als die "explizite Verwendung wissenschaftlicher Forschungsmethoden und -techniken für den Zweck der Durchführung einer Bewertung" definiert werden kann (Wittmann, 1985, S3). Je nach Zielsetzungen und zeitlichem Ablauf lassen sich entsprechende genauere Gliederungen vornehmen.

Wesentliche Fragestellungen, welche im Rahmen einer Evaluierung unabhängig von der genauen Ausrichtung zu klären sind, lassen sich dabei grundsätzlich folgendermaßen gliedern (vgl. Sommerlad 1992):

1. Konzeptuelle Fragen: Aufbau, Modell, Rahmenbedingungen des Evaluationsgegenstandes sowie Positionen und Interessen der beteiligten Akteure
2. Technische Fragen: Methodische Vorgangsweise, Einheiten der Analyse, Datenerhebung und Qualitätskontrolle, Art des Berichtes und Zielgruppe.
3. Operationale Fragen: Personen, die für Durchführung der Evaluation verantwortlich und beteiligt sind sowie zeitliche Gliederungen.

Handbücher bieten für diese Fragestellungen Konzepte für Vorgehensweisen und potentielle Methoden und Instrumente an, zumeist für spezifische Evaluierungsfelder (z.B. Programmevaluierung, Evaluierung der Arbeitsmarkt-, Gesundheitspolitik, etc.) aufbereitet. Die methodischen Empfehlungen bieten dabei sowohl qualitative als auch quantitative Ansätze im Sinne einer Toolbox an, wobei potentielle Einsatzfelder beschrieben werden. Diese Unterschiede werden in Handbüchern meistens auf der Ebene der Methoden im Rahmen ein und derselben potentiellen Toolbox abgehandelt, wobei die zugrundeliegenden Paradigmen und Annahmen nicht reflektiert werden. Die Auswahl wird jedoch einerseits direkt mit den Präferenzen der am Entscheidungsprozess beteiligten Personen beeinflusst, sollte aber auch durch die Fragestellungen des Evaluationsgegenstandes begründet sein.

Dabei haben insbesondere AuftraggeberInnen aber auch die Beforschten häufig eine Vorstellung davon, wie eine Evaluation bzw. geeignete Methode auszusehen hat und erwarten ein entsprechendes Verhalten der Evaluatoren, wodurch bestimmte Vorgaben bereits festgelegt sind. Ein Grundwissen aber auch intensivere Vertrautheit mit speziellen Methoden, kann dabei durch Evaluatoren vorgeschlagene Abweichungen von dieser gelernten Norm verhindern.

Die Evaluatoren wiederum haben jeweils eine bestimmte Ausbildung genossen und daraus abgeleitet nicht nur sehr unterschiedliches methodisches Know-how sondern auch eine spezifischen Zugang zu Fragestellungen welche letztendlich die vorgeschlagenen Vorgangsweise festlegen.

Grundsätzlich kann ein funktionalistischer von einem interpretativen Zugang unterschieden werden (nach Burrell und Morgan, 1979). Bei den pragmatischen Überlegungen darüber, welche Ansätze bessere Ergebnisse liefern könnten, wird dabei entweder auf quantitative Daten und entsprechende statistische Verfahren zurückgegriffen oder aber auf Beobachtungen und unterschiedliche Formen von Interviews und Textanalysen im Rahmen eines interpretativen Zugangs, wobei zumeist die Wahrnehmungen und Meinungen der betroffenen Akteure miteinbezogen werden. Tendenziell ist damit im ersten Fall eine eher an den Zielsetzungen der Programmdesigner, im zweiten Fall aber eine an den Erfahrungen der Zielgruppen ausgerichtete Perspektive verbunden.

Prinzipiell kann jedoch für jede Perspektive und Art der Evaluation zwischen beiden Zugängen gewählt werden, wenn auch die Art der Fragestellung entscheidend für die Wahl sein sollte. Diese sind aus den Zielsetzungen der Evaluation zu entwickeln. Durch die starke Verbindung von Methodenwahl mit Personen sowie deren Erkenntnisinteressen kommt es jedoch zu grundlegenden Einschränkungen. Eine weitere Verengung des Verständnisses resultiert aus der Dominanz sozialwissenschaftlicher Forschungstechniken. In der Literatur wird Evaluation als Bestandteil der Sozialwissenschaften betrachtet (vgl. Rossi et al., 1988).

Auch wenn Evaluationen breiter aufgefasst werden kann, welche Tätigkeitsbereich wie Management, Beratung, Konfliktregelung etc. umfasst, so kann doch festgestellt werden, dass die sozialwissenschaftliche Zuordnung sich als eine einfache Gemeinsamkeit der Untersuchungsgegenstände begründen lässt: dass sie durch menschliches Handeln entstanden sind und auch auf eine Beeinflussung desselben abzielen. Viele der verwendeten Ansätze haben ihren Ursprung daher in den Sozialwissenschaften und sind somit von der dortige Diskussion über die „richtige Art der Forschung“ stark mitbeeinflusst worden.

Bei der heute durchaus üblichen Verwendung sogenannter Toolboxen, welche keine scharfe Trennung sondern im Gegenteil eine Mischung von Instrumenten unterschiedlichster Herkunft darstellen, mag eine fundierte Auseinandersetzung mit der Entstehungsgeschichte und vor allem mit den daraus abgeleiteten unterschiedlichen Auffassungen über Art des Forschungsgegenstandes und angemessener Herangehensweise unnötig erscheinen. Tatsache ist, dass die Kenntnis ebendieser Grundannahmen wesentlich für die Wahl des richtigen Verfahrens und der gelungenen Umsetzung sind. Wo diese nicht gegeben ist, werden dafür häufig Methoden tatsächlich ohne ausreichendem Wissen über die notwendigen Voraussetzungen mit angewandt, wodurch die Ergebnisse, insbesondere durch ein mangelndes ineinander greifen der Teilaspekte hinter den Erwartungen zurückbleiben können.

In weiterer Folge sollen daher die den in quantitativen und qualitativen Methoden zugrundeliegenden Paradigmen erläutert werden. Dabei wird zum besseren Verständnis kurz auf die historische Dimension eingegangen und auf die aus der Diskussion der Sozialwissenschaften entstandene Kritik am quantitativen Vorgehen. Die quantitative Sozialforschung orientiert sich am Ideal der Naturwissenschaften, während der

methodologische Gegenstrom dem interpretativen Paradigma verpflichtet ist. Hierunter werden Theorieansätze verstanden, welche die Analyse des Sinnverstehens von Handlungen und die Rolle der Kommunikation in der Forschung betonen und die mit der naturwissenschaftlichen Orientierung nicht verträglich sind. Wesentliche Grundannahme ist eine Abgrenzung der Sozial- von den Naturwissenschaften, die erforderlich ist, weil in den Sozialwissenschaften "Gesetze" nach dem Muster der Naturwissenschaft nicht aufgestellt werden können, und weil "Handeln", das nicht durchgängig als rational geordnet betrachtet werden kann, kaum in kausalen Zusammenhängen erklärt werden kann (vgl. Hoffmann-Riem 1980). Nach dem interpretativen Paradigma wird gesellschaftliche Wirklichkeit nicht als objektiv gegeben angenommen, sondern - über Bedeutungszuschreibungen - als gesellschaftlich konstruiert. Individuellem Handeln liegt ein Sinn zugrunde, der sich nicht aus der Beobachtung allein erschließen lässt, sondern der in der Regel über Sprache vermittelt werden muss. Dieser methodologische Gegenstrom kann über die Frankfurter Schule (Kritische Theorie/Positivismusstreit/Adorno u.a.) und die verstehende Soziologie auf eine aristotelische Tradition zurückgeführt werden, d.h. auf einen verstehenden Erklärungstyp. Der Hauptstrom dagegen kann auf eine galileische Tradition verweisen, die kausale Erklärungsmodelle verfolgt und sich auf positivistische und kritisch rationalistische Theorien stützt. Haupt- und Gegenstrom haben also beide eine lange und bewegte Vergangenheit, wechseln sich in ihrer Bedeutung gelegentlich ab oder bestehen auch mehr oder weniger gleichberechtigt nebeneinander. Qualitative Methoden mit der Betonung des Datengewinns (gegenüber der Betonung der Datenanalyse beim Hauptstrom) sind also auf diesem Hintergrund als Methoden zu verstehen, die den kommunikativen und interaktiven Aspekt im Forschungsprozess betonen und ihm durch verschiedene Verfahrensweisen gerecht zu werden versuchen (Wilson 1982).

Grundsätzliche Streitpunkte sind/waren vor allem:

- Inwieweit kann man naturwissenschaftliches Vorgehen auf soziale Sachverhalte übertragen ("Atome reagieren, Menschen handeln sinnhaft")?
- Inwieweit kann man Menschen oder soziale Phänomene aus ihrer Einbettung isolieren kann, ohne den Gesamtzusammenhang zu zerstören? Ist das Ganze mehr als die Summe seiner Teile?
- Ist Verstehen oder nur Erklären Ziel der Wissenschaft?

Außerdem ordnen kritische Rationalisten oft qualitatives Vorgehen in den "Entdeckungszusammenhang" und damit in den vorwissenschaftlichen Bereich ein, ohne jedoch ihre Theorie- und Hypothesenbildung ausreichend zu reflektieren und zu explizieren.

1.1.1 Die naturwissenschaftliche, analytisch-nomologische Position

Quantitative Verfahren in den Sozialwissenschaften haben ihren Ursprung im "Positivismus" von Auguste Comte, welcher u.a. durch Emile Durkheim sowie später Karl Popper und Hans Albert aufgegriffen wurde.

Hans Albert lehnt eine Dualität von Sozial- und Naturwissenschaften ab und ist um eine Einführung quantitativer Methoden bzw. eine analytisch-nomologische Position bemüht. Die geisteswissenschaftlichen Verfahren müssen abgelegt werden, damit die Sozialwissenschaften einen positivistischen Stil erhalten, der ihnen wahre Gesetzesaussagen ermöglicht. Wissenschaft sei eine rein sachliche Angelegenheit, die zwar auf "sozialer Aktivität" beruht und ein "humanes Unternehmen" ist, nichts desto trotz aber der Rationalität unterliegt. Wissenschaft sei keine "Offenbarung des Glaubens, der Vernunft oder der Sinne".

Der Kritische Rationalismus enthält als methodologische Leitmotiv das Prinzip der Falsifikation. Theorien sollen ständigen Falsifikationsversuchen ausgesetzt werden, da keine Theorie beweisbar ist. Die empirische Basis der objektiven Wissenschaft ist also nichts 'Absolutes'. Der Schluss von einzelnen Erfahrungen auf allgemeine Gesetze ist für Popper nicht zulässig. Bei noch so vielen Beobachtungen gleicher Art könne man nicht sicher auf eine Regelmäßigkeit schließen, wogegen man eine vermeintliche Regelmäßigkeit mit einer einzigen Unregelmäßigkeit umstürzen kann.¹

- Folgende Charakteristika kennzeichnen dabei einen "positivistischen Stil":
- Ablehnung der "inneren Erfahrung", wissenschaftliche Erkenntnis muss auf beobachtbaren Tatsachen basieren
- Positives Denken untersucht Gesetzmäßigkeiten, die die Phänomene regieren, fragt aber nicht nach den letzten Ursachen der Phänomene oder ihrem eigentlichen Sinn. Das "Wie" wird untersucht, nicht das "Warum".
- Von Interesse ist die Verknüpfung der Gesetze der Phänomene und ihre Abfolge (ermöglicht Theoriebildung).
- Wirklichkeit ist nie absolut entschlüsselbar (verifizierbar), Erkenntnis nähert sich ihr nur relativ an.
- Soziale Tatsachen bestehen unabhängig von dem erfahrenden und erkennenden Subjekt, Einflüsse auf den Gegenstand lassen sich ausschalten bzw. ohne Schaden ignorieren.
- Soziale Wirklichkeit ist wie unbelebte Natur zu behandeln und zu analysieren ("soziale Physik").

Der kritische Rationalismus sucht im quantitativen Vorgehen nach Gesetzmäßigkeiten. In Anlehnung an die Naturwissenschaften (Physiker, Chemie) werden dafür hochgradig standardisierte Methoden, und genau vorgeschriebene Vorgangsweisen, um die Ergebnisse nachvollziehbar zu machen. Man geht also immer von Hypothesen oder Theorien aus, die auf konkrete Sachverhalte der sozialen Realität angewendet oder "übersetzt" werden (Deduktion). Aus der allgemeinen Hypothesen werden überprüfbare

¹ Vgl. das bekannte Beispiel der Schwäne: Die Behauptung alle Schwäne wären weiß kann nicht durch noch so viele beobachtete weiße Schwäne zum Gesetz werden, denn es könnte zu irgendeiner Zeit und einem Ort einen Schwan geben der nicht weiß ist. Dagegen reicht ein einziger schwarzer Schwan aus, um die Behauptung zu widerlegen.

singuläre Aussagen, sog. Basissätze, abgeleitet. Wichtig dabei: Es sollten am besten solche Basissätze abgeleitet werden, die am wahrscheinlichsten die vorausgesetzte Theorie widerlegen. Mit der Wahl von Indikatoren werden die Bestandteile der Basissätze (Variablen) operationalisiert, einer Messung zugänglich gemacht. Anhand statistischer Methoden werden die postulierten Beziehungen zwischen den Variablen überprüft.

1.1.2 Die geisteswissenschaftlich hermeneutisch-interpretative Position

Auch qualitative Verfahren können auf eine lange Tradition verweisen, welche bereits bei John Stuart Mill (einer der Väter der Wirtschaftswissenschaft) mit der Unterscheidung von "natural sciences" (den Wissenschaften der unbelebten Natur) und "moral sciences" (den Sozialwissenschaften) nachvollzogen werden kann.

Die Methoden der Sozialforschung, die sich unter diesem Paradigma entwickelt haben und die üblicherweise unter der Kennzeichnung qualitative Verfahren zusammengefasst werden, entstanden aus und in der Auseinandersetzung mit den traditionellen quantifizierenden Verfahren. Bis Mitte der 60er Jahre wurde in einem weitgehenden Konsens ihre Qualität vielfach daran gemessen, inwieweit sie sich einem von den Naturwissenschaften übernommenen Modell anzunähern vermochten.

Gegen die hinter diesen Modellen stehenden wissenschaftstheoretischen Grundpositionen wurden insbesondere im sogenannten "Positivismusstreit" u. a. von Vertretern der sogenannten Frankfurter Schule (Adorno, Habermas; Luhmann) Kritik formuliert. Diese Kritikpunkte gaben den Anstoß für die Entwicklung von neuen Konzepten der empirischen Sozialforschung in den letzten 40 Jahren. Beeinflusst von phänomenologischen, hermeneutischen sowie sprachphilosophischen Positionen innerhalb der Geistes- und Sozialwissenschaften hat sich das interpretative Paradigma seitdem zu einer eigenständigen Richtung innerhalb der empirischen Sozialwissenschaft entwickelt.

Die Hermeneutik Hans-Georg Gadamers, die verstehende Soziologie Max Webers, die phänomenologisch ausgerichtete Soziologie von Alfred Schütz, aber auch die Arbeiten des symbolischen Interaktionismus und - noch jüngeren Datums - der Ethnomethodologie waren und sind für diese Richtung der empirischen Sozialwissenschaft von großer Bedeutung.

- Die zentrale Grundannahmen lassen sich hierbei wie folgt zusammenfassen:
- Soziale Wirklichkeit ist schon immer durch die in ihr handelnden Menschen vorinterpretiert.
- Uninterpretierte Erfahrungen gibt es nicht, also auch keine subjektunabhängig analysierbare Realität
- Soziale Phänomene werden durch die Handlungen von Menschen geschaffen, die diese Handlungen als sinnvoll erleben.
- Diese Vorinterpretiertheit des Untersuchungsgegenstandes gilt es systematisch berücksichtigen

- Der Handelnde also, den es nicht nur zu beobachten, sondern auch zu verstehen gilt, ist nicht passiv und einflusslos den sozialen Tatsachen ausgesetzt, die seinem Willen widerstreben, vielmehr konstituiert er aktiv den Erkenntnisgegenstand, indem er handelt. Diesen Vorgang der Bedeutungsverleihung gilt es nachzuvollziehen, um verstehend deuten zu können.

Die Vertreter des Interpretativen Paradigmas lehnen die kritisch-rationale "atomistische" Betrachtung des Menschen und seiner Sozialität ab, weil sie meinen, dass eine solche Sichtweise dem Menschen in seiner Komplexität, im Netz seiner sozialen Beziehungen, seiner psychischen Verfasstheit, seinen situativen Umständen nicht gerecht wird ("Atome reagieren, Menschen handeln" - eine Übertragung naturwissenschaftlicher Methoden in den Bereich der Sozialwissenschaften erscheint daher nicht sinnvoll).

Die Methodik des Kritischen Rationalismus reproduziert nur die Vorannahmen der WissenschaftlerInnen, die Wissenschaft entferne sich dabei von ihrem Erkenntnisobjekt (also der sozialen Realität) und beschäftige sich nur noch mit systemkonformen Untersuchungsgegenständen anstatt offen für Neues zu sein.

Die wissenschaftstheoretischen Grundannahmen des Interpretativen Paradigmas haben Folgen für die Methoden, mit denen die darauf basierende Forschung arbeitet. Insbesondere in den Sozialwissenschaften gilt es, den Handlungssinn und die vom Handelnden damit verbundenen Motive und Absichten zu verstehen. Die Wirklichkeit ist dabei eine Folge von Konstruktionen die durch den Alltagsverstand ausgesucht und interpretiert werden. Dies hat zur Konsequenz, dass die wissenschaftliche Begriffsbildung in den Sozialwissenschaften nach anderen Kriterien erfolgen muss als in den Naturwissenschaften. SozialwissenschaftlerInnen sind gezwungen, ihre Konstruktionen und Begriffe auf die Konstruktionen der Alltagshandelnden abzustimmen, während NaturwissenschaftlerInnen sich nicht darum kümmern müssen.

Interpretative ForscherInnen nähern sich dem Forschungsfeld grundsätzlich theorie- und hypothesenlos (aber mit einem forschungsleitenden Interesse) und nehmen dort die Position "pragmatisch-neutralen Beobachtens" ein, das heißt, sie versuchen, die Sinnsetzungen, die Handlungen, die Wirklichkeitskonstruktionen der Menschen in ihren Zusammenhängen zu deuten und zu verstehen (deshalb: Verstehende/Interpretative Sozialforschung). Dafür verwenden sie sehr flexible Methoden der Datenerhebung und -analyse ("Qualitative Methoden", z.B. Leitfadeninterviews, Konversationsanalyse) und halten den Forschungsablauf so offen wie möglich, berücksichtigen also alles, was Ihnen zur Beschreibung ihres Gegenstandes nützlich erscheint. Die Flexibilität der explorativen Vorgehensweise bedeutet nicht, dass die Untersuchung richtungslos vonstatten ginge, aber es bedeutet, dass der Blickwinkel zunächst weit ist, aber erst im Verlauf der Untersuchung fortschreitend zugespitzt wird. (vgl. Lamnek S, 1995, S. 3-38). Ziel der Forschung ist die Abstraktion vom Einzelfall, d.h. eine Typenbildung, die in einer Theorie des Forschungsfeldes enden sollte (es handelt sich also um ein induktives Vorgehen, bei dem von Einzelfällen auf allgemeineres geschlossen wird).

1.1.3 Natur- und geisteswissenschaftliche Begriffspaare

ERKLÄREN	VERSTEHEN
Nomothetisch: strebt nach dem Aufstellen von Gesetzmäßigkeiten	Idiographisch: beschreibt die Sinnkonstitutionen und damit einhegende Realitätskonstruktionen aktiv handelnder Subjekte
Theorieprüfend: ausgehend von einer Theorie werden Hypothesen aufgestellt, die messbar gemacht und geprüft werden	Theorieentwickelnd: aus den qualitativ erhobenen Daten der Sozialwelt, also den Selbstbeschreibungen und -thematizierungen der Handelnden (= Konstruktionen erster Ordnung) werden Typen entwickelt und zu einer Theorie des Untersuchungsgegenstands verdichtet (= Konstruktionen zweiter Ordnung)
Deduktives Vorgehen: von einem allgemein angenommenen Zusammenhang wird auf einen speziellen Zusammenhang in der beobachtbaren Realität geschlossen (vom Allgemeinen zum Speziellen)	Induktiv: Einzelfälle werden als exemplarisch für einen bestimmten Untersuchungsgegenstand gehalten und erlauben damit den Schluss auf allgemeine, "typische" Strukturen (Schluss vom Speziellen aufs Allgemeine)
Objektiv: es gibt eine außerhalb des Individuums liegende Realität, die (nur) wissenschaftlich erfassbar und beschreibbar ist	Subjektiv: alle Realität kommt erst durch aktive Konstruktion der handelnden Subjekte zustanden, diese Konstruktionsprinzipien gilt es zu erfassen und zu beschreiben
Ahistorisch: Das "Vorher" des Untersuchungsgegenstandes ist nicht erfassbar und wird deshalb für vernachlässigbar gehalten	Historisierend: Die Entwicklung des Untersuchungsgegenstandes ist für seine Erklärung und deren Gültigkeit bedeutend
Geschlossen: Untersuchungsausschnitt ("Sample") und Vorgehen ist genau vorgeschrieben, muss vorab geplant und genau befolgt werden (lineare Strategie ²)	Offen: Untersuchungsausschnitt soll möglichst theorieelos und daher unvoreingenommen wahrgenommen werden, Forschungsablauf folgt pragmatischen Gesichtspunkten und ist offen für Unerwartetes (zirkuläre Strategie)
Prädetermination durch ForscherInnen: ForscherIn ist vororientiert durch Theorie in ihrem/seinem Kopf, nur sie/er ist kompetent zum "richtigen" Einschätzen der Situation	Relevanzsystem der Betroffenen: Wirklichkeitssicht und Selbstthematizierungen der Untersuchten bilden Grundlage der Interpretation
Distanz: Kontakt des ForscherInnen zum Erforschten wird als "Störquelle" empfunden	Identifikation: Kontakt der ForscherInnen mit Erforschtem ist nicht zu vermeiden wird

² Zu Unterscheidung vom *linearen (quantitativen)* und *zirkulären (qualitativen)* Modell des Forschungsprozesses vgl. Flick, 1995

und ist auszuschalten	als konstitutiv für Forschungsprozess gesehen "Teilhabe mit fremdem Blick"
Statisch: Forschungsablaufplan wird vorher festgelegt und ist genau zu befolgen, d.h. erst nach der Datenerhebung folgt die Analyse	Dynamisch-prozessuale Datenerhebung: wann wo welche Daten erhoben werden, liegt im Ermessen des/der ForscherIn und wird durch sein Forschungsinteresse bestimmt Datenerhebung und Datenanalyse parallel
Partikularisch: Soziale Realität wird ausschnittsweise erfasst, alle Faktoren außerhalb dieses definierten Ausschnittes werden als konstant angenommen	Holistisch: Versucht einer ganzheitlichen Erfassung seines Forschungsobjektes, also so viel Kontext wie möglich mit in die Analyse einzubeziehen
Zufallsstichprobe: eine Stichprobe nach dem Zufallsprinzip gilt als "repräsentativ"	Theoretical Sampling: ausgehend von seinen ersten Analyseergebnissen suchen ForscherInnen weitere Untersuchungsfälle, um "blinde Flecken" seiner vorläufigen Theorie zu erhellen
Unterschiede: Falsifikationen bestimmter Basissätze führen zum Verwerfen einer ganzen Theorie oder zu ihrer raum-zeitlichen Einschränkung	Gemeinsamkeiten: innerhalb verschiedener Einzelfälle wird nach Gemeinsamem (z.B. Strukturen) gesucht: "Typenbildung"
Reduktive Datenanalyse: Soziale Phänomene werden in ihren Ausprägungen zählbar gemacht und in Zahlen verschlüsselt statistisch analysiert	Explikative Datenanalyse: Statt Daten zu reduzieren, werden sie deutend expliziert (mit Kontext "angereichert"), um sie zu verstehen
Hohes Messniveau: Viele Fälle werden zur Erklärung herangezogen, mit Hilfe von Messskalen lassen sich Unterschiede relativ genau (z.T. sogar prozentual) angeben	Niedriges Messniveau: Vergleichsweise wenige Fälle werden ausführlich analysiert und sind verständnisleitend, Unterschiede und Gemeinsamkeiten lassen sich nicht in Zahlen oder Prozenten angeben
Schlussfolgerungen	
Überprüfbar durch: Standardisierung des Ablaufs und der Methoden	Überprüfbar durch: Offenlegung des Vorgehens, eigener Annahmen, logische Konsistenz
Problem: Annahme, Realität kann über Zahlenmaterial ohne Zusatzinterpretation gesetzmäßig abgebildet werden. Es müssen aber Vorannahmen getroffen werden. Kein Nutzen ohne qualitative Interpretation! (Campbell 1974) "All research is a form of story telling" (Kidder & Fine 1987)	Problem: Annahme, Realität kann durch die Darstellung von Beziehungen abgebildet werden. Auch qualitatives Material muss selektiert, strukturiert, geordnet, bewertet und vor allem reduziert werden. Somit ist jede Forschung auch Quantifizierung
Vorteil: präziser, genauer, Modellbildung	Vorteil: Weniger Vorannahmen, offener
Für große Programme / Umfragen geeignet (aber bei Erhebung: achten auf Rücklauf,	Case Studies, für Samples, die für qualitative Analyse zu klein sind.

Verzerrung in Antworten, Non-Response, Kontrollgruppen → Zeit und Kosten!)	
Verbindungspotential: Grenzen der qualitativen Untersuchung festlegen, strukturelle Aspekte erfassen	Verbindungspotential: Rahmenbedingungen, Intentionen der Akteure verstehen (Giddens 1984)

1.2 Der Positivismusstreit und seine Folgen

1.2.1 die Kritik an der quantitativen Sozialforschung

Gegen die hinter den Modellen der Kritischen Rationalisten stehenden wissenschaftstheoretischen Grundpositionen wurden in Deutschland vor allem im sogenannten "Positivismusstreit" u. a. von Vertretern einer kritisch-dialektischen Gesellschaftstheorie (Adorno; Habermas; Luhman) Kritik formuliert.

Die Hauptkritikpunkte waren:

1. Bei der Messung geht es in der Sozialforschung immer um die Umformung kommunikativer Erfahrungen in Daten. Da es uninterpretierte Erfahrungen aber nicht gibt, stellt der Umformungsprozess des Messens selbst eine Interpretation dar. ForscherInnen greifen bei der Zuordnung von Zeichen zu Erfahrungen nach bestimmten Messregeln, auf ihr alltagsweltliches Vorverständnis zurück, ohne allerdings diesen Rückgriff auf den "common sense" zu explizieren (Cicourel, 1970). Dieser Prozess wird im wissenschaftlichen Vorgehen in keiner Weise systematisch reflektiert. Gerade dies soll aber bei den sogenannten qualitativen Verfahren der Fall sein.
2. Wissenschaftliches Erkenntnisinteresse und der Bereich potentieller Forschungsgegenstände wird von vornherein auf das "Positive" und das tatsächlich Gegebene beschränkt. Nicht unmittelbar sichtbare Phänomene, latente Sinnstrukturen bleiben vom Erkenntnisprozess ausgegrenzt (u.a. bei der quantitativen Inhaltsanalyse nach Berelson).
3. Erfahrung als Überprüfungs- und Bewährungsinstanz wissenschaftlicher Aussagen wird nur sehr eingeschränkt berücksichtigt. Zugelassen zur empirischen Prüfung von Hypothesen sind nur jene Erfahrungsdaten, die in irgendeiner Form standardisierbar (quantifizierbar) sind. Alltagswissen und lebensweltlicher Erfahrungsschatz werden damit als Quelle, Gegenstand und als Verifikationskriterium wissenschaftlicher Erkenntnis ausgeschlossen.
4. Ein einmal ausgebildetes methodisches Instrumentarium wird zur Verfolgung der verschiedensten Fragestellungen immer wieder herangezogen, ohne dass genügend auf die Eigenart der Forschungsgegenstände Rücksicht genommen wird. Oder überspitzt formuliert: Es wird nur untersucht, was vermeintlich mit den verfügbaren Methoden exakt erfasst werden kann, andere Forschungsfragen

werden ausgeschlossen, anstatt sie zum Anlass zu nehmen, für sie adäquatere Methoden zu entwickeln.

Die Kritik an quantitativen Methoden richtet sich also vornehmlich gegen deren Empirieferne, die eine für die Sozialforschung unvertretbare Distanz zum Gegenstandsbereich bewirkt. So machte Filstead den Trend zur Quantifizierung und die Vorherrschaft deduktiver Theoriebildung "mit ihrem Labyrinth der verschiedenen logischen, mathematischen und technischen Anhänge" verantwortlich für ein "verminderte(s) Verständnis der empirischen sozialen Welt" und eine "künstliche Auffassung von Realität" (Filstead 1979, S.31). Anstatt die soziale Realität zu erforschen, sei man durch die Verwendung quantitativer Methoden dazu übergegangen "die empirische soziale Welt so lange zu biegen, umzuformen und zu verdrehen, bis sie auf das Modell passte, dass zu ihrer Erforschung verwendet wurde." (ebd. S.32)

Blumer kommt zu dem Schluss, dass es mit einer deduktiven Methodologie, die sich im wesentlichen statistischer Forschungsmethoden, der Konstruktion logischer und mathematischer Modelle nahezu unmöglich sei, Prinzipien dafür aufzuzeigen, wie Probleme, Daten, Beziehungen, Konzeptionen und Interpretationen unter Berücksichtigung der Beschaffenheit der zu untersuchenden empirischen Welt entwickelt werden können (Blumer, 1979). Blumer wirft der quantitativ und deduktiv orientierten Sozialwissenschaft vor, sie habe sich "von der direkten Überprüfung der empirischen Welt" (ebd., S.115) abgewandt, überrascht auf den ersten Blick, da die Anwendung einer solchen Forschungsmethodologie die Überprüfung empirischer Hypothesen ja gerade ermöglichen soll. Blumer wirft der quantitativen Forschung jedoch nicht vor, dass ihr keine angemessenen Verfahren der Hypothesenüberprüfung zur Verfügung steht, sondern dass sie keinerlei Verfahrensvorschriften zur Generierung relevanter Hypothesen umfasst. Weil es im Rahmen deduktiver Forschung nur beschränkt möglich ist, relevante Hypothesen vor der Datenerhebung überhaupt zu formulieren, sind alle raffinierten Strategien der Hypothesentestung letztendlich nicht zielführend. Gerdes bringt eine solche Kritik folgendermaßen auf den Punkt:

"Die vorherrschende Methodologie geht implizit davon aus, die Forscher hätten in genügendem Maße inhaltsreiche Vorstellungen über die untersuchten Wirklichkeitsbereiche zur Verfügung; als problematisch erscheint hier lediglich, ob diese Vorstellungen tatsächlich der empirischen Wirklichkeit entsprechen. Fraglich bleibt dann jedoch immer noch, ob sich denn aus diesen Hypothesen - selbst wenn sie sich empirisch bestätigen lassen! - eigentlich ein zutreffendes Bild der sozialen Wirklichkeit zusammenfügen lässt: Sachverhalte, über die der Forscher keine Vorstellungen hat, weil er den betreffenden Wirklichkeitsbereich nicht umfassend kennt, können nämlich in seinen Hypothesen gar nicht auftauchen, werden also auch nicht getestet und fehlen folglich im wissenschaftlichen Bild dieses Wirklichkeitsbereichs. Sind solche Sachverhalte konstitutiv für den untersuchten Bereich, bleibt die wissenschaftliche Darstellung ohne ausreichenden Bezug zur Wirklichkeit - und zwar selbst dann, wenn sie sich ausschließlich auf empirisch bestätigte Hypothesen stützen kann." (Gerdes 1979, S.5)

Das entscheidende Argument gegen eine deduktive, hypothesentestende und für eine hypothesenfindende Forschungsstrategie ist dabei nicht, dass der Stand der Theoriebildung hinreichend gehaltvolle Hypothesen noch nicht zulässt. Vielmehr wird postuliert, dass eine Formulierung empirisch gehaltvoller Hypothesen in den Sozialwissenschaften grundsätzlich eine Vertrautheit mit der untersuchten sozialen Welt erfordert³. Dieser Zugang kann auf die Tradition der phänomenologischen Soziologie von Schütz und dem symbolischen Interaktionismus von Mead zurückgeführt und als interpretative Soziologie (vgl. Giddens, 1984) bezeichnet werden. Gemeinsam ist diesen Theorieansätzen, dass in Abgrenzung zu funktionalistischen Strömungen die Orientierungs- und Interpretationsleistungen der Akteure explizit in die Erklärung sozialen Handelns mit einfließen sollen.

1.2.2 Grundlagen interpretativer Verfahren

Gegen die qualitative Forschungsmethoden wurde im Gegenzug der Vorwurf mangelnder Reliabilität und unzureichender Repräsentativität, erhoben. In den fünfziger und sechziger Jahren wurden daher eine Reihe von Vorschlägen zur Systematisierung und Entwicklung qualitativer Methoden gemacht, und das auch von Seiten stärker quantitativ orientierter ForscherInnen; so z.B. das fokussierte Interview (Merton, Kendall 1979), oder aber teilnehmende Beobachtung. Gleichzeitig wurden Auswertungsverfahren für qualitative Daten entwickelt und beschrieben, welche die dargestellten Beschränkungen einer deduktiven Methode durch ein methodisch kontrolliertes induktives Vorgehen überwinden sollten (vgl. insbesondere Gaser, Strauss 1967).

In seiner Monographie über die interpretative Soziologie hebt Giddens hervor, dass diese Ansätze trotz vielfältiger Unterschiede und Gegensätze eine gemeinsame Perspektive teilen, nämlich das Verständnis der Produktion von Gesellschaft, als einen Prozess, dessen Ergebnis "eine auf Fertigkeiten beruhende, vom Menschen getragene und 'geschehen gemachte' Leistung" ist (Giddens 1984, S.16).

SozialwissenschaftlerInnen sind nur dann in der Lage, angemessene Theorien über eine soziale Lebensform vor dem Kontakt mit dem Gegenstandsbereich, d.h. vor der empirischen Phase des Forschungsprozesses zu formulieren, wenn ihnen das entsprechenden Wissen der dort Handelnden zur Verfügung steht, also in der Regel nur dann, wenn sie selber im Rahmen dieser Lebensform sozialisiert wurden. Weder die Konstruktion von Messinstrumenten (Fragebögen, Skalen etc.) noch die Klassifikation und Kategorisierung sozialer Phänomene durch soziologische Beobachter ist denkbar,

³ Nach Hoffmann-Riem (1980) liegt dem normativen Paradigma die Annahme zugrunde, einzelne Handlungen von Individuen könnten deduktiv aus der Kenntnis eines gemeinsamen Symbolsystems, d.h. eines kulturell geteilten Handlungsmusters, abgeleitet werden und kontextunabhängig beschrieben werden. Im psychologischen Behaviorismus z.B. führte die positivistische Orientierung zu einer Beschränkung auf beobachtbares "Verhalten" und zu einem Verzicht auf die Untersuchung sinnvermittelter "Handlungen".

ohne dass diese permanent auf ihr eigenes common-sense Wissen zurückgreifen. Die Fähigkeit, adäquate Hypothesen für einen Gegenstandsbereich zu formulieren, hängt damit vom Umfang des (praktischen) Vorwissens über diesen Gegenstandsbereich ab - ein Sachverhalt, der auch von quantitativ orientierten SoziologInnen erkannt und beschrieben wird (vgl. Esser 1987). Dieser Umstand verliert seine vermeintliche Trivialität, wenn man sich vergegenwärtigt, dass die gemeinhin im Kontext deduktiver Forschung getroffene Annahme, dass das Wissen der/des UntersucherIn für die Formulierung relevanter Hypothesen ausreicht, auf zwei weiteren, stillschweigenden Annahmen beruht:

1. Das für die Forschungsfragestellung bedeutsame Alltagswissen von UntersucherIn und Untersuchten ist identisch.
2. Die Untersuchten befolgen die in dem Gegenstandsbereich geltenden sozialen Regeln einheitlich und eindeutig, es existiert kein Spielraum der Norminterpretation: Regeln bspw. der Bedeutungszuschreibung sind für alle in der Sozialwelt Lebenden eindeutig - alle Akteure kennen sie und können sie in allen Situationen in der gleichen Weise anwenden.

1.3 Qualitative Verfahren im Sprachgebrauch

Die Verwendung des Begriffs qualitativen Methoden im Rahmen einer Evaluation bezieht sich zumeist auf völlig unterschiedliche Aspekte der interpretativen Sozialforschung welche im folgenden erläutert werden sollen.

1.3.1 (quantitative) Analyse qualitativer Daten

Häufig wird der Begriff verwendet, um die Anwendung von qualitativen Verfahren zur Gewinnung von Daten (z.B. offene, halboffene, narrative Interviews, teilnehmende Beobachtung usw.) zu bezeichnen, die jedoch traditionell quantitativ ausgewertet werden, d.h. die mit qualitativen Verfahren gewonnenen qualitativen Daten werden in quantitative Daten transformiert und mit den üblichen statistischen Methoden verarbeitet.

Eine zweite Variante betrifft die besonderen Auswertungsverfahren, die auf Daten angewendet werden können, die z.B. nur dem Nominal- oder Ordinalskalenniveau entsprechen und die häufig als qualitative Daten bezeichnet werden. Zu diesen Auswertungsverfahren zählen viele "nichtparametrische" Verfahren und spezielle Weiterentwicklungen wie Kontingenzanalyse, Strukturanalyse usw. (Rudinger et al. 1985).

1.3.2 explorative Voruntersuchung zur Generierung von Hypothesen

Unter diese zweite Kategorie qualitativer Forschung fallen alle jene Forschungen, die qualitative Methoden nur für Vorversuche zur Generierung von Hypothesen einsetzen. Qualitative Forschung wird hier nicht als eigenständige Methode angesehen, sondern nur für die Vorbereitungsphase in einem Forschungsprozess akzeptiert. Deswegen werden auch keine hohen Ansprüche hinsichtlich der üblichen Kriterien wie Objektivität, Zuverlässigkeit und Gültigkeit gestellt. Die qualitative Forschung - in dieser Phase der

Untersuchung eingesetzt - muss aber in der Lage sein, Neues und Unerwartetes zutage zu fördern, das dann in der eigentlichen Forschungsphase erst wissenschaftlich "exakt" untersucht werden muss.

1.3.3 sinnverstehende Vorgehen

Beim sinnverstehenden Vorgehen werden individuelle Bedeutungszuschreibungen, situative Kontexte und gesellschaftliche Rahmenbedingungen in den Forschungsprozess mit einbezogen und durch kommunikatives Handeln zugänglich gemacht. Diese dritte Kategorie von qualitativer Forschung wird als Alternative zur traditionellen Forschungsauffassung verstanden. Es wird eine Angemessenheit zwischen den gewählten Methoden und den untersuchten Fragestellungen angestrebt, und die Fragestellungen ergeben sich meist aus gesellschaftlichen Problemstellungen und Bedürfnissen.

Qualitative Verfahren im Sinne interpretative Sozialforschung kommen dann zum Einsatz, wenn ein bestimmtes Erkenntnisinteresse vorliegt, das mit quantitativen Methoden nicht erreicht werden kann. Gründe hierfür können sein:

1. Der Zugang zum Untersuchungsgegenstand: Bestimmte Einrichtungen, Institutionen oder auch Personen lassen sich z.B. auf Fragebogenuntersuchungen kaum ein (z.B. Alternative Betriebe, Führungskräfte usw.). Bei offenen Gesprächen, freien Interviews o.ä. gibt es solche Ablehnungen nicht in gleichem Maße. Andere haben ein so großes (berechtigtes) Misstrauen gegenüber jeder Art von wissenschaftlicher Untersuchung, dass erst nach gründlicher Klärung der Zielsetzung der Untersuchung und persönlicher Bekanntheit der UntersucherInnen eine Mitarbeit zugesagt wird. Durch die Methode der teilnehmenden Beobachtung ist diese Hürde manchmal zu nehmen. Auch die Vereinzelung in der Untersuchungssituation wird bisweilen abgelehnt, vor allem dann, wenn dies keine vertraute Alltagssituation ist. Ein Gespräch zu mehreren kann angemessener sein.
2. Die Relevanz der Aussagen der Befragten: Oftmals sind die Meinungen, Einschätzungen, Bewertungen oder Erinnerungen der Befragten nicht direkt abrufbar. Nur durch hilfreiche oder geschickte Unterstützung durch die UntersucherInnen können Reflexionen und Erinnerungen in Gang gesetzt werden, die dann zu Antworten führen, die mehr sind als Widerspiegelungen der Oberfläche oder mehr als das, was die Befragten für sozial erwünscht halten. Auch kann nur dann mit einer angemessenen Darstellung von Seiten der Befragten gerechnet werden, wenn Ihnen der Sinnzusammenhang der Untersuchung und damit auch jeder einzelnen Frage in der Untersuchung einleuchtet und nachvollziehbar ist. Anderenfalls werden die Antworten so beliebig sein, wie die Fragen den Befragten erscheinen.
3. Das richtige Verständnis und die richtige Einordnung der Aussagen: Durch die Betonung des zirkulären und Prozesshaften bei der qualitativen Forschung ergeben sich vielfältige Möglichkeiten für die UntersucherInnen, sich

rückzuversichern, ob sie das Gesagte richtig verstanden und eingeordnet haben: Durch die intensive Benutzung von Sprache in der Untersuchung, durch die Möglichkeiten der Rückfrage, durch die Verzahnung von Durchführung und Auswertung - und der damit ermöglichten frühzeitigen Rückkoppelung von Ergebnissen - und, sofern teilnehmende Beobachtung als Verfahren eingesetzt wird, durch die Erfahrungen am eigenen Leibe.

4. Die Breite des Untersuchungsansatzes: Das 'Nichtfestgelegtsein' auf bestimmte Hypothesen, die Verzahnung von Durchführung und Auswertung und die maximale (strukturelle) Variation der Perspektiven erlauben eine umfassende und flexible Auseinandersetzung mit dem zu untersuchenden Problem. Die Tatsache, dass die Fragestellung im Laufe der Untersuchung präzisiert und eingegrenzt werden kann, macht eine 'umfassende' Untersuchung handhabbar und ökonomisch vertretbar.
5. Die Überwindung des Vorverständnisses der UntersucherInnen: Fehlende oder falsche Vorstellungen der UntersucherInnen von ihrem Untersuchungsfeld sind bei der Verwendung von qualitativen Methoden kein Handicap. Sofern es den UntersucherInnen gelingt, die geforderte 'Offenheit' zu bewahren, und sie keine zu starken inhaltlichen und strukturellen Vorgaben machen, enthalten qualitative Vorgehensweisen genügend Korrekturmöglichkeiten, die es ermöglichen, dass die UntersucherInnen ihr falsches bzw. unvollständiges Vorverständnis entdecken und korrigieren können. Unterstützt wird dies auch durch Diskussionen der UntersucherInnen untereinander, sofern sichergestellt werden kann, dass diese Diskussionen einigermaßen 'herrschaftsfrei' ablaufen.
6. Die Entdeckung relevanter Fragestellungen und angemessener Vorgehensweisen: Durch die Verzahnung von Durchführung und Auswertung, durch die frühzeitige Rückvermittlung von Ergebnissen an die Untersuchten und durch die wachsende Kenntnis des Feldes ist es den UntersucherInnen möglich, die Relevanz der gewählten Fragestellung zu beurteilen. Damit ist nicht gemeint, dass die Untersuchten ihnen sagen könnten oder sollten, welche Fragestellungen für sie relevant sind, noch, dass die ForscherInnen - z.B. ausgehend von einer wissenschaftlichen Theorie - die Frage der Relevanz entscheiden könnten, sondern dass nur gemeinsam eine Fragestellung ausgehandelt werden kann.

Welche Vorgehensweisen angemessen sind zeigt sich u.U. schon daran, welche von den Untersuchten überhaupt akzeptiert werden (s.o.), aber auch daran, welche Vorgehensweisen im Sinne des Untersuchungszieles ergiebig sind. Durch die Verzahnung von Durchführung und Auswertung ersparen sich die UntersucherInnen zumeist unerfreuliche Überraschungen, wenn sie bei u.U. großer Datenmenge feststellen müssen, dass ihre Daten nicht auswertbar sind, wie es sich bei strikter Trennung von Durchführung und Auswertung leicht ergibt.

1.4 Die Bedeutung der interpretativen Sozialforschung für die Evaluation

In interpretativen Ansätzen werden die zu untersuchenden bzw. zu evaluierenden Personen als informierte "MitarbeiterInnen" verstanden und nicht als „Forschungs-subjekte“. Aber auch die Evaluatoren werden „überprüft“. Eine klare Trennung der Rollen erscheint nicht sinnvoll, da alle erklären, argumentieren, interagieren. Die Evaluatoren sind für Rückmeldungen und Kritik der zu Evaluierenden offen und sind auf diese Weise um Qualität sowie laufende Entwicklung bemüht. Zielsetzung ist ein vertieftes Verständnis für alle Beteiligten zu bewirken, welches sich im Laufe des Evaluations-Prozesses entwickelt. Eine derartige Zugangsweise ist relativistisch und die Methodologie daher qualitativ. Die Vorgangsweise ist jedoch reglementiert, um jederzeit nachvollziehbar zu sein und um eine Überprüfung der Ergebnisse zu ermöglichen.

Guba und Lincoln (1989) fordern in ihrem „Fourth-Generation-Evaluation“-Ansatz explizit den Einsatz interpretativer im Gegensatz zu qualitativen Methoden. Sie charakterisieren die vorangegangenen Generationen durch die jeweilige Orientierung an Messung, Beschreibung und Urteilsfindung. Die 4. Generation hingegen liefert ausverhandelte Bilder der sozialen Realität, welche allen Beteiligten nicht nur zugänglich sind, sondern ihnen gehören, um deren Ergebnisse für weitere Entwicklungen zu verwenden.

Guba und Lincoln definieren für eine Evaluation der vierten Generation Kriterien für faire, bildende, ontologische, katalytische, und taktische Authentizität und lenken damit die Aufmerksamkeit auf ethische und beziehungsrelevante Fragen:

Fairness meint dabei das Ausmaß, in dem alternative Konstruktionen des Forschungsprozesses und Ergebnisse, welche durch die zu Evaluierenden angeregt wurden auch als solche in einem eigenen Teil des Berichtes ausgewiesen werden. Eine Evaluierung hat auf die Beteiligten Akteure insofern eine bildende Wirkung, als diese lernen, die Interessen anderer zu verstehen, und zwar jenseits ihrer eigenen Stakeholder-Gruppe. Ontologisch, katalytisch und taktische Authentizität beziehen sich auf die Fähigkeit der Beteiligten ihre eigene Situation besser zu erfassen und zu verstehen, durch den Evaluierungsprozess motiviert zu sein, ihre Situation zu verbessern und tatsächlich ermächtigt zu werden dies auch zu tun.

1.5 Die Verbindung quantitativer und qualitativer Methoden

In der Gegenüberstellung qualitativer und quantitativer Methoden in der Sozialforschung versuchte u.a. Wilson die besonderen Merkmale beider Herangehensweisen herauszustellen und die jeweiligen Verkürzungen und Schwachstellen aufzuzeigen. Er plädiert für eine ausgewogene Verwendung beider Methoden. (Wilson 1982) Wie bereits ausführlich erläutert verbindet sich mit der quantitativen Methoden das Vorbild der Naturwissenschaften, d.h. es sollen allgemeingültige nomothetische Erklärungen mittels deterministischer oder prohabilitischer Gesetze gefunden und quantitativ formuliert werden. Die Weiterentwicklung qualitativer Methoden wiederum ist aus der Ablehnung

des naturwissenschaftlichen Vorbilds und die Betonung der Kontextabhängigkeit der sozialen Erscheinungen forciert worden. Soziale Erscheinungen sollen möglichst detailreich beschrieben werden mit ihren Beziehungen untereinander und zu ihrem übergeordneten Ganzen. Kausalen Verallgemeinerungen wird kein Wert beigemessen. Wilson betont, dass methodische Auffassungen nicht von philosophischen Doktrinen abhängen sollten, sondern von den Erscheinungen, bzw. Tatsachen bestimmt werden sollten, die man untersuchen will: "Methodologische Auffassungen werden zwar nicht durch die Tatsachen festgelegt, doch sie müssen auf die Tatsachen Rücksicht nehmen." (Wilson 1982, S. 489).

Eine Kombination von quantitative and qualitative Methoden ist jedoch bei sorgfältiger Planung als sinnvoll zu bewerten (vgl. dazu auch Miles, 1983). Generell ist der Einsatz qualitativer Instrumente besser geeignet für die explorative Erforschung noch relativ ungenau definierter Fragestellungen bzw. offener Problemfelder innerhalb dieser (vgl. Smith, Manning 1982), da sie die Entwicklung von Konzepten unterstützen. Quantitative Instrumente erlauben eine stärkere Fokussierung und damit eine strengere Analyse von Detailfragen. Smith und Manning definieren Forschung einen Prozess, welcher von Paradigmen ausgeht, die zuerst durch qualitative und anschließend durch quantitative Forschung formalisiert werden. Gerade für die notwendige Berücksichtigung von Feedback-Schleifen sind dabei qualitative Ansätze hilfreich.

Kleining entwickelte eine Methodologie, die quantitative und qualitative Methoden nicht als unvereinbar gegenüberstellte, sondern beiden Ausrichtungen ihren Platz zuwies in einem umfassenden Modell aller sozialwissenschaftlichen Methoden (Kleining 1982).

Kleining geht davon aus, dass die Alltagstechniken (die alltäglich genutzten Strategien zur Auseinandersetzung mit unserer Umwelt) das Reservoir für alle sozialwissenschaftlichen Methoden sind. Aus diesen Alltagstechniken sind die sozialwissenschaftlichen Methoden (zur Erkenntnis der Umwelt) durch Ausgrenzung, Absonderung aus ihrem Alltagszusammenhang und durch Abstraktion entwickelt, und zwar die qualitativen Verfahren als eine erste, die quantitativen als eine zweite Stufe der Abstraktion.

Hieraus leitet Kleining drei Folgerungen ab:

1. **Einheit der Methoden:** Alle Verfahren stehen miteinander in Beziehung.
 2. **Abfolge der Verfahren:** Die Alltagsstrategien sind früher als die sozialwissenschaftlichen Verfahren (die durch Ausgliederung und Abstraktion aus ihnen entstanden sind), die qualitativen früher als die quantitativen, weil sie einen geringeren Abstraktionsgrad besitzen (früher = logisch / genetisch, nicht historisch zu verstehen). Dies impliziert, dass eine wissenschaftliche Behandlung eines Themas mit Techniken zu beginnen hat, die mit niedrigerem Abstraktionsgrad aus den Alltagstechniken entstanden sind und nicht mit solchen eines hohen Abstraktionsniveaus.
- Qualitative Forschung muss in jedem Fall der quantitativen vorausgehen, braucht aber nicht von ihr gefolgt zu werden. Umgekehrt gilt: Die quantitative Forschung

braucht als Vorstufe die qualitative Forschung, ohne die sie Gefahr läuft, Sinnlosigkeiten zu produzieren.

- 3. Qualitative und quantitative Verfahren sind gegensätzlich:** Beide Verfahren gehören verschiedenen Abstraktionsebenen an und sie unterscheiden sich in der Art der Abstraktion: Qualitative Verfahren abstrahieren die Technik des Vergleichs auf Gemeinsamkeiten, quantitative Verfahren abstrahieren die Technik des Vergleichs auf Unterschiede (Genauer: Die qualitative Sozialforschung arbeitet in Richtung auf das Erfassen von Gemeinsamkeiten, die quantitative in Richtung auf das Erfassen von Unterschieden). Beide Verfahren haben unterschiedliche Zielsetzungen: Qualitative Verfahren zielen auf das Aufdecken von Bezügen, quantitative Verfahren auf das Messen unterschiedlicher Ausprägungen schon bekannter Bezüge. In ihrer reinen Form sind qualitative Verfahren das genaue Gegenteil der quantitativen Verfahren.

Ein 'Methodenverschnitt' - wie er leider recht häufig zu beobachten ist - resultiert nach Kleinig aus der unglücklichen Kombination beider Strategien: der Verwendung qualitativer Verfahren in einer linearen Strategie (siehe dazu die eingangs beschriebenen Begriffspaare zur Charakterisierung der Wissenschaften). Hier kann kaum etwas Neues gefunden werden, das Verfahren ist extrem unökonomisch und die Auswertung solcher Daten erweist sich in der Regel als außerordentlich frustrierend, weil im Nachhinein - wenn es zu spät ist - klar wird, an welchen Stellen die Weichen anders hätten gestellt werden müssen.

Sommerlad (1992) argumentiert, dass Methoden so gewählt werden müssen, dass sie sowohl der Situation angemessen sind, welche durch Stakeholder, Ressourcen, Kenntnissen der Methoden und politischer Interessen gegeben ist. In dieser Situation werden Fragen gestellt, deren Beantwortung bestimmten Zielgruppen dienen soll. Allzu häufig werden Instrumente definiert, bevor diese wesentlichen Rahmenbedingungen geklärt sind.

Eine Systematisierung des Einsatzes unterschiedlicher methodologischer Zugänge im Rahmen eines Forschungsprojektes wurde von Denzin vorgenommen. Er bezeichnet diese Möglichkeiten als Triangulationen und unterscheidet vier Arten, deren Kombinationen von ihm ausführlich diskutiert werden (Denzin 1970):

1. Triangulation der Daten: Die Verwendung unterschiedlicher Datenquellen in einer Studie (zu unterschiedlichen Zeitpunkten und von unterschiedlichen Personen)
2. Triangulation der Untersuchenden: Unterschiedliche Forscher oder Evaluatoren beschäftigen sich mit der gleichen Fragestellung
3. Triangulation der Theorie: Unterschiedliche Perspektiven werden genutzt um die Kategorien der Analyse festzulegen
4. Triangulation der Methoden: unterschiedliche Methoden werden zur Analyse derselben Fragestellung eingesetzt

Ein derartige Betrachtungsweise kann jedoch beim Versuch der Umsetzung teuer werden und als Ergebnis eine Serie von schlecht implementierten Methoden liefern anstelle eine mit Sorgfalt gewählte und durchgeführte.

1.6 Fragenkatalog zum Evaluationsdesign

Aus den grundlegenden Unterschieden ergeben sich folgende Differenzierungen bzw. Verbindungen der Methoden im Bereich der Evaluation:

1. Formative Evaluation wird eher mit qualitativen Methoden arbeiten, summative kann auch quantitative einsetzen.
2. Zur Hypothesenbildung und Methodenabklärung eignen sich qualitative Erhebungen und Methoden der Interpretation
3. Ebenso können quantitative Ergebnisse für das bessere Verständnis durch qualitative Erhebungen ergänzt werden.
4. Eine weitere Möglichkeit besteht darin, an Hand einzelner Fragestellungen bei der Festlegung des Evaluationsdesigns zu einem geeigneten Methodenmix zu kommen.

Für eine sukzessive Annäherung an optimales Evaluationsdesign sind im folgenden Fragen aufgelistet, welche eine möglichst genaue Beschreibung des vorliegenden Problemfeldes erlauben. Bei der Entwicklung des Design ist es von Bedeutung, zwischen den Möglichkeiten der Datenerhebung und der Auswertung zu unterscheiden. Es sei in diesem Zusammenhang auf die dementsprechenden obigen Ausführungen zur begrifflichen Klärung qualitativer Verfahren noch einmal verwiesen. Zu beachten gilt, dass dem Wunsch nach optimaler Lösung in jedem Fall auch alltägliche Begrenzungen wie Kosten und Zeitpläne gegenüberstehen.

1.6.1 Fragestellungen zu Rahmenbedingungen und Evaluationsgegenstand

Art der Evaluation: formativ, summativ, Projekt/Programm.

Arten der formativen Evaluierung: Implementierung, Fortschritt bzw. Verlauf

Wie ist das Evaluationsobjekt entstanden?

Welche Problemstellung wird behandelt?

Wer ist an dem Projekt beteiligt? (Netzwerk der aktiven und passiven AkteurInnen)

Was war/ist die Zielsetzung?

Überwiegen monetäre Rechtfertigungen, sollen Strukturen verbessert werden, aus der systematischen Betrachtung der Erfahrungen gelernt werden, ...

Welche Zielsetzungen sind operationalisierbar, messbar?

Abklärung der Interessen von Stakeholdern, Auftraggeber, Zielgruppe:

- Was sind ihre Grundsätzlichen Ziele?

- Welche Fragen werden gestellt? Welche Prioritäten können zugeordnet werden?
- Sind die Fragen deskriptiv, interpretativ, normativ
- Welche Art der Information wird benötigt?

Budgetrahmen des Eval-Objektes?

Budgetrahmen für Evaluierung?

Zeitraumen des Projektes und der Evaluation? (Wann beginnt die Evaluation, welche zeitlichen Abstände gibt es.)

1.6.2 Prinzipielle Vorgangsweise für die Design-Entwicklung

Aus den allgemeinen Fragenkatalog sind die spezifischen Formulierungen für das jeweilige Evaluations-Objekt abzuleiten.

Danach führen folgende Arbeitsschritte zum Evaluations-Design:

1. Festlegung der benötigten Information zur Beantwortung der Fragen
2. Festlegung der Erhebungsmethoden
3. Erhebung der Information bzw. Daten
4. Analyse der Daten
5. Berichtslegung für die Zielgruppe bzw. den Auftraggeber

Überprüfung von Budget und Zeitplan können zu Anpassung des Designs führen. Ebenso können erste Interviews, erhaltenes Material etc. Adaptionen notwendig machen.

1.6.3 Fragen zur Klärung der Methoden

Was waren/sind die Zielsetzungen?

Welche Zielsetzungen sind operationalisierbar, messbar?

- Welche Fragen werden gestellt?
- Sind die Fragen deskriptiv, interpretativ, nomothetisch?
- Welche Art der Information wird benötigt?
- Wer wird die Information verwenden?
- Welche Auswertung wird dem Informationsgehalt des Materials gerecht?

Schema zur Methodenklärung					
Frage	Art d. Frage (deskr., interpr., nomoth.)	Erhebung (was, wie, woher)	Auswertung	Zeitplan	Kosten

Budgetrahmen des Evaluationsgegenstandes?

Budgetrahmen für Evaluierung?

Zeitraumen des Projektes und der Evaluation?

Wann beginnt die Evaluation, welche zeitlichen Meilensteine sind zu formulieren?

Welche Daten werden benötigt? Welche Daten sind bereits vorhanden?

Wie können die fehlenden Daten am besten erhoben werden?

Wie „einzigartig“ ist das Evaluations-Objekt?

Sollen die Ergebnisse generalisiert werden?

(Je einzigartiger die Situation desto sinnvoller ist die Verwendung qualitativer Verfahren, da eine Generalisierung nicht möglich ist und Nutzen, Wirkung auf TeilnehmerInnen besser beschrieben werden kann. Vgl. dazu Lincoln 1989)

1.6.4 Checkliste für Erhebungsmethoden

Bei Erhebungen auf Grund von Fragebögen, welche einer quantitativen Analyse zugeführt werden sollen, gilt es potentielle Probleme zu vermeiden welche aus der Ziehung der Stichprobe, dem Antwortverhalten, der Notwendigkeit von Non-Response-Analysen sowie von Kontrollgruppe resultieren können.

Bei Qualitativen Erhebungen ist die Auswahl der Fälle mit großer Sorgfalt vorzunehmen: Sollen es extreme Positionen (worst/best case) oder durchschnittliche Fälle sein?

Welche Sekundärdaten sind vorhanden und können eingesetzt werden?

1. Dokumente: Korrespondenz, Proposals, Projektbeschreibungen, Berichte,
2. Daten: vorhandenes Zahlen-Material zu Projekten, Personen, Web-Seiten

Daten können in weiterer Folge erhoben werden von:

1. Direkt von Individuen
 - 1.1. Fragebögen,
 - 1.2. Tagebücher
2. Durch BeobachterInnen
 - 2.1. Beobachtungen
 - 2.2. Interviews
 - 2.2.1. Strukturiert, offen, Tiefeninterviews
 - 2.2.2. (Fokus-) Gruppen (Achtung auf Wirkung der Personen aufeinander!):
 - 2.2.3. Moderiert,
3. Mittels technischer Hilfen:
 - 3.1. Video, Tonband, Fotografien
 - 3.2. Bibliographische Methoden

1.7

Literatur

Berelson, B. (1952). <i>Content Analysis in Communications Research</i> , Glencoe ILL: The Free Press.
Blumer, H. (1954). What is wrong with social research. <i>American Sociological Review</i> , 14, 3-10.
Blumer, H. (1979). Methodologische Prinzipien empirischer Wissenschaft. In K. Gerdes (Hrsg.), <i>Explorative Sozialforschung</i> . Stuttgart: Enke.
Burrell G., Morgan G. (1979): <i>Sociological Paradigms and Organisational Analysis</i> . London: Heinemann.
Checkland, Peter (1981). <i>Systems Thinking, Systems Practice</i> . Chichester: John Wiley
Cicourel A. V. (1970), <i>The Acquisition of Social Structure. Toward a Development Sociology</i> , in: <i>Understanding Everyday Life</i> , hg. von J. D. Douglas, Aldine Publ., Chicago, 136-163
Cook, T.D., Campbell, D.T. (1979). <i>Quasi-experimentation: Design and Analysis Issues or Field Settings</i> , Chicago, IL: Rand McNally.
Denzin, N. K., Lincoln Y. S. (1998): <i>Collecting and Interpreting Qualitative Materials</i> , Thousand Oaks, CA.: SAGE
Denzin, N.K. (1970) "Chapter 12: Strategies of Multiple Triangulation", from Norman K. Denzin, <i>The Research Act: A Theoretical Introduction to Sociological Methods</i> , Aldine Pub. Co.: Chicago, pp. 298–313.
Devereux, G. (1984). <i>Angst und Methode in den Verhaltenswissenschaften</i> . Frankfurt/Main: Suhrkamp.
Esser, H. (1987) : Warum die Routine nicht weiterhilft. Überlegungen zur Kritik an der "Variablen-Soziologie", in: Müller, N. H. (Hrsg.): <i>Problemlösungsoperator Sozialwissenschaft</i> , Bd. 1. Stuttgart: Enke, S. 230-245.
Dodds, A. E., Lawrence, J. A., & Guiton, P. (1984). University students' perceptions of influences on external study. <i>Distance Education</i> , 5(2), 174–185.
Filstead, W. J. (1970). <i>Qualitative methodology</i> . Chicago: Markham. Gage, N. L. (1989). The paradigm wars and their aftermath. <i>Educational Researcher</i> , 18(7), 4–10.
Fitz-Gibbon & Morris (1987): <i>How to Design a Program Evaluation</i> . Newbury Park, CA: Sage
Flick, U. (1995). <i>Qualitative Forschung. Theorie, Methoden, Anwendung in Psychologie und Sozialwissenschaften</i> . Reinbek: Rowohlt, 318
Fuchs, W. (1970). Empirische Sozialforschung als politische Aktion. <i>Soziale Welt</i> , 21/22, 1-7.
Gergen, Kenneth J. (1994) <i>Toward Transformation of Social Knowledge</i> (2nd ed.). Sage, Thousand Oaks, CA.
Giddens A. (1984): <i>The Constitution of Society</i> . Berkeley and Los Angeles: University of California Press.
Glaser, B. G., & Strauss, A. L. (1967). <i>The discovery of grounded theory</i> . Chicago: Aldine.
Guba E., Lincoln Y. (1989): <i>Fourth Generation Evaluation</i> . Newbury Park, California: Sage.
Hoffmann-Riem, C. (1980). Die Sozialforschung in einer interpretativen Soziologie - Der Datengewinn. <i>Kölner Zeitschrift für Soziologie und Sozialpsychologie</i> , 32, 339-372.
House, E. (1980). <i>Evaluating with Validity</i> . Beverley Hills: Sage.

Kelle U.(1994): Empirisch begründete Theoriebildung. Zur Logik und Methodologie interpretativer Sozialforschung, Weinheim: Deutscher Studien Verlag, S. 44 – 53.
Kidder L.H., Fine M. (1987): Qualitative (and Quantitative) Research: When Stories Converge. In Mark M.M., Shotland R.L. (eds): Multiple Methods in Program Evaluation. New Directions for Program Evaluation, no. 35, San Francisco: Jossey-Bass.
Kleining, Gerhard (1982). Umriß zu einer Methodologie qualitativer Sozialforschung. Kölner Zeitschrift für Soziologie und Sozialpsychologie, 34, 224-253.
Kuhn, Thomas S. (1976), Die Struktur wissenschaftlicher Revolutionen. Frankfurt a. M.
Lamnek, S, (1995): Qualitative Sozialforschung. Bd. 1 Methodologie. Weinheim
Lincoln Y.S. (1991): The Arts and Sciences of Programm Evaluation. Evaluation Practice 12 (1), 1-7.
Miles, M. B. (1983). Qualitative data as an attractive nuisance: The problem of analysis. In J. Van Maanen (Ed.), <i>Qualitative methodology</i> . Beverly Hills, CA: Sage.
Mills, C Wright (1959). The Sociological Imagination. Oxford: Oxford University Press.
Mingers, John (1992). "Technical, practical and critical OR - past, present and future?". In Mats Alvesson and Hugh Willmott (eds), <i>Critical management studies</i> . London: Sage.
Patton M.Q. (1990): Qualitative Evaluation and Research Methods. Newbury Park, CA: Sage.
Patton, M. Q. (1981). Creative Evaluation. Beverley Hills: Sage.
Rossi P.H., Freeman H.E., Hofmann G. (1988): Programm-Evaluation. Einführung in die Methoden der angewandten Sozialforschung, Stuttgart: Enke.
Rudinger, G., Chaselon, F., Zimmermann, E. J., Henning, H. J. (1985). Qualitative Daten. München: Urban & Schwarzenberg.
Schütz, A. (1932) "Der sinnhafte Aufbau der sozialen Welt. Eine Einleitung in die verstehende Soziologie" .
Schütz, A. (1971): Gesammelte Aufsätze I. Das Problem der sozialen Wirklichkeit. Den Haag, S. 39-54.
Smith R. B., Manning P. K. (1982). <i>A handbook of social science methods, volume 2: Qualitative methods</i> . Cambridge, MA : Ballinger.
Sommerlad, E. (1992). <i>A Guide to Local Evaluation</i> , Report no. HE/L62/1186, London: Evaluation Development and Review Unit, Tavistock Institute of Human Relations London: Employment Department.
Stern, Eliot (1991). "The Evaluation of Policy and the Politics of Evaluation". In Tavistock Institute Annual Review 1990, pp.28-30.
Stern, Eliot (1993). "The challenge of 'real-time' evaluation". In Tavistock Institute Annual Review 1992/3, pp.38-41.
Taylor, Peter C. (1996) Constructivism and interpretive research. Paper presented at seminar, Science Education Centre, National Taiwan Normal University, Taipei; 7 May 1996.
Thomas P. Wilson (1973): Theorien der Interaktion und Modelle soziologischer Erklärung.
Van Maanen, John (1988) <i>Tales of the Field: on Writing Ethnography</i> . The University of Chicago Press, Chicago.
Wilson, T. P. (1973). Theorien der Interaktion und Modelle soziologischer Erklärung. In Arbeitsgruppe Bielefelder Soziologen. Alltagswissen, Bd. 1, S. 54 - 79

Wilson, T. P. (1982). Qualitative oder quantitative Methoden in der Sozialforschung. Kölner Zeitschrift für Soziologie und Sozialpsychologie, 34, 487-508.
Wittmann W.W. (1985) Evaluationsforschung. Aufgaben, Probleme und Anwendungen. Berlin: Springer
Lamnek, S. (1995): Qualitative Sozialforschung. Bd 2: Methoden und Techniken 3. Aufl. München: Psychologie Vlg. Union.
Merton, R. K, Kendall, P. L. (1979): Das fokussierte Interview, in: Hopf, Ch./Weingarten, E. (Hrsg.) 1979: Qualitative Sozialforschung, Stuttgart: Klett-Cotta

2 KAPITEL ZWEI KOSTEN NUTZEN ANALYSE UND FISKALANALYSE

Dorothea Sturn, Mirjam Novakovic

2 KAPITEL ZWEI	
KOSTEN NUTZEN ANALYSE UND FISKALANALYSE	1
2.1 KOSTEN NUTZEN ANALYSE	2
2.1.1 <i>Bedeutung der Methode für die Arbeit des Instituts</i>	2
2.1.2 <i>Kurzcharakterisierung der Methode in seiner klassischen Anwendung:</i>	2
2.1.2.1 Die volkswirtschaftliche Dimension der Kosten Nutzen Analyse.....	2
2.1.2.2 Arbeitsschritte in der Kosten Nutzen Analyse	3
2.1.3 <i>Anwendungen im Rahmen der Evaluation von Innovations- und Technologiepolitik</i>	6
2.1.3.1 Volkswirtschaftliche Bewertung technologiepolitischer Programme	6
2.1.3.2 Privater versus sozialer Nutzen.....	7
2.2 FISKALANALYSE	10
2.2.1 <i>Grundsätze der Fiskalanalyse</i>	10
2.2.1.1 Fragestellungen.....	10
2.2.1.2 Fiskalische Effekte am Beispiel kommunaler Beschäftigungsförderung	11
2.2.1.3 Kosten-Nutzen Indikatoren	12
2.2.1.4 Untersuchungsdesign	14
2.2.1.5 Grenzen und Möglichkeiten.....	17
2.2.2 <i>Ein konkretes Beispiel: Fiskalanalyse der arbeitsmarktpolitischen Maßnahme „Office Management Centre“</i>	17
2.2.2.1 Datenmaterial und Methodik der Fiskalanalyse	18
2.2.2.2 Schritt 1: Die Kalkulation von Bruttokosten und Bruttoerträgen	19
2.2.2.3 Schritt 2: Modellierung eines Referenzszenarios	20
2.2.2.4 Schritt 3: Szenario nach Jahren:.....	21
2.2.2.5 Schritt 4: Kalkulation der Nettokosten und Nettoerträge nach Jahren	22
2.2.2.6 Schritt 5: Aggregation und Diskontierung der Nettokosten und Nettoerträge.....	24
2.2.2.7 Schritt 6: Berechnung des Break Even Points.....	26
2.2.2.8 Schritt 7: Sensitivitätsanalyse	26
2.2.2.9 Zusammenfassung.....	28

2.1 **Kosten Nutzen Analyse**

2.1.1 Bedeutung der Methode für die Arbeit des Instituts

Die **Kosten-Nutzen-Analyse** ist eine Analysemethode, die in ihrer klassischen Form auf die Bewertung (ex-ante sowie ex-post) von großen, investiven Projekten zugeschnitten ist. In dieser Form ist sie sehr teuer und wird im Institut nur selten zur Anwendung kommen – wenn auch die Analyse sozio-ökonomischer Kosten und Nutzen für große Projekte im Rahmen der Strukturfonds sowie für alle Projekte im Rahmen des Kohäsionsfonds verlangt wird. Von besonderem Interesse ist die Methode aber vor allem aufgrund der Anwendbarkeit einzelner Aspekte auch im Rahmen kleinerer Wirkungsanalysen. Dies betrifft insbesondere die Kalkulation indirekter Effekte, die Abschätzung intangibler Effekte und die Möglichkeiten der Quantifizierung von Wirkungen einer Intervention: Das kosten-nutzen-analytische "Korsett" erzwingt, dass explizite Aussagen in bezug auf die Art der (eingetretenen oder zu erwartenden) Kosten und Nutzen und ihre ungefähre Größenordnung getroffen werden.

