

BERICHT
INSTITUT FÜR
TECHNOLOGIE- UND REGIONALPOLITIK

*QUANTITATIVE METHODEN DER
EVALUIERUNG TECHNOLOGIEPOLITISCHER
PROGRAMME*

JOANNEUM RESEARCH

INSTITUT FÜR
TECHNOLOGIE- UND REGIONALPOLITIK

QUANTITATIVE METHODEN DER EVALUIERUNG TECHNOLOGIEPOLITISCHER PROGRAMME

IM AUFTRAG DES BUNDESMINISTERIUMS
FÜR VERKEHR, INNOVATION UND TECHNOLOGIE
RTG.1999.AF.006-01

Projektleitung: W. Polt und D. Sturn
W. Pointner, B. Woitech: (Joanneum Research)
A. Geyer (ARC Seibersdorf), G. Hutschenreiter (WIFO)
C. Rammer (ZEW)

Wien, April 02



WIFO



INHALTSVERZEICHNIS

EINLEITUNG	IV
TEIL A	1
1 EVALUIERUNG VON DIFFUSIONSFÖRDERPROGRAMMEN BEISPIELE AUS DER ANALYSE DER MANUFACTURING EXTENSION PARTNERSHIP (Wolfgang Pointner)	2
1.1 DIE MANUFACTURING EXTENSION PARTNERSHIP (MEP)	2
1.2 DYNAMISCHES MODELL DES ENTSCHEIDUNGSPROZESSES VON MARKTTILNEHMERN	3
1.3 ZIELE UND VERFAHREN VON EVALUIERUNGEN	7
1.3.1 <i>Monitoring der Vollzugskontrolle</i>	8
1.3.2 <i>Fallstudien</i>	9
1.3.3 <i>Teilnehmerbefragung und deskriptive Auswertung</i>	9
1.3.4 <i>Ökonometrische Analysen</i>	11
1.3.5 <i>Verwendung von Sekundärdatenbasen</i>	14
1.3.6 <i>Experimentelle Methoden</i>	15
1.4 LITERATUR	17
2 EVALUIERUNG DES CIM-AKTIONSPROGRAMMS DER SCHWEIZER BUNDESREGIERUNG (Christian Rammer)	18
2.1 EINLEITUNG.....	18
2.2 DAS CIM-AKTIONSPROGRAMM DER SCHWEIZER BUNDESREGIERUNG	20
2.3 DER EVALUIERUNGSANSATZ	21
2.4 ERFAHRUNGEN MIT DER ANWENDUNG DES EVALUIERUNGSANSATZES IN DER SCHWEIZ.....	24
2.5 ERFAHRUNGEN MIT DER ANWENDUNG DES EVALUIERUNGSANSATZES IN ÖSTERREICH	28
2.6 SCHLUSSFOLGERUNGEN	30
2.7 LITERATUR	32
3 EVALUIERUNG DIREKTER FÖRDERUNG AM BEISPIEL DES „ADVANCED TECHNOLOGY PROGRAMS“ (Wolfgang Polt, Birgit Woitech)	33
3.1 GESCHICHTE UND ORGANISATION, ZIELE UND AUSRICHTUNG.....	33
3.2 DAS EVALUIERUNGSDESIGN VON ATP	37
3.3 EVALUIERUNG DER ÖKONOMISCHEN AUSWIRKUNGEN VON F&E-PROGRAMMEN	42
3.4 EVALUATION AUF DER BASIS VON TEILNEHMERINNENBEFRAGUNGEN	45
3.5 EVALUATION MITTELS KOSTEN-NUTZEN-ANALYSEN (KNA).....	48
3.5.1 <i>Fallstudie I: Medizinische Innovationsprojekte</i>	50
3.5.2 <i>Fallstudie II: Digitale Speichersysteme (Digital Data Storage - DDS)</i>	54
3.5.3 <i>Möglichkeiten und Grenzen der KNA</i>	58
3.6 ANWENDBARKEIT DES EVALUIERUNGSANSATZES IN ÖSTERREICH	59
3.7 LITERATUR	61
4 METHODEN DER EVALUATION ÖKONOMISCHER EFFEKTE VON TECHNOLOGIE-PROGRAMMEN: DIE B.E.T.A METHODE ZUR BEWERTUNG VON ESA PROGRAMMEN (Dorothea Sturn)	63
4.1 EINLEITUNG.....	63
4.2 ESA: DATEN UND ZIELE	63
4.2.1 <i>Ziele der ESA</i>	63

4.2.2	<i>Finanzierung der ESA Programme</i>	64
4.2.3	<i>Bewertung des Zielsystems</i>	64
4.3	ÜBERBLICK ÜBER METHODEN DER EVALUIERUNG VON RAUMFAHRTAKTIVITÄTEN	65
4.3.1	<i>Fallstudien</i>	65
4.3.2	<i>Ökonometrische Modelle</i>	66
4.3.3	<i>Vorher-Nachher-Vergleiche</i>	67
4.4	DER B.E.T.A. ANSATZ	67
4.4.1	<i>Ziele und Umfang der Evaluation</i>	67
4.4.2	<i>Methode</i>	68
4.4.3	<i>Datenerfordernis und Aufwand</i>	70
4.5	ERGEBNISSE DER EVALUATION	71
4.6	STÄRKEN UND SCHWÄCHEN DER METHODE	72
4.6.1	<i>Stärken der Methode</i>	72
4.6.2	<i>Schwächen bei der konkreten Anwendung der Methode</i>	73
4.6.3	<i>Prinzipielle Schwächen bzw. Probleme des Ansatzes</i>	73
4.7	WEITERENTWICKLUNGSMÖGLICHKEITEN UND ANWENDUNGSGBIETE	76
4.7.1	<i>Einsatzmöglichkeiten der Methode</i>	76
4.7.2	<i>Einsatzmöglichkeiten bei österreichischen Programmen</i>	77
4.8	LITERATUR	77
5	EVALUIERUNG VON NIST-LAB PROGRAMMEN (Anton Geyer)	79
5.1	ORGANISATION UND TÄTIGKEIT DES NIST	79
5.1.1	<i>Organisatorische Gliederung</i>	79
5.1.2	<i>Programmplanungs- und Evaluierungssystem</i>	79
5.2	ÖKONOMISCHE IMPACT-STUDIEN FÜR NIST-FORSCHUNGSPROGRAMME	81
5.2.1	<i>Ziel der Impact-Studien</i>	82
5.2.2	<i>NIST-Methodik zur Impact-Evaluierung</i>	82
5.2.3	<i>Quantitative Beurteilungsmaße</i>	84
5.3	ABSCHÄTZUNG DER ÖKONOMISCHEN IMPACTS AM BEISPIEL DES NIST-FORSCHUNGSPROGRAMMS „ALTERNATIVE KÄLTEMITTEL“	86
5.3.1	<i>Hintergrund des NIST Forschungsprogramms „Alternative Kältemittel“</i>	86
5.3.2	<i>Modell für die Impact Abschätzung</i>	87
5.3.3	<i>Durchführung der empirischen Erhebung</i>	88
5.4	ANWENDUNGSGBIETE, DISKUSSION UND FOLGERUNGEN	90
5.4.1	<i>Erfahrungen des NIST mit dem Impact-Modell</i>	90
5.4.2	<i>Nutzen des Modells für das NIST</i>	91
5.4.3	<i>Grenzen des Verfahrens zur Impact-Abschätzung</i>	91
5.4.4	<i>Anwendbarkeit des Modells in Österreich</i>	92
5.5	LITERATUR	93
6	EIN MODELL ZUR ABSCHÄTZUNG DER KOSTENEFFEKTIVITÄT: DIE EVALUATION VON INNOVATIONSFÖRDERPROGRAMMEN DES BRITISCHEN DEPARTMENT OF TRADE (DTI) (Anton Geyer)	94
6.1	EINLEITUNG UND HINTERGRUND	94
6.2	DAS NAO-MODELL ZUR ABSCHÄTZUNG DER RELATIVEN KOSTENEFFEKTIVITÄT	95
6.2.1	<i>Kriterien und Indikatoren für die Ermittlung der Programmergebnisse</i>	96
6.2.2	<i>Bestimmung der Kosteneffektivität verschiedener Programme</i>	96
6.2.3	<i>Untersuchte DTI-Programme</i>	98

6.2.4	<i>Evaluierungsergebnisse</i>	99
6.2.5	<i>Diskussion des NAO-Modells</i>	100
6.3	ERFAHRUNGEN UND ANWENDUNGSMÖGLICHKEITEN	102
6.3.1	<i>Mögliche Anwendungsgebiete des Modells in Österreich</i>	102
6.4	LITERATUR	103
7	COMEVAL EVALUIERUNGSTOOLKIT (Anton Geyer)	104
7.1	COMEVAL ENTWICKLUNGSGESCHICHTE	104
7.1.1	<i>Ziele des COMEVAL-Toolkits</i>	104
7.1.2	<i>Beschreibung der COMEVAL-Methode</i>	105
7.1.3	<i>Entwicklungsphasen und Test des COMEVAL-Toolkits</i>	106
7.1.4	<i>Komplementarität von COMEVAL mit anderen Ansätzen und Hilfsmitteln</i>	107
7.2	MODULE, INHALTE UND BEWERTUNGSVERFAHREN DES COMEVAL-TOOLKITS	108
7.2.1	<i>Module des Toolkits</i>	108
7.2.2	<i>Darstellung von Projektergebnissen und Effekten im COMEVAL-Toolkit</i>	110
7.2.3	<i>Der Evaluierungsbericht</i>	113
7.3	ERFAHRUNGEN UND MÖGLICHE ANWENDUNGSGEBIETE	116
7.3.1	<i>Anwendung des COMEVAL-Toolkits in der Europäischen Kommission</i>	116
7.3.2	<i>Anwendungsmöglichkeiten in Österreich</i>	117
7.4	LITERATUR	117
	TEIL B	116
8	EVALUIERUNG STEUERLICHER ANREIZE FÜR FORSCHUNG UND ENTWICKLUNG (Gernot Hutschenreiter)	119
8.1	STEUERLICHE ANREIZE FÜR F&E	119
8.2	VOR- UND NACHTEILE STEUERLICHER ANREIZE	119
8.3	STEUERLICHE ANREIZE FÜR F&E: EIN INTERNATIONALER ÜBERBLICK	121
8.4	EVALUIERUNGEN STEUERLICHER ANREIZE FÜR F&E	122
8.4.1	<i>Evaluierungskriterien</i>	122
8.4.2	<i>Quantitative Methoden der Evaluierung</i>	122
8.4.3	<i>Empirische Evidenz</i>	123
8.5	BEST PRACTICES DER EVALUIERUNG	124
8.5.1	<i>Australien: Kosten-Nutzen-Analyse</i>	124
8.5.2	<i>Niederlande: Ein pragmatischer Mix</i>	126
8.6	DIE INDIREKTE F&E-FÖRDERUNG IN ÖSTERREICH	127
8.6.1	<i>Der Forschungsfreibetrag vor der Steuerreform 2000</i>	127
8.6.2	<i>Vorschläge zur Reform der indirekten F&E-Förderung</i>	128
8.6.3	<i>Die Steuerreform 2000</i>	129
8.7	ANWENDUNG DER INTERNATIONALEN EVALUIERUNGSERFAHRUNGEN AUF ÖSTERREICH	129
8.7.1	<i>Die Ziele der steuerlichen F&E-Förderung in Österreich</i>	130
8.7.2	<i>Design/Operationalisierung</i>	131
8.7.3	<i>Wirkungsanalyse</i>	134
8.7.4	<i>Datenerfordernisse</i>	135
8.8	SCHLUSSFOLGERUNGEN	135
8.9	LITERATUR	137
	TEIL C	139
9.	PROGRAMM UND FOLIEN DER WORKSHOP	

Aufbau des vorliegenden Berichts

In jüngster Zeit wurden international erhebliche Anstrengungen unternommen, um zu einer Bestandsaufnahme und Bewertung der bisherigen Praktiken bei der Evaluierung technologiepolitischer Instrumente, Programme und Institutionen zu gelangen. Besonders hervorzuheben ist dabei der Versuch zur Identifikation von "good" oder "best practices" der Evaluierung. Auch in Österreich wurden diese internationalen Erfahrungen reflektiert.

Der vorliegende Projektbericht versteht sich als Beitrag zur weiteren Konkretisierung und Entwicklung der österreichischen Evaluierungspraxis. Er steht damit im **Kontext vielfältiger Bemühungen des Bundesministeriums für Verkehr, Technologie und Innovaton**, in Österreich eine systematische Evaluierungskultur für den FTE-Bereich aufzubauen.

Neben der vom Ressort selbst, von Joanneum Research, der Technologie Impulse Gesellschaft, dem österreichischen Forschungszentrum Seibersdorf und dem österreichischen Institut für Wirtschaftsforschung initiierten und organisierten **„Plattform Technologie Evaluierung“** sind hier folgende Forschungsprojekte zu nennen:

- „Evaluierung von Forschungs- und Technologiepolitik: 'Best Practices' in der Implementierung der Resultate von Evaluierungen“ (FTSP österreichisches Forschungszentrum Seibersdorf)
- „Evaluierung von innovativen Aktionen in der Technologie-, Struktur- und Arbeitsmarktpolitik: Entwicklung von Methoden, Indikatoren und „Good Practice“ (EvinA)“ (FTSP Joanneum Research)

Das Projektteam hat von Beginn an den engen Zusammenhang zwischen dem gegenständlichen Forschungsprojekt zum Thema „Quantitative Methoden der Evaluierung technologiepolitischer Programme“ und den beiden FTSPs thematisiert.

Aufgrund dessen fand im Rahmen des Projektes ein Workshop statt, bei dem die Ergebnisse der **beiden FTSPs** vorgestellt und diskutiert wurden. Programm und Folien des Workshops finden sich in **Teil C** des vorliegenden Berichts.

Im Herbst des Jahres 1999 wurde in Österreich eine mögliche Neugestaltung der steuerlichen Anreize für Forschung und Entwicklung diskutiert. Aus diesem Grund hatte der im Oktober 1999 gelegte Zwischenbericht des Projektes eine Diskussion **der Evaluierung indirekter F&E Förderung** zum Thema. Eine aktualisierte Version dieses Beitrags von Gernot Hutschenreiter (WIFO) findet sich in **Teil B** des vorliegenden Berichts.

Teil A bildet den Schwerpunkt des Berichts und behandelt quantitative Methoden der Evaluation technologiepolitischer Programme. Im folgenden wird ein Überblick gegeben:

Quantitative Evaluierungsmethoden im Überblick

Quantitative Evaluierungsmethoden können im wesentlichen in drei Gruppen eingeteilt werden:

Beschreibende und vergleichende statistische Analysen mittels quantitativer Science&Technology-Indikatoren (z.B. Patenthäufigkeit), generiert aus Unternehmensbefragungen, Daten des Programm-Monitorings oder sekundärstatistischem Datenmaterial, stellen die grundlegendste quantitative Methodengruppe dar.

Ökonometrische Modelle sind die vermutlich am weitesten verbreitete Methode zur Messung der ökonomischen Effekte von F&E-Programmen. Sie werden sowohl auf der Unternehmensebene wie auch der sektoralen und der gesamtwirtschaftlichen Ebene angewandt, um Zusammenhänge zwischen Eingriffen der öffentlichen Hand und den davon ausgehenden Wirkungen zu testen bzw. quantitativ abzuschätzen. Ökonometrische Methoden können in umfassende Modelle wirtschaftlicher und innovatorischer Prozesse eingebettet sein, welche die einzelnen Glieder der vermuteten Wirkungskette im Detail darstellen; oder aber sie beschränken sich auf eine (oder wenige) Schätzgleichung(en), um den Einfluss erklärender Variablen (z.B. die Teilnahme an einem technologiepolitischen Programm) auf einen Wirkungsindikator (z.B. Innovationsgrad des Unternehmens) zu testen.

Die **Kosten-Nutzen-Analyse** als methodischer Rahmen zur strukturierten Untersuchung technologiepolitischer Programme bringt besonders deutlich zum Ausdruck, worum es bei öffentlichen Förderungsmaßnahmen eigentlich gehen sollte: 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.

Der Schwerpunkt der Beiträge liegt bei jenen Methoden, die Quantifizierungen von Programmwirkungen im Rahmen eines (ökonometrischen) Modells oder eines Kosten-Nutzen-analytischen Ansatzes vornehmen. Das heißt, es werden vor allem Ansätze vorgestellt, die den Einfluss eines Programms auf quantitative oder quantifizierbare Variablen bei den programmteilnehmenden Firmen und Forschungseinrichtungen (oder auch darüber hinaus in der gesamten Volkswirtschaft) untersuchen.

Daher handelt es sich bei den vorgestellten Evaluationen in der Regel um **reine Wirkungsevaluationen**, d.h. andere Fragestellungen wie z.B. die Relevanz der Politikintervention, die Effizienz der Programmumsetzung u.ä. werden in der Regel nicht gehandelt.

Ein grundlegendes Problem aller quantitativen Methoden ist die **Datengrundlage**. Datenquellen können unterschieden werden nach:

- Primärerhebungen in der Regel in Form von **Teilnehmerbefragungen** (schriftlich oder in Interviewform)
- Nutzung von **Sekundärdatenbasen**: Sekundärdatenbasen liegen vor allem in aggregierter Form vor, und sind daher eher komplementär als substitutiv zu Primärerhebungen zu sehen. Für Mikrodaten sind teilweise Unternehmensdatenbanken nutzbar, allerdings unterliegen diese Datenbasen in der Regel statistischen Verzerrungen und geben daher die Grundgesamtheit der Unternehmen nicht repräsentativ wieder (vgl. Beitrag 1).

- Nutzung von Daten aus dem **Programm-Monitoring** (z.B. existieren für die Innovationsprogramme des DTI detaillierte Daten aus dem in Großbritannien verbindlichen ROAME Verfahren – vgl. hierzu Beitrag 7)
- Primärerhebung **im Kontext einer anderen, breiteren Primärerhebung**: Dies stellt im Prinzip einen Idealfall dar. Die für die Evaluation notwendigen Daten können in diesem Fall entweder direkt aus den Rohdaten der Primärerhebung genommen oder mit evaluationsspezifischen Zusatzerhebungen kombiniert werden. Auch lassen sich auf diese Weise relativ einfach und genau Kontrollgruppen definieren. Beispiele sind das Schweizer CIM Programm (vgl. Beitrag 2) oder das amerikanische GMEA (vgl. Beitrag 1).