2.1.2 Kurzcharakterisierung der Methode in seiner klassischen Anwendung:

Durch einen Vergleich der volkswirtschaftlichen Kosten und Nutzen eines Projekts soll ermittelt werden, ob es durchführenswert ist; bzw. soll ein Vergleich der Vor- und Nachteile mehrerer Projekte den Nachweis erbringen, welches das relativ beste ist. Diese Ermittlungen lassen sich ex ante – als Grundlage für Planungsentscheidungen – oder ex post – als Methode der Evaluation – anstellen. Die KNA ist damit eine volkswirtschaftliche Methode zur Selektion oder Bewertung öffentlicher Projekte, mit der versucht wird, die Kosten und Nutzen eines Projekts in monetären Größen auszudrücken. Auf diese Weise wird ein einheitlicher Bewertungsmaßstab zum Vergleich von Projektkosten und Projektnutzen geschaffen. Die KNA berücksichtigt das Zeitprofil der Kosten und Nutzen, d.h. die Tatsache, dass insbesondere bei größeren Projekten Kosten und Nutzen in unterschiedlichen Zeiträumen anfallen, indem sie den Gegenwartswert der volkswirtschaftlichen Kosten und Nutzen ermittelt und miteinander vergleicht. Die angestrebte Maximierung des Gegenwartswerts aller Nutzen vermindert um den aller Kosten, soll der Rationalisierung der Verwendung finanzieller Mittel im öffentlichen Sektor dienen.

2.1.2.1 Die volkswirtschaftliche Dimension der Kosten Nutzen Analyse

Die Besonderheit der KNA besteht darin, dass die Effekte eines Projekts für die gesamte Volkswirtschaft versucht wird zu ermitteln. Dies hat die folgenden Konsequenzen:

- **Externalitäten** werden mit berücksichtigt. Das heißt, die nicht marktbewerteten Auswirkungen des Projekts auf Dritte jenseits des Projektträgers bzw. der betrachteten geförderten Firma oder Person sind zu ermitteln. Wie später noch gezeigt wird, sind in

der technologiepolitischen Diskussion Externalitäten in Form von Wissens-Spill-Overs von besonderer Bedeutung.

- Die Bewertung der erfassten Kosten und Nutzen erfolgt gemäß der **volkswirtschaftlichen Opportunitätskosten**. Die Opportunitätskosten geben an, welchen Nutzen die durch das Projekt gebundenen Ressourcen bei anderweitiger Verwendung gestiftet hätten. Immer dort, wo Marktunvollkommenheiten existieren, weichen die Opportunitätskosten von den beobachtbaren Preisen ab und man berechnet „Schattenpreise“, welche um die Verzerrungen korrigieren sollen. In der Praxis hat dies einige triviale Konsequenzen, wie z.B. dass von den direkten Kosten die Steuern abgezogen werden müssen, es folgen aber auch einige hochproblematische Bewertungen aus der Verwendung von Schattenpreisen: Beispielsweise ist die Bewertung von Zeit immer problematisch und die Verwendung eines hypothetischen Lohnsatzes für die betroffene Gruppe kann nur dann legitimiert werden, wenn der mit dem Projekt verbundene Zeitverlust wirklich alternativ als Arbeitszeit genutzt worden wäre. Umgekehrt ist die Verwendung von „Schattenlöhnen“ – im Grenzfall in der Höhe von Null - anstelle angefallener Lohnkosten im Fall hoher Arbeitslosigkeit schwierig und kontrovers.
- **Verteilungsaspekte** spielen eine Rolle. Werden unterschiedliche Gruppen durch das Projekt unterschiedlich belastet und begünstigt, so sind diese anhand eines geeigneten Wohlfahrtskriteriums zu aggregieren.
- Die **Diskontrate** soll im Idealfall sozial definiert sein, d.h. ein Maß der Zeitpräferenz darstellen. Wie hoch diese ist, darauf hat die ökonomische Theorie bislang keine eindeutige und operationalisierbare Antwort gefunden. Häufig wird die Effektivverzinsung langfristiger Staatsanleihen verwendet oder auch die langfristige reale Wachstumsrate der Volkswirtschaft als Indikator der durchschnittlichen sozialen Opportunitätskosten.

2.1.2.2 Arbeitsschritte in der Kosten Nutzen Analyse

Schritt 1: Definition und Abgrenzung der zu berücksichtigenden Kosten und Nutzen

Reale und pekuniäre Kosten und Nutzen: Prinzipiell sind nur reale Kosten und Nutzen zu berücksichtigen, d.h. projektinduzierte Änderungen der physischen Produktionsmöglichkeiten sowie Änderungen des Nutzens. Pekuniäre Nutzen und Kosten entstehen im Zusammenhang mit Änderungen in den relativen Preisen, die sich ergeben wenn die Märkte sich an die neue Situation (mit Projekt) anpassen. Problematisch wird die Unterscheidung immer dann, wenn man davon ausgehen kann, dass sich die Märkte nicht perfekt anpassen können – sei es weil Externalitäten vorliegen, oder auch Verteilungswirkungen zu vermuten sind.

Direkte und indirekte Kosten und Nutzen: Direkte Nutzen und Kosten sind diejenigen, die auf das Projektziel bezogen sind, indirekte Kosten und Nutzen lassen sich als „Nebenprodukte“ oder „Sekundäreffekte“ definieren. Wie weiter unten noch genauer gezeigt

wird, spielen in der technologiepolitischen Diskussion indirekte Kosten und Nutzen eine ganz erhebliche Rolle. In die KNA sollten beide Kategorien eingehen.

Interne und externe Kosten und Nutzen: Interne Kosten belasten den Projektträger und müssen von ihm getragen werden, externe Kosten belasten Dritte. Analoges gilt für den Nutzen. Eine KNA berücksichtigt im Unterschied zu betriebswirtschaftlichen Rentabilitätsrechnungen auch externe Kosten und Nutzen (Externalitäten). Dies sind meist schwer zu bewerten, da es sich in der Regel um Intangible handelt wie z.B. die Senkung der Lebensqualität durch Lärmbelastung, aber auch soziale Effekte wie z.B. die Wahrung sozialer Strukturen bei regionalen Projekten etc..

Tangible und intangible Kosten und Nutzen: Tangible Kosten und Nutzen können immer dann, wenn keine offensichtlichen Marktunvollkommenheiten vorliegen, zu Marktpreisen bewertet werden, ansonsten behilft man sich mit Schattenpreisen. Für intangible Kosten und Nutzen existieren keine Marktpreise. Soweit möglich werden intangible Effekte mit Hilfe von Opportunitätskosten bewertet, häufig sind diese Effekte allerdings so schwierig bewertbar, dass sie lediglich am Ende der Analyse genannt werden und eine qualitative Argumentation die Bedeutung dieser Effekte unterstreicht. Beispiele sind intangible Nutzen, die im Zusammenhang mit Bildungsprojekten auftreten wie allgemeine Steigerung der Lebensqualität (?) oder die positiven Auswirkungen eines insgesamt höheren Bildungsstandes auf die Demokratie.

Bei diesem Arbeitsschritt können die folgenden Fehler auftreten:

- Falsche Abgrenzung des „Projekts“, d.h. wesentliche Abhängigkeiten bleiben unerfasst. Auch können Zurechnungsprobleme auftreten, wenn bestimmte Effekte das Ergebnis mehrerer schwierig voneinander trennbarer Projekte ist.
- Der Zeithorizont kann zu lang oder zu kurz gewählt werden.
- Bei der Abgrenzung der Kosten und Nutzen ist ein sehr häufiger Fehler, dass pekuniäre Kosten oder Nutzen mit einbezogen werden. Oft bleiben externe Effekte und / oder intangible Effekte unberücksichtigt.

Schritt 2: Quantifizierung der physischen Komponenten von Kosten und Nutzen

In diesem zweiten Schritt wird das sogenannte „Mengengerüst“ der KNA erstellt. Es werden Prognosen errechnet, wie viele Flächen, Gebäude, Personen etc. von dem betrachteten Projekt im festgelegten Zeitraum betroffen sind.

Vor allem bei ex-ante-Bewertungen sind Prognosen möglichst konservativ zu berechnen und es empfiehlt sich, Prognose-Varianten zu rechnen.

Schritt 3: Bewertung der relevanten Kosten und Nutzen

Eine Besonderheit der KNA ist, dass Kosten und Nutzen durch Abbildung in monetären Größen kommensurabel gemacht werden.

Auf Seiten der **Kosten** nimmt man bei hinreichend vollkommenen Märkten Marktpreise, ansonsten Schattenpreise. Wie bereits erwähnt, ist die Bestimmung von Preisen für auftretende externe Effekte besonders schwierig. Theoretisch werden für solche Effekte die

sozialen Kosten des externen Effekts eingesetzt: Welche Summen an Geld müssten den betroffenen Personen gezahlt werden, um sie für den negativen Effekt (z.B. für die Lärmbelästigung) zu kompensieren? Selbst wenn Märkte oft Anhaltspunkte geben, wie hoch diese Kosten sind (z.B. sinkende Mieten in Flugplatznähe geben einen Anhaltspunkt für die sozialen Kosten der Lärmbelästigung), so steht man hier häufig auch vor dem Problem, den Beobachtungszeitraum richtig zu bestimmen. So können bei negativen Umweltauswirkungen von Projekten die heutigen sozialen Kosten gering sein, die Kosten für folgende Generation jedoch erheblich.

Der **Nutzen** des Projektoutputs wird prinzipiell in Einheiten der Zahlungsbereitschaft für die Inanspruchnahme des Projektoutputs gemessen. Dieses Konzept setzt daher eine annähernde Kenntnis von Nachfrage nach dem Projektoutput voraus, der Nutzen kann dann definiert werden als die Konsumentenrente (Integral unter der Nachfragekurve abzüglich des Projektumsatzes). Auch die Form der Angebotskurve ist von Belang, denn anfallende Produzentenrenten sind mit zu berücksichtigen.

Bei diesem Arbeitsschritt können die folgenden Fehler auftreten:

- Überschätzung der Kosten fallen vor allem bei existierenden Marktverzerrungen an, d.h. wenn die beobachtbaren Preise über den Opportunitätskosten liegen (z.B. bei Arbeitslosigkeit).
- Unterschätzung der Kosten ergibt sich häufig durch eine unvollständige Berücksichtigung von negativen Externalitäten
- Eine Über- oder Unterschätzung des Nutzens resultiert in erster Linie aus einer falschen Ermittlung von Nachfrage- und Angebotsfunktionen. Häufig werden hier betriebswirtschaftlich orientierte Ersparnis Ansätze zu Grunde gelegt anstatt Zahlungsbereitschaften ermittelt.

Schritt 4: Aggregation der Kosten- und Nutzenströme, Wahl einer Diskontrate

In der Regel fallen Kosten und Nutzen eines Projekts zu unterschiedlichen Zeitpunkten in verschiedenem Ausmaß an. Um die Ströme über die Zeit hinweg aggregieren zu können, benötigt man einen Gewichtungsfaktor, der als sozialer Diskontsatz die gesellschaftliche Zeitpräferenz abbilden sollte. Die Wahl der Diskontrate ist vor allem dann von Bedeutung, wenn unterschiedliche Projekte mit verschiedenen zeitlichen Ausgabenprofilen verglichen werden. Beispielsweise kann ein Projekt mit einem zunächst sehr hohen Investitionsaufwand und in der Folge geringen laufenden Kosten den gleichen Nutzen stiften wie ein Projekt mit annähernd konstanten Kostenströmen über die Zeit hinweg. In diesem Fall wird mit einer niedrigen Diskontrate eher das Projekt mit dem hohen Investitionsaufwand präferiert, mit einer hohen Diskontrate eher das Projekt mit den konstanten Kostenströmen. Je höher die Diskontrate, desto höher bewertet man heutige Kosten und Nutzen im Vergleich zu künftig anfallenden.

Das Ergebnis der Aggregation sind Gegenwartswerte der Kosten und Gegenwartswerte der Nutzen.

Schritt 5: Wahl eines Entscheidungskriteriums

Werden – vor allem im Rahmen von ex-ante-Evaluationen – alternative Projekte evaluiert, so ist schließlich ein Entscheidungskriterium zu wählen, woran sich die Auswahl der Alternativen orientieren soll. Gängige Kriterien sind:

- Die Maximierung des Kapitalwerts (Gegenwartswert der Nutzen minus Gegenwartswert der Kosten), d.h. jene Alternative ist zu wählen, die den höchsten Kapitalwert erbringt
- Die Maximierung des internen Zinssatzes, d.h. jene Alternative ist zu wählen, die den höchsten internen Zinssatz erbringt.

Wird nur ein Projekt bewertet – wie das bei ex-post-Evaluationen häufig der Fall ist - so ist die zeitliche Darstellung des Saldos aus Nutzen und Kosten sinnvoller, da das Bild aussagekräftiger ist als eine einzelne Zahl. Es kann davon ausgehend ermittelt werden, in welchem Ausmaß künftige Nutzenströme anfallen müssten, damit das Projekt einen positiven Kapitalwert erbringt. Alternativ dazu kann die Berechnung des internen Zinssatzes zeigen, wie niedrig die soziale Diskontrate jedenfalls anzusetzen ist, damit das Projekt einen positiven Kapitalwert erbringt.

Schritt 6: Durchführung einer Sensitivitätsanalyse

Sensitivitätsanalysen sollen zeigen, wie robust die Berechnungen sich gegenüber Änderungen zentraler Annahmen verhalten. Meist ist es die soziale Diskontrate, die einer Sensitivitätsanalyse unterzogen wird: Je unterschiedlicher die Zeitstruktur von Kosten- und Nutzenströmen, desto stärker werden die Ergebnisse bei unterschiedlichen Diskontratsätzen voneinander abweichen.

2.1.3 Anwendungen im Rahmen der Evaluation von Innovations- und Technologiepolitik

2.1.3.1 Volkswirtschaftliche Bewertung technologiepolitischer Programme

Kosten-nutzen-analytische Überlegungen spielen in diesem Kontext immer noch eine wichtige Rolle, wo es um „social returns“ oder **volkswirtschaftliche Effekte technologiepolitischer Programme** geht. Dies heißt im wesentlichen, dass nicht nur Nutzen und Kosten der direkten Programmteilnehmer erfasst und quantifiziert werden, sondern auch jene Kosten- und Nutzenkategorien, die an anderen Stellen der Volkswirtschaft auftreten. Bei der Erfassung des volkswirtschaftlichen Effekts treten die folgenden prinzipiellen Probleme auf, welche unterschiedliche Ansätze unterschiedlich lösen:

- **Zurechnungsproblem:** Welcher Anteil der – kosten- und nutzenstiftenden - technologischen Veränderung ist auf das Programm zurückzuführen?
- **Volkswirtschaftliche Dimension:** Sind alle Nutzenkategorien aller Marktteilnehmer erfasst? Wie können – positive und negative – Effekte abgeschätzt werden, die nicht bei den direkten Programmteilnehmern auftreten?

- Wie kann man mit **intangiblen Effekten** umgehen (Bewertung von Zeit, von Sicherheit, von Lebensqualität, von Wissenszuwachs)? Können Schattenpreise ermittelt werden?
- **Zeitstruktur:** Zu welchen Zeiten fallen Kosten an, zu welchen Zeiten fallen Nutzenströme an? Was ist aufgrund dessen der richtige Zeitpunkt für die Evaluation? Wie lassen sich die Effekte diskontieren?

Im folgenden konzentrieren sich die Ausführungen auf das Problem der Ermittlung sozialer Kosten und Nutzen, da dies bei der Bewertung technologiepolitischer Programme von besonderer Bedeutung ist. Einer der zentralen Gründe für öffentliche Intervention bei Forschungs- und Innovationsprozessen ist, dass aufgrund hoher Spill-overs bzw. positiver Externalitäten der soziale Nutzen weit größer ist als der private, woraus eine – gemessen am sozialen Optimum – private Unterinvestition in F&E erfolgt. Dient diese Argumentation als zentrale Legitimation für technologiepolitische Programme, so sollte man zumindest eine grobe Vorstellung über das Ausmaß des sozialen Nutzens haben, der mit Innovationsprozessen verbunden ist.

2.1.3.2 Privater versus sozialer Nutzen

Im Rahmen von Forschungs- und Innovationsprozessen können zwei Wege der Weitergabe von Nutzen an andere unterschieden werden:

- Das neue Wissen bzw. der technologische Wandel ist **„Embodied“**, d.h. er materialisiert sich in neuen Produkten oder neuen Prozessen. In diesem Fall wird die „Neuheit“ über Marktbeziehungen an die Nachfrage bzw. vor- und nachgelagerte Branchen weitergegeben.
- Das neue Wissen bzw. der technologische Wandel ist **„Disembodied“**: In diesem Fall wird Wissen an andere via Veröffentlichungen, Personaltransfer etc. weitergegeben, auch ohne dass Marktbeziehungen vorhanden sind. Das neue technologische Wissen kann sich in diesem Fall jeder aneignen, in der ökonomischen Theorie wird dieser Fall „Manna-Fortschritt“ genannt.

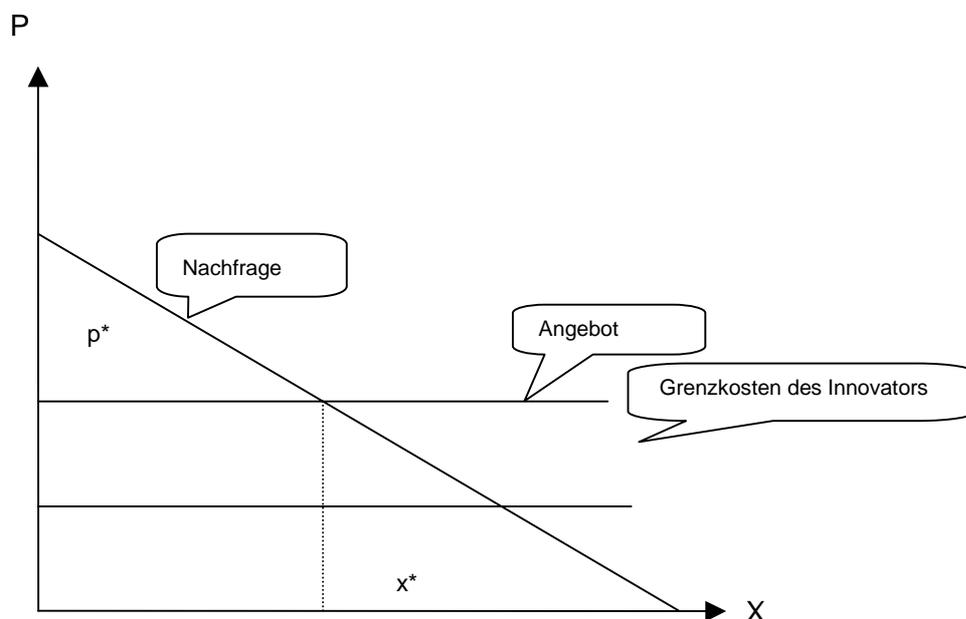
Ein in der neueren Diskussion immer wichtiger gewordene Spezialfall von „disembodied“ technologischem Wandel stellt jene Situation dar, bei dem sozialer Nutzen durch komplementäres Know-how, kritische Massen und Netzwerkeffekte generiert wird.

Es sollen nun drei Fälle unterschieden werden, bei denen jeweils unterschiedliche Regimes der Weitergabe von Nutzen an andere – trotz völlig gleicher Marktbedingungen – zu sehr unterschiedlichen volkswirtschaftlichen Bewertungen führen. Der Einfachheit halber sei angenommen, ein technologiepolitisches Programm wirke allein auf eine Verbesserung der Produktionstechnik einer am Programm teilnehmenden Firma, was sich in sinkenden Durchschnitts- und Grenzkosten¹ niederschlägt. Die Marktnachfrage bleibt unverändert.

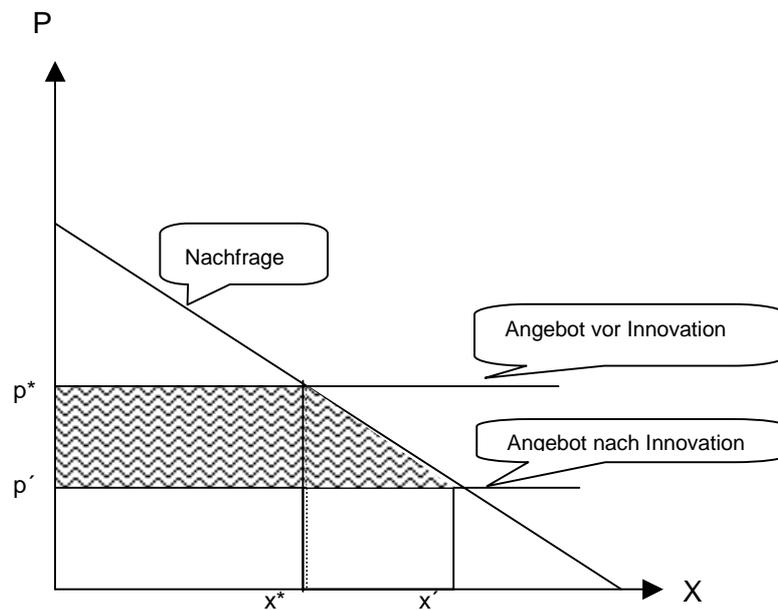
¹ Es seien konstante Durchschnitts- und Grenzkosten angenommen, damit ist die Angebotskurve eine Horizontale.

Fall 1: Nutzen fällt lediglich beim Innovator an

Angebot und Nachfrage bilden bei der Preis-Mengen-Kombination $p^* x^*$ ein Marktgleichgewicht. Die erfolgreiche Prozessinnovation des Innovators führt dazu, dass sich seine Grenzkosten unter der Marktangebotskurve befinden, da nur er selbst diese neue Technologie anwenden kann – beispielsweise weil er in der Lage ist, die Erfindung zu schützen – bleibt das Marktgleichgewicht unverändert. Da der Innovator selbstverständlich zum Marktpreis p^* verkauft, erwirtschaftet er eine Produzentenrente im Ausmaß der schraffierten Fläche. Diese Fläche stellt gleichermaßen den privaten wie auch den sozialen Nutzen der Innovation dar. Die Situation ist nicht stabil: Der Innovator wird einen hohen Anreiz haben, in dem betrachteten Markt Monopolmacht zu erlangen.

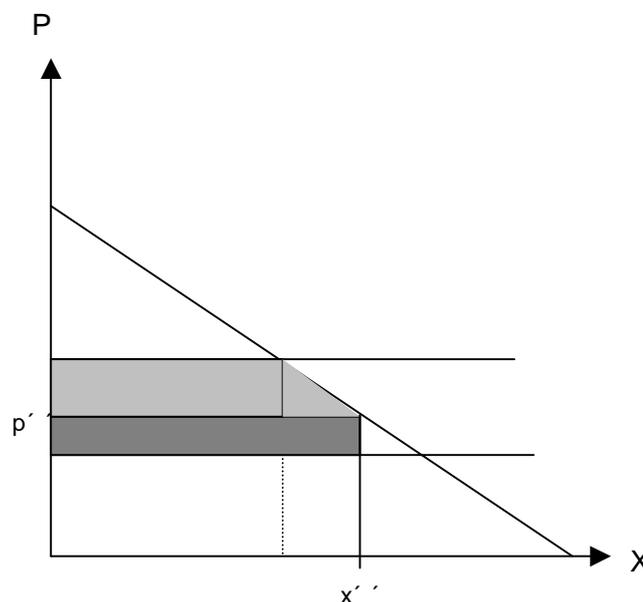
**Fall 2: „Disembodied“: Wissen eignen sich alle anderen an**

Fall 2 zeigt die andere Extremsituation: Das neue technologische Wissen hat hier den Charakter eines öffentlichen Gutes, jeder kann die neue Technologie verwenden. Damit wird die Grenzkostenkurve des neuen Produktionsprozesses zur neuen Marktangebotskurve und der Marktpreis sinkt entsprechend auf die neue Gleichgewichtssituation $p' x'$. Die Produzentenrente ist Null, die zusätzliche Konsumentenrente jedoch erheblich, sie entspricht der schraffierten Fläche. Der soziale Nutzen ist hier wesentlich größer als in Fall 1, die Innovationsanreize für den einzelnen allerdings gering, bzw. in diesem Extremfall gar nicht vorhanden.



Fall 3: „Embodied Wissen“: Nutzen wird teilweise an die Nachfrage weitergegeben

Dieser dritte Fall zeigt eine Situation zwischen den beiden zuvor geschilderten Extremsituationen: Eine Gruppe von Innovatoren kann zu niedrigeren Grenzkosten produzieren, das Wissen um die neue Technologie ist allerdings kein öffentliches Gut sondern in der Technologie „embodied“, lediglich die Innovatoren selbst können es anwenden – oder möglicherweise auch verkaufen. Die Konkurrenz der Innovatoren untereinander wird dazu führen, dass der Preis sinkt, in der Regel aber nicht bis zum Niveau der neuen Grenzkosten, sondern nur bis etwa p' . Der soziale Nutzen der Innovation enthält damit sowohl Elemente der Konsumentenrente (hell schraffierte Fläche) als auch Elemente der Produzentenrente (dunkel schraffierte Fläche).



Der Vergleich dieser einfachen, stilisierten Fälle zeigt, wie falsch der in Evaluationen häufig gewählte Ansatz ist, den Nutzen der Innovation in Einheiten von Umsatzzuwachs oder Gewinnzuwachs der Innovatoren zu messen, ohne den gesamten Markt im Auge zu haben. Gemessen am privaten Nutzen der Innovatoren stellt Fall 1 die beste Situation dar und Fall 2 die schlechteste. Gemessen am sozialen Nutzen ist es genau umgekehrt: Fall 2 maximiert den gesellschaftlichen Nutzen, das heißt die Summe aus Produzenten- und Konsumentenrente.

2.2 Fiskalanalyse

In der *Fiskalanalyse* sollten einerseits die Kosten und andererseits die Erträge, die durch eine Intervention für die öffentliche Hand entstehen, beziffert und einander gegenübergestellt werden, um daraus Aussagen über die fiskalische Effizienz der Intervention abzuleiten.

Die Fiskalanalyse stellt damit eine Unterform der Kosten-Nutzen-Analyse (KNA) – der Be-“Wertung” bzw. E-“Valuierung” im engeren Sinne – dar. Im Gegensatz zu individuellen oder gesamtgesellschaftlichen Kosten-Nutzen-Bilanzen interessieren hier jene Kosten und Nutzen, die für die öffentliche Hand relevant sind, während andere Effekte nicht berücksichtigt werden.

2.2.1 Grundsätze der Fiskalanalyse

"Eine Fiskalanalyse umfasst die Quantifizierung der monetären Effekte einer Förderaktion auf die öffentlichen Budgets." (Fehr-Duda, Neudorfer 1994, S.48). Ziel ist es, Fördermaßnahmen der öffentlichen Hand nach fiskalischen Gesichtspunkten zu beurteilen. Im folgenden werden die Grundsätze der Fiskalanalyse am Beispiel arbeitsmarktpolitischer Maßnahmen illustriert.

2.2.1.1 Fragestellungen

Wozu dient das Instrument der Fiskalanalyse z.B. im Bereich der Arbeitsmarktpolitik? Eine zentrale Forschungsfragestellung lautet: Lohnen sich die Aufwendungen, die einer Region durch Finanzierung und Durchführung von Maßnahmen regionaler oder nationaler Beschäftigungsförderung entstehen, in Relation zu erwartbaren fiskalischen und sozialen Nutzeneffekten?

Damit wird die Grundsatzproblematik des volkswirtschaftlichen Saldos zwischen den Aufwendungen für Arbeitslosigkeit und Existenzsicherung einerseits sowie den Erträgen und Einsparungen durch öffentlich geförderte Erwerbstätigkeit andererseits berührt. Letztlich geht es also um die Generalthematik: "Finanzierung von Arbeit oder Finanzierung von Arbeitslosigkeit". Diese volkswirtschaftliche Grundsatzfragestellung, die den gesamtfiskalischen Aspekt öffentlicher Beschäftigungsförderung anspricht, ist oft der Hintergrund solcher Analysen.

2.2.1.2 Fiskalische Effekte am Beispiel kommunaler Beschäftigungsförderung

Zu den fiskalischen Effekten zählen einerseits die Aufwendungen, die für die Finanzierung entsprechender Projekte zur arbeitsmarktlichen Integration von Erwerbslosen bereitgestellt werden, auf der anderen Seite sind Einnahmen (durch Steuern), Einsparungen (Von Sozialhilfe), Wertschöpfung (mittels Arbeit) zu verzeichnen, die durch Beschäftigung statt Arbeitslosigkeit entstehen.

Fiskalische Effekte:

- **Kommunale Aufwendungen**

Darunter werden jene Mittel verstanden, die im städtischen Verwaltungshaushalt zur (Mit)Finanzierung von Beschäftigungs- und Qualifizierungsprojekten bereitgestellt werden.

- **Multiplikationseffekte**

Die Multiplikationseffekte beschreiben das Verhältnis zwischen den kommunalen Aufwendungen für die Projekte zu sogenannten Drittmitteln (nicht kommunale Finanzierung). Kommen z.B. 60% der Projektfinanzierung aus nicht-kommunalen Quellen, erhält die Stadt für ATS 1.—nochmals ATS 1,6.-- zusätzlich aus fremden Kassen.

- **Einnahmefeffekte**

Damit sind die Erträge gemeint, die durch den kommunalen Anteil am Gesamtaufkommen der Lohn- bzw. Einkommenssteuer entstehen.

- **Einsparungseffekte**

Das sind jene Wirkungen, die sich durch die Vermeidung kommunaler Sozialleistungen bei Beschäftigung im Sinne von Opportunitätskosten monetär beschreiben lassen.

- **Wertschöpfungseffekte**

Die Analyse dieser als „Wertschöpfungseffekte“ bezeichneten Wirkungen richtet sich darauf, ob und inwieweit monetär faßbare Leistungen, z.B. in der Form von produzierten Gütern oder erstellten Dienstleistungen, durch die Beschäftigung von Arbeitslosen speziell für die Kommune im Sinne eines fiskalischen Zuwachses entstehen, ohne bereits durch die städtischen Aufwendungen für diese Projekte abgedeckt zu sein.

- ***Psycho-soziale Effekte:***

Da eine ausschließliche fiskalische Bilanzierung von Kosten und Nutzen arbeitsfördernden Maßnahmen nicht voll gerecht wird, wird empfohlen auch Individualanalysen durchzuführen, in deren Rahmen materielle, soziale, gesundheitliche und psychische Be- und Entlastungseffekte bei den Projektteilnehmern untersucht werden.

Eine zentrale Forschungsfragestellung lautet hier: Lohnen sich die Aufwendungen für Maßnahmen der örtlichen Beschäftigungsförderung, und zwar im Hinblick darauf, dass mit

Arbeitslosigkeit vermutlich verbundene Belastungen psycho-sozialer Art durch die kommunalen Maßnahmen kompensierbar oder zumindest teilkompensierbar sind?

- Materielle Be/Entlastungseffekte

Unter materiellen Belastungseffekten sind das verfügbare Pro-Kopf-Haushaltseinkommen und die Konsumausgaben sowie der Aspekt der Verschuldung gemeint.

- Soziale Be/Entlastungseffekte

Zu den sozialen Belastungseffekten zählen Kontaktprobleme, Einsamkeit und Rückzugtendenz.

- Gesundheitliche Be/Entlastungseffekte

Obwohl die meisten gesundheitlichen Effekte immer eng mit psycho-somatischen Aspekten verbunden sind, wird hier nur der allgemeine Gesundheitszustand betrachtet (Durchschnittliche Anzahl von Arzt- und Krankenhausbesuchen pro Jahr pro Person).

- Psychische Be/Entlastungseffekte

Dazu werden jene Krankheitssymptome gezählt, die primär aus seelische Belastungen zurückzuführen sind z.B. „Ärger mit den Nerven“, „Schlafstörungen“ und „Ermüdungserscheinungen“.

- Gesellschaftspolitische Integration/ Desintegration

Untersuchungsgegenstände sind hier „Verlust von Politikinteresse“, „Staatsverdruß“, „Wahlabstinenz“ etc.

2.2.1.3 Kosten-Nutzen Indikatoren

Kosten

Fiskalisch:

- Investitionskosten
- Personalkosten und Sachkosten
- Opportunitätskosten der Verwaltung
- Ansprüche an das System der sozialen Sicherung während der Maßnahme
- Neue Ansprüche an das System der sozialen Sicherung nach der Maßnahme

Psycho-sozial:

- Beeinträchtigung des Selbstvertrauens
- Beeinträchtigung sozialer Kontakte und Kommunikationsfähigkeiten
- Beeinträchtigung des persönlichen Aktivitätsniveaus, der Entscheidungs/Handlungsfähigkeit und des persönlichen Zeitmanagement...

Erträge**Fiskalisch**

- Einsparung an Transferleistungen
- Rückflüsse an die Sozialversicherung
- Direkte Steuerleistungen
- Indirekte Steuerleistungen

Psycho-sozial:

- Stärkung des Selbstvertrauens
- Erweiterung sozialer Kontakte und Kommunikationsfähigkeiten
- Stärkung der gesundheitlichen und psychischen Konstitution...

Beispiel:**Kosten**

1. Investitionskosten
2. Personalkosten und Sachkosten
3. Opportunitätskosten der Verwaltung
4. Ansprüche an das System der sozialen Sicherung während der Maßnahme
5. Neue Ansprüche an das System der sozialen Sicherung nach der Maßnahme

Erträge

6. Einsparung an Transferleistungen
7. Rückflüsse an die Sozialversicherung
8. Direkte Steuerleistungen
9. Indirekte Steuerleistungen

Die Investitionskosten (1) beinhalten die durch die öffentliche Hand geförderten Leistungen. Bei den Personal- und Sachkosten (2) werden Nettobeträge errechnet, da die darauf entfallenden Steuern und Sozialversicherungsbeiträge direkt der öffentlichen Hand zufließen. Als Opportunitätskosten (3) der Verwaltung werden die Kosten eingerechnet, die durch die Verwaltung der Maßnahme entstehen, auch hier werden Nettobeträge gerechnet. Bei vielen Programmen haben die TeilnehmerInnen den Anspruch (4) auf Deckung des Lebensunterhaltes. Dieser ist ebenfalls als Nettobetrag auf der Kostenseite zu verbuchen. Gleichzeitig entfallen jedoch Transferleistungen wie Arbeitslosengeld oder Notstandshilfe, die von den DLU-Kosten abzuziehen sind. Durch die Teilnahme am Programm können neue (5) Ansprüche an das System der sozialen Sicherung entstehen.

Durch die Aufnahme eines Beschäftigungsverhältnisses erlischt der Anspruch auf Arbeitslosengeld (6), diese Beträge sind als Nettobeträge auf der Ertragsseite zu verbuchen. Sobald der Teilnehmer am Programm ein höheres Einkommen erhält als vorher, steigen auch seine Sozialversicherungsbeiträge (7). Auch die Lohnsteuersätze sind auf der Ertragsseite zu verbuchen (8). Indirekte Steuern (9): Unter der Annahme, dass die Teilnehmer das gesamte Einkommen während und nach der Maßnahme verausgaben, wird das durchschnittliche Konsumbündel mit $x\%$ versteuert.

2.2.1.4 Untersuchungsdesign

Je nach Komplexität der Forschungsfragestellung muss dementsprechend ein mehr oder weniger vielschichtiges Untersuchungsdesign gewählt werden. Es können sowohl individuelle als auch kollektive Effekte der Beschäftigungsförderung erforscht werden. Dazu sind Informationen auf unterschiedlichen Aggregationsebenen zu erheben. Es kann zwischen der Mikro-Ebene (Bürger-Ebene), der Meso-Ebene (Projektebene) und Makro-Ebene (Stadtebene) unterschieden werden:

Untersuchungsebenen:		
Mikro-Ebene (Bürger-Ebene): <ul style="list-style-type: none"> • Individualbefragung von Arbeitslosen • Individualbefragung von Beschäftigten 	Meso-Ebene (Projektebene): <ul style="list-style-type: none"> • Erhebung von Projektdaten z.B.: Maßnahmekosten, Wertschöpfung durch Arbeit) • Erhebung von Aggregatdaten auf Projektebene z.B. Lohnsteueraufkommen der Maßnahme 	Makro-Ebene (Stadtebene): <ul style="list-style-type: none"> • Erfassung von Gebührensätzen / Leistungsentgelten • Erfassung kommunaler Finanzierungsanteile

Quelle: Trube Achim (1995) „Fiskalische und soziale Kosten-Nutzen-Analyse örtlicher Beschäftigungsförderung“

Untersuchungsmethoden

Je komplexer die Ausgangsfragestellung ist, desto vielfältiger gestalten sich die Methoden zur Daten- und Informationsbeschaffung, um in angemessener und zuverlässiger Weise das notwendige empirische Material zur Überprüfung der Kosten-/Nutzen-Hypothesen zusammenzutragen. Eine Mehr-Methoden-Konzeption lässt sich wie folgt systematisieren:

Methoden	Einsatzbereiche	Themenbereiche
Interviews anhand von standardisierten Fragebögen	Einzelbefragung von <ul style="list-style-type: none"> • Arbeitslosen Sozialhilfeempfängern • Vormalen arbeitslosen Sozialhilfeempfängern • Teilnehmer in speziellen Projekten örtl. Beschäftigungsförderung (Projektgruppe) • Arbeitslose aus Arbeitslosentreffen (Kontrollgruppe) 	<ul style="list-style-type: none"> • Demographische Merkmale • Ausbildung, Beruf, Arbeitslosigkeit • Einkommen, Schulden, materielle Situation • Gesundheitliche Situation • Soziale Situation • Psychische Situation • Gesellschaftspolitische Situation
Unstandardisierte Experteninterviews	<ul style="list-style-type: none"> • Einzelbefragung von Mitarbeitern aus 	<ul style="list-style-type: none"> • Kostenaufwand der Kommune

	<p>Beschäftigungsprojekten und von Vertretern der Projektträger</p> <ul style="list-style-type: none"> • Einzelbefragung von Mitarbeitern der Kommunalverwaltung 	<ul style="list-style-type: none"> • Kostenaufwand für örtliche Sozialleistungen • Kosten der Kommune bei Vergabe von öffentl. Aufträgen an Beschäftigungsprojekte • Leistungsentgelte/ Gebühren der Kommune • Kommunale Einnahmefekte durch Steuern • Einsparungen örtl. Sozialleistungen • Wertschöpfungseffekte
Teil-standardisierte Intensiv-Interviews (anhand eines Leitfadens)	<ul style="list-style-type: none"> • Einzelbefragungen von Personen in Maßnahmen der Beschäftigungsförderung • Einzelbefragung von Langzeitarbeitslosen mit und ohne Aussicht auf Beschäftigung 	<ul style="list-style-type: none"> • Arbeitslosen-/Beschäftigten-Biographien • Integrationsformen und -verläufe • Verarbeitungsmuster von Arbeitslosigkeit und Beschäftigungsförderung
Sekundärstatistische Erfassungen	<ul style="list-style-type: none"> • Bewilligungsstellen des örtlichen Sozialhilfeträgers • Projekte Wirtschaftspläne, Statistiken..) 	<ul style="list-style-type: none"> • Laufende und einmalige Sozialleistungen der Kommune • Finanzierung, Steueraufkommen, Kalkulationen von Projekten
Dokumenten-/ Inhaltsanalyse	<ul style="list-style-type: none"> • Projekte (Konzepte, Geschäftsberichte) • Städtische u. öff. Stellen (Verordnungen, Förderrichtlinien) 	<ul style="list-style-type: none"> • Aufbau, Merkmale, Leistungen • Kommunale Leistungsentgelte u. Gebühren • Auftragsvergabe/ Auftragskonditionen • Löhne/ Gehälter, Steuersätze • Ziele, Umfang, Träger

Quelle: Trube Achim (1995) „Fiskalische und soziale Kosten-Nutzen-Analyse örtlicher Beschäftigungsförderung“

2.2.1.5 Grenzen und Möglichkeiten

Die Fiskalanalyse stellt durch die objektive Bewertung der Maßnahme und der expliziten Berücksichtigung von Annahmen eine wertvolle Erweiterung der Wirkungsanalyse dar. Doch zugleich handelt es sich dabei um ein enges Konzept, da nur eng definierte Alternativen (meist die Einführung der Maßnahme im Vergleich um Zustand ohne Maßnahme) verglichen werden können.

Die Qualität der Fiskalanalyse hängt wesentlich von den Informationen über die Wirkungen der Maßnahme ab. Daher sollten komplexere Maßnahmen auch nur in Verbindung mit einer umfassenderen Evaluation (zumindest Wirkungsanalyse, aber besser auch Policy-Evaluation, Implementationsanalyse) durchgeführt werden.

Die Probleme der Bewertung, Diskontierung und Aggregation können teilweise durch unterschiedliche Szenarien einer Sensitivitätsanalyse gelöst werden. In die Sensitivitätsanalysen sollten unterschiedliche Szenarien zur Wirkung der Maßnahme, unterschiedliche Diskontraten, Grade der Nachhaltigkeit der Wirkungen, Rationalisierungsmaßnahmen etc. eingehen. So können jene Faktoren identifiziert werden, deren Veränderung die größte Auswirkung auf die Fiskalanalyse zeigt.

2.2.2 Ein konkretes Beispiel: Fiskalanalyse der arbeitsmarktpolitischen Maßnahme „Office Management Centre“

Im Rahmen des Projekts „Evaluation des abz wien“, ein arbeitsmarktpolitischer Träger in Wien, der verschiedene Maßnahmen zur Erleichterung des Wiedereinstiegs von Frauen in den Arbeitsmarkt anbietet, wurde für eine der Maßnahmen eine Fiskalanalyse durchgeführt. Bei der Maßnahme („**abzwien.office management center (omc)**“) handelt es sich um eine Kombination aus Qualifizierungs- und Beschäftigungsprojekt, d.h. die teilnehmenden Frauen sind ein Jahr lang im abz beschäftigt, arbeiten dort an Projekten und werden für die Erfordernisse moderner Büroorganisation weiterqualifiziert.

Die Besonderheit der Evaluation liegt darin, dass das Design sowohl qualitative als auch quantitative Methode integriert. Damit umfasst die Studie die folgenden Teile:

- Als Hintergrund und Einstieg dient eine kurze Analyse der Arbeitsmarktsituation für Frauen in Österreich und Wien, es folgt eine Kurzcharakterisierung des politischen Systems sowie eine Übersicht über die aktuelle Arbeitsmarktpolitik in Österreich und Wien.
- Im Anschluss daran gilt es, abz wien und zwar sowohl als gesamte Organisation als auch für die einzelnen Maßnahmen zu diesem Kontext in Bezug zu setzen. Dabei kristallisierten sich Konsens, kritische Punkte, aber auch offene Fragen heraus, die – so

hoffen wir – Basis für eine weitere Auseinandersetzung und Kommunikation zwischen AMS und abz wien sein können. Die vielen und vielfältigen Maßnahmen des *abz wien* wurden dann auch im einzelnen einer qualitativen Analyse unterzogen, bei der auch einige Firmen zu Wort kamen.

- Der quantitative Teil der Evaluation analysiert zunächst Beschäftigungsverläufe auf Grundlage eines 10jährigen Beobachtungszeitraums der Karrieren aller Kundinnen aus zwei zentralen Maßnahmen des abz wien: Aus dem **abzwien.office management center** (omc) und aus den **abzwien.orientierungskursen** (bo). Die Studie schließt mit einer Fiskalanalyse für das *omc*.

2.2.2.1 Datenmaterial und Methodik der Fiskalanalyse

In einem ersten Schritt werden die zu berücksichtigenden Kosten und Erträge identifiziert. Zu den fiskalischen **Kosten** zählen

- personenbezogene Kosten, welche durch Zahlungen an die Kurs- bzw. Projektteilnehmerinnen entstanden sind,
- weiterhin die Kosten für Schlüsselkräfte und Trainerinnen, die direkt der Maßnahme zuzurechnen sind und die anteiligen Overheadkosten (Personen- und Sachleistungen für die Administration der Maßnahme - soweit öffentlich finanziert).
- Ein dritter zu berücksichtigender Kostenfaktor sind die erworbenen Transferleistungen in Fällen, in denen eine teilnehmende Person neue Ansprüche an die Sozialversicherung erworben hat und nach ihrem Austritt aus dem Projekt höhere Transferleistungen als vor der Teilnahme bezieht.

Erträge für die öffentliche Hand resultieren aus

- den eingesparten Transferleistungen, die ohne die Maßnahme zu zahlen gewesen wären sowie aus
- zusätzlichen Rückflüssen an die öffentliche Hand in Form von Steuern und Abgaben, welche bei gestiegenen Einkommen der Teilnehmerinnen zustande kommen können.

Wichtig ist, dass Erträge nur dann in die Kalkulation mit aufgenommen werden können, wenn sie mit hoher Wahrscheinlichkeit der Maßnahme kausal zuzurechnen sind. Dies ist immer dann schwierig, wenn - wie hier der Fall - aufgrund fehlender Kontrollgruppe alle Annahmen darüber, was mit den Teilnehmerinnen ohne Maßnahme passiert wäre, auf arbiträren Annahmen basiert. Es wurde daher versucht, plausible Szenarien für den hypothetischen Karrierenverlauf der Teilnehmerinnen zu entwerfen, der zustande gekommen wäre, hätten sie an der Maßnahme nicht teilgenommen.

Es ergeben sich damit die folgenden Kosten- und Ertragskomponenten:

Kosten

K1 Direkte Teilnehmerinnenkosten

K2 Direkte Projektkosten (Trainerinnen etc.)

K3 Overheadkosten
K4 Erworbene Transferleistungen
Erträge
E5 Transfereinsparungen
E6 Zusätzliche Rückflüsse aus Steuern und Abgaben

Die Ermittlung der jeweiligen Kosten- und Ertragsgrößen basiert auf

- einem Datensatz, der aus einer Kombination von Daten aus dem EDV-System des Arbeitsmarktservice und aus den Versicherungsverläufen des Hauptverbandes der Sozialversicherungsträger besteht. Diese Datenquellen beinhalten zum einen Grundcharakteristika von Personen und zum anderen deren Versicherungsverläufe.
- von AMS Wien zur Verfügung gestelltem Datenmaterial (Förderverträge etc.),
- von *abz wien* zur Verfügung gestelltem Datenmaterial (Genaue Aufstellung der Teilnehmerinnen, Aufschlüsselung der Overhead Kosten, etc.),
- sekundärstatistischem Material, das insbesondere zur Simulation einer Vergleichsgruppe herangezogen werden musste.

2.2.2.2 Schritt 1: Die Kalkulation von Bruttokosten und Bruttoerträgen

K1 Direkte Teilnehmerinnenkosten

Für die Ermittlung der direkten Teilnehmerinnenkosten werden die Jahresbruttogehälter herangezogen. Steuern und Sozialversicherungsbeiträge müssen abgezogen werden, da diese der öffentlichen Hand sofort wieder zurückfließen und daher nicht als Kosten geltend gemacht werden können.

$$K1 = \Sigma \text{ Jahresnetto Gehalt} = (\text{Bruttogehalt} * 14) - (\text{DN-Beitrag zur SV} * 14) - (\text{Lohnsteuer} * 12)$$

K2 Direkte Projektkosten

Zu den direkten Projektkosten zählen zum einen Personalkosten für Schlüsselkräfte - soweit dem Bereich OMC zurechenbar, Ausbildungskosten und Sachkosten.

$$K2 = \text{Personalkosten Schlüsselkräfte netto} + \text{Ausbildungskosten netto} + \text{Sachkosten netto}$$

K3 Overheadkosten

Die Overheadkosten gliedern sich in Personal- und Sachkosten.

Die Personalkosten betreffen ausschließlich Schlüsselkräfte, bei den Sachkosten wurde wiederum ein durchschnittlicher MWST-Steuersatz von 15% angenommen.

$$K3 = \text{Overhead Personalkosten netto} + \text{Overhead Sachkosten netto}$$

K4 Erworbene Transferleistungen

Erworbene Ansprüche treten auf, wenn die Teilnehmerinnen nach der Maßnahme Transferleistungen in Form von Arbeitslosengeld oder Notstandshilfe beziehen, und sich den Anspruch auf diese Leistungen durch die Teilnahme erwarben. Zur Ermittlung dieser Fälle wurde die Zahl der Frauen berechnet, die

- in den Jahren 1996-1998 an OMC teilnahmen,
- vorher keinen Anspruch auf Arbeitslosengeld oder Notstandshilfe hatten
- und im Jahr nach der Maßnahme Arbeitslosengeld bzw. Notstandshilfe bezogen.

In die Berechnung einbezogen wurde jeweils nur ein Jahr nach Maßnahmenende. Arbeitslosengeld oder Notstandshilfebezüge im zweiten Jahr nach der Maßnahme können dem während des OMC Besuchs erworbenen Anspruchs nicht mehr eindeutig zugerechnet werden, da in allen Fällen neue Zeiten der Erwerbstätigkeit dazwischenlagen.

$$K4 = \Sigma \text{AL Tage im Jahr nach der Maßnahme} * \text{öS } 202,20$$

E5 Transfereinsparungen

Die Ersparnis von Arbeitslosenunterstützung oder anderen Transferleistungen (z.B. Notstandshilfe) welche von den Teilnehmerinnen vor dem Projekt bezogen wurden, wird als Ertragsfaktor berücksichtigt. Eingesparte Transferleistungen bedeuten insofern Opportunitätserträge. Zur Ermittlung der Anzahl der Teilnehmerinnen, bei denen vor der Maßnahme Transferzahlungen vorlagen, wurde lediglich der unmittelbare Status vor der Maßnahme untersucht:

$$E5 = \Sigma (\text{Maßnahmendauer} * \text{öS } 250,-) * 42\%.$$

E6 Zusätzliche Rückflüsse an Steuern und Abgaben

Wenn die teilnehmenden Personen Einkommenssteigerungen zu verzeichnen hatten, dann kann davon ausgegangen werden, dass sich in der Folge erhöhte öffentliche Einnahmen aus Steuern und Abgaben ergaben. Diese zusätzlichen Rückflüsse betreffen im wesentlichen indirekte Steuern aus dem Konsum eines erhöhten Einkommens sowie direkte Steuern und Sozialversicherungsabgaben.

2.2.2.3 Schritt 2: Modellierung eines Referenzszenarios

Bislang unberücksichtigt in der Darstellung blieben Mitnahmeeffekte. Diese sind als jener Teil der Erträge zu verstehen, die auch ohne die Maßnahme zustande gekommen wäre und daher der Maßnahme nicht zugerechnet werden dürfen. So hätte auch ohne OMC ein gewisser Prozentsatz der teilnehmenden Personen eine Beschäftigung gefunden. Diese Personen haben zwar Kosten verursacht, die Erträge können aber nicht als Folge des Projektes interpretiert werden und müssen daher abgezogen werden.

Die Berechnung von Mitnahmeeffekten ist immer fehlerbehaftet, da ja die hypothetische Situation „was wäre wenn genau die Teilnehmerinnen nicht teilgenommen hätten“ nicht messbar ist. Eine Kontrollgruppe, die mit den Teilnehmerinnen in wesentlichen Merkmalen übereinstimmt, kann als beste Annäherung an diese hypothetische Situation verstanden

werden. Liegt keine Kontrollgruppe vor – wie es hier aufgrund der technischen Unmöglichkeit eine Kontrollgruppe über einen so langen Maßnahmenzeitraum mit gleichzeitig vergleichsweise wenigen Teilnehmerinnen zu berechnen der Fall ist – behilft man sich mit möglichst gutem sekundärstatistischem Datenmaterial sowie mit Vorher-Nachher-Vergleichen.

Als **Vergleichsgruppe** werden hier langzeitarbeitslose Frauen in Wien mit Mobilitätseinschränkungen als Vermittlungerschwernis herangezogen. Aus der Teilnehmerinnenstatistik ist bekannt, dass diese Gruppe an Frauen im Durchschnitt schlechter qualifiziert – im Sinne der höchsten abgeschlossenen Ausbildung – ist als die Teilnehmerinnen. Die hier vorgenommene Kalkulation unterschätzt daher möglicherweise das wirkliche Ausmaß an Mitnahmeeffekten. Auf der anderen Seite kommen etwa die Hälfte aller Frauen direkt „Out of Labour Force“ in die Maßnahme – d.h. sie waren meist sehr lange Zeiten sehr weit weg vom Arbeitsmarkt, und für diese Gruppe ist der Weg in die Beschäftigung ausgesprochen schwer.

Folgende Daten sind für die Vergleichsgruppe von Belang:

- Das Verhältnis der Abgänge vom Bestand (hier liegt sekundärstatistische Information vor)
- Die durchschnittliche Beschäftigung im Jahr der Maßnahme, die auch ohne Maßnahme zustande gekommen wäre (hier lagen keine sekundärstatistischen Informationen darüber vor, wie lange Personen, welche aus der Arbeitslosigkeit in Beschäftigung abgehen, durchschnittlich beschäftigt sind, daher wurden Daten aus den Vorkarrieren der betrachteten Frauen herangezogen)

2.2.2.4 Schritt 3: Szenario nach Jahren:

In einem nächsten Schritt mussten die Zeiten für Beschäftigung, für Arbeitslosigkeit (AL) und für Out of Labour Force (OLF) bei Fortschreibung der oben ermittelten Abgangswahrscheinlichkeiten für die Folgejahre berechnet werden.

	Beschäftigung	AL	OLF
Jahr 1	97	157	113
Jahr 2	123	121	121
Jahr 3	141	112	112

Die Tabelle zeigt die resultierenden Zeiten an Beschäftigung, Arbeitslosigkeit und OLF des Referenzszenarios unter den folgenden Annahmen:

- von den 28,8% Arbeitsaufnahmen bleibt 1/3 über das nächste Jahr in Beschäftigung (Beschäftigungstyp 1), aus diesen Beschäftigungsverhältnis fallen 18% nach einem Jahr wieder in unregelmäßige Beschäftigung²,
- 1/3 der Beschäftigungsverhältnisse wird das nächste Jahr nicht überdauern (Beschäftigungstyp 2), sind aber noch aufrecht ,
- das restliche Drittel teilt sich gleichermaßen auf OLF und AL auf (50:50).
- Der Einfachheit halber wird zusätzlich angenommen, dass Personen in kurzen Beschäftigungsverhältnissen (Beschäftigungstyp 2), OLF und AL die gleiche Beschäftigungserwartung für das Jahr haben

Das Referenzszenario beschreibt, welchen Status die Frauen in den jeweiligen Jahren gehabt hätten, wären sie nicht in die Maßnahme gegangen. Für alle weiteren Berechnungen werden **Beschäftigung und auch Einkommen der Teilnehmerinnen als die zusätzliche Beschäftigung gegenüber dem Referenzszenario definiert.**

2.2.2.5 Schritt 4: Kalkulation der Nettokosten und Nettoerträge nach Jahren

Die Nettokosten und Nettoerträge ergeben sich aus den jeweiligen Bruttowerten unter Abzug der Kosten und Erträge des Referenzszenarios. Dabei wird wie folgt vorgegangen: Aus dem Datensatz der Teilnehmerinnen erhält man die Informationen über die tatsächlichen Beschäftigungszeiten. Diese Zeiten werden entsprechend zu den Zeiten der Vergleichsgruppe in Bezug gesetzt.

Die entsprechenden Daten der **Teilnehmerinnen** sind:

	Beschäftigung	AL	OLF
Jahr 1	31	1	2
Jahr 2	252	77	26
Jahr 3	292	40	29

Entsprechend dieser Daten werden für jedes Jahr anfallende Kosten und Erträge berechnet, die Differenzen zu den Kosten und Erträgen des Alternativszenarios gebildet und aufsummiert.

Jahre der Maßnahme (= Jahre 1)

- Im Jahr 1 erfolgt eine Korrektur der direkten Teilnehmerinnenkosten um die Tage der durchschnittlichen Beschäftigung bereits im Jahr der Maßnahme, damit erhält man die Nettokosten für K1.

² Die Wahrscheinlichkeit, dass ein 1 Jahr dauerndes Beschäftigungsverhältnis das nächste Jahr überlebt beträgt etwa 82%.

- K2 und K3 bleiben unverändert
- E5 bleibt unverändert
- Bei E6 fallen lediglich indirekte Steuern an, da die in K1 berechneten direkten Teilnehmerinnenkosten bereits Nettoeinkommen darstellen. Im Hinblick auf die indirekten Steuern wird angenommen, dass das gesamte zusätzliche Einkommen (d.h. das Einkommen, welches über dem Einkommen des Referenzszenarios liegt) konsumiert wird und das durchschnittliche Konsumbündel mit einem mittleren MWST Satz von 15% belastet ist. Die indirekten Steuern berechnen sich demnach:

$$E6 = \Sigma (15\% * \text{zusätzliches Einkommen gegenüber der Gruppe ohne Maßnahme})$$

Jahr 2: Jahr nach der Maßnahme

Im Jahr nach der Maßnahme fallen Kosten nur noch in Form erworbener Transfers an (K4). Erträge entstehen dagegen durch erhöhte Beschäftigung von Maßnahmenteilnehmerinnen gegenüber dem Referenzszenario ohne Maßnahme. Diese setzen sich zusammen aus:

- Eingesparten Transferleistungen (Arbeitslosengeld oder Notstandshilfe), die im Fall geringerer Beschäftigungszeiten angefallen wären (=E5).
- Zusätzlich entstehen Rückflüsse aufgrund direkter Steuern sowie aufgrund der Besteuerung höherer Konsumausgaben infolge des höheren erzielten Einkommens während Zeiten der Beschäftigung (=E6). Wie bereits kurz erwähnt, entsteht hier das höhere Einkommen lediglich durch den Mengeneffekt, d.h. durch mehr Beschäftigung. Mangels genauerer Informationen wird in konservativer Schätzung unterstellt, dass die Durchschnittseinkommen bei Beschäftigung sich nicht voneinander unterscheiden.

Für E6 (zusätzliche Rückflüsse an Steuern und Abgaben) fallen hier sowohl direkte als auch indirekte Steuern an. Die MWST-Rückflüsse berechnen sich:

$$\text{MWST-Rückflüsse: (Differenz der Beschäftigungsdauer} * \text{Durchschnittseinkommen netto} + \text{Differenz der AL Dauer} * 250) * 15\%$$

Die direkten Steuern und Abgaben entsprechen

$$\text{LNK} = 21,65\% * \text{Bruttoeinkommen (Sozialversicherung} + 3\% * \text{Bruttoeinkommen (Kommunalsteuer} + 4,5\% * \text{Bruttoeinkommen (FLAF) = Dienstgeberabgabe}$$

$$\text{SVDN} = \text{Bruttoeinkommen} * 17,6\% \text{ (Dienstnehmerbeitrag zur Sozialversicherung)}$$

$$\text{LSt} = \text{üblicher Steuersatz auf (Bruttoeinkommen} - \text{SVDN) = Lohnsteuer}$$

$$\text{Steuern und Abgaben: LNK} + \text{SVDN} + \text{LSt}$$

Um Nettoerträge zu erhalten, werden diese Rückflüsse wiederum nur für jene Zeiten berechnet, in denen die Frauen mit Maßnahme mehr in Beschäftigung waren als die Frauen ohne Maßnahme.

Jahr 3: Zwei Jahre nach der Maßnahme

Die Kalkulation der Nettoerträge für zwei Jahre nach der Maßnahme funktioniert analog: Aus dem Datensatz erhält man wieder die Informationen über die tatsächlichen Beschäftigungszeiten, wobei die Frauen aus dem Maßnahmenjahrgang 1998 hochgerechnet wurden. Diese Zeiten werden entsprechend zu den Zeiten der Vergleichsgruppe in Bezug gesetzt, Kosten fallen ab diesem Jahr keine mehr an.

Aufgrund der gegenüber dem Jahr zuvor gestiegenen Beschäftigung auf 292 Tage (=80%) fallen hier für alle Kategorien höhere Beträge an als in der Kalkulation für das Jahr 2 (Jahr nach Maßnahme), obwohl der Anteil der Beschäftigungstage auch in der Vergleichsgruppe angestiegen ist.

Hochrechnung für Folgejahre

Um für ein weiteres Jahr kalkulieren zu können, werden für die Teilnehmerinnen die gleichen Zeiten herangezogen, für die Vergleichsgruppe wird ein weiterer Beschäftigungsanstieg gemäß der oben 9.2.8 definierten Annahmen kalkuliert. Damit ergeben sich wieder leicht geringere Erträge, da nur für die Vergleichsgruppe ein weiterer Beschäftigungsanstieg angenommen ist, nicht aber für die Teilnehmerinnen.

2.2.2.6 Schritt 5: Aggregation und Diskontierung der Nettokosten und Nettoerträge

Unter diesen Annahmen ergeben sich gesamte Kosten und Erträge für 4 Jahre. 1997 steht hier für das Jahr der Maßnahme und bildet einen Durchschnitt aus den untersuchten Jahren 1996-1998. Gleiches gilt für 1998 und 1999: Auch hier verbergen sich jeweils Durchschnitte aus drei Jahren, für das Jahr 2000 gelten die im vorigen Abschnitt definierten Bedingungen.

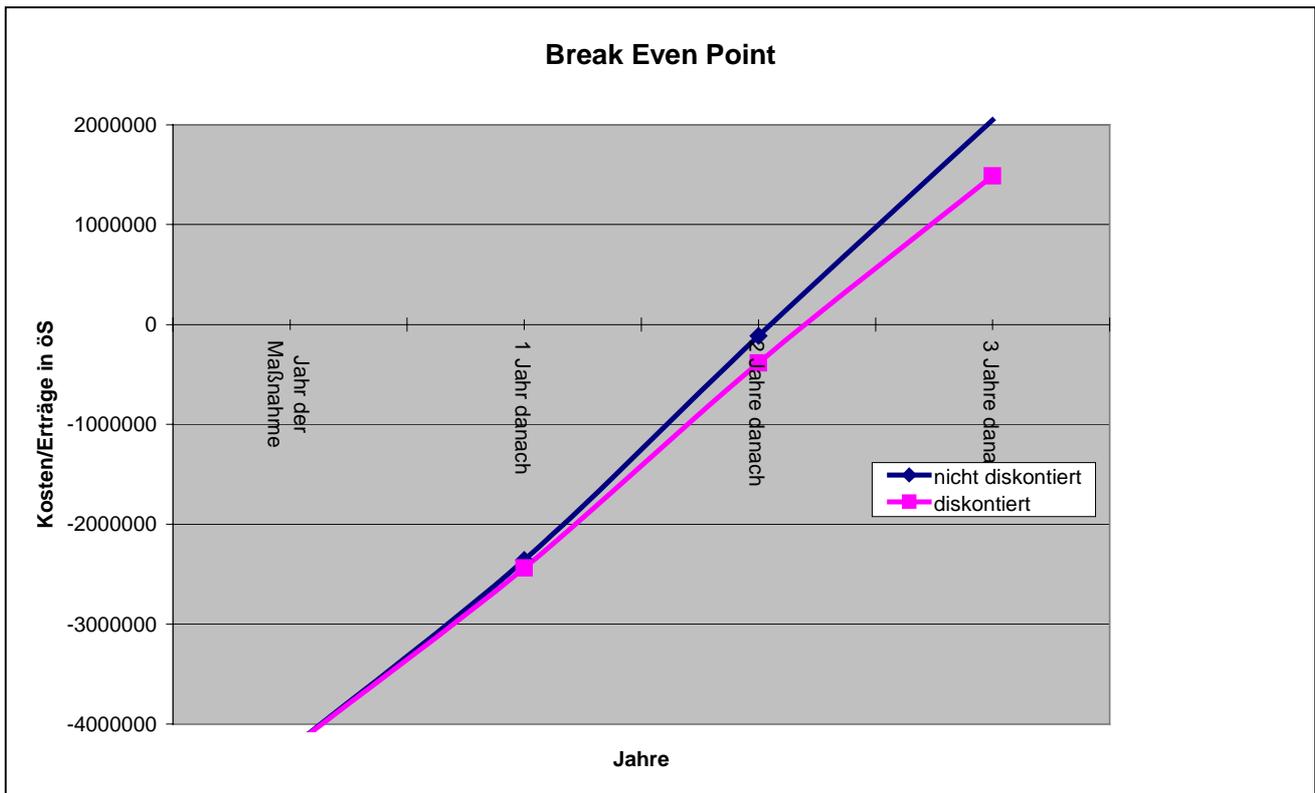
Kosten und Erträge in den Jahren 1998 bis 2000 wurden auf den Gegenwartswert von 1997 abdiskontiert. Grundlage der Diskontierung sind die Sekundärmarktrenditen (Bundesanleihen mit einer Laufzeit von 10 Jahren), wobei der Wert für das Jahr 2000 einer Prognose des WIFO entnommen ist²².

		1997	1998	1999	2000
Erträge	Eingesparte Transfers	1.677.513	248.685	322.193	308.381
	indirekte Steuern	112.264	153.418	154.140	148.936
	direkte Steuern		1.566.919	1.765.426	1.704.113
Summe		1.789.777	1.969.022	2.241.760	2.161.430
Kosten		6.034.264	81.271		

²² Demnach gilt für 1998 4,4%, für 1999 4,7% und für das Jahr 2000 5,6%.

Nettoströme	-4.244.487	1.887.751	2.241.760	2.161.430
kumuliert	-4.244.487	-2.356.737	- 114.977	2.046.453
diskontiert	-4.244.487	-2.436.297	- 385.409	1.487.126

2.2.2.7 Schritt 6: Berechnung des Break Even Points



Die Abbildung zeigt den Break Even kurz nach Ablauf des zweiten Jahres nach der Maßnahme, mit Diskontierung entsprechend etwas später. Das heißt, die Maßnahme amortisiert sich sehr schnell.

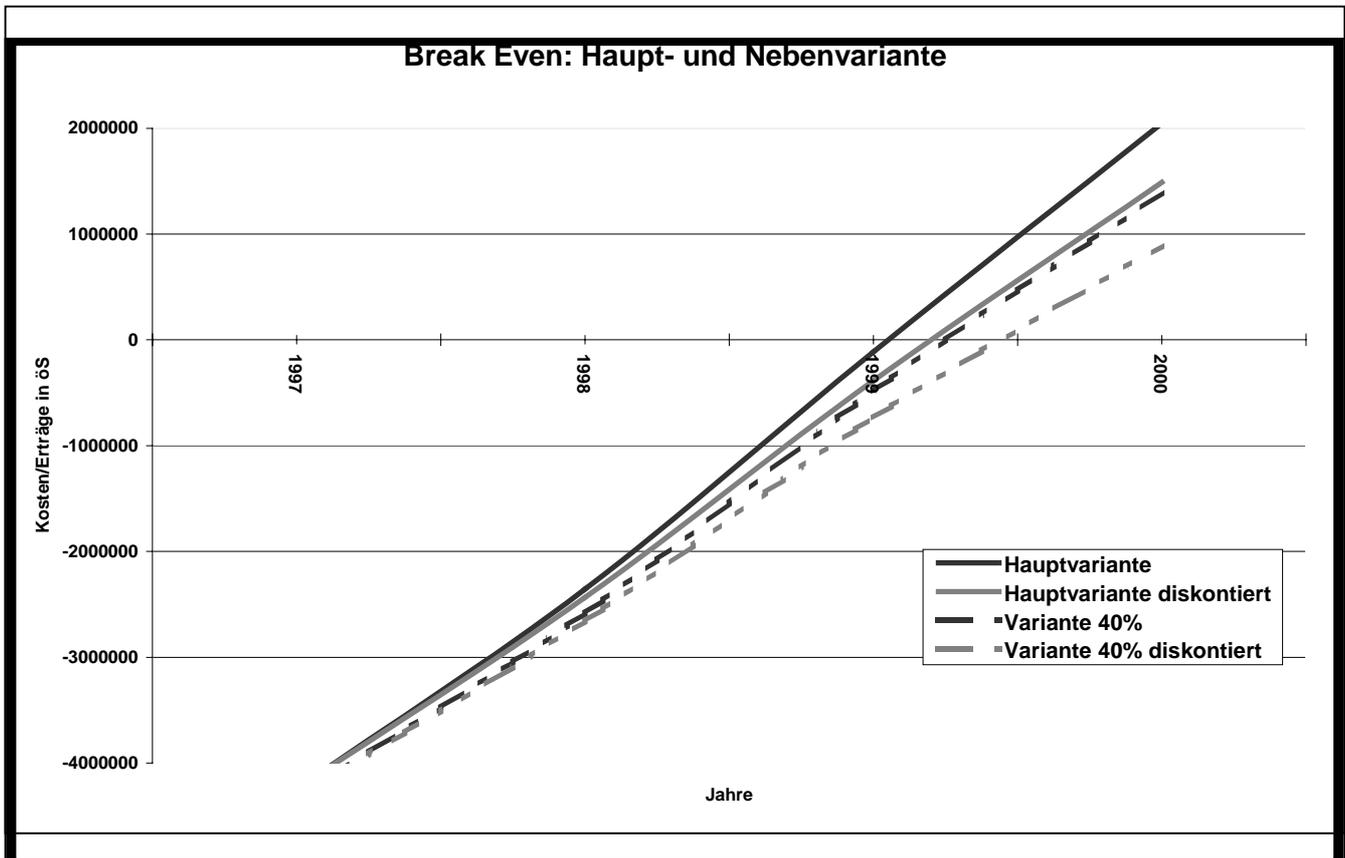
2.2.2.8 Schritt 7: Sensitivitätsanalyse

Um zu überprüfen, wie robust sich das Ergebnis gegenüber Änderungen zentraler Annahmen verhält, wird eine Nebenvariante gerechnet. Zuvor waren sekundärstatistische Informationen über die Gruppe langzeitarbeitsloser Frauen mit Mobilitätsbeschränkungen in Wien Grundlage für die Modellierung der Vergleichsgruppe. Nun sind aber die OMC Teilnehmerinnen im Durchschnitt besser qualifiziert als diese Gruppe und man kann aufgrund dessen annehmen, dass sie auch ohne Maßnahme bessere Beschäftigungschancen gehabt hätten als eben diese Gruppe.

Daher ist hier unterstellt, dass nicht wie zuvor 28,8% der arbeitslosen Frauen in Beschäftigung abgehen, sondern stattdessen 40%. Dies erhöht die Beschäftigungszeiten im Referenzszenario nicht unerheblich und mindert daher die Erträge aus der Mehrbeschäftigung der Teilnehmerinnen.

Allerdings zeigt die nachstehende Graphik, dass sich auch unter dieser Annahme der Break Even nicht wesentlich verändert. Auch in dieser Nebenvariante amortisiert sich die Maßnahme relativ schnell, der Break Even liegt mit diskontierten Daten immer noch im dritten Jahr nach der Maßnahme.

Dies heißt, obige Kalkulationen sind relativ robust gegenüber Änderungen der Annahmen, oder anders ausgedrückt: Selbst wenn das Klientel des OMC deutlich bessere Beschäftigungschancen hat als der Durchschnitt arbeitsloser Frauen mit Mobilitätsbeschränkungen, die Maßnahme lohnt sich in jedem Fall.



2.2.2.9 Zusammenfassung

Eine vollständige Fiskalanalyse wurde lediglich für das OMC gerechnet. War es angesichts einer fehlenden Kontrollgruppe bereits für das OMC schwierig genug, Erträge die nach Maßnahmenende anfallen, der Maßnahme eindeutig zuzurechnen - für die BO wäre es schlicht unseriös gewesen: Zielsetzung und Design der BO orientiert sich explizit nicht an einer möglichst schnellen Integration der Teilnehmerinnen in den ersten Arbeitsmarkt, dies wäre ein verfehltes Evaluationskriterium für den Maßnahmenerfolg. Vielmehr ist die BO eine Maßnahme, die sich hoch komplementär zu Folgemaßnahmen verhält, denn letztlich verbessert sie deren Erfolg. Der spätere Integrationserfolg – oder auch Misserfolg – ist in diesen Fällen nicht mehr eindeutig der BO oder der Folgemaßnahme zurechenbar.

Die Fiskalanalyse berechnet zunächst in der üblichen Weise die definierten Kategorien. Um mit dem Problem einer fehlenden Kontrollgruppe seriös umgehen zu können – was gerade angesichts dessen, dass hohe Mitnahmeeffekte immer als eines der Hauptargumente gegenüber den Leistungen des *abz wien* ins Feld geführt wird, besonders schwer wiegt, wurde auf Grundlage sekundärstatistischer Materials ein **Referenzszenario mit Hilfe einer Vergleichsgruppe** modelliert. Als Vergleichsgruppe werden hier langzeitarbeitslose Frauen in Wien mit Mobilitätseinschränkungen als Vermittlungerschwernis herangezogen. Für diese Gruppe wird kalkuliert- zunächst auf der Basis existierender Daten und in der Folge dann auf der Basis von Annahmen – wie hoch der Anteil der Frauen ist, der aus der Arbeitslosigkeit heraus in Beschäftigung abgeht und wie lange sie im Durchschnitt in Arbeit bleiben. Aufgrund der Besonderheit, dass viele Frauen im OMC nicht aus der Arbeitslosigkeit – im Sinne von Bezug von Arbeitslosengeld – sondern aus Out of Labour Force kommen, berücksichtigt das Referenzszenario auch Abgänge und Zugänge in Out of Labour Force.

Ergebnis des Szenarios ist eine Entwicklung des jeweiligen Verbleibs in den drei Kategorien Arbeitslos, Beschäftigung und Out of Labour Force wie folgt:

Dieses Referenzszenario wird der tatsächlichen Entwicklung gegenübergestellt, d.h. es werden für jedes Jahr anfallende Kosten und Erträge berechnet, die Differenzen zu den Kosten und Erträgen des Alternativszenarios gebildet und aufsummiert. Der den Verlaufsanalysen zugrundeliegende Datensatz gibt gesicherte Informationen für zwei Jahre nach Ende der Maßnahme. Für das erste Jahr nach der Maßnahme sind noch alle Frauen aus den hier betrachteten Kursen 1996 bis 1998 enthalten, schon im zweiten Jahr jedoch müssen die Frauen aus dem Maßnahmenjahr 1998 hochgerechnet werden.

Es war geplant, diese Daten über die Dauer in den jeweiligen Kategorien für die Folgejahre plausibel zu extrapolieren, die Ergebnisse auf Gegenwartswerte zu diskontieren und auf dieser Grundlage einen Break Even Point zu ermitteln. Diese Arbeitsschritte erübrigten sich allerdings weitgehend, da **bereits kurz nach Ablauf des zweiten Jahrs nach Maßnahmenende Kosten und Erträge der Maßnahme ausgeglichen waren**. Dennoch wurde ein zusätzliches Jahr hochgerechnet und auf das Ausgangsjahr 1997 (als Durchschnitt aus 1996-1998) diskontiert.

Die wesentlichen Gründe für die sehr kurze Amortisationsdauer des Projekts sind:

- Die direkten Teilnehmerinnenkosten sind zwar für AMS in der Tat hoch, nach Abzug aller Steuern und Abgaben, die an die öffentliche Hand zurückfließen, aber durchaus moderat.
- Die Frauen befinden sich nach einem und vor allem nach zwei Jahren zu einem so hohen Ausmaß in Beschäftigung, dass relativ hohe Rückflüsse an Steuern und Abgaben resultieren.

Diese Ergebnis ist umso bemerkenswerter, als dass zwei Elemente des Analysedesigns die Erträge eher moderat schätzen ließen: Erstens wurde im Gegensatz zum sonst üblichen Vorgehen berücksichtigt, dass nicht für alle Frauen Alternativkosten in Höhe der ansonsten ausgezahlten Transfers (Arbeitslosengeld oder Notstandshilfe) anfallen. Dies deshalb, weil nicht nur zwischen Beschäftigung und Arbeitslosigkeit differenziert wurde, sondern auch die Kategorie Out of Labour Force einfließt. Zweitens wurde angenommen, dass die Frauen in der Vergleichsgruppe, wenn sie in Beschäftigung sind, das gleiche Einkommen beziehen, wie die Frauen aus dem OMC. Da die Durchschnittsverdienste dieser Frauen – wie wir aus der Verlaufsanalyse wissen - relativ hoch sind, werden hier die Erträge eher unterschätzt.

Um zu überprüfen, wie robust sich das Ergebnis gegenüber Änderungen zentraler Annahmen verhält, wurde eine **Nebenvariante** gerechnet. Zuvor waren sekundärstatistische Informationen über die Gruppe langzeitarbeitsloser Frauen mit Mobilitätseinschränkungen in Wien Grundlage für die Modellierung der Vergleichsgruppe. Nun sind aber die OMC Teilnehmerinnen im Durchschnitt besser qualifiziert als diese Gruppe und man kann aufgrund dessen annehmen, dass sie auch ohne Maßnahme bessere Beschäftigungschancen gehabt hätten als eben diese Gruppe. Daher wird für die Nebenvariante unterstellt, dass nicht wie zuvor 28,8% der arbeitslosen Frauen in Beschäftigung abgehen, sondern statt dessen 40%. Dies erhöht die Beschäftigungszeiten im Referenzszenario nicht unerheblich und mindert daher die Erträge aus der Mehrbeschäftigung der Teilnehmerinnen. Dennoch verschiebt sich der Break Even Point um nicht einmal ein halbes Jahr, er liegt immer noch im dritten Jahr nach der Maßnahme. Obige Kalkulationen verhalten sich damit relativ robust gegenüber Änderungen der Annahmen, oder anders ausgedrückt: Selbst wenn das Klientel des OMC deutlich bessere Beschäftigungschancen hat als der Durchschnitt arbeitsloser Frauen mit Mobilitätseinschränkungen, die Maßnahme lohnt sich in jedem Fall.

Dies heißt letztlich, dass trotz eines relativ konservativen Designs vieles darauf hindeutet, dass es sich beim **OMC um eine sehr effiziente Maßnahme für die öffentliche Hand handelt**. Im Vergleich zu anderen Evaluationsergebnissen arbeitsmarktpolitischer Maßnahmen errechnen wir hier mit zwei Jahren nach der Maßnahme einen **extrem frühen Break Even**.

3 KAPITEL DREI EVALUIERUNG VON NETZWERKEN: BEWERTUNG VON LERNEN UND KOOPERIEREN IM RAHMEN EINER CLUSTERORIENTIERTEN TECHNOLOGIEPOLITIK

Christian Hartmann

3 KAPITEL DREI	
EVALUIERUNG VON NETZWERKEN: BEWERTUNG VON LERNEN UND KOOPERIEREN	
IM RAHMEN EINER CLUSTERORIENTIERTEN TECHNOLOGIEPOLITIK	1
3.1 EINLEITUNG	2
3.2 DER DERZEITIGE STAND DER FORSCHUNG	2
3.2.1 <i>Probleme bei der Bewertung von Kooperationen</i>	3
3.3 DIE BEWERTUNG VON NETZWERKEN AM BEISPIEL DER CLUSTERBEOBACHTUNG IN DER STEIERMARK	4
3.3.1 <i>Einleitung</i>	4
3.3.2 <i>Die verwendete Methodik</i>	4
3.3.3 <i>Verwendete Indikatoren</i>	5
3.4 SCHLUßFOLGERUNGEN	10
3.5 LITERATUR.....	11

3.1 Einleitung

Das Projektmodul „Evaluierung von Netzwerken“ hat sich zum Ziel gesetzt, einen Indikatorenrahmen zur Beurteilung des Kooperations- und Lernverhaltens von Betrieben zu erarbeiten, bzw. bestehende Projekte in diesem Bereich auf ihren Erfahrungshintergrund hin zu untersuchen. Ein weiteres Ziel ist dabei die Ableitung von Schlußfolgerungen hinsichtlich der zukünftigen Konzeption von Netzwerkevaluierungen.

In einem ersten Schritt wird dazu eine ausführliche Sichtung der vorhandenen Literatur vorgenommen, wobei vier Initiativen und deren Evaluierungsansatz als wesentlich anzusehen sind und daher im Detail vorgestellt werden. Ein zweiter Schritt dient dann dazu, anhand der Clusterbeobachtung in der Steiermark eine mögliche Methodik für die Evaluierung von Netzwerkaktivitäten vorzustellen und in ihren Ergebnissen kurz zu diskutieren. Eine abschließende Betrachtung soll dann anhand von Schlußfolgerungen mögliche Hinweise für die zukünftige Gestaltung von Netzwerkbewertungsprojekten anarbeiten.

3.2 Der derzeitige Stand der Forschung

Die Förderung und damit verbunden die Evaluierung von Netzwerken und Kooperationen ist ein relativ neues Thema innerhalb der Technologiepolitik. Erst seit Beginn der 90er Jahre sind gezielte Programme und Strategien zur Stimulierung und Unterstützung von Kooperationen (insbesondere zwischen KMU) vorzufinden. Dementsprechend gering ist daher auch der verfügbare Bestand an Forschungsarbeiten zur Bewertung von Netzwerken und Kooperationen auch wirtschaftspolitischer Sicht. Die derzeit verfügbaren relevanten Arbeiten lassen sich auf vier Projekte bzw. Untersuchungen reduzieren:

- **Die dänische Netzwerkprogramm Initiative:** Dieses Programm stellt nicht nur den ersten Versuch dar, Netzwerke aktiv über den Broker-Ansatz zu stimulieren¹, sondern bietet durch die Zwischenevaluierung der Initiative von Gelsing im Jahr 1993² auch bis dato eine der umfassendsten Bewertungen eines derartigen Förderungsprogramms. Neben der Nachhaltigkeit der Netzwerke wurden die Wirkungen der Netzwerke auf die Kosten, den Umsatz, Produktinnovationen sowie die internationale Wettbewerbsfähigkeit der Betriebe untersucht.

¹ Rosenfeld, S. A. (1996): Does cooperation enhance competitiveness? Assessing the impacts of inter-firm collaboration. *Research Policy* 25, 247-263.

² Gelsing, L. (1993): Evaluating programs promoting networks. In: *Significant Others: Background Papers. Regional Technology Strategies*, Chapel Hill. NC.

- **Die Netzwerkinitiative des Staates Oregon:** Nach dänischen Vorbild wurde 1991 in Oregon eine ähnliche Förderungsinitiative gestartet, welche dann 1993 von Rosenfeld evaluiert wurde. Die Bewertung musste sich dabei vor allem auf die technische und organisatorische Umsetzung des Programms konzentrieren – ökonomische Wirkungen konnten nur in Gestalt der Erwartungen der Beteiligten abgefragt werden.
- **Die Netzwerkinitiative der Nothwest Area Foundation:** Die Initiative, die 1990 gestartet wurde, zielte anstelle eines aktiven Brokeransatzes auf eine Netzwerkentwicklung durch die Schaffung entsprechender Rahmenbedingungen ab. Das Projekt wurde dabei von Rosenfeld begleitet, der 1993 eine abschließende Evaluierung erstellte. Es wurde dabei fünf Fragekomplexe bearbeitet: 1. Der Prozess der Netzwerkformierung, 2. die positiven Effekte für die beteiligten Betriebe (Umsatz, Produktivität, Wettbewerbsfähigkeit), 3. Regionale Effekte, 4. die Nachhaltigkeit der Netzwerke und 5. Der Einfluss auf qualitative Indikatoren wie z.B. die Arbeits- und Lebensqualität der Betroffenen.
- **USNet:** Ausgehend von der gleichnamigen Initiative, die 1994 von Institutionen der nationalen Innovationsinfrastruktur ins Leben gerufen wurde, hat Shapira 1998³ eine abschließende Evaluierung des Programms durchgeführt. Die Untersuchung hatte dabei drei zentrale Schwerpunkte: 1. die Effektivität der von USNet durchgeführten Aktivitäten bzw. der angebotenen Dienstleistungen, 2. der Impact auf die Kompetenzen, Strategien und Leistungsangebote der beteiligten Institutionen und 3. Der Impact auf kooperierende KMU (Strategien, Wettbewerbsfähigkeit und wirtschaftliche Performanz).

3.2.1 Probleme bei der Bewertung von Kooperationen

Alle Evaluatoren, die bisher Netzwerkentwicklungen untersucht und bewertet haben geben an, dass diese Arbeit mit massiven Problemen behaftet ist. Diese Schwierigkeiten umfassen dabei die folgenden Bereiche:

- **Langfristigkeit des Beobachtungszeitraums:** Alle bisher bewerteten Netzwerkinitiativen konnten nur über einen Zeitraum von drei Jahren beobachtet werden. Längerfristige Effekte, seien dies Indikatoren der betrieblichen Wettbewerbsfähigkeit oder seien dies qualitative Indikatoren, können innerhalb dieses Zeitraums nur unvollkommen erfasst und bewertet werden. Darüber hinaus besteht die Möglichkeit, dass bestimmte Fragen wie etwa die Nachhaltigkeit der Netzwerkentwicklung erst lange nach dem Ende der Initiative wirklich bewertet werden könnten.
- **Bewertung von schwer erfassbaren Effekten wie Lernen:** Bestimmte Effekte sind mit herkömmlichen Methoden der Evaluierung und des Monitorings nur sehr

³ Shapira, P. (1998): The Evaluation of USNet: Overview of Methods, Results and Implications. Final Report. Regional Technology Strategies, Chapel Hill. NC.

unvollkommen oder fast gar nicht erfassbar. Beispiele sind hier etwa der Wissenstransfer zwischen Organisationen oder die gemeinsamen Lernprozesse.

- **Diffuse Zielsetzungen:** Die meisten der untersuchten Netzwerkinitiativen verfügen über keinen klaren oder operationalisierbaren Zielkatalog. Es ist diesbezüglich zu vermerken, dass in den Arbeiten die Ziele der Förderprogramme, insbesondere in Bezug auf Art und Ausrichtung der angestrebten Netzwerke nicht systematisch dargestellt werden.

3.3 Die Bewertung von Netzwerken am Beispiel der Clusterbeobachtung in der Steiermark

3.3.1 Einleitung

Ähnlich wie in den weiter oben bereits dargestellten Netzwerkevaluierungsprogrammen ist das Ziel der Clusterbeobachtung in der Steiermark die Beurteilung der Entwicklung der betroffenen Cluster. Auch hier ist festzuhalten, dass eine klare Zielkonzeption für die Cluster von Seiten der Wirtschaftspolitik fehlte und somit auch nicht überprüft werden konnte. Was jedoch das steirische Beispiel von den dänischen und amerikanischen Projekten deutlich unterscheidet, ist die explizite Einbeziehung von Effekten wie Lernen. Auch hinsichtlich des Beobachtungszeitraumes bietet das steirische Beispiel einen längeren Zeithorizont – er liegt hier bei insgesamt fünf Jahren.

Im folgenden wird nun die spezielle Methodik der Clusterbeobachtung kurz vorgestellt, dann erfolgt eine Präsentation der verwendeten qualitativen Indikatoren und abschließend wird versucht allgemeine Schlußfolgerungen aus dem Vergleich der dänischen und amerikanischen Initiativen mit dem steirischen Projekt zu ziehen.

3.3.2 Die verwendete Methodik

Die steirische Clusterbeobachtung hat die Cluster (bzw. konstituierende Netzwerke) auf zwei Ebenen untersucht:

- Auf Ebene von Liefer- und Bezugsverflechtungen
- Auf Ebene der Kooperations-, Wissensaustauschs- und Lernbeziehungen

Einerseits sind tatsächliche erfolgte Lieferungen von intermediären Gütern- und Dienstleistungen zwischen Sektoren erfasst worden, andererseits wurden einzelbetriebliche Kooperationsformen, ihre Intensität und Qualität erhoben. Beide empirische Analysen beziehen sich auf Betriebe mit Standort in der Steiermark.

Es wird die Verknüpfung eines "top-down"- mit einem "bottom-up"-Ansatz geleistet und somit eine Kombination von aggregierten Wettbewerbsfaktoren (im Sinne von Porter) mit regionsspezifischen Kooperationsbeziehungen angeboten.

Ebene 1: Analyse über alle Wirtschaftssektoren – Liefer- und Bezugsverflechtungen

Als Basis dazu dient die Analyse von Input-Output-Verflechtungen, abgeleitet aus einer regionalen Input-Output-Tabelle, die auf Sekundärdaten und einer umfassenden Primärerhebung in der steirischen Wirtschaft basiert.

Fragestellungen:

- Welche Wirtschaftsbereiche sind für die Steiermark von besonderer Bedeutung?
- Wie stark sind die Lieferverflechtungen dieser Wirtschaftsbereiche?
- Welche ökonomisch-statistischen Cluster lassen sich daraus ableiten?

Ebene 2: Analyse von Clusterbeziehungen und -strukturen

Neben der Analyse über alle Wirtschaftssektoren konzentriert sich die Clusterbeobachtung auf die Detailanalyse von 5 Clustern. Hier werden folgende Fragestellungen mit einbezogen:

- Wie groß sind die Cluster (Zahl der Beschäftigten, der Betriebe, der Produktionswerte und der Produktion pro Beschäftigten) ?
- Wie haben sich die Cluster hinsichtlich der obigen Indikatoren in der Vergangenheit entwickelt (bis in die jüngere Vergangenheit hinein – Jahr 1998) ?
- Welche Kooperationsmuster finden sich in den Clustern? Dominieren etwa eher wissensintensive oder zwischenbetriebliche Kooperationen?
- Unterscheiden sich die Clusterunternehmen in ihrem Kooperationsverhalten von anderen Unternehmen? Sind sie kooperationsbereiter als der Durchschnitt der steirischen Industrie?
- Wie haben sich die Kooperationen in der Vergangenheit (fünf Jahre) entwickelt, mit welchen Partnern haben sich die Kooperationen intensiviert? Welche Kooperationsinhalte werden mit welchen Partnern gepflegt? Wie haben sich diese Inhalte in den vergangenen fünf Jahren entwickelt?
- Welche Strategien wären mit Hinblick auf die bestehenden Chancen und Gefahren sinnvoll?

3.3.3 Verwendete Indikatoren

Unter dem Begriff „Kooperation“ wird in diesem Kontext folgendes verstanden:

- Kooperation ist die auf Dauer angelegte aktive Zusammenarbeit zwischen zwei (oder mehreren) Partnern, wobei gemeinsame Ziele und klare Regeln zwischen allen Beteiligten gegeben sein müssen. Regelmäßige Geschäftsbeziehungen sind hingegen allein noch keine Kooperationen – hier fehlen klare gemeinsame Ziele und oftmals auch „Spielregeln“. Auch informelle Kontakte alleine reichen nicht aus, unter Kooperation wird eine also aktive Zusammenarbeit verstanden.

Konkrete Beispiele für Kooperationen sind etwa Einkaufsgemeinschaften, gemeinsame Marketing- oder F&E-Aktivitäten, Qualifizierungsverbände, gemeinsame Dachmarken etc.

In Bezug auf Kooperationen werden zwei besondere Schwerpunkte untersucht:

Die Analyse der Kooperationen und Partner

Sie zeigt einerseits die grundsätzliche Kooperationsneigung der Clusterbetriebe. Andererseits wird jedoch auch die räumliche Orientierung (regional/überregional) der Kooperationen untersucht. Die Analyse der Partner schließlich ermöglicht es, die Lernfähigkeit des Clusters darzustellen: Dominieren etwa wissensintensive Kooperationen (mit Universitäten und anderen F&E-Einrichtungen) so ist von einem inter-organisationalen Lernen im Cluster auszugehen. Umgekehrt zeigen intensive Kooperationsbeziehungen mit regionalen Betrieben eine starke regionale Einbettung in organisationaler Hinsicht an. Insgesamt wurden hier die folgenden Indikatoren zum Einsatz gebracht:

Räumliche Kooperationsorientierung	Zusammenarbeit <ul style="list-style-type: none"> - lokal - regional - überregional
Funktionale Kooperationsorientierung	Zusammenarbeit mit <ul style="list-style-type: none"> - Lieferanten - Kunden - Konkurrenzunternehmen - Firmen anderer Branchen
Kooperationsintensität	Zusammenarbeit <ul style="list-style-type: none"> - permanent - oft - manchmal
Wissensintensive Kooperationen	Zusammenarbeit mit <ul style="list-style-type: none"> - Universitäten

	<ul style="list-style-type: none"> - Fachhochschulen - Außeruniv. F&E-Einrichtungen
--	---

Die Analyse der Kooperationsinhalte

Hier steht zum einen die Betrachtung der jeweiligen Kooperationsgegenstände im Vordergrund, zum anderen soll jedoch auch dargestellt werden, wie sich diese Kooperationen in den vergangenen fünf Jahren entwickelt haben (gewachsen, gleichgeblieben, abgenommen). Insgesamt wurden hier die folgenden Indikatoren zum Einsatz gebracht:

Entwicklungsverläufe der Kooperationen	Zusammenarbeit hat in den letzten fünf Jahren <ul style="list-style-type: none"> - zugenommen - abgenommen - gleichgeblieben
Kooperationsinhalte	Zusammenarbeit im Bereich <ul style="list-style-type: none"> - Wissensaustausch - Pooling von betrieblichen Funktionen - Projektmanagement - Strategische Allianzen

Das Beispiel der steirischen Cluster in den Bereichen Automobil, Metall, Holz, IT und Pharma

Die Zusammenschau der fünf analysierten steirischen Cluster zeigt ein abwechslungsreiches Bild⁴. So weisen manche Cluster eine überaus starke positive Dynamik auf, während andere eher stagnieren oder sogar leicht schrumpfen. Zugleich gibt es Cluster, die aufs engste mit der regionalen Wissensinfrastruktur verbunden sind, aber wenig regionale Lieferverflechtungen besitzen und andererseits verfügen andere Cluster über hohe regionale Verflechtungen und wenig wissensintensive Kooperationsbeziehungen.

Abbildung 1 zeigt die untersuchten Cluster in der Zusammenschau ihrer jeweiligen Größe und Entwicklungsdynamik. Dabei wird einerseits die Entwicklung der

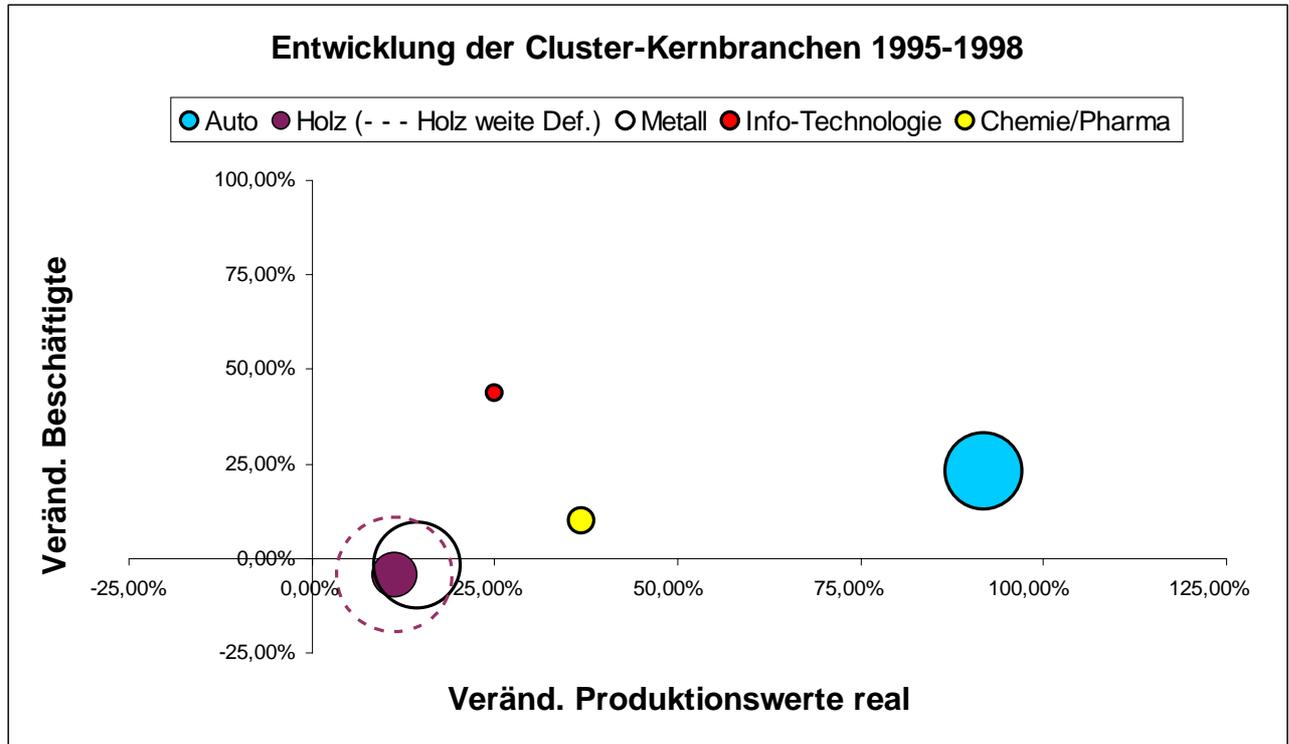
⁴ Die Clusteranalyse bezieht sich auf fünf Cluster, die im Detail untersucht wurden. Zwei weitere für die Steiermark relevante Branchenbereiche sind

- „Baustoffe“ (ÖNACE 26: v.a. Ziegeleien, Beton- und Zementindustrie)
- „Elektrotechnik/Elektronik“ (ÖNACE 30/32, 31, 33)

Über den angebotenen Leistungsumfang hinausgehend, werden zu diesen Branchenbereichen im Anhang einige wesentliche Eckdaten bereitgestellt, die ihre Größe, regionale Einbettung und Entwicklung beschreiben.

Beschäftigtenzahl sowie der Bruttoproduktionswerte gemeinsam dargestellt. Andererseits gibt die Fläche der Kreise jeweils die Bruttoproduktion im Jahr 1998 wieder.

Abbildung 1:

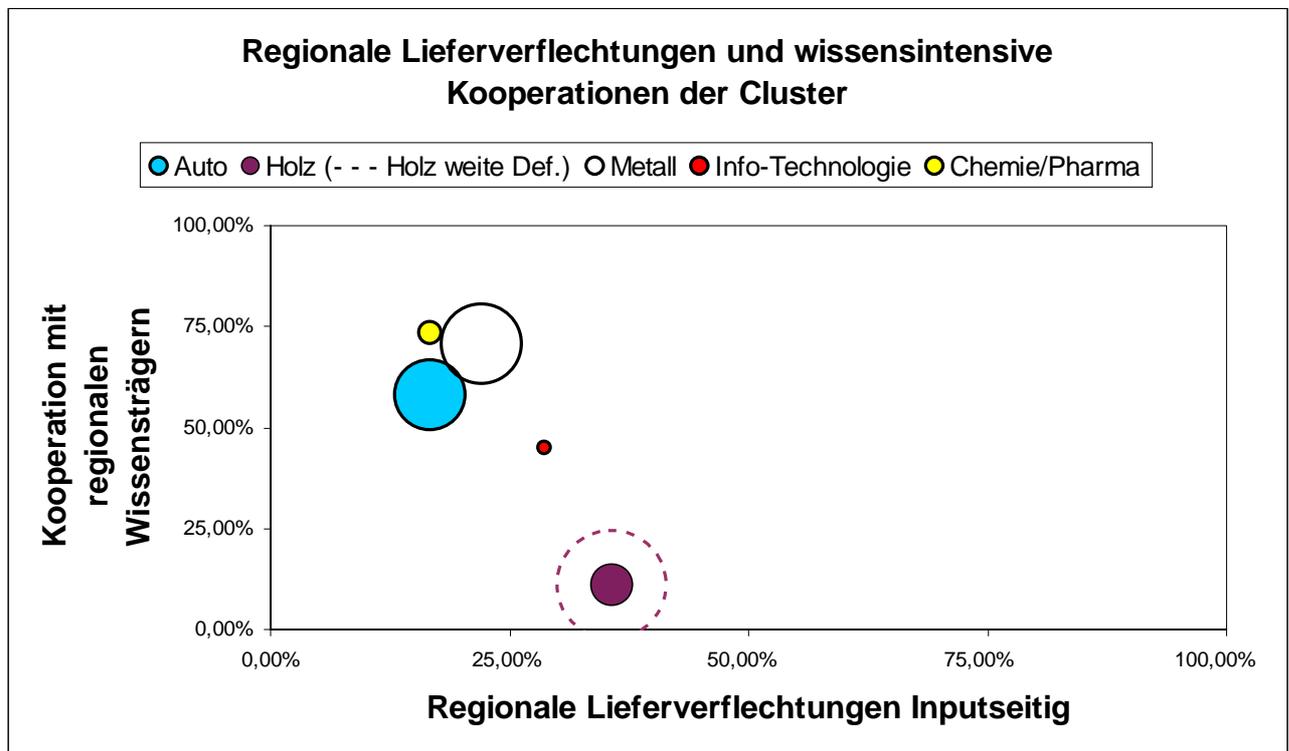


Die Entwicklung der Kernbranchen (Veränderung der Beschäftigten / reale Veränderung der Produktionswerte) zeigt, dass besonders der Automobilcluster in den vergangenen Jahren eine überaus beachtliche Dynamik aufweist: Von 1995 bis 1998 hat seine Beschäftigtenzahl um +23% zugenommen (Ø Österr. nur +6%), der Produktionswert steigerte sich sogar um 92% (Ø Österr. nur +25,6%). Doch auch der Chemie/Pharma Cluster weist mit einer Zunahme der Bruttoproduktion um +37% (Ø Österr. nur +9,5%) eine deutliche positive Entwicklung auf. Hinsichtlich seiner Beschäftigungsentwicklung liegt er mit einem plus von 10,3% sogar konträr zum österreichischen Trend (-2,8%). Der Metall-Cluster hingegen hat in den letzten Jahren an Beschäftigten verloren (-1,6%) und verliert dabei etwas stärker als der österreichische Durchschnitt (-0,8%). Auch das Wachstum der Bruttoproduktion ist im Metall-Cluster im Österreichvergleich unterdurchschnittlich (+14,5% gegenüber +20,2%). Der Holzcluster hat in den vergangenen Jahren mit -4,2% ebenfalls einen stärkeren Rückgang der Beschäftigten erlebt als der österreichweite Branchendurchschnitt (-1,2%). Jedoch verfügt er über ein im Österreichvergleich etwas stärkeres Wachstum der Bruttoproduktion (+11% gegenüber +8%). Der Informationstechnologiecluster weist von 1995 bis 1998 eine positive

Beschäftigungsentwicklung auf (+43,9%), er liegt damit jedoch etwas unter der gesamtösterreichischen Entwicklung (+50,6%).⁵

Eine Besonderheit der gegenständlichen Cluster-Untersuchung liegt darin, dass sowohl die „harten“ Lieferbeziehungen als auch die „weichen“ Kooperationsbeziehungen quantitativ erfasst werden. Abbildung 2 zeigt, in wie weit die untersuchten Cluster entweder über eine starke regionale Einbettung – also über hohe Inputs (Zulieferungen) aus der Steiermark – verfügen, oder aber über Kooperationen eng mit der regionalen Wissensinfrastruktur verbunden sind.

Abbildung 2:



Der Holzcluster verfügt von allen untersuchten Clustern über die stärkste regionale wirtschaftliche Einbettung. Seine Unternehmen beziehen 35,7% ihrer Vorleistungen aus der Steiermark. Zugleich besitzt dieser Cluster die geringste Kooperationsintensität mit regionalen Wissensträgern: Nur ca. 11% der Unternehmen geben an, mit Universitäten, Fachhochschulen und außeruniversitären F&E-Einrichtungen zusammenzuarbeiten. Der Informationstechnologiecluster weist ebenfalls eine gute regionalwirtschaftliche Einbettung auf. Er bezieht 28,7% seines Inputs von Unternehmen aus der Steiermark; zugleich arbeiten 45% der Clusterunternehmen nach eigenen Angaben regelmäßig mit regionalen Wissensträgern zusammen. Der Metall-Cluster ist etwas geringer regional eingebettet – 22% des Inputs für den Cluster stammen aus steirischen Betrieben.

⁵ Die Entwicklung der Bruttoproduktion im IT-Cluster kann aufgrund mangelnder Datengrundlagen im Dienstleistungssektor nur grob geschätzt werden, sie dürfte im Zeitraum 1995-1998 bei etwa +25% liegen.

Um so stärker ist hingegen die Präsenz wissensintensiver Kooperationen. Über 70% der Betriebe geben an, mit regionalen Wissensträgern regelmäßig zusammenzuarbeiten. Der Chemie/Pharma Cluster ist hinsichtlich wissensintensiver Beziehungen noch etwas stärker vernetzt: Über 73% geben an, regelmäßig mit der regionalen Wissensinfrastruktur zusammenzuarbeiten. Hinsichtlich der regionalwirtschaftlichen Einbettung ist festzuhalten, dass der Chemie/Pharma Cluster nur 16,7% seines Inputs aus der Steiermark bezieht und somit geringere wirtschaftliche Verflechtungen aufweist als die anderen Cluster. Gleiches lässt sich auch für den Automobilcluster sagen, der ebenfalls nur 16,7% seines Inputs aus der Region bezieht. Zugleich verfügt der Automobilcluster jedoch über intensive Kooperationen mit regionalen Wissensträgern: Über 58% der Unternehmen kooperieren regelmäßig mit Universitäten, Fachhochschulen oder außeruniversitären F&E-Einrichtungen.

3.4 **Schlußfolgerungen**

Aufbauend auf dem vorgestellten Beispiel bzw. den erhobenen Defiziten bestehender Evaluierung von Netzwerken lassen sich für die Neukonzeption solcher Projekte die folgenden Hinweise bzw. Schlußfolgerungen ableiten:

- **Gemischter Approach notwendig:** Die Evaluierung von Netzwerken und Clustern sollte immer sowohl einen quantitativen als auch einen qualitativen Ansatz umfassen. Beide Ansätze sollten dabei so kombiniert werden, dass sich synergistisch ergänzen. Im vorliegenden Beispiel stand dabei für die quantitative Erhebung eine Input-Output-Tabelle zur Verfügung, die noch mit Ergebnissen der schriftlichen Befragung ergänzt wurde. Erst das Zusammenspiel beider Perspektiven macht eine Bewertung der jeweiligen Initiative möglich.
- **Zielkonzeptionen bereits in Förderungsprogrammen operationalisieren:** Um die Vernetzung von Unternehmen untereinander und/oder mit Forschungseinrichtungen besser bewerten zu können, wäre eine operative Zielsetzung in den jeweiligen Programmen sinnvoll. Diese Zielsetzungen könnten dabei einerseits die Kooperationsorientierung (z.B. wissensintensiv) oder auch die Intensitäten (Anzahl permanent kooperierender Betriebe) umfassen. Mögliche Benchmarks sind hier ebenfalls festzulegen.
- **Beobachtungszeitpunkte:** Die meisten Evaluierungen von Netzwerkinitiativen setzen zeitlich gleich nach dem Ende des jeweiligen Programms an. Zumindest in Ergänzung dazu sollte jedoch eine weitere Erhebung etwa fünf Jahre nach Abschluß des Programms durchgeführt werden.
- **Lernen über mittelbare Indikatoren erfassen:** Lernprozesse lassen sich innerhalb von Befragungen nur sehr schlecht erfassen. Zugleich bieten aber mittelbare Indikatoren - die auch relativ leicht erfassbar sind - eine gute Möglichkeit, Netzwerk-Lernen zu erheben. Im vorliegenden Beispiel aus der

Steiermark wurde dabei die Erkenntnis, dass wissensintensive Kooperationen zu Lernprozesse führen, aktiv genutzt.

3.5 Literatur

Camagni R (1991): Introduction: from the local ‚milieu‘ to innovation through cooperation networks. In: Camagni, R. (Hrsg.): Innovation Networks. Spatial Perspectives. Belhaven Press: London, New York, 1-9.

Cooke, P.; Morgan, K. (1990), Learning Through Networking: Regional Innovation and the Lessons of Baden Württemberg, Regional Industrial Research Report Nr. 5, Cardiff.

Dodgson, M. (1996): Learning, Trust and inter-firm technological linkages: some theoretical associations, in: Coombs, R. et. al. (ed.) (1996), Technological Collaboration - The Dynamics of Cooperation in Industrial Innovation, Cheltenham.

Gelsing, L. (1993): Evaluating programs promoting networks. In: Significant Others: Background Papers. Regional Technology Strategies, Chapel Hill. NC.

Koschatzky K., Muller, E.; (1997): Firm Innovation and Region – Theoretical and Political Conclusions on Regional Innovation Networking. European Regional Science Association, 37th European Congress, Italy (CD-Rom).

Rosenfeld, S. A. (1996): Does cooperation enhance competitiveness? Assessing the impacts of inter-firm collaboration. Research Policy 25, 247-263.

Schatz, R. (ed.) (1996): Netzwerke als Basis der lernenden Organisation, Bonn.

Shapira, P. (1998): The Evaluation of USNet: Overview of Methods, Results and Implications. Final Report. Regional Technology Strategies, Chapel Hill. NC.

Shrivastava, P. (1983): A typology of organizational learning systems, in: Journal of Management Studies, 1/1983, S. 7-28.

Staehele, W. (1991): Management, Eine verhaltenswissenschaftliche Einführung, 6. erweiterte Auflage, München, Vahlen.

Steiner, M.; Hartmann, C.M. (1998), Learning with Clusters: A Case Study from Upper Styria, in: Steiner M. (Hg.), Clusters and Regional Specialisation, European research in regional science, 8, S. 211-225.

Steiner, M.; Hartmann, C.M. (1999), Interfirm Co-operation and Learning within SME-Networks: Two Case Studies from the Styrian Automotive Cluster, in: Fischer, M.; Suarez-Villa, L.; Steiner, M. (Hg.), Innovation, Networks and Localities, S. 85-106.

4 KAPITEL VIER EVALUATION IM RAHMEN DER REGIONALEN STRUKTURPOLITIK DER EUROPÄISCHEN UNION

Markus Gruber

4	KAPITEL VIER EVALUATION IM RAHMEN DER REGIONALEN STRUKTURPOLITIK DER EUROPÄISCHEN UNION	I
4.1	EVALUATION IM RAHMEN DER STRUKTURPOLITIK DER EUROPÄISCHEN UNION - ZUSAMMENFASSUNG	1
4.2	EINLEITUNG	13
4.3	STRUKTURPOLITIK DER EU	13
4.4	BEGLEITUNG UND BEWERTUNG IM RAHMEN DER EUROPÄISCHEN STRUKTURPOLITIK	16
4.4.1	<i>Entwicklung der Bedeutung von Begleitung und Bewertung im Rahmen der Europäischen Strukturpolitik</i>	16
4.4.2	<i>Rahmenbedingungen</i>	18
4.4.3	<i>Herausforderungen von Strukturfondsevaluationen</i>	21
4.4.4	<i>Thematische Ausrichtung der Evaluation</i>	21
4.5	ZIELE UND INDIKATOREN	23
4.5.1	<i>Die allgemeine Interventionslogik</i>	23
4.5.2	<i>Indikatoren</i>	26
4.5.2.1	Programmindikatoren: Input, Output, Ergebnis und Wirkung	27
4.5.2.2	Kontextindikatoren	28
4.5.2.3	Basisindikatoren	28
4.6	METHODISCHE ANSÄTZE ZUR EVALUIERUNG VON STRUKTURFONDS-PROGRAMMEN DER EU	33
4.6.1	<i>Überblick über Evaluierungsmethoden</i>	33
4.6.2	<i>Erfahrungen aus der Evaluierungspraxis</i>	36
4.6.2.1	Review der Europäischen Kommission	36
4.6.2.2	Methoden zur Evaluation von strukturpolitischen Programmen der EU in der Praxis	38
4.7	MONITORING	41
4.7.1	<i>Monitoring in Großbritannien</i>	42
4.7.2	<i>Context</i>	42
4.7.3	<i>Monitoring Frameworks and systems</i>	45
4.7.3.1	Monitoring systems in England and Wales	46
4.7.3.2	Regional case study: West Midlands	49
4.7.3.3	Monitoring systems in Scotland	51
4.7.3.4	Regional case study: Western Scotland	54
4.7.4	<i>Monitoring Indicators</i>	56
4.7.4.1	Monitoring indicators in England	56
4.7.4.2	Monitoring indicators in Scotland	64
4.7.5	<i>Conclusions</i>	69
4.8	KRITISCHE REFLEXION DER KONZEPTIONELLEN ÜBERLEGUNGEN ZUR EVALUIERUNG DER EU- STRUKTURFONDS-PROGRAMME	70
4.9	ANHANG	74

4.9.1	<i>Ebenen der Evaluation: Fragestellungen der Evaluierung von Strukturfondsprogrammen</i>	74
4.9.1.1	Ex-ante Evaluierung	74
4.9.1.2	Interim Evaluierung	76
4.9.1.3	Ex-post Evaluierung	77
4.9.2	<i>Verknüpfung von Zielhierarchien und Indikatoren – Indikorentableaus für Ziel-2-Interventionen in Österreich</i>	78
4.9.3	<i>Review of Scottish Structural Fund Evaluation Studies – interim evaluations- June 2000</i>	82
4.9.4	<i>Case studies: EPRC/FAI-Study</i>	87
4.9.5	<i>Definition von Indikatoren</i>	97
4.9.6	<i>Indikatoren für die Begleitung und Bewertung: Kommissions-Vorschläge</i>	99
4.10	LITERATUR.....	107

4.1 Evaluation im Rahmen der Strukturpolitik der Europäischen Union - Zusammenfassung

Zur EU-Strukturpolitik

Die EU-Strukturpolitik ist dem regionalen Disparitäten-Ausgleich verpflichtet. Die schrittweise Erweiterung der Gemeinschaft und die Verstärkung und Beschleunigung des Integrationsprozesses hat dieser Zielsetzung im Laufe der Zeit immer mehr an Gewicht gegeben. Die EU-Strukturpolitik wird legitimiert durch inakzeptable hohe Disparitäten sowie durch den Integrationsprozess, der nicht nur zu Gewinnen und Gewinnern führt, sondern auch Kosten verursacht, die jedoch unter den Mitgliedsländern und ihren Regionen nicht gleich verteilt sind. In der einfachsten Interpretation, wird das Problem als Core-periphery Problem der Union gesehen mit einem Kernraum von hoch entwickelten Nationen und Regionen und der weniger stark entwickelten Peripherie. Damit läßt sich argumentieren, daß die supranationale Ebene der Europäischen Union am besten geeignet ist, diesen negativen Wirkungen der Integration entgegenzuwirken und Mittel für eine „fairere“ Verteilung der Integrationsvorteile zu verteilen.

Der EU-Strukturpolitik wurde daher in den letzten zwei Jahrzehnten eine wachsende Bedeutung beigemessen. 1988 kam es zu einer einschneidende Reform mit der Einführung der Förderungs-Prinzipien wie der Konzentration der Mittel auf prioritäre Ziele, der Einführung einer mehrjährigen, partnerschaftlichen – d.h. zwischen Union und Mitgliedsländern durchgeführten – Programmplanung sowie der Additionalität, die besagt, daß Strukturfonds-Mittel zusätzlich sein müssen und nicht nationale Mittel und Budgetansätze verdrängen dürfen. Nach einer deutlichen Erhöhung der für die Strukturfonds bereitgestellten Mittel bis 1992, deren weiteren Ausweitung im Zuge des Maastricht-Vertrages und der Einrichtung des Kohäsionsfonds, liegt heute der Anteil der Strukturfonds am EU-Budget bei 35%. Im Rahmen der Agenda 2000 kam es nochmals zu einer Reform, die auf eine Konzentration der Mittel auf die prioritären Ziele, eine Konzentration der geförderten Gebiete und der gesetzten Interventionen abzielt sowie einen Rückzug der Kommission aus der Umsetzung der Programme und Verlagerung ihrer Kompetenzen auf die Planungsphase und Begleitung der Programme vorsieht. Die Begleitung in Form von Monitoring und Evaluierung hat durch das neue Strukturfonds-Regulativ nochmals an Bedeutung gewonnen. Der vollständige Evaluierungszyklus von ex-ante, Zwischen- und ex-post-Evaluierung wurde nunmehr in der Strukturfonds-Verordnung EG 1260/1999 aufgenommen.

Zum Stellenwert der Begleitung und Bewertung strukturpolitischer Programme der EU

Begleitung und Evaluierung hatten seit der Reform 1988 einen wachsenden Stellenwert eingenommen. Nach anfänglicher Kritik an fehlenden Evaluierungsstandards und sehr unterschiedlicher Interpretationen des Strukturfonds-Regulativs durch die einzelnen Mitgliedsländer und Regionen hat die Kommission versucht, einen Beitrag zur Entwicklung einer europäischen Evaluationskultur zu leisten und Standards für die Evaluierung ihrer Programme zu entwickeln: Dies äußerte sich in der Einrichtung von Evaluierungseinheiten innerhalb der Kommission, in der Einrichtung des MEANS-Programmes, der Entwicklung

von Leitlinien zur Evaluation sowie eines verstärkten Informations- und Erfahrungsaustausches zwischen Programm-Managern sowie Evaluatoren.

Aus heutiger Sicht gelten daher folgende Rahmenbedingungen:

- Hohe Verbindlichkeit und Stellenwert durch Verankerung des gesamten Evaluierungszykluses in den Verordnungen;
- Umfassendes Design von Konzeptebene, Programmvollzug, Wirkungsanalyse einschließlich der Möglichkeit internationaler Vergleiche;
- Orientierungsrahmen für die Programm-Entwicklung und Evaluation durch Systematisierungsversuche der Kommission und einen organisierten Erfahrungsaustausch.

Allerdings zwingt die Vergleichbarkeit zu sehr schematischen Ansätzen und engt den Raum für innovative Ansätze deutlich ein. Zudem führt der Rechtfertigungsdruck der EU zu einer übermäßigen Betonung der Zahl der geschaffenen Arbeitsplätze sowie zu Quantifizierungsversuchen, auch dort wo dies kaum gerechtfertigt erscheint und auch Arbeitsplatzeffekte kaum direkt zu messen und erfassen sind.

Zu thematischen Schwerpunkte der Evaluation strukturpolitischer Programme der EU

Das Evaluationsdesign im Rahmen der Strukturfonds-Programme ist umfassend und beinhaltet:

- Konzeptebene: Die Beurteilung der Interventionen hinsichtlich ihres Designs, der gesetzten Zielsetzungen, ihrer Kohärenz und ihrer Begründung aufgrund der Problemlage;
- Durchführungsebene bzw. operationale Ebene: Die Beurteilung der für die Umsetzung des Programmes vorgesehenen bzw. implementierten Mechanismen wie Programm-Management, Projektauswahl, Monitoring, Publizitätsvorschriften etc.;
- Programm-Vollzug: Die Beurteilung der Programm-Performance anhand von finanziellen und physischen Umsetzungsindikatoren;
- Wirkungsebene: Die Beurteilung der Wirkung des Programmes im Hinblick auf die definierten Ziele.

Eine besondere Beachtung finden jeweils die sogenannten horizontalen Prioritäten der Europäischen Union: Umwelt, Chancengleichheit, Beschäftigung, Innovation/F&E, Kleine- und Mittlere Betriebe. Siehe dazu im Detail Übersicht 1:

Übersicht 1 Thematische Schwerpunkte der Strukturfonds-Evaluation

Schwerpunkt	Inhalt	Ex-ante	Interim	Ex-post
Erfahrungen über Wirkungen bisheriger Programme:	Beurteilung der Erfahrungen über Effektivität und Effizienz der Programme aus vorangegangenen Programm-Perioden und deren Berücksichtigung im aktuell vorliegenden Programm.			
Analyse der Ausgangssituation:	Beurteilung der sozio-ökonomischen Ausgangssituation und Ableitung eines Stärken/Schwächen-Profiles. Speziell zu beachten sind folgende Politikschwerpunkte: Umweltsituation, Chancengleichheit, Beschäftigung, Innovation, KMU.			
Strategiekonzept	Beurteilung der vorgelegten Strategien hinsichtlich ihrer Adäquanz und Kohärenz; Beurteilung des Strategiekonzeptes hinsichtlich der Politikbereiche: Umwelt, Chancengleichheit, Beschäftigung, Innovation, KMU.			
Zielsetzungen	Beurteilung der gewählten Zielsetzungen sowohl hinsichtlich der Kohärenz mit der Analyse als auch der Frage ob die Zielsetzungen ausreichend quantifiziert wurden und die gewählten quantitativen Ziele als realistisch zu beurteilen sind.			
Maßnahmen	Beurteilung der einzelnen Maßnahmen und des Maßnahmen-Mix auf Adäquanz und Kohärenz.			
Externe Kohärenz mit EU / nationaler Politik	Beurteilung der Kohärenz des Programmes hinsichtlich der EU-Politik-Bereiche sowie der nationalen Politiken.			
Finanztableaus Gewichtung von Strategien und Maßnahmen	Beurteilung der Adäquanz der gewählten finanziellen Gewichtung der einzelnen Programm-Teile und Maßnahmen.			
Indikatoren	Beurteilung der Aussagekraft, Konsistenz und Adäquatheit der bereits vorliegenden Indikatoren			
Partnerschaftsprinzip	Beurteilung, ob im Rahmen der Vorbereitung / Umsetzung des Programmes die Regelungen hinsichtlich der Partnerschaft entsprechend eingehalten wurden.			
Operationelle Rahmenbedingungen	Beurteilung der (geplanten) Verfahren der Programm-Umsetzung. <ul style="list-style-type: none"> - Beurteilung der Organisationsstrukturen hinsichtlich Klarheit, Transparenz, Informationsflüsse - Beurteilung der eingesetzten Mechanismen der Umsetzung (Ausschreibungen, Antragsverfahren, Publizität und Öffentlichkeitsarbeit) - Beurteilung der Projektselektions-Kriterien und -mechanismen - Beurteilung des Begleit(=Monitoring)systems. 			
Programm-Vollzug - Inputfaktoren - Finanzielle Indikatoren	Beurteilung des Standes der Mittelbindungen und Auszahlungen nach Programm / Prioritäten / Maßnahmen und Feststellen von Gründen für Abweichungen zu den geplanten Finanztabellen.			
Programm-Vollzug - Output-Faktoren	Beurteilung des Programmfortschritts hinsichtlich der Erreichung der festgelegten Output-Ziele. Rückschlüsse auf die Zielerreichung; Spezielle Beachtung der horizontalen Schwerpunkte der Union Umwelt, Beschäftigung, Chancengleichheit, KMU, Innovation.			
Wirkungsanalyse	Beurteilung der Programm-Wirkungen auf Basis der Programm-Ziele; Abschätzung von Brutto- bzw. Netto-Effekten.			

Quelle: eigene Zusammenstellung basierend auf VO 1260/99 sowie Arbeitspapieren der EU-Kommission

	Schwerpunkt der Evaluation
	Zu behandelndes Thema

Zur Verknüpfung von Zielen und Indikatoren und dem Indikatorenset

Formal bauen die Überlegungen auf das Logical framework auf, wodurch in der Planung ausgehend vom Globalziel auf unterschiedlichen Ebenen die dazugehörigen Indikatoren sozusagen von oben nach unten festgelegt werden. Bei der Begleitung wird in umgekehrter Reihenfolge – also praktisch von unten nach oben - die Zielerreichung überprüft.

Alle Interventionen der Programme werden in einem spezifischen Kontext in bezug auf ein bestimmtes globales Ziel geplant. Dieses Ziel ist wiederum für die Interventionsstrategie ausschlaggebend und führt zu einer Reihe spezifischer, im wesentlichen auf die Förderschwerpunkte ausgerichteter Ziele. Jedes spezifische Ziel wird zur Erreichung der operationellen Ziele mit Hilfe konkreter Maßnahmen umgesetzt. Die entsprechenden Zielsetzungen werden – soweit dies möglich ist – auf unterschiedlichen Ebenen quantifiziert. Nach anfänglichen Schwierigkeiten in den Programmen Anfang der 90er Jahre sind hinsichtlich der Spezifizierung und Quantifizierung der Ziele deutliche Verbesserungen zu erkennen (Bachtler/Taylor 1999). Durch diese Vorgangsweise soll ein Zusammenhang zwischen Zielen und Indikatoren-Ketten entwickelt werden. In der Begleitung bzw. Bewertung erfolgt die Betrachtung in umgekehrter Reihenfolge:

- Die Maßnahmen werden unter Einsatz von (personellen, finanziellen) Mitteln (**Input**) durchgeführt.
- Sie führen zu einem physischen **Output** (z.B. Kilometer gebauter Straße, bereitgestellte Ausbildungsplätze usw.), der die Fortschritte bei der Durchführung der Maßnahme verdeutlicht.
- Damit verbunden sind „**Results**“, die die (unmittelbaren) Auswirkungen der finanzierten Maßnahmen (z.B. verkürzte Fahrtzeiten, verringerte Transportkosten oder höhere Zahl “erfolgreicher” Ausbildungsteilnehmer) auf die Begünstigten der Maßnahme darstellt.
- Die Leistungen werden schließlich anhand ihrer **Wirkung** im Hinblick auf die Erreichung der globalen und der spezifischen Ziele des Programms gemessen. Sie bilden die wichtigste Grundlage für die Bewertung des Erfolg oder Mißerfolgs der in Frage stehenden Intervention. Zu den **spezifischen** Wirkungen zählen u.a. eine Erhöhung des Warenverkehrs oder eine bessere qualitative Anpassung an die Anforderungen des Arbeitsmarktes. Der Input wird im Vergleich zum Output und damit in bezug auf die erzielten Ergebnisse und Wirkungen dargestellt.

Demzufolge werden die:

- **operationellen Ziele** anhand des Outputs ausgedrückt (z.B. Bereitstellung von Schulungsmaßnahmen für Langzeitarbeitslose);
- **spezifischen Ziele** anhand der Leistung ausgedrückt (z.B. Verbesserung der Beschäftigungsfähigkeit von Langzeitarbeitslosen durch Schulungsmaßnahmen);
- **globalen Ziele** anhand der Wirkung ausgedrückt (z.B. Verringerung der Zahl der Langzeitarbeitslosen).

Es wird damit einerseits eine systematische Verknüpfung von Zielen und Indikatoren versucht und andererseits sollen „Wirkungsketten“ innerhalb des Indikatorensystems dargestellt werden. Demnach unterscheidet man zwischen (siehe dazu Indikatoren-Liste im Anhang):

- **Input-Indikatoren** geben an wieviele finanzielle Mittel eingesetzt wurden.
- **Output-Indikatoren** beziehen sich auf die unmittelbare Fördertätigkeit. Sie werden häufig in physikalischen oder finanziellen Einheiten gemessen (z.B. Kilometer gebauter Straße, Zahl der finanziell unterstützten Unternehmen, Zahl der bereitgestellten Ausbildungsplätze usw.)
- **Ergebnisindikatoren** beziehen sich auf die direkten und unmittelbaren Auswirkungen eines Programms. Die Indikatoren können physischer (Verringerung der Fahrtzeiten, Zahl der Auszubildenden mit erfolgreichem Abschluß) oder finanzieller Art sein (Investitionen des Privatsektors, Verringerung der Transportkosten usw.).
- **Wirkungsindikatoren** beziehen sich auf die Wirkungen des Programms über die unmittelbaren Auswirkungen auf die Endbegünstigten hinaus, die zum Teil erst mit einem time-lag beobachtbar werden.

Die Kommission hat in ihren Arbeitspapieren diesbezüglich auch Listen von Indikatoren vorgeschlagen, die in den Programmen 2000-2006 zur Anwendung kommen sollen (siehe dazu Anhang).

Zum Monitoring

Monitoring gilt als integraler Bestandteil und Grundlage für Berichterstattung und die Evaluation. Während in der Vergangenheit die Regelungen hinsichtlich des Monitorings von den Mitgliedsstaaten und den Regionen sehr unterschiedlich interpretiert wurden, sind nunmehr deutliche Verbesserungen festzustellen (EPRC 1999). Die Monitoring-Systeme gehen nunmehr über ein reines Finanzmonitoring hinaus und umfassen in zunehmenden Maße auch physische Indikatoren (Output und Result), die in Zusammenarbeit mit der Kommission in Form von sogenannten Core-Indicators festgelegt werden. Als good practice Beispiele für umfangreiche und auf die Programme abgestimmte Monitoringsysteme innerhalb der Strukturfonds-Ansätze sind jene in Großbritannien, insbesondere auch in Schottland zu erwähnen. Allerdings bleiben zahlreiche Probleme aufrecht, die u.a. umfassen: einseitige Ausrichtung auf die Zählung von Arbeitsplätzen mit zum Teil unklaren Definitionen und der Gefahr von Doppelzählungen, fehlenden qualitativen Informationen.

Zu den Evaluierungsmethoden

Bei der Betrachtung der Evaluierungsmethoden steht man schließlich vor der Herausforderung der Evaluation eines multi-sektoriellen, partnerschaftlichen entwickelten und umgesetzten Programmes. Prinzipiell scheint damit zwar eine breite Palette an verfügbaren Tools zur Datenerhebung und –analyse anwendbar zu sein: Allerdings ergeben sich sowohl konzeptionell als auch in der praktischen Umsetzung erhebliche

Einschränkungen, die in der Evaluierungspraxis letztlich zur Anwendung eines relativ einfachen Methodenspektrums führen:

- Durch die Tatsache, dass multi-sektorielle und partnerschaftliche Programme zu evaluieren sind, werden – neben den über das Monitoring gesammelten Daten - von der Kommission bzw. im MEANS-Programm vorwiegend qualitative Ansätze der Datengenerierung und Analyse wie Fokus-Gruppen, Case-Studies, Experten-Interviews etc. vorgeschlagen. Standard-Methoden der Evaluation wie multivariate Verfahren, Regressions-Analysen sind demnach kaum – und wenn – dann eben nur für spezielle Fragestellungen innerhalb der Gesamtevaluation einsetzbar.
- Des weiteren empfiehlt die Kommission einen bottom-up-approach, in dem man sich über die Einzelprojektergebnisse (bereitgestellt über das Monitoring oder generiert durch gesonderte Erhebungen) an die Beantwortung der Evaluationsfragen annähert. Top-down-Ansätze im Sinne des Einsatzes von makro-ökonomischen Modellen sowie Input-Output-Modellen werden eher nur für große Programme bzw. Regionen (im Sinne von Ziel-1-Gebieten) empfohlen. Zum Teil wurden solche Modelle für Ziel-1-Regionen auch von der EU-Kommission beauftragt, wie das dynamische Input-Output-Modell von Beutel oder das HERMES bzw. daraus für Irland entwickelte HERMIN-Modell (siehe dazu Anhang).

Einerseits wird argumentiert, dass die Programme vom Umfang her zu gering seien, um eine signifikante Wirkung auf das Wirtschaftswachstum zu entfalten. Der Bremer Ausschuß für Wirtschaftsforschung, der dieser Meinung folgt, zeigt, daß selbst in Nordrhein-Westfalen, in dem über 50 % der deutschen Ziel-2-Mittel (nach der Zieldefinition-alt) ausgegeben werden, der Förderungsanteil gemessen am BIP weit unter einem Promille, gemessen an der Wertschöpfung bei etwa 2 Promille liegen (BAW 1994). Zum anderen wird aber vor allem argumentiert, dass aus Kosten-Nutzen-Gesichtspunkten sowie aus Fragen der Datenverfügbarkeit der Einsatz von Modellen für kleinere Programme als nicht geeignet erscheint. Allerdings wird diesem Argument weder von EPRC/FAI (2000) noch in der vorliegenden Studie gefolgt: Sofern ein regionales Modell entwickelt vorliegt – wie es beim sozio-ökonomischen Modell Steiermark (STYR-I-O) der Fall ist – welches auf entsprechend geeigneten Datengrundlagen erstellt wurde und vor allem regionale Export-Ströme berücksichtigen kann – ist dies auch für die Programm-Evaluierung entsprechend einsetzbar. Allerdings bedarf es einer Überführung der angebotsseitigen Impulse in nachfrageseitige Effekte.

Reflexion der konzeptionellen Überlegungen

Die „empirischen“ Erfahrungen im Rahmen der Entwicklung, Umsetzung und Bewertung von Strukturfonds-Programmen zeigen jedoch eine deutliche Relativierung der konzeptionellen Überlegungen.

- Tatsächlich zeigt sich, dass das „logical framework“ in der Planung und Evaluierung der Programme kaum in die Praxis umgesetzt wird. Die Ziele sind – wenn auch deutliche Verbesserungen aufgrund der Vorschriften der Kommission aber auch aufgrund der Erfahrungen der Regionen in der Programmierung zu beobachten sind – häufig unspezifisch und wenig operationalisiert. Dies hat auch system-immanente Ursachen, versuchen sich die Regionen doch so wenig wie möglich in der Programm-Planung zu

binden und durch unspezifische Zielsetzung eine größtmögliche Flexibilität für die Umsetzungsphase zu erhalten.

- Daraus resultieren auch häufig Probleme in der Verknüpfung der strategischen Zielsetzungen mit den Indikatoren – die übermäßig stark auf Arbeitsplatzeffekte konzentriert sind. Ein als good practice zu bezeichnendes Beispiel ist hier jedoch der von Bremen verfolgte Ansatz der Verknüpfung von Zielhierarchien mit Indikatoren (siehe dazu Anhang – Indikatorentableaus).
- Das Evaluierungs-Instrumentarium wurde vor allem in einer Zeit entwickelt, als die Strukturfonds-Programme sehr stark auf traditionelle Infrastruktur- und angebotsseitige Maßnahmen sowie einfache Investitionen ausgerichtet waren. Spätestens seit Mitte der 90er Jahre ist jedoch ein Übergang auf eher „weiche“, nachfrageorientierte Maßnahmen wie Kooperationen, Netzwerke, Cluster, F&E etc. zu beobachten. Sowohl methodische Anforderungen als auch die Indikatoren sind jedoch noch zu sehr an den traditionellen Interventionen orientiert, während für diese „neuen“ Maßnahmenfelder noch nicht adäquat angepaßt wurden.
- Die Evaluationen konzentrieren sich meist auf qualitative Fragestellungen während wirklich Wirkungsabschätzungen – auch in ex-post-Analysen - kaum zu beobachten sind (EPRC/FAI 2000). Vor allem sind jedoch die Fragestellungen der Netto-Effekte und dabei insbesondere die Bereinigung der Brutto-Effekte um Mitnahme-, Verlagerungs- und Substitutionseffekte ungelöste Probleme. Die Analyse der Evaluationsstudien, von „Meta-Studien“ sowie methodischen Konzepten der EU-Kommission¹ zeigte:
 - Aussagen über Mitnahme- und Verlagerungseffekten beruhen meist auf Surveys, die an die Förderungsempfänger adressiert sind. Die Frage lautet, ob das Projekt auch ohne Förderung durchgeführt, zu einem späteren Zeitpunkt durchgeführt oder in kleinerem Umfang durchgeführt worden wäre. Auf Basis der Surveys wird versucht, Koeffizienten für Mitnahme- und Verlagerungseffekte zu schätzen.
 - Als Alternative werden häufig Koeffizient und benchmarks für Mitnahme- und Verlagerungseffekte praktisch „off the shelf“ verwendet, was prinzipiell als problematisch einzuschätzen ist, wenn nicht berücksichtigt wird, daß diese nicht nur nach Sektoren stark variieren, sondern vom Design der Maßnahmen und dem Auswahlprozess der Projekte stark abhängig sind.²
 - Kontrollgruppen-Ansätze werden kaum verwendet, wenn meist nur im Rahmen der ESF-Förderung (Ziel 3). Lediglich eine Evaluierungsstudie im Ziel-1-Gebiet Irland, welche einen Kontrollgruppenansatz verwendete, um Netto-Effekte für Maßnahmen des EFRE war zu identifizieren.

D.h., zusammenfassend, daß die Evaluierungspraxis sich eher auf bescheidenem methodischen Niveau bewegt, was vor allem auch auf Zeit- und Kostenrestriktionen im Rahmen der Programm-Evaluierungen zurückzuführen ist.

¹ EPRC/FAI 2000; Ernest&Young 1999, MEANS Collection (1999), EK (1995), Gemeinsame Leitlinien für die Begleitung und die Zwischenbewertung

² Allerdings wurde in einer hier als good practice zu bezeichnenden Studie wurden 250 Evaluations-Studien analysiert, um daraus eine Liste von Koeffizienten zur Berechnung von Netto-Effekten bei Arbeitsplatzzahlen zu generieren. Diese Liste wurde in der Folge in einer Evaluierungsstudie eingesetzt (Yorkshire and Humberside Ziel-2-Evaluierung).

Zu den Stärken und Schwächen im Überblick**Übersicht 2 Stärken / Schwächen in der Evaluation von Strukturfonds**

	Stärke	Schwäche
Evaluationsverfahren	hohe Verbindlichkeit, hoher Stellenwert	
Struktur und Inhalte	umfassendes Design mit allen wesentlichen Elementen (Konzept. Design, Wirkungsanalyse), Int. Vergleichbarkeit	schematisch, wenig Raum für innovative methodische Ansätze
Zielformulierung	Versuch einer systematischen Verknüpfung (Zielhierarchien)	in Praxis: breite, unspezifische Zielsetzungen (alle lesen sich gleich)
Zusammenhang Ziele - Evaluationskriterien	<ul style="list-style-type: none"> - Kriterien werden bei Programmplanung festgelegt - Kausalitätsketten: Output, Outcome, Impact 	<ul style="list-style-type: none"> - enorme Datenmengen ohne Informationsgehalt - Praxis: unsystematischer Indikatorenmix - eher auf "traditionelle" Interventionen
Methodische Probleme	Methodische Diskussionen auf EU-Ebene (z.B. MEANS)	<ul style="list-style-type: none"> - Wirkungsindikatoren auf Monitoring Ebene - Nettoeffekte - Additionalität

-

Quelle: eigene Zusammenstellung

Aus den im Rahmen des Projektes gewonnenen Einschätzungen über die konzeptionellen Ansätze und die Realität der Evaluierungspraxis sowie der Kenntnis der Rahmenbedingungen der Evaluationen in Österreich läßt sich ableiten, daß es Weiterentwicklungspotential einerseits in der Anwendung von qualitativen Methoden durch die Einbeziehung von Stakeholdern mit in den Evaluierungs-Prozeß (bspw. Concept mapping of impacts bzw. Matrix of cross impacts etc. wie in der MEANS Collection Volume 4 näher beschrieben) und andererseits in der Umsetzung von Modell-Anwendungen (wie dies auch von EPRC/FAI vorgeschlagen wurde) bspw. im Rahmen einer Evaluierung des Ziel-2-Programmes Steiermark gegeben ist. Dazu ist festzuhalten, daß der Einsatz eines Modells nur ergänzend erfolgen kann, um Gesamtwirkungen auf Größen wie Beschäftigung und Wertschöpfung abschätzen zu können. Wesentliche Dimensionen der Programme, die zur Stärkung der Wettbewerbsfähigkeit beitragen sollen müssen außerhalb des Modells beurteilt werden.

Einsatz des Sozio-ökonomischen Modells STYR-I-O zur Evaluierung von strukturpolitischen Programmen

Das Sozio-ökonomische Modell STYR-I-O

Das am Institut entwickelte regionalökonomische Input-Output Modell STYR-I-O ist ein ökonomisches Prognose- und Analysemodell mit dem die zukünftige regionale Wirtschaftsentwicklung prognostiziert und auch die Auswirkungen wirtschaftspolitischer Maßnahmen, Veränderungen der institutionellen Rahmenbedingungen etc. auf Sektoren der steirischen Wirtschaft analysiert werden können. Dies geschieht auf der Grundlage der seit 1976 beobachteten wirtschaftlichen Trends und Entwicklungen in unterschiedlichen Branchen auf nationaler und regionaler Ebene sowie einer regionalen Input-Output Tabelle der Steiermark.

Die steirische Wirtschaft wird im Rahmen von STYR-I-O in 42 Sektoren gegliedert. Die Aktivitäten jedes dieser Sektoren werden auf Grundlage von drei ökonomischen Variablen gemessen: Beschäftigung (selbstständig sowie unselbstständig Beschäftigte), Produktionswert und Wertschöpfung. Die Veränderungen dieser drei Variablen werden auch prognostiziert.

Während die Input-Output Tabelle die Verflechtungen zwischen den Sektoren abbildet, d.h. ihre wechselseitigen Handelsbeziehung erfaßt, wird in den ökonometrischen Gleichungen versucht, den empirischen Zusammenhang zwischen

- Produktion und Beschäftigung,
- Beschäftigung und Wertschöpfung,
- Wertschöpfung und Nachfrage,

jeweils auf Ebene der einzelnen Sektoren, in Abhängigkeit von den exogenen, das sind im Fall von STYR-I-O vor allem nationale, Variablen, zu schätzen. Da die Verfügbarkeit sowie die Qualität regionaler Daten stark eingeschränkt ist, wird auf eine komplexere, stärker theoriegeleitete Modellierung der Schätzgleichungen (z.B. die Ableitung der Gleichungen aus ökonomischen Optimierungsproblemen) verzichtet.

Im folgenden soll auf den möglichen Einsatz von STYR-I-O oder ähnlicher Modelle im Rahmen von Evaluierungen eingegangen werden und dabei sowohl dessen Vorteile wie auch die Nachteile bzw. die Verwendungsbeschränkungen diskutiert werden.

Der Einsatz von STYR-I-O im Rahmen von Evaluierungen regionaler Förderprogramme

Messung indirekter und induzierter Programmwirkungen

Neben den direkten Wirkungen eines Förderprogramms, wie den unmittelbar ausgelösten wirtschaftlichen Effekten in geförderten Betrieben oder der zusätzlichen Beschäftigung durch die Wiedereingliederung von Arbeitslosen etc., können auch die indirekten und induzierten Effekte des Programms einen beträchtlich Umfang erreichen und sollten deshalb in die Evaluierung miteinbezogen werden. Der Hauptnutzen eines makroökonomisch orientierten Modells besteht darin, die quantitative Dimension dieser Effekte zu abzuschätzen. Handelt es sich zudem um ein sektoral disaggregiertes Modell wie STYR-I-O, kann auch die sektorale Verteilung der Effekte erfasst und damit regionale Spezialisierungsmuster berücksichtigt werden. Dies ist auch ein wesentlicher Vorteil gegenüber einfachen, sogenannten „export-base“ orientierten Multiplikatoren.

Bei der Berechnung der indirekten und induzierten Programmwirkungen wird in STYR-I-O, im Gegensatz zu reinen Input-Output Modellen, von keinem linearen Zusammenhang zwischen, zum Beispiel, Beschäftigung und Produktion ausgegangen. Auch die statische Natur von Input-Output Modellen, welche die Lieferbeziehungen zwischen den Sektoren nur zu einem, oft schon weit zurückliegenden Zeitpunkt abbildet, wird in STYR-I-O überwunden, indem die Veränderung der Input-Output-Koeffizienten modelliert und damit struktureller Wandel zugelassen wird.