Die Beiträge im Einzelnen

Der **erste Beitrag** (Kapitel 1: Evaluierung von Diffusionsförderprogrammen Beispiele aus der Analyse der Manufacturing Extension Partnership) einen Überblick über verschiedene quantitative Methoden der Evaluation von Förderprogrammen, welche die Diffusion neuer Technologien adressieren. Als Beispiel gilt das Manufacturing Extension Partnership (MEP) Programm, ein sehr umfangreiches und facettenreiches Diffusionsprogramm in den U.S.A., welches in jedem Bundesstaat unterschiedliche Schwerpunkte und Ausprägungen definiert. Ausgehend von einer modellhaften Darstellung des Diffusionsprozesses werden verschiedene Methoden der Evaluation mit ihren jeweiligen Vor- und Nachteilen vorgestellt, die alle bei der Evaluation des MEP – z.T. in verschiedenen Bundesstaaten – zum Einsatz kamen.

Das **dargestellte Modell** bildet Entscheidungsgrundlagen und Entscheidungsprozesse von Unternehmen ab und zeigt, welche Schritte im Prozess bzw. welche Variablen von den Wirkungen eines Diffusionsprogramms betroffen sind. Diese modellhafte Abbildung der Programmwirkung ist nicht nur ein hilfreicher analytischer Schritt bei der Konzeption einer Evaluation – dadurch werden Variablen definiert, auf die im Verlauf der Evaluation Bezug genommen werden - sondern auch notwendig für die Hypothesenbildung von Seiten der Programminitiatoren: Welche Wirkungen erwarte ich und welche Fragen habe ich aufgrund dessen an die Evaluation? Es zeigt sich hier bereits deutlich, dass die Wahl der Methode und die Spezifikation des Evaluationsmodells abhängig sind von den Hypothesen hinsichtlich der Wirkungsweise des Programms.

Bei der Darstellung der verschiedenen quantitativen Methoden liegt der Schwerpunkt auf **deskriptiven Analysen auf der Basis von Teilnehmerbefragungen** sowie auf **einfachen ökonometrischen Analysen**. An dieser Stelle wird deutlich, dass ökonometrische Schätzungen sehr genaue Spezifikationen der Variablen benötigen, damit die Programmwirkung, welche zu messen intendiert war, nicht überlagert wird von anderen Einflüssen.

Zu genau diesem Problem zeigt der **zweite Beitrag** (Kapitel 2: Evaluierung des CIM-Aktionsprogramms der Schweizer Bundesregierung) eine sehr exakte Lösung. Hier wird eine **mikroökonomische Schätzung** vorgestellt, die sehr genaue und detaillierte Aussagen hinsichtlich der direkten Wirkungen eines Diffusionsprogramms auf das Adoptionsverhalten der Zielgruppe erlaubt. Neben einer sehr überlegten Definition der Variablen ist ein Grund für die Robustheit und Exaktheit des Ansatzes, dass die abhängige Variable (Adoptionsneigung) rein technisch definiert wird, eine Wirkung auf ökonomische Größen wie Outputentwicklung und Beschäftigung wird nicht untersucht.

Avanciert ist der Ansatz in Bezug auf zwei Punkte: Erstens wird ein elaborierter Kontrollgruppenansatz vorgestellt, zweitens werden zwei Verhaltensgleichungen, eine „Adoptionsgleichung“ und eine „Politikgleichung“ simultan geschätzt. Dies erlaubt eine Abbildung von **Mitnahmeeffekten** bzw. eine Darstellung der Kausalität des Zusammenhangs zwischen Programmteilnahme und intendierter Verhaltensänderung.

Die dargestellten Vorteile des Ansatzes sind allerdings damit verbunden, dass viele Effekte unberücksichtigt bleiben, oder anders ausgedrückt, die **Aussagen sind zwar exakt aber**

begrenzt. Indirekte Effekte und zwar sowohl auf der Ebene der untersuchten Unternehmen (z.B. Lerneffekte, Wirkung auf Innovationsstrategie, Innovationsmanagement, auf Folgeinnovationen, Wettbewerbsposition etc.) als auch auf volkswirtschaftlicher Ebene wird nicht untersucht.

Die Anforderungen an Datenumfang und **Datenqualität** sind hier extrem hoch. Dieses Problem konnte im konkreten Fall etwas gemindert werden, da die Primärerhebung gemeinsam mit der allgemeinen Schweizer Innovationserhebung durchgeführt wurde und so die programmteilnehmenden Firmen lediglich einen Ergänzungsbogen zur Standardbefragung ausfüllen mussten.

Die folgenden vier Beiträge beschäftigen sich alle mit der **Ermittlung sozialer Erträge von Technologieprogrammen**. 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 Nutzens treten die folgenden prinzipiellen Probleme auf, welche unterschiedliche Ansätze unterschiedlich lösen:

- 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?
- 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?
- Welcher Anteil der – kosten- und nutzenstiftenden - technologischen Veränderung ist auf das Programm zurückzuführen?

Im **dritten Beitrag** (Kapitel 3: Evaluierung direkter Förderung am Beispiel des „Advanced Technology Programs“ ATP) wird zunächst das ausgebaute und sehr differenzierte Evaluierungssystem des amerikanischen „Advanced Technology Programs“ (ATP) beschrieben. Im Rahmen dieses umfassend und mit verschiedenen Ansätzen evaluierten Programms wurden auch Kosten-Nutzen-analytische Verfahren zur Anwendung gebracht. Diese werden an zwei Beispielen (Medizinische Technologien und Techniken zur digitalen Datenspeicherung) dargestellt, wobei versucht wird, den sozialen Ertrag der jeweiligen Projekte abzuschätzen. Dabei werden verschiedene Nutzungskategorien (Verbilligung der Produkte, neue Qualitäten etc.) systematisch identifiziert, den Kosten des Projekts und mit gegenübergestellt und mit verschiedenen Verfahren (Kapitalwert, interner Zinsfuß) „Gegenwartswerte“ berechnet.

Diese können dem privaten Ertrag gegenübergestellt und mit diesem verglichen werden. Unter anderen wurden dabei Methoden zur Abschätzung des Konsumentennutzens auch bei Fehlen einer empirischen Nachfragekurve ermittelt – und zwar mittels eines Kosten-Index-Verfahrens, das die Zahlungsbereitschaft der Konsumenten abschätzt.

Bei den untersuchten Projekten ergab sich in der Mehrzahl ein sozialer Ertrag, der substantiell über dem privaten lag, was ein Hinweis auf die Sinnhaftigkeit der öffentlichen Förderung durch das ATP ist.

Der Vorzug dieser Verfahren liegt in ihrer Systematik mit der Kosten und Nutzen verglichen werden. Da insbesondere bei der Abschätzung von Nutzen und Erträgen große Unsicherheiten bestehen, empfehlen sich Szenarien- bzw. Bandbreitenabschätzungen der zu erwartenden Nutzen. Mit diesen können dann verschiedene Projekte verglichen werden.

Der vierte Beitrag (Kapitel 4: Methoden der Evaluation ökonomischer Effekte von Technologieprogrammen: Die B.E.T.A. Methode zur Bewertung von ESA Programmen) beschäftigt sich mit der Messung ökonomischer Effekte von Raumfahrtprogrammen. Hier werden nicht die direkten Effekte des Programms ermittelt, sondern der „Sekundärnutzen“ berechnet, d.h. es wird der Versuch unternommen, die industrielle Nutzung von jenen F&E-Ergebnissen zu messen, die sich aus der Teilnahme an Raumfahrtprogrammen ergeben haben. Dieser Nutzen wird gemessen, indem für eine Vielzahl möglicher Effekte (Technologische Effekte, Kommerzielle Effekte, Effekte auf Organisation und Methoden und Effekte den Faktor Arbeit betreffend) in vielfältigen **Interviews mit dem Management** der programmteilnehmenden Organisation zwei Koeffizienten ermittelt werden: erstens die Wirkung des jeweiligen Effekts auf den Umsatz und zweitens der Anteil des Effektes, welcher auf die Programmteilnahme zurückzuführen ist.

Der Nutzen der Programmteilnahme wird damit gebildet als die Differenz zwischen der tatsächlichen Struktur der Unternehmen und einer hypothetischen Situation (ohne Programm). Damit wird ein ambitionierter Schritt weg von einem schlichten Vorher-Nachher-Vergleich hin zu einem **echten "Mit (hypothetisch) Ohne-Vergleich"** unternommen. Eine weitere Stärke des Ansatzes ist, dass die Erfassung und Befragung **unterschiedlicher Funktionsbereiche und Funktionsträger** innerhalb der betroffenen Firmen Einblick gewährt in die umfassenden firmeninternen und funktionsübergreifenden Auswirkungen technologischer Änderungen. Allerdings ist zu vermuten, dass einige Mängel der Methode zu einer deutlichen Überschätzung des Nutzenniveaus führen. Auch gehen in die Kalkulation nur die Effekte bei den Programmteilnehmern selbst sowie bei ihren direkten Lieferanten ein, andere **Wirkungen auf die Volkswirtschaft** bleiben vernachlässigt. So zeigt der Beitrag beispielhaft, dass eine Situation, bei welcher der Umsatz der programmteilnehmenden Organisationen unverändert bleibt oder sogar abnimmt, einen großen volkswirtschaftlichen Gesamtnutzen stiften kann.