Generierung von „Benchmarks“

Ein dynamisches Modell, das die Veränderungen einer regionalen Wirtschaft über mehrere Jahre hindurch statistisch erfasst und nach Vorgabe von Prognosewerte für die exogenen Variablen auch in die Zukunft projizieren kann, bietet dem Evaluator somit prinzipiell auch die Möglichkeit, die durch das Programm ausgelösten Veränderungen des regionalwirtschaftlichen Entwicklungspfades der ohne Programm vorhergesagten Entwicklung gegenüberzustellen, d.h. einen „Benchmark“ zu generieren. Damit können die Programmwirkungen (etwa gemessen an der zusätzlich hervorgebrachten regionalen Wertschöpfung) direkt mit den Programmkosten verglichen werden.

Durch ein solches „Benchmarking“ soll vor allem auch die im Rahmen von Evaluierungen immer wieder diskutierte Frage der Additionalität gelöst werden. Allerdings muss dazu das Programm einen Umfang aufweisen, der auch erkennbare Wirkungen generiert, und die Evaluierung kann nur ex-post bzw. einige Zeit nach Programmbeginn durchgeführt werden. Zudem muss die Frage der Kausalität geklärt werden. Ist der Programmumfang zu gering, muss der Evaluator selbst eine Einschätzung der zusätzlichen direkten Effekte (z.B. in Form von neuen Arbeitsplätzen) treffen, die dann als Input in die Modellberechnungen eingeht.

Eine Analyse der wirtschaftlichen Effekte eines Förderprogramms lässt meist die Opportunitätskosten, d.h. den Nutzen aus dem Einsatz der Mittel für alternative Verwendungen (sonstige Förderprogramme, Steuersenkungen etc.), außer Acht. Die Wirkungen bekannter Alternativen könnten jedoch in Modellberechnungen simuliert werden.

Modelleinsatz in ex-ante Evaluierungen

Modelle wie STYR-I-O werden meist für ex-post Evaluierungen verwendet. Ist es dem Evaluator möglich, eine realistische Einschätzung der Programmwirkungen ex-ante vorzunehmen oder kann er zumindest über deren sektorale Richtung schon genauer Auskunft geben, so können sie jedoch auch für ex-ante Evaluierungen eingesetzt werden. Dabei könnten zum Beispiel Szenarien gerechnet werden, die einen hypothetischen „Break even“ Punkt des Programms errechnen: Geht man von bekannten Kosten des Programms aus, so kann etwa die Zahl der zusätzlichen Beschäftigten geschätzt werden, ab der sich ein Programm im volkswirtschaftlichen Sinn „rechnet“, d.h. die eingesetzten Mittel von den Einkommensgewinnen übertroffen werden.

Modelle als abstraktes Bild der Wirklichkeit

Jedes ökonomische Modell, das für empirische Analysen verwendet wird, muss die komplexen ökonomischen Zusammenhänge auf jene Beziehungen reduzieren, die als wesentlich für die Entwicklung der Wirtschaft angesehen werden. Zudem liegen über wirtschaftliche Aktivitäten, über das Verhalten der Wirtschaftssubjekte (Betriebe, Konsumenten etc.) und den dieses Verhalten bestimmenden Faktoren nur unzureichende Informationen vor. Aus diesem Grund sind in allen Modellen zahlreiche Annahmen enthalten, welche die Vorstellungen des Modellbauers über ökonomische Prozesse widerspiegeln bzw. auf Ergebnissen der theoretischen ökonomischen Forschung basieren. Viele dieser Annahmen haben jedoch wesentlichen Einfluss auf die Modellergebnisse, was bei der Interpretation dieser Ergebnisse im Rahmen der Evaluierungsstudien berücksichtigt werden muss.

STYR-I-O als nachfrageorientiertes Modell

STYR-I-O als regionalökonomisches Input-Output Modell enthält ebenso zahlreiche Annahmen, welche nicht nur die Modellergebnisse beeinflussen, sondern auch die Einsatzmöglichkeiten des Modells beschränken.

Obwohl durch STYR-I-O versucht wird, die Veränderungen der Input-Output Koeffizienten und somit technologischen Wandel bzw. Veränderungen in den interregionalen Lieferbeziehungen zu berücksichtigen, ist es dennoch ein stark nachfrageorientiertes Modell, das die Angebotsseite nur unzureichend erfasst. Wird durch ein Programm das Aktivitätsniveau eines Sektors erhöht, so kommt es über eine Erhöhung der intermediären Nachfrage sowie, über Einkommenssteigerungen, auch der Endnachfrage zu Produktionsimpulsen für andere Sektoren und damit zu positiven wirtschaftlichen Effekten in der gesamten Region. Diese positiven Effekte werden jedoch durch ein nachfrageorientiertes Modell, das die Angebotsseite vernachlässigt, tendenziell überschätzt: Zum einen werden oft Kapazitätsbeschränkungen nicht berücksichtigt (wie auch in STYR-I-

O), zum anderen werden Preisveränderungen nicht explizit modelliert, welche die zusätzlich Nachfrageerhöhung dämpfen könnten. Es wird also von einem unendlich elastischen Angebot (bei „durchschnittlichen“ Preisen) ausgegangen.

Darüber hinaus sind mit dem Einsatz makroökonomischer Modelle im speziellen Kontext der Evaluierung noch weitere Probleme verbunden: Programme zielen oft auf eine Verbesserung der betrieblichen Wettbewerbsfähigkeit ab, indem zum Beispiel die Produktivität erhöht oder die Innovationsfähigkeit gesteigert werden soll. In einem makroökonomischen Modell wie STYR-I-O können solche betrieblichen Verbesserungen nur dann abgebildet werden, wenn sie in Veränderungen des Produktionsniveaus einzelner Sektoren „übersetzt“ werden können. Die Modellierung des Zusammenhangs zwischen Kostensenkungen, Innovationen etc. und Veränderungen in der Produktion ist jedoch äußerst schwierig bzw. mit großer Unsicherheit behaftet und muss auf jeden Fall außerhalb des Makro-Modells, im wesentlichen durch ein weiteres, z.B. ökonometrisches Modell, durchgeführt werden.

Generell gilt, dass Evaluatoren große Sorgfalt bei der Generierung des Modellinputs und der dabei getroffenen Annahmen, etwa über die regionalen Direktimporte, walten lassen müssen, da diese weitgehend die durch das Modell berechneten Effekte beeinflussen. Bei fast allen Modellanwendungen stehen weniger die Rechenvorgänge mit dem Modell im Vordergrund, als vielmehr die Generierung eines brauchbaren und auch realistischen Modellinputs. Im Kontext von Evaluierung besteht somit die größte Herausforderung für den Evaluator darin, die Programmeffekte in Modellinputs zu übersetzen.

Ein weiteres Problem in der Modellanwendung besteht in der räumlichen Orientierung der Modelle, die sich oft von der Zielregion des Programms unterscheidet. Richtet sich ein Programm auf eine Subregion innerhalb des durch das Modell erfassten Gebiets, so können die Wirkungen auf diese Subregion nur schwer aus den Modellergebnissen abgeleitet werden. Die Anpassung des Modells an die Gegebenheiten und Strukturen der Subregion ist allerdings nur mit hohen Kosten möglich bzw. unterliegt auch, je nach Größe dieser Region, noch weitergehenden Datenbeschränkungen.

4.2

Einleitung

Im folgenden Abschnitt im Rahmen des EVINA-Schwerpunktes Regionale Strukturpolitik erfolgt eine kritische Auseinandersetzung mit konzeptionellen Überlegungen und praktischen Erfahrungen im Rahmen der Evaluierung strukturpolitischer Programme der EU. Dabei stehen folgende Fragestellungen im Mittelpunkt:

- Welche Rahmenbedingungen werden für die Evaluierung strukturpolitischer Programme der EU gesetzt?
- Welche Fragestellungen stehen im Mittelpunkt?
- Anhand welcher Indikatoren werden Evaluierungen vorgenommen?
- Welche methodischen Ansätze werden verwendet?
- Wie erfolgt das Zusammenspiel von Monitoring und Evaluation?

Beschreibende und analytische Aussagen beziehen sich in erster Linie auf Programme des Zieles-2 (alt) der EU Regionalpolitik.³ Es handelt sich dabei um strukturpolitische Programme zur Umstellung der von der rückläufigen industriellen Entwicklung schwer betroffenen Regionen. Aus innovationspolitischen Gesichtspunkten sind die Ziel-2 (alt)-Gebiete aufgrund der notwendigen Erneuerung ihrer industriellen Basis von besonderem Interesse. Als Grundlage der Analyse wurde gewählt:

- Publikationen, methodische Arbeitspapiere der EU-Kommission;
- Publikationen des MEANS Programmes;
- Eine Sichtung von Evaluierungsstudien, insbesondere in Großbritannien, Deutschland;
- Die Berücksichtigung von Erfahrungen aus der Evaluierung von Strukturfondsprogrammen in Österreich.

Das European Policies Research Centre (EPRC) der Universität Strathclyde in Glasgow / Schottland verfügt über umfangreiches Wissen im Themenbereich EU-Strukturpolitik und leitet das IQ-Netzwerk (Improving the Quality of Structural Fund Programming through Exchange of Experience) mit 12 Partner-Regionen. Aus diesem Grund wurde im Rahmen des Projektteiles eine Kooperation mit EPRC eingegangen. Neben Informationsaustausch und der Bereitstellung von Unterlagen wurden konkrete Ausarbeitungen im Schwerpunkt Monitoring durch EPRC vorgenommen.

4.3 Strukturpolitik der EU

Die EU Strukturpolitik ist dem regionalen Disparitäten-Ausgleich verpflichtet. Die schrittweise Erweiterung der Gemeinschaft und die Verstärkung und Beschleunigung des

³ Unter Ziel-2 (alt) wird die seit der Strukturfondsreform 1988 und bis 1999 gültige Zielgebietsdefinition verwendet. Im Rahmen der Agenda 2000 wurden für die Periode 2000-2006 die Ziel-2-Gebiete erweitert um ehemalige Ziel-5b-Gebiete sowie städtische Problemgebiete. Da sich die Analyse jeweils auf Maßnahmen des Europäischen Fonds für Regionale Entwicklung bezieht, haben die Ausführungen ihre Gültigkeit auch für das Ziel-2 der Strukturfondsperiode 2000-2006.

Integrationsprozesses hat dieser Zielsetzung im Laufe der Zeit immer mehr an Gewicht gegeben. Durch die regionale Struktur- und Kohäsionspolitik sollte einer Verstärkung der Disparitäten entgegengesteuert werden.

Warum bedarf es überhaupt einer Regionalpolitik auf der Ebene der Union? Nach Fitzgerald/Michie (1997) steht im Zentrum die Argumentation, die inakzeptable Höhe der sozio-ökonomischen Disparitäten innerhalb der Europäischen Union. In der einfachsten Interpretation, wird das Problem als Core-periphery Problem der Union gesehen mit einem Kernraum von hoch entwickelten Nationen und Regionen und der weniger stark entwickelten Peripherie. Durch die schrittweise Erweiterung wurden diese Disparitäten verstärkt und manifestierten sich in einem starken Nord-Süd-Gefälle.

Verstärkt wird dieses Gefälle durch den Integrationsprozess, der nicht nur Gewinne, sondern auch Kosten verursacht und diese Kosten nicht gleich über alle Mitgliedsländer verteilt. Damit läßt sich argumentieren, daß die supranationale Ebene der Europäischen Union am besten geeignet ist, diesen negativen Wirkungen der Integration entgegenzusteuern und Mittel für eine „fairere“ Verteilung der Integrationsvorteile zu verteilen.⁴ In diesem Zusammenhang wurde auch vielfach – insbesondere von Jacques Delors - das Argument der Solidarität bemüht.

Bereits in den Römer Verträgen wurden 1957 festgelegt, daß auf eine ‚harmonische‘ Entwicklung durch die Reduktion regionaler Disparitäten und den Rückstand benachteiligter Regionen zu achten ist. Allerdings dauerte es bis Mitte der 70er Jahre bis spezifische Politikmaßnahmen gesetzt bzw. finanzielle Mittel bereitgestellt wurden, genauer gesagt bis 1975 als der Europäische Fonds zur Regionalen Entwicklung (EFRE) eingerichtet wurde. Seit dieser Zeit wurden der Regionalpolitik der Europäischen Union zunehmend mehr Mittel gewidmet. Im wesentlichen können wir drei Phasen der EU-Strukturpolitik unterscheiden:

- In der **Phase 1** wurde der EFRE eher als Kompensationsmechanismus für jene Länder gesehen, die überdurchschnittlich stark zum EU-Budget beigetragen haben. Die Förderung erfolgte weitgehend unkoordiniert in Form von Einzelprojekten. Kritik an der mangelnden Koordination und Effizienz der Förderung führte zur Reform der EU-Strukturpolitik im Jahre 1988.
- Mit dieser **einschneidenden Reform 1988 (Beginn der Phase 2)** erfolgte eine Neu-Definition der Aufgaben der Strukturfonds, ihrer Koordination sowie der Koordination mit weiteren Instrumenten wie der Europäischen Investment Bank. Es wurden gemeinsame Kriterien für die Abgrenzung von Problemregionen, Umfang und Form der Unterstützung sowie Richtlinien für das Reporting entwickelt. Folgende Prinzipien der EU-Regionalförderung wurden eingeführt. Die Förderung wurde nunmehr:
 - auf 6 prioritäre Ziele fokussiert;
 - für integrierte, mehrjährige regionale Entwicklungsprogramme vergeben;
 - die im partnerschaftlichen Prozeß zwischen den nationalen Mitgliedsstaaten und der EU-Kommission erarbeitet werden;

⁴ Vgl. Armstrong 1994, Dinan 1994, Swann 1992, Molle 1994, Tsoukalis 1992.

- Die bereitgestellten Mittel der EU unterliegen dem Prinzip der Additionalität, sie dürfen also keine nationalen Budgetansätze ersetzen, sondern müssen zusätzlich für die Förderungsregionen bereitgestellt werden;
- Die Strukturpolitik muß mit anderen Politiken (bspw. Wettbewerbspolitik, Umweltpolitik) im Einklang stehen.

Die Abwicklung der EU-Strukturpolitik erfolgt nunmehr über die drei Fonds dem Europäischen Fonds für regionale Entwicklung (EFRE), dem Europäischen Sozialfonds (ESF) und dem Europäischen Fonds für Ausgleich und Garantie in der Landwirtschaft (EAGFL). Die von den Fonds bereitgestellten Mittel müssen jeweils um nationale Mittel ergänzt werden. Neben den Zielgebietsprogrammen werden über die Strukturfonds auch sogenannte Gemeinschaftsinitiativen finanziert, in denen die Kommission selbst Initiative ergreift und entsprechende Richtlinien entwirft, auf Basis deren von den Mitgliedsländern Programme eingereicht werden (bspw. Interreg, Urban, Leader).

Der Ansatz zur Programmplanung bedeutete einen Wechsel von der Unterstützung von Einzelprojekten zu einer Programm-Finanzierung. Die Reform brachte auch eine Verdoppelung der Mittel der Strukturfonds (von 6,3 Milliarden ECU 1987 zu 14,1 Milliarden ECU 1993) mit sich. Dabei sollten 80% der Mittel auf die am stärksten benachteiligten Regionen (sogenannte Ziel-1-Gebiete) konzentriert werden. Anfang der 90er Jahre kam es – im Zuge der Diskussionen über die Weiterentwicklung in Richtung einer Währungsunion - zu einer weiteren Aufwertung der Strukturfonds. So wurden die Mittel von 14,1 Milliarden ECU (1993) auf 27,4 Milliarden ECU (1999) aufgestockt. Im Zuge des Maastricht Vertrages kam es auch zur Einrichtung des Kohäsions-Fonds, der vor allem jene Länder (auf nationaler Ebene) unterstützen sollte, die weniger als 90% des BIP/je Einwohner des EU-Durchschnitts erreichten.

- Während die Reform 1993 nur geringe Änderungen mit sich brachte, markieren die mit der **AGENDA 2000** verbundenen Neuregelungen der EU-Strukturfonds den Beginn der **Phase 3**. Die Neuordnung der EU-Strukturpolitik beruht auf drei Eckpfeiler:⁵
 - Konzentration der Strukturfondsinterventionen auf prioritäre Ziele, Reduktion der Zahl der Gemeinschaftsinitiativen, der Größe der geförderten Gebiete sowie der geförderten Interventionen;
 - Vereinfachte und dezentralisierte Durchführung der Fonds mit einer klareren Aufgabenteilung zwischen der EU-Kommission und den Regionen;
 - Stärkung der Effizienz und Kontrolle durch den Rückzug der Kommission auf strategische Steuerung bei gleichzeitiger Intensivierung der Begleitung und Bewertung;

⁵ 21. Juni 1999 wurde die Reform der Strukturfonds formell vom Europäischen Rat beschlossen. Europäische Kommission (EK) (1999), Reform der Strukturfonds 2000-2006, Eine vergleichende Analyse, Juni 1999.

- Die mittlerweile auf die Zahl sieben angewachsenen Ziele der EU⁶ werden auf drei und die Zahl der Gemeinschaftsinitiativen wurde ebenso deutlich reduziert. Die drei vorrangigen Ziele umfassen:
- Ziel 1: Förderung der Entwicklung und der strukturellen Anpassung der Regionen mit Entwicklungsrückstand;
 - Ziel 2: Unterstützung der wirtschaftlichen und sozialen Umstellung der Gebiete mit strukturellen Schwierigkeiten (im allgemeinen ehemalige Ziel 2 bzw. Ziel 5b Gebiete, erweitert um städtische Problemgebiete);
 - Ziel 3: Unterstützung der Anpassung und Modernisierung der Bildungs-, Ausbildungs- und Beschäftigungspolitiken und –systeme (im allgemeine die ehemaligen Ziele 3 und 4).

Das neue Ziel 2 vereinigt die Ziele 2 und 5b der vergangenen Periode 1994-1999 und bezieht neu auch städtische Problemgebiete und vom Dienstleistungssektor stark abhängige Gebiete in die Umstellung ein. Es wurde zu einem strategischen Instrument zur Bewältigung des Strukturwandels ausgebaut. Durch das Ziel 2 werden insgesamt 18% der Gemeinschaftsbevölkerung erfaßt, was gegenüber der Periode von 1994-1999 eine Reduzierung um mehr als ein Viertel bedeutet. Zudem verhindert eine Übergangsförderung („phasing out“) aus dem Fonds für Regionalentwicklung bis Ende 2005 ein abruptes Ende der Strukturfondsförderung in diesen Gebieten.

Die für die EU-Strukturpolitik bereitgestellten Finanzmittel wurden in diesen Entwicklungsphasen kontinuierlich gesteigert. Die Ausgaben der Struktur- und Kohäsionspolitik machen nun mehr rund 35% des gesamten EU-Haushaltes aus: Für die Periode 1994-1999 wurden rund 169 Mrd. Euro, für die Periode 2000-2006 sind 195 Mrd. Euro vorgesehen.

4.4 Begleitung und Bewertung im Rahmen der Europäischen Strukturpolitik

4.4.1 Entwicklung der Bedeutung von Begleitung und Bewertung im Rahmen der Europäischen Strukturpolitik⁷

Im Folgenden wird ein kurzer Überblick über die Entwicklung der Ansätze sowie der Bedeutung der Begleitung von Strukturfonds-Programmen der Europäischen Union gegeben. Dabei zeigte sich, dass seit der Strukturfonds-Reform 1988 signifikante Änderungen, im Sinne der Entwicklung einer Evaluationskultur, zu beobachten war.

Bis in die späten 80er Jahre konzentrierten sich Evaluationen im Rahmen regionaler strukturpolitischer Maßnahmen auf den Norden Europas, Großbritannien und Deutschland, während in anderen Regionen kaum Evaluationen eingesetzt wurden. Insbesondere seit 1988 ist für beide Aspekte der Begleitung – dem Monitoring und der Evaluierung – eine

⁶ Mit dem Beitritt Österreichs, Finnlands und Schwedens wurde ein zusätzliches Ziel, die Förderung extrem dünn besiedelter Gebiete (Ziel 6) eingerichtet.

⁷ Bachtler/Taylor 1999.

starke Weiterentwicklung zu beobachten. Dies ist einerseits zurückzuführen auf die in den Verordnungen gestellten erhöhten Anforderungen, andererseits aber auch auf ein wachsendes Bewußtseins über die Notwendigkeit der Analyse der Effektivität und Wirkung der implementierten Förderungspolitik. Es gilt als weitgehend akzeptiert, daß die vor 1988 – also vor der Strukturfonds Reform – eingesetzten Begleit- und Bewertungsmechanismen inadäquat waren. Dies war auf mehrere Faktoren zurückzuführen:

- Fehlende Koordination innerhalb der Kommission und fehlende Evaluierungsstandards
- Delegation der Hauptverantwortung für Evaluationen an die Mitgliedsstaaten
- Relativ unkoordinierte Förderung von Einzelprojekte etc.

Durch die Reform 1988 wurden die Begleit- und Bewertungsmechanismen ausgedehnt und intensiviert, es wurde ein umfangreiches System zur Begleitung der Programme eingeführt: Dies äußerte sich in der Implementierung von Monitoringsystemen und von partnerschaftlich (zwischen EU- und Mitgliedsstaaten) besetzten Begleitausschüssen. Erstmals vorgesehen wurde auch eine ex-ante und ex-post Evaluierung. Während der Förderungsperiode 1989-1993 wurden diese Vorschriften erstmals umgesetzt. Allerdings waren sowohl Monitoring als auch Evaluierung jene Aspekte der Strukturfonds-Reform, denen in der tatsächlichen Umsetzung am wenigsten Beachtung geschenkt wurde.

Erst zu Beginn der 90er Jahre war ein Wendepunkt erkennbar. 1992 „situated at a turning point between the end of the first actions lasting several years and the preparation of the ensuing Community interventions, evaluation work received fresh impetus“ (MEANS 1993 zit. nach Bachtler/Taylor 1999). Über die Anforderungen der Verordnungen hinausgehend wurde sowohl von Seiten der Kommission als auch weiterer Institutionen der Gemeinschaft ein Bedarf nach umfangreicheren Informationen über die Implementierung der Programme und ihren Effekten erkannt. Letztlich wurde auch auf Basis des Maastricht-Vertrages dem Monitoring sowie der ex-ante und ex-post Evaluierung zunehmende Beachtung geschenkt.

Von Seiten der Kommission wurden Maßnahmen unternommen, um die Effektivität des Evaluierungsprozederes von 1994 an zu verbessern. Durch die Schaffung spezieller Evaluierungseinheiten innerhalb der Kommission wurden die Evaluierungsanstrengungen koordiniert. Zudem wurde das MEANS-Programm „Evaluation Methods for Actions of a Structural Nature“ ins Leben gerufen, um eine europäische Evaluierungskultur durch die Verbesserung der Qualität und Nutzbarkeit der Evaluierungen zu entwickeln.

Heute kann man inzwischen auf umfangreiche Erfahrung in der Programm-Evaluierung zurückgegriffen werden, ebenso wie auf Unterlagen, in denen versucht wurde, Evaluierungsstandards zu entwickeln. Darüber hinaus kann auch im Rahmen der Partnerschaft ein Lernprozess konstatiert werden, der sich in einem zunehmenden Bewußtsein über die Integration von Begleitung und Evaluierung als Steuerungsinstrument bei den in der Programmplanung und umsetzung beteiligten Akteuren äußert.

4.4.2 Rahmenbedingungen

Die Begleitung und Bewertung der Strukturfondsprogramme ist in den jeweiligen Verordnungen (VO (EG) Nr. 1260/1999 mit den allgemeinen Bestimmungen über die Strukturfonds und VO (EG) Nr. 1783, 1784, 1257 für die Strukturfonds) geregelt. Die Begleitung und Bewertung ist auf zwei Ebenen vorgesehen:

- Die Auswirkungen der Strukturfonds auf die Ziele der Europäischen Union werden durch die Kommission beurteilt. Dazu ist folgende Berichterstattung vorgesehen:
 - Jahresberichte über die Durchführung der Strukturpolitik und
 - Dreijahresberichte über die erzielten Fortschritte bei der Verwirklichung des wirtschaftlichen und sozialen Zusammenhaltes.
- Auf Ebene der einzelnen Programme erfolgt die Begleitung
 - durch partnerschaftlich besetzte Begleitausschüsse, die als Steuerungsgremium dienen;
 - eine Berichterstattung in Form von Jahres- oder Fortschrittsberichten;
 - ein Monitoringssystem, das Auskunft über Fortschritt der Programm-Umsetzung sowohl hinsichtlich der Mittelbindung, als auch des angestrebten Programm-„Outputs“ und die Programm-Wirkung geben soll.

Durch die Neuregelung der regionalen Strukturförderung der EU im Rahmen der AGENDA 2000 wurden die Anforderungen an die Mitgliedsstaaten sowohl hinsichtlich Klarheit und Kohärenz der Programme als auch hinsichtlich Begleitung, Bewertung deutlich gesteigert. Als wichtige Änderung gegenüber den früheren Programm-Perioden sind die Erweiterung der Partnerschaft auf regionale und lokale Behörden sowie die Wirtschafts- und Sozialpartner,⁸ die Einführung einer Verwaltungsbehörde, der das Programm-Management obliegt sowie die Definition und Festlegung der Zuständigkeit für die drei Bewertungsarten Ex-ante, Interim- und Ex-post-Bewertung in der Strukturfondsverordnung VO 1260/99. Hinsichtlich der Zuständigkeit für die Bewertungsverfahren sieht die Verordnung 1260/99 folgendes vor:

- Die Ex-ante Bewertung liegt in der Verantwortung der für die Vorbereitung der Pläne zuständigen Stellen in den Mitgliedstaaten.
- Die Interim Evaluierung liegt in der Verantwortung der Verwaltungsbehörde des Programms, in Zusammenarbeit mit der Kommission. Die Bewertung erfolgt durch einen externen Gutachter
- Die Ex-post-Bewertung liegt in der Verantwortung der Europäischen Kommission in Zusammenarbeit mit dem Mitgliedsstaat und der zuständigen Verwaltungsbehörde.

⁸ Die Partnerschaft gilt der Vorbereitung, der Finanzierung, der Begleitung und der Bewertung der Interventionen. Diese umfassendere Sichtweise von Partnerschaft bedeutete eine Neuerung gegenüber den früheren Vorschriften, wo Partnerschaft auf die Stufen der Programmplanung beschränkt war und die Mitglieder der Partnerschaft von den einzelnen Mitgliedsstaaten zu benennen waren.

Ex-ante-Bewertung (VO 1260/99 Artikel 41)

(1) Die Ex-ante-Bewertung dient als Grundlage für die Ausarbeitung der Pläne, der Interventionen und der Ergänzung zur Programmplanung und ist Teil von diesen. Die Ex-ante-Bewertung wird unter der Verantwortung der Behörden vorgenommen, die für die Ausarbeitung der Pläne, der Interventionen und der Ergänzung zur Programmplanung zuständig sind.

(2) Während der Ausarbeitung der Pläne und der Interventionen dient die Ex-ante-Bewertung der Analyse der Stärken, Schwächen und Möglichkeiten des betreffenden Mitgliedstaates, der Region oder des Sektors. Sie beurteilt mit Blick die Kohärenz der Strategie und der gewählten Ziele mit den Merkmalen der betreffenden Regionen oder Gebiete, einschließlich der Entwicklung ihrer Bevölkerung, sowie die voraussichtliche Wirkung der geplanten Aktionsprioritäten, wobei sie die spezifischen Ziele, wenn ihrer Art nach möglich, im Vergleich zur Ausgangssituation quantifiziert.

Die Ex-ante-Bewertung berücksichtigt insbesondere Lage im Bereich der Wettbewerbsfähigkeit und Innovation, der kleinen und mittleren Unternehmen, der Beschäftigung sowie des Arbeitsmarktes im Hinblick auf die europäische Beschäftigungsstrategie, der Umwelt sowie der Gleichstellung von Männern und Frauen, und umfaßt insbesondere:

a) eine Ex-ante-Bewertung der sozioökonomischen Situation, insbesondere der Tendenzen des Arbeitsmarktes, einschließlich in den Regionen mit besonderen Beschäftigungsproblemen, sowie der globalen Strategie im Bereich der Entwicklung der Humanressourcen und ferner der Art und Weise, wie diese Strategie mit der in den nationalen Aktionsplänen dargelegten Beschäftigungsstrategie verknüpft ist;

b) eine Ex-ante-Bewertung des Zustands der Umwelt in der betreffenden Region, insbesondere jener Umweltbereiche, die durch die Intervention voraussichtlich erheblich beeinflusst werden; die Vorschriften zur Einbeziehung der Umweltdimension in die Intervention und deren Kohärenz mit auf nationaler, regionaler und lokaler Ebene vorhandenen kurz- und langfristigen Zielsetzungen (z. B. Umweltmanagementpläne); die Vorschriften zur Sicherstellung der Einhaltung der Gemeinschaftsregelung im Umweltbereich. Die Ex-ante-Bewertung gibt eine soweit wie möglich quantifizierte Beschreibung der gegenwärtigen Umweltsituation und eine Abschätzung der erwarteten Auswirkungen der Strategie und Interventionen auf die Lage der Umwelt;

c) eine Ex-ante-Bewertung der Lage bezüglich der Chancengleichheit von Männern und Frauen auf dem Arbeitsmarkt und Gleichbehandlung am Arbeitsplatz, einschließlich der spezifischen Zwänge, denen jede der beiden Gruppen unterworfen ist; eine Abschätzung der erwarteten Auswirkungen der Strategie und der Interventionen, insbesondere bezüglich der Eingliederung der Männer und Frauen in den Arbeitsmarkt, der allgemeinen und beruflichen Bildung, des Unternehmertums der Frauen und der Vereinbarkeit von Familien- und Berufsleben.

Die Ex-ante-Bewertung umfaßt die Prüfung der Relevanz der vorgesehenen Durchführungs- und Begleitmodalitäten sowie der Kohärenz mit den Gemeinschaftspolitiken und der Berücksichtigung der indikativen Leitlinien. Dabei werden die Ergebnisse der Bewertungsarbeiten der vorangegangenen Programmplanungszeiträume berücksichtigt.

(3) Die Bewertung der in der Ergänzung zur Programmplanung vorgesehenen Maßnahmen dient dazu, ihre Kohärenz mit den Zielen der entsprechenden Schwerpunkte nachzuweisen, ihre spezifischen Ziele zu quantifizieren, wenn ihre Art dies zuläßt, und zu einem späteren Zeitpunkt gemäß Artikel 35 Absatz 3 Buchstabe b) die Relevanz der Auswahlkriterien zu prüfen.

Halbzeitbewertung (VO 1260/99 Artikel 42)

(1) Die Halbzeitbewertung mißt unter Berücksichtigung der Ex-ante-Bewertung die ersten Ergebnisse der Interventionen, ihre Relevanz und die Verwirklichung der angestrebten Ziele. Sie beurteilt außerdem die Verwendung der Finanzmittel sowie die den Verlauf der Begleitung und Durchführung.

(2) Die Halbzeitbewertung wird unter Verantwortung der Verwaltungsbehörde in Zusammenarbeit mit der Kommission und dem Mitgliedstaat vorgenommen. Sie betrifft die einzelnen gemeinschaftlichen Förderkonzepte und Interventionen. Die Halbzeitbewertung wird von einem unabhängigen Bewertungssachverständigen durchgeführt, dem Begleitausschuß des gemeinschaftlichen Förderkonzepts oder der Intervention vorgelegt und sodann in der Regel drei Jahre nach Genehmigung des gemeinschaftlichen Förderkonzepts oder der Interventionen der Kommission, spätestens aber am 31. Dezember 2003, im Hinblick auf die Überprüfung übermittelt.

(3) Die Kommission prüft die Relevanz und die Qualität der Bewertung auf der Grundlage von zuvor einvernehmlich zwischen Kommission und Mitgliedstaat festgelegten Kriterien im Hinblick auf die Revision der Intervention und die Zuweisung der Reserve.

(4) In Anknüpfung an die Halbzeitbewertung wird für jedes gemeinschaftliche Förderkonzept und jede Intervention eine Aktualisierung der Halbzeitbewertung vorgenommen. Mit Blick auf die Vorbereitung späterer Interventionen werden die Bewertungsarbeiten spätestens am 31. Dezember 2005 abgeschlossen.

Ex-post-Bewertung (VO 1260/99 Artikel 43)

(1) Die Ex-post-Bewertung soll unter Berücksichtigung der bereits vorliegenden Bewertungsergebnisse Aufschluß über die Verwendung der Mittel, die Wirksamkeit und Effizienz der Interventionen und ihre Auswirkungen geben, damit daraus Lehren für die Politik des wirtschaftlichen und sozialen Zusammenhalts gezogen werden können. Sie bezieht sich auf die Faktoren, die für den Erfolg oder Mißerfolg der Durchführung verantwortlich sind, sowie auf die ausgeführten Aktionen und die erzielten Ergebnisse, einschließlich ihrer Nachhaltigkeit.

(2) Die Ex-post-Bewertung wird unter Verantwortung der Kommission in Zusammenarbeit mit dem Mitgliedstaat und der Verwaltungsbehörde vorgenommen. Sie betrifft die Interventionen und wird von unabhängigen Bewertungssachverständigen durchgeführt. Die Ex-post-Bewertung wird spätestens drei Jahre nach Ablauf des Programmplanungszeitraums abgeschlossen.

4.4.3 Herausforderungen von Strukturfondsevaluationen

Versucht man nun die unterschiedlichen Rahmenbedingungen für die Evaluation von Strukturfonds-Programmen zusammenzuführen, so kann dies wie folgt zusammengefaßt werden:

Rahmenbedingungen

- Ein hoher Mitteleinsatz (35% des EU-Budgets) erfordert eine Überprüfung der Wirksamkeit der eingesetzten Programme. Umgekehrt bedeutet dies einen hohen politischen Rechtfertigungsdruck seitens der EU-Kommission insbesondere hinsichtlich des Beitrags zum Abbau der Disparitäten und der Schaffung von Beschäftigung
- Das Thema Evaluation hat folglich eine hohe Verbindlichkeit und einen hohen Stellenwert durch entsprechende Festlegungen in den Verordnungen; vorgesehen ist das Durchlaufen des gesamten Evaluierungszykluses: Ex-ante, Interim-, Ex-post-Evaluierung
- Aufgrund fehlender Vergleichbarkeit der Evaluationen und stark schwankender Qualität wurden umfangreiche Anstrengungen hinsichtlich von Systematisierungsversuchen und der Entwicklung von Standards (Leitlinien, MEANS-Programm etc.) unternommen, um damit Orientierungen für Evaluationen zu schaffen.

Anforderungen an die Evaluierung

- EU-Programme sind komplexe multi-sektorale Programme mit einer großen Zahl von Maßnahmen - wie in Österreich – zahlreichen den Programmen zugrunde liegenden Förderungsrichtlinien aus der Technologie-, Gründungs-, Tourismusförderung, Kultur, Forschung etc.
- Es ist eine große Zahl von Stakeholdern involviert, die Programme bauen auf einer Partnerschaft zwischen regionalen/lokalen Vertretern, Vertretern der nationalen Ebene und Kommissionsebene auf.
- Die Programme wirken keineswegs isoliert. Parallel – in Österreich auch manchmal über die selben Förderungsrichtlinie – werden weitere Maßnahmen für die selben Fördergebiete eingesetzt.

Diese Anforderungen kombiniert, mit jenen der EU-Kommission (die in die unten dargestellten Fragestellungen münden) führt zu einer schier nicht zu bewältigenden Komplexität von Evaluierungen von EU-Strukturfonds-Programmen.

4.4.4 Thematische Ausrichtung der Evaluation

Eine genauere Analyse der in den Verordnungen vorgesehenen Inhalte der unterschiedlichen Evaluierungsphasen einerseits sowie der von der Kommission vorgelegten Arbeitspapiere⁹ führt zu folgenden Themenstellungen, die in den Evaluationen der Strukturfondsprogramme zu behandeln sind:

⁹ Arbeitspapiere der Kommission: Die Ex-Ante-Bewertung der Strukturfondsinterventionen sowie: Indikatoren für die Begleitung und Bewertung: Eine indikative Methode sowie Verordnung (EG) Nr. 1260/1999 des Rates vom 21. Juni 1999 mit allgemeinen

Übersicht 3 Thematische Schwerpunkte der Strukturfonds-Evaluation

Schwerpunkt	Inhalt	Ex-ante	Interim	Ex-post
Erfahrungen über Wirkungen bisheriger Programme:	Beurteilung der Erfahrungen über Effektivität und Effizienz der Programme aus vorangegangenen Programm-Perioden und deren Berücksichtigung im aktuell vorliegenden Programm.			
Analyse der Ausgangssituation:	Beurteilung der sozio-ökonomischen Ausgangssituation und Ableitung eines Stärken/Schwächen-Profiles. Speziell zu beachten sind folgende Politikschwerpunkte: Umweltsituation, Chancengleichheit, Beschäftigung, Innovation, KMU?			
Strategiekonzept	Beurteilung der vorgelegten Strategien hinsichtlich ihrer Adäquanz und Kohärenz; Beurteilung des Strategiekonzeptes hinsichtlich der Politikbereiche: Umwelt, Chancengleichheit, Beschäftigung, Innovation, KMU.			
Zielsetzungen	Beurteilung der gewählten Zielsetzungen sowohl hinsichtlich der Kohärenz mit der Analyse als auch der Frage ob die Zielsetzungen ausreichend quantifiziert wurden und die gewählten quantitativen Ziele als realistisch zu beurteilen sind.			
Maßnahmen	Beurteilung der einzelnen Maßnahmen und des Maßnahmen-Mix auf Adäquanz und Kohärenz.			
Externe Kohärenz mit EU / nationaler Politik	Beurteilung der Kohärenz des Programmes hinsichtlich der EU-Politik-Bereiche sowie der nationalen Politiken.			
Finanztableaus Gewichtung von Strategien und Maßnahmen	Beurteilung der Adäquanz der gewählten finanziellen Gewichtung der einzelnen Programm-Teile und Maßnahmen.			
Indikatoren	Beurteilung der Aussagekraft, Konsistenz und Adäquatheit der bereits vorliegenden Indikatoren			
Partnerschaftsprinzip	Beurteilung, ob im Rahmen der Vorbereitung / Umsetzung des Programmes die Regelungen hinsichtlich der Partnerschaft entsprechend eingehalten wurden.			
Operationelle Rahmenbedingungen	Beurteilung der (geplanten) Verfahren der Programm-Umsetzung. <ul style="list-style-type: none"> - Beurteilung der Organisationsstrukturen hinsichtlich Klarheit, Transparenz, Informationsflüsse - Beurteilung der eingesetzten Mechanismen der Umsetzung (Ausschreibungen, Antragsverfahren, Publizität und Öffentlichkeitsarbeit) - Beurteilung der Projektselektions-Kriterien und -mechanismen - Beurteilung des Begleit(=Monitoring)systems. 			
Programm-Vollzug - Inputfaktoren - Finanzielle Indikatoren	Beurteilung des Standes der Mittelbindungen und Auszahlungen nach Programm / Prioritäten / Maßnahmen und Feststellen von Gründen für Abweichungen zu den geplanten Finanztabellen.			
Programm-Vollzug - Output-Faktoren	Beurteilung des Programmfortschritts hinsichtlich der Erreichung der festgelegten Output-Ziele. Rückschlüsse auf die Zielerreichung; Spezielle Beachtung der horizontalen Schwerpunkte der Union Umwelt, Beschäftigung, Chancengleichheit, KMU, Innovation.			
Wirkungsanalyse	Beurteilung der Programm-Wirkungen auf Basis der Programm-Ziele; Abschätzung von Brutto- bzw. Netto-Effekten.			

Quelle: eigene Zusammenstellung basierend auf VO 1260/99 sowie Arbeitspapieren der EU-Kommission

	Schwerpunkt der Evaluation
	Zu behandelndes Thema

Bestimmungen über die Strukturfonds jeweils abrufbar unter: http://www.inforegio.cec.eu.int/wbdoc/docoffic/working/sf2000_de.htm;

4.5 Ziele und Indikatoren

4.5.1 Die allgemeine Interventionslogik

Im Rahmen der Programmplanung erfolgt die Festlegung der Ziele der Strukturfondsprogramme sowie die Zuweisung der Fördermittel auf die verschiedenen Maßnahmen. Zwischen der Entscheidung über die Mittelverteilung und den Zielen besteht eine logische Verknüpfung. Diese kann auf zwei Weisen betrachtet werden: von oben nach unten oder von unten nach oben. In der Praxis werden bei der Programmplanung beide Betrachtungsweisen berücksichtigt:

Der enge Bezug zwischen Programmplanung und Begleitung wird graphisch in der folgenden Übersicht dargestellt und in Anlehnung an die Leitlinien der Kommission weiter ausgeführt.¹⁰

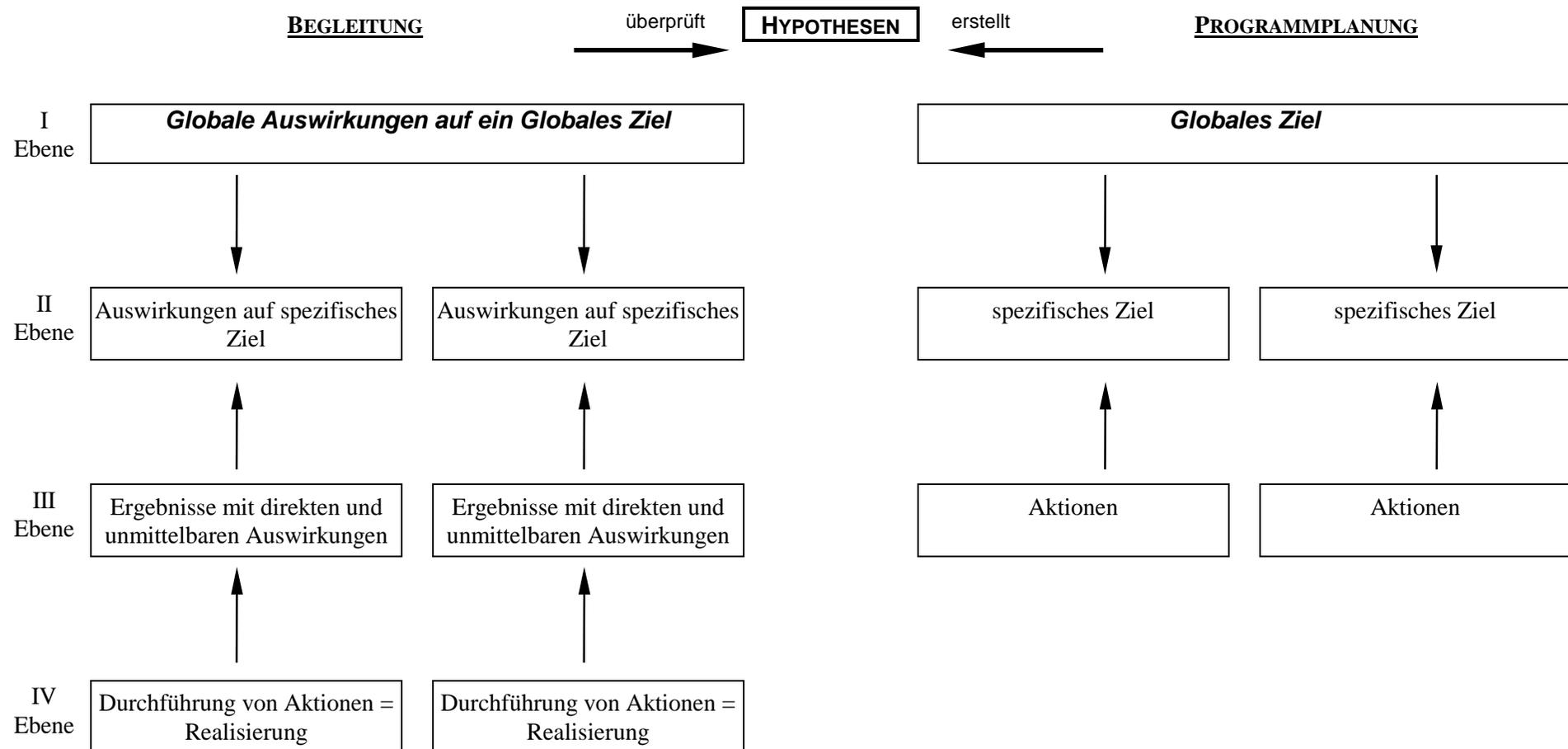
In der Phase der **Programmplanung** werden "die angestrebten globalen und spezifischen Ziele, die geeignetsten Aktionen und Maßnahmen, durch die sich diese Ziele erreichen lassen, sowie die zu erwartenden Ergebnisse dieser Aktionen und Maßnahmen"¹¹ definiert. Das Schema der Programmplanung entwickelt sich aus einer Betrachtungsweise von oben nach unten. Die generelle Ausrichtung der im Programm festgeschriebenen einzelnen Interventionen wird durch die Definition eines globalen Ziels bestimmt. Die genauere Festlegung und die spezifischen Eigenschaften jeder Intervention wird durch eine Anzahl spezifischer Ziele weiter charakterisiert. Auf der untersten Ebene befinden sich eine Anzahl von Aktionen, die zu positiven Ergebnissen führen sollten, um das jeweilige spezifische Ziel erreichen zu können. Die Programm-Planung erfolgt auf Basis von Hypothesen über Zusammenhänge von Maßnahmen und zu erwartenden Ergebnissen, die ihrerseits wieder zur Zielerreichung beitragen.

Ausgangspunkt für die **Begleitung** ist eine Betrachtungsweise einerseits von unten nach oben und andererseits von oben nach unten. Auf der untersten Ebene erfolgt die Durchführung bestimmter Aktionen und die Realisierung festgelegter Aufgaben. Die Wirksamkeit dieser Aktionen wird aufgrund der erzielten Ergebnisse beurteilt. Die Ergebnisse haben direkte und indirekten Auswirkungen auf das angestrebte spezifische Ziel. Die Auswirkungen insgesamt haben globalere Auswirkungen auf die globale Zielsetzung. Zum Zeitpunkt der Begleitung/Evaluation werden die für die Wirkungszusammenhänge entwickelten Hypothesen überprüft. Durch das vorne dargestellte Schema (logical framework) soll die Kohärenz der unterschiedlichen Interventionsebenen und deren stufenweise Verknüpfung dargestellt werden.

¹⁰ Diese schematische Darstellung erfolgt in Anlehnung an: Europäische Kommission (1995): Strukturfonds der EU 1994 - 1999. Gemeinsame Leitlinien für die Begleitung und die Zwischenbewertung. (DE/05/95/52070100.p00(FR)). S. 10 -11.

¹¹ Europäische Kommission (1995): Strukturfonds der EU 1994 - 1999. Gemeinsame Leitlinien für die Begleitung und die Zwischenbewertung. (DE/05/95/52070100.p00(FR)). S. 8.

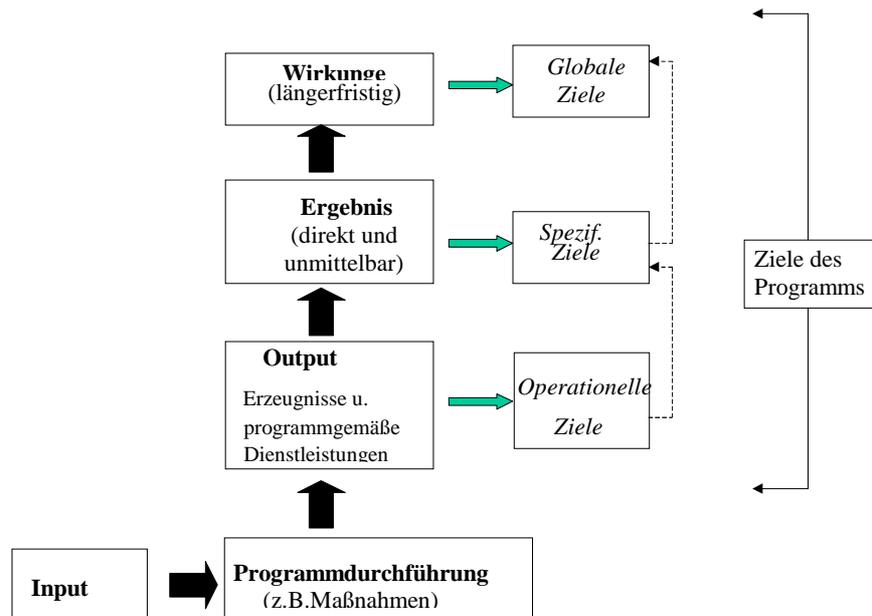
Übersicht 4: Begleitung und Programmplanung



Quelle: Europäische Kommission (1995): Strukturfonds der EU 1994 - 1999. Gemeinsame Leitlinien für die Begleitung und die Zwischenbewertung. (DE/05/95/52070100.p00(FR))

Anders formuliert bedeutet dies: Alle Interventionen werden in einem spezifischen Kontext in bezug auf ein bestimmtes **globales Ziel** geplant. Dieses Ziel ist wiederum für die Interventionsstrategie ausschlaggebend und führt zu einer Reihe **spezifischer**, im wesentlichen auf die Förderschwerpunkte ausgerichteter Ziele. Jedes spezifische Ziel wird zur Erreichung der operationellen Ziele mit Hilfe konkreter Maßnahmen umgesetzt.

Übersicht 5 Die Interventionslogik eines Programms



Quelle: EK (1999) Indikatoren für die Begleitung und Bewertung – eine indikative Methode

In der Begleitung bzw. Bewertung erfolgt die Betrachtung in umgekehrter Reihenfolge:

- Die Maßnahmen werden unter Einsatz von (personeller, finanzieller) Mitteln (**Input**) durchgeführt.
- Sie führen zu einem physischen **Output** (z.B. Kilometer gebauter Straße, bereitgestellte Ausbildungsplätze usw.), der die Fortschritte bei der Durchführung der Maßnahme verdeutlicht.
- Damit verbunden sind „**Results**“, die die (unmittelbaren) Auswirkungen der finanzierten Maßnahmen (z.B. verkürzte Fahrtzeiten, verringerte Transportkosten oder höhere Zahl „erfolgreicher“ Ausbildungsteilnehmer) auf die Begünstigten der Maßnahme darstellt.

- Die Leistungen werden schließlich anhand ihrer **Wirkung** im Hinblick auf die Erreichung der globalen und der spezifischen Ziele des Programms gemessen. Sie bilden die wichtigste Grundlage für die Bewertung des Erfolg oder Mißerfolgs der in Frage stehenden Intervention. Zu den **spezifischen** Wirkungen zählen u.a. eine Erhöhung des Warenverkehrs oder eine bessere qualitative Anpassung an die Anforderungen des Arbeitsmarktes. Der Input wird im Vergleich zum Output und damit in bezug auf die erzielten Ergebnisse und Wirkungen dargestellt.

Demzufolge werden die:

- **operationellen Ziele** anhand des Outputs ausgedrückt (z.B. Bereitstellung von Schulungsmaßnahmen für Langzeitarbeitslose);
- **spezifischen Ziele** anhand der Leistung ausgedrückt (z.B. Verbesserung der Beschäftigungsfähigkeit von Langzeitarbeitslosen durch Schulungsmaßnahmen);
- **globalen Ziele** anhand der Wirkung ausgedrückt (z.B. Verringerung der Zahl der Langzeitarbeitslosen).

4.5.2 Indikatoren

Die konkrete Begleitung und Evaluation erfolgt mit Hilfe von Indikatoren in quantitativer, deskriptiver oder qualitativer Form. Nach Nohlen/Nuscheler (1974) ist ein „Indikator eine Größe die auf etwas anderes hinweist, etwas anderes indiziert; in der Regel ist er (...) ein Teil oder eine Ersatzmaßeinheit oder ein indirektes Maß für nicht-meßbare Faktoren oder Ziele“. Es muß damit ein funktionalisierbarer, kausal fundierter Zusammenhang zwischen einer Kenngröße und einer Aussage bestehen.¹²

Auf eine gewünschte Quantifizierung der Ziele, trotz der Schwierigkeiten in der Umsetzung, wird seitens der Kommission besonders hingewiesen. "Soweit möglich, ist es immer wünschenswert, quantifizierte Messungen zu nutzen und die abgeschlossene Ausführung, erzielte Ergebnisse und Auswirkungen in quantitativer Form mit Hilfe von Indikatoren zu präsentieren."¹³ Die Komplexität sozioökonomischer Probleme erfordert jedoch gleichfalls deskriptive oder qualitative Informationen als ergänzende Erklärung. Im folgenden werden unterschiedliche Arten von Indikatoren, die in dieser Form im Rahmen von Evaluationen zum Einsatz gebracht werden, dargestellt.

¹² Zitiert nach Bremer Zeitschrift für Wirtschaftsforschung 4/94 S 52.

¹³ Europäische Kommission (1995): Strukturfonds der EU 1994 - 1999. Gemeinsame Leitlinien für die Begleitung und die Zwischenbewertung. (DE/05/95/52070100.p00(FR). S. 16.

4.5.2.1 Programmindikatoren: Input, Output, Ergebnis und Wirkung¹⁴

Um die Durchführung eines Programms begleiten und seine Performance im Hinblick auf die gesetzten Ziele beurteilen zu können, wird eine Reihe von Indikatoren benötigt, die im Vorfeld oder zu einem möglichst frühen Zeitpunkt der Programmdurchführung festgelegt werden müssen, um die entsprechenden Daten sammeln zu können. In den meisten Fällen werden ihnen quantitative Zielvorgaben zugewiesen. Es werden verschiedene Arten von Indikatoren unterschieden, die einerseits mit den Zielsetzungen auf unterschiedlichen Ebenen verknüpft werden und andererseits „Wirkungsketten“ abbilden sollen.

Es wird damit einerseits eine systematische Verknüpfung von Zielen und Indikatoren versucht und andererseits sollen „Wirkungsketten“ innerhalb des Indikatorensystems dargestellt werden. Demnach unterscheidet man zwischen:

- **Input-Indikatoren** geben an wieviele finanzielle Mittel eingesetzt wurden.
- **Output-Indikatoren** beziehen sich auf die unmittelbare Fördertätigkeit. Sie werden häufig in physikalischen oder finanziellen Einheiten gemessen (z.B. Kilometer gebauter Straße, Zahl der finanziell unterstützten Unternehmen, Zahl der bereitgestellten Ausbildungsplätze usw.)
- **Ergebnisindikatoren** beziehen sich auf die direkten und unmittelbaren Auswirkungen eines Programms. Die Indikatoren können physischer (Verringerung der Fahrtzeiten, Zahl der Auszubildenden mit erfolgreichem Abschluß) oder finanzieller Art sein (Investitionen des Privatsektors, Verringerung der Transportkosten usw.).
- **Wirkungsindikatoren** beziehen sich auf die Wirkungen des Programms über die unmittelbaren Auswirkungen auf die Endbegünstigten hinaus, die zum Teil erst mit einem time-lag Zeit beobachtbar werden.

Die Kommission hat in ihren Arbeitspapieren diesbezüglich auch Listen von Indikatoren vorgeschlagen, die in den Programmen 2000-2006 zur Anwendung kommen sollen (siehe dazu Anhang).

Mit Hilfe der definierten Indikatoren können Faktoren wie Wirksamkeit und Effizienz gemessen werden.

- **Wirksamkeit:** Hier wird das, was getan wurde, mit dem verglichen, was ursprünglich geplant war, d.h. es werden die

¹⁴ European Commission (EC) (2000), The New Programming period 2000 – 2006: methodological working papers, Working paper 3, Indicators for Monitoring and Evaluation, An indicative methodology sowie MEANS-Collection (1999), Volume 2.

tatsächlich erzielten Ergebnisse mit dem erwarteten oder erhofften Output, den Ergebnissen und/oder Wirkungen verglichen.

- **Effizienz:** Hier wird das Verhältnis zwischen dem erzielten Output, Ergebnis und/oder Wirkung und dem Input (insbesondere Finanzmittel) untersucht, der zu diesem Zweck eingesetzt wurde.

Wirksamkeit und Effizienz, d.h. Output, Ergebnis und Wirkung, können folglich für jede Phase des Programms oder Maßnahme gemessen werden.

4.5.2.2 Kontextindikatoren

Die allgemeine Strukturfondsverordnung legt fest, daß die Programme eine quantifizierte Beschreibung der derzeitigen sozio-ökonomischen Situation und des Entwicklungsrückstands bzw. -potentials für die betreffende Region enthalten sollen. Dieser Anforderung wird mit Hilfe von Kontextindikatoren Rechnung getragen, die Teil des Programmplanungsverfahrens sind. Ferner bilden sie die Grundlage für:

- die sozioökonomischen und strategischen Analysen (z.B. SWOT¹⁵-Analysen), die die Strategie-Findung für das Programm unterstützen;
- die Beobachtung der allgemeinen wirtschaftlichen Entwicklungen einer Region im Rahmen der Begleitung der Programme;
- die Festlegung quantifizierter Ziele;
- die Bewertung der sozioökonomischen Wirkungen des Programms.

Unter **Basisdaten** werden die Ausgangswerte verstanden, die als Grundlage für die Kontext- oder Wirkungsindikatoren herangezogen werden. Sie sollten mit Bezug auf die jeweiligen Programmziele festgelegt werden und können bspw. Angaben zur Zahl der Industriearbeitsplätze oder des Investitionsvolumens in bestimmten Sektoren umfassen.

4.5.2.3 Basisindikatoren

Aufgrund der großen Zahl von Interventionen sowie deren Heterogenität ist korrespondierend in der Regel eine entsprechend große Zahl von Indikatoren für die Begleitung und Bewertung festzulegen. Um einerseits die Begleitung der Programme zu vereinfachen und andererseits die Ergebnisse bestimmter Interventionen sowie die Ergebnisse des Gesamtprogrammes vergleichbar zu machen, wird von der EU-Kommission empfohlen, sogenannte Basis-Indikatoren einzusetzen.

¹⁵ Strengths – Weaknesses – Opportunities – Threats (Stärken - Schwächen - Chancen - Gefahren)

Basisindikatoren sind Indikatoren, die für Vergleiche zwischen ähnlichen Programmen oder Maßnahmen herangezogen werden können. Sie können in manchen Fällen auf eine höhere Ebene aggregiert werden. Von Seiten der EU-Kommission wird vor allem darauf gedrängt, für die von der Europäischen Union forcierten Prioritäten (bspw. Beschäftigung, Umweltwirkungen, Chancengleichheit, Innovation, Technologie, KMU) Basis-Indikatoren zur Anwendung zu bringen, um vergleichbare Informationen über die Förderungsprogramme hinweg sammeln zu können.

Generell wird vorgeschlagen, die Anzahl von Basisindikatoren gering zu halten, um sicherzustellen, daß sie für die Zwecke der Programmbegleitung und für vergleichende oder thematische Analysen geeignet und anwendbar sind. Sie können sich auf den Output, die Ergebnisse und/oder die Wirkungen beziehen.

Im folgenden wird eine Auswahl von von der Kommission vorgeschlagener Basisindikatoren dargestellt.

Beschäftigung

Die Beschäftigung ist eines der wichtigsten Ziele der Strukturfondsinterventionen. Bei der Quantifizierung der Wirkungen der Interventionen auf die Beschäftigung muß berücksichtigt werden, daß die Strukturförderung sowohl direkte (z.B. Arbeitsplätze, die von einem geförderten KMU geschaffen werden) als auch indirekte Auswirkungen hat (z.B. Arbeitsplätze, die durch eine neue Infrastruktur induziert werden). Zur genaueren Bewertung und Gegenüberstellung der tatsächlichen Beschäftigungseinflüsse sollten idealerweise die Nettoeffekte ermittelt werden. Diese werden anhand der Bewertung der Bruttoeffekte (von den Begünstigten gemeldete Gesamtwirkung) unter Berücksichtigung der Mitnahmeeffekte (ohne eintretender Beschäftigungsgewinn), der Verlagerungseffekte (Beschäftigungsverluste innerhalb desselben Zielgebiets) und der Multiplikatoreffekte (indirekte Effekte) geschätzt. Dies würde auch eine Doppelzählung vermeiden. Um die Genauigkeit dieser Schätzungen zu verbessern, sollte zwischen erhaltenen Arbeitsplätzen (die ohne die Intervention verlorengegangen wären) und neuen Arbeitsplätzen, die durch die Intervention geschaffen werden, unterschieden werden.

Als ein gemeinsamer Bezugsrahmen für eine genauere Bewertung der Beschäftigungseffekte sowie eine bessere Quantifizierung der Zielvorgaben für die erwarteten Beschäftigungswirkungen der Programme und Maßnahmen wird die Beachtung folgender Aspekte empfohlen:

- Umrechnung in FTE-Einheiten (Full Time Equivalentents)¹⁶
- Unterscheidung zwischen Brutto-/Nettobeschäftigungseffekten¹⁷
- Aufschlüsselung zwischen neu geschaffenen und erhaltenen Arbeitsplätzen
- weitere Aufschlüsselung zwischen Männern und Frauen, um Auswirkungen auf die Priorität „Chancengleichheit“ erfassen zu können.

Die Informationen liegen im Prinzip auf Projektebene vor, die dann auf Maßnahmen- und Programmebene aggregiert werden.

Die Entscheidung, ausschließlich Basis-Wirkungsindikatoren für die Bewertung der Beschäftigungswirkungen einzusetzen, hängt mit den mittel- bis langfristigen Zielen der Strukturfonds zusammen. Dennoch kann Beschäftigung auch anhand der direkten und unmittelbaren Auswirkungen auf Projektebene gemessen werden.

Basisindikatoren - Wirkung

<i>Indikatorart</i>	<i>Definition</i>	<i>Messung</i>
Beschäftigung (neu geschaffen)	Zusätzliche Arbeitsplätze in einem Unternehmen oder einer Einrichtung, die es ohne das Programm, die Maßnahme oder das Projekt nicht geben würde.	Anzahl (% der Zielvorgabe). Brutto/Netto FTE-Einheiten Männer/Frauen
Beschäftigung (erhalten)	Die Arbeitsplätze, die aufgrund des Projekts/Programms im Unternehmen oder in der Einrichtung erhalten bleiben und die ohne die Maßnahme verlorengegangen wären.	Anzahl (% der Zielvorgabe). Brutto/Netto FTE-Einheiten Männer/Frauen

Telekommunikation und Informationsgesellschaft

Basisindikatoren - Output

<i>Indikatorart</i>	<i>Definition</i>	<i>Messung</i>
Digitalisierungsq ote	Anzahl digitaler Telefonleitungen je 1000 Einwohner	Anzahl und Anstieg in %
Neugründungen im Bereich Informations- technik	Neugegründete Unternehmen, die Dienstleistungen im Bereich Informationstechnik anbieten (online-Dienste, elektronischer Handel usw.)	Anzahl

¹⁶ Teilzeitarbeitsplätze können auf der Basis zwei zu eins in Vollzeitarbeitsplätze (Full Time Equivalentents - FTE-Einheiten) umgerechnet werden. Arbeitsplätze werden als dauerhaft bezeichnet, wenn sie nach Abschluß der Intervention oder zumindest für einen im Vorfeld festgelegten Zeitraum nach Abschluß der Intervention erhalten bleiben.

¹⁷ Allerdings wird nicht definiert, wie diese Unterscheidung zwischen Brutto- und Netto-Effekte methodisch umgesetzt werden soll, sondern nur festgehalten, dass eine solche Unterscheidung getroffen werden sollte.

Basisindikatoren - Ergebnis

<i>Indikatorart</i>	<i>Definition</i>	<i>Messung</i>
Dienstleistungen im Bereich Informationstechnik	KMU, die Dienstleistungen im Bereich Informationstechnik entwickeln und anbieten	Anzahl
Dienstleistungen im Bereich Informationstechnik	Internet Einwahlpunkte ("Points of Presence") je Ortszone	Anzahl

Basisindikatoren - Wirkung

<i>Indikatorart</i>	<i>Definition</i>	<i>Messung</i>
Beschäftigung (neu geschaffen)	Weitere Arbeitsplätze, die aufgrund des geförderten Projekts im Bereich Informationsgesellschaft im Unternehmen oder in der Einrichtung geschaffen werden.	Anzahl und % aller geschaffenen Arbeitsplätze (Männer / Frauen)

Forschung & Entwicklung, Technologie und Innovation (FuE/I)Basisindikatoren - Output

<i>Indikatorart</i>	<i>Definition</i>	<i>Messung</i>
Vernetzung	Gemeinsame FuE-Projekte (Förderung von Kooperationsprojekten zwischen Unternehmen und Forschungsinstituten)	Anzahl
FuT/I-Projekte und Technologieerwerb	Unternehmen, die Fördermittel für FuT/I-Projekte und den Erwerb von Technologien erhalten	Anzahl

Basisindikatoren - Ergebnis

<i>Indikatorart</i>	<i>Definition</i>	<i>Messung</i>
Induzierte FuE/I-Investitionen	Induzierte Investitionen in FuE/I durch Unternehmen, die an geförderten Gemeinschaftsprojekten teilnehmen	Anstieg in %

Basisindikatoren - Wirkung

<i>Indikatorart</i>	<i>Definition</i>	<i>Messung</i>
Innovation	Neue Produkte / Prozessen, die von geförderten entwickelt und vermarktet werden	Anzahl
Beschäftigung (neu geschaffen)	Zusätzliche Arbeitsplätze, die aufgrund des geförderten FuE/I-Projekts im Unternehmen oder in der Einrichtung geschaffen werden.	Anzahl und % aller Arbeitsplätze (Männer / Frauen)

KMUBasisindikatoren - Output

<i>Indikatorart</i>	<i>Definition</i>	<i>Messung</i>
KMU-Förderung (bestehend)	Bestehende KMU, die finanzielle Unterstützung erhalten (d.h. Zuschüsse, Darlehen oder Kapitalbeteiligung, <u>außer</u> Beratung/Information)	Anzahl (davon mit weiblichem/männlichem Inhaber)
KMU-Förderung (neu)	Neue KMU, die finanzielle Unterstützung erhalten (d.h. Zuschüsse, Darlehen oder Kapitalbeteiligung, <u>außer</u> Beratung/Information)	Anzahl (mit weiblichem/männlichem Inhaber), Größe (Mikro / klein u. Mittlere) ¹⁸
Grenzüberschreitende KMU-Vernetzung	KMU, die an grenzüberschreitenden Projekten beteiligt sind	Anzahl

Basisindikatoren - Ergebnis

<i>Indikatorart</i>	<i>Definition</i>	<i>Messung</i>
KMU-Investitionen (Leverage-Effekt)	Direkte Privatinvestierung in Unternehmen, die finanzielle Unterstützung erhalten.	Mio. Euro und % der gesamten Fördermittel

Basisindikatoren - Wirkung

<i>Indikatorart</i>	<i>Definition</i>	<i>Messung</i>
Beschäftigung (neu geschaffen und erhalten)	Arbeitsplätze, die aufgrund der KMU-Förderung im Unternehmen oder in der Einrichtung geschaffen oder erhalten werden.	Anzahl und % aller Arbeitsplätze (Männer / Frauen)
Überlebensrate	Neue KMU, die finanziell unterstützt werden und nach 18 Monaten noch bestehen	Anzahl und % aller geförderten neuen KMU, (Männer / Frauen)

¹⁸ Als Mikro werden Unternehmen bezeichnet, die weniger als 10 Beschäftigten haben. Als Klein / Mittlere werden Unternehmen bezeichnet, die weniger als 250 Beschäftigten oder einen Umsatz von bis 40 Meuro haben. (Empfehlung der Kommission vom 30. April 1996 im OJ 107/30.04.96, S. 4).

4.6 Methodische Ansätze zur Evaluierung von Strukturfonds-Programmen der EU

4.6.1 Überblick über Evaluierungsmethoden

Für den folgenden Überblick über Evaluierungsmethoden wurden als Startpunkt die MEANS-Unterlagen herangezogen. Dabei zeigt sich, daß grundsätzlich eine breite Palette an möglichen Tools zur Evaluierung regionaler strukturpolitischer Programme zur Verfügung stehen – und im Einzelfall auch eingesetzt werden. Allerdings wurde durch die folgenden „empirischen“ Erfahrungen dieses Bild relativiert. In der tatsächlichen Evaluierungspraxis zeigt sich, dass ein deutlich eingeschränktes Spektrum an Methoden zur Datenerhebung sowie Datenanalyse zur Anwendung kommt.

Die folgende Zusammenstellung der Evaluierungsmethoden wird zeigen, dass aufgrund der Komplexität der Programme tendentiell stärker qualitative Ansätze verfolgt werden und eher ein bottom-up-approach befürwortet wird (vgl. dazu MEANS-Collection Volume 3). Man nähert sich also über die Einzelprojektergebnisse (die nach Möglichkeit aggregiert werden) an die Evaluation an, während Modell-Ansätze (makro-ökonomische Modelle, Input-Output-Modelle) kaum – und wenn, dann im Rahmen der Evaluierung von größeren Programmen eingesetzt werden. Aufgrund der Komplexität der Programm-Evaluation können zahlreiche Standard-Ansätze der Evaluation sowohl hinsichtlich der Datenerhebung (standardisierte Befragungen) als auch der Auswertung (multivariate Analyseverfahren, Vergleichsgruppen-Ansätze) nicht eingesetzt werden. Der folgende Überblick orientiert sich an den Standard-Schritten einer Evaluierung – wie sie im Means-Programm Volume 3 dargestellt wurden:

Tabelle 1 Standard-Schritte einer Evaluierung

<i>Strukturierung</i>	<ul style="list-style-type: none"> • Klärung der zu evaluierenden Effekte und Definition der Kriterien
<i>Datenerhebung / Beobachtung</i>	<ul style="list-style-type: none"> • Auswahl eines Instruments zur Datenerhebung • Abgrenzung des Evaluierungsraumes • Datenerhebung
<i>Analyse</i>	<ul style="list-style-type: none"> • Auswertung der Daten • Abschätzung von Effekten
<i>Bewertung</i>	<ul style="list-style-type: none"> • Bewertung vor dem Hintergrund unterschiedlicher Kriterien • Allgemeine Schlußfolgerungen

Im Folgenden wird ein Überblick über die im Rahmen des MEANS-Programmes vorgeschlagenen Evaluierungsmethoden strukturiert nach den einzelnen

Evaluierungsschritten gegeben. Für detailliertere Erläuterungen sei auf die MEANS-Collection Volume 3 „Principal evaluation techniques and tools“ verwiesen.

Tabelle 2 Methoden zur Strukturierung des Evaluierungsfeldes

SWOT-Analysis	Ableitung eines Stärken-Schwächen-Profiles für die Region. Meist ex-ante für die Definition der strategischen Ausrichtung des Programmes eingesetzt. Kann auch „interim“ eingesetzt werden, zur Überprüfung der Programm-Prioritäten.
Concept mapping of impacts	Definition der zu evaluierenden Effekte bzw. der verwendeten Indikatoren. Vor allem geeignet, bei komplexen Programmen, multiplen oder auch unklaren Zielsetzungen. Erhöht die Evaluierbarkeit der Programme. Ist vor allem geeignet, wenn ein gemeinsamer Konsens zwischen den Trägern des Programmes wichtig ist. Einsetzbar Ex-ante zur Definition der Zielsetzungen, aber auch Interim, Ex-post zur Präzisierung der Ziele und Evaluierungs-Indikatoren. Vergleichbar dazu „Colour vote“.
Impact Matrix	Einsetzbar, um einen einfachen Überblick über ein komplexes Programm zu erhalten. Maßnahmen und erwartete Wirkungen werden horizontal und vertikal aufgetragen. Jede Zelle wird mit einer Bewertung belegt, die die Bedeutung der Wirkungskomponente widerspiegelt. Kann eingesetzt werden, um sicherzustellen, daß möglichst alle Wirkungen erfaßt werden, zur Auswahl von Maßnahmen zur Detailevaluierung dienen. Darüber hinaus kann dieser interaktive Evaluierungsansatz auch herangezogen werden oder um nach erfolgter empirischer Arbeit zu einer „partnerschaftlichen“ Gesamteinschätzung der Programmwirkungen zu gelangen.