In **fünften Beitrag** (Kapitel 5: Evaluierung von NIST-lab Programmen) wird ein sehr aufwendiger Ansatz zur Evaluation der ökonomischen Auswirkungen von Programmen des National Institute of Standards and Technology (NIST) vorgestellt. Diese Wirkungsanalyse stellt nur **eines von vier Evaluationselementen** des NIST dar, sie werden ergänzt um eine externe ExpertInnenbegutachtung (Peer Review), quantitative Output-Indikatoren und der regelmäßigen Erhebung der Kundenzufriedenheit.

Im Prinzip wird hier versucht, unter Verwendung von Sekundärdaten, Indikatoren und Befragungen den gesellschaftlichen Nutzen des Programms zu bestimmen und damit drei Standardwerte zu bestimmen: Die Internal Social Rate of Return, die Implied Rate of Return

und das Kosten-Nutzen-Verhältnis. Interessant ist die Ermittlung des Nutzens bei Herstellern und Anwendern durch die Fragestellung: „Wie hoch wäre Ihr (privater) Aufwand gewesen, wenn das NIST kein Programm durchgeführt hätte“? Damit kann – zumindest in Ansätzen – eine Aussage über die **relative Effizienz zwischen privaten und öffentlichen Investitionen** gemacht werden, wodurch bessere Informationen hinsichtlich der alternativen Verwendung öffentlicher Mittel zur Verfügung stehen als bei dem einfachen Vergleich unterschiedlicher Kosten-Nutzen-Relationen. Handelt es sich bei dem betrachteten Programm jedoch um Schaffung von Infrastruktur oder allgemein zugänglichem Wissen mit entsprechendem öffentlichen-Gut-Charakter, so ist diese Frage irreführend: In solchen Fällen wäre der kollektive Aufwand aller die Ergebnisse bzw. die Infrastruktur nutzenden Einrichtungen gemeinsam ein geeigneterer Referenzpunkt.

Der sechste Beitrag (Kapitel 6: Ein Modell zur Abschätzung der Kosteneffektivität: Die Evaluation von Innovationsförderprogrammen des Britischen Department of Trade DTI) stellt ein Modell zur **Abschätzung der relativen Kosteneffektivität** verschiedener Technologieentwicklungs- und Diffusionsprogramme des britischen Department of Trade and Industry (DTI) vor. Ausgehend von den Daten des in Großbritannien verbindlichen ROAME Verfahrens wurden Indikatoren definiert, die Hinweise auf Kosten und Nutzen der Förderprogramme des DTI geben: Für die Ermittlung der Kosten werden 5 Indikatoren definiert, beispielsweise die Kosten der Programmadministration als monetäre Kosten, die Anzahl der abgelehnten Anträge als nicht-monetäre Kosten. Die Nutzenkategorien gliedern sich in 5 Ergebniskriterien (z.B.: „Verwertung der Ergebnisse“, „Forschung und Entwicklung“), denen insgesamt 14 Indikatoren zugeordnet sind. Den einzelnen Indikatoren werden Werte von 0 bis 100 zugeordnet, um eine Vergleichbarkeit zwischen den Programmen zu erreichen – eine Monetarisierung wird hier nicht benötigt. Um ein eindeutiges Ranking der Programme hinsichtlich ihrer Kosteneffektivität zu erreichen, werden die einzelnen Kriterien gewichtet und zusammengezählt. Methodisch handelt es sich bei diesem Ansatz eigentlich nicht um ein Modell sondern um ein **Indikatoren-Aggregations-Normierungs- und -gewichtungsverfahren**, das **Vergleiche zwischen verschiedenen Programmen** ermöglicht. Unter zwei Voraussetzungen, nämlich erstens dem Vorhandensein einer entsprechend guten Datengrundlage aus dem Monitoring und zweitens ähnliche Zielsetzungen der Programme, so dass die Definition eines gemeinsamen Indikatorenbündels möglich ist, können mit dieser Methode zwei Aussagen getroffen werden, die üblicherweise mit Evaluationen unbeantwortet bleiben: Wo sind Schwachstellen bei einzelnen Programmen im Vergleich zu anderen zu orten und wie können öffentliche Mittel in bester Weise auf alternative Verwendungszwecke alloziert werden.

Der siebente Beitrag (Kapitel 7: COMEVAL Evaluierungstoolkit) schließlich beschäftigt sich mit dem **Evaluierungstoolkit COMEVAL**, das Mitte der 90er Jahre für die Europäische Kommission als allgemeine und einheitliche Evaluierungsmethodik für F&E Programme entwickelt wurde. Die Methode ist **hoch komplex** und hat sich wahrscheinlich auch aus diesem Grund nie breit durchgesetzt. Sie erlaubt die Beschreibung und graphische Darstellung von Zielen und Ergebnissen der F&E-Projekte und ermöglicht ein Nachvollziehen des „**Programmprofils**“ **über den gesamten Projektzyklus** vom Start bis zum Ende der Implementierungsphase mit Hilfe entsprechender Softwaretechnologien. Das Toolkit besteht aus acht verschiedenen Modulen, die von den NutzerInnen für ihre jeweiligen Bedürfnisse entsprechend unterschiedlich konfiguriert werden können. Es hat damit Ähnlichkeiten mit

einem ExpertInnensystem für Projektverfolgung und Monitoring. Projektergebnisse und – effekte werden im Rahmen eines Moduls auf Basis eines hierarchischen Indikatorensystems (Projektqualitätsindikatoren) identifiziert. Wenn auch gerade die Komplexität des Toolkits besticht, für eine konkrete Anwendung – gerade auf nationalstaatlicher Ebene – müssten deutliche Vereinfachungen und Anpassungen vorgenommen werden.

Programm	Land	Politik-orientierung	Ansatz / Inhalt	Bearbeitung
1. Manufacturing Extension Programme (MEP) - und regionale Sub-Programme	USA	Diffusion	Überblick über verschiedene Methoden - Deskriptive Analyse von Teilnehmerbefragungen - Einfache mikroökonomische Analyse Modellhafte Darstellung des Diffusionsprozesses	Wolfgang Pointner (Joanneum)
2. Advanced Manufacturing Technologies – Computer-Integrated Manufacturing	CH	Diffusion	Avancierte mikroökonomische Schätzung - Kontrollgruppenansatz - Abbildung von Mitnahmeeffekten	Christian Rammer (Seibersdorf)
3. Advanced Technology Programme (ATP), - diverse Subprogramme	USA	F&E-Ausgaben (direkt)	Fallstudien (Soziale) Kosten-Nutzen-Analyse	Wolfgang Polt, Birgit Woitech (Joanneum)
4. ESA	INT	Missionen	Kosten-Nutzen-Analyse basierend auf Mikrodaten (Interviews) Messung von „Spin-offs“ – ökonomischer Sekundärnutzen von Raumfahrtprogrammen	Dorothea Sturn (Joanneum)
5. NIST – Labs - Unterprogramm „Alternative Kältemittel“	USA	Institutionen Technologie- transfer	Aufwendiger Evaluationsansatz mit 4 Elementen, eines davon Wirkungsanalyse Bestimmung des gesellschaftlichen Nutzens von Programmen (Kosten-Nutzen-Analyse)	Toni Geyer (Seibersdorf)
6. National Audit Office / DTI	UK		Verwendung von Indikatoren aus dem Programm-Monitoring (Indikatoren-Aggregations-Normierungs- Gewichtungsverfahren) Abschätzung der relativen Kosteneffektivität Vergleich verschiedener Programme	Toni Geyer (Seibersdorf)
7. COMEVAL	EU	Evaluierungs Tool Kit	Avanciertes, kompliziertes Tool Kit Erfordernisse an Monitoring	Toni Geyer (Seibersdorf)

TEIL A

1 EVALUIERUNG VON DIFFUSIONSFÖRDERPROGRAMMEN BEISPIELE AUS DER ANALYSE DER MANUFACTURING EXTENSION PARTNERSHIP

*Wolfgang Pointner
Joanneum Research*

1.1 Die Manufacturing Extension Partnership (MEP)

In den USA existierten Ende der 90er Jahre 380.000 SMEs mit weniger als 500 Beschäftigten. Diese Unternehmen haben aufgrund ihrer meist geringeren Ressourcenausstattung (insbes. Kapital oder Qualifikation der Mitarbeiter) häufig Schwierigkeiten beim Zugang zu den neuesten technischen Entwicklungen ihrer Branchen, aber auch zu Management- oder Organisationsinnovationen, durch die ihre Produktionsprozesse effizienter und damit wettbewerbsfähiger gestaltet werden können. Um diese Zugangsbeschränkung auszugleichen initiierte das National Institute of Standards and Technology (NIST) in Kooperation mit einzelnen bundesstaatlichen Behörden die Manufacturing Extension Partnership (MEP)¹. Im Jahr 1995 beliefen sich die Bundesausgaben für die MEP auf 138,4 Millionen USD, die Gesamtaufwendungen betragen zumindest das Doppelte dieser Summe, da die Auflage besteht, dass einzelne Staaten oder Kommunen die MEP in gleicher Höhe unterstützen wie die Bundesverwaltung.

Die MEP unterhält ein Netzwerk von 400 Manufacturing Technology Centres (MTCs) mit Niederlassungen in jedem Bundesstaat, dem District of Columbia und Puerto Rico. Die MTCs beschäftigen Experten mit einschlägiger Berufserfahrung und stellen den SMEs Beratung und Entscheidungshilfen zur Verfügung, in so unterschiedlichen Arbeitsfeldern wie:

- Prozessautomatisierung
- Qualitätsmanagement
- Betriebsorganisation
- Weiterbildung von Mitarbeitern
- Marketing
- Energiesparmaßnahmen oder
- E-Commerce

Durch ihre bundesweite Verteilung sind die MTCs für alle SMEs leicht erreichbar und verfügen über Informationen über das regionale/lokale Wirtschaftsgeschehen. Die Vernetzung mit dem NIST hat den Vorteil, dass auch komplexere Fragestellungen, die über die Kompetenz des lokalen Beraterteams hinausgehen, von sachkundigen Experten beantwortet werden können. Die Kontakte des NIST ermöglichen den SMEs auch den Zugang zu den technologischen Innovationen von staatlichen Forschungseinrichtungen. Auch bei der betrieblichen Implementierung solcher Innovationen stehen die MTCs beratend zur Seite.

Im Bundesstaat Georgia wurde im Rahmen der MEP 1994 die Georgia Manufacturing Extension Alliance (GMEA) eingerichtet. Das führende Institut in der Organisation der GMEA ist das Georgia Institute of Technology (Georgia Tech), das über eine langjährige Erfahrung im Bereich der industriellen Diffusionsförderung verfügt. Als operative Zielsetzung der GMEA wurden folgende Punkte definiert:

- Vergrößerung des Klientenkreises durch Eröffnung neuer Büros in den Clusterregionen Georgias
- Erweiterung des Leistungsangebots auf Basis einer Erhebung der Bedürfnisse von betroffenen Unternehmen
- Erleichterung des Zugangs zu neuen Technologien durch Bildung integrierter Netzwerke.

GMEA war 1997 landesweit in 18 Regionalbüros tätig, die dabei erbrachten Leistungen entsprechen den Arbeitsfeldern der MTCs. Bereits in den ersten beiden Jahren des Bestehens von GMEA nahmen 21% aller sachgüterproduzierenden Unternehmen in Georgia diese Leistungen in Anspruch. Da Georgia Tech auch im Bereich von Evaluierungen tätig ist, wurde schon in der Anfangsphase der GMEA ein ambitioniertes Evaluierungskonzept erstellt und Sorge getragen, dass die nötigen Daten und Informationen gesammelt wurden.

1.2 Dynamisches Modell des Entscheidungsprozesses von Marktteilnehmern

Ausgangspunkt dieses Vergleichs verschiedener Evaluierungsmethoden ist ein Modell von Jarmin und Jensen aus 1997, das den bei Unternehmen abbilden soll, die miteinander im Wettbewerb stehen. Auch wenn Unternehmen der selben Branche angehören oder die selben Technologien anwenden, besitzen sie noch genügend relevante Eigenschaften, die ganz unterschiedlich ausgeprägt sein können und einen entscheidenden Einfluss auf den Diffusionsprozess haben. Dazu zählen einerseits beobachtbare Merkmale wie die Anzahl der Beschäftigten, Investitionsvolumen oder eingegangene Kooperationen mit Forschungsinstituten, die sich zumeist auch sinnvoll quantifizieren lassen; andererseits gibt es aber auch qualitative Unterschiede zwischen Unternehmen, die sich nur schwer feststellen und kaum oder gar nicht messen lassen, wie etwa die Managementqualitäten der Unternehmensleitung oder die Aufgeschlossenheit der Mitarbeiter gegenüber betrieblichen Innovationen. Von all diesen Eigenschaften werden die Fähigkeiten der Informationsgewinnung und -verarbeitung im Modell besonders behandelt. Neben den firmenspezifischen Charakteristika lassen sich Unternehmen auch noch nach den Strategien, die sie verfolgen, und „idiosynkratischen Schocks“, also unvorhersehbaren, betriebsinternen Veränderungen im Produktionsprozess, die nur das einzelne Unternehmen betreffen, unterscheiden.

Die Heterogenität der Unternehmen führt dazu, dass diese im Falle eines allgemeinen Schocks unterschiedlich betroffen sind und auch unterschiedlich reagieren. Die Implementierung eines öffentlichen Förderprogramms kann in diesem Zusammenhang als allgemeiner Schock verstanden werden. Daher ist es interessant, bei der Evaluierung solcher Programme Daten auf Unternehmens- oder Betriebsebene zu betrachten, da durch eine Aggregation interessante Aspekte der Heterogenität verloren gehen.

¹ Eine detailliertere Darstellung findet sich auf der MEP-Homepage: www.mep.nist.gov.

Als theoretischer Rahmen der Evaluierung fungiert ein dynamisches Modell des Entscheidungsprozesses in einem Unternehmen. Am Beginn einer Periode verfügen die Unternehmen über ein Set an Informationen, auf deren Grundlage sie Erwartungen bilden über vergangene Ereignisse, die ihnen nicht bekannt sind oder solche, die noch nicht eingetreten sind. Informationen und Erwartungen beeinflussen das Verhalten der Unternehmen, dieses Verhalten wiederum wirkt sich auf andere Unternehmen aus, auf die vorhandene Wissensbasis und auf den „state of the world“ ganz allgemein. Die Informationen über diese Auswirkungen werden dann zur Grundlage im Entscheidungsprozess der nächsten Periode.

Die Informationen I_{it} , über die ein Unternehmen i am Anfang der Periode t verfügt, setzen sich zusammen aus dem Vektor Y_{it-1} , der die Charakteristika des Unternehmens i zum Zeitpunkt t enthält. Zu diesen Charakteristika zählen Umsätze, Beschäftigte, Gewinne, Branchenzugehörigkeit, Anzahl der Niederlassungen, etc. In diesem Modell wird davon ausgegangen, dass jedes Unternehmen über die eigenen Charakteristika vollständig informiert ist.

Der Vektor Y_{jt-1}^i enthält alle Daten, die das Unternehmen i über die Charakteristika anderer Unternehmen sammeln kann. Dieser Vektor ist nicht so detailliert wie Y_{it-1} , da das Unternehmen i als Außenstehender nicht alle Charakteristika der anderen Unternehmen beobachten kann. Darüber hinaus erhält ein Unternehmen auch nicht die vollständige Einsicht in die betrieblichen Abläufe seiner Konkurrenten und daher ist die Wahrnehmung der Daten in Y_{jt-1}^i mit einem Fehler behaftet. Die Werte in Y_{jt-1}^i sind also nur Näherungswerte der tatsächlichen Realisierungen von Y_{jt-1} . Eine andere Teilmenge von Y_{jt-1} wird von öffentlichen Einrichtungen wie Forschungsinstituten oder Statistikämtern aufgezeichnet. Diese Daten Y_{jt-1}^g stehen bei der Evaluierung von staatlichen Förderprogrammen zur Verfügung.

Ein Unternehmen ist auch vollständig über die Aktionen und Entscheidungen, die in der letzten Periode getroffen wurden, informiert. Dieser Vektor A_{it-1} enthält Informationen über Preissetzung, Investitionshöhe und andere strategische Entscheidungen. Auch die Entscheidungen der anderen Unternehmen A_{jt-1}^i können beobachtet werden, allerdings gilt auch hier wieder, dass diese Beobachtungen nicht unter der Annahme der vollständigen Information, sondern nur verzerrt durchgeführt werden.

Weiters sammeln die Unternehmen Daten über den Zustand der Welt S_t sowie das gesamte verfügbare technische Wissen K_t der jeweiligen Periode t . Die Informationen eines einzelnen Unternehmens i S_t^i und K_t^i sind wieder nicht vollständig, sondern nur fehlerbehaftete Teilmengen der tatsächlichen Ausprägungen. Die Entscheidungsdynamik eines Unternehmens wird durch folgendes Optimierungsproblem modelliert:

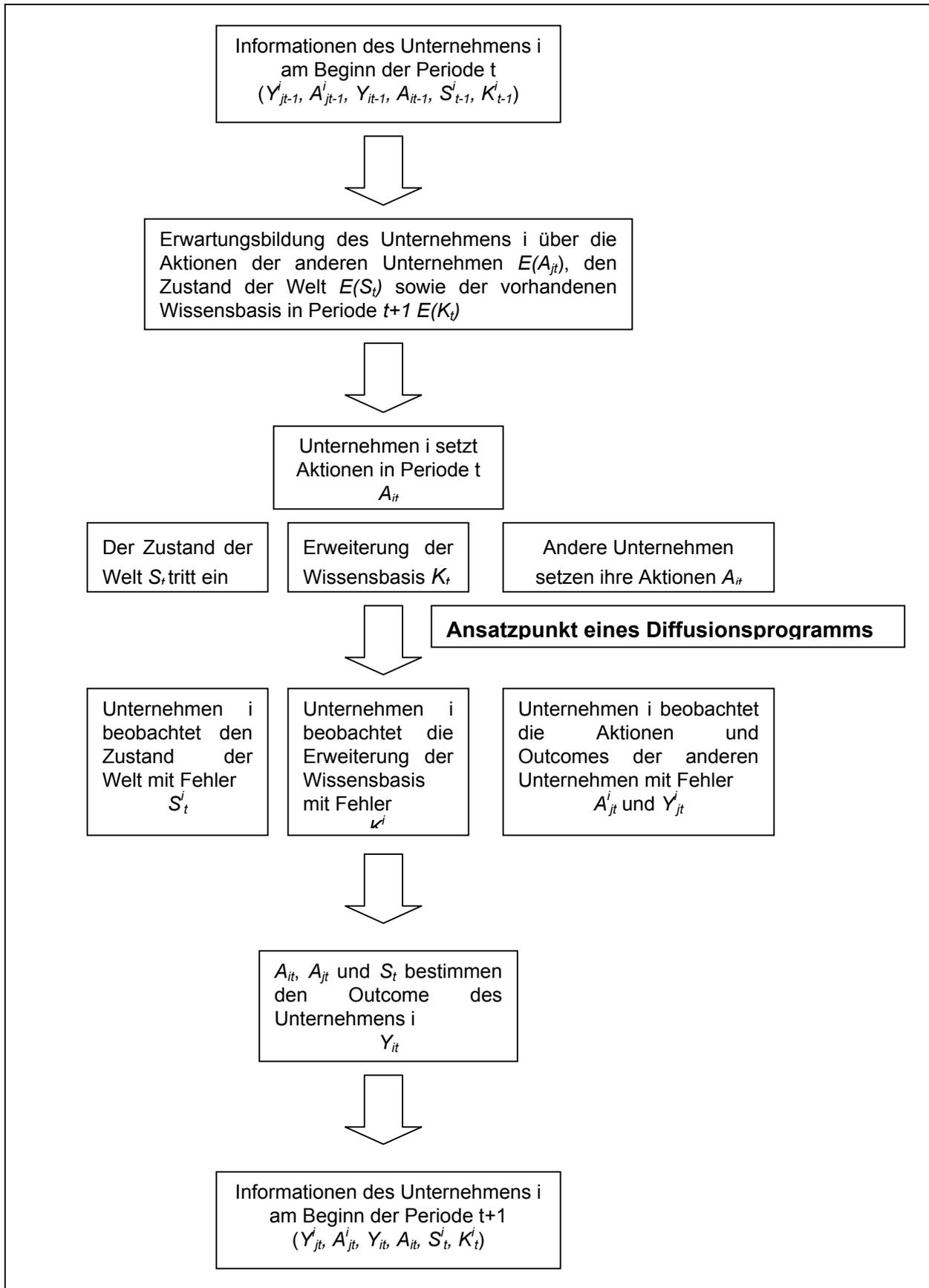
$$\begin{aligned} \max V_{it} &= \pi_{it}(A_{it}, A_{jt}, S_t, Y_{it}, I_{it}) + \delta_i V_{it+1} \\ \text{s.t. } Y_{it} &= f(Y_{it-1}, A_{it-1}, A_{jt-1}, S_{t-1}, K_{it}) \\ I_{it} &= g(I_{it-1}, Y_{jt-1}^i, A_{jt-1}^i, Y_{it-1}, A_{it-1}, S_{t-1}^i, K_{t-1}^i) \\ K_t &= h(K_{t-1}, A_{it}, A_{jt-1}, S_t) \end{aligned}$$

Dabei steht V_{it} für den Wert der Profitströme, π_{it} für die Gewinnfunktion des Unternehmens i in Periode t , und δ_i für den Diskontfaktor, mit dem die zukünftigen Gewinne auf den Gegenwartswert abgezinst werden. Die Funktionen f und g beschreiben, wie sich die

Charakteristika bzw. die Informationsmenge des Unternehmens über die Zeit entwickelt, in h wird die Entwicklung der gesamten verfügbaren Wissensbasis abgebildet.

Die MEP als diffusionsorientiertes Programm setzt bei der Perzeption der Unternehmen bezüglich der technischen Wissensbasis K_t an, diese wird im Rahmen dieses Modells durch die Variable K_t^i dargestellt. Das Programmdesign geht davon aus, dass SMEs strukturell einen erschwerten Zugang zu neuen Technologien und produktionsrelevanten Innovationen haben. Die Verbesserung von K_t^i soll dann über eine Steigerung der technologischen Leistungsfähigkeit zu einer besseren Wettbewerbssituation der Betriebe führen. Um diesen Impact zu überprüfen, müssen Kriterien definiert werden, die als Indikator der Programmeffektivität dienen können. Diese Kriterien Y_{it}^p bilden eine Teilmenge von Y_{it}^s , den öffentlich beobachtbaren Charakteristika der Unternehmen. Im angesprochenen Fall der MEP-Evaluierung wäre ein mögliches Kriterium, das als Indikator der Wettbewerbsfähigkeit dienen und aus öffentlichen Statistiken abgeleitet werden kann, die Veränderung der Produktivität (Arbeitsproduktivität, Kapitalproduktivität oder TFP) im Zeitraum der Programmteilnahme und unmittelbar danach.

Abbildung 1: Diffusion und betrieblicher Entscheidungsprozess



1.3 Ziele und Verfahren von Evaluierungen

Die den MEP-Evaluierungen zugrunde liegenden Problemstellungen zielen einerseits in Richtung Wirkungsanalyse und andererseits darauf ab, die Qualität der Programmabwicklung zu beurteilen. Hinsichtlich der Wirkungsanalyse ergeben sich dabei folgende Fragen: In welchem Ausmaß hat ein Programm die damit verbundenen Erwartungen erfüllt? Besteht zwischen der Verwirklichung bestimmter Ziele und den im Programm eingesetzten Mitteln ein kausaler Zusammenhang? Hat die Implementierung des Programms zu nicht-intendierten Effekten geführt? Die Qualität der Programmabwicklung lässt sich am ehesten an Kriterien festmachen wie: Zufriedenheit der Teilnehmer insgesamt, Schwierigkeiten bei der Beantragung der Teilnahme oder Kosten des Programmmanagements.

Die Antworten auf all diese Fragen ergeben ein gutes Bild von allen Aspekten der Effektivität und der Effizienz eines Förderprogramms, allerdings muss bei jedem Evaluierungsschritt noch geklärt werden, inwieweit die vorgesehene oder durchgeführte Untersuchungsmaßnahme einen hinreichend zuverlässigen Eindruck der tatsächlich erreichten Effekte liefert. Dies hängt ganz wesentlich von der jeweils angewandten Methode ab, ob diese der Problemstellung und/oder den verfügbaren Daten- und Informationsquellen adäquat ist.

Die angesprochenen Themenkomplexe, die in einer Evaluierung behandelt werden, korrespondieren auch mit den Interessen der beteiligten Institutionen. Während die mit der operativen Abwicklung befassten eher an der qualitativen Beurteilung ihrer Arbeit interessiert sind und daran, ob und wie stark ihre Tätigkeit auch direkte Effekte in den beteiligten Unternehmen ausgelöst hat, liegt der Fokus auf politischer Ebene mehr an der Wirkungsanalyse und der Messung der Zielerreichung.

Im folgenden sollen nun fünf Evaluierungsmethoden dargestellt werden, die sich sowohl methodisch als auch in den zur Durchführung benötigten Daten unterscheiden, wobei sich die einzelnen Verfahren nicht gegenseitig ausschließen, sondern teilweise sogar aufeinander aufbauen.

Tabelle 1

Methoden	Vorteile	Nachteile
Vollzugskontrolle	<ol style="list-style-type: none"> 1. Genaue Informationen über die erbrachten Leistungen 2. Vergleichbarkeit zu anderen Datensätzen 3. Voraussetzung für Teilnehmerbefragung 	<ol style="list-style-type: none"> 1. Keine Kontrollgruppe 2. Beschränkung auf die vom Projektträger erbrachten Leistungen 3. Kaum Vergleiche von verschiedenen Projekten möglich
Fallstudien	<p>Genaue Informationen über die Wirkung der erbrachten Leistungen auf die untersuchten Teilnehmer</p>	<ol style="list-style-type: none"> 1. Keine Kontrollgruppe 2. Oft nur Darstellung von „success stories“
Teilnehmerbefragung	<ol style="list-style-type: none"> 1. Fragen können mit den Interessen auf allen Programmebenen abgestimmt werden 2. Konkretes Feedback für die operative Ebene 	<ol style="list-style-type: none"> 1. Keine Kontrollgruppe 2. Statistische Verzerrung bei Rücklauf 3. Teuer
Ökonometrische Analysen	<ol style="list-style-type: none"> 1. Kontrollgruppe 2. Nutzung externer Datenquellen (billig) 3. Impactanalyse möglich 	<ol style="list-style-type: none"> 1. „Selection bias“ 2. Relevante Daten müssen in externen Quellen vorhanden sein oder teuer selbst beschafft werden
Experimentelle Methoden	<ol style="list-style-type: none"> 1. Kontrollgruppe 2. Kein „Selection bias“ 	Nicht praktikabel

1.3.1 Monitoring der Vollzugskontrolle

Die Daten, die bei der Vollzugskontrolle eines Programms gesammelt werden, bilden oft einen ersten Orientierungsrahmen bei einer Evaluierung. Dazu zählen etwa Informationen über den rechtlichen Status der involvierten Einheiten oder deren Adressen, Aufzeichnungen über die von den Programmabwicklern erbrachten Leistungen sowie deren Kosten. Diese Art von Daten ergibt sich oft schon aus Bewerbungsunterlagen, Vertragsabschlüssen oder der Ausführung von programmspezifischen Aufgaben durch die abwickelnden Institutionen. Die Erhebung dieser Daten ist also mit sehr geringen zusätzlichen Kosten verbunden.

Dem Vorteil der geringen Erhebungskosten von solchen Verwaltungsdaten steht der Nachteil gegenüber, dass bei der Aufzeichnung solcher Daten nicht oder kaum an ihre spätere Verwendung im Zusammenhang mit einer Evaluierung gedacht wird. Bei reinen Statusvariablen wie Wohnort oder Betriebsform spielt dieser Unterschied keine Rolle, sehr wohl aber bei der Beurteilung der in einzelnen Projekten erbrachten Leistungen. Aus verwaltungstechnischen Motiven werden dabei meist nur Art und Umfang der Aktivitäten aufgezeichnet. Dies ist zwar eine notwendige, aber noch keine hinreichende Datenbasis für eine Evaluierung, da den Aufwendungen auch die damit erzielten Wirkungen

gegenübergestellt werden müssen, bzw. erst einmal die Frage geklärt werden muss, ob das in Frage stehende Leistungsangebot überhaupt einen kausalen Effekt hatte.