Tabelle 3 Tools zur Datenerhebung

Interviews	Interviews werden vor allem eingesetzt, um qualitative Einschätzungen und Meinungen von jenen Personen zu erhalten, die von einem bestimmten Programm, Projekt oder ihren Kontext, ihrer Implementierung oder Wirkungen betroffen sind. Wird meist für formative Evaluierungen eingesetzt bzw. meist mit explorativen Charakter.
Focus Group	Fokusgruppen beinhalten mehrere Personen mit ähnlichen Merkmalen und liefern qualitative Ergebnisse im Rahmen einer zielgerichteten, gelenkten Diskussion. Sie werden zunehmend mehr eingesetzt, vor allem im Zusammenhang mit konflikt-behafteten Themenbereichen. Die Integration der Programm-Träger in den Evaluierungsprozess über Fokus-Gruppen erhöht im allgemeinen die Akzeptanz von Evaluierungen.
Case Studies	Detail-Analysen an spezifischen Evaluierungsobjekte (Projekte, Entscheidungsprozesse, etc.), eingesetzt meist um ein umfassendes Verständnis von komplexen Situationen zu erlangen. Kann sowohl für explorative Arbeiten (Identifizierung von Wirkungen), als auch zur Überprüfung von Hypothesen (bspw. Überprüfung von erwarteten Wirkungen) eingesetzt werden ebenso wie zur Informations-Diffusion (good-practice-Beispiele). Im Rahmen von komplexen Programmen scheinen Case Studies das einzige adäquate tool zu sein, Effekte bei Projektträgern zu beobachten, wenn man die Komplexität der unterschiedlichen Programm-Komponenten adäquat berücksichtigen will.
Kaum geeignet:	Klassische Erhebungsinstrumente wie schriftliche Befragungen scheinen bei multi-sektoralen, komplexen Programmen aus Kosten-Nutzen-Überlegungen

Questionnaire Survey	kaum einsetzbar, da häufig sehr unterschiedliche Typen von Projekten bzw. Projektträgern vorliegen.
-----------------------------	---

Tabelle 4 Tools für die Datenanalyse

Geographische Informationssysteme	Ermöglichen die Verknüpfung und graphische Aufbereitung unterschiedlicher Datensätze. Sind vor allem geeignet, wenn die unmittelbare Lokalisierung der geförderten Projekten von Bedeutung ist.
Case Studies	Erlauben durch die Verknüpfung unterschiedlicher Datenquellen auch für komplexe Programme eine Analyse von Wirkungen.
Shift-Share-Analysis	Wird seit den 60er Jahren eingesetzt. Erfolgt durch eine Projektion von nationalen Entwicklungstrends auf die betrachtete Region. Wird eingesetzt, um eine „policy-off“ Situation und durch den Vergleich mit der tatsächlichen Entwicklung die Gesamtwirkungen eines Programmes abzuschätzen. Wird daher eingesetzt um die Gesamtwirkung einer „Poitik“ in einer Region abzuschätzen. Kann nur retrospective, im allgemeinen ex-post eingesetzt werden. ¹⁹
Input-Output-Modelle und makroökonomische Modelle	Input-Output Modelle und Makro-Ökonomische Modelle werden verwendet um die ökonomische Entwicklung von Regionen zu simulieren. Diese Tools werden in der Regel ex-ante eingesetzt und versuchen eine Abschätzung der Entwicklung mit bzw. ohne Programmimplementierung und liefern einen Gesamtabeschätzung der makroökonomischen Wirkungen. Der Vorteil von Input-Output-Modellen liegt in der nach Sektoren dissaggregierten Abschätzung von Wirkungen.
Kaum geeignet: Multivariate Analyseverfahren	Viele Standard-Analyseverfahren (bspw. multivariate Analysen) sind aufgrund der Heterogenität der Maßnahmen und der unterschiedlichen Einflußgrößen nur sehr eingeschränkt geeignet für die Evaluierung komplexer multi-sektoraler Programme.

Tabelle 5 Tools für die Gesamtbewertung

Expert Panels	Ermöglicht die Entwicklung einer synthetischen Beurteilung des Programmes. Unterschiedliche unabhängige Experten entwickeln gemeinsam eine Einschätzung über die Bewertung des Programmes. Es wird vor allem eingesetzt um einen Konsens über komplexe Fragestellungen zu erreichen. Expert Panels erhöhen im allgemeinen die Akzeptanz der Schlußfolgerungen der Evaluation da unterschiedliche Gesichtspunkte berücksichtigt und ein Konsens darüber entwickelt wird.
Multicriteria Analysis	Multicriteria Analysis wird eingesetzt, um eine vergleichende Bewertung heterogener Maßnahmen oder Projekte zu erhalten. Sie ist besonders hilfreich bei der Entwicklung einer Bewertung komplexer Probleme. Sie wird meist ex-ante bei der Auswahl von Projekten eingesetzt, weniger häufig bei Programm-Bewertungen.

¹⁹ Allerdings ist aufgrund der Tatsache, daß die Einflüsse in der Entwicklung von Makro-Größen wie Beschäftigung oder Wertschöpfung weit über das Förderprogramm hinausgehen, die Anwendung dieser Methode höchst umstritten. Sie sollte – wenn überhaupt - eher als unterstützende Methode eingesetzt werden.

Cost-Benefit-Analysis und Cost-Effectiveness Analysis	Beides sind traditionelle Bewertungskriterien, die jedoch für die Bewertung von komplexen, multi-sectralen Programmen kaum einsetzbar sind.
--	---

4.6.2 Erfahrungen aus der Evaluierungspraxis

Im Folgenden wird versucht in erster Linie auf Arbeiten des MEANS-Programmes bzw. der EU-Kommission beruhende methodische Überblick, im Lichte der Evaluierungspraxis auf seine praktische Relevanz hin einzuschätzen. Dies erfolgt vor dem Hintergrund, dass sich zwar für die im MEANS-Programm angeführten methodischen Ansätze jeweils vereinzelte Anwendungsbeispiele in einer Art Pilotanwendung finden, jedoch der state of the art in der Evaluierungspraxis deutlich davon abweicht. Dazu werden einerseits ein interner Review der Europäischen Kommission über Evaluierungsmethoden – der allerdings nicht nur Strukturpolitische Programme umfaßt – sowie eine Studie des Europäischen Policies Research Centre in Kooperation mit dem Fraser of Allender Institute im Auftrag des Scottish Office herangezogen. Dies wird weiters ergänzt um die Sichtung von Evaluierungsstudien aus Großbritannien und Deutschland sowie österreichischen Evaluationsstudien.

4.6.2.1 Review der Europäischen Kommission

In einer internen Analyse der EU-Kommission zum Vergleich von Evaluierungsmethoden (EC 2000b), in der 31 Evaluierungsberichte aus unterschiedlichen Politikbereichen analysiert wurden, kam man zu folgenden zusammenfassenden Einschätzungen:

- Die Evaluierungen fokussieren sehr stark auf Surveys und Interviews als zentrales Element der Datensammlung
- Die Angaben über die verwendeten Methoden und Datenquellen variieren nicht nur sehr stark, sondern weisen erhebliche Defizite auf; methodische Konzepte werden zum Teil falsch bezeichnet oder überhaupt falsch eingesetzt
- Es gibt erhebliche Defizite bei Performance-Indikatoren
- Avancierte methodische Ansätze wie ökonometrische Analysen werden kaum eingesetzt (in 2 von 31 Fällen).

Eine der gezogenen Schlußfolgerungen war letztlich, daß der Einsatz von quantitativen Methoden (insbesondere auch ökonometrischen Ansätzen) weiter verstärkt werden sollte

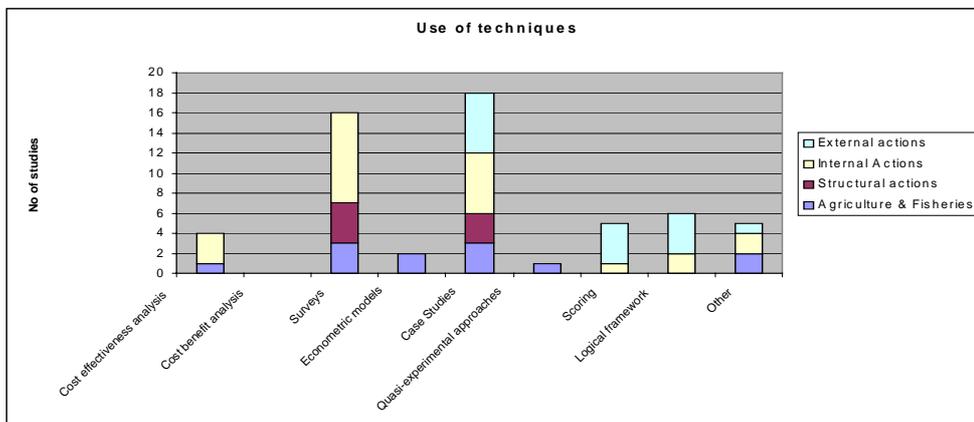
Interviews und Surveys sind bei weitem die am häufigsten eingesetzten Tools zur Datenerhebung und Auswertung, vielfach (nämlich 15 von 31 Evaluationen) verwendeten ausschließlich Surveys und Interviews als Grundlage der Analyse.

Allerdings beschränken sich die Erhebungen fast ausschließlich auf Programm-Manager oder Förderungsempfänger. Unabhängige Informanten, Mitbewerber oder auch nicht erfolgreiche Bewerber werden jedoch kaum in das Erhebungskonzept einbezogen. Im Rahmen der Anwendung von Survey/Interviews wurden folgende Schwächen festgestellt:

- Nur wenige Berichte zeigten, „that the rational behind opinions has been checked“;
- Es werden kaum subjektive Einflußfaktoren diskutiert, die die Antworten beeinflussen können;
- Die Zahl der Interviews ist häufig sehr gering; meist fehlen überhaupt Angaben über das Sampling.

Vor allem jedoch ist kaum nachvollziehbar inwieweit es sich um die präsentierten Einschätzungen / Ergebnisse um jene von Einzelpersonen (Einzelmeinungen) handelt, vom Evaluator aus einer Reihe von unterschiedlichen Beobachtungen abgeleitet wurde oder eine übereinstimmende Meinung einer Vielzahl von interviewten Personen darstellt.

Surveys und case studies sind die bei weitem am häufigsten eingesetzten Methoden. Lediglich im Rahmen der Evaluierung agrarpolitischer Fragestellungen werden eher quantitative Methoden eingesetzt.



Quelle: European Commission (2000b)

4.6.2.2 Methoden zur Evaluation von strukturpolitischen Programmen der EU in der Praxis²⁰

Durch das European Policies Research Centres wurde in Kooperation mit dem Fraser of Allander Institute der Universität Strathclyde im Auftrag des Scottish Executive eine Analyse unter dem Titel „Methodologies used in the Evaluation of Effectiveness of European Structural Funds: A Comparative Assessment“ durchgeführt. Ziel war es, ein Screening der Evaluierungspraxis im Rahmen der Strukturfonds-Förderungen durchzuführen und daraus Empfehlungen für die Handhabung der Evaluierungspraxis abzuleiten.

Die unten noch etwas ausführlicher dargestellten Ergebnisse lassen sich in folgende Punkte zusammenfassen:

- Die Evaluierungskultur im Rahmen der Bewertung von Strukturfonds-Interventionen ist in Großbritannien und in Schottland im speziellen im europäischen Vergleich als good practice anzusehen.
- Tow-down-Analysen (ökonometrische Modelle etc.) werden kaum und wenn dann meist nur in größeren Regionen (vorwiegend Ziel-1-Regionen) eingesetzt.
- Weitere methodische Ansätze bspw. Kontrollgruppen für die Wirkungsabschätzung im Rahmen des EFRE werden kaum verwendet. Es wurde lediglich ein Beispiel für einen Kontrollgruppenansatz in Irland identifiziert, während im ESF Kontrollgruppenansätze üblicher sind; Lediglich in einem Falle wurde als „avanciertere“ Methode jene der Shift-Share-Analyse eingesetzt.
- Die Abschätzungen von Impacts bauen in der Regel auf Surveys auf. Über die Surveys wird versucht, Abschätzungen von Mitnahme und Verlagerungseffekten vorzunehmen.

Zur Evaluierungspraxis in Großbritannien/Schottland

Schottland und Großbritannien können durchaus als positives Beispiel für die Umsetzung von Strukturfonds-Evaluierungen genannt werden. Innerhalb gemeinsamer Rahmenbedingungen für Großbritannien wurde seit mehr als fünf Jahren eine systematische Evaluierungsstrategie umgesetzt. Nichts desto trotz gibt es Defizite in der Wirkungsabschätzung. Allerdings mußte auch festgestellt werden, daß in keinem Mitgliedsstaat die Abschätzung der tatsächlichen Wirkungen der Strukturfonds-Förderungen in zufriedenstellendem Maße vorgenommen wurde.

Zur Abschätzung von Netto-Effekten

²⁰ Bachtler J./Polverari, L./Taylor S./Ashcroft B./Swales, K., Methodologies used in the Evaluation of Effectiveness of European Structural Funds: A Comparative Assessment, Study of the European Policies Research Centre and Fraser of Allander Institute, University of Strathclyde, Glasgow 2000

Ein Review von Ernst & Young von 75 Evaluierungsstudien zeigte, daß für die Periode 1989-1993 nur 18 Prozent, und für die Periode 1994-1999 nur 36 Prozent bezug zum Thema „Nettoeffekte“ genommen haben. Nur 28% versuchten Verlagerungseffekte, 37% positive indirekte Effekte abzuleiten. In Großbritannien werden offensichtlich intensivere Anstrengungen in diese Richtung gesetzt: in 42 % der Fälle (11 von 26 Studien) wurde versucht sowohl Verlagerung als auch positive indirekte Effekte abzuleiten.

Folgende Vorgangsweise wird in der Regel gewählt:

- Aussagen über Mitnahme- und Verlagerungseffekten beruhen meist auf Surveys, die an die Förderungsempfänger adressiert sind. Die Frage lautet dann, ob das Projekt auch ohne Förderung durchgeführt, zu einem späteren Zeitpunkt durchgeführt oder in kleinerem Umfang durchgeführt worden wäre. Auf Basis der Surveys wird versucht, Koeffizienten für Mitnahme- und Verlagerungseffekte zu schätzen.
- Als Alternative werden häufig Koeffizient und benchmarks für Mitnahme- und Verlagerungseffekte praktisch „off the shelf“ verwendet, was prinzipiell als problematisch einzuschätzen ist, wenn nicht berücksichtigt wird, daß diese nicht nur nach Sektoren stark variieren, sondern vom Design der Maßnahmen und dem Auswahlprozess der Projekte stark abhängig sind. Allerdings wurde in einer hier als good practice zu bezeichnenden Studie wurden 250 Evaluations-Studien analysiert, um daraus eine Liste von Koeffizienten zur Berechnung von Netto-Effekten bei Arbeitsplatzzahlen zu generieren. Diese Liste wurde in der Folge in einer Evaluierungsstudie eingesetzt (Yorkshire and Humberside Ziel-2-Evaluierung).
- Kontrollgruppen-Ansätze werden kaum verwendet, wenn meist im ESF (Ziel 3). Der Ernst & Young review über Evaluierungsstudien weist lediglich eine Evaluierungsstudie im Ziel-1-Gebiet Irland aus, welche einen Kontrollgruppenansatz verwendete, um Netto-Effekte für Maßnahmen des EFRE zu identifizieren.

Zu den Top-down-Analysen

Die Analyse zeigte – und dies stimmt prinzipiell mit Empfehlungen der EU-Kommission bzw. des MEANS-Programmes überein - daß Top-down-Analysen (ökonometrische Modelle etc.) vor allem in größeren Regionen (vorwiegend Ziel-1-Regionen) zum Einsatz kommen. Vielfach wurde die Entwicklung entsprechender Modelle auch gesondert von der EU-Kommission in Auftrag gegeben (siehe dazu Anhang). Diese Vorgangsweise stimmt mit den Empfehlungen seitens der EU-Kommission bzw. des MEANS-Programmes überein, die den Einsatz von Modellen vor allem für größere Programme empfehlen. Dabei wird argumentiert, daß vor allem das Programm-Volumen eine entsprechende Größe erreichen muß (ca. 1% des BIP)

und Modelle vor allem in kleinen „offenen“ Regionen ungeeignet seien. Die weiteren Argumente beziehen sich eher auf die Frage der „Wirtschaftlichkeit“ des Einsatzes als auf methodische Aspekte, da makro-ökonomische Modelle eher auf nationaler Ebene vorliegen, aufgrund der hohen Kosten und der Datenverfügbarkeit, werden sie auch eher auf dieser Ebene eingesetzt.

Der Bremer Ausschuß für Wirtschaftsforschung, der in seiner umfassenden Evaluierungsstudie zur Erarbeitung des Ansatzes für die Evaluierung der Ziel-2-Programme Bremen grundsätzliche methodisch-theoretische Überlegungen anstellte, kommt zu einer ähnlichen Schlußfolgerung. Dabei wird gezeigt, daß selbst in Nordrhein-Westfalen, in dem über 50 % der deutschen Ziel-2-Mittel (nach der Zieldefinition-alt) ausgegeben werden, der Förderungsanteil gemessen am BIP weit unter einem Promille, gemessen an der Wertschöpfung bei etwa 2 Promille liegen (BAW 1994).

Relativierend muß hier jedoch hinzugefügt werden, daß es sich bei dieser Betrachtung lediglich um jene der öffentlichen Förderungsgelder handelt. Die Relation zwischen öffentlichen Mitteln und eingesetzten privaten Mitteln lag beispielsweise in Österreich in der letzten Programmperiode in den Ziel-2-Gebieten zwischen 1:7 (Steiermark, Niederösterreich, Oberösterreich) und 1:12 (Vorarlberg) lag (ÖROK 1998). Wird dies mit berücksichtigt, so steigt das Programm-Volumen beträchtlich an.

EPRC/FAI kommen in ihrer Studie zur Einschätzung, daß wohl nicht die geographische Größe die bestimmende Größe für die Einsetzbarkeit eines Modell-Ansatzes sei. Zwar besteht - wie von der Kommission argumentiert - in größeren Regionen, insbesondere wenn das Fördergebiet mit einer ganzen Nation übereinstimmt wie bei den Kohäsions-Ländern, natürlicherweise eine verbesserte Datenlage und es werden mit höherer Wahrscheinlichkeit – aus welchen Gründen auch immer – Makro-Modelle vorhanden sein. Der Vorteil von Modell-Ansätzen liege in jedem Falle darin, daß ein getestetes Modell über ökonomische Zusammenhänge einschließlich von Markt-Interaktionen und spill-over Effekten den Überlegungen zu Grunde liegt. EPRC/FAI kommen daher zum Schluß, daß top-down-Ansätzen mehr Beachtung geschenkt werden sollte, die weniger in Konkurrenz als vielmehr als Ergänzung zu bottom-up-Ansätze eingesetzt werden sollten.

Im Anhang wird basierend auf der Studie von EPRC/FAI ein Überblick über die verwendeten methodischen Ansätze in Evaluierungsstudien in Schottland gegeben sowie Case-Studies über den Einsatz von Modellen in der Evaluierung von Strukturfonds dargestellt.

4.7

Monitoring

Das Programm-Monitoring stellt einen integralen Bestandteil der Begleitung der Strukturfondsprogramme dar und unterstützt die Programm-Umsetzung in vielfältiger Art und Weise:

- Die über das Monitoring bereitgestellten Informationen dienen im wesentlichen als Grundlage für die Jahres- und Fortschrittsberichte sowie die Berichterstattung an den Begleitausschuß.
- Das Monitoring sichert den Überblick über den Programm-Fortschritt und bildet damit auch mit eine Informationsgrundlage für Programm-Änderungen (bspw. Umschichtung finanzieller Mittel zwischen Programm-Teilen).
- Durch das Monitoring wird Transparenz über den Programm-Fortschritt geschaffen, was gerade vor dem Hintergrund partnerschaftlicher, multi-sektoraler Programme eine nicht zu unterschätzende Funktion darstellt.
- Die über das Monitoring bereitgestellten Informationen dienen als wesentliche Grundlage für Evaluationen.

Allerdings bestehen von den Programm-Verantwortlichen im allgemeinen vielfältige Vorbehalte, insbesondere gegenüber die durch das Monitoring zu erfassenden zusätzlichen Informationen und dem damit verbundenen administrativen Aufwand (am häufigsten wird mit einer nicht zumutbaren Belastung für die Betriebe argumentiert). Zum Teil entstehen gewisse Reserviertheiten aus der Angst vor zu viel Transparenz, insbesondere gegenüber der EU-Kommission. Daher werden die Vorschriften hinsichtlich Begleitung auch sehr unterschiedlich von den einzelnen Mitgliedern der „Partnerschaft“ interpretiert. Der nationale administrative Kontext hat im allgemeinen einen großen Einfluß darauf, wie die Vorschriften für Begleitung und Evaluierung interpretiert und angewendet werden können. Gleichzeitig variiert die Handhabung innerhalb der Mitgliedstaaten und ihrer Regionen ebenso stark.

Während sich die *finanzielle* Begleitung in den letzten Jahren in den meisten Ziel-2-Regionen deutlich verbessert hat, war die Erfassung von Performance-Indikatoren (Output, Results) in der Vergangenheit nach wie vor wenig zufriedenstellend (Bachtler/Taylor 1999). Nicht nur im Zusammenhang mit der Evaluierungspraxis, sondern auch mit Monitoring-Systemen können Großbritannien und ihre Regionen als good practice herangezogen werden. Im folgenden wurde daher in Zusammenarbeit mit dem European Policies Research Centre in Glasgow ein Review hinsichtlich des Einsatzes von Monitoring in den Strukturfondsprogrammen in Großbritannien durchgeführt. Im speziellen wurden zwei Case Studies für Western Scotland und West Midlands durchgeführt.

4.7.1 Monitoring in Großbritannien

The United Kingdom has long been considered as one of the leading Member States in the field of monitoring and evaluation. Great emphasis has been placed on assessing the efficiency and effectiveness of national regional policy interventions since the 1970s, and there is a well-established 'evaluation culture' in the UK. This culture has been carried over into the Structural Funds field, where the UK has actively complied with the EU regulatory requirements for *ex ante*, interim and *ex post* evaluation. At the Edinburgh Council in 1992, the UK took the lead in arguing for a higher profile to be given to monitoring and evaluating the Structural Funds, which subsequently led to an increased profile for evaluation across the EU.²¹

The UK commitment has been more evident in the field of evaluation than with respect to monitoring. While financial monitoring has always been considered important, the monitoring of physical outputs has not been systematic nor integrated. Over successive programming periods, evaluators across the UK have been highly critical of monitoring systems, the indicators contained in the SPDs, the interrelationships between indicators at different levels and the collection of data.

However, it is evident that the situation is improving. Since the mid-1990s, UK Government Departments have progressively developed guidance on the indicators and monitoring arrangements used by regions; particularly notable is the concept of 'rolling evaluation' initiated by the Scottish Office. Successive evaluation reports record the investment being made and the resultant improvements in the quality of indicators and data available.

The new monitoring requirements have acted as a further stimulus for change. Across the UK, new, integrated systems for monitoring physical outputs are being developed with emphasis on core indicators and common definitions. Devolution to Scotland and Wales has been a further imperative, requiring effective monitoring not just to satisfy the EU requirements but also to ensure transparency and accountability to the new political structures.

4.7.2 Context

The process of developing a monitoring and evaluation framework has been largely 'top down', driven by UK Government Departments, initially with a view to establishing a common base for the spatial Structural Fund programmes with common indicators and baselines. While the national authorities have had a key role in setting the framework and providing guidance, it has been the regional Government Offices (in England) and the Programme Management Executives (Scotland, Wales) that have given practical effect to the guidance and the increasing commitment of resources to evaluation. Arguably, for much of the past decade, the

²¹Bachtler J and Michie R (1995) A New Era in EU Regional Policy Evaluation: The Appraisal of the Structural Funds, *Regional Studies*, 29 (8).

exercise of evaluation in the UK has been driven by the need to fulfil the regulatory requirements and meet EC expectations.

Prior to 1994, Structural Fund monitoring and evaluation activity was limited and relatively weak in the UK. Early programmes lacked physical monitoring information, the data was inconsistent and not collected systematically. Objectives tended to be vague and difficult to assess, with a lack of intermediate objectives and targets capable of being used in performance measurement. The *ex post* evaluation of the 1989-93 programme in *North-West England* typified the general picture across the UK:²²

“Whilst financial monitoring information is largely complete and consistent (albeit with some retrospective reconciliation), data relating to jobs and other physical outputs from projects is patchy, inconsistent in its collection across the programming period and largely unverified. Databases of variable quality were maintained by the different governments...but difficulties have been experienced in trying to combine and centralise these.... Local databases are still the best source of information for parts of the programme [but] much of the information about output targets and claimed achievements...is only available by reference to raw data taken manually from available project files.”

In this early period, the UK government approach to Structural Fund implementation was to ensure that EC resources were fully utilised. Monitoring focused on financial monitoring to ensure that the audit trail was clear and that all available funds would be spent. *“The Programmes were not predicated on how projects would perform, on setting objectives and targets for individual projects or on evaluating the outputs and impacts at project and Programme level”*²³. There was also a lack of monitoring guidance for partners and applicants, leading to variations in the monitoring data collected within and across programmes. The quality of the monitoring data supplied was constrained by the ability of project managers to report accurately on their activities. Core terminology and concepts, even for some basic issues (eg. definitions of tourism and differentiation between jobs created and safeguarded), were used differently by different implementers and programme managers.

Many of these problems persisted in the 1994-96 programmes. Evaluators subsequently commented on the poor and problematic evaluation frameworks provided by the UK SPDs, most of which were flawed and limited. Examples of criticisms (taken from the *West Midlands*) were that the SPD was unclear about the output definitions used (eg. new/safeguarded jobs, direct/indirect employment), the

²² CRED (1996) *Ex Post Evaluation of the Objective 2 Programmes (1989-93) for North West England*, Report by CRED Research Unit, University of Liverpool and European Policies Research Centre to DG XVI, European Commission.

²³ Williamson E (1996) *The Development of the Concept of Rolling Evaluation of European Structural Funds in Scotland*, Paper Presented to the 1996 Berlin Means/DG XII Regional Policy Conference on Structural Fund Evaluation, 2-3 December 1996.

use of ill-suited performance indicators and no explanation of the derivation of indicator values.²⁴

However, over the 1994-96 period, efforts were made to improve responses in the UK both in line with the requirements of the new regulatory framework and the importance accorded to evaluation by the UK Government. A group of UK Government Departments worked together in the context of 'GAMESF', first producing guidance on prior appraisal and project appraisal and then on interim and *ex post* evaluations.²⁵ The latter work was undertaken jointly by the Scottish Office and Department of Environment, which together produced (in 1996) a common framework for the monitoring and evaluation of European Structural Fund interventions which took into account the Commission's guidance.

In *England*, the regional frameworks for monitoring and evaluation began to improve, although at different rates and without a common approach to issues such as definitions. In the *West Midlands*, the evaluators noted that "the output framework for the 1997/99 SPD represents a significant advance on its predecessor, although it too is some way from being a model framework for evaluators."²⁶ Outputs were set out more clearly within a more accessible framework, and the split between different job types was much more evident. However, several problems remained, notably the definition of non-job indicators and a lack of logical linkage between activity, output and impact indicators. The assessment of the 1997-99 *Greater Manchester* SPD was similar: the overall structure of the SPD is much clearer than its predecessor with a logical Measure structure which refers to the projected levels of activity to be supported, the associated outputs and anticipated impacts. Yet, the 1997-99 document retains the same inherent weaknesses as that for 1994-96 – it is not possible to track a defined path from the amount of resource through to activities, outputs and targets."²⁷

During this period, other developmental work was taking place in *Scotland*, involving the Scottish Office and the programme secretariats collectively to develop a 'rolling programme of evaluation'.²⁸ It began with definitional issues, aiming to construct a taxonomy that could be applied to all programmes. An 'evaluation plan' was agreed for each programme, indicating what would be done, when and how in relation to monitoring and evaluation. These plans indicated that it was necessary to produce baseline information against which progress could be tested, and that economic statistics would be collected at top-down (macro) level and project information at bottom-up (micro) level.

²⁴ DTZ Pidea (1999) *An Evaluation of the Objective 2 Programmes in the West Midlands, 1994-99*, Final Report to the West Midlands Objective 2 Programme Monitoring Committee, Birmingham.

²⁵ DoE/SO (1996) *Framework for the Evaluation of Area Based European Structural Funds*, Department of the Environment (Regeneration Directorate) and Scottish Office (European Funds Division).

²⁶ DTZ Pidea (1999) *op. cit.*

²⁷ Pion Economics/Roger Tym & Partners (1999) *Evaluation of GMLC Objective 2 Programme 1994-99, Manchester*.

²⁸ EPRC/FAI (2000) (2000) *Methodologies Used in Evaluating the Effectiveness of the Structural Funds: A Comparative Assessment*, Report to the Scottish Executive, European Policies Research Centre and Fraser of Allander Institute, University of Strathclyde, Glasgow.

Project information would be gathered in the first instance from application forms and subsequently from quarterly claim forms, for which a single format was to be elaborated across all seven programmes. For each programme, measure-level activity and intermediate output indicators were extrapolated. A uniform definition of activities and related intermediate outputs applicable to all programmes was identified, so that it would be possible to collect the same information - with the same meaning - across Scotland. The same exercise was carried out for macro-economic impact indicators, of which a list was drafted and circulated to all Monitoring Committees. Specific problems were also addressed, such as the difference between jobs created and safeguarded, and the problem of double counting of impacts.

At the end of this exercise, which took about two years (1994-96), a common language had been established and an effective monitoring system implemented for the Scottish programmes. The work undertaken made it possible to implement a monitoring system common to all programmes which could form the basis for a comparative assessment of programme performance across Scotland. Consistent core data were collected across all programmes which enabled the implementation of the planned rolling evaluation and, as a result, an overall global picture of the outputs, and in part the results, which had been achieved with Structural Fund support. Recent evaluations confirm that the indicators, baselines and targets used in the 1997-99 generation of Scottish programmes (eg. *Western Scotland* and *Eastern Scotland*) were relatively robust and built on sound rationale, avoiding some of the revisions and additional work required in other parts of the UK.²⁹

Looking across the UK, the 1994-99 programming periods saw a progressive improvement in the monitoring frameworks, the indicators used and data collection systems. However, so-called 'good practice' was still limited to a small number of regions. In most cases, the monitoring of physical outputs and impacts was inconsistent and rudimentary.

4.7.3 Monitoring Frameworks and systems

For the new round of programmes, as in other countries, the UK draft SPDs have a standard text for the implementing provisions which include a section on monitoring. Drawing on Articles 34-36 of the Council Regulation No.1260, these state that:

“Monitoring will be carried out by the Managing Authority assisted by the Programme Monitoring Committee. This monitoring will ensure the quality and effectiveness of implementation through assessment of progress towards achievement of the financial, physical and

²⁹ EKOS Ltd (1999) *Interim Evaluation of the Western Scotland Objective 2 1997-1999 Programmes*, Report for Strathclyde European Partnership Ltd, Glasgow. EKOS Ltd (1999) *Interim Evaluation of the Eastern Scotland Objective 2 1997-1999 Programmes*, Report for Eastern Scotland European Programme Executive, Glasgow.

impact indicators defined in the SPD. It will involve the organisation and co-ordination of the data relating to the financial, physical and impact indicators and those concerning qualitative aspects of implementation.”

In the case of the Scottish programmes this provision is modified by making monitoring the “joint responsibility of the Managing Authority, the Monitoring Committee and the PME and the Partners”, while in Wales, monitoring will be carried out by the Welsh European Funding Office on behalf of the Management Authority”. As in other areas of Structural Funds implementation, the approach to monitoring differs between the constituent parts of the UK. In *England*, the emphasis is on ensuring that project monitoring is carried out efficiently and effectively, so that national government can ensure compliance with the financial management, monitoring and control requirements of the EU. An integrated system for reporting physical outputs is currently being developed, but regions vary greatly in the sophistication of their monitoring arrangements and the range (and definition) of monitoring information collected. In *Wales*, the emphasis also appears to be on regulatory compliance with respect to financial and physical monitoring. By contrast in *Scotland*, the greater transparency required as a result of devolution has led to significant investment in new frameworks and systems for monitoring to collate standardised information, as well as a new evaluation strategy. These efforts are not only ensure compliance with the EU regulations but also provides accountability to the Scottish parliament and provides an effective programming tool.

4.7.3.1 Monitoring systems in England and Wales

At UK level, responsibility for the administration of the Structural Funds is split between the Department of Trade & Industry (DTI) and the Department of Environment, Transport & the Regions (DETR). While the DTI has lead responsibility for negotiation and liaison with the European Commission on Structural Funds, the DETR has operational management responsibility for implementation, and, as noted above, has the formal Managing Authority role. With respect to monitoring, DTI has an oversight and reporting role (eg. it produces a regular national monitoring report on the performance of programmes). The DETR is responsible for ensuring that the monitoring (and evaluation) of Structural Fund programmes in the English regions is carried out to specification and on time, and that it meets the requirements of the EU regulations, enabling the UK Government to provide reports to the European Commission. In response to the more stringent and demanding EU monitoring requirements, two sets of developments are currently underway at UK level: the development of a common system for collating information; and the provision of more detailed guidance to regions.

First, the DETR is developing a common structure for the collation of output information across the English regions. Since 1998, the UK has had an operational IT system for collecting claim information and financial data; this is able to provide data at every level (project, measure, priority, programme) and forms the basis for

the quarterly reports prepared by DTI on the financial progress of programmes. The system is now being extended to encompass output information. This involves drawing up a set of common 'core indicators' (see Section 4.7.4 below) for all of the Government Offices in the English regions to ensure consistent information on outputs. These are to be fed into a new, expanded IT system currently being prepared. Regions will be expected to feed in programme data, for both financial and physical output indicators, from the start of the programme period and keep it up-to-date on an on-going basis:

As yet, these procedures are still in their early stages. Only the first Objective 1 regions (eg. *Merseyside, South West England*) have entered new programme data onto the system. Once fully operational, central government should have a 'real time' overview of the progress of progress with respect to financial and physical outputs and the degree to which they are meeting targets. Access to the whole system will be restricted to DETR and DTI (each region will only have access to its own information), which will make regular reports on the national picture to the EC. Each region will also continue to send its own reporting information to the EC, according to a common template, covering the wider and more sophisticated range of monitoring information collected by many regions.³⁰

Second, more guidance is being issued to regions. This begins with the detailed description of management and monitoring arrangements included in the standard implementing provisions issued to all Government Offices in England (and adopted with some modifications in Scotland and Wales). Separate guidance has been issued on financial monitoring, including the Commission's audit manual, the financial control and audit requirements.

The main development concerning guidance, however, are the efforts being made to improve the overall framework for Structural Fund implementation by producing a publicly available 'Structural Funds Manual'. Drawn up on an inter-departmental basis, led by the Department of Trade & Industry, the Manual intends to codify existing best practice and to incorporate practical guidance on new policy. It is applicable to the whole of the UK, although parts are explicitly tailored to reflect the requirements of the English regions.

With respect to monitoring, the focus of the Manual is primarily on the project level. First, the Manual provides guidance on the outputs required in project applications. It requires outputs to be clearly stated, consistently defined, quantified wherever possible and capable of practical measurement. Targets must avoid double counting and ensure that the offer letter retains the power to withdraw grant if essential overall aims are not met. Second, the manual provides guidance on the way that project

³⁰The output information for Structural Fund programmes are also included in the annual 'service level agreements' between central government and each of the regional Government Offices which contain details on the anticipated expenditure and output targets across all areas of government spending in the regions. These agreements, revised each year and subject to quarterly reports, also contain a chapter on ERDF and ESF where each government department states what outputs are expected.

monitoring should be carried out. The principle objectives of project monitoring are defined as:

- to check that grant is used for the purposes for which it is made available and that the terms and conditions of grant are complied with;
- to ensure that grant is paid when it is clear that a claim meets the terms set in the offer letter and that the project is proceeding and can be expected to proceed according to plan;
- to deal with projects which fail to meet forecasts; and
- to provide the basis from which the evaluation of the benefits achieved by ERDF support can be assessed.

The conduct of project monitoring is left to the (English) regional Government Offices. The Manual sets out some general principles on the depth of monitoring (requiring that all projects are checked for compliance with the offer letter and that targets are met), the process of keeping track of the project, the benefits of visits to projects to check claims for payment, and the importance of recording information. Government Offices are reminded that “in addition to maintaining up-to-date information on individual projects, it is important that offices maintain and regularly update a database for project and programme outputs as an essential aid to project assessment and project management”. The integrated database will allow forecast, monitoring and outturn economic outputs for programmes to be recorded.

Lower down in the process, the guidance provides advice on monitoring at the project level. It advises on the range of information to be collected by ERDF and ESF projects as well as issues that need to be considered in relation to project objectives, baselines, indicators, targets, data collection and project management and systems. Critically, it emphasises that monitoring requirements need to be considered during the design of a project, otherwise “it will be an “afterthought” and may not be effective”. This emphasis on integrating project monitoring during the design stage of project development contrasts with the advice provided in the UK ‘Structural Funds Manual’; this focuses more on the formal principles and actions required for effective project monitoring, noting that “Monitoring should commence immediately following the applicant’s acceptance of the offer letter”.

Looking beyond monitoring, the auditing arrangements have also changed. The three main UK implementing departments have brought together their independent audit teams to create a single ‘Government Office Audit Team’ (GO-AT). Working as a full-time team, GO-AT will examine audit regional Government Office, approximately on a two-year cycle, to assess the way in which the Structural Fund programmes are being managed. While the main focus will be on financial flows, the team will also be looking at value-for-money issues (eg. examining project visits,

sample of claims etc) and make recommendations for management and monitoring of programmes.

4.7.3.2 Regional case study: West Midlands

It was noted earlier that the approach of central government has been to establish a common, integrated system for financial information and a set of core indicators for physical outputs, as well as providing guidance on monitoring procedures. It is, however, very much up to individual regional Government Offices to determine the extent of monitoring arrangements and the systems used for tracking and recording programme and project progress. Among the English regions, the most progress has been made by the bigger programmes, which have had more of an incentive to develop effective monitoring systems. Of these, the West Midlands is considered to be the leading English region in the sophistication of its monitoring arrangements.

As part of the preparation process for the 2000-06 programming period, the Government Office for the West Midlands (GO-WM) has spent 2-3 years developing a comprehensive monitoring system, partly in response to the criticisms of evaluators about the quality of monitoring systems and indicators in the last two SPDs (see Section 4.7.2).

The first step was to draw up a hierarchy of indicators for outputs, results and impacts. As noted below (see Section 4.7.4.1), these were based on the Commission guidance and the EKOS report, although they have been constrained by the lack of a finalised DETR list of core indicators. The second step was to prepare accurate baselines; in some cases, these were based on data sets already used in previous programmes, in other cases primary research was commissioned. Establishing baselines has not always been easy. Some data is not available at the programme level; county-level data, for example, does not reflect what is happening in the programme area, and some is not available in the form required, necessitating the use of proxies.³¹

The third stage was to establish methodologies for measuring impacts. Working together with other Government Offices (eg. East Midlands), GO-WM looked at previous programme performance in the 1994-96 and 1997-99 periods to establish what the programme should be able to deliver. The result is a set of tables for each action listing the projected results, outputs and impacts, as well as an assessment of overall programme impact according to a set of five indicators.

The final part of the development process involves creating the analytical framework for processing the data, for example to translate gross jobs into net jobs and the calculation of indirect jobs by assessing supply chains. Working with EKOS and GO-WM, the GO-WM have been developing methodologies (currently being validated) for each field of intervention that can exploit applicant data.

³¹ Some regions have established a regional observatory to generate appropriate data; this is being considered by West Midlands.

To operationalise this system, considerable effort has been made to improve the collection and recording of data. Originally this began in 1997-98 with a pilot project for Objective 3 where applicants were required to make monthly returns detailing the money spent, the beneficiaries (including a breakdown by category, such as by ethnic group, and comparison with projections) and the outputs (what was being delivered). In 1999, this was extended to Objective 2, although the workload soon began to cause problems for the secretariat. While applicants are required to make monthly returns to co-funders, it became clear that this frequency was too much of a burden for the GO-WM and the monitoring reports became quarterly.

In addition, the process of collecting monitoring information became more pro-active. Like other secretariats, the GO-WM has always undertaken desk-based monitoring, comparing claim data with application projections. The new system also now involves site visits to check on project progress. The aim was to have 100 percent coverage, ie. to visit every project, although this has had to be scaled back to one visit per applicant. The visits are proving popular with both GO-WM staff, who get to see projects 'on the ground' and have the opportunity to identify and disseminate good practice, and it is popular with applicants, who get an opportunity for more personal interaction with GO-WM staff and can discuss problems at an early stage. This process will be important for the new programme where GO-WM want to build up a good relationship with applicants to ensure they understand the more onerous financial management and monitoring requirements and ensure that decommitment does not occur.

The site visits are supported by a computerised database which contains all the monitoring information for all projects, as well as the assessments of the appraisal officers and an assessment of the 'capacity of applicants to deliver'. The system is relatively transparent in that the applicant is asked to counter-sign the GO-WM assessments. In future, this approach will be extended with a GO-WM risk assessment of every applicant. GO-WM will make a judgement of 'capability to deliver' based on three levels of risk:

- *acceptable*, where the applicant has a good record of delivery;
- *unknown*, where there is inadequate information on delivery, as in the case of new applicant organisations or old organisations which have been restructured or have new staff; and
- *unacceptable*, where the applicant is known as not having the capability or where there are significant concerns.

The GO-WM are expecting all applicants to go through this complicated process of risk assessments (involving checks on the records and record-keeping systems of the applicant and meetings with applicant organisations), but in general most applicants will only have to go through the process once. The appraisal officers will undertake the assessments in the first instance, although the accountants, financial audit and monitoring staff may also be involved.

In resource terms, GO-WM estimate that about 30 people have some responsibility for monitoring to varying degrees, perhaps equating to a full-time resource input of 10-15 members of staff.

Two final points are worth noting concerning the relationship between the monitoring of Structural Funds intervention and the monitoring of other, national economic development programmes in the region. First, monitoring by the European secretariat in GO-WM has benefited from survey research conducted in the region. For ESF, the GO-WM derived some insights from the two-yearly household surveys of c.25,000 people in the West Midlands, conducted by the Training & Enterprise Councils (TECs), with good sample coverage of different income and employment groups at a local level. The TECs also conducted a business survey of the views and expectations of businesses. These have not been conducted for Structural Funds purposes but have been part of the general monitoring activities of the TECs. Although the TECs are being replaced by new economic development organisations, the GO-WM hope to exploit the survey potential of intermediaries such as chambers of commerce by requiring a survey database to be created as a condition of business aid.

A second, wider issue is the interest on the part of the new regional development agency, Advantage West Midlands (AWM), has been keen to exploit the Structural Fund expertise of GOWM to develop a region-wide, comprehensive monitoring and evaluation framework encompassing all interventions. This initiative is still in its early days, but the AWM has developed a basic 'AWM framework' for the region, within which 'nests' the more sophisticated framework for the Objective 2 programme (which, in reality, is driving the AWM framework). AWM are encouraging colleagues in the region to develop their own 'nested' systems within the AWM framework. In the first instance, the intention is to have commonality, from the top downwards in: definitions of output; the data used; and how the data is collected. It is anticipated that this will make it easier to disentangle the outputs associated with different policy interventions. The ultimate aim is to get real impact evaluation, and to separate out the impacts of parallel socio-economic interventions such as the Structural Funds and the domestic Single Regeneration Budget. However, the proposal is still in its early days and a long way short of this goal. For example, as yet there is no input-output framework for the West Midlands. Nevertheless, over the next six years, the framework should certainly deliver what is needed for the management and delivery of the Objective 2 programme and improve the quality of monitoring and evaluation.

4.7.3.3 Monitoring systems in Scotland

In Scotland, there have been major changes to the context, framework and operational systems for monitoring. First, the strategic context has changed, with more political emphasis on the need for accountability. Second, a new framework for monitoring and evaluation is being developed by the Scottish Executive, which involves a new set of core indicators and monitoring systems. Third, more emphasis

is being placed on generating reliable data by working with applicants to improve project monitoring.

The devolution of government means that strategic and operational aspects of Structural Funds implementation are subject to much greater scrutiny and transparency than before. Over the past year, the Scottish Executive has undertaken a review of the implementation of Structural Funds through the Programme Management Executives and the Scottish Parliament's European Committee has reviewed each of the draft SPDs prior to their submission to the European Commission as well as conducting an inquiry into the implementation of the Funds in Scotland. It is anticipated that, in addition to the annual review of the programmes conducted by the Commission, the Scottish Parliament will carry out its own annual review of the progress and performance of programmes. Lastly, a high-level Scottish Structural Funds Forum is to be established, chaired by a government minister, to review on a regular basis how the programmes are being delivered and what adjustments need to be made to ensure that they link to policy developments.

Scottish ministers have made public political commitment to improving the effectiveness of the new programmes.³²

“Scottish Ministers have stated their aim of ensuring that Structural Funds are used effectively, efficiently and leave a lasting legacy in a way that complements their overall policy objectives.”

Furthermore, in response to the review of PMEs which advocated that more emphasis on reporting and monitoring³³, the Minister for Finance noted:³⁴

“The new processes will need to put more emphasis on re-view and evaluation than was the case in the past. To this end it will be important to develop effective monitoring and reporting systems. This will require a shift of resources to underpin the new systems.....the aim should be to do so within largely existing resources by simplifying the process insofar as possible, particularly in relation to application and subsequent project management. The more emphasis that is put on clear thinking about the value of the programmes and the less on routine paper work, the better for the way the new programmes are delivered.”

At an operational level, the Scottish Executive has been developing an overall evaluation strategy for the five Scottish Structural Fund programmes. This is intended to put monitoring activities into context and is based on the themes of

³² Scottish Executive (2000) *Framework for Economic Development in Scotland*, Edinburgh, p.16.

³³ Scottish Executive (2000) *Report of the Steering Committee of the Review of the Programme Management Executives*, 6 March 2000, Edinburgh.

³⁴ Formal Response by the Minister for Finance to the Report of the Steering Committee of the Review of the Programme Management Executives, <http://www.scotland.gov.uk/esf/response-00.asp>.

regulatory requirements, political accountability, overarching priorities and continuous feedback. It is important to note that monitoring and evaluation is not just being driven by the need to comply with Structural Fund regulations but also by other objectives:

- to provide 'proper accountability' to the Scottish Parliament and the wider population;
- to provide an 'essential tool' in achieving effective programme and project design and implementation;
- to influence future decisions on Structural Fund projects; and
- to enable effective programme review and evaluation through having consistent monitoring information.

Over the past six months, comparative research on Structural Fund evaluation methodologies has been undertaken³⁵ and work has been done on the strengths and weaknesses of monitoring systems in Scotland, supported by the publication of guidance for people and organisations involved in delivering Structural Funds³⁶ and a public political commitment to improving the effectiveness of the new programmes. An important part of the dissemination of information is provided by the new 'monitoring handbook' which sets out the up-to-date position on project monitoring and evaluation, an explanation of the 'core indicators' for the Scottish programmes (see below), the collection of information on the horizontal themes of the programmes, case studies of project monitoring and a checklist for partners, applicants and others (see Table 4.1).

³⁵ EPRC/FAI (2000) *op. cit.*

³⁶ Scottish Executive (2000) *Measuring Progress: A Handbook for Monitoring European Structural Fund Projects*, Edinburgh.

Table 4.1: Monitoring checklist for Scottish Structural Fund programmes

<ul style="list-style-type: none"> • Monitoring starts at design stage not at implementation stage. Collect all relevant information during the design phase which will provide your starting point on which to monitor performance. • Consider and establish the required baselines (ensuring that they will assist in monitoring against objectives) • Carefully chose the most relevant and appropriate indicators (again remembering they have to enable progress towards objectives to be monitored) • The Monitoring Framework tells everybody: <ul style="list-style-type: none"> • what will be monitored; • how it will be monitored; • who will undertake it (roles and responsibilities); and • when to carry out monitoring and gather data (considering perhaps milestones) • Liaise and make arrangements for data collection with partners - it is unprofessional and unwise to expect bodies to answer regular & perhaps lengthy requests for data if they have no prior knowledge of your requirements. • Consider the costs and systems required, e.g. new IT systems, new filing systems, any data which will have to be purchased, staff time etc • Ensure the project or scheme managers are encouraged to accept monitoring as a positive and useful tool • If operating a scheme, make sure that client groups understand fully what information will be requested from them, when it will be requested and in what form they will be expected to provide it • Make sure you incorporate the requirements of any external funders into the Monitoring Framework - check you are keeping the information and data in a way that can easily be extracted for all purposes and users • Review what you are doing to ensure the project is being delivered: 	<p style="text-align: center;"><i>According to schedule</i></p> <p style="text-align: center;"><i>To the appropriate clients</i></p>	<p style="text-align: center;"><i>To the required quality</i></p> <p style="text-align: center;"><i>In planned volumes</i></p>
---	--	--

4.7.3.4 Regional case study: Western Scotland

As in the UK as a whole, monitoring of the Western Scotland programme was almost non-existent prior to 1994. During the 1994-96 programme period monitoring began to improve, although was still rather sporadic and tended to respond to problems eg. non-activity at project levels. It was not until 1997 that a more strategic approach was undertaken; a key step was ensuring that the claim process was standardised (before 1997, organisations tended to submit claims in a format and over a timescale of their own choosing) with the quarterly submission of claims and payment in arrears (for ERDF)³⁷. This allowed programme managers to assess, more readily,

³⁷ ESF monitoring needed to be different because of the upfront nature of payments.

the rate of expenditure, the rate of outputs and the relationship between the two. Further, a programme of monitoring visits was undertaken, covering a minimum of five percent of projects by measure, all projects over one million euro and every project from a new organisation. This work revealed that, while financial expenditure could usually be verified, those projects that could verify physical performance were usually in the minority. Many of the problems associated with measuring performance – which are still present today – were attributable to lack of awareness or understanding of monitoring requirements, inadequate systems and/or a tendency to base reporting on forecasts. Following monitoring visits, reports were made to the Scottish Executive and were theoretically meant to be followed up by programme managers; in practice, there were/are few sanctions available in the event of underachievement or lack of monitoring information. (unless expenditure was/is ineligible).

For the new programmes (2000-2006), a new monitoring framework has been established at Scotland level (as noted above). In the Western Scotland programme baseline indicators have been established, drawn from work by the consultants EKOS. Programme managers have also looked at the average and best performance of projects in the 1997-99 period (for those indicators where data is available) and they have also identified 'exemplar' projects. A hierarchy of indicators has been established for each measure, with assumptions identified, collectively contributing to the impacts at priority and programme levels (see below).

The starting point for the new system will be a better informed partnership. All applicants receive the Scottish Executive monitoring handbook, and the monitoring and evaluation framework is described in detail in the SPD. When applicants prepare applications, they will be required to submit forecasts for the indicators relevant to the measure under which they are applying. In a first 'filter', programme managers will assess the adequacy of the indicator information provided and compare it with benchmarks; lower forecasts will be sent back. In the second filter, the applications will be scrutinised by the advisory groups, which (among other criteria) will consider whether the forecasts are credible. The successful applications will get an offer letter, on the condition that they abide with the monitoring compliance requirements. Project managers will be required to submit a claim every quarter (with 10 day's grace, or a penalty of suspension of the grant) and report on the expenditure and performance at each stage.

The system is intended to work electronically, via the Web, with drop-down menus that specify the precise wording and definition of each indicator (at application and claim stages) so that information inputs are completely standardised. This would allow 'real time' monitoring of both financial and physical indicators at measure, priority and programme levels (and potentially even Scotland level). However, there have been major problems with the IT system, and applications are currently based on Excel sheets with applicants typing in their own indicators, using their own

language, thus yielding a mix of information. It has been estimated that the IT system may not be operational until 2003.

4.7.4 Monitoring Indicators

4.7.4.1 Monitoring indicators in England

In England, as noted earlier, the DETR is drawing up a set of common 'core indicators' for the English Structural Fund programmes (see Table 4.2). This will be the first time that a consistent set of indicators has been drawn up and represents a wider range of standard outputs than has been used before. The indicators are based on the Commission working paper on monitoring indicators³⁸ and the core indicators being negotiated by the Commission with each Member State. Importantly, the regional Government Offices may also have their own - more extensive and different - set of indicators as determined by the partnerships and the nature of the interventions proposed in the SPDs. Also, they can use their own definitions for the indicators. However, regardless of the regional approach to monitoring, the various regional indicator lists must link into the DETR-specified core indicators.

Table 4.2: Proposed core indicators for Structural Fund programmes in England

<i>Definition</i>	<i>Unit</i>
Outputs	
SMEs support > £ 2000	Number
SMEs advised > 5 person days	Number
SMEs supported > £ 1000	Number
SMEs advised > 3 person days	Number
SMEs supported < £ 1000	Number
SMEs advised < 3 person days	Number
Large businesses - financial support	Number
Potential SMEs/SEs advised/supported	Number
New or refurbished floor space	Square metres
Rehabilitation of derelict land	Hectares
Buildings renovated	Number of buildings
Finance for environmental technology or products	Number of firms
Environmental audits	Number of audits
Environmental advice	Number of firms

³⁸CEC (1999) *Indicators for Monitoring and Evaluation: An indicative methodology*, Working Paper 3, The New Programming period 2000-2006: methodological working papers, DG Regio, Commission of the European Communities, Brussels.

Active labour market	Number of beneficiaries
Action on social exclusion	Number of beneficiaries
Lifelong learning	Number participating
Workforce flexibility	Number of employees in training
New road	Km
Improved road	Km
New rail	Km
Improved rail	Km
Financial inputs	
Private sector direct contributions	Thousand pounds
Private sector leverage	Thousand pounds
Results	
Additional jobs created	Gross permanent FTE M/F Measured as jobs filled (30 hours)
Jobs safeguarded	Gross permanent FTE M/F Measured as jobs filled (30 hours)
Increase in SME turnover	Gross increase in turnover of SMEs advised/supported
New start-ups	Of those advised/supported the gross number starting trading
Active labour market	Number of positive outcomes of which jobs
Active on social exclusion	
Lifelong learning	Number gaining a qualification
Workforce flexibility	Number gaining a qualification
Day visitors	Number
Night visitors	Number
Additional temporary jobs	Gross FTEs M/F excluding construction jobs
Additional construction jobs	Gross FTEs M/F
Additional sales	Thousand pounds (Gross)

Each English region is expected to input programme data when the SPD is approved. Then, as project applications are approved, the data is entered into the system on a continuous basis - first commitments, then claims. In addition to the standard categories for recording information (project, measure, priority), data input is required to record the 'fields of intervention' specified in the Commission guidance (ie. productive environment, human resources, basic infrastructure, technical assistance and innovative actions) by category and sub-category. Recording of fields of intervention will be determined by the regional secretariat not the applicant; this may be somewhat arbitrary in some cases where a project cuts across two fields of

intervention and the secretariat staff are expected to make a subjective split between fields in entering the data.

Beyond these core indicators, regions have a much wider set of monitoring indicators. The West Midlands' framework is populated by two sets of indicators.³⁹ First, there are so-called *context or condition indicators*. These must capture the conditions that the Objective 2 Programme objectives are designed to influence for the better. They will, therefore, have to be couched at the level of the West Midlands as a whole or at the sub-regional level (e.g. the Objective 2 area); or, sometimes be more narrowly defined in terms of localities and communities where these are targeted by the Objective 2 Programme. This category includes three sets of indicators that need to be established:

- **headline indicators**, which reflect the strategic vision for the region set out in the Objective 2 Programme. All actions taken in the name of the strategy should have direct or indirect, quantified or qualitative impacts in terms of these indicators;
- **pillar indicators**, appropriate to the strategic 'pillars' (or objectives) that underpin the Objective 2 Programme, which need to reflect the particular regional conditions whose improvement will contribute in distinct ways to the achievement of the WMES strategic vision; and
- **framework indicators**, relating to the Objective 2 Programme cross-cutting themes of operating values that should permeate every action taken in the name of the strategy and include sustainability and equal opportunities.

Second, there are response indicators needed to measure the direct and indirect consequences of the interventions made to deliver Objective 2 Programme objectives - i.e. the response to regional conditions. They will, therefore, have to capture the Objective 2 Programme inputs and their consequential outputs, results and impacts. These measure the *inputs* associated with the actions undertaken in the Objective 2 Programme, their associated *outputs* and the *results* and *impacts* they generate.

The following tables illustrate the range of indicators at programme level (showing impact by priority and horizontal theme) and the indicators for one of the pillars (priorities) - a diverse and dynamic business base (see Table 4.2 and Table 4.3).

³⁹ Draft West Midlands Objective 2 SPD, May 2000.

Übersicht 6 Programme Level Impacts by Priority and Horizontal Theme

Priority Indicator	Diverse and Dynamic Business Base	Learnin g and Skilful region	Infrastructur e and Conditions for Growth	Regeneratin g Communitie s	Total	Innovation	Sustainabilit y	Equal Opportunitie s	Total
Net Jobs Created	12,550	1,098	10,360	1,814	25,822	18,000	7,000	750	25,750
Net Jobs Safeguarded	4,340	555	6,774	574	12,243	9,000	3,000	250	12,250
Additional Net Value Added	640	149	481	160	1,430	1,150	150	130	1,430
Net Change in NVA per Capita	TBA	TBA	TBA	TBA	TBA	TBA			
Energy Cost Savings	TBA	TBA	TBA	TBA	TBA		TBA		
Increase in Training Participation		TBA			TBA			TBA	
Increase in SMEs Training		TBA			TBA	TBA		TBA	
New Products and Processes	TBA				TBA	TBA			
Targeted Beneficiaries into Employment				8,260	8,260			(8,260)	(8,260)
Funding Allocated (ME)	669,249	376,543	578,952	301,765	1,926,509				

Übersicht 7 Pillar indicators and baseline values

PILLAR INDICATORS AND BASE-LINE VALUES			
Pillar: Diverse, dynamic business base	Base-line values	Source	Commentary
% of SMEs introducing new products into the market (1)	1996 Spring: Objective 2=9%; WM=17% 1998 Autumn: Objective 2=15.4%; WM=14.6%	WMBS	Check for new Objective 2
Number of Internet Point of presence per local call area (3)		Indicator suggested by the EC	Availability of indicator for WM being investigated
Performance of global businesses in the region (5)	High growth firms per capita=25.6 in WM; 41.4 in SE; and 29.8 in Scotland (1997)	Trends Business Research	Scottish Enterprise Global Companies Enquiry 1998
Proportion of organisations (50+ employees) with IIP recognitions (4)	WM=13% (1997); 23% (1999); UK=10% and 22%	IIP UK	
Gross value-added per head in manufacturing (and services when available) and in target sectors (1)(4)	WM=£ 26.8k (1995); £ 30.3k (1997)=87%; 92% of UK;	Census of Production	Annual Business Inquiry will produce labour productivity estimates for services
R&D spend intensity in manufacturing and target sectors (5)	Manufacturing business R&D as % value-added: WM=3.7%; UK=6.7% (1997)	DTI Competitiveness Indicators	Availability of indicator for WMES target sectors being investigated
Investment in new development per capita (2)			Birmingham Economic Strategy being reviewed for WM indicator values
% employment and investment in target sectors (1)(5)			Availability of indicator for WMES target sectors being investigated
Start-ups & SMEs developing and commercialising IT services (3)		Indicator suggested by the EC	Availability of indicator for WM being investigated
% companies introducing new technology in last 6 months (1)	Prod.=37% (1996); 31% (1998) Design=13% (1996); 13% (1998) Office auto=53% (1996); 57% (1998)	WMBS	Check for new Objective 2
Export to sales/output ratios in target sectors (5)			Availability of indicator for WM being investigated
The following are the sources for the proposed indicators:			
(1) West Midlands Objective 2 (1997-99); (2) Birmingham Economic Strategy; (3) European Commission Working Paper 3; (4) DETR Core RDA indicator; (5) work in hand by SQW to specify the indicator and to provide base-line values; (6) DETR supplementary guidance on Skills Development Fund; (7) West Midlands Regional Planning guidance (Monitoring Report)			

The Commission services have been keen to ensure that the impact of the cross-cutting themes (gender equal opportunities and environmental sustainability) are addressed by the SPDs. In some cases, specialised monitoring indicators have been defined for these. Among the more notable examples is the draft *West Midlands* SPD which identifies three horizontal themes of innovation, environmental sustainability and equality of opportunity “that permeate all aspects of thinking” within the programme. With respect to sustainability, the SPD focuses on encouraging eco-efficiency within industry and business reducing their costs and increasing their competitiveness and increasing their competitive edge; developing sustainable transport and IT projects; and encouraging the best practicable approach to conserving water in situ. These objectives are not allocated particular measures but are dependent on actions in various parts of the programme. The ‘environmental sustainability’ outputs and impacts, therefore, are generic indicators, whose value are dependent on the emphasis placed on the various actions in the programme.

Table 4.3: Environmental Sustainability - Outputs, Results and Impacts in the West Midlands SPD

Input	Output		Results		Impact	
From other vertical priority funding streams	Urban renewal projects	57	New Jobs	20,000	Net Jobs Created	7,000
	Ha of Land reclaimed	157	Safeguarded Jobs	10,800	Net Jobs Safeguarded	3,000
	Rural renewal projects	70	New Sales	£ 600M	Net Value Added	£ 150M
	Sustainable transport Solutions	44	Safeguarded sales	£ 360M	Net Change in GVA	TBA
	Environmental Enhancements	63	Businesses improved	6,400	Energy Cost Savings	TBA
	ICT solutions for environmental issues	4,240	New Products	TBA		
	Development of Community Enterprises based on environmental issues	120	Sustainable Processes Introduced	TBA		
	Projects based on innovation in response to new environmental opportunities	50	Renewable Energy	TBA		
	Projects to co-develop business efficiency and environmental improvements	50	Reduction in Energy Costs	TBA		
			Reduction in water Consumption	TBA		
			Increase in Passenger Volumes on Public Transport	TBA		
			Remote/Home Workers	2,000		

The approach to equal opportunities is the same. The West Midlands SPD is proposing to target the: under-representation of disadvantaged groups in 'growing' and new industries, and the over-representation in vulnerable industries; the over-representation in certain jobs, characterised by lower rates of pay, lower skills, part-time employment and lower levels of training; lower activity rates and poor 'employability' as defined by key skills; and

barriers to access. Again, monitoring of equal opportunities is based on a generic set of outputs and impacts shared with measures being undertaken across the programme.

Table 4.4: Equal Opportunities - Outputs, Results and Impact in the West Midlands SPD

Strand	Output		Results		Impact	
Under-representation of disadvantaged groups in growing and new industries	Projects to improve qualifications and skills of disadvantaged groups to enable them to apply for jobs in growth industries	60	Disadvantage Groups with improved qualifications	4,000	Net increase of disadvantaged groups employed in growth industries	
	Projects to promote strategic links between employers, education and training establishments	20	EBPs supported	20		
Low Activity Rates and Poor Employability	Basic Skills training	Obj 3	NVQs gained	4,000	Net increase in level of qualification	4,000
	Confidence and Awareness raising	Obj 3	Trainees progressing to further positive outcomes	4,000		
Access Barriers to employment, training and provision of services	Projects to increase choice and relevance of training	20	Increased Training provision of direct relevance to the disadvantaged groups	20	Net increase in numbers of disadvantaged undertaking training	4,000
	Projects to improve access to training, employment and services through transport solutions and new technology	2000	Number of people accessing training, employment or services through the use of alternative solutions	5,000		
	Projects to assist employers change outdated attitudes, increase awareness and eliminate structural barriers and negative perceptions	20	Number of employers introducing and implementing equal opportunities programmes	2,000		

4.7.4.2 Monitoring indicators in Scotland

In Scotland, a key element of the evolving monitoring system is the development of 'Core Indicators' defined by the Scottish Executive in partnership with the Programme Management Executives. The purpose of the monitoring indicators is to enable achievements in each area to be put into context, to ensure that it is possible to measure 'like with like' and to make it possible to compare and contrast results. In the past each programme tended to have its own performance indicators. In order to make evaluation more effective it has been decided to establish a national list of core indicators. Five main themes are covered: social inclusion, economic and industrial indicators, demographics, labour market and the environment. A comprehensive system is being introduced across Scotland to ensure that these core indicators are collected comparably: in all areas; to the same level; using the same units of measurement; with each indicator defined in exactly the same way; and that practitioners, partners and other users are doing the same thing. The Core Indicators cover both ERDF and ESF and have been designed to take account of the horizontal themes, equal opportunities and sustainable development, as well as the Scottish Executive's priority of social inclusion.

Table 4.5: Core indicators for ERDF in Scotland

Project Level Core Indicators Indicator Name	Social Inclusion	Equal Opp- ortunities	Enviro. Activity
Jobs and Employment			
Gross new jobs created.	✓	✓	✓
Number of jobless people securing employment.	✓		
Number of young people under 25 years of age placed in employment.	✓		
Number of gross jobs created through self-employment.		✓	
Gross number of jobs safeguarded.		✓	
Assistance to Business			
Number of businesses assisted and instances of assistance, showing SMEs separately.	✓	✓	✓
Increase in sales in assisted businesses.	✓		
Hectares of serviced land created or enhanced, split by greenfield / brownfield sites.			✓ ✓
Square metres of business space created or enhanced – occupied after 18 months/3 years.	✓		
Private sector finance levered in by new assisted projects.			
Value of investment in R&D by assisted SMEs.			
Number of patents/intellectual property rights registrations by assisted SMEs.			✓
Number of organisations taking up e-commerce trading.			
Increase in visitor numbers.			
Number and SQM of new training/learning facilities constructed or upgraded.	✓		
Social Inclusion and Equal Opportunities			
Gross number of assisted organisations introducing active equal opportunity policies.	✓	✓ ✓	✓
Number of childcare facilities and places created.	✓	✓ ✓	
Number of organisations achieving recognised quality awards.	✓		✓
Environment			
K/W of new renewable energy infrastructure installed.			✓ ✓
Number of SMEs undertaking environmental management.	✓		✓ ✓
Number of strategic environmental/forestry partnerships funded.	✓		✓ ✓
Hectares of natural habitat under management.			✓ ✓
Increase in volume of waste recycled or reused.			✓ ✓

Table 4.6: Core Indicators for ESF in Scotland

<i>Project Level Core Indicators</i>	Social Inclusion	Equal Opportunities	Enviro. Activity
Indicators Number of beneficiaries of ESF assistance, show separately: <ul style="list-style-type: none"> • % of persons aged 16-24 receiving assistance • % of persons in this group unemployed for less than 6 months before assistance began • % of persons aged 25 and over receiving assistance • % of this group unemployed for less than 12 months before assistance began 	✓	✓	
% of those who complete their course	✓	✓	
% of those leaving ESF funded training for positive outcomes, split by end destination: <ul style="list-style-type: none"> • into full or part-time employment • self employment • full time FE/training/other Government training schemes 	✓	✓	✓
% of leavers who gain a qualification or part qualification	✓	✓	✓
% of beneficiaries who receive assistance specifically geared towards self employment	✓	✓	✓
Number of existing companies given direct assistance from ESF	✓	✓	✓
% of parents with children under 5, who are in employment 6 months after ESF assistance	✓	✓ ✓	

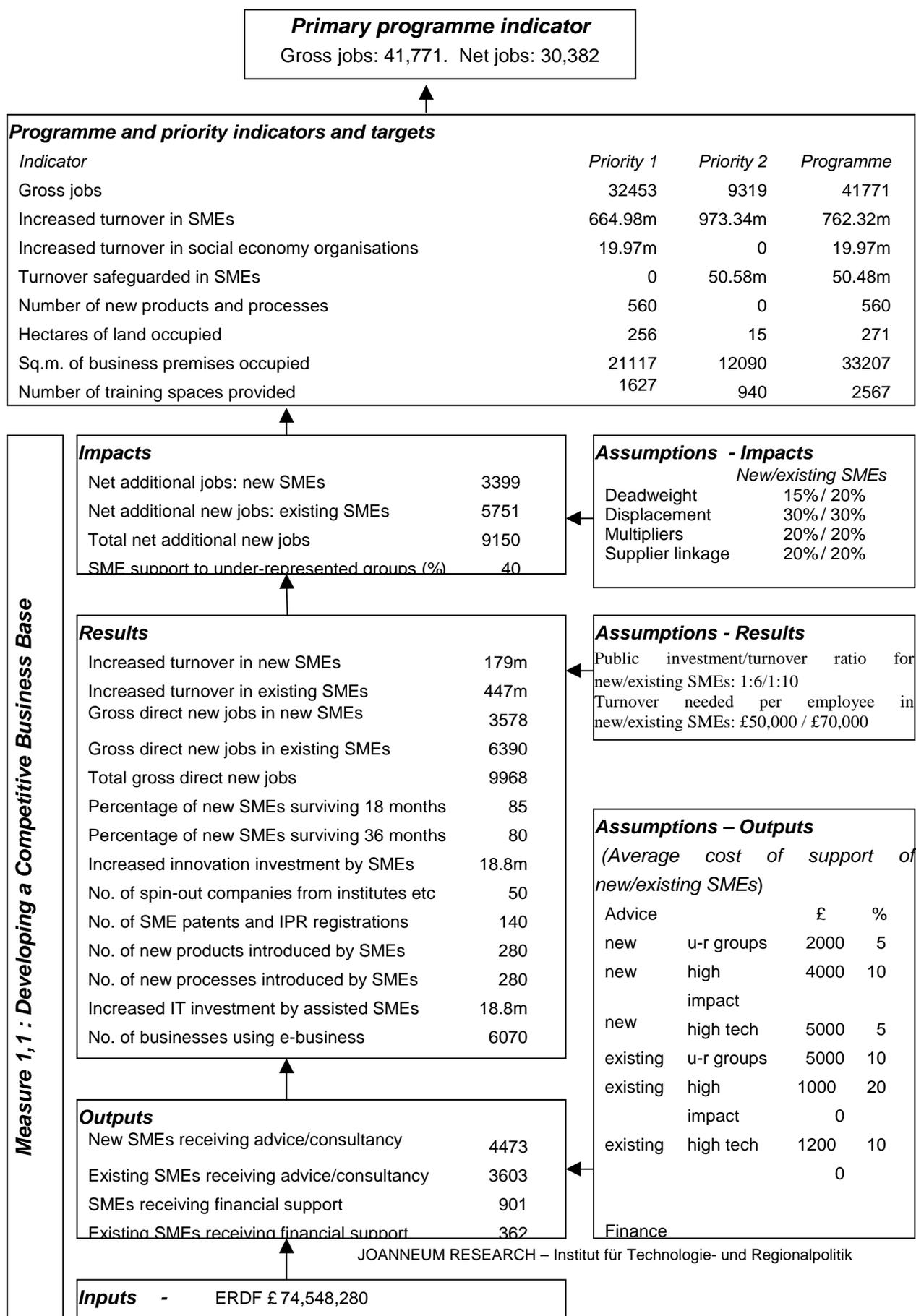
Using this system as a basis, and taking account of the EC guidance, the analysis of previous programmes, the monitoring requirements and resources of partner organisations and the EKOS Revised Indicators Study, the *Western Scotland SPD* presents a framework of indicators and targets developed for the programme (see Figure 1):

- At *programme level*, the primary indicator of Structural Fund intervention is job creation, and all projects supported under the programmes will report against gross jobs created. The programme is anticipated as creating 41,771 gross jobs. Taking account of effects such as dead-weight, displacement, multipliers and supplier linkages, the SPD projects net job creation of 30,382 jobs
- At both *programme and priority levels*, a range of other indicators has been identified – SME turnover, turnover in social economy organisations, new products and

processes, occupation of business premises, provision of training spaces, employment of residents in social inclusion areas, persons gaining training qualifications and guidance given to unemployed people.