Die Vollzugskontrolle wird üblicherweise auf Ebene der Projekte oder der Teilnehmer durchgeführt. Das hat zur Folge, dass oft für die einzelnen Projekte sehr heterogene Datensätze vorliegen, was natürlich die Evaluierung des gesamten Programms nicht gerade erleichtert. Die Möglichkeit der standardisierten Erhebung von Verwaltungsdaten hängt aber von der Homogenität der Projekte und damit auch vom Programmdesign ab.

1.3.2 Fallstudien

In einer Fallstudie werden die Beziehungen und einzelne Arbeitsschritte zwischen Programmteilnehmern und den mit der Abwicklung betrauten Einheiten detailliert aufgezeigt. Dieser Evaluierungstyp wird meist auf Projektebene eingesetzt und beinhaltet eine Darstellung der Ausgangslage des betreffenden Teilnehmers, der von den Abwicklern gesetzten Aktivitäten und der dadurch induzierten Veränderungen sowie deren Auswirkungen. In der eingangs skizzierten Modellstruktur ließe sich eine Fallstudie aus dem MEP-Programm so darstellen, dass ausgehend von der Anfangsausstattung I_{it} eines Unternehmens dieses Informationsset durch die Teilnahme am Programm erweitert wird und zwar besonders hinsichtlich des verfügbaren technischen Wissens K_t . Diese Erweiterung des Informationsstands und die daraus resultierenden Aktivitäten A_{it} des Unternehmens werden in einer Fallstudie dokumentiert. Aus der Untersuchung von erfolgreichen Teilnehmern lassen sich möglicherweise konkrete Hinweise ableiten, welche Instrumente eines Programms für bestimmte Situationen besonders geeignet sind. Die öffentlichkeitswirksame Darstellung solcher „success stories“ dient auch der Bestärkung teilnahmewilliger Unternehmen, sich für eine Partizipation zu entscheiden².

Mit dieser Werbewirksamkeit von Fallstudien ist auch einer ihrer Nachteile verbunden: es werden kaum Projekte untersucht, die einen geringen oder gar einen negativen Einfluss auf die teilnehmenden Unternehmen hatten. Das liegt zum Teil auch daran, dass Fallstudien mit sehr hohem Aufwand für alle Beteiligten verbunden sind, und daher nur eine Auswahl der durchgeführten Projekte auf diese Weise untersucht wird. Bei dieser Auswahl kann es leicht zu einer positiven Verzerrung kommen. Da im Zuge einer Programmevaluierung meist nur Fallstudien von teilnehmenden Unternehmen durchgeführt werden, lassen sich deren Ergebnisse nicht mit einer Kontrollgruppe kontrastieren.

1.3.3 Teilnehmerbefragung und deskriptive Auswertung

Eine häufig eingesetzte Evaluierungsmethode ist die Erhebung mittels Fragebogen. Dabei werden die Unternehmen während oder nach ihrer Teilnahme an einem Programm nach ihren Erfahrungen befragt. Abhängig von den zur Verfügung stehenden Mitteln und der Größe des Programms wird die Grundgesamtheit aller Teilnehmer einbezogen oder nur eine Stichprobe daraus. Eine Fragebogenerhebung erlaubt das detaillierte Abklären von unterschiedlichen Aspekten eines Programms, die mit den unterschiedlichen Interessen der verschiedenen Programmverantwortlichen korrespondieren. Die Durchführung einer solchen

² Vgl. die Beschreibung solcher „success stories“ auf der Homepage des MEP.

Befragung ist per Telefon möglich oder auf postalischem Weg, die Auswertung erfolgt zumeist computerunterstützt. Da in Teilnehmerbefragungen oft auch eine qualitative Bewertung der im Programm erbrachten Leistungen enthalten ist, sollte die Aussendung der Fragebögen nicht durch eine programmverantwortliche Partei erfolgen, sondern durch den externen Evaluator, da sich ansonsten eine Verzerrung im Antwortverhalten der Teilnehmer ergeben könnte.

Die Ergebnisse einer Teilnehmerbefragung sind ein wichtiges Feedback für die Programmverantwortlichen auf der operativen Ebene. Je nach erfolgter Fragestellung lässt sich mehr oder weniger deutlich überprüfen, wie die eingesetzten Instrumenten von den Unternehmern beurteilt werden und ob das Angebotsspektrum auch mit den Bedürfnissen der Teilnehmer übereinstimmt. Daraus können Indikatoren für die Effizienz der Programmabwicklung abgeleitet werden.

Auch auf der Ebene der strategischen Zielsetzungen können Teilnehmerbefragungen sinnvoll eingesetzt werden. Wenn zwischen Programmteilnahme und Befragung ein gewisser Zeitraum verstrichen ist, können die Veränderungen der relevanten Zielvariablen erhoben werden. Die Bestimmung des optimalen Zeitraums zwischen Programmteilnahme und Befragung unterliegt einem trade-off, da nur wenige Programme bereits nach einigen Monaten zu signifikanten Ergebnissen führen werden, andererseits nach ein paar Jahren die Informationen über den Programmverlauf in den teilnehmenden Betrieben häufig nur mehr rudimentär vorhanden sein werden. Außerdem sinkt mit zunehmenden Umfang des Fragebogens die Bereitschaft, diesen auch zu beantworten.

Da die Rücklaufquote einer Befragung kaum 100% ausmachen wird, besteht die Gefahr, dass es bei der Auswertung zu einem „recall bias“ kommt. Um dem vorzubeugen, ist es erforderlich, durch eine non-response-Analyse die Übereinstimmung der erhaltenen Fragebögen mit der Grundgesamtheit abzutesten.

Ein weiterer Nachteil von Teilnehmerbefragungen ist das Fehlen einer Kontrollgruppe; diese wäre wichtig, um die reinen Programmeffekte zu schätzen, indem die Ergebnisse der Teilnehmer denen der Kontrollgruppe gegenübergestellt werden. Um diesem Problem zu begegnen, könnte die Fragebogenerhebung auch auf eine Stichprobe von Nicht-Teilnehmern ausgedehnt werden. Dadurch erhöht sich aber der Aufwand der Evaluierung deutlich.

Zur Auswertung einer Teilnehmerbefragung steht eine Reihe von Methoden zur Auswahl. Oft werden in einem ersten Schritt deskriptive Verfahren (z.B.: Darstellung der Häufigkeitsverteilungen) angewendet, da diese mit den gängigen Statistikprogrammen einfach durchzuführen sind und bereits einen guten Überblick über die gesammelten Daten vermitteln. In der Evaluierung der GMEA wurden sowohl die erwarteten als auch die tatsächlich eingetretenen Wirkungen abgefragt, die Antworten wurden in einer Matrix zusammengefasst. Die möglichen Programmwirkungen bilden die Spalten, die Art der durchgeführten Projekte (=Instrumente) ist in den Zeilen aufgelistet. Über die Zeilen wurde die durchschnittliche Wirkung auf 1 normiert, sodass überdurchschnittliche Auswirkungen eines Instruments mit einem Wert größer 1 und unterdurchschnittliche mit einem Wert kleiner 1 repräsentiert werden. Auf diese Weise wird die Effektivität der eingesetzten Instrumente

auf Unternehmensebene deutlich und die operative Ebene kann die Auswirkungen ihrer Aktivitäten besser einschätzen.

Tabelle 2

Projektart/Instrument	Umsatz- steigerung	Kapital erhöhend	Lager- ersparnis	Kosten- ersparnis	Neue Jobs schaffend	Jobs erhaltend	Zeit intensiv
EDV	0,90	1,41	2,55	1,21	1,02	1,22	1,21
Betriebsgestaltung	1,18	1,20	1,23	1,22	1,34	1,28	0,95
Umweltschutz	0,35	0,86	0,30	0,78	0,38	0,78	0,57
Qualifizierung	0,80	0,75	1,29	1,18	1,10	1,54	0,82
Marketing	1,66	0,65	0,21	0,07	2,20	0,80	0,64
Materialtests	0,65	0,81	0,26	0,73	0,81	0,50	0,57
Management	1,37	1,15	2,17	1,10	0,85	2,27	0,76
Prozessverbesserung	1,24	1,37	1,18	1,07	0,96	0,80	1,02
Energie	0,27	1,34	0,36	1,59	0,34	0,35	0,81
Produktentwicklung	1,64	0,87	0,35	0,73	1,18	0,67	1,01
Qualitätskontrolle	1,09	0,67	1,09	1,05	1,07	0,87	1,99

Da die Realisierung von Programmwirkungen eine bestimmte Zeitspanne benötigt, deren Länge von der Art der eingesetzten Instrumente abhängt, ist es ratsam, eine Wirkungsanalyse nicht bereits während oder unmittelbar nach Abschluss eines Programms durchzuführen. Bei der GMEA-Evaluierung wurden die Teilnehmer zweimal befragt, einmal etwa einen Monat nach dem Ende ihrer Teilnahme und noch einmal etwa ein Jahr später. Durch dieses zweistufige Verfahren können einerseits bereits eingetretene Wirkungen des Programms bewertet werden und andererseits lassen sich so die Erwartungen der Unternehmen mit den tatsächlich eintretenden Effekten vergleichen. Nach einem Jahr kann auch schon erhoben werden, wie viele der Teilnehmer aufgrund der Beratungen durch die GMEA entsprechende Aktivitäten gesetzt haben. Von den 1994 teilnehmenden Unternehmen gaben 85% unmittelbar nach ihrem Ausscheiden aus dem Programm an, dass sie beabsichtigen, ihre neuen Erfahrungen auch umzusetzen. Nach einem Jahr hatten 68% konkrete Handlungen unternommen und 12% waren noch unentschlossen ob oder wie sie diese Umsetzung in Angriff nehmen sollten.