- At *measure level*, the SPD then provides details for each measure relating to:
 - ❑ inputs (ERDF and ESF expenditure);
 - ❑ outputs eg. number of SMEs receiving support;
 - ❑ results eg. increased turnover in new SMEs;
 - ❑ impacts eg. net additional new jobs in new SMEs; and
 - ❑ the assumptions on which the calculations of outputs, results and impacts are based.

Figure 1: Monitoring and evaluation framework – Western Scotland Draft SPD



4.7.5 Conclusions

The monitoring of Structural Funds has made significant strides over the past decade, particularly over the last programming period. New frameworks are being introduced, driven not just by regulatory requirements but also by the desire for better programming and the need to use monitoring to ensure the effective and efficient implementation of programmes. In Scotland and Wales, there is a further dimension with the need for political accountability to the new devolved governmental structures. There is a 'step change' in the consistency of information to be collected with core or common indicators at UK and Scotland levels, although with varying degrees of standardisation of definitions and problems with the availability of data at the right spatial scales and for the right time periods. Hierarchies of indicators have been established – based on the European Commission's recommendations – with a separation of results, output and impact indicators, and new, integrated systems for data collection and analysis are being introduced. These changes will potentially improve significantly the quality of data available to evaluators, but it relies greatly on the willingness of partners and applicants (especially the important partners) to comply with the requirements and expectations for reporting and for the major technical problems associated with new systems to be overcome.

4.8

Kritische Reflexion der konzeptionellen Überlegungen zur Evaluierung der EU-Strukturfonds-Programme

Im Rahmen der europäischen Strukturpolitik liegen mittlerweile langjährige Erfahrungen bei der Förderung regional-politischer Programme vor, gekoppelt an intensiven Versuchen der Systematisierung von Aktivitäten zur Begleitung, Bewertung von Programmen. Umfangreiche methodische Arbeiten wurden dazu im Rahmen des MEANS-Programmes durchgeführt. Entsprechend Indikatoren-Sets wurden von der Kommission aufgebaut. Formal liefert das logical framework den notwendigen Rahmen, um eine systematische Verknüpfung der unterschiedlichen Zielebenen zu ermöglichen. Allerdings zeigt die Praxis der Programm-Planung sowie Evaluierung ein deutlich differenzierteres Bild:

- In der Praxis ist diese systematische Verknüpfung zwischen Zielhierarchien und Indikatoren kaum beobachtbar. Die Ziele der Programme werden meist sehr breit gewählt und sind häufig relativ unspezifisch. Diese Zielauswahl ist system-immanent, denn die Mitgliedsstaaten und ihre Regionen versuchen sich in der Programm-Planung so wenig wie möglich zu binden und größtmögliche Flexibilität für die Umsetzungsphase zu bewahren.
- Die zugeordneten Indikatoren sind nicht immer geeignet die Zielerreichung tatsächlich anzuzeigen. Häufig liegt eine Mischung aus Wirkungsindikatoren – mit ungeklärten methodischen Problemen – und einfachen Output-Indikatoren vor. Betrachtet man die Entwicklung im Zeitablauf der Strukturfonds-Förderungen sind zwar deutliche Verbesserungen erkennbar. Diese beschränken sich jedoch vor allem auf Regionen mit langer Erfahrung in der Umsetzung von Strukturfondsprogrammen (Bachtler/Taylor 1999):
 - In der ersten Förderperiode nach der Reform der Strukturfonds fehlte es weitgehend an in den Programmen festgelegten Zielen und Indikatoren
 - In der darauffolgenden Periode 1994-1996 konnten deutliche Verbesserungen erkannt werden: einige Programme stellten sogar Indikatoren auf drei Ebenen (Programme, Prioritäten, Maßnahmen) bereit. Allerdings gab es weiter große Defizite; unvollständige Indikatorenlisten, fehlende Quantifizierungen von Zielen, unklare Zielsetzungen etc.
 - Ein Review der 1994-1996 Programme (C3E 1995 zitiert in Bachtler/Taylor 1999) zeigte, daß die meisten Programme zwar auf positive Beschäftigungseffekte Bezug genommen haben, allerdings fehlten letztlich doch bei knapp der Hälfte der Programme entsprechende Indikatoren, sie waren nicht quantifiziert oder unvollständig. Wenn Ziele angegeben wurden, wurden diese nicht entsprechend spezifiziert. Angaben bezogen sich ausschließlich auf Brutto-Effekte. Ein Vergleich von den Planzahlen hinsichtlich Arbeitsplätze je eingesetzte Mittel führte zu extrem unterschiedlichen

Werten auch in vergleichbaren Programmen innerhalb eines Mitgliedsstaates.

- Nochmals deutliche Verbesserungen wurden von der Kommission in einem Review der Programme 1997-1999 festgestellt (EC 1997) vor allem hinsichtlich der Spezifizierung von Beschäftigungszielen und der Verwendung einer einheitlichen Terminologie. Allerdings wurde auch klar, daß die Erfahrungen letztlich noch sehr unterschiedlich sind und nach wie vor die verwendeten Indikatoren und die Quantifizierung von Zielsetzungen unbefriedigend blieb.
- Eine interne Analyse von Evaluierungen der Europäischen Kommission (2000) zeigte, daß die explizite Nutzung von Indikatoren auch Ende der 90er Jahre noch sehr stark variierte. Im allgemeinen wurden input, output und outcome Indikatoren eher implizit eingesetzt oder in 12 von 31 Fällen gar nicht verwendet. Es gibt nur wenige Beispiele von Evaluierungen, die diese Systematik explizit verwendet haben (bspw. Evaluierung der Gemeinschaftsinitiative LEADER). Allerdings dürften hier durch die Vorgaben der Kommission wesentliche Verbesserungen in der Programm-Periode 2000-2006 zu erwarten sein.
- Darüber hinaus ist zu beobachten, daß sich die Programme in den letzten 10 Jahren deutlich verändert haben: Dominierten in den ersten Programm-Perioden noch Infrastrukturmaßnahmen, so ist spätestens ab 1994 ein Übergang auf nachfrageorientierte, „weiche“ Maßnahmen zu erkennen. In zahlreichen Programmen finden sich Technologietransfer-Maßnahmen, Clusterbildung, Aufbau von Netzwerken und Förderung von Kooperationen etc. Allerdings hat der methodische Ansatz und die verwendeten Indikatoren mit den qualitativen Veränderungen der regionalen Strukturpolitik nicht schrittgehalten. Für neue, innovativere Ansätze sind in den Programmen so gut wie keine Indikatoren zu finden. Auch die umfangreichen methodischen Arbeiten bauen eher auf Indikatoren für traditionelle Infrastruktur oder betriebliche Investitionsmaßnahmen auf, während jedoch Ansätze für die Bewertung modernerer Förderungen nach wie vor weitgehend fehlen.

Einen interessanten Ansatz zur Evaluation regionaler strukturpolitischer Programme zeigt die Region Bremen. Hierarchische Zielsysteme wurden mit einem System qualitativer und quantitativer Indikatoren systematisch verknüpft. Es dominiert ein bottom-up-Ansatz, in dem die Wirkungen der einzelnen Projekte aggregiert werden.

Case Study Indikatoren-Tableaus Bremen

Ein good practise-Beispiel für eine gelungen Verknüpfung von Zielen unterschiedlicher Ebene und den Indikatoren stellt der von Bremen entwickelte und verfolgte Ansatz dar, der gleichsam auch durch ein entsprechendes Monitoring-System abgestützt wird.

Der Ansatz baut auf ein hierarchisches Zielsystem auf, welches in der Programm-Planung entwickelt wurde. Darauf aufbauend wird ein Indikatorensystem definiert, welches folgenden Anforderungen genügt:

- es muß sowohl quantitative wie auch qualitative Aussagen ermöglichen;
- die Indikatoren müssen aggregierbar sein;
- und die zeitliche Reichweite einzelner Auswirkungen berücksichtigen (nach kurz-, mittel- und langfristiger Dauer unterscheiden);
- es sollte dennoch „leicht zu handhaben“ sein - die Daten müssen also ohne übermäßigen Aufwand zu beschaffen sein.

Die bisher verwendeten Indikatoren-Tableaus folgen dabei noch nicht der Seitens der Kommission vorgeschlagenen Strukturierung in Input, Output, Ergebnis und Impact-Indikatoren. Da jedoch der Strukturierung diese Logik implizit zugrunde liegt, ist eine entsprechende formale Anpassung als unproblematisch einzuschätzen. Im Anhang werden die entsprechend des Bremer-Systems auf die österreichischen Programm-Schwerpunkte übertragenen Indikatoren-Tableaus dargestellt.

- Im Prinzip können im Rahmen der Strukturfonds-Evaluierungen eine breite Palette an Standardansätze der Evaluierung zum Einsatz kommen und zwar sowohl im quantitativen (Kosten-Nutzen, Korrelationen, Input-Output, ökonometrische Modelle etc.) als auch im qualitativen Bereich (Interviews, Case Studies, Focus groups, surveys, Delphi etc.). In der Praxis sind jedoch wesentliche Einschränkungen zu beachten: Aufgrund der Evaluation von Multi-Sektoren-Programmen können zahlreiche methodische Ansätze, die sich für die Beantwortung spezifischer Fragestellungen eignen mögen, kaum oder nur mit übermäßigem Aufwand eingesetzt werden. Was vielleicht theoretisch möglich wäre, ist in der Praxis aufgrund von Kosten- und Zeitbudgets vielfach nicht mehr umsetzbar. Folglich gibt es nur jeweils vereinzelte Beispiele für den Einsatz von anspruchsvolleren Methoden. Es sind kaum tatsächliche Wirkungsabschätzungen vorgenommen worden und wenn, bauen diese in erster Linie auf Surveys – im günstigsten Falle bei den Projektträgern - auf. Zudem werden zahlreiche methodische Probleme der Wirkungsanalyse zwar in den Grundlagen-Papieren diskutiert – sind jedoch fernab von einer Lösung. Messung von Mitnahmeeffekten und damit Netto-Wirkungen seien hier angesprochen.

- Die Evaluationen konzentrieren sich meist auf qualitative Fragestellungen während eine wirklich Wirkungsabschätzung – auch in ex-post-Analysen kaum zu beobachten sind (EPRC/FAI 2000). Vor allem sind jedoch die Fragestellungen der Netto-Effekte und dabei insbesondere die Bereinigung der Brutto-Effekte um Mitnahme-, Verlagerungs- und Substitutionseffekte ungelöste Probleme.
- Aussagen über Mitnahme- und Verlagerungseffekten beruhen meist auf Surveys, die an die Förderungsempfänger adressiert sind. Die Frage lautet, ob das Projekt auch ohne Förderung durchgeführt, zu einem späteren Zeitpunkt durchgeführt oder in kleinerem Umfang durchgeführt worden wäre. Auf Basis der Surveys wird versucht, Koeffizienten für Mitnahme- und Verlagerungseffekte zu schätzen.
 - Als Alternative werden häufig Koeffizient und benchmarks für Mitnahme- und Verlagerungseffekte praktisch „off the shelf“ verwendet, was prinzipiell als problematisch einzuschätzen ist, wenn nicht berücksichtigt wird, daß diese nicht nur nach Sektoren stark variieren, sondern vom Design der Maßnahmen und dem Auswahlprozess der Projekte stark abhängig sind.⁴⁰
 - Kontrollgruppen-Ansätze werden kaum verwendet, wenn meist nur im Rahmen der ESF-Förderung (Ziel 3). Lediglich eine Evaluierungsstudie im Ziel-1-Gebiet Irland, welche einen Kontrollgruppenansatz verwendete, um Netto-Effekte für Maßnahmen des EFRE war zu identifizieren.

⁴⁰ Allerdings wurde in einer hier als good practice zu bezeichnenden Studie wurden 250 Evaluations-Studien analysiert, um daraus eine Liste von Koeffizienten zur Berechnung von Netto-Effekten bei Arbeitsplatzzahlen zu generieren. Diese Liste wurde in der Folge in einer Evaluierungsstudie eingesetzt (Yorkshire and Humberside Ziel-2-Evaluierung).

4.9 Anhang

4.9.1 Ebenen der Evaluation: Fragestellungen der Evaluierung von Strukturfondsprogrammen

Im folgenden wird ein Überblick über jene Fragestellungen gegeben, die im allgemeinen im Rahmen von Evaluierungen von regionalen strukturpolitischen Programmen, die von EU kofinanziert werden, anzusprechen sind. Die Fragestellungen wurden abgeleitet aus der den Strukturfonds-Verordnungen sowie den Arbeitspapieren der EU-Kommission.

4.9.1.1 Ex-ante Evaluierung

Inhaltlicher Schwerpunkt	Fragestellung
Erfahrungen über Wirkungen bisheriger Programme:	<ul style="list-style-type: none"> • Welche Erfahrungen über Effektivität und Effizienz der Programme konnte in früheren Programm-Periode gewonnen werden und inwieweit wurden diese adäquat in den Programmen berücksichtigt?
Analyse der Ausgangssituation: Sozio-ökonomische Analyse Stärken/Schwächen-Profil	<ul style="list-style-type: none"> • Ist das Analysekonzept adäquat? • Wurden die relevanten Entwicklungen erfaßt? • Bilden die identifizierten Stärkefelder und Strukturdefizite eine ausreichende Grundlage zur Strategiefindung? • Können ergänzende „Baseline Daten“ für Kontextindikatoren (siehe dazu Indikatoren) bereitgestellt werden? • Gibt es entsprechende Analysen zu den horizontalen Schwerpunkten der Union: Umweltsituation, Chancengleichheit, Beschäftigung, Innovation, KMU
Strategiekonzept	<ul style="list-style-type: none"> • Wie sind die Strategien im einzelnen und der Mix der Strategien zu beurteilen? • Ist das Strategiekonzept mit der Analyse kohärent? • Sind die Strategien so konzipiert, daß synergetische Wirkungen erwartet werden können? • Wie ist das Strategiekonzept im Lichte der horizontalen Schwerpunkten der Union: Umweltsituation, Chancengleichheit, Beschäftigung, Innovation, KMU zu beurteilen?
Zielsetzungen	<ul style="list-style-type: none"> • Sind die gewählten Ziele aussagekräftig? • Lassen sich die Ziele aus dem Strategiekonzept ableiten? • Soweit Quantifizierungen vorliegen: Sind diese als realistisch einzuschätzen
Maßnahmen	<ul style="list-style-type: none"> • Wie sind die einzelnen Maßnahmen bzw. geplanten Instrumente sowie der Maßnahmenmix zu beurteilen? • Sind die Maßnahmen kohärent mit Zielen und

	<p>Strategiekonzept?</p> <ul style="list-style-type: none"> • Sind die Maßnahmen so konzipiert, daß synergetische Wirkungen erwartet werden können?
Kohärenz mit EU / nationaler Politik	<ul style="list-style-type: none"> • Inwieweit korrespondiert das Programm mit nationaler / EU-weiter Politik? • Greifen die unterschiedlichen Ebenen (EU / national / regional) in einander?
Finanztableaus Gewichtung von Strategien und Maßnahmen	<ul style="list-style-type: none"> • Ist die finanzielle Gewichtung der Schwerpunkte als adäquat einzuschätzen?
Indikatoren	<ul style="list-style-type: none"> • Welche Aussagekraft haben die bereits vorliegenden Indikatoren • Wie konsistent ist das Zielsystem und damit die innere Kohärenz des Programms • Sind sie aussagekräftig und geeignet den Programmfortschritt bzw. -wirkung abzubilden
Additionalität	<ul style="list-style-type: none"> • Wird im Programm entsprechend dargestellt wie das Prinzip der Additionalität eingehalten wird? Werden entsprechende Finanzdaten über Regionalpolitische Maßnahmen in der Vorperiode bzw. außerhalb des Programms angeführt?
Partnerschaftsprinzip	<ul style="list-style-type: none"> • Wie wurde das Programm entwickelt? • Welche Institutionen auf regionaler, nationaler Ebene wurden eingebunden?
Operationelle Rahmenbedingungen	<ul style="list-style-type: none"> • Wie sind die geplanten Verfahren der Programm-Umsetzung im Lichte der Strategien und des Maßnahmenbündels zu beurteilen? <ul style="list-style-type: none"> - Welche Organisationsstruktur liegt der Umsetzung zu Grunde (Beurteilung von Klarheit, Transparenz, Informationsflüssen) - Welche Mechanismen zur Umsetzung sind vorgesehen (Ausschreibungen, Antragsverfahren, Publizität und Öffentlichkeitsarbeit) - Welche Projektselektionsmechanismen sind vorgesehen - Wie kann das Begleit(=Monitoring)system beurteilt werden

Quelle: eigenen Zusammenstellung

4.9.1.2 Interim Evaluierung

Inhaltlicher Schwerpunkt	Fragestellung
Kontext-Analyse Sozio-ökonomische Analyse Stärken/Schwächen-Profil Politik-Rahmenbedingungen	<ul style="list-style-type: none"> • Wie hat sich die sozio-ökonomische Situation seit Programm-Start verändert? Hat das Stärken/Schwächen-Profil nach wie vor seine Gültigkeit? • Gibt es bedeutende politische, administrativ-institutionelle Änderungen, die Auswirkungen auf das Programm-Management haben?
Strategiekonzept und Zielsetzungen	<ul style="list-style-type: none"> • Sind die gewählten Zielsetzungen und Strategien im Lichte der Analyse im Punkt 1 nach wie vor gültig?
Programm-Vollzug - Inputfaktoren - Finanzielle Indikatoren	<ul style="list-style-type: none"> • Wie stellt sich der Stand von Mittelbindung und Auszahlungen nach Programm / Prioritäten / Maßnahmen / Projekten dar? • Wie kann die (inner)regionale Verteilung der Umsetzung beurteilt werden • Welche Projektstruktur kann identifiziert werden: Groß-versus Kleinprojekte; Öffentliche versus private Projektträger etc. • Welche Gründe für abweichenden Programm-Vollzug können identifiziert werden? • Lassen sich Rückschlüsse über eine fristgerechte und programmkonforme Umsetzung ziehen?
Programm-Vollzug - Output-Faktoren	<ul style="list-style-type: none"> • Welcher Programmfortschritt bei der Erreichung der festgelegten Output-Ziele ist zu beobachten? • Wie ist die Effizienz im Sinne von Output/Input-Größen zu beurteilen? • Lassen sich Rückschlüsse darüber ziehen, ob die festgelegten Ziele erreicht werden können? • Spezielle Beachtung der horizontalen Schwerpunkte der Union Umwelt, Beschäftigung, Chancengleichheit, KMU, Innovation
Operationelle Rahmenbedingungen	<ul style="list-style-type: none"> • Wie sind die angewendeten Verfahren der Programm-Umsetzung im Lichte der Strategien und des Maßnahmenbündels zu beurteilen? <ul style="list-style-type: none"> - Welche Organisationsstruktur liegt der Umsetzung zu Grunde (Beurteilung von Klarheit, Transparenz, Informationsflüssen) - Welche Mechanismen zur Umsetzung sind vorgesehen (Ausschreibungen, Antragsverfahren, Publizität und Öffentlichkeitsarbeit) - Welche Projektselektionsmechanismen sind vorgesehen - Wie kann das Begleit(=Monitoring)system beurteilt werden

Quelle: eigenen Zusammenstellung

4.9.1.3 Ex-post Evaluierung

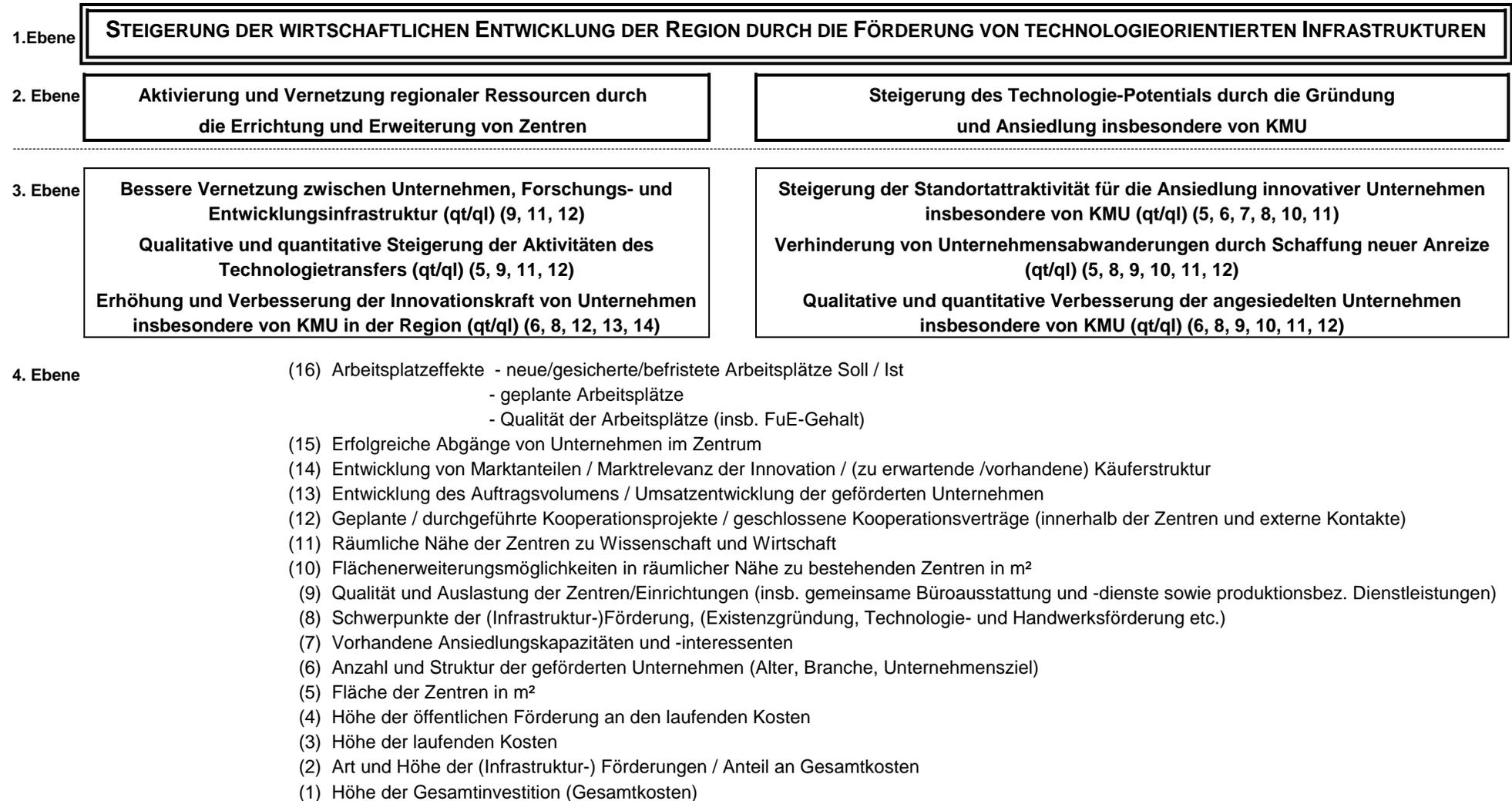
Inhaltlicher Schwerpunkt	Fragestellung
Kontext-Analyse Sozio-ökonomische Analyse Stärken/Schwächen-Profil	<ul style="list-style-type: none"> • Wie hat sich die sozio-ökonomische Situation während der Programm-Laufzeit verändert? • Konnte Stärken ausgebaut / Schwächen abgebaut werden?
Programm-Vollzug - Inputfaktoren - Finanzielle Indikatoren	<ul style="list-style-type: none"> • Wie stellt sich der Stand von Mittelbindung und Auszahlungen nach Programm / Prioritäten / Maßnahmen / Projekten dar? • Wie kann die (inner)regionale Verteilung der Umsetzung beurteilt werden • Welche Projektstruktur kann identifiziert werden: Groß- versus Kleinprojekte; Öffentliche versus private Projektträger etc. • Welche Gründe für abweichenden Programm-Vollzug können identifiziert werden?
Programm-Vollzug - Output-Faktoren	<ul style="list-style-type: none"> • Inwieweit wurden die festgelegten Ziele erreicht? • Wie ist die Effizienz im Sinne von Output/Input-Größen zu beurteilen? • Spezielle Beachtung der horizontalen Schwerpunkte der Union Umwelt, Beschäftigung, Chancengleichheit, KMU, Innovation
Wirkungsanalyse	<ul style="list-style-type: none"> • Welche Wirkungen auf die definierten Ziele lassen sich abschätzen? • Welche Aussagen können zu Brutto- bzw. Netto-Effekten gemacht werden? • Welche Faktoren hinsichtlich Erfolg/Mißerfolg von Maßnahmen können identifiziert werden?

Quelle: eigenen Zusammenstellung

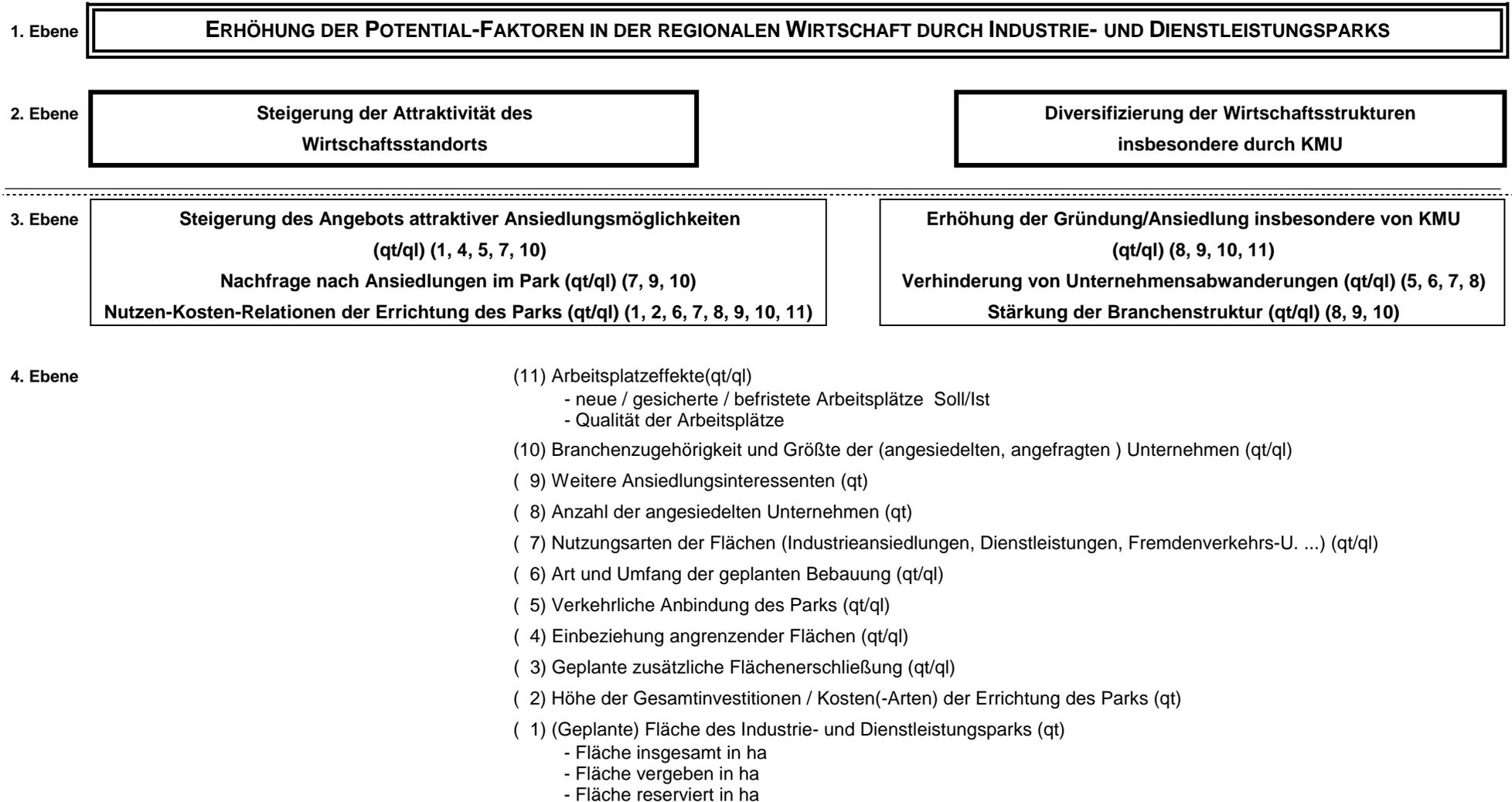
4.9.2 Verknüpfung von Zielhierarchien und Indikatoren – Indikatorentableaus für Ziel-2-Interventionen in Österreich⁴¹

⁴¹ Österreichischen Raumordnungskonferenz, Zwischenbewertung der Interventionen der Ziel-2-Programme, des RESIDER-II- und des RECHAR-II-Programmes in der Programmperiode 1995-99 in Österreich, Studie der JOANNEUM RESEARCH in Kooperation mit dem Bremer Ausschuß für Wirtschaftsforschung und dem European Policies Research Centre, Schriftenreihe Nr. 140, Wien 1998

Förderung von Impulszentren (Gründer-, Technologietransfer-, Innovations-, Sciencezentren, Technologie- und Forschungspark)



Industrie- und Dienstleistungsparks



Bildungsinfrastruktur: Aus- und Weiterbildung

1. Ebene

**BEITRAG ZUR DIVERSIFIZIERUNG UND MODERNISIERUNG DER REGIONALEN WIRTSCHAFTSSTRUKTUR DURCH FLANKIERENDE QUALIFIZIERUNG
HIER: INFRASTRUKTUREN ALS BASIS**

2. Ebene

**Beseitigung von Qualifizierungsengpässen zur Flankierung des
strukturellen Wandels**

**Promotion und Beschleunigung des strukturellen Wandels durch
perspektivische Qualifizierung auf zukünftige Bedarfe**

3. Ebene

Programm-Integration: Aktive Flankierung der investiven Schwerpunkte durch Qualifizierung (ql/qt) (1, 2, 3, 4, 6)
Größenanalyse: Stärkung insbesondere von KMU durch Aus- und Weiterbildung (ql/qt) (4, 5, 6)
Sektorenanalyse: Anteil der Engpaß- und der Wachstumssektoren als Zielbetriebe (ql/qt) (4, 5, 6)

4. Ebene

- (6) Sektoren- und Betriebsgrößenanalyse der potentiellen Nutzer
- (5) Zielgruppenanalyse: Berufsstand, Qualifizierungsniveau, Alter, Geschlecht, reg. Herkunft der Teilnehmer
- (4) Art der Aus-, Fort- und/oder Weiterbildung
- (3) Förderungsanteile/sonstige Finanzierung der Infrastruktur/ergänzend: Finanzierung der laufenden Kosten
- (2) Errichtungskosten der Infrastruktur (ergänzend: laufende Kosten)
- (1) Physische Beschreibung der Infrastruktur (qm, Anzahl der Räume etc, sonst. Ausstattung an Geräten, Maschinen etc...)

4.9.3 Review of Scottish Structural Fund Evaluation Studies – interim evaluations- June 2000

Report & Consultancy	Overall Approach	Targets & Benchmarks	Indicators & Data	Additionality & Dead-weight	Secondary Impacts	Efficiency	Effective-ness
<i>Highlands & Islands Objective 1 Programme 1994-99</i>	<p><i>Bottom-up</i> – sample survey of 101 projects; 19% interviews with project officers of sponsoring body; use of files for applications and claims data.</p> <p><i>Top-down</i> – overview of change in main aggregate indicators; comparison with Scotland, UK and EU on some indicators. No top-down assessment.</p>	No specific recommendations.	Need for revised programme monitoring system.	Survey of projects asked whether project would have gone ahead wholly or partially without programme funding. Some attempt to account for additionality and deadweight. Use of judgement. Allowance for double counting. Use NAO research on percent net job creation after allowance for non-additionality, displacement & multipliers.	Not examined directly.	Not considered.	Judgement on progress towards targets and objectives.
<i>Eastern Scotland Objective 2 Programme 1997-1999</i>	<p><i>Bottom-up</i> – review of other Obj 2 SPDs for 1997-99 and previous SPD, review of ERDF and ESF data from PME: 40 interviews with Programme partners, staff within PE and officials of DG XVI using structured questionnaire; 10 policy workshops to review particular themes. Telephone survey plus monitoring data (via claim forms).</p> <p><i>Top-down</i> – no socio-economic review since economic trends examined in review of 94-96 SPD.</p>	Targets and indicators changed since the beginning of the Programme limiting Programme assessment.	No specific recommendations.	Impacts to be assessed in ex-post evaluation.	Not considered.	Not considered.	Judgement on progress towards targets and objectives.

Quelle: EPRC/FAI 2000

Report & Consultancy	Overall Approach	Targets & Benchmarks	Indicators & Data	Additionality & Dead-weight	Secondary Impacts	Efficiency	Effective-ness
---------------------------------	-------------------------	---------------------------------	------------------------------	--	--------------------------	-------------------	-----------------------

<p><i>Western Scotland Objective 2 Programme 1997-99</i></p>	<p><i>Bottom-up</i> - review of other Obj 2 SPDs for 1997-99 and previous SPD; review of ERDF and ESF data from PME; 30 interviews with Programme partners using structured questionnaire; postal survey (census) of 130 partners.</p> <p><i>Top-down</i> - no socio-economic review since economic trends examined in review of the 94-96 SPD.</p>	<p>Targets based on 'sound rationale'. Earlier revision of some targets following independent review for DGXV1.</p>	<p>No specific recommendations.</p>	<p>Impacts to be assessed in <i>ex-post</i> evaluation.</p>	<p>Not considered.</p>	<p>Not considered.</p>	<p>Judgement on progress towards targets and objectives.</p>
<p><i>Evaluation of Borders Objective 5b Programme 1994-99</i></p>	<p><i>Top-down</i> – overview of aggregate indicators before and after start of SPD. Comparisons with Scotland. Use of shift-share analysis of employment change.</p> <p><i>Bottom-up</i> – sample of 30 projects for case study, interviews, grossing up.</p>	<p>No specific recommendations.</p>	<p>No specific recommendations.</p>	<p>Judgement.</p>	<p>Judgement.</p>	<p>Judgement.</p>	<p>Judgement.</p>
<p><i>Dumfries & Galloway Objective 5b Programme 1994-99</i></p>	<p><i>Bottom-up</i> - 47 project case studies 25% of total; analysis of application and claims forms; interviews with project officer of sponsoring body.</p> <p><i>Top-down</i> – socio-economic overview only.</p>	<p>No specific Recommendations.</p>	<p>Confusion between activities and impacts. New target output measures proposed and appropriate indicators.</p>	<p>Judgement from case studies.</p>	<p>Not examined directly.</p>	<p>Not considered.</p>	<p>Judgement on progress towards targets and objectives.</p>

Quelle: EPRC/FAI 2000

Report & Consultancy	Overall Approach	Targets & Benchmarks	Indicators & Data	Additionality & Dead-weight	Secondary Impacts	Efficiency	Effective-ness
<i>Evaluation of North and West Grampian Objective 5b Programme 1994-99</i>	<p><i>Top-down</i> – overview of limited indicators before and after start of SPD. Comparisons with Scotland and GB. Too early to gauge SPD impact - used as check on current relevance of SPD priorities & measures.</p> <p><i>Bottom-up</i> - project files for application and claim form data for ERDF. ESF projects using SO info. 25 interviews with partners and PME. Use of semi-structured questionnaire. Project delays.</p>	A number of targets and their numerical value less than ideal'. Reported project outputs and impacts often do not relate to SPD targets. Need to review and amend initial targets.	<p>Programme area embraces only part of specific local authority or TTWA areas, so data for the Programme area is limited.</p> <p>Not all indicators were found to be reasonable. Too many indicators. Need to identify impact indicators for each Measure.</p>	Focus on gross (physical) outputs and intermediate impacts (direct). Based on case studies, review of project files and judgement. Judgement offered on extent of deadweight in general.	Not specific to projects. General judgement offered.	Not considered.	Only suitability of targets and measures considered.
<i>Evaluation of Rural Stirling & Upland Tayside Objective 5b Programme 1994-99</i>	<p><i>Top-down</i> – overview of aggregate indicators before and after start of SPD. Comparisons with Scotland and GB. Too early to gauge SPD impact – used as check on current relevance of SPD priorities & measures.</p> <p><i>Bottom-up</i> - project files for application and claim form data for ERDF. ESF projects using Scot Office info. 25 interviews with partners and PME. Use of semi-structured questionnaire. Project delays.</p>	Several targets 'less than ideal' - 'only small proportion of project activity captured within SPD targets. No impact targets provided for a number of measures.	<p>Output data from project applicants is of 'variable quality'</p> <p>Lack of clarity in definition of some indicators and problem of double counting were key issues to be resolved in monitoring arrangements.</p>	Clarification of gross outputs ie. activities associated – and estimate of direct impact based on judgement from "experience of other projects and programme evaluation".	Not considered.	Not considered.	Only suitability of targets and measures considered.

Quelle: EPRC/FAI 2000

Report & Consultancy	Overall Approach	Targets & Benchmarks	Indicators & Data	Additionality & Dead-weight	Secondary Impacts	Efficiency	Effective-ness
<i>Evaluation of the Eastern Scotland CSFs 1989-93</i>	<p><i>Bottom-up</i> – developed programmed-base from SO's financial monitoring data plus data on output & impacts from each of 1032 ERDF supported projects. Plus analysis of ESF projects, 65 case study subsets plus 32 interviews with P/ship using semi-structured questionnaire.</p> <p><i>Top-down</i>: use of non-EU funded comparator areas in Nottingham & Derby plus GB average using S/S analysis. Scotland also used as comparator for some indicators. No direct link between 'b-u' & 't-d' approaches but compares CSF jobs outcome overall of 17,000 jobs or 4.5% of employment.</p>	No comments.	No comments.	Application of assumed factors for each. Variation according to priority. Grossing up from case studies.	Assumed displacement factor and standard multiplier applied variation according to priority. Grossing up from case studies.	Cost per Job.	Judge-ment.
<i>Evaluation of the Galloway CSF 1989-93 Objective 5b</i>	<p><i>Bottom-up</i> – representative sample of CSF projects: project records, interviews with project managers and sponsors. Grossing up.</p> <p><i>Top-down</i> – aggregate indicators, pre-CSF and trends for policy-off and some comparator areas plus national change. Use of S/S analysis of employment. Judgemental linking of b-u and t-d approaches.</p>	Some targets. Little meaning in development terms.	Lack of systematic monitoring for CSF as whole.	Judgement.	Drawn from secondary sources & judgement.	Cost per Job.	Judge-ment. Spend did not relate well to objectives.

Quelle: EPRC/FAI 2000

Report & Consultancy	Overall Approach	Targets & Benchmarks	Indicators & Data	Additionality & Dead-weight	Secondary Impacts	Efficiency	Effectiveness
<i>Eastern Scotland Objective 2 Programme 1994-96</i>	<p><i>Bottom-up</i> - review of project financial and monitoring data from PME; review of evaluations of projects funded by the 1994-96 Programme; telephone survey using structured questionnaire of sample beneficiaries.</p> <p><i>Top-down</i> - socio-economic review; some assessment; no attempt to reconcile 'b-u' with 't-d'.</p>	Lack of a clear hierarchy of targets at Programme, Priority and Measure levels.	Inconsistency between indicators and targets Poor quantification of some targets. Need for more data to assess progress towards meeting targets.	Estimates of additionality derived from 'benchmark data' for Obj 2 programmes produced by consultant; from specific evaluation data; and from review of earlier evaluations of Programme projects.	Application of 'off the shelf' displacement and multiplier factors.	Cost per job created or safeguarded.	Judgement on progress towards targets and objectives but limited by narrow range of indicators.
<i>Western Scotland Objective 2 Programme 1994-96</i>	<p><i>Bottom-up</i> - review of other Obj 2 SPDs on financial allocation and Priority & Measure structures; review of evaluations of projects funded by the 1994-96 Programme; project application data; no further survey work for ERDF; for ESF projects a survey of beneficiaries.</p> <p><i>Top-down</i> - socio-economic review; some assessment; no attempt to reconcile 'b-u' with 't-d'.</p>	Progress in developing targets from activities to output and impacts. Need for appropriate Targets for 'softer' activity such as networking between companies. Over-ambitious targets at project level.	Monitoring framework not sufficiently comprehensive to identify routes to impact of measures. Poor quantification of some targets.	Estimates of additionality derived from 'benchmark data' for Obj 2 programmes produced by consultant; from specific evaluation data; and from review of earlier evaluations of Programme projects.	Application of 'off the shelf' displacement and multiplier factors.	Cost per job created or safeguarded.	Judgement on progress towards targets and objectives.

Quelle: EPRC/FAI 2000

4.9.4 Case studies: EPRC/FAI-Study

CASE STUDY: USE OF CONTROL GROUPS TO ASSESS THE IMPACT OF ESF FUNDING - UNITED KINGDOM

The Pieda interim evaluation of the UK ESF Objective 3 programme, 1994-96, is a good example of the measurement of the net impact of ESF projects, where the recipients are often not firms but individuals. For part of the evaluation of net impacts, Pieda made use of a control group, where participants in one of the programme-supported schemes were compared with a control group of non-assisted individuals with 'similar' characteristics. The control group was constituted using the JUVOS database, which provides a continuous record for a group of claimant unemployed at particular junctures.

Drawing two cohorts of people who had been unemployed for one year at two different points in time for each cohort, the control group could be tracked in terms of their employment history and compared with the performance of those receiving training through the ESF scheme. The use of two separate time periods for the control group cohorts allowed for the effects of varying national unemployment circumstances to be taken into account. Two comparison dates were used to compare the control groups with those benefiting from the training: six months and twelve months after completion of the training programme.

CASE STUDY: USE OF INTERVIEWS AND EVALUATORS JUDGEMENT TO ESTIMATE ADDITIONALITY AND DISPLACEMENT - GIBRALTAR

In the final evaluation of the 1994-96 Objective 2 programme and the interim assessment of the 1997-99 programme for Gibraltar, face-to-face and telephone interviews were conducted by ECOTEC using a sample of projects to measure additionality among other impacts. The researchers' questionnaires had a set of four boxes for different degrees of additionality: 'exhibiting low additionality <25 percent', '25-49 percent', '50-74 percent' and 'very high additionality 75 percent+'. The researchers were requested to probe interviewees about the impacts of their projects and the extent to which they could be attributed to SF assistance: the researchers were then to use their own judgements to decide which category best described the project. Similar judgements were employed in determining displacement effects but not in the actual interviews.

CASE STUDY: DYNAMIC INPUT-OUTPUT ANALYSIS - BEUTEL

Jorg Beutel developed an evaluation system for DG XVI. This system uses dynamic Input-Output (I-O) analysis to evaluate the impact of Structural Funds on the economies of Greece, Spain, Ireland and Portugal. There are two elements to this system. First, Beutel develops a method for updating Input-Output tables. Second, he has used the relevant national I-O tables to simulate the impacts of Structural Fund expenditure in both *ex ante* and *ex post* evaluation exercises.

Updating I-O tables

One problem in using I-O analysis is that the construction of I-O tables is costly and time consuming. This means that I-O tables are typically available only for particular years and are published only with a time lag. Therefore, one can be working with a very dated table. A mechanism for using available information to update existing tables is clearly very valuable and improves the credibility of the analysis. However, in the Scottish context this is less important in that, at the moment, the Scottish Executive is producing annual I-O tables and these are relatively up to date.

Simulating the impact of the Structural Funds

Beutel uses I-O analysis to quantify the impact of the Structural Funds on the following economic variables:

- the rate of growth;
- other economic aggregates and the industrial structure;
- employment; and
- capital stock.

The I-O system is used to identify three separate effects. These are:

- the demand impacts generated by the construction of the infrastructure financed by the Structural Funds;
- the wages and salaries supported by the Structural Fund; and
- the additional induced investment generated as a result of the expansion in local activity.

Infrastructure

The expansion in infrastructure is the most straightforward to deal with. Here, this expenditure is simply treated as an expansion in final investment demand, and its impact calculated in the normal way via the Leontief inverse.

Wages and salaries supported by the Social Fund

The impact of the Social Fund expenditure seems to be treated as allowing an expansion in resources to particular resource-constrained sectors. This will lead to subsequent expansion in indirect and induced outputs and incomes. Whether this treatment is strictly compatible with the way in which the I-O model is used to analyse the other elements of Structural Fund expenditure is questionable.

Induced investment

In this case, the model is dynamised through linking net investment to changes in output via an accelerator mechanism. This means that sectoral capital stocks adjust, through investment, to changes in sectoral output, giving an added boost to the economy in terms of further increase in final demand. The predicted change in local activity now evolves over time as the economy more gradually adjusts to the injection delivered by the Structural Funds.

Once the changes in gross output have been determined, the subsequent impact on total employment and capital stock can be calculated through the use of the capital and employment matrices. Employment is a key policy variable and changes in capital stock might be taken to generate supply side improvements in the local economy. Structural changes in the economy can also be tracked, as can variables such as the balance of payments.

Strengths and weaknesses of the dynamic I-O approach

Strengths

- It is based upon a formal modelling approach with a clear analytical framework.
- It stresses sectoral disaggregation.
- It incorporates the interaction between economic sectors and elements of final demand.
- It uses existing I-O tables and is relatively low cost.

Within a Scottish context, an I-O based technique has the advantage that the economy is relatively well-served in terms of available I-O tables and certainly the existing information embodied in those tables could be more effectively utilised for evaluation purposes.

Weaknesses

- I-O is not a supply-side model. It is a demand driven model that has a passive (permissive) supply side. This has two key implications.
- First, the supply-side stimuli that the Structural Funds aim to produce have to be converted to a demand side stimuli to be captured by I-O. This has to be done in an essentially *ad hoc* way.
- Second, there are no supply constraints built into the model as it normally operates.
- I-O tables are not available for existing programme areas so that evaluation at that level is more problematic.

Whilst there are strengths to this type of modelling it is an unsatisfactory way, if used in isolation, to model the impact of Structural Fund expenditure. It is unable to capture the effect of improved competitiveness and efficiency that are central aims of Structural Fund intervention. However it is the case that at present the bottom-up Scottish (and UK) approach is really a hybrid. 'Industrial survey' methods are used to derive the direct outputs, and an export-base multiplier is employed

to calculate the full impact. In this respect, at the very least, Scottish evaluation could make more use of the national and sub-national I-O tables that now exist in Scotland. This would imply a sectoral disaggregation of outputs which at present generally is not produced in evaluation studies. This in itself would be a beneficial change.

CASE STUDY: ECONOMETRIC TECHNIQUE - THE PARADISE MODEL

Eddy Blaas and Peter Nijkamp have developed an econometric technique for evaluating the impact of the European Regional Development Fund. The technique has the acronym PARADISE, Policy Assessment of Regional Achievements and Developments Induced by Stimuli of ERDF. The aim is to generate an evaluation method that can be applied in a consistent manner across regions in all the EU countries. This would give the Commission an indication of the effectiveness of ERDF expenditure as a whole, and also reveal the countries and types of region in which this expenditure had been particularly effective.

Over the time period covered by Blaas and Nijkamp, infrastructure takes the major share of ERDF expenditure, and they take the major economic variable affected by this expenditure to be private investment. They are, therefore, attempting to identify and quantify statistically the link between the levels of ERDF expenditure and private investment in the same region. They are mindful that employment is the variable that has the strongest policy relevance and, therefore, also attempt to measure the relationship between private investment and employment.

The method involves a two-step approach. The first step is an ‘exploratory frequency analysis’. Imagine that the relevant impact variable is A and the policy (or control) variable is B and that we have observations on these variables across all regions. These variables are standardised for regional size, so that they are expressed, for example, per head of population. The average values of A and B across the whole population under consideration are A^* and B^* . Each region is then classified as having an above or below average value for A and B. This places the region in one of the four categories identified in the table below. If there is expected to be a positive relationship between the policy and impact variable, we expect the bulk of the observations across regions to lie in the quadrants I and IV.

Cross-classification frequency table:

	$B > B^*$	$B < B^*$
$A > A^*$	I	II
$A < A^*$	III	IV

The frequency analysis is taken to be a flexible tool that is used in an exploratory way to identify possible relationships. Where a time series of cross-sectional (regional) data is available this analysis can be done year by year, with and without lags or for moving averages. Given the usual time lags that are thought to occur

between improved infrastructure and increased private activity plus the lumpy nature of these investments, these options are valuable. Also a number of variables can be tried out. Further, if the number of regions permits, statistical testing (for example χ^2 analysis) can be undertaken at this level of analysis.

The model that Blaas and Nijkamp test to determine private investment, I_p , has additional variables alongside the ERDF expenditure, I_c . These are the expenditure on other public investment, I_{o-c} , and the change in gross value added, ΔGVA . The model therefore takes the form:

$$I_p = I_p(I_c, I_{o-c}, \Delta GVA)$$

The first step in the PARADISE procedure is to undertake extensive forms of this pairwise frequency analysis between the dependent and independent variables, taken one at a time. These are interpreted as revealing a positive relationship between the level of private investment in a region and all the independent variables.

The second stage of the PARADISE method is to perform a full-blown multivariate econometric analysis. Within this framework a lot of experimentation was done regarding appropriate time lags, etc. For an experiment with Dutch data, using moving averages the coefficient on the ERDF expenditure term was positive and significant, though in many of the other regressions it was not.

Strengths and weaknesses of the PARADISE model

Strengths

- It is econometric analysis, which is, in practice, the most rigorous form of testing with economic data.
- It develops an evaluation method that can potentially be used across all the EU regions with commonly-available data.

The statistical testing of the relationship between an independent policy variable and a dependent impact variable is in principle a powerful tool and this is a general methodological approach most strongly favoured in economics. The advantages that the regression results can be compared across European regions is perhaps less relevant in domestic evaluations.

Weaknesses

- It is only an *ex post* technique. Further, in so far as the technique depends upon time series data, results are generated some time after programmes are completed.
- The evaluation method does not distinguish explicitly between individual elements of the programme. Potentially if ERDF expenditure is differentially effective between regions and countries, inferences might be drawn linking this with the differential nature of the programmes in the different spatial areas, but this is very indirect.
- This approach does not focus on individual partnerships.

- The rigour of the existing econometric analysis is not clear. For example, no diagnostic statistics are reported. Here the model specification is restricted by the data availability for the least data-rich country.

This approach is appropriate for more long-run strategic decision taking at the EU level. It does not fulfil any of the short and medium run monitoring functions of present evaluations. It is, therefore, properly seen as complementary to many of the alternative bottom-up and top-down approaches.

Relevance for Scotland

The idea of attempting to identify the aggregate effects of Structural Fund policy through direct econometric methods is appealing though a more sophisticated model that this should be attempted.

CASE STUDY: THE HERMIN MODEL

Hermin is a development of the Hermes model. The Hermes model was designed by the Commission in the late 1970s primarily to identify and quantify the impacts of the oil price shocks on the individual economies of the EC. The Hermin models were subsequently built for the peripheral economies of Greece, Ireland, Portugal and Spain. Their main focus was to capture the impacts of the CSF and specifically the effects of policy aimed at enhancing infrastructure, human capital formation and other elements of industrial policy. Here we concentrate on the Irish Hermin model: the corresponding models for the three other peripheral economies are similar but differ in some details.

In the Hermin model there are four sectors:

- the traded sector (manufacturing);
- the non-traded sector (market services, utilities and construction);
- the public sector (public administration, education and health); and
- the agricultural sector (agriculture, forestry and fishing).

The traded sector

In the Irish Hermin, the output of the manufacturing (= traded) sector is driven by external demand. Specifically, the Irish traded sector is a pure price taker on world markets. Output is therefore dependent on the exogenous world price and the endogenous cost competitiveness.

The non-traded sector

The output of the non-traded sector (= sheltered or market-services sector) is determined by final demand. In this case, there is a wider range of sources of final demand, and the price is set as a mark-up on unit costs.

The public sector

Output in the public sector is treated as exogenous (ie. driven by government policy). However, the treatment of borrowing and debt accumulation is endogenous with the option of enforcing a PSBR borrowing constraint.

The agricultural sector

Here again output is primarily policy-determined through the operation of the CAP.

In all sectors, a KLEM (capital, labour, energy and materials) constant-returns production function is used. Factor demands are a function of sectoral output and relative factor prices. Also, in each sector capital is fixed in each period but is updated between periods through net investment.

In the Hermin model, the traded sector plays a crucial role in determining the performance of the whole economy. The small country assumptions are imposed so that exchange rates and interest rates are fixed exogenously. In the traded sector, the wage rate is determined by a bargaining procedure. The wage is therefore a function of commodity prices, direct and indirect tax wedges, the unemployment rate and productivity in the traded sector.

This wage is then transferred to the non-traded sectors (that is to say, the non-traded sectors experience the same percentage changes in the wage rate as does the traded sector).

Over the longer run, the unemployment rate, and thereby the manufacturing wage rate, will be affected by migration, which is assumed to take a Harris-Todaro form. This simply means that net migration flows are driven by variations in Irish – UK real wage and unemployment rate differentials.

Externalities

The Structural Funds focus on three types of expenditure which are thought to have externalities which will improve the overall performance of the recipient economy and encourage regional convergence. These are:

- factor productivity externalities;
- industrial composition externalities; and
- labour market externalities.

Factor productivity externalities

These are externalities experienced by the private sector which take the form of increased productivity. They are linked to two types of programme:

- public infrastructure; and
- education and training.

Evidence suggests that efficiency in the private sector is positively related to the stock of public infrastructure. Similarly, private sector productivity is positively related to the stock of human capital. Therefore in the Hermin model, the efficiency parameters on the private-sector production functions are taken to be increasing in the stock of public infrastructure and the number of trainees. Any increase in the efficiency of production has a positive impact on the level of economic

activity through an increase in competitiveness and a higher real wage (encouraging immigration). However, an increase in labour productivity can reduce the demand for labour if the stimulus to activity is insufficiently low.

Industrial composition externalities

This stands for the stimulus to the level of foreign direct investment that is thought to be linked to expenditure on the Structural Funds. In the Hermin model, the proportion of the world's output of traded goods (here manufactures) produced in Ireland depends upon Irish cost-competitiveness. However, it is also argued that the attractiveness of Ireland for foreign-owned plants will additionally depend on the level of infrastructure capital and the skill level of the local labour force. Therefore, domestic output in the traded sector is again made dependent upon the stock of public expenditure and the number of trainees.

Labour market externalities

Programmes under the European Social Fund focus on the long-term unemployed and school leavers. These groups have only a small impact on the operation of the labour market and unemployment amongst these sections of the labour market has little impact on wage determination. If expenditure by the Structural Funds make some of these workers more active participants in the labour market, this will have an effect on the wage bargaining mechanism, making the wage more sensitive to variations in the unemployment rate. Again this is incorporated in the Irish Hermin model.

Strengths and weaknesses of the Hermin approach

Strengths

- It generates an explicit model of the recipient economy. This model is open to scrutiny and based on prior theoretical reasoning so that the causal mechanisms are transparent.
- It captures economy-wide impacts.
- It incorporates supply-side effects, such as competitiveness and labour market effects.
- Once the model has been set up, additional simulation is inexpensive, particularly as against extensive face-to-face interviews.

These are very significant strengths. Any evaluation process requires some theoretical framework. Often these are both implicit and simplistic. The formal approach of the Hermin model makes the underlying model explicit (and therefore subject to challenge and debate) and also tailors the model to address the key evaluation issues raised by the Structural Funds. The model will allow sensitivity analysis to reveal the extent to which results change with a change in the values of key parameters.

Weaknesses

- Formal economic modelling involves a high set-up cost.
- The data requirements are high.

It would seem unlikely that one would set up a formal economic model of the Hermin type simply to carry out evaluations of Structural Funds. Hermin is a national model and therefore can be used in many other circumstances so that the fixed costs involved in setting up the model is spread over a wider range of work. Also if such a model is to be primarily econometrically parameterised, there probably is insufficient data at the regional level and certainly there would be difficulty applying this approach to individual partnerships. However, a similar model does now exist for Scotland. This is the Fraser of Allander Institute's AMOS Computable General Equilibrium model, a version of which has been used to evaluate Scottish Enterprise policy at the Scottish and UK level.

CASE STUDY: THE CSF EX ANTE EVALUATION FOR THE MEZZOGIORNO REGION OF ITALY

The Objective 1 CSF aims to encourage economic growth through its impact on the productivity, and therefore volume, of investment. The intention of the programme is to generate 'breaks' (rotture) in the supply- and demand-side behaviour of key economic agents. A traditional supply-side model, in which the productivity of private capital is expressed as a function of the stock of public capital, has been deemed inappropriate. This is because the parameters of this function would reflect what has been characterised as inefficient past public intervention in the Mezzogiorno. Indeed, the future improved quality of public investments has been interpreted as a direct additional effect on the productivity of the system. The macro-economic model therefore has been built around a total factor productivity function, which has as arguments the 'break variables' that the programme targets. Four variables are endogenous to the model: export, investments, net imports and the activity rate. The other variables are exogenous and simulate the positive externalities that the programme will have on the economy of the Mezzogiorno. Through these variables, the positive externalities on the economy of the Mezzogiorno are simulated and measured under alternative hypothesis.

To some extent the model is similar to the Hermin model, but is differentiated along the following lines:

- It does not distinguish between different types of public investments
- The explicit generation of prices is not considered
- It has no sectoral disaggregation
- It estimates the elasticity of some of the determinants of the function of the total productivity of factors.
- It is an explicitly regional model and deals separately with trade to and from the rest of Italy and the rest of the World.

The approach adopted by the Italian government has allowed it to verify the coherence of the objectives of the programme in relation to three main elements. These are:

- the compatibility with the overall macro-economic framework,
- the existence of, and transmission mechanisms through which, the hypothesised externalities act and,
- lastly, the size and dynamics of these externality effects on the productivity and, consequently on the growth, of the economic and social system of the Mezzogiorno.

Strengths and weaknesses

These are best measured against the Hermin model.

Strengths

- This is a regional model, which takes into account regional interaction
- Parameterised on regional data

Weaknesses

- Less sectoral disaggregation
- Less-well articulated supply side
- The model is particularly orientated toward the evaluation of infrastructure expenditure

4.9.5 Definition von Indikatoren

Indikatoren entsprechend der Vergleichbarkeit von Informationen unterscheidet man:

- **Generic Indicators:** Ein „generic Indicator“ ermöglicht die Beurteilung unterschiedlicher Interventionen innerhalb eines Programms. Er ermöglicht einen internen Vergleich sowie eine Aggregation auf Programme-Ebene. Zum Teil können „generic indicators“ auch nur Teile von Programme erfassen. Beispiele sind: Anteil der KMUs an den Förderungsempfängern, Kosten je neuen/erhaltenen Arbeitsplatz etc.
- **Key Indicators:** Indikatoren, die sowohl einen Vergleich innerhalb des Programms als auch zwischen Programmen ermöglicht. Entscheidend ist eine einheitliche Definition. Ist dies gewährleistet, können „key indicators“ auch auf EU-Ebene aggregiert werden. Beispiele sind: Gebundene und ausbezahlte Mittel je Maßnahmentyp, Kosten je geschaffenen/erhaltenen Arbeitsplatz.

Entsprechend des Umfangs an Informationen werden unterschieden:

- **Programme Indicators:** Über „programme Indicators“ werden ausschließlich die durch das Programm erzielten direkten und indirekten Effekte erfaßt. Bspw. Zahl jener Unternehmen, die durch die Unterstützung des Programms zusätzliche Exporte ermöglicht wurden.
- **Context Indicators:** „Context indicators“ erfassen sozio-ökonomische Aspekte des gesamten Förderungsgebietes bspw. Zahl der exportierenden Unternehmen, Zahl und Entwicklung der Zahl der Arbeitslosen etc.

Hinsichtlich der Programmumsetzung werden unterschieden:

- **Resource Indicators:** „Resource Indicators“ stellen Informationen über die in der Programm-Umsetzung eingesetzten Mittel bereit. Dazu zählen insbesondere eingesetzte Finanzmittel, aber auch Humankapital, etc.
- **Output Indicators:** Die im Abtausch zu den eingesetzten Finanzmitteln erzielten Programm-Aktivitäten, bspw. Zahl der Beratungs-Stunden, Zahl der geschaffenen m2 an Technologieparks etc.
- **Result Indicators:** „Result indicators“ weisen eine unmittelbare Änderung bei den Förderungsempfängern aus bspw. erreichte Qualifikationen, Einsparungen bei Reisezeiten, Zufriedenheit mit den beanspruchten Beratungsleistungen etc.
- **Impact Indicators:** „Impact indicators“ messen über die unmittelbar aufgrund der Programm-Umsetzung beobachtbaren Änderungen hinausgehenden, längerfristig anhaltenden Wirkungen bspw. Beschäftigungsrate nach 12 Monaten, Überlebensraten bei Unternehmensgründungen etc. Impact indicators können erst mit einem bestimmten Zeitabstand beobachtet werden und werden in der Regel nicht im Monitoring, sondern im Rahmen von Evaluierungen erhoben und beurteilt.

Hinsichtlich der Evaluierungskriterien werden unterschieden:

- **Relevance Indicators:** „Relevance Indicators“ versuchen die Programm-Ziele mit der Bedarfssituation einer Region zu verknüpfen. Als Beispiele können angeführt werden: Die Zahl der geplanten Beratungsprojekte für Exporte im Vergleich zu der Zahl der nicht-exportierenden Unternehmen oder die Zahl der angebotenen Schulungseinheiten im Vergleich zu der Zahl der Langzeitarbeitslosen.
- **Efficiency Indicators:** „Efficiency Indicators“ setzen die Outputs, Ergebnisse bzw. Wirkungen in Bezug zu den eingesetzten Mitteln. Mit welchem Mitteleinsatz wurden die jeweiligen Ergebnisse erzielt? Beispielsweise Arbeitsplätze je eingesetzter Mittel.
- **Effectiveness Indicators:** „Effectiveness indicators“ setzen die erreichten Outputs, Ergebnisse und Wirkungen in Bezug mit den im Programm angegebenen erwarteten Werten.

Entsprechend der Quantifizierung und Nutzung der Information werden unterschieden:

- **Monitoring Indicators:** „Monitoring indicators“ werden durch das Programm-Management im Rahmen des Monitorings im Kontakt mit den Förderungsnehmern gesammelt. Diese sollte zumindest die Bereiche Input, Output, Results umfassen, während Impacts im allgemeinen über Evaluationen erfaßt werden.
- **Evaluation Indicators:** „Evaluation Indicators“ werden nicht über das Monitoring, sondern in erster Linie über die Evaluation erfaßt. Es handelt sich dabei im allgemeinen um Impact Indikatoren.

Quelle: MEANS Collection Volume 2

4.9.6 Indikatoren für die Begleitung und Bewertung: Kommissions-Vorschläge⁴²

BEISPIELE FÜR INDIKATOREN FÜR DIE BEGLEITUNG UND BEWERTUNG

1. WIRTSCHAFTSBEREICH

Anm.: Die Zahlen für die geschaffene oder erhaltene Brutto-/Nettobeschäftigung sollten nach Möglichkeit nach Männern/Frauen aufgeschlüsselt werden.

Interventionsbereiche	Output	Ergebnis	Wirkung
13 Förderung der Anpassung und Entwicklung ländlicher Gebiete⁴³			
1306 Erneuerung und Entwicklung von Dörfern und ländlichen Gebieten und Erhalt des ländlichen Kulturguts	<ul style="list-style-type: none"> ● m² erneuerter Dorfplätze/-straßen ● Zahl erneuerter Häuser ● Zahl der geförderten Projekte 	<ul style="list-style-type: none"> ● Zahl der Einwohner, die in der Nähe (unter 1 km) der erneuerten Gebiete leben ● Zahl der Unternehmen / Geschäfte im Fördergebiet 	<ul style="list-style-type: none"> ● Zahl der geschaffenen/erhaltenen Brutto-/Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze) ● % der Einwohner, die in den nächsten fünf Jahren bleiben wollen
1309 Entwicklung und Verbesserung der Infrastrukturen in Verbindung mit der Entwicklung der Landwirtschaft	<ul style="list-style-type: none"> ● Zahl der geförderten Projekte 		
15 Beihilfen für Großunternehmen			
151 Materielle Investitionen (Einrichtungen und Ausstattungen, Beteiligung an staatlichen Beihilfen)	<ul style="list-style-type: none"> ● Zahl der geförderten Großunternehmen 	<ul style="list-style-type: none"> ● Direkte Privatinvestitionen der Unternehmen, die finanzielle Unterstützung erhalten (in Meuro und als % der Gesamtinvestition) ● Gebaute/sanierte Bodenfläche (m²) 	<ul style="list-style-type: none"> ● Zahl der geschaffenen/erhaltenen Brutto-/ Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze) ● Anstieg des Umsatzes der Unternehmen, die finanzielle Unterstützung erhalten haben, nach zwei Jahren (%)
152 Umwelttechnologien, saubere und wirtschaftliche Energietechnologien	<ul style="list-style-type: none"> ● Zahl der geförderten Umweltaudits ● Zahl der Unternehmen, die Umweltberatung von Fachleuten erhalten 	<ul style="list-style-type: none"> ● Zahl begünstigter Unternehmen, die im Umweltsektor neu gegründet wurden ● % Unternehmen, die für eine Zertifizierung nach 	<ul style="list-style-type: none"> ● Anstieg des Umsatzes der im Umweltsektor geförderten Unternehmen nach zwei Jahren (in %) ● Zahl der eingeführten neuen Produkte/Verfahren

⁴² European Commission (EC) (2000), The New Programming period 2000 – 2006: methodological working papers, Working paper 3, Indicators for Monitoring and Evaluation, An indicative methodology

⁴³ Die Liste betrifft nur die von EFRE geförderten Interventionen.

	(mindestens 5 Tage)	bestimmten	● Verringerung der
	● Zahl der Unternehmen, die finanzielle Beihilfen zur Einführung von Umwelttechnologien bzw. zur Entwicklung von Ökoprodukten erhalten	Umweltnormen in Frage kommen	Schadstoffemissionen (CO ₂ , NO _x usw. in %)
Interventions bereiche	Output	Ergebnis	Wirkung
153 Wirtschaftliche Beratungsdienste (einschließlich Internationalisierung, Export, Umweltmanagement, Technologieerwerb)	<ul style="list-style-type: none"> ● Zahl der geförderten Unternehmen ● Anzahl der bereitgestellten Beratungsdienste 	<ul style="list-style-type: none"> ● Zahl der Unternehmen, die ins Exportgeschäft einsteigen ● Zahl der Unternehmen, die in neue Märkte exportieren ● % der Unternehmen, die mit den bereitgestellten Diensten zufrieden sind 	<ul style="list-style-type: none"> ● Exportgeschäft der geförderten Unternehmen nach 18 Monaten in % ● Zunahme des geschaffenen Mehrwerts nach 18 Monaten ● Zahl der geschaffenen/erhaltenen Brutto-/ Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze)
155 Finanzengineering	<ul style="list-style-type: none"> ● Zahl/Höhe der Garantiefonds, die finanziell gefördert werden ● Zahl der finanziell geförderten Leasing-Pläne 	<ul style="list-style-type: none"> ● Zahl der Unternehmen, die mit der Förderung zufrieden sind 	<ul style="list-style-type: none"> ● Zahl der geschaffenen/erhaltenen Brutto-/ Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze)
16 Beihilfen für KMU und Handwerksbetriebe			
161 Materielle Investitionen (Einrichtungen und Ausstattungen, finanzielle Beteiligung an staatlichen Beihilfen)	<ul style="list-style-type: none"> ● Zahl der finanziell geförderten KMU (mit Männern/Frauen Inhabern) ● Zahl neuer KMU, die finanziell gefördert werden (davon von Frauen gegründet) 	<ul style="list-style-type: none"> ● Neue/gestiegene Umsätze in KMU (Euro) ● Direkte Privatinvestitionen der Unternehmen, die finanzielle Unterstützung erhalten (in Mio. Euro und % der Gesamtinvestitionen) ● Zahl der Frauen als Projektleiter in Privatprojekte (% aller Projekte) 	<ul style="list-style-type: none"> ● Überlebensrate von finanziell geförderten Neugründungen nach 18 Monaten (in %) ● Zahl der geschaffenen/erhaltenen Brutto-/ Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze) ● Umsatzsteigerung von finanziell geförderten Unternehmen nach zwei Jahren
Interventions bereiche	Output	Ergebnis	Wirkung
162 Umwelttechnologien, saubere und wirtschaftliche Energietechnologien	<ul style="list-style-type: none"> ● Zahl der Unternehmen, die finanzielle Beihilfen 	<ul style="list-style-type: none"> ● Zahl begünstigter Unternehmen, die im Umweltsektor neu 	<ul style="list-style-type: none"> ● Umsatz der im Umweltsektor geförderten Unternehmen nach zwei Jahren (%)

	<p>zur Einführung von Umwelttechnologien bzw. zur Entwicklung von Ökoprodukten erhalten</p> <ul style="list-style-type: none"> ● Zahl der geförderten Umweltaudits ● Zahl der KMU, die in grenzüberschreitenden Projekten tätig sind ● Zahl der KMU, die Umweltberatung erhalten (mindestens 5 Tage) 	<p>gegründet wurden</p> <ul style="list-style-type: none"> ● % der Unternehmen, die für eine Zertifizierung nach bestimmten Umweltnormen in Frage kommen ● Verringerung der Energiekosten in KMU (%) ● Direkte Privatinvestitionen der Unternehmen, die finanzielle Unterstützung erhalten (in Mio. Euro und % der Gesamtinvestitionen) 	<ul style="list-style-type: none"> ● Zahl der eingeführten neuen Produkte/Verfahren ● Anstieg des Verkaufs umweltfreundlicher Produkte (%) ● Zahl der geschaffenen/erhaltenen Brutto-/Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze) ● Verringerung der Schadstoffemissionen (CO₂, NO_x usw. in %)
163 Wirtschaftliche Beratungsdienste Investitionen in Humanressourcen (Information, Unternehmensplanung, Beratungsdienste, Marketing, Management, Design, Internationalisierung, Export, Umweltmanagement, Technologieerwerb)	<ul style="list-style-type: none"> ● Zahl der KMU (mit männlichen / weiblichen Inhabern), die Umweltberatung erhalten (mindestens 5 Tage) 	<ul style="list-style-type: none"> ● Zahl der KMU, die ins Exportgeschäft einsteigen ● Zahl der KMU, die in neue Märkte exportieren ● % der KMU, die mit den bereitgestellten Diensten zufrieden sind 	<ul style="list-style-type: none"> ● Exportgeschäft⁴⁴ der geförderten Unternehmen nach 18 Monaten in % ● Anstieg des geschaffenen Mehrwerts nach 18 Monaten ● Zahl der geschaffenen/erhaltenen Brutto-/ Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze)
164 Gemeinsame Dienste für Unternehmen (Unternehmensparks, Gründerzentren, Förderung, Promotionskampagnen, Vernetzung, Konferenzen, Messen)	<ul style="list-style-type: none"> ● Zur Verfügung gestellte Industrieflächen in ha ● Zahl der finanziell unterstützten Projekte 	<ul style="list-style-type: none"> ● Direkte Privatinvestitionen der Unternehmen, die finanzielle Unterstützung erhalten (in Mio. Euro und % der Gesamtinvestitionen) ● Zufriedenheit der Begünstigten (Männer/Frauen) 	<ul style="list-style-type: none"> ● Anstieg des geschaffenen Mehrwerts nach 18 Monaten ● Zahl der geschaffenen/erhaltenen Brutto-/ Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze) ● % der regionalen Unternehmen (davon KMU), die nach 18 Monaten als Zulieferer für geförderte Unternehmen tätig sind ("Anstoßwirkung")
165 Finanzengineering	<ul style="list-style-type: none"> ● Zahl/Höhe der Wagnis- und Gründungskapitalfonds, die finanziell gefördert werden ● Zahl/Höhe der Garantiefonds, die finanziell gefördert 	<ul style="list-style-type: none"> ● Zahl neu gegründeter/aufgebauter Unternehmen (mit männlichen / weiblichen Inhabern) ● Zahl der KMU, die mit der Förderung zufrieden sind 	<ul style="list-style-type: none"> ● Zahl der geschaffenen/erhaltenen Brutto-/ Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze)

⁴⁴ Unter Exportgeschäft werden alle Verkäufe oder Verträge zwischen den Unternehmen der Region und Verbrauchern, Gesellschaften und Organisationen außerhalb der betreffenden Region verstanden. Ein weiterer Indikator kann der Wert der Verkäufe über die Landesgrenzen hinaus sein.

werden		(Männer/Frauen)	
Interventions bereiche	Output	Ergebnis	Wirkung
17 Fremdenverkehr			
171 Materielle Investitionen (Informationszentren, Beherbergung, Gaststätten, Ausstattung)	<ul style="list-style-type: none"> ● Zahl der neuen/modernisierten Fremdenzimmer ● Zahl der gebauten/umgebauten Hotels ● Zahl der geschaffenen/verbesserten Attraktionen ● Zahl der neu gegründeten Fremdenverkehrs betriebe (Männer/Frauen) 	<ul style="list-style-type: none"> ● % der neuen/modernisierten Fremdenzimmer ● Zahl der Übernachtungen pro Jahr in geförderten Beherbergungsbetrieben (nach einem Jahr) ● Kundenzufriedenheit (Männer/Frauen) in % 	<ul style="list-style-type: none"> ● Geschaffener Mehrwert pro Jahr (in %) ● Zahl der geschaffenen/erhaltenen Brutto-/ Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze)
172 Immaterielle Investitionen (Planung und Organisation eines touristischen Angebots, Aktivitäten in den Bereichen Sport, Kultur und Freizeit)	<ul style="list-style-type: none"> ● Zahl der geförderten Wirtschaftseinheiten (mind. fünf Tage) ● Zahl der geförderten Festivals und Veranstaltungen 	<ul style="list-style-type: none"> ● Durchschnittliche Kosten eines Aufenthalts (Euro pro Person) ● Durchschnittliche Zahl der Besucher pro Tag 	<ul style="list-style-type: none"> ● Zahl der Besucher pro Jahr in den geförderten Einrichtungen ● Geschaffener Mehrwert (in %) ● Zahl der geschaffenen/erhaltenen Brutto-/ Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze)
173 Gemeinsame Dienste für Unternehmen im Fremdenverkehrsbereich (einschl. Promotionskampagnen, Vernetzung, Konferenzen, Messen)	<ul style="list-style-type: none"> ● Zur Verfügung gestellte Fläche (m²) ● Zahl der geförderten neuen Marketing-Initiativen / Programme ● Zahl der organisierten Konferenzen/ Ausstellungen 	<ul style="list-style-type: none"> ● % der Konferenzen / Ausstellungen im Zusammenhang mit der lokalen Wirtschaftstätigkeit ● Zufriedenheit der Begünstigten (Männer/Frauen in %) 	<ul style="list-style-type: none"> ● Zahl der gewerblichen Kontakte von örtlich ansässigen Unternehmen aufgrund der Aktivitäten des Zentrums nach einem Jahr ● Zahl der Unternehmen, die das Promotionszentrum genutzt haben (nach einem Jahr)
174 Berufliche Bildung	<ul style="list-style-type: none"> ● Zahl der verfügbaren anerkannten Qualifikationen ● Umfang der Ausbildung (Stunden x Teilnehmer) 	<ul style="list-style-type: none"> ● % der Teilnehmer mit erfolgreichem Abschluß (Männer/Frauen) 	
Interventions bereiche	Output	Ergebnis	Wirkung
18 Forschung, technologische Entwicklung und Innovation (FuE/I)			
181 Forschungsprojekte an Hochschulen und in Forschungsinstituten	<ul style="list-style-type: none"> ● Zahl der geförderten Forschungsprojekte ● Zahl der geförderten Wissenschaftler / Forschungsstipendiaten (davon Männer/Frauen) 	<ul style="list-style-type: none"> ● % der erfolgreich abgeschlossenen Projekte (Veröffentlichungen, usw.) ● Zahl der geförderten Doktoranden (davon Frauen/Männer) 	<ul style="list-style-type: none"> ● Zahl der Patente aus innovativen Entwicklungen ● Zahl der Unternehmensgründungen durch Akademiker

			<ul style="list-style-type: none"> ● Zahl der Beschäftigten im Bereich FuE/I (Zahl und % aller Arbeitsplätze, Frauen/Männer)
182 Innovation und Technologietransfer, Vernetzung von und Partnerschaften zwischen Unternehmen und/oder Forschungszentren	<ul style="list-style-type: none"> ● Zahl der Unternehmen, die für FuE/T-Projekte und Technologieerwerb finanziell unterstützt werden ● Zahl der Beratungen/ Schulungsstunden ● Zahl der geförderten KMU ● Zahl der Kooperationsprojekte zwischen geförderten Unternehmen und Forschungs instituten 	<ul style="list-style-type: none"> ● Zahl der örtlichen Unternehmen, die sich an geförderten gemeinsamen Forschungsprojekten beteiligen (davon KMU) ● Anstieg der FuE/I-Investitionen durch Unternehmen, die an geförderten gemeinsamen Forschungsprojekten teilnehmen ● % der KMU, die mit dem Dienst zufrieden sind 	<ul style="list-style-type: none"> ● Zahl der Kooperationsabkommen zwischen Forschungsinstituten und geförderten Unternehmen nach einem Jahr ● Zahl der regionalen Unternehmen, die nach 18 Monaten positive Nebeneffekte melden (davon KMU) ● Zahl der geförderten Unternehmen, die Patente oder Lizenzen erwerben oder sich an Kooperationsprojekten beteiligen (nach einem Jahr) ● Zahl der neuen Produkte/Verfahren, die von geförderten Unternehmen vermarktet werden ● geschaffener Mehrwert nach zwei Jahren ● Zahl der geschaffenen/erhaltenen Brutto-/Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze)
183 FuE/I-Infrastrukturen	<ul style="list-style-type: none"> ● Zur Verfügung gestellte Fläche (ha) ● gebaute/sanierte Fläche (m²) ● Zahl der bereitgestellten gemeinsamen Dienste 	<ul style="list-style-type: none"> ● Zahl der geschaffenen FuE-Arbeitsplätze (FTE-Einheiten, davon Frauen/Männer) ● Zahl der KMU mit Zugang zu den gemeinsamen Diensten 	<ul style="list-style-type: none"> ● Zahl der Kleinunternehmen, die sich im Park ansiedeln (nach einem Jahr) ● Zahl kleiner High-Tech-Unternehmen, die sich im Park ansiedeln (nach einem Jahr) ● Zahl der geschaffenen/erhaltenen Brutto-/Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze)

2 HUMANRESSOURCEN

Anm.: Die Zahlen über die geschaffenen oder erhaltenen Brutto-/Nettoarbeitsplätze sollten nach Möglichkeit nach Männern/Frauen aufgeschlüsselt werden.

Interventionsbereiche	Output	Leistung	Wirkung
21 Arbeitsmarktpolitische Maßnahmen			
Unterstützung an Einzelpersonen (Ausbildung, Beratung und Anleitung, Beschäftigungshilfe, integrierte Maßnahmen...)	●Zahl der Begünstigten	●Stellungsquote der Begünstigten in Beschäftigung (%).	●Stellungsquote der Begünstigten nach 1 Jahr. ●Arbeitslosigkeitsreduzierung der Zielbevölkerung (%)
Unterstützung an Strukturen und Systeme (Lehrerausbildung - Beratungs- und Anleitungsdienste...)	●Anzahl von angebotenen Plätzen ●Anzahl ausgebildeter Ausbilder ●Anzahl von Projekten	●Arbeitslose, die zusätzliche Beratungsdienste (%) nutzen ● Abdeckungsquote der Bezugsgruppe (%)	●Stellungsquote der Begünstigten nach 1 Jahr
22 Soziale Integration			
Unterstützung an Einzelpersonen (Bahnen zu Integration, integrierte Maßnahmen, bestimmte Ausbildungsmaßnahmen.)	●Zahl der Begünstigten	●Anstieg der Dauer von Berufserfahrung (Durchschnitt/Begünstigte) ●Zunahme der Qualifikationen (Zahl der Begünstigten, die ein Diplom oder Zertifikat erhalten haben,) ●Zufriedenheit der Begünstigten (%)	●Stellungsquote der Empfänger nach 1 Jahr. ●Verringerung der Arbeitslosigkeit in der Zielgruppe (%)
Unterstützung an Strukturen und Systeme (soziale Begleitung, Informationen, lokale Initiativen für die Entwicklung der Beschäftigung...)	●Zahl lokaler Initiativeprojekte ●Zahl von Projekten	●Zahl von Vereinigungen sozioökonomischer Partner im Rahmen örtlicher geschaffenen Pakte ● Abdeckungsquote der Bezugsgruppe (%)	●Nachhaltigkeit der Vereinigungen (%; die noch 2 Jahre nach Ende von Unterstützung existieren,)
23 Ausbau der allgemeinen und beruflichen Bildung (Einzelpersonen, Unternehmen)			
Unterstützung an Einzelpersonen und Unternehmen (Laufbahntwicklungswege für frühe Studienabbrecher, wenig ausgebildete Erwachsene, Schulungsmaßnahmen für Ausbildern...)	● Zahl der geschaffenen Ausbildungsplätze (Stunden, Tage) ● Zahl der Projekten	● Zahl der Ausbilder/Berater... nachdem ihre Qualifikationen zugenommen haben Abdeckungsquote der Bezugsgruppe (%)	● Verminderung der frühen Studienabbrecher (%) ● Zunahme der Qualifikationen in der Zielgruppe (%)

Interventionsbereiche	Output	Leistung	Wirkung
24 Anpassungsfähigkeit der Arbeitskräfte, Unternehmergeist und Innovationsfähigkeit, Informations- und Kommunikations-technologien			
Unterstützung Einzelpersonen, Unternehmen	<ul style="list-style-type: none"> ● Zahl der Angestellten in Schulungsprogrammen (Art, Dauer) ● Zahl der KMU, die finanzielle Unterstützung für Schulungsprogramme erhalten (Umfang, Art, Dauer) ● Zahl der Begünstigten 	<ul style="list-style-type: none"> ● Anstieg des KMU-Budget für Schulungsmaßnahmen (Anstieg in %) 	<ul style="list-style-type: none"> ● Zahl der begünstigten Personen, die nach 2 Jahren ein Unternehmen gegründet haben (Männer / Frauen) ● Anstieg des Mehrwerts nach 18 Monaten ● Zunahme der Produktivität der Beschäftigten aufgrund verbesserter Fähigkeiten (Anstieg des Umsatzes in %/Angestellte) ● Zahl der geschaffenen/erhaltenen Brutto-/Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze)
Unterstützung an Strukturen und Systemen (Sozialbegleitung, Informationen, lokale Entwicklungsinitiativen, Sozialbereich...)	<ul style="list-style-type: none"> ● Zahl der für KMU Beratungsdienste, ● Zahl der Projekten 		
25 Positive Beschäftigungsmaßnahmen für Frauen			
Unterstützung an Personen (Laufbahnentwicklung von Frauen, Unternehmergeist unter Frauen...)	<ul style="list-style-type: none"> ● Zahl der Begünstigten ● Betrag der Wirtschaftszuschüsse / Darlehen ● Stunden in 	<ul style="list-style-type: none"> ● Stellungsquote der Begünstigten (%) 	<ul style="list-style-type: none"> ● Zunahme der Aktivitätsrate der Frauen am Arbeitsmarkt (%) ● Zunahme der Beschäftigungsrate der Frauen am Arbeitsmarkt nach 1 Jahr ● Anstieg der Männer/Frauen die männlich/weiblich-beherrschten Sektoren und Arbeitsplätze nach 2 Jahren beschäftigt sind (%)
Unterstützung an Strukturen und Systemen	<ul style="list-style-type: none"> ● "Bestimmte Aktion" Schulungsmaßnahmen 	<ul style="list-style-type: none"> ● Anteil der Zielgruppe, die an "Bestimmten Aktionen" 	

(Sensibilisierungsveranstaltungen, Vernetzung.)	<ul style="list-style-type: none"> ● verbraucht ● Zahl der Sensibilisierungsveranstaltungen ● Zahl der geförderten neuen Netze ● Zahl der Projekten 	teilgenommen hat	<ul style="list-style-type: none"> ● Lebensdauer der neuen Netze (2 Jahre nach Abschluß der Förderung noch bestehende Netze in %)
32 Infrastrukturen im Bereich Telekommunikation und Informationsgesellschaft			
322 Informations- und Kommunikationstechnologie (einschließlich Maßnahmen für größere Sicherheit und sichere Datenübertragung)	<ul style="list-style-type: none"> ● Zahl und Anstieg in % der digitalen Telefonanschlüsse je 1000 Einwohner ● Zahl der ISDN-Teilnehmer je 1000 Einwohner ● Länge des gebauten Breitbandnetzes (Glasfaser-km) ● Zahl der Internet-Hosts je 1000 Einwohner 	<ul style="list-style-type: none"> ● Rückgang der Zahl der Netzausstürze ● Zahl der geschaffenen Dienstleistungen (Internet-Zugang) ● Zahl der KMU und Großunternehmen, die Dienste der Informationsgesellschaft entwickeln und anbieten ● Verbindungsstunden insgesamt / Monat (nach 6 Monaten) 	<ul style="list-style-type: none"> ● Geschaffene/erhaltene Brutto-/Netto-Arbeitsplätze (Anzahl und % aller Arbeitsplätze)
Interventionsbereiche	Output	Leistung	Wirkung
323 Dienste und Anwendungen für den Bürger	<ul style="list-style-type: none"> ● Zahl der Umschulungsmaßnahmen ● Zahl der geschaffenen online-Dienste ● Zahl der Schulungsstunden (Stunden x Teilnehmer) ● Zahl der Teilnehmer (Männer/ Frauen) 	<ul style="list-style-type: none"> ● Zahl der Nutzer/Schulungsteilnehmer ● Zufriedenheit der Nutzer/Schulungsteilnehmer (%) 	<ul style="list-style-type: none"> ● % der Teilnehmer, die innerhalb von 6 Monaten vermittelt werden konnten (davon Frauen/Männer)
324 Dienste und Anwendungen für KMU (elektronischer Geschäftsverkehr, Aus-/Weiterbildung, Vernetzung)	<ul style="list-style-type: none"> ● Zahl der neu gegründeten Firmen, die Dienste der Informationsgesellschaft anbieten (online, e-Geschäftsverkehr usw.) 	<ul style="list-style-type: none"> ● Zahl der KMU mit Zugang zu den angebotenen Diensten (Internet-Zugang) ● Zahl der KMU und Großunternehmen, die Dienste der Informationsgesellschaft entwickeln und anbieten ● Zahl der Einwahlpunkte ("Points of Presence") je Ortszone 	<ul style="list-style-type: none"> ● Zahl der geschaffenen/erhaltenen Brutto-/Netto-Arbeitsplätze nach 2 Jahren (Anzahl und % aller Arbeitsplätze)

4.10

Literatur

- Bachtler J./Michi R., A New Era in EU Regional Policy Evaluation? The Appraisal of Structural Funds, in: *Regional Studies*, Vol. 29.8, pp. 745 – 751.
- Bachtler J./Polverari, L./Taylor S./Ashcroft B./Swales, K., *Methodologies used in the Evaluation of Effectiveness of European Structural Funds: A Comparative Assessment*, Study of the European Policies Research Centre and Fraser of Allander Institute, University of Strathclyde, Glasgow 2000
- Bachtler, J./Taylor S., *Objective 2: Experiences, Lessons and Policy Implications*, European Policies Research Centre, Glasgow, Juli 1999
- Bachtler, J./Taylor S./Kearney C., *Extended Synthesis of agreed Single Programming Documents in Objective 2 Areas, 1994-1996*, European Policies Research Centre, University of Strathclyde, Glasgow, 1996
- Bachtler, J./Taylor S./Kearney C., *Monitoring and Evaluation in Objective 2 Programmes: Progress and Challenges*, IQ-net thematic paper Series 1, No 4, European Policies Research Centre, University of Strathclyde, Glasgow 1996,
- BAW-Monatsberichte, *Evaluierung Europäischer Strukturfonds-Programme, Zum Stand der Kunst – Teil II*, Bremer Ausschuß für Wirtschaftsforschung, Bremen 4/1997
- Bremer Zeitschrift für Wirtschaftsforschung, *Evaluierungsstudie - Theoretische Grundlagen zur Methodik der Evaluierung*, Bremer Ausschuß für Wirtschaftsforschung 4/1994
- C3E (1995) Centre for European Evaluation Expertise, Information Note
- European Commission (EC) (2000a), *The New Programming period 2000 – 2006: methodological working papers*, Working paper 3, Indicators for Monitoring and Evaluation, An indicative methodology
- European Commission (EC) (2000b), *Annual Evaluation Review 1999, Highlights of results from evaluation of expenditure programmes*, Working paper of the Commission departments; e-mail from Mr. Timo Takkula;
<http://europa.eu.int/comm/budget/evaluation/en/bilan/evaluationreview1999.html#TOC475502965>
- European Commission (EC) (1999a), *Evaluating socio-economic programmes*, MEANS Collection Volume 1-6; Luxemburg, 1999
- European Commission (EC) (1999b), *Thematic Evaluation of Structural Fund Impacts on SMEs*, 1999
- European Commission (EC) (1999c), *The Thematic Evaluation of the Partnership Principle*, 1999
- European Commission (EC) (1997), *The New Regional Programmes 1997-99 under Objective 2 of the Community's Structural Policies – focusing on Job Creation*, COM (97), 524 final, Commission of the European Communities, Brussels
- Europäische Kommission (EK) (1999), *Reform der Strukturfonds 2000-2006, Eine vergleichende Analyse*, 1999
- Europäische Kommission (1995): *Strukturfonds der EU 1994 - 1999. Gemeinsame Leitlinien für die Begleitung und die Zwischenbewertung.* (DE/05/95/52070100.p00(FR).

- European Policies Research Centre (EPRC 2000), At the Starting Block: Review of the new Objective 2 Programmes, Glasgow 2000
- Ernst & Young "Ex-Post Evaluation of Objective 2 Programmes 1989-93: Synthesis Report for the European Commission, DG XVI, 1997
- Holzinger E., Programm-Evaluation, Theoretische Grundlagen und Anwendungsmöglichkeiten in Raumordnung und Regionalpolitik, ÖIR, Wien
- Michi, R./Fitzgerald R., (1997), The Evolution of the Structural Funds, in Bachtler J./Turok I., The Coherence of EU Regional Policy, Contrasting Perspectives on the Structural Funds, London 1997
- Morgan, K. (1996) The Information Society, Opportunities for SMEs in Objective 2 Regions, Department of City and Regional Planning, University of Wales, Cardiff, paper prepared for the Objective 2 Desk Officers, 1996
- Österreichischen Raumordnungskonferenz, Zwischenbewertung der Interventionen der Ziel-2-Programme, des RESIDER-II- und des RECHAR-II-Programmes in der Programmperiode 1995-99 in Österreich, Studie der JOANNEUM RESEARCH in Kooperation mit dem Bremer Ausschuss für Wirtschaftsforschung und dem European Policies Research Centre, Schriftenreihe Nr. 140, Wien 1998
- Scottish Executive, Measuring Progress, A Handbook for Monitoring European Structural Fund Projects, 2000
- Taylor, S., RTD/Innovation Policies in Objective 2 Programmes, Challenges and best practice, IQ-net thematic paper Series 1, No 2, European Policies Research Centre, University of Strathclyde, Glasgow 1996

Programme und Evaluierungen aus den Regionen:

- Western Scotland
- West Midlands
- Bremen
- Nordrhein Westfalen
- Österreich

5 KAPITEL FÜNF DIE EVALUATION VON ARBEITSMARKTPOLITIK

Birgit Woitech, Klaus Zinöcker

KAPITEL FÜNF DIE EVALUATION VON ARBEITSMARKTPOLITIK.....	1
EINLEITUNG.....	2
ANSATZPUNKTE DER EVALUIERUNG VON ARBEITSMARKTPOLITIK IN ÖSTERREICH.....	2
<i>Warum Evaluierung von arbeitsmarktpolitischen Entscheidungen?.....</i>	3
<i>Funktionen der Evaluierung.....</i>	3
<i>Annäherung Evaluationsgegenstand: Was wird eigentlich evaluiert?.....</i>	5
<i>Annäherung Evaluationstypus: Evaluierung aus Sicht der arbeitsmarktpolitischen Terminologie.....</i>	6
5.2.4.1 Evaluierung von arbeitsmarktpolitischen Projekten.....	6
5.2.4.2 Evaluierung von Trägern arbeitsmarktpolitischer Maßnahmen.....	7
5.2.4.3 Die Evaluierung arbeitsmarktpolitischer Programme.....	7
<i>Annäherung seitens der Evaluierungsdimensionen.....</i>	11
5.2.5.1 Zeitliche Evaluierungsdimension.....	11
5.2.5.2 Inhaltliche Evaluierungsdimension.....	11
<i>Zusammenfassung.....</i>	13
EIN BEZUGSRAHMEN FÜR EINE IDEALTYPISCHE EVALUIERUNG DER ARBEITSMARKTPOLITIK.....	15
<i>Zur Kontextanalyse.....</i>	17
5.3.1.1 Externe Rahmenbedingungen.....	17
5.3.1.2 Rechtlich-institutionelle Rahmenbedingungen.....	17
5.3.1.3 Beurteilung der arbeitsmarktpolitischen Strategie.....	18
5.3.1.4 Fragestellungen bei Evaluation abwien.....	18
<i>Zur Problemanalyse.....</i>	19
5.3.2.1 Phase I der Problemanalyse: Angebotsseitige Analyse.....	19
5.3.2.2 Phase II der Problemanalyse: Nachfrageseitige Analyse.....	19
5.3.2.3 Phase III der Problemanalyse: Mismatch-Analyse.....	19
5.3.2.4 Phase IV der Problemanalyse: Programmbewertung.....	20
<i>Zum Monitoring.....</i>	20
<i>Zur Wirkungsanalyse.....</i>	21
5.3.4.1 Was misst Wirkungsanalyse?.....	22
5.3.4.1.1 Bewertungs- und Erfolgskriterien.....	23
5.3.4.2 Verfahren und Methoden der Wirkungsanalyse.....	25
5.3.4.2.1 Mikroökonomische Evaluationsstudien.....	26
5.3.4.2.2 Matching Pairs Analyse.....	26
5.3.4.2.3 Vorher-Nachher Vergleich.....	28
5.3.4.3 Analyse des Integrationserfolges im Rahmen der Evaluation abwien.....	29
<i>Bilanzierung.....</i>	29
VERGLEICH.....	30
ANFORDERUNGEN AN ARBEITSMARKTPOLITISCHE EVALUIERUNGEN IN ÖSTERREICH.....	33
ZUSAMMENFASSUNG.....	40
.....	40
5.7 LITERATUR.....	41

5.1 Einleitung

Der Themenkomplex Evaluation hat sich im Bereich der Arbeitsmarktpolitik in den letzten Jahren sowohl methodisch als auch von der Breite des Einsatzes stark weiterentwickelt. Auch wenn sehr oft kritisiert wird, dass die europäische "Evaluationskultur" der US-amerikanischen hinterherhinkt, vor allem was den Einsatz fortgeschrittener ökonometrischer Methoden betrifft, kann man dennoch feststellen, dass die Zahl der Evaluationsstudien kontinuierlich wächst. Dies ist zum einen darauf zurückzuführen, dass die Ausgaben für aktive Arbeitsmarktpolitik in den letzten Jahren stetig gewachsen sind, damit auch der Legitimationsdruck gegenüber der Öffentlichkeit und zum anderen vermehrt Anstrengungen unternommen werden, die Datenlage zu verbessern, um ein anspruchsvolleres - unter Umständen auch aussagekräftigeres - Studiendesign möglich zu machen.

Dennoch ist die Evaluierung von Maßnahmen der aktiven Arbeitsmarktpolitik noch ein relativ junges Feld. „Nichtsdestotrotz zeigt ein sorgfältiger Überblick über den bestehenden Erfahrungsschatz [über Evaluierungen von Arbeitsmarktmaßnahmen] , dass diese Forschung noch in den Anfängen steckt.“¹ Häufig wird dabei auf das "Vorbild" USA verwiesen, in der seit langem beträchtliche Mittel für die wissenschaftliche Evaluation von Arbeitsmarktprogrammen bereitgestellt werden. Dabei handelt es sich sehr häufig um sozialwissenschaftliche Experimente, mittels derer die Wirksamkeit von Arbeitsmarktprogrammen im Sinne einer Verbesserung der wirtschaftlichen Situation der TeilnehmerInnen überprüft wurde.

Bevor jedoch die Frage nach der Bestimmung der Wirksamkeit von arbeitsmarktpolitischen Maßnahmen geklärt werden kann, ist es notwendig, grundlegende Fragestellungen der Evaluierung von Arbeitsmarktpolitik zu erörtern. Im ersten Abschnitt werden daher die Ansatzpunkte der Evaluierung der Arbeitsmarktpolitik in Österreich dargestellt. Dabei geht es vor allem darum abzuklären, welche Funktionen Evaluierungen zu erfüllen haben, was der eigentliche Gegenstand der Evaluierung ist und innerhalb welcher Dimensionen (inhaltlich und zeitlich) sie sich bewegen. Dabei wird versucht, sich an einem österreichischen Beispiel für Best-Practice im Bereich der Evaluierung der Technologiepolitik anzulehnen. Im Anschluss daran wird eine Einordnung der Evaluierungspraxis für Österreich vorgenommen. Darauf aufbauend wird eine "typische" (hinsichtlich Zeitpunkt, Inhalt und Methode) Arbeitsmarktevaluation konstruiert, welche im dritten Abschnitt dann einem Beispiel einer idealtypischen Evaluierung gegenübergestellt wird, um danach die österreichische Evaluierungspraxis anhand aktueller Ausschreibungen des Arbeitsmarktservices analysieren zu können. Das Schwergewicht dieses Teils wurde dabei auf die Darstellung einer idealtypischen Evaluierung und den ihr zugrundeliegenden Analyseschritten, mit denen die Wirksamkeit von Programmen (und Maßnahmen) überprüft werden soll, gelegt.

5.2 Ansatzpunkte der Evaluierung von Arbeitsmarktpolitik in österreich

... oder die Frage, wovon spricht man eigentlich, wenn man von arbeitsmarktpolitischer Evaluierung in Österreich spricht?

Im folgenden soll versucht werden, an Hand mehrerer Faktoren bzw. Fragestellungen die Thematik der arbeitsmarktpolitischen Evaluierung (in Österreich) enger zu fassen. Warum

¹ Christoph M Schmidt in DIW [2000] , Seite 428

eigentlich die Evaluierung von arbeitsmarktpolitischen Entscheidungen? Welche Ziele werden damit verfolgt? Zu welchem Zweck? Was ist der Evaluierungsgegenstand und in welchen inhaltlichen und zeitlichen Dimensionen bewegt sich arbeitsmarktpolitische Evaluierung? Die Analyse der verschiedenen Evaluierungsdimensionen orientiert sich an Fritz, O., Hutschenreiter G., Storn D.: Evaluierung von FTE-Programmen zwischen Best Practice Entwicklung und österreichischen Strukturen. Auf Grundlage der aus der theoretischen Evaluierungsliteratur bekannten Definitionen und methodischen Ansätzen soll ein Bild vom 'state of the art' arbeitsmarktpolitischer Evaluierung in Österreich entworfen werden.

5.2.1 Warum Evaluierung von arbeitsmarktpolitischen Entscheidungen?

Die typischerweise hohe Kosten, die mit arbeitsmarktpolitischen Eingriffen verbunden sind, erfordern ziel- und erfolgsorientierten Einsatz der Mittel, um auch gesellschaftlich gerechtfertigt zu sein. Die Erwartungen an die Arbeitsmarktpolitik können dabei sehr unterschiedlich formuliert werden und ergeben sich aus der allgemeinen Politikorientierung und theoretischen Grundpositionen. Dennoch lassen sich gemeinsame Handlungsfelder bzw. Begründungen für die Forderung nach einer wissenschaftlichen Evaluierung arbeitsmarktpolitischer Entscheidungen identifizieren:

- Besondere Sorgfaltspflicht in der Verwendung öffentlicher Gelder. Daraus resultiert eine Pflicht zur Transparent-Machung dieser Verwendung und darüber hinaus ein Druck auf die Verwaltung zur Rechtfertigung ihrer Allokationsentscheidungen.
- Kontrollpflicht, Mängelidentifikation, Aufdeckung von Schwachstellen im Verwaltungsablauf.
- Beschaffung, Aufbereitung, Interpretation arbeitsmarktpolitisch relevanter Fakten².

5.2.2 Funktionen der Evaluierung

Grundsätzlich sind drei Funktionen von Evaluierung unterscheidbar:

- die Legitimationsfunktion,
- die Informationsfunktion (auch eine informationsorientierte Lernfunktion) und schließlich
- die Lenkungsfunktion im Sinne einer zielorientierten Kontrollfunktion.

Erfüllt Evaluierung all diese Funktionen, so ist von einem hohen Stellenwert auszugehen, der einem Evaluierungssystem zukommt. Die Ergebnisse der Auswertungen in den folgenden Kapiteln lassen den Schluss zu, dass arbeitsmarktpolitische Evaluierungen nur zum Teil jede dieser Funktionen erfüllen:

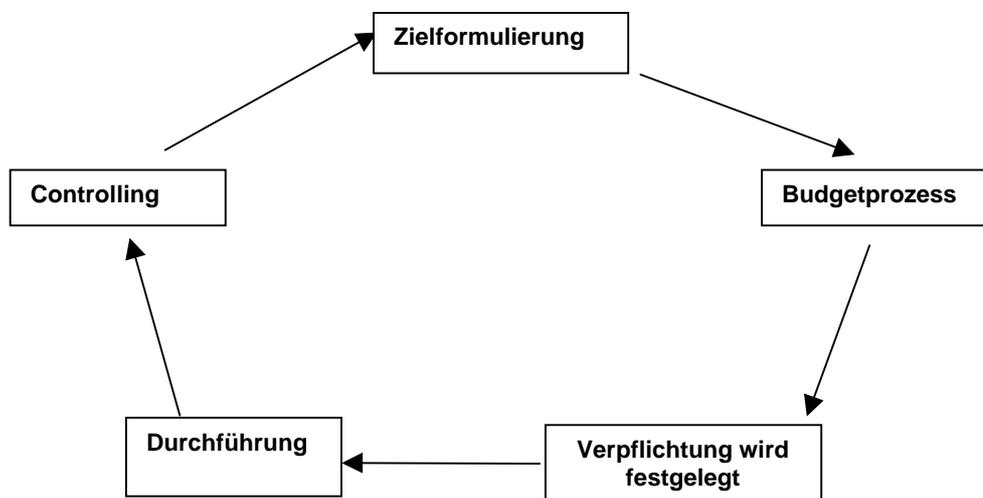
Diese Evaluierungsfunktionen für die Arbeitsmarktpolitik spiegeln sich in der Organisationsstruktur des zentralen Akteurs, dem Arbeitsmarktservice AMS, und dem innerhalb dieser Organisation gewählten Form des Verwaltungsablaufs, einer Ziel- und Ergebnissteuerung. (Abbildung 1) wider. Das Arbeitsmarktservice sieht sich heute selbst als „das führende kundenorientierte Dienstleistungsunternehmen am Arbeitsmarkt in Österreich“³. Neben der Vermittlung von

² Fritz, Hutschenreiter u. Storn [1997] schreiben von einer operativen und einer strategischen Informationsfunktion von Evaluierung.

³ Vgl. dazu <http://www.ams.or.at/amsallg/index.htm>

Arbeitssuchenden auf offene Stellen gehören zu den „Produkten“ des Dienstleistungsunternehmens AMS auch die Beratung, Information, Qualifizierung und finanzielle Förderung von Arbeitssuchenden und Unternehmen (vgl. dazu Tabelle 1, Seite 6). Mit dem Arbeitsmarktservicegesetz aus dem Jahre 1994 wurde die Arbeitsmarktverwaltung (AMV) aus dem Bundesministerium für Arbeit, Gesundheit und Soziales ausgegliedert und das Arbeitsmarktservice (AMS) als Dienstleistungsunternehmen des öffentlichen Rechts konstituiert. Heute basieren die Steuerungsmethoden des Arbeitsmarktservice auf einen Methodenmix aus Normierung, Ressourcensteuerung und Zielvorgaben, wobei die aktive Arbeitsmarktpolitik (Dienstleistungen und Förderungen) durch Instrumentennormierung und Zielvorgaben gelenkt wird.⁴ Jährlich werden 8 Bundesziele entwickelt, flankiert werden diese von je einem Landesziel und einem regionalen Ziel. (Vgl. dazu Tabelle 2, Seite 9) Die Organisation Arbeitsmarktservice ruht auf den drei Säulen Bundesgeschäftsstelle (BGS), Landesgeschäftsstelle (LGS) und regionaler Geschäftsstelle (RGS).

Abbildung 1: Prinzipskizze einer Ziel- Ergebnissteuerung



Quelle: Gruber, Zinöcker et al. Seite 46 [nach Frieder Naschold: Ergebnissteuerung, Wettbewerb, Qualitätspolitik, edition sigma 1995]

Aus den in Abbildung 1 dargestellten Etappen des Steuerungsablaufs können verschiedene Funktionen einer Evaluierung illustriert werden:

- Als Legitimation des Trägers gegenüber dem Auftraggeber LGS, aber auch im Verhältnis LGS gegenüber der darüber liegenden Organisationseinheiten (BGS, Ministerium).
(--> Festlegung der Verpflichtung)
- Als Informations- und Kontrollmittel für die auftraggebenden Stellen in der LGS
(--> Controlling)
- Als Lenkungsfunktion für zukünftige Gestaltungen des Zielsystems des AMS bzw., damit verbunden, für die zukünftige Ausgestaltung der

⁴ Vortrag von Herbert Buchinger, aus: ÖSB Tagungsbericht 19/20 Oktober 1999. Unter „Steuerung über Normen“ versteht Buchinger Anweisungen der Zentrale an die Ausführenden, was wie zu tun ist, wobei die Zentrale für das Ergebnis verantwortlich sei; bei einer „Steuerung über Ressourcen“ steuert die Zentrale per Zuteilung von Ressourcen, die dorthin gelenkt werden, wo beste Effekte erreicht werden können. Bei einer Steuerung über Ziele schließlich gibt die Zentrale Ziele bzw. zu erreichende Effekte vor, die Maßnahmen zur Zielerreichung werden von den Ausführenden selbst bestimmt.

Maßnahmenplanung.
(--> Zielformulierung)

Evaluierungen von Arbeitsmarktpolitik erfüllen in Österreich auf unterschiedlichen Ebenen (BGS – LGS – RGS) in unterschiedlichen Etappen (Ziel- Ergebnis- Steuerung) unterschiedliche Funktionen. Hierbei entzieht es sich der Kenntnis der AutorInnen, in welchem Ausmaß diese Erfüllung passiert, wie weit Funktion über den auftraggebenden Personenkreis hinaus erfüllt wird und in wie weit innerhalb der Organisation AMS es zu einem sinnvollen Zusammenführen dieser Ergebnisse kommt (Schnittstellenproblematik).

5.2.3 Annäherung Evaluationsgegenstand: Was wird eigentlich evaluiert?

Evaluierungen werden nicht in allen Feldern und von allen Akteuren der Arbeitsmarktpolitik – oder weiter gefasst, der Arbeitsmarkt- und Beschäftigungspolitik - in gleicher Art und Weise und im gleichen Ausmaß eingesetzt. Haupteinsatzgebiet ist die aktive Arbeitsmarktpolitik, die von der Beschäftigungspolitik bzw. von Maßnahmen der passiven Arbeitsmarktpolitik im nachfolgenden abgegrenzt werden soll.

Unter **Arbeitsmarktpolitik** ist die:

- Summe aller staatlichen Maßnahmen, die unmittelbar das Angebots- und Nachfrageverhalten auf dem Arbeitsmarkt beeinflussen - ohne Umweg über Güter- oder Kapitalmärkte und die
- Gesamtheit aller Aktivitäten, der für Arbeit zuständigen Ressorts auf Bundes- und Landesebene und der mit ihnen verbundenen Verwaltungseinheit, dem Arbeitsmarktservice

zu verstehen. Es werden drei Grundformen der aktiven Arbeitsmarktpolitik unterschieden, die

- Förderung von Humankapital,
- Veränderung von Anreizen,
- Bereitstellung von Arbeitsplätzen.

Das österreichische Arbeitsmarktservice unterteilt den Katalog seiner möglicher Maßnahmen der aktiven Arbeitsmarktpolitik in Beratung, Qualifizierung und Beschäftigung (vgl. dazu Tabelle 1 weiter unten).

Im Gegensatz dazu werden unter dem Begriff der passiven Arbeitsmarktpolitik alle Maßnahmen verstanden, auf die ArbeitnehmerInnen, Arbeitslose oder einfach Berechtigte einen rechtlichen Anspruch besitzen. Im allgemeinen können die Maßnahmen einer passiven Arbeitsmarktpolitik als eher traditionell angesehen werden.

Beispiele dafür sind:

- Ausgaben für Arbeitslosengeld,
- Sozialhilfe,
- Vorruhestandsmaßnahmen.

Eine wesentliche Unterscheidung zwischen aktiver und passiver Arbeitsmarktpolitik liegt somit im Ermessensspielraum der zuständigen Akteure: Bei den Maßnahmen der passiven

Arbeitsmarktpolitik besitzen diese keinerlei Ermessensspielraum zur Ver- und Zuteilung, bei jenen der aktiven Arbeitsmarktpolitik jedoch sehr wohl.

Detaillierter werden die Möglichkeiten der aktiven und passiven Maßnahmengestaltung der Arbeitsmarktpolitik im Rahmen des OECD Employment Outlook, OECD Taxonomie der Maßnahmen der Arbeitsmarktpolitik dargestellt⁵.

Beschäftigungspolitik ist im Vergleich zur Arbeitsmarktpolitik der weitere Begriff: Beschäftigungspolitik ist die Summe aller staatlichen Maßnahmen, die auf das Ziel der Vollbeschäftigung ausgerichtet sind. Ihre Instrumente können fiskalpolitischer, geldpolitischer, aber auch arbeitsmarktpolitischer Natur sein: Arbeitsmarktpolitik ist also ein Teilbereich der Beschäftigungspolitik.

Vorrangiger Gegenstand arbeitsmarktpolitischer Evaluierungen in Österreich sind Maßnahmen der aktiven Arbeitsmarktpolitik.

5.2.4 Annäherung Evaluationstypus: Evaluierung aus Sicht der arbeitsmarktpolitischen Terminologie

Bei der Sichtung einschlägiger Literatur zum Thema Evaluierung von Arbeitsmarktpolitik war auffällig, dass von keiner gemeinsamen Terminologie ausgegangen werden kann. Daher soll – um ein gemeinsames Verständnis in den einzelnen Abschnitten sicherzustellen – versucht werden, einige Grundbegriffe der Evaluierung im arbeitsmarktpolitischen Kontext genau zu umreißen, um etwaige Verwirrungen von Anfang an gezielt ausschließen zu können. Zielrichtung dieser Definition soll in erster Linie die Evaluationsart sein: Projekt-, Programm- bzw. Trägerevaluierung.

5.2.4.1 Evaluierung von arbeitsmarktpolitischen Projekten

Im Sprachgebrauch der Akteure der Arbeitsmarktpolitik ist ein „arbeitsmarktpolitisches Projekt“ mit einer „arbeitsmarktpolitischen Maßnahme“ gleichzusetzen: Ein Projekt ist die kleinste mögliche Einheit eines Evaluierungsgegenstandes in der Arbeitsmarktpolitik. Beispiele für solche Projekte bzw. Maßnahmen sind etwa die Beschäftigungsgesellschaft Liezen, ein metallverarbeitender Betrieb mit dem vorrangigen Ziel der Reintegration von Langzeitarbeitslosen oder die Beratung von arbeitslosen Jugendlichen.

Grundsätzlich unterscheidet der Katalog möglicher Maßnahmen des Arbeitsmarktservice folgende Projekte:

Tabelle 1: Mögliche Maßnahmen einer aktiven Arbeitsmarktpolitik des AMS

Beratung	Qualifizierung	Beschäftigung
Allgemeine Beratung	Allgemeine Qualifizierung	Allgemeine Beschäftigungsprojekte
Berufsorientierung	Förderung des Lehrabschlusses	Beschäftigungsgesellschaften
Externe Beratung	Weiterbildung	Gemeinnützige Arbeitskräfteüberlassung

⁵ Innerhalb der EU kommt eine kompatible Taxonomie zum Einsatz: EU, Common Indicators for monitoring the Employment Guidelines, Mai 99

Beratung	Qualifizierung	Beschäftigung
Aktivgruppen	Arbeitsmarktpolitische Selbstorganisationsprojekte	Gemeinnützige Beschäftigungsprojekte
Job-Coaching	Arbeitsstiftungen	Sozialökonomische Betriebe
Gründerberatung		Förderung von KBE

Quelle: Gruber, Zinöcker et al. [2000], Seite 19 (Nach dem Arbeitsmarktpolitischen Handbuch des Arbeitsmarktservice Steiermark)

Die in Tabelle 1 dargestellten Maßnahmen stellen die Gesamtheit aller möglichen Evaluierungsgegenstände arbeitsmarktpolitischer Projekte innerhalb des österreichischen Arbeitsmarktservices dar; auf die genaue Ausgestaltung und den Anforderungskatalog dieser Evaluierungen wird in Kapitel 5.5, Anforderungen an arbeitsmarktpolitische Evaluierungen, eingegangen, da sie, wie ausgeführt werden wird, die Hauptform der Evaluierung der Arbeitsmarktpolitik in Österreich darstellen.

5.2.4.2 Evaluierung von Trägern arbeitsmarktpolitischer Maßnahmen

Mit der Durchführung von arbeitsmarktpolitischen Maßnahmen werden in der Regel Trägerorganisationen durch das AMS beauftragt. Die vorrangige Organisationsform dieser Trägerorganisationen sind Vereine (als Beispiele möge hier „Rettet das Kind“, „pro mente“ oder die Volkshilfe dienen) und GmbHs (Bei den GmbHs handelt es sich zumeist um regional verankerte Gesellschaften, in denen sich etwa die Gemeinden eines Arbeitsmarktbezirkes zusammengeschlossen haben). Zwar ist die Konzentration eines Trägers auf eine einzelne Maßnahme möglich (ein Träger führt nur eine bestimmte Maßnahme durch), doch lassen die Entwicklungen der letzten Jahre einen Trend zu professionellen Multimaßnahmenträgern erkennen, die z.B. überregional ein gewisses Produkt des AMS oder diverse Maßnahmen gleichzeitig an verschiedenen Orten anbieten.

Die Auswertung von Evaluierungen in Kapitel 5.5 berücksichtigt keine Evaluierungsform, die sämtlich Maßnahmen eines Trägers umfasst. Allerdings ist die Unterscheidung zwischen der Evaluierung einer Maßnahme und eines Trägers genau dann schwierig, wenn dieser Träger genau eine arbeitsmarktpolitische Maßnahme durchführt. Die Erkenntnisse aus Kapitel 5.5 lassen den Schluss zu, dass es sich in jenen Fällen, wo Träger und Maßnahme deckungsgleich sind, es sich sehr wohl um Projektevaluierungen handelt.

5.2.4.3 Die Evaluierung arbeitsmarktpolitischer Programme

Der Begriff der Programmevaluierung wird in der Literatur am missverständlichsten verwendet: Unter einem „Programm“ kann sehr wohl eine klar definierte, kleinste Einheit eines Evaluierungsgegenstandes verstanden werden, „Programmevaluierung“ kann also synonym für die Evaluierung eines Projektes stehen. Als „Programmevaluierung“ kann aber auch die Evaluierung aller Projekte innerhalb eines arbeitsmarktpolitischen Zieles zusammengefasst werden.

Grundsätzlich ist die Zielplanung des AMS zentral vorgegeben, allerdings werden die Bundesziele von Landeszielen und regionalen Zielsetzungen flankiert. Tabelle 2⁶ katalogisiert nun Schwerpunkte, diesen Schwerpunkte zugeordnete operativen Ziele bzw. die zugeordneten konkreten Maßnahmen. Die Evaluierung eines arbeitsmarktpolitischen Programms wäre nun in Anlehnung an die nachfolgende Tabelle als Evaluierung eines Schwerpunktes des AMS illustrierbar, etwa eine Evaluierung der Tätigkeiten im Rahmen des Schwerpunktes „Verbesserung der Chancen von Frauen am Arbeitsmarkt“. Denkbar wäre auch die Evaluierung eines operativen Ziels in diese Kategorie zu fassen, z.B. „Langzeitarbeitslose in Arbeit bringen“.

Die Erkenntnisse aus Kapitel 5.5 lassen den Schluss zu, dass der Regelfall einer arbeitsmarktpolitischen Evaluierung die Evaluierung eines Projektes ist. Ausnahmen hierfür sind die in den Auswertungen des zitierten Kapitels erwähnte "Evaluierung des Territorialen Beschäftigungspaktes für Wien" oder die „Begleitende Bewertung der Umsetzung des Nationalen Aktionsplanes für Beschäftigung im Jahr 1999“⁷.

⁶ Nicht berücksichtigt werden in dieser Darstellung die Ziele der Arbeitsmarkregionen bzw. der regionalen Geschäftsstellen des AMS.

⁷ Begleitende Evaluierung des Nationalen Aktionsplanes für Beschäftigung in Österreich im Jahr 1999 (1999), Hofer H., Lassnigg L., Leitner A., Prenner P., Steiner M., Wroblewski A. (IHS), Biffel G., Mahringer H. (WIFO), im Auftrag des BMAGS und BMwA.

Tabelle 2: Zielsetzungen des AMS Steiermark 1998

Schwerpunkt	Operatives Ziel	Zugeordnete Maßnahmen
I Verhinderung von dauerhafter Ausgrenzung aus dem Beschäftigungssystem:	1 <i>Langzeitarbeitslosigkeit verhindern</i>	<ul style="list-style-type: none"> • Aktivgruppen • Förderung der regionalen Mobilität • Gründerberatung • Gemeinnützige Arbeitskräfteüberlassung • Schulung Arbeitsloser • Regionale Initiativen und Maßnahmenpakete
	2 <i>Langzeitarbeitslose in Arbeit bringen</i>	<ul style="list-style-type: none"> • Arbeitstrainingszentren • Gemeinnütziger Beschäftigungsprojekte, Beschäftigungsgesellschaften, Gemeinnützige Eingliederungsbeihilfe • Sozialökonomische Betriebe • Geschützte Arbeit • Betriebliche Eingliederungsbeihilfe

Schwerpunkt	Operatives Ziel	Zugeordnete Maßnahmen
II Verbesserung der Chancen von Frauen auf dem Arbeitsmarkt	3 <i>Integration von Frauen mit Betreuungspflichten in den Arbeitsmarkt</i>	<ul style="list-style-type: none"> • Förderung von Kinderbetreuungseinrichtungen • Kinderbetreuungsbeihilfe • Frauenqualifizierungsprojekte
III Optimierung des Arbeitsmarktmatching	4 <i>Einschaltung des AMS am Stellenmarkt sichern</i> Landesziel: Steigerung der Effizienz der Vermittlung 5 <i>Verbesserung der Nutzung des Beschäftigungspotentials im Inland</i>	
IV Optimale Bereitstellung der behördlichen Dienstleistungen	6 <i>Prompte Anweisung von finanziellen Leistungen</i>	
V Prävention von Jugendarbeitslosigkeit	7 <i>Integration v. vorgemerkten Lehrstellensuchenden in die Berufsausbildung sichern</i> 8 <i>Berufseinstieg von Jugendlichen sichern</i>	<ul style="list-style-type: none"> • Berufsvorbereitung • Lehrstellenförderung

5.2.5 Annäherung seitens der Evaluierungsdimensionen

Nachdem weiter oben zwischen den möglichen Gegenständen einer arbeitsmarktpolitischen Evaluierung unterschieden, also eine Abgrenzung hin zur Beschäftigungspolitik und zur passiven Arbeitsmarktpolitik sowie eine Differenzierung Programm – Projekt - Institution versucht wurde, werden nun die zeitlichen und inhaltlichen Rahmen zu definieren sein, innerhalb deren sich diese Evaluierungen aufspannen ließen.

Unberücksichtigt bleibt eine etwaige weitere Unterscheidung nach räumlichen Differenzierungskriterien im Sinne der Stellung des Evaluators zum Evaluationsgegenstand.

5.2.5.1 Zeitliche Evaluierungsdimension

Eine gebräuchliche Einteilung von Evaluierungen setzt am Zeitpunkt ihrer Durchführung – in Bezug zum Programmverlauf – an. Folgende Typen können unterschieden werden:

- **Ex-ante Evaluierungen** sind Evaluierungen, die vor Projektbeginn einsetzen und eine erweiterte „rationality in programmes“ [Schmid] ermöglichen. Für Ex-ante Evaluierungen im arbeitsmarktpolitischen Kontext sind methodische Ansätze wie etwa die systematische Informationsaufbereitung über Größe und Struktur potentieller Zielgruppen und die wahrscheinliche Anzahl von TeilnehmerInnen wichtig; Voraussetzung ist der Zugang zu Methoden der Prognostizierung.
- **Interim Evaluierungen**, die während der Maßnahme bzw. des Projektes durchgeführt werden, bieten die Möglichkeit des direkten Zugangs zu Implementations- und zum Teil Administrationsprozessen. Somit ermöglichen ihre Ergebnisse den Eingriff in den Ablauf einer Maßnahme zur Verbesserung oder zur Nachjustierung.
- **Ex-post Evaluierungen** starten nach Programmende und betrachten die Programmentwicklung, Impacts und Kosten bzw. Nutzen retrospektiv.

Eine zeitliche Trennung zwischen Interims- und Ex-post- Analyse ist nicht immer exakt realisierbar, da viele Projekte kein definiertes Ende haben. Durch die Effekte der Kameralistik hätte vordergründig jedes arbeitsmarktpolitische Projekt ein wohldefiniertes Ende: mit Ende des Kalenderjahres laufen alle arbeitsmarktpolitischen Projekte a priori aus. Da dieses Auslaufen jedoch in der Kameralistik und nicht im Design der Projekte begründet liegt, schießt eine etwaige Diskussion um dieses vordefinierte Ende am Ziel vorbei. Als *ex-post* ist entweder ein Zeitpunkt nach Beendigung eines Projektes anzusehen, *ex-post* ist aber auch die Evaluierung eines Projektes, dessen Ende nicht vorhersehbar ist und dessen Bestandslänge die methodischen Voraussetzungen der Evaluierung erfüllt. Arbeitsmarktpolitische Evaluierungen in Österreich sind in der Regel *ex-post* Evaluierungen.

5.2.5.2 Inhaltliche Evaluierungsdimension

Tabelle 3 beinhaltet allgemeine, nicht projektspezifische Evaluierungsinhalte, die Teil einer Evaluierung sein könnten. Abhängig von den Evaluierungsfunktionen, vom Projekttyp, der methodischen Herangehensweise ist die Programmrealität nicht standardisierbar. Die Inhalte

einer Projektevaluierung werden in Konzept-, Design-, Prozess-, und Wirkungsanalyse unterschieden. In Tabelle 3: Inhalte einer Projektevaluierung und typische Fragestellungen werden diese inhaltlichen Dimensionen dargestellt:

Tabelle 3: Inhalte einer Projektevaluierung und typische Fragestellungen

<i>Konzeptevaluierung</i>
<p><i>Überprüfung der des:</i></p> <ul style="list-style-type: none"> • Problem- und Ursachenbeschreibung • Interventionslegitimation • sozio- techno- ökonomischen Rahmenbedingungen • Ziele • Zielgruppenfestlegung • potentielle Zielkonflikte • Lösungsansatzes und möglicher Alternativen • Zeithorizont • Programmannahmen
<i>Designevaluierung</i>
<p><i>Überprüfung der des:</i></p> <ul style="list-style-type: none"> • operationalen Ziele • Interventionsmodells • Organisatorischen Struktur: Maßnahmen, Programmablauf, -kosten
<i>Prozessevaluierung</i>
<p><i>Überprüfung der Implementierung:</i></p> <ul style="list-style-type: none"> • Programmrealisierung vs. Programmplanung • administrative und operative Effizienz und Effektivität
<i>Wirkungsevaluierung</i>

Überprüfung der/ des:

- Zielerreichung
- direkte und indirekte Effekte
- kurze und lange Frist
- Kausalitäten
- ökonomische, soziale, technologische, wissenschaftliche und Umwelteffekte
- Effizienz

Darstellung folgt Fritz, Hutschenreiter u. Sturn [1997], Seit 11

5.2.6 Zusammenfassung

Der Auswertung aus „Tabelle 5: Ausgewählte Evaluierungsprojekte von AMS Wien und WAFF“ und den nachfolgenden Darstellungen vorgehend kann vorläufig eine Beantwortung jener Frage versucht werden, die eingangs des Kapitels aufgeworfen wurde: Wovon spricht man eigentlich, wenn man über Evaluierung von Arbeitsmarktpolitik in Österreich diskutiert?

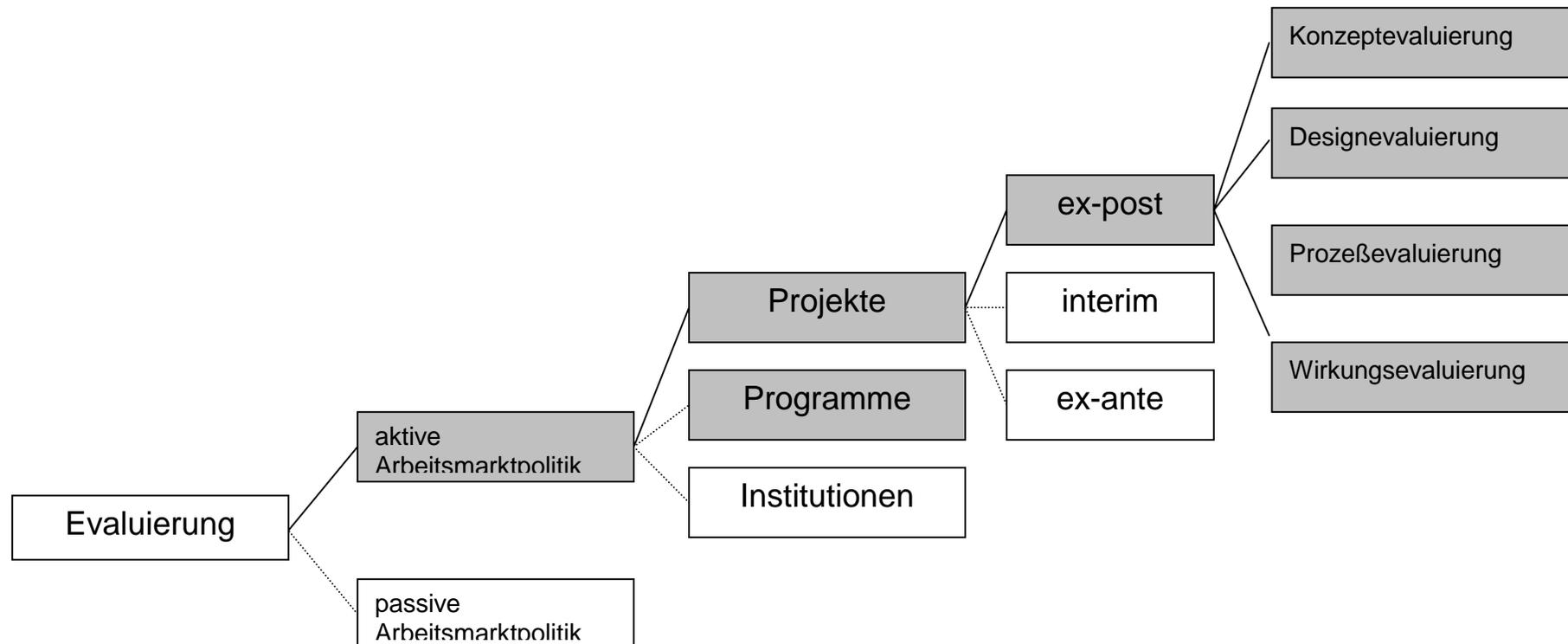
Zur nachfolgenden Darstellung ist *einschränkend* zu erwähnen, dass es natürlich zu Durchbrechungen dieses Schemas kommen kann. Diese Darstellung soll allerdings den *vorherrschenden Typus* von Evaluierungen der Arbeitsmarktpolitik in Österreich darstellen:

Gegenstand einer typischen, österreichischen arbeitsmarktpolitischen Evaluierung ist ein Projekt der aktiven Arbeitsmarktpolitik. Alle möglichen Evaluierungsinhalte werden in einer ex-post Analyse behandelt.

Abbildung 2: Über typische Evaluierungen der Arbeitsmarktpolitik in Österreich



Eigene Darstellung, © InTeReg Wien



5.3 Ein Bezugsrahmen für eine idealtypische Evaluierung der Arbeitsmarktpolitik

Obige Darstellung einer typischen Evaluierung der Arbeitsmarktpolitik in Österreich soll im folgenden nun - in enger Anlehnung an eine Analyse von Günther Schmid et al. [1997⁸] - in einen inhaltlich idealtypischen Evaluierungsrahmen eingebettet werden, um unter Berücksichtigung der vorangegangenen Überlegungen einen möglichst vollständigen Evaluierungsprozess darstellen zu können. Freilich wird eine Umsetzung aller nachfolgender Evaluierungsschritte eher selten zur Anwendung gelangen. Ergänzt werden einzelne Phasen durch Erfahrungen und Fragestellungen, die sich aus der Evaluation des *abzwien*⁹ ableiten. Zentrale Punkte einer Evaluierung von Maßnahmen der (aktiven) Arbeitsmarktpolitik sind jene nach Effektivität und Effizienz: Haben die Ausgaben für Arbeitsmarktpolitik ihre Ziele erreicht? Wurden die erwünschten Effekte mit möglichst geringen Kosten erzielt? Hierin werden sich arbeitsmarktpolitische Projekte nicht von Projekten aus anderen Politikbereichen unterscheiden. Schmid et al. gehen auf dem Weg zur Beantwortung dieser Fragen einen umfassenden Weg und unterscheiden fünf Etappen eines vollständigen Evaluierungsprozesses der Arbeitsmarktpolitik:

Tabelle 4: Etappen der Evaluierung der Arbeitsmarktpolitik und ihre Inhalte

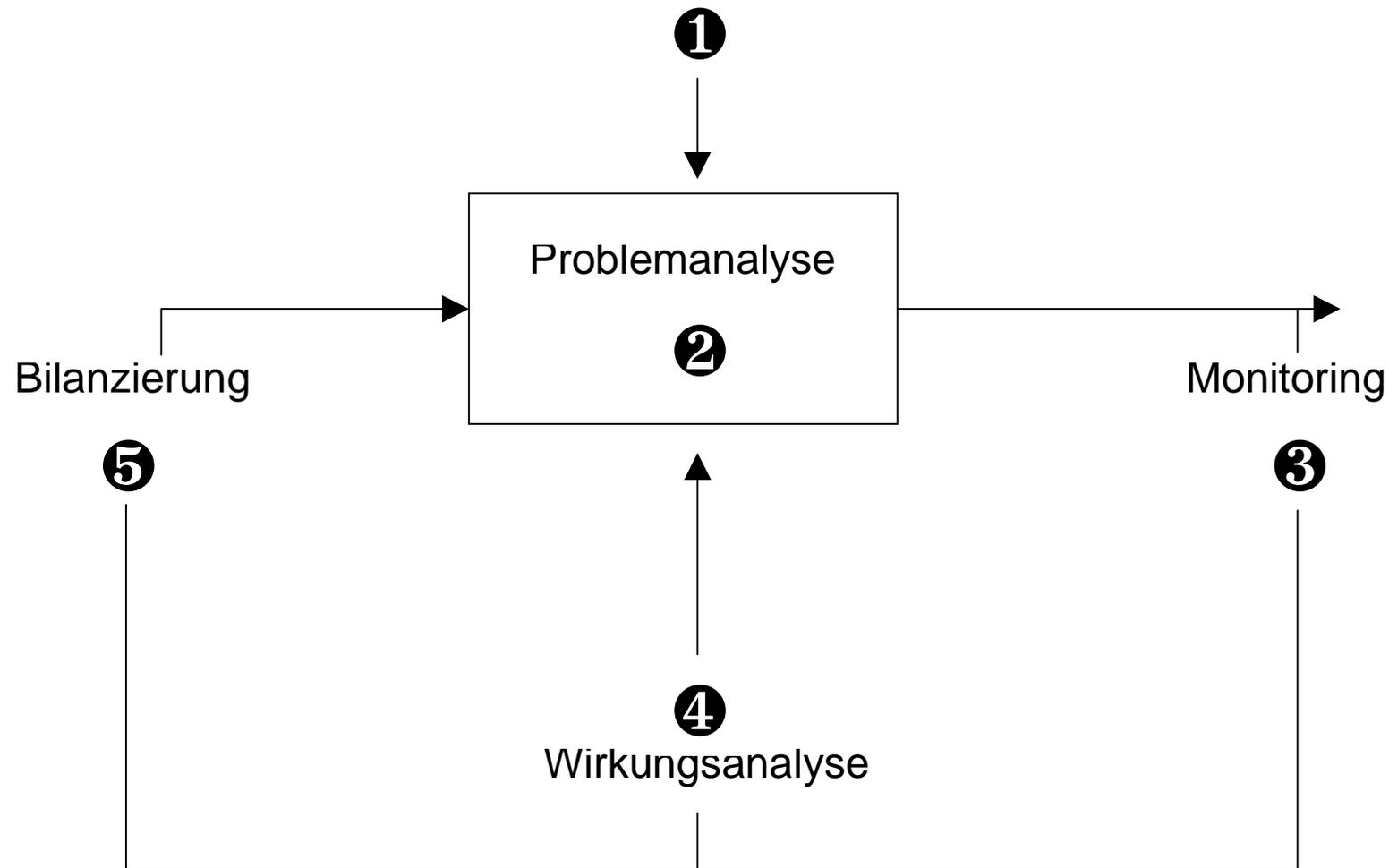
	Etappe	Inhalte
Phase I	Kontextanalyse	Erfassung der Rahmenbedingungen des Arbeitsmarktes und der Arbeitsmarktpolitik
Phase II	Problemanalyse	Präzisierung des arbeitsmarktpolitischen Handlungsbedarfs und der einzusetzenden Instrumente
Phase III	Monitoring	Laufende Kontrolle des Erfolges (Zielkontrolle) während der Umsetzung von Programmen
Phase IV	Wirkungsanalyse	Bestimmung des tatsächlichen Effektes (Outcome) der Arbeitsmarktpolitik auf operativen Zielgrößen
Phase V	Bilanzierung	Messung von Nutzen und Kosten (soweit wie möglich) in monetären Größen und Gegenüberstellung in einer Nutzen-Kosten-Analyse

Quelle: Schmid et al. 1997, Seite 2f.

⁸ G. Schmid/K. Schöman/H. Schütz, Evaluierung der Arbeitsmarktpolitik, WZB Discussion Paper, 1997

⁹ Evaluierung der Joanneum Research im Auftrag des AMS Wien einer arbeitsmarktpolitischen Einrichtung für Frauen in Wien - Konzeption, Programme, Teilnehmerinnen: Wirkungs- und Fiskalanalyse.

Abbildung 3: Etappen der Evaluierung der Arbeitsmarktpolitik und ihre Abfolge
Kontextanalyse



5.3.1 Zur Kontextanalyse

Die Kontextanalyse beinhaltet die Erfassung der Rahmenbedingungen des Arbeitsmarktes und der Arbeitsmarktpolitik. Sie "soll Aufschluss über die zu ergreifenden arbeitsmarktpolitischen Strategien geben. Strategien sind [...] klare Vorstellungen über das zu erreichende Ziel und die geeignete Kombination zu ergreifender Maßnahmen angesichts gegebener Rahmenbedingungen [...] " (Schmid 1997). Von Strategien kann also nur dann gesprochen werden, wenn „Handelnde eine bestimmte Entwicklung nicht völlig beherrschen, ihr aber auch nicht ohne jede Einflussmöglichkeit gegenüberstehen“. Das Ergebnis einer Strategie hängt sowohl ab

- von der Entwicklung der unbeeinflussbaren Faktoren als auch
- vom Einsatz der beeinflussbaren Faktoren.

Eine Kontextanalyse zielt nun vor allem auf die unbeeinflussbaren (externen) Faktoren und deren Entwicklung ab, da auf diese kurzfristig nicht eingewirkt werden kann, sie jedoch die Rahmenbedingungen der Arbeitsmarktpolitik mit beeinflussen. Durch ihre laufende Beobachtung kann man den Akteuren der Arbeitsmarktpolitik jedoch gleichzeitig die Entwicklung von koordinierten Strategien ermöglichen.

Eine Kontextanalyse muss laut Schmid zumindest grob folgende Fragen beantworten¹⁰:

1. Welches sind die externen Rahmenbedingungen, die den Arbeitsmarkt beeinflussen?
2. Welches sind die rechtlichen und institutionellen Rahmenbedingungen, die den Arbeitsmarkt bestimmen?
3. Welche Veränderungsprozesse sind z. Zt. zu beobachten?
4. Mit welchen neuen Rahmenbedingungen ist in absehbarer Zeit zu rechnen?

5.3.1.1 Externe Rahmenbedingungen

Externe Rahmenbedingungen lassen sich in Angebots- und Nachfragebedingungen unterscheiden. Zu den Angebotsbedingungen gehören vor allem Bevölkerungsentwicklung, Wanderungsbilanz, Sozialstruktur der Bevölkerung, Alters- und Qualifikationsstruktur, aber auch Arbeitszeitregelungen oder infrastrukturelle Faktoren wie z. B. die Verfügbarkeit von Kinderbetreuungseinrichtungen; zu den Nachfragebedingungen Wirtschaftswachstum, Produktivität, Strukturwandel, Qualifikationsbedarf, private und öffentliche Investitionen, nationale und internationale Marktentwicklungen.

5.3.1.2 Rechtlich-institutionelle Rahmenbedingungen

Die rechtlichen und institutionellen Rahmenbedingungen in Österreich sind bestimmt durch die Verknüpfung der Arbeitsmarktpolitik mit Wirtschafts- und Sozialpolitik¹¹, der föderalistische Aufbau der Republik und damit verbunden die organisatorische

¹⁰ Einschränkung ist hier zu erwähnen, dass Schmid in seinem Papier nicht eingrenzt, zu welchem Zeitpunkt er bzw. seine Co-Autoren Evaluierung ansetzt – ex post oder ex ante. Da er aber als Ziel der Kontextanalyse angibt, den Akteuren der Arbeitsmarktpolitik ein koordiniertes Vorgehen zu ermöglichen, ist wohl von einem ex ante Ansatz auszugehen.

¹¹ Als institutionelle Entsprechung dieser Verknüpfung möge die Zusammenlegung der Agenden von Wirtschafts- und Arbeitsmarktpolitik in einem Ministerium gelten.

Ausgestaltung des österreichischen Arbeitsmarktservices, budgetäre Rahmenbedingungen (wie sie im Koalitionsabkommen der jeweiligen Regierungen mehr oder minder ablesbar sind) sowie den Europäischen Sozialfonds (ESF) als Co- Finanzierungsinstrument der Europäischen Union und seinen Bestimmungen und Richtlinien.

Die rechtliche Ausgestaltung der Arbeitsmarktpolitik geschieht im wesentlichen durch das AMSG (Arbeitsmarktservicegesetz), das AMFG (Arbeitsmarktförderungsgesetz), das AMPFG (Arbeitsmarktpolitik-Finanzierungsgesetz) sowie das ALVG (Arbeitslosenversicherungsgesetz).

5.3.1.3 Beurteilung der arbeitsmarktpolitischen Strategie

„Am Schluss einer Kontextanalyse wäre die arbeitsmarktpolitische Strategie angesichts der identifizierten Kontextbedingungen zu beurteilen“.¹² Zu den zentralen Fragen in diesem Zusammenhang gehören für Schmid jene nach der Klarheit, der Konsistenz, Komplementarität (Verknüpfbarkeit) sowie der Quantifizier- und Kontrollierbarkeit der Ziele.

5.3.1.4 Fragestellungen bei Evaluation abzwien

1. Welches sind die externen Rahmenbedingungen, die den Arbeitsmarkt beeinflussen?
2. Welches sind die rechtlichen und institutionellen Rahmenbedingungen, die den Arbeitsmarkt bestimmen?
3. Welche Veränderungsprozesse sind z. Zt. zu beobachten?
4. Mit welchen neuen Rahmenbedingungen ist in absehbarer Zeit zu rechnen?

Zur Beantwortung obiger Fragestellungen wurde in einem ersten Schritt die Arbeitsmarktsituation von Frauen in Österreich (und in Wien) dargestellt. Entwicklungen von Erwerbs- und Arbeitslosenquoten, Veränderungen in der Struktur der Erwerbstätigkeit (Zunahme der Erwerbstätigkeit von Frauen im mittleren Alter, Art der Beschäftigungsverhältnisse), Dauer der Arbeitslosigkeit wurden ebenso berücksichtigt wie qualifikatorische Veränderungen oder geänderte wirtschaftliche Bedingungen (z. B. starkes Wachstum von IKT und Multimedia Sektors). Darin spiegeln sich zum Teil rechtliche und institutionelle Rahmenbedingung wider wie auch (beobachtbare) Veränderungsprozesse. Ergänzt wurde diese Phase durch eine Kurzcharakterisierung der aktuellen Arbeitsmarktpolitik in Österreich (und Wien), wobei hier vor allem auch versucht wurde, die Arbeitsmarktpolitik insgesamt zu speziellen Förderprogrammen für Frauen (im Rahmen des Nationalen Aktionsplanes) in Bezug zu setzen, um Aussagen über die Konsistenz mit höher geordneten strategischen Zielen ("Beschäftigung durch Qualifikation sichern") treffen zu können. Darüber hinaus wurde auf Basis von ExpertInneninterviews (MitarbeiterInnen des Arbeitsmarktservices, PolitikerInnen, SozialpartnerInnen, frauenpolitischen AkteurInnen) versucht, frauen- und arbeitsmarktpolitische Handlungsfelder aufzuzeigen, um Informationen darüber zu erhalten, welche Maßnahmen vorrangig zu ergreifen wären, um Ziele wie Chancengleichheit, eine Abnahme weiblicher Arbeitslosigkeit und eine Erhöhung der Frauenbeschäftigung zu erreichen.

¹² Schmid et al. 1997, Seite 13

5.3.2 Zur Problemanalyse

In Tabelle 4 werden die Inhalte einer Problemanalyse als Präzisierung des arbeitsmarktpolitischen Handlungsbedarfs und der einzusetzenden Instrumente zusammengefasst. Fragen wie „Welche Personen sind besonders benachteiligt“ oder „welche Qualifikationen werden nachgefragt“ sollen im Rahmen einer solchen Analyse beantwortet werden. In einem ersten Schritt sollen angebotsseitige, in einem zweiten Schritt nachfrageseitige Überlegungen angestellt werden (ähnlich der Betrachtungen zu den externen Rahmenbedingungen in der Kontextanalyse), in einem dritten Schritt wären die quantitativen und qualitativen Diskrepanzen („mismatch“) zwischen Angebot und Nachfrage festzustellen, in einem vierten die theoretische Beurteilung vorzunehmen, ob Ziel- und Instrumentenmix der Problemstruktur angemessen sind. (Schmid 1997, Seite 15. Nachfolgende Darstellung der Phasen der Problemanalyse folgt der Schmid'schen Darstellung eng)

5.3.2.1 Phase I der Problemanalyse: Angebotsseitige Analyse

Eine Analyse der Probleme der Angebotsseite erstreckt sich vorrangig auf eine Analyse der Arbeitslosigkeit. Zu betrachtende Indikatoren sind Arbeitslosenquoten, Dauer der Arbeitslosigkeit, Mehrfacharbeitslosigkeit, Zugangsquoten der Arbeitslosigkeit und nach Herkunftsort¹³, Abgangsquoten der Arbeitslosigkeit¹⁴, jeweils nach soziodemographischen Gesichtspunkten gegliedert¹⁵. Veränderungsdaten dieser Indikatoren können als Frühwarnsystem herangezogen werden.

5.3.2.2 Phase II der Problemanalyse: Nachfrageseitige Analyse

Die Nachfrageseite ist vorrangig über eine Analyse von offenen Stellen ergründbar: Offene Stellen nach Sektoren, Berufen und Qualifikationsmerkmalen, Dauer der offenen Stellen und Zugangsquoten nach Herkunftsort (Sektor oder Beruf, Teilzeit) bzw. nach unternehmensdemographischen Gesichtspunkten.

Registrierte offene Stellen decken jedoch nicht das gesamte Spektrum der Nachfrage ab. Zusätzlich wären Wachstumsraten der Beschäftigung nach Tätigkeitsmerkmalen, Qualifikation bzw. Berufen zu untersuchen; Analyse der Stellenanzeigen, regelmäßige Befragungen bzw. Panels bei Betrieben wäre ein Mittel, diese Frage zu beantworten.

5.3.2.3 Phase III der Problemanalyse: Mismatch-Analyse

Mismatch bezeichnet qualifikationsbedingte Unterschiede zwischen Arbeitsangebot und Arbeitsnachfrage. Mittels einer Gegenüberstellung der Zahl offener Stellen zur Zahl der Arbeitslosen kann ein erster Indikator für das Ausmaß des Mismatch herangezogen werden. Jedoch ist dieser Indikator mit zwei Schwächen behaftet, die das Ergebnis deutlich verzerren: Zum einen entspricht die Zahl der gemeldeten offenen Stellen häufig nicht dem tatsächlichen Stellenpotential und zum anderen unterschätzt die Zahl der registrierten Arbeitslosen das tatsächlich zur Verfügung stehende Arbeitsangebot.

¹³ Schmid et al. [1997] nennen Sektor oder Beruf, Berufsrückkehr, Erstausbildung aus Teil- oder Vollzeitbeschäftigung, befristeter / unbefristeter Beschäftigung, aus einer arbeitsmarktpolitischen Maßnahme.

¹⁴ Schmid et al. [1997] nennen Teil- oder Vollzeitbeschäftigung, arbeitsmarktpolitische Maßnahme, längere Krankheit oder Inaktivität, Rente, Bildungssystem.

¹⁵ Schmid et al. [1997] nennen Alter, Geschlecht, ethnische Herkunft, Familienstatus (z.B. ob im Haushalt noch andere Einkommensbezieher vorhanden sind)

Da dieser Indikator also nicht ausreichend ist, um das in bestimmten Regionen existierende Ungleichgewicht von angebotener und nachgefragter Qualifikation zu messen, empfehlen Schmid et al. die Erfahrungen mit vierteljährlichen Befragungen von Betrieben und Erwerbspersonen in Holland im Sinne eines 'best practice case' genauer zu studieren und für die jeweiligen regionalen Verhältnisse auszuwerten¹⁶.

5.3.2.4 Phase IV der Problemanalyse: Programmbewertung

Den Abschluss sollte eine Feinabstimmung zwischen Programm- und Problemstruktur bilden. Schmid et al. fassen die zentralen Fragen dieser Phase wie folgt zusammen:

1. Ist nach Kenntnis der Problemlage und theoretischer Zusammenhänge die Programmstruktur im einzelnen mit der Problemstruktur kongruent?
2. Werden die richtigen Zielgruppen anvisiert?
3. Sind die verfolgten Ziele im einzelnen konsistent?
4. Sind sie klar formuliert?
5. Passen die Instrumente zu den Zielen?
6. Ist die Organisation der Arbeitsmarktpolitik in der Lage, das Programm effektiv umzusetzen?
7. Sind genügend Mittel vorgesehen?

5.3.3 Zum Monitoring

Monitoring von Arbeitsmarktpolitik wurde in den letzten Jahren immer wichtiger; Gründe hierfür sind :

- Budgetbeschränkungen, die Forderung nach Lean Management und die Reform des öffentlichen Sektors im Rahmen eines New Public Management.
- Dezentralisierungstendenzen in der Arbeitsmarktpolitik. Durch die Abgabe von Verantwortlichkeiten an niedrigere Organisationsebenen (im Falle des AMS von BGS an LGS bzw. von LGS an RGS) wird das Monitoring-System einer Organisation zunehmend komplex. Dezentralisierung bewirkt einerseits, dass Informationen dort gebündelt werden können wo sie auch wirklich benötigt werden, andererseits versetzt sie die zentrale(re) Organisationseinheit in die Lage, Informationen über ihre untergeordneten Abteilungen zu sammeln und diese zu vergleichen.
- Anforderungen der EU - Kommission seit 1994 bzw. im Rahmen der ESF- Co-Finanzierung.
- (Technische) Vereinfachung von, beispielsweise Datenaustausch, hat die Implementation von Monitoringsystemen vereinfacht und somit gefördert.

Ein idealtypisches Monitoring-System¹⁷ bestünde aus folgenden Elementen:

- Zielvorgaben hinsichtlich Input, Output und Outcome, die in einem bottom-up Prozedere von den politischen und administrativen Entscheidungsträgern auf nationaler, regionaler und lokaler Ebene vorgegeben werden;
- Die Definition von Indikatoren, die es erlauben, Ziele zu messen;

¹⁶ Schmid et. al, 1997, S. 21

¹⁷ Nach Peter Auer in WZB: "Progress in Evaluation Research", Seite 7

- Die Hinzuziehung von statistischen Informationen zur Abrundung dieser Zielvorgaben und Indikatoren; und
- Feedback-Schleifen, die garantieren, dass beobachteten Abweichungen adäquat entgegen getreten wird.

Ein optimales Design eines Monitorings von Arbeitsmarktpolitik beginnt mit der Bestimmung von Politikfeldern ("policy formation"). Schon in diesem frühen Stadium betont Auer (ebenda) die Bedeutung der klaren Zielformulierung für die effektive Einführung von Monitoringsystemen. "[...] *effective monitoring must be written into laws and ordinances*". Nach der Implementierung eines arbeitsmarktpolitischen Programms, sollte augenblicklich ein kontinuierlicher Prozess der Beobachtung beginnen, wobei die Intervalle zwischen den Beobachtungszeiträumen von Maßnahme zu Maßnahme variieren können. Auer bezeichnet ein vierteljährliches Beobachtungsintervall als guten Kompromiss, wobei für kurzfristige Programme diese Intervalle durchaus kürzer sein können.

"The core function of monitoring is to detect indicators turning "red" and to initiate subsequent remedial action (feedback)." Das Feedback zwischen Monitoring und der Programmimplementierung ist stärker als zwischen Monitoring und der Bestimmung von Politikfeldern, in denen die Programmmaßnahmen zu Ergebnissen führen sollen. Allerdings sollten Monitoring-Ergebnisse, die Standardergebnisse der Maßnahmen widerspiegeln, an die Politikebene weitergegeben werden. Auer: "Evaluation, in which deadweight, substitution effects and other factors are also taken into consideration, offers a more profound assesment of programme impacts than monitoring, and thus affects policy formation more strongly than monitoring,"

Auer bezeichnet als Mindestanforderung an ein einfaches Monitoring-System folgende Inhalte:

- Gegenüberstellung von geplanten Budget und tatsächlich ausgeschütteter Mittel (nach Verwendung, Zielgruppen o. ä.),
- Geplante Teilnahme und tatsächliche Teilnahme,
- Kosten pro TeilnehmerIn auf Basis der bereits erwähnten Punkte.

Unter diesen Voraussetzungen, so Auer, würden Monitoring-Systeme nützliche Hinweise über die laufende Performance der Beobachteten geben, gleichzeitig könnten "under performers" per Vergleich entdeckt werden.

5.3.4 Zur Wirkungsanalyse

„[...] Die bisherige [arbeitsmarktpolitische] Evaluierungspraxis hat es fast vollständig versäumt, eine Vergleichssituation zu konstruieren, die hinreichend genau beschreibt, welcher Zustand ohne den zu evaluierenden Eingriff eingestellt hätte.“¹⁸

¹⁸ Christoph M Schmidt in DIW [2000] , Seite 425

In Tabelle 4 wurden die Inhalte der Wirkungsanalyse als „Bestimmung des tatsächlichen Effektes (Outcome) der Arbeitsmarktpolitik auf operativen Zielgrößen“ umrissen. Eine solche Analyse stellt also den Outcome einer Maßnahme in Vergleich

- zu den gesetzten Zielen und
- zu jenem Ergebnis, das ohne Maßnahme (vermutlich) zustande kommen wäre.

Eine solche Gegenüberstellung hat also sowohl den Charakter einer Messung, als auch jenen einer Schätzung. Eine vollständige Wirkungsanalyse zieht auch nichtintendierte Effekte in Betracht, die sowohl positiv als auch negativ sein können. Die hypothetische Aussage darüber, was passiert wäre, wenn das Programm/die Maßnahme nicht durchgeführt worden wäre, wird auch als die kontrafaktische Situation bezeichnet. Sie entzieht sich jedoch aufgrund ihrer Unbeobachtbarkeit einer Messung im üblichen Sinn und muss aus anderweitig vorhandenen Informationen konstruiert werden (5.3.4.2). „Die Analyse des Kontrafaktischen ist [daher] die eigentliche Kunst wissenschaftlicher Evaluierung¹⁹.

5.3.4.1 Was misst Wirkungsanalyse?

Wirkungsanalysen sollen in erster Linie Effektivität messen. Schmid et al. [1997] bezeichnen eine Maßnahme dann als effektiv, wenn ihre Wirkung abzüglich der negativen nichtintendierten Wirkungen größer ist als das Ergebnis ohne Maßnahme. Aktive Arbeitsmarktpolitik strebt vor allem die (Wieder-)Eingliederung von Arbeitslosen und zwar insbesondere der "Problemgruppen" (Langzeitarbeitslose, Ältere, Geringqualifizierte) in den Arbeitsmarkt an. Kann die Arbeitsmarktsituation der TeilnehmerInnen an einer Maßnahme der aktiven Arbeitsmarktpolitik tatsächlich verbessert werden und sind die damit verbundenen nichtintendierten negativen Effekte nicht zu stark (Verdrängungseffekte, Mitnahmeeffekte, Substitutionseffekte), so wäre eine einzelne Maßnahme bzw. ein Programm als effektiv zu bezeichnen.

Um jedoch explizite Aussagen zum Zielerreichungsgrad (Welcherart ist die Verbesserung der Arbeitsmarktsituation) treffen zu können, ist die Festlegung konkreter Bewertungs- bzw. Erfolgskriterien, also spezifischer, messbarer und zeitlich fixierter Ziele eine Voraussetzung. Diese sollten sich wiederum aus den Zielen der einzelnen Instrumente der Arbeitsmarktpolitik ergeben²⁰; im idealtypischen Fall sich also aus Kontext- und Problemanalyse ableiten lassen. In der Folge stellt sich somit auch die Frage nach der Abwägung verschiedener Ziele und sich daraus ergebener Zielkonflikte bzw. konkurrierender Ziele. Beispielsweise kann das Ziel der Rückführung in Beschäftigung vor allem bei Wiedereinsteigerinnen durchaus mit dem (höher geordneten) Ziel des Arbeitslosenabbaus konfliktieren. Werden Frauen, die Out of Labour Force waren in die Beschäftigung rückgeführt und dort mit dem durchschnittlichen Risiko der Arbeitslosigkeit konfrontiert, dann wird dies – unter sonst gleichbleibenden Bedingungen - die Arbeitslosenquoten erhöhen; einfach weil erwerbstätige Frauen im Durchschnitt öfter und länger arbeitslos sind als Männer.

¹⁹ Schmid et al. [1997] , Seite 31

²⁰ Viktor Steiner, 2000

Darüber hinaus ergibt sich auch ein Konflikt aus den Forderungen nach besonderer Förderung der Zielgruppen am Arbeitsmarkt einerseits und dem Anspruch auf Qualität, auf Leistung, auf geringe Abbruchzahlen bei den Maßnahmen, auf hohe Eingliederung in Arbeit und möglichst unbefristete und dauerhafte sozialversicherungspflichtige Beschäftigung andererseits. Beide Ziele sind nicht gleichzeitig erreichbar, denn je mehr Risikomerkmale eine Person auf sich vereinigt, umso schlechter sind ihre Chancen auf Rückführung in Beschäftigung.

Bei der Festlegung von Bewertungskriterien für den Erfolg bzw. die Effektivität einer Maßnahme bedarf es daher einer genauen Abwägung verschiedener Ziele für die jeweiligen Maßnahmen. So detailliert das Zielsystem des AMS auch festgeschrieben sein mag, so wenig ermöglicht es die gleichzeitige Integration unterschiedlicher Zielebenen, die etwa auch aus der Vermischung von maßnahmen- und programmspezifischen Zielen resultieren können.

5.3.4.1.1 Bewertungs- und Erfolgskriterien

Welche Beobachtungseinheiten und -zielgrößen relevant sind, wird durch die Art der Maßnahme – beispielsweise auf welchem Aggregationsniveau sie eingerichtet ist - und die Datenlage bestimmt. Darüber hinaus gibt es natürlich noch unterschiedliche Vorstellungen darüber, was mit aktiver Arbeitsmarktpolitik erreicht werden kann. In den meisten Evaluationsausschreibungen des Arbeitsmarktservice wird vorab aus der Verbindung von Zielen mit Indikatoren eine entsprechende Auswahl von Erfolgskriterien getroffen. Implizit oder explizit ist dabei immer von Kausalitäten die Rede²¹, zum Beispiel wenn als Ziel einer Qualifizierungsmaßnahme die Aufnahme eines regulären, dauerhaften Beschäftigungsverhältnisses formuliert wird. Die tatsächliche Aufnahme einer Beschäftigung wird dann als der "kausale Effekt der Maßnahmenteilnahme" gesehen, der durch eine "ökonomische Analyse des Erfolges von Maßnahmen nach deren Ende festgestellt wird" (Fitzenberger, 1999, Seite 1).

Für die EvaluatorInnen stellt sich nun einerseits die Frage nach dem Zeitpunkt der Evaluation und zum anderen danach, ob die Beschränkung auf das Beschäftigungsziel nicht zu kurz greift. Die Teilnahme an Qualifizierung führt nicht nur zu einer Aufstockung des Humankapitals, sie kann auch die Möglichkeit bzw. den Zugang (Information) zu einer zielgerichteten Weiterbildung schaffen, die den jeweiligen Personen wesentlich bessere Einstiegschancen eröffnet, da unter Umständen der Mismatch zwischen Arbeitsplatzanforderungen und dem Qualifikationsprofil der Arbeitslosen verringert werden kann²². Da sich hier jedoch nur schwer eindeutige (und kausale) Verbindungen zur Maßnahme herstellen lassen, bleiben diese flankierenden Aspekte oft unberücksichtigt. Darüber hinaus würde es bedeuten, dass man einen sehr langen Nachbeobachtungszeitraum wählen müsste²³, in dem es möglich ist, den Karriereverlauf der Teilnehmer genau zu verfolgen. Dies scheitert jedoch meist an der Verfügbarkeit entsprechender Daten.

²¹ Blaschke/Plath, 2000, Seite 464

²² Blaschke/Plath, 2000, Seite 465

²³ Denkbar wäre eine fortlaufende Beobachtung derselben Person im Rahmen von Paneluntersuchungen

Als Indikatoren für den Erfolg einer Maßnahmen werden daher sehr häufig solche herangezogen, die mit verfügbaren Datenmaterial möglich und messbar sind. Dazu zählen neben den bereits angeführten Beschäftigungseffekten auch Einkommenseffekte (Kann nach der Maßnahmenteilnahme ein höheres Einkommen erzielt werden), die Dauerhaftigkeit und Stabilität der Beschäftigung sowie daraus resultierend auch die Verkürzung von Arbeitslosigkeitszeiten (Beschäftigungsdauer nach der Maßnahme, Verringerung der Diskontinuitäten im Erwerbsverlauf) aber auch psychosoziale Effekte wie die Steigerung des Selbstvertrauens, die Stabilisierung der Lebensumstände (geregeltes Einkommen sowie Arbeitsalltag), sozialer Aufstieg. Letztere entziehen sich jedoch einer monetären Bewertung, sollten jedoch als Ergänzung der quantitativen Kriterien in jedem Fall berücksichtigt werden. Zu überlegen wäre weiters, ob man durch Interviews oder schriftliche Befragung versuchen soll, nicht-monetäre Arbeitsbedingungen (Arbeitsplatzzufriedenheit, beruflicher Aufstieg etc.) zu bestimmen.

Aus dem Bruttobeschäftigungseffekt einer Maßnahme kann jedoch nicht geschlossen werden, dass diese auch zur Gänze einen positiven Nettoeffekt bewirken, da mit einer Maßnahme auch nicht intendierte wirkungsmindernde Effekte (Verdrängungs-, Substitutions- und Mitnahmeeffekte) verbunden sein können.

Mitnahmeeffekte ergeben sich dann, wenn Arbeitslose an Maßnahmen teilnehmen, bei denen sich vermuten lässt, dass sie auch ohne diese eine neue Beschäftigung gefunden hätten oder wenn Unternehmen eine Subvention für die Beschäftigung eines Jugendlichen erhalten, den sie auch ohne staatlicher Transferleistung eingestellt hätten. Gleichzeitig bedeutet dies auch, dass ein Teil der Ausgaben für diese Maßnahme keinerlei Auswirkung auf die angestrebten Erfolgskriterien hat, der Nettobeschäftigungseffekt der Maßnahme also geringer ist, als der eigentliche Bruttoeffekt.

Substitutionseffekt bezeichnet die Möglichkeit, dass Arbeitgeber ihre Arbeitnehmer aufgrund "einer Änderung relativer Lohn- bzw. Lohnstückkosten gegen staatlich geförderte Arbeitslose austauschen" (Steiner, 2000, S. 39). Die reguläre nicht-geförderte Beschäftigung sinkt und es ist somit keine Auswirkung auf die Gesamtbeschäftigtenzahl feststellbar. Schmid et al. 1997 führen darüber hinaus auch noch die Möglichkeit der Programmsubstitution an: Durch die Gleichzeitigkeit von Programmen, die gleiche oder ähnliche Wirkungen haben, kann die Erweiterung nur eines dieser Programme ein anderes substituieren, wenn dieses nicht auch ebenso erweitert oder gar zurückgenommen wird. Tatsächliche (Beschäftigungs-)Effekte werden nur dann wirksam, wenn das substituierte Programm weniger effektiv war. Ansonsten verändern sich nur die relativen Eingliederungserfolge der Programme zueinander, die Gesamtbeschäftigung würde jedoch keine Erhöhung erfahren.

Verdrängungseffekte werden typischerweise im Zusammenhang mit deren Wirkungsmechanismus auf die Gütermärkte diskutiert. Beispielsweise können Unternehmen mit subventionierter Beschäftigung ihren Marktanteil zu Lasten von Unternehmen, deren Beschäftigung nicht staatlich gefördert wird, ausweiten und somit Wettbewerbsvorteile erhalten. Dies führt jedoch zu einer Verdrängung regulärer Arbeitsplätze, zum anderen "können sogar private Investitionen oder Unternehmensgründungen verhindert werden, was langfristige negative Auswirkungen auf die reguläre Beschäftigung zur Folge hätte" (Steiner, 2000, Seite 40).

Empirisch können sowohl Verdrängungs- als auch Substitutionseffekte mittels **aggregierter** Wirkungsanalysen aufgedeckt werden²⁴, indem überprüft wird, ob die Ausgaben für die aktive Arbeitsmarktpolitik Auswirkungen haben sowohl auf die Abgänge aus Arbeitslosigkeit in Beschäftigung als auch auf die aus Beschäftigung in Arbeitslosigkeit. Auf der Maßnahmenebene sind solche Effekte nur sehr schwer messbar und seriöserweise auch kaum mittels einer Befragung feststellbar.

Neben diesen wirkungsmindernden Faktoren kann die Wirkung einer Maßnahme durch *Multiplikatoreffekte* auch gesteigert werden. Schmid et al. unterscheiden hier zwischen direkten und indirekten Effekten. Direkte Multiplikatoreffekte entstehen durch Zusatzmaterialien zur Durchführung einer Maßnahme, beispielsweise wenn durch den Aufbau eines Beschäftigungsprojektes Maschinen oder Geräte (EDV-Anlagen, Büroeinrichtung...) für die Herstellung bzw. Entstehung bestimmter Produkte benötigt werden. Indirekte Multiplikatoreffekte können sich durch Preisvorteile oder Wertsteigerungen ergeben, die den Standort für die Ansiedlung von anderen Betrieben attraktiver machen (Bereitstellung spezifischer Qualifikationen z. B. für IKT-Betriebe).

Für die konkrete Evaluierung stellt sich somit die Frage, welche Verfahren für den Nachweis der Effektivität einer Maßnahmen am geeignetsten erscheinen, dies vor allem auch unter Berücksichtigung der Generierung und Aufbereitung bzw. überhaupt der Verfügbarkeit von entsprechenden Daten und somit der Machbarkeit bestimmter Analysemethoden.

5.3.4.2 Verfahren und Methoden der Wirkungsanalyse

Das Ziel der Wirkungsanalyse ist, wie oben ausgeführt, der Vergleich der tatsächlichen mit der kontrafaktischen Situation. Die Voraussetzung für einen seriösen Vergleich ist, eine Vergleichssituation zu schaffen, die sich idealerweise lediglich in einem Umstand von der beobachteten Situation unterscheiden sollte: durch die hypothetische Aussage darüber, was mit der Person passiert wäre, wenn die Maßnahme nicht durchgeführt worden wäre. Die kontrafaktische Situation kann jedoch durch rein betriebswirtschaftliche Methoden nicht dargestellt werden; Sie kann, so Schmidt [2000, Seite 429] , „schlichtweg nicht beobachtet werden.[...] Dieses Phänomen wird daher in der Literatur als das fundamentale Evaluierungsproblem bezeichnet.“ Mittels sozialen Experimenten sowie mikroökonomischen Evaluationsstudien wird daher versucht, diese schwierige Problematik des eigentlich Unbeobachtbaren zu lösen oder damit zumindest fundiertere Aussagen über die Wirkungen von Maßnahmen und Programmen treffen zu können.

Die Konstruktion der hypothetischen Vergleichssituation, die für die Bewertung einer Maßnahme unabdingbar erforderlich ist, kann dabei auf unterschiedliche Weise erfolgen. Eine Möglichkeit dazu ist der Aufbau eines **Experiments** (ähnlich der Bewährungskontrolle bei Medikamenten), in der ein Zufallsverfahren über die Teilnahme an der Maßnahme entscheidet. Unter idealen Bedingungen wird das Selektionsproblem durch die Randomisierung gelöst. Die Wirkung einer Maßnahme ist dann einfach das Ergebnis des "Treatments" durch die Maßnahme abzüglich des Ergebnisses ohne Programm. Solche soziale Experimente setzen jedoch eine Reihe von Bedingungen voraus, die in der Realität

²⁴ Steiner, 2000

nicht gegeben sind. Ist es mittels Laborexperimenten möglich, künstliche Wirklichkeiten zu schaffen, in denen jeder Teilnehmer den gleichen Einflussgrößen ausgesetzt ist, so ist dies bei sozialwissenschaftlichen Experimenten nicht möglich. Arbeitslose sind nicht identisch, sie verfügen beispielsweise nicht über gleiche Lernvoraussetzungen oder Fähigkeiten. Ohne Annahmen über unbeobachtete Einflussfaktoren, die sowohl mit dem Maßnahmenerfolg als auch mit der Maßnahme selbst korreliert sind, können kausale Effekte nur schwierig identifiziert werden²⁵. Darüber hinaus werden indirekte Effekte, die sich aus Substitutionsbeziehungen zwischen geförderten MaßnahmenteilnehmerInnen und NichtteilnehmerInnen ergeben (siehe 5.3.4.1.1), ausgeschlossen.

Sozialwissenschaftliche Experimente nehmen vor allem in den USA eine dominante konzeptionelle Rolle ein, als diejenige Evaluationsstrategie, die in der Lage ist, die glaubwürdigsten und stabilsten Ergebnisse zu erzielen²⁶. Beispielsweise wurde das wichtigste Ausbildungsprogramm der 80er Jahre, der Job Training Partnership Act (JTPA), mit einem groß angelegten sozialwissenschaftlichen Experiment evaluiert. Doch auch in Europa werden zunehmend methodisch anspruchsvolle Verfahren der Wirkungsanalyse, sog. nicht-experimentelle mikroökonomische Ansätze, eingesetzt.

5.3.4.2.1 Mikroökonomische Evaluationsstudien

Für die Evaluation des Maßnahmeerfolgs genügt es nicht, nur die geförderten Teilnehmer nach einer bestimmten Zeitspanne auf ihren Beschäftigungsstatus zu untersuchen, da diese möglicherweise auch ohne Förderung eine Beschäftigung gefunden hätten. Das zentrale Konzept stellt dabei die *Vergleichsgruppe* dar, die sich von der Teilnehmergruppe nur durch die Nichtteilnahme an einer bestimmten Maßnahme unterscheidet, sonst aber bezüglich der Merkmale, die den Erfolg einer Maßnahme bestimmen, identisch ist²⁷.

Die Grundidee bei mikroökonomischen Evaluationsstudien besteht also darin, zu der Gruppe der Teilnehmer eine Gruppe von vergleichbaren Nichtteilnehmern, die *Vergleichs- oder „Kontrollgruppe“*, zu bilden und das Erfolgskriterium in den beiden Gruppen zu vergleichen. Dabei soll die Vergleichsgruppe bezüglich aller Merkmale, die den Erfolg einer Maßnahme beeinflussen, möglichst ähnlich sein. Eine vollständige Vergleichbarkeit ist insofern logisch nicht möglich, als man die identischen Personen niemals gleichzeitig in der Teilnehmer- und in der Vergleichsgruppe beobachten kann²⁸. Die Evaluationsmethoden unterscheiden sich daher vor allem danach, wie eine Vergleichsgruppe auf Basis nicht-experimenteller Daten abgebildet wird bzw. welche Annahmen zur Lösung des Selektionsproblems getroffen werden. Selektionsprobleme treten dann auf, wenn es sich bei der Gruppe der Teilnehmer nicht um eine rein zufällig ausgewählte Gruppe handelt, sondern die Teilnahme an einer Maßnahme mit beobachteten und in den Daten unbeobachteten Faktoren korreliert ist.

5.3.4.2.2 Matching Pairs Analyse

Die Wirkungen einer Politikmaßnahme wären, wie bereits ausgeführt, am besten dokumentiert, wenn die Zustände oder Outcomes der Betroffenen nach dem Einsatz dieser

²⁵ Steiner, 2000, Seite 43

²⁶ IZA Compact, Dezember 2000

²⁷ Steiner, 2000, S. 44

²⁸ Steiner, 2000 ebd.

Maßnahme den Zuständen oder Outcomes der selben Objekte ohne Maßnahme gegenübergestellt würden. Ein derartiger Analyseablauf würde bedeuten, dass zum Beispiel Betriebe für bestimmte Aktivitäten eine öffentliche Förderung erhalten und nach der Dokumentation der Ergebnisse in einem neuen Anlauf die selben Aktivitäten ohne öffentliche Unterstützung durchführen müssten. Die Differenz der Ergebnisse in beiden Fällen wäre dann die direkt messbare Wirkung der Förderung. Eine derartige Evaluierung ist im Feld der Sozialwissenschaften kaum vorstellbar.

Um im Evaluierungsverfahren dennoch Aussagen treffen zu können, die in ihrer Präzision über die herkömmlichen Kontrollgruppenvergleiche hinausgehen, kann ein „Matching Pairs“-Verfahren angewendet werden. Eine „Matching Pairs“-Analyse ist eine in den USA sehr verbreitete Evaluierungsmethode²⁹, die vor allem im Bereich der Arbeitsmarktforschung zur Anwendung kommt. Auch im EU-Raum wurden bereits Programmevaluierungen mit diesem Ansatz³⁰ durchgeführt. Bei Evaluierungen im Bereich der Arbeitsmarktforschung besteht zumeist allerdings der Vorteil, dass beim Matchen auch die zeitliche Struktur der Variablen berücksichtigt werden kann, da über Sozialversicherungen oder ähnliche Datenquellen Art und Umfang von Beschäftigungsverhältnissen im Zeitverlauf klar nachverfolgt werden können. Für Betriebe oder Unternehmen existieren leider keine vergleichbaren Datenbanken und in einer Primärbefragung (wie im vorliegenden Fall) können aus erhebungstechnischen Gründen nur eingeschränkt Daten erfasst werden, die Rückschlüsse auf die Dynamik bestimmter relevanter Variablen erlauben.

Die Grundidee dieser Methode ist folgende: Geförderte und nicht-geförderte Einheiten (wie Personen oder Betriebe), die einander in relevanten Strukturmerkmalen so ähnlich wie möglich sind, werden miteinander verglichen. Die Relevanz der Merkmale wird dabei hinsichtlich ihres Einflusses auf die Teilnahme am Förderprogramm sowie auf die vom Programm intendierten Ergebnisse definiert. Damit soll die Signifikanz des Vergleichs gestärkt werden. Einerseits werden nur Einheiten verglichen, die zur Zielgruppe des Programms zählen. Bei der Evaluierung eines Qualifikationsprogramms etwa sollten nur Personen verglichen werden, für die eine derartige Maßnahme Sinn ergibt; im Falle eines Technologiediffusionsprogramms kommen nur Betriebe als Vergleichseinheiten in Frage, bei denen die betreffende Technologie auch zur Anwendung kommen könnte.

Andererseits werden jene Ergebnisse, die als Indikatoren für den Erfolg eines Programms verwendet werden, auch von anderen Faktoren als der Förderung beeinflusst. Um die Wahrscheinlichkeit eines systematischen Unterschieds in diesen Merkmalen zwischen den Programmteilnehmern und der Kontrollgruppe zu minimieren, wird die Vergleichsgruppe auch nach ihrer Ähnlichkeit hinsichtlich dieser Faktoren ausgewählt. Es wäre also kein zulässiges Evaluierungsdesign, wenn etwa langzeitarbeitslose Alkoholiker in ländlichen Krisenregionen mit arbeitslosen Jugendlichen in urbanen Ballungsgebieten hinsichtlich ihrer Vermittlungschancen verglichen würden.

²⁹ Vgl. Heckman, Smith (1996)

³⁰ Ein interregionaler Vergleich der Wettbewerbsfähigkeit irischer und englischer KMUs durch O'Farrel, Hitchens (1989) sowie eine Untersuchung aktiver Arbeitsmarktpolitik in Rheinland-Pfalz durch Almus, Egel, Lechner, Pfeiffer, Spengler (1998). In Österreich kam ein vergleichbarer Evaluierungsansatz auch in Zilian, Lassnigg, Wroblewski (1999) zur Anwendung

Jedem Teilnehmer wird das Element der Vergleichsgruppe zugeordnet, das sich am wenigsten von ihm unterscheidet; beide ergeben somit ein „matched pair“. In der weiteren Analyse kann dieses Paar so betrachtet werden, als handle es sich um ein und dieselbe Einheit, einmal mit Förderung und einmal ohne. In diesem Sinn wird auch von einem „quasi-experimentellen“ Evaluierungsverfahren gesprochen.

5.3.4.2.3 Vorher-Nachher Vergleich

Trotz vieler Schwächen ist der Vorher-Nachher Vergleich in der Praxis ein sehr häufiges Kriterium zur Beurteilung des Erfolgs arbeitsmarktpolitischer Maßnahmen. Diese Methode zeichnet sich zum einen durch ihre Einfachheit aus und zum anderen kann sie ohne einer direkten Vergleichsgruppe auf der Basis einer Stichprobe von Teilnehmern an einer Maßnahme durchgeführt werden. Dafür ist es nicht unbedingt notwendig, über Individualdaten zu verfügen, das Verfahren könnte auch mittels gruppierter Daten von Teilnehmern zur Anwendung gelangen.

Liegen für jede Person oder eine hinreichend große Anzahl von Personen in der Teilnehmergruppe mindestens eine Beobachtung vor der Teilnahme und eine Beobachtung nach der Teilnahme an einer Maßnahme vor, kann deren kausaler Effekt durch die mittlere Beschäftigungsquote der Teilnehmer vor und nach der Maßnahme geschätzt werden³¹. Dies setzt die Annahme voraus, dass die Beschäftigungsquote der Gruppe der Teilnehmer vor der Maßnahme im Mittel der hypothetischen Beschäftigungsquote dieser Gruppe nach der Maßnahme entspricht, sofern diese Gruppe an der Maßnahme nicht teilgenommen hätte. Als Vergleichsgruppe fungiert bei dieser Methode also die Teilnehmergruppe selbst, allerdings bezogen auf die Periode vor Beginn der Maßnahme. Die implizite Annahme dabei ist, dass die Beschäftigungsquote vor der Teilnahme durch diese nicht beeinflusst wird. Damit diese Annahme erfüllt sein kann, darf die Beschäftigungsquote der Teilnehmergruppe nicht durch Faktoren beeinflusst werden, die über die Zeit variieren. Sofern diese Zeiteffekte durch beobachtete Variable, wie beispielsweise die Zahl der offenen Stellen, konjunkturelle Schwankungen oder dem gesamtwirtschaftlichen Rückgang der Beschäftigungsquoten erfasst werden, kann dafür statistisch kontrolliert werden. Zeitvariierende unbeobachtete Effekte, wie zum Beispiel individuelle Verhaltensänderungen, sind mit der obigen Bedingung jedoch nicht vereinbar. Deren Vernachlässigung resultiert sehr wahrscheinlich in einer verzerrten Schätzung des Maßnahmenerfolges.

In Erwartung der Teilnahme an einer arbeitsmarktpolitischen Maßnahme kann sich beispielsweise die Suchintensität vor der Maßnahme deutlich einschränken. Arbeitslose, die noch eine Chance auf Beschäftigung gehabt hätten, werden ihre Suchanstrengungen völlig minimieren. Nimmt man nun die Beschäftigungsquote unmittelbar vor Maßnahme als relevanten Vergleichsmaßstab, so kann sich der Erfolg der Maßnahme als ein sehr deutlicher herausstellen, da die Beschäftigungsquote unmittelbar vor Maßnahme aufgrund des individuellen Verhaltens der potentiellen Teilnehmer relativ niedrig war. Würde man den Zeitpunkt der Beobachtung vor Maßnahme weiter in die Vergangenheit schieben, so würde man unter Umständen eine höhere Beschäftigungsquote feststellen, der Erfolg der Maßnahme würde sich dementsprechend niedriger darstellen³². Da der Zeitpunkt der

³¹ Steiner, 2000, S. 52

³² Dieser Effekt wird als 'Ashenfelter's Dip' bezeichnet (Steiner, 2000, S. 53)

Antizipation der Teilnahme an einer arbeitsmarktpolitischen Maßnahme jedoch nicht bekannt ist, muss er mehr oder weniger willkürlich festgelegt werden; und damit auch der geschätzte Maßnahmenerfolg.

Statistische Evaluationen alleine würden wahrscheinlich zu kurz greifen³³. Sie ermöglichen Effektivitätsüberprüfungen auf Basis vorhandenen Datenmaterials und gewählter Erfolgskriterien. Um jedoch Ansatzpunkte für steuernde Eingriffe zu gewinnen, etwa Veränderungen im Maßnahmendesign um eine Verbesserung der Zielgruppenorientierung zu erreichen, bedarf es umfassenderer Analysen und einer entsprechenden Einbeziehung weiterer, meist vorgelagerter Evaluierungsebenen.

Um den vorangegangenen Abschnitt durch ein Beispiel zu illustrieren, wird im folgenden die Analyse des Integrationserfolges im Rahmen der Evaluation des *abzwien* kurz dargestellt.

5.3.4.3 Analyse des Integrationserfolges im Rahmen der Evaluation abzwien

Grundlage der Verlaufsanalyse war ein Datensatz, der aus einer Kombination von Daten aus dem EDV-System des Arbeitsmarktservice und aus den Versicherungsverläufen des Hauptverbandes der Sozialversicherungsträger bestand. Diese Datenquellen beinhalten zum einen Grundcharakteristika von Personen und zum anderen deren Versicherungsverläufe. Zur Verfügung standen Verlaufsdaten für den Zeitraum 1.1.1990 bis 31.12.1999. Die Verlaufsanalyse konzentrierte sich auf zwei zentrale Maßnahmen des *abzwien*; deren Effektivität wurde anhand folgender Kriterien überprüft:

- **Integrationserfolg:** Konnten die Teilnehmerinnen langfristig stabile Beschäftigungsverhältnisse aufnehmen?
- **Einkommensvergleiche:** Um wie viel besser waren die Jobs nach der Maßnahme?
- **Diskontinuitäten:** Waren die Maßnahmen in der Lage, Diskontinuitäten in den Erwerbsverläufen zu reduzieren?

Ein zentrales Problem, welches die Interpretation der gesamten Analyse (und in der Folge auch der Fiskalanalyse) erschwert, ist die nicht gesicherte Kausalität, mit der man Effekte der Maßnahme zurechnen kann. Solche Vorher-Nachher-Vergleiche sind daher, wie bereits oben erwähnt, nur ein sehr unzureichender Ersatz für die richtige Fragestellung „was wäre ohne Maßnahme passiert“, zumal sich hier viele Einflüsse mischen: Konjunkturelle Schwankungen, die persönliche Lebensgeschichte der Frauen, die Kindererziehungszeiten beenden, etc. Ganz bewusst wurde daher von einer Darstellung etwa von Weiterbeschäftigungsquoten (also des sich in Beschäftigung Befindens unmittelbar nach dem Ende der Fördermaßnahme) Abstand genommen, da diese Kennzahl schließlich nichts über die Dauer der Beschäftigung aussagt.

5.3.5 Bilanzierung

Werden nun dem Nutzen einer Maßnahme (z. B. eingesparte Arbeitslosenunterstützung, höhere Einkommen, höhere Steuerleistung) die Kosten der Maßnahme gegenübergestellt, so erhält man ein Maß für die Effizienz der Maßnahme. Eine Maßnahme ist dann effizient,

³³ Blaschke/Plath, 2000, S. 463

wenn ihre Nutzen-Kosten-Relation größer ist als die aller anderen denkbaren Alternativen (strenges Kriterium der **Paretoeffizienz**³⁴). Dies würde also bedeuten, dass Kosten und Nutzen einer Maßnahme mit anderen Maßnahmen der aktiven und passiven Arbeitsmarktpolitik verglichen werden sollte. Eine solche Vorgangsweise wird jedoch kaum je gewählt, was sich vor allem durch die unbefriedigende Verfügbarkeit von Datenmaterial bzw. aufgrund des enormen Datenerhebungsaufwandes (Recherchen bei verschiedensten Maßnahmenträgern notwendig) erklären lässt³⁵ und auch auf grundsätzlich methodischen Problemen der Kosten-Nutzen-Analyse beruht (siehe dazu auch den Teilabschnitt zur Kosten-Nutzen-Analyse).

Darüber hinaus erscheint es nicht sinnvoll, wahllos Maßnahmen für einen Vergleich heranzuziehen. Die Maßnahmen der aktiven Arbeitsmarktpolitik unterscheiden sich nicht nur in ihren Inhalten sondern auch nach den jeweiligen Zielgruppen und somit deren jeweiligen Problemlagen. Würde man eine Qualifizierungsmaßnahme für Wiedereinsteigerinnen mit einer für Langzeitarbeitslose vergleichen, so wäre dieser Vergleich streng genommen zulässig und richtig; in der Konzeption der Kosten-Nutzen-Rechnung müssten "...wenn man sie richtig durchführt, [auch] implizite Wirkungen explizit gemacht werden... (Wildavsky 1970, S. 373)". Die genannten Problemgruppen sind jedoch am regulären Arbeitsmarkt mit unterschiedlichen (exogen bestimmten) Vorurteilen konfrontiert, weswegen der Einstieg für Wiedereinsteigerinnen unter Umständen leichter zu bewältigen ist als für Langzeitarbeitslose, denen häufig das Stigma des "Problemfalls" (psychische und physische Beeinträchtigung, Alkoholprobleme,..) anhaftet. Bei einer streng monetären Bewertung würden unterschiedliche Beschäftigungschancen daher unberücksichtigt bleiben, weil sie einer solchen nicht zugänglich sind. Sie sind jedoch sicher wichtiger Bestandteil des Nutzen bzw. ein nicht vernachlässigbarer Einflussfaktor für den Erfolg einer Maßnahme. Werden diese Wirkungen daher nur qualitativ gegenübergestellt, wird damit das Prinzip der Paretoeffizienz durchbrochen, da eine unmittelbare Vergleichbarkeit nicht mehr gegeben ist. Daraus kann man ableiten, dass der Vergleich einer Maßnahme der aktiven Arbeitsmarktpolitik mit anderen nur dann zu validen und seriösen Ergebnissen führen, wenn sowohl Gegenstand als auch Zielgruppenorientierung möglichst ähnlich gestaltet sind.

Die Be-'wertung' der Ergebnisse einer Maßnahme erfolgt meist im Rahmen einer Kosten-Nutzen-Rechnung. Auf die Bestimmung des Begriffes und Konzeptes der Kosten-Nutzen-Analyse wird in einem späteren Kapitel noch eingegangen werden.

5.4 Vergleich

In diesem Abschnitt wird nun versucht, die zuvor dargestellten Modelle zur Evaluierung von Programmen oder Projekten gegenüberzustellen bzw. miteinander zu verknüpfen, um so zu einem idealtypischen Evaluierungsrahmen zu gelangen, der sowohl inhaltliche als auch zeitliche Dimensionen miteinander vereint. Dafür ist es notwendig, die einzelnen Phasen des Evaluierungsschemas nach Schmid et. al [1997] in Bezug zur inhaltlichen Evaluierungsdimension in Kapitel 5.2.5.2 oben zu stellen.

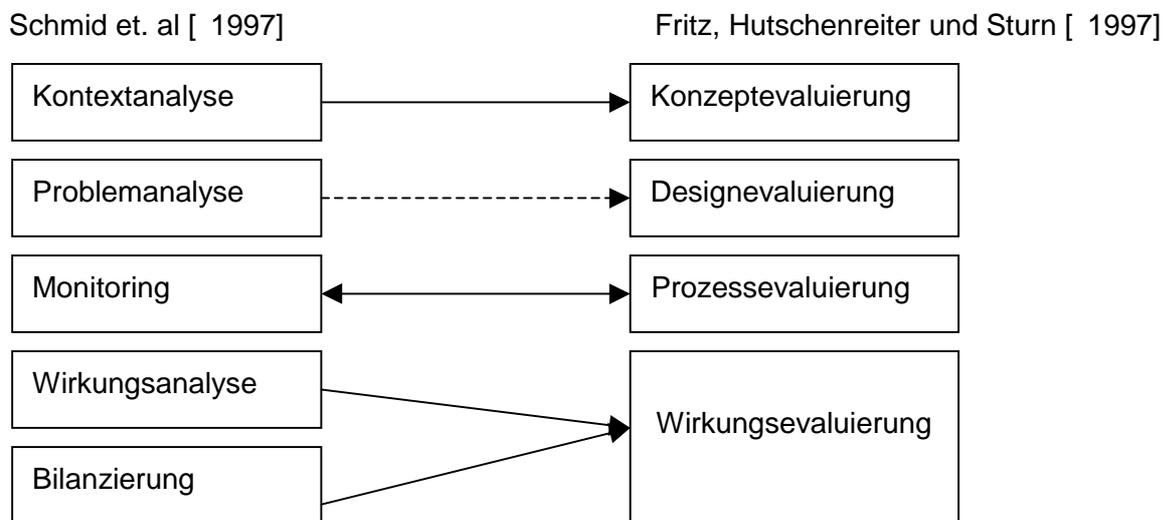
³⁴ Schmid et al., 1997, S. 32

³⁵ Viktor Steiner, 2000, S. 21

Fritz, Hutschenreiter und Sturn [1997] unterscheiden zwischen Konzept- Design-, Prozess-, und Wirkungsevaluierung – eine Unterscheidung, die aus der langen Tradition technologiepolitischer Evaluierung stammt und auf den ersten Blick nicht direkt auf jene des Schmid’schen Schemas umlegbar ist. Grund hierfür ist, dass Schmid die Evaluierungsphasen einem gewissen Zeitpunkt zuordnet, während die Evaluierungsarten in 5.2.5.2 oben lediglich nach der Zielrichtung der in ihnen zusammengefassten Fragestellungen bündeln: Schmid et al. integrieren also Inhalte und Evaluierungszeitpunkte. In 5.2.6 wurde bereits der Schluss gezogen, dass die in Österreich vorherrschende Form arbeitsmarktpolitischer Evaluierungen die ex-post Analyse von Analyse von Programmen oder Maßnahmen der aktiven Arbeitsmarktpolitik ist. Von ihrer zeitlichen Dimension ist sie gleichzeitig auch die einzige, bei der alle Evaluierungsphasen nach Schmid et al. anwendbar sind. Ex-ante Evaluierung hingegen würden auf Kontext- und eventuell Problemanalyse beschränkt bleiben. Eine ähnlich klare zeitliche Zuordnung ist für das Evaluierungsschema nach Fritz, Hutschenreiter und Sturn [1997] nicht machbar, obwohl natürlich auch hier implizite zeitliche Dimensionen enthalten sind.

Bricht man die Phasen jedoch auf ihnen zugeordnete Fragestellungen herunter, so können erste Evidenzen für eine Vergleichbarkeit gefunden werden. Nachfolgende Darstellung versucht, die Evaluierungsphasen nach Schmid et al. den Evaluierungsarten nach Fritz et al. zuzuordnen:

Abbildung 4: Evaluierungsphasen und Evaluierungsarten



eigene Darstellung © InTeReg

Schmid et al. ordnen der **Kontextanalyse** die Erfassung der Rahmenbedingungen (externe sowie rechtlich -institutionelle) des Arbeitsmarktes und der Arbeitsmarktpolitik sowie die Einbettung der arbeitsmarktpolitischen Strategie innerhalb dieses Kontextes zu, wobei dies vor allem der Überprüfung der Klarheit der Zielformulierung, der Konsistenz und Komplementarität der Ziele dienen soll. Auch Fritz et al. formulieren ähnlich Inhalte für die Phase der **Konzeptevaluierung**; sozioökonomische, wirtschaftliche und technologische Rahmenbedingungen sowie die Interventionslegitimation sollen ebenso überprüft werden, wie die Vereinbarkeit von Schwerpunkten (siehe Tabelle 2: Zielsetzungen des AMS

Steiermark 1998) bzw. (übergeordneten) Zielen einerseits untereinander und andererseits auf der Ebene der operativen Ziele. Dabei ist es wichtig, potentielle Zielkonflikte aufzuzeigen, um bereits in dieser Phase Lösungsansätze und/oder mögliche Alternativen anzudenken, die jedoch erst nach Abschluss der Designevaluierung konkreter ausformuliert werden können.

In der bereits angesprochenen Phase der **Designevaluierung** sollen gemäß Fritz et al. das Interventionsmodell, die operativen Ziele (z. B. Zielgruppenerfassung/-erreichung) und die organisatorische Struktur überprüft werden. Die Problemanalyse bei Schmid et al. hingegen soll arbeitsmarktpolitischen Handlungsbedarf sowie einzusetzende Instrumente präzisieren. Angebotsseitige (Welche Personen sind besonders von Arbeitslosigkeit betroffen?) sowie nachfrageseitige (Analyse der offenen Stellen, Ergünden des Qualifikationsbedarfes) Überlegungen fließen hier ebenso ein wie eine daran anschließende Analyse des Mismatch für die zu untersuchende Region. In der letzten Phase soll dann die Abstimmung zwischen Programmstruktur und Problemdruck erfolgen; es soll eine Beurteilung dahingehend vorgenommen werden, ob Ziele und Instrumenteneinsatz mit den Problemlagen übereinstimmen bzw. diesen überhaupt angemessen sind. Diese als "**Programmbewertung**" bezeichnete Phase entspricht unseres Erachtens nach der Designevaluierung bei Fritz et al. Die Problemanalyse bei Schmid et al. ist wesentlich umfassender gestaltet und beinhaltet auch eine meso- und mikroökonomische Beschreibung des Arbeitsmarktes, welcher in der Designevaluierung hingegen kaum eine Bedeutung zukommt bzw. nicht explizit behandelt wird.

Monitoring nach Schmid et al. beschäftigt sich mit der laufenden Kontrolle des Erfolges (Zielkontrolle) während der Umsetzung von Programmen. **Prozessevaluierung** nach Fritz et al. soll die Mechanismen der Durchführung von Programmen im Detail untersuchen, um so zu einer Überprüfung der (möglichen) Abweichungen zwischen Programmrealisierung und Programmplanung zu gelangen. Die Einhaltung von organisatorischen Aufbau und Ablauf soll dabei ebenso erfasst werden, wie die adäquate Verwendung von Budgetmittel (z. B. Werbekosten etc.). Diese Evaluierungsschritte sind komplementär und sollten daher zu einer sogenannten **Implementationsevaluierung** zusammengefasst werden. Implementationsevaluierung bezieht sich auf die Überprüfung der Umsetzung einer Maßnahme im Hinblick auf Organisation, Inhalt und TeilnehmerInnenstruktur unter Berücksichtigung der Ziel- und Planungsvorgaben sowie der internen Effizienz³⁶. Dabei kann man zwischen verschiedenen Strategien unterscheiden: Monitoring, Steuerung durch Leistungskennzahlen und Prozeß-Evaluation.

In der Literatur wird immer wieder betont, dass Implementierungsanalysen für die Weiterentwicklung von laufenden Programmen/Projekten unbedingt notwendig sind, da die klassischen Wirkungsanalysen oft nur Ja/Nein Aussagen ergeben. Allerdings ist dieses Feld noch sehr wenig entwickelt und wird als das "big missing link in most evaluations studies" (Schmid 1996, S. 227) bezeichnet³⁷. Es wurden zwar noch keine Standard-Methoden entwickelt, es "setzt sich jedoch zunehmend die Auffassung durch, dass die Durchführung von Programmen auf Maßnahmenebene grundsätzlich mit internen Berichts- und Informationssystemen (Monitoring) und Feed-back Mechanismen (Leistungskennzahlen)

³⁶ Lassnigg/Leitner/Wroblewski/Steiner/Steiner/Mayer (IHS), Schmid/Schömann (WZB), Evaluationsschema für Maßnahmen der aktiven Arbeitsmarktpolitik in Wien, Projektbericht im Auftrag des AMS Wien und WAFF, EQUI, IHS 2000.

³⁷ Lassnigg et al., Schmid et al. 2000, S. 32

verbunden sein sollte, die den beteiligten Akteuren laufende Informationen über die Durchführung liefern können" (Lassnigg et al., Schmid et al. 2000, S. 32). Unseres Erachtens sollte daher Abbildung 3: Etappen der Evaluierung der Arbeitsmarktpolitik und ihre Abfolge insofern modifiziert werden, als dass die Phase des Monitorings um die Prozessevaluierung ergänzt wird und beide gemeinsam unter dem Begriff der Implementierungsanalyse untergeordnet werden sollten.

Wirkungsanalyse und **Bilanzierung** sind im Evaluierungsschema von Fritz et al. nicht voneinander getrennt. **Wirkungsevaluierung** beinhaltet daher sowohl die Beurteilung der Effektivität als auch der Effizienz einer Maßnahme.

5.5 Anforderungen an arbeitsmarktpolitische Evaluierungen in Österreich

Manche Akteure der Arbeitsmarktpolitik in Österreich greifen relativ häufig auf Evaluierungen als Entscheidungshilfe zurück; Abbildung 2 hat bereits (Teil-)Ergebnisse des gegenständlichen Kapitels vorweggenommen und dargestellt, in welchen Bereichen der Arbeitsmarktpolitik Evaluierungen eingesetzt werden und in welchen nicht bzw. sehr eingeschränkt. Ziel des Kapitels 5.5 wird es sein, diese Evaluierungspraxis einer genaueren Betrachtung zu unterziehen.

Auftraggeber einer Evaluierung im arbeitsmarktpolitischen Konnex sind im Regelfall die Landesgeschäftsstellen des Arbeitsmarktservices (vgl. dazu auch Tabelle 5). Auch bei den MitarbeiterInnen der Maßnahmen (TrägervertreterInnen bzw. sogenannten Schlüsselkräften) besteht natürlich Interesse an den Evaluierungsergebnissen.

Seit 1997 hat das Arbeitsmarktservice Österreich bzw. seine Landesgeschäftsstellen ihre Einladung zur Offertelegung von Evaluierungen von arbeitsmarktpolitischen Maßnahmen immer mehr standardisiert; in Wien agiert als Einladender neben der LGS des AMS Wien auch der Wiener ArbeitnehmerInnenförderungsfonds, WAFF. Nachfolgend sollen die Terms of References auf gewünschte Inhalte, Methoden bzw. Anforderungen an die Anbotlegenden ausgewertet werden.

Die Einladungen sind in der Regel immer gleich strukturiert: Nach einer Vorstellung des *Arbeitsmarktservices* bzw. des *WAFF* werden in prägnanter Form *Hintergrundinformationen* zu der zu evaluierenden Maßnahme geboten. Anschließend werden *zentrale Fragestellungen* formuliert, der gewünschte *Methodenmix* und die *Anforderungen an das angebotslegende Institut* dargestellt. Die Bestandteile und Struktur der Angebotsunterlagen werden umrissen, abschließend die Eckdaten des Forschungsvorhabens bzw. die Dotierung festgehalten. In nachfolgender Auswertung fanden Evaluierungen Berücksichtigung, zu denen Joanneum Research ab Dezember 1999 eingeladen wurde, ein Angebot zu legen. Es handelt sich hierbei jedoch nicht um ein vollständiges Bild der arbeitsmarktpolitischen Evaluierungen in Wien in diesem Zeitraum:

Tabelle 5: Ausgewählte Evaluierungsprojekte von AMS Wien und WAFF

Projekttitlel	Art der Maßnahme	Art der Evaluierung	Auftraggeber	Laufzeit	Dotierung ³⁸
Evaluierung der Maßnahmen im Rahmen des Territoralen Beschäftigungspaktes 1999 für Wien		Programm-evaluierung	AMS Wien/WAFF	12/00 – 12/01	ATS 1,250.000.-
Evaluierung des Sozialökonomischen Betriebes Pa-Re	Beschäftigung	Projekt-evaluierung	AMS Wien	6/00 – 10/00	ATS 300.000.-
Evaluierung von ausgewählten Betrieben mit Lehrstellenförderung in Wien	Beschäftigung	Projekt-evaluierung	AMS Wien	03/00 – 10/00	ATS 350.000.-
Evaluierung der stiftungsähnlichen Maßnahme WiederIn	Beschäftigung	Projekt-evaluierung	AMS Wien/WAFF	12/99 – 07/00	ATS 400.000.-
Evaluierung des Pilotprojektes „Regionales Wirtschaftsservice“	Beratung	Projekt-evaluierung	WAFF	06/00 – 10/00	ATS 400.000.-
Evaluierung der arbeitsmarktpolitischen Maßnahmen „Berufsorientierung und Individualbetreuung durch „Die Berater“ und „Job Now“	Beratung	Evaluierung eines Projektbündels	AMS Wien/WAFF	12/99 – 6/00	ATS 800.000.-
Evaluierung der arbeitsmarktpolitischen Maßnahme Comeback	Beschäftigung	Projekt-evaluierung	AMS Wien/WAFF	10/99 – 04/00	ATS 320.000.-

Quelle: Eigene Auswertung Anhand der Einladungen zur Offertlegung

Auffallend ist hier, dass vorrangig Maßnahmen aus dem Bereich Beschäftigung evaluiert wurden, keinerlei Maßnahmen aus dem Bereich Qualifizierung. Der Grund hierfür dürfte jedoch nicht im Bereich des AMS zu suchen sein, sondern lediglich im Auswahlprozess für die vorliegende Auswertung.

³⁸ Die Dotierung versteht sich als Maximalsumme inkl. aller Abgaben. Unter „Abgaben“ wird auch die Umsatzsteuer verstanden.

Tabelle 6 gibt eine Übersicht über die in den Terms of References der Einladungen zur Offertlegung gewünschten Methodenmixes. Hier fällt auf, dass die Evaluierung der Lehrstellenförderung eine Ausnahme in der gängigen Anforderungsstruktur bildet: Auch der gewählte Vergabemodus dieses Projektes (im Gegensatz zu allen anderen Projekten war neben einem schriftlichen Angebot auch eine Vorstellung des Projektvorhabens gewünscht) lässt den Schluss zu, dass es sich hierbei um ein Evaluierungsvorhaben abseits normierter Bahnen handelte.

Einige der gewünschten Methoden sind als Standardanforderung zu charakterisieren: Literaturrecherche und –analyse, Vergleich mit ähnlichen Maßnahmen, Aktenanalyse. Auch die Analyse von AMS-Daten und Erhebungen bei den TeilnehmerInnen sind als solche zu identifizieren. Eine Präzisierung erfolgte bei der Anforderung nach ExpertInneninterviews: Hier wird in den Terms of References nun der zu interviewende Personenkreis genauer umrissen.

Tabelle 6: Terms of References: gewünschter Methodenmix

	Evaluierung TEP Wien	Evaluierung SöB Pa-Re	Evaluierung Lehrstellenförderung Wien	Evaluierung WiederIn	Evaluierung „Regionales Wirtschaftsservice“	Evaluierung „Die Berater“ und „Job Now	Evaluierung Comeback
Gewünschter Methodenmix:							
Literaturrecherche und –analyse	✓	✓	✓	✓	✓	✓	✓
Aktenanalyse	✓	✓		✓	✓	✓	✓
Vergleich mit ähnlichen Maßnahmen		✓		✓	✓	✓	✓
Quantitative Analyse von AMS- Einzeldatensätzen	✓	✓		✓	✓	✓	
Quantitative Analyse von AMS- Aggregatdaten	✓	✓		✓	✓	✓	
Leitfadengestützte Totalerhebung der TeilnehmerInnen		✓					
Qualitative Exploration							
anhand der Befragung des/r Träger/s	✓	✓	✓				
anhand der Befragung der Schlüsselkräfte		✓	✓				
anhand der Befragung der mit der Maßnahme befassten Akteure	✓	✓	✓				
anhand von ExpertInneninterviews				✓	✓	✓	✓
Quantitative Primärerhebung ³⁹ bei den TeilnehmerInnen			✓	✓	✓	✓	✓
Begleitgruppe					✓		✓

Quelle: Eigene Auswertung Anhand der Einladungen zur Offertlegung

³⁹ In einigen Fällen wird unter einer quantitativen Primärerhebung eine Totalerhebung verstanden.

Tabelle 7: Terms of References: Anforderungen an den Auftragnehmer

	Evaluierung TEP Wien	Evaluierung SöB Pa-Re	Evaluierung Lehrstellenförderung	Evaluierung WiederIn	Evaluierung „Regionales Wirtschaftsservice“	Evaluierung „Die Berater“ und „Job Now	Evaluierung Comeback
Anforderungen an den Auftragnehmer:							
Einschlägige sozialwissenschaftliche Forschungstätigkeit	✓	✓	✓	✓	✓	✓	✓
Einschlägige wirtschaftswissenschaftliche Forschungstätigkeit			✓				
Erfahrung auf dem Feld der Arbeitsmarktforschung	✓	✓	✓	✓	✓	✓	✓
Erfahrung auf dem Gebiet der Evaluierung AMP-Maßnahmen	✓			✓	✓	✓	✓
Einschlägige Evaluierungserfahrung		✓					
Erfahrung mit der Evaluierung von frauenspezifischen Maßnahmen		✓					
Anwendungserfahrung quantitativer Methoden	✓			✓	✓	✓	✓
Anwendungserfahrung qualitativer Methoden	✓	✓		✓	✓	✓	✓
Erfahrung auf dem Gebiet der Strukturanalyse			✓				

Quelle: Eigene Auswertung Anhand der Einladungen zur Offertlegung

Die Auswertung der Terms of References nach Anforderungen an den Auftragnehmer spiegeln die Auswertungen in der vorrangegangenen Tabelle 7: Terms of References: Anforderungen an den Auftragnehmer wider. Erfahrung auf den Gebieten der Arbeitsmarktforschung und der Evaluierung sind Standards, nur in Einzelfällen werden dem Charakter der zu evaluierenden Maßnahme entsprechend Zusatzqualifikationen eingefordert: Beispiel hierfür sei die Erfahrung mit der Evaluierung von frauenspezifischen Projekten oder Programmen.

In Tabelle 8 werden nun die primären Fragestellungen der Terms of References den einzelnen Phasen einer idealtypischen Evaluierung zugeordnet:

Tabelle 8: Terms of References der Evaluierungsausschreibungen von AMS und WAFF: Primäre Fragestellungen

Primäre Fragestellungen	Kontextanalyse	Problemanalyse	Implementierungsanalyse	Wirkungsanalyse	Bilanzierung
Gegenüberstellung: Maßnahme Anforderungen des Arbeitsmarkts	✓				
Überprüfung der regionalen Settings	✓				
Bewertung der Maßnahmen im Spannungsfeld zwischen Arbeitsmarktpolitik, sozialarbeiterischen Ansprüchen und betriebswirtschaftlichen Notwendigkeiten.	✓				
Darstellung der Struktur der Grundgesamtheit der Zielgruppe	✓				
Gegenüberstellung der Strukturdaten von Zielgruppe und TeilnehmerInnen		✓			
Gegenüberstellung: Maßnahme – Bedürfnisse der Zielgruppe		✓			
Überprüfung der Anstrengungen des Maßnahmenträgers bei der Zieldefinition und der Zielerreichung		✓			
Treffsicherheit Population – Zielgruppe		✓			
Darstellung des Verhältnisses aller beteiligten Akteure zueinander.		✓	✓		
Untersuchung der didaktischen und inhaltlichen Konzeption in allen Phasen der Durchführung			✓		
Sozialökonomische Analyse der TeilnehmerInnen			✓		
Untersuchung des gesamten Maßnahmenumfeldes und –ablaufes			✓		
Untersuchung der administrativen Effizienz im Hinblick auf die beteiligten Institutionen, die Organisation			✓		

Primäre Fragestellungen	Kontextanalyse	Problemanalyse	Implementierungsanalyse	Wirkungsanalyse	Bilanzierung
der Maßnahme					
Überprüfung des Qualitätsmanagements des Maßnahmenträgers			✓		
Schnittstellenanalyse zur Durchführung der Maßnahme			✓		
Drop-Out Raten der TeilnehmerInnen			✓	✓	
Erhebung der TeilnehmerInnenzufriedenheit				✓	
Integrationsleistungen der Maßnahme				✓	
Karriereverläufe der TeilnehmerInnen				✓	
Analyse der Effekte für die beteiligten Unternehmen				✓	
Analyse der quantitativen und qualitativen Effekte auf die Beschäftigung.				✓	
Abschätzung von Nutzen und Kosten				✓	✓
Erarbeitung von Vorschlägen für die Weiterentwicklung der Maßnahme, Beratung und Vermittlung					✓
Nachhaltigkeit der Wirkungen					✓
Veränderungswünsche der TeilnehmerInnen bezüglich der Maßnahme					✓
Bewertung der Effizienz der Maßnahmen im Spannungsfeld zwischen Arbeitsmarktpolitik, sozialarbeiterischen Ansprüchen und betriebswirtschaftlichen Notwendigkeiten.					✓

Quelle: Eigene Auswertung Anhand der Einladungen zur Offertlegung

5.6 Zusammenfassung

Mit vorliegendem Papier wurde versucht, das Feld der arbeitsmarktpolitischen Evaluation für Österreich enger zu fassen: Mögliche Gegenstände, zeitliche und inhaltliche Dimensionen wurden zuerst allgemein umrissen, um schließlich ein idealtypisches Evaluierungskonzept für Arbeitsmarktpolitik vorzustellen, welches in bestimmten Punkten mit einem aktuellen Projektbeispiel illustriert wurde.

Als elementare Bestandteile eines umfassenden Evaluierungsansatzes können Kontextanalyse, Problemanalyse, Implementationsanalyse, Wirkungsanalyse und schließlich Bilanzierung identifiziert werden. Kontextanalyse hat die Erfassung der Rahmenbedingungen des Arbeitsmarktes und der Arbeitsmarktpolitik zum Ziel, eine Problemanalyse die Präzisierung des arbeitsmarktpolitischen Handlungsbedarfs und der einzusetzenden Instrumente, eine Implementationsanalyse die Kontrolle des Erfolges während der Umsetzung von Programmen und ihrer Entstehung, eine Wirkungsanalyse die Bestimmung des tatsächlichen Effektes (Outcome) der Arbeitsmarktpolitik auf operativen Zielgrößen heruntergebrochen, die Bilanzierung schließlich die Messung von Nutzen und Kosten (soweit wie möglich) in monetären Größen und Gegenüberstellung in einer Nutzen- Kosten-Analyse.

Gegenstand einer typischen, österreichischen arbeitsmarktpolitischen Evaluierung ist fast immer ein Projekt der aktiven Arbeitsmarktpolitik. Alle möglichen Evaluierungsinhalte werden in der Regel in einer ex-post Analyse behandelt.

Aus den Auswertungen im vorangegangenen Kapitel (insbesondere in Tabelle 8) der herrschenden Evaluierungspraxis in der Arbeitsmarktpolitik wird klar, dass Evaluierungen

- keine „Black-Box Evaluationen“⁴⁰ im Sinne einer abstrakten Gegenüberstellung der Inputs und der Outputs von Maßnahmen mehr sind, sondern umfassende Designansätze bieten müssen.
- nach wie vor Projektevaluationen sind; Projektbündel oder Programme sind eher selten Gegenstand von Evaluierungen.

Nichtsdestotrotz erzeugen Terms of References implizit durch den Auftraggeber im Evaluierungsdesign eine Gewichtung zwischen den inhaltlichen Blöcken einer Evaluierung. Ungeachtet dieser Vorselektion muss jedoch, um einem hohen inhaltlichen Evaluierungsanspruch entsprechen zu können, eine Balance zwischen den einzelnen Modulen, also zwischen Kontextanalyse bis hin zur Bilanzierung, gefunden werden.

5.7

⁴⁰ Lassnig et al., Seite 11, Schmid G., O'Reilly J., Schömann K. Seite 3

Literatur

- Almus, M.; Egel, J., Lechner, M.; Pfeiffer, F.; Spengler, H.; Die gemeinnützige Arbeitnehmerüberlassung in Rheinland-Pfalz – eine ökonometrische Analyse des Wiedereingliederungserfolgs, ZEW working paper, 1998
- Blaschke, D., Plath, H.-E., Möglichkeiten und Grenzen des Erkenntnisgewinns durch Evaluation aktiver Arbeitsmarktpolitik. In: Mitteilungen aus der Arbeitsmarkt- und Berufsforschung (MittAB), 33. Jg./2000, Heft 3
- Deutsches Institut für Wirtschaftsforschung DIW : Evaluation im Spannungsfeld zwischen Wissenschaft und Politik, Vierteljahresshefte zur Wirtschaftsforschung, Heft 3, 69. Jahrgang, 2000
- Fritz, O., Hutschenreiter G., Sturn D.: Evaluierung von FTE-Programmen zwischen Best Practice Entwicklung und österreichischen Strukturen, Studie im Auftrag des Bundesministeriums für Wissenschaft, Verkehr und Kunst, Wien – Graz 1997
- Gruber, M., Zinöcker K., Gerhardter G., Mahringer H., Woitech, B.: Steuerungsformen des AMS und ihre Rückwirkungen auf arbeitsmarktpolitische Träger anhand des Beispiels Obersteiermark, Studie im Auftrag des AMS LGS Steiermark., Wien - Graz 2000
- Heckman, J., Smith, J., Experimental and Nonexperimental Evaluation. In: International Handbook of Labour Market Policy and Evaluation, Schmidt, O` Reilly, Schömann (eds.); E. Elgar, 1996, pp.37-88
IZA Compact, Dezember 2000
- Lassnigg, L., Leitner, A., Wroblewski, A., Steiner M., Steiner V., Mayer K., (IHS), Schmid G., Schömann K. (WZB): Evaluationsschema für Maßnahmen der aktiven Arbeitsmarktpolitik in Wien, Projektbericht im Auftrag des AMS Wien und WAFF, EQUI, IHS 2000.
- O'Farrel, P.N., Hitchens, D., Small Firm Competitiveness and Performance, Gill and Macmillan, Dublin 1989
- Schmid G., Auer p., Mosley H., Schömann K. (Eds.): Progress in Evaluation Research: Documentation of Two Transfer Workshops on the "International Handbook of Labour Market Policy and Evaluation" Discussion Paper FS I 97-203, Wissenschaftszentrum Berlin für Sozialforschung 1997
- Schmid G., O'Reilly J., Schömann K.: Theory and Methodology of Labour Market Policy and Evaluation: An Introduction. In: Schmid G. et al.: International handbook of labour market policy and evaluation, Cheltenham [u.a.] , Elgar, 1997
- Schmid, G., Schömann K., Schütz H.: Evaluierung der Arbeitsmarktpolitik. Ein analytischer Bezugsrahmen am Beispiel des Arbeitsmarktpolitischen Rahmenprogramms in Berlin, Discussion Paper FS I 97-204, Wissenschaftszentrum Berlin für Sozialforschung 1997
- Steiner, V., Hagen, T.: Von der Finanzierung der Arbeitslosigkeit zur Förderung der Arbeit – Analysen und Empfehlungen zur Steigerung der Effizienz und Effektivität der Arbeitsmarktpolitik in Deutschland. ZEW Wirtschaftsanalysen, Nomos Verlag, Baden-Baden 2000
- Sturn D., Woitech B., Mahringer H., Evaluierung des abzwien, Studie im Auftrag des AMS Wien, Wien 2000
- Zilian, H. G., Lassnigg, L., Wroblewski, A., Arbeitslosenschulung in der flexibilisierten Wirtschaft, Hampp Verlag, München, 1999