1.3.4 Ökonometrische Analysen

Durch eine Evaluierung mittels ökonometrischer Schätzung sollen die Effekte eines Programms auf die Teilnehmer hinsichtlich bestimmter Zielvariablen im Vergleich zu einer Kontrollgruppe bewertet werden. Durch einfaches Vergleichen der Entwicklung der Zielvariable bei den Teilnehmern und der Kontrollgruppe kann nicht auf den Effekt eines Programms geschlossen werden, da sich die beiden Gruppen in Bezug auf andere relevante Einflussgrößen ebenfalls unterscheiden können, sodass ein Unterschied auch darin begründet sein kann.

Um eine ökonometrische Analyse durchführen zu können, ist es wichtig, möglichst viele dieser relevanten Merkmale eines Unternehmens Y_{it} aufzuzeichnen. Der Zusammenhang zwischen einem dieser Merkmale, der abhängigen Variable Y^p_{it} , und dem Vektor der

unabhängigen oder erklärenden Variablen X_{it} wird in einer Schätzung bestimmt, die dabei ermittelten Koeffizienten geben an, wie stark der Einfluss einer einzelnen Variable X_{it} auf Y_{it}^p ist. Die Auswahl der relevanten erklärenden Variablen sowie die bei der Spezifikation der Schätzgleichung postulierte Kausalität lassen sich aus den Modellen der ökonomischen Theorie ableiten.

Als abhängige Variable wird dasjenige Merkmal („Performance“) in die Schätzung eingehen, das mit dem zu evaluierenden Programm gefördert werden sollte; wenn sich dieses Merkmal nicht eindeutig messen lässt, wird ein Indikator an dessen Stelle treten. Wichtige erklärende Variable, die der Vektor X_{it} in jedem Fall enthalten sollte, sind Angaben über die Größe der betreffenden Unternehmen, ihre Branchenzugehörigkeit oder den Unternehmensstandort. Die verwendete Schätzgleichung könnte z.B. folgende lineare Form haben:

$$Y_{it}^p = \alpha + \beta X_{it} + \gamma Z_{it} + \varepsilon_{it}$$

Die Variable Z_{it} in obiger Gleichung ist ein Indikator für die Programmteilnahme. Bezüglich Messung der Teilnahme stehen mehrere Varianten zur Auswahl, Z_{it} könnte eine Dummy-Variable sein, die im Fall der Nicht-Teilnahme den Wert 0 annimmt und bei Teilnahme des Unternehmens i den Wert 1. Es lassen sich aber auch mehrere Ausprägungen von Z_{it} modellieren, die verschiedene Intensitäten der Teilnahme oder der im Programm erbrachten Leistungen widerspiegeln. Je besser die operative Abwicklung eines Programms in den Daten der Vollzugskontrolle dokumentiert ist, desto genauer kann auch die Teilnahmevariable modelliert werden. Der Koeffizient γ misst den Effekt der Programmteilnahme Z_{it} auf die Zielvariable Y_{it}^p ; ε_{it} ist ein Störterm.

Ökonometrische Analyse des GMEA

Für das GMEA wurde 1996 eine derartige ökonometrische Evaluierung durchgeführt (Youtie, Shapira, 1997). Basierend auf dem Georgia Manufacturing Survey, einer Vollerhebung aller Betriebe mit mehr als 10 Beschäftigten, wurde aus den Daten der 1002 antwortenden Betriebe eine Schätzgleichung berechnet, in der die Veränderung der Wertschöpfung pro Beschäftigten zwischen 1994 und 1996 als abhängige Variable Y_{it}^p fungierte. Als erklärende Variablen wurden einerseits die Teilnahme an einem GMEA-Projekt Z_{it} und andererseits eine Reihe von Unternehmensmerkmalen X_{it} herangezogen. Diese Merkmale umfassten:

- Die Wachstumsrate der Beschäftigten 1994-1996
- Die Veränderungsrate von 1994-1996 bei denjenigen Beschäftigten, die an ihrem Arbeitsplatz Computer oder andere programmierbare Maschinen verwendeten
- Ob es sich um ein Ein-Betriebs-Unternehmen handelte (Dummy)
- Zweistellige Branchenklassifikation (Dummy)
- Beschäftigtengrößenklassen (Dummy)
- Ob der Betriebssitz in einem Ballungsraum lag (Dummy)
- Ob im Verwaltungsbezirk des Betriebs ein GMEA-Büro lag (Dummy)
- Leistungen einer privaten Beratungsfirma (Dummy)
- Leistungen einer anderen öffentlichen Einrichtung als GMEA (Dummy)

- Kooperationen mit anderen Betrieben in Bezug auf Produktentwicklung, Design, Weiterbildung, Qualitätssicherung oder Marketing

Die statistischen Ergebnisse dieser Studie sind in folgender Tabelle zusammengefasst:

Tabelle 3

Variable	Koeff.	Signifikanz
Wachstumsrate der Beschäftigten	-0,1008	0,01
Veränderungsrate von Beschäftigten mit Computer	-0,0003	
GMEA-Klient	0,0026	0,1
Ballungsraum	0,0051	0,01
Verwaltungsbezirk mit GMEA-Büro	-0,0016	
Private Beratungsfirma	-0,0033	0,01
Anderen öffentlichen Einrichtung als GMEA	-0,0008	
Kooperationen mit anderen Betrieben	-0,0003	
Ein-Betriebs-Unternehmen	-0,0041	0,01
Textil	0,0071	0,01
Holz	-0,0065	0,05
Möbel	-0,0074	0,05
Elektronik	0,0058	0,1
Metallverarbeitung	0,0061	0,01
1994 – 10-19 Beschäftigte	0,0025	
1994 – 20-49 Beschäftigte	0,0036	0,05
1994 – 50-99 Beschäftigte	-0,0001	
1994 – 100-249 Beschäftigte	0,0053	0,01

Diese Werte lassen sich so interpretieren, dass die Teilnahme an einem GMEA-Projekt die Wachstumsrate der Wertschöpfung pro Beschäftigten um 0,26% erhöht. Dabei ist zu bedenken, dass es sich beim MEP um ein sehr umfassendes Diffusionsprogramm handelt, das auch Leistungen in Bereichen wie Betriebsorganisation oder Marketing anbietet. Programme mit einem engeren Fokus, etwa auf die Diffusion bestimmter Prozesstechnologien, könnten kaum in diesem Rahmen untersucht werden. Ziel solcher Programme ist im wesentlichen die Verbreiterung der Wissensbasis von Unternehmen und die Beschleunigung der Adoption neuer Technologien. Bis eine erfolgreiche Diffusion zu Effekten wie etwa einer höheren Wertschöpfung pro Beschäftigten führt, vergeht tendenziell eine größere Zeitspanne, in der das erworbene Wissen oder die neue Technologie erst einmal in den betrieblichen Produktionsprozess implementiert werden muss.

Außerdem ist die Steigerung der Wertschöpfung auch von betrieblichen Charakteristika wie eben Organisation oder Marketing abhängig, die aber von solchen Programmen weitgehend unbeeinflusst bleiben und sich nur schwer in einem quantitativen Modell abbilden lassen³. So könnte die Teilnahme an einem Technologieprogramm auch ein Indikator für die besseren

³ Vgl. Mowery, 1995, S. 516

Managementfähigkeiten der Betriebsführung sein und der Koeffizient der Schätzgleichung würde dann auch den Effekt dieser höheren Fähigkeiten messen und nicht nur die Wirkung der Programmteilnahme. Der Einsatz eines ökonometrischen Verfahrens ist also nur dann sinnvoll, wenn zwischen der fraglichen abhängigen Variable und der Programmteilnahme ein möglichst direkter kausaler Zusammenhang gegeben ist. Andernfalls besteht die Möglichkeit, dass der in einer Schätzung berechnete Effekt der Programmteilnahme zum Teil auch die Wirkungen anderer Faktoren beinhaltet und daher überschätzt wird. Dies ist wegen der bereits angesprochenen Schwierigkeit, qualitative Merkmale wie die Managementfähigkeit sinnvoll zu messen und in ein ökonometrisches Modell zu integrieren, nicht unwahrscheinlich.

1.3.5 Verwendung von Sekundärdatenbasen

Bei der Analyse eines bestimmten Programms kann eine Kontrollgruppe aus der Gesamtheit aller Nicht-Teilnehmer definiert werden, um aus dieser Gruppe eine Zufallsstichprobe zu ziehen und diese dann mit den Teilnehmern zu vergleichen. Diese Vorgangsweise ist allerdings mit hohem Aufwand verbunden und stellt daher nur für ausreichend dotierte Evaluierungen einen gangbaren Weg dar.

Eine Alternative zur teuren Ausweitung der Erhebung auf Nicht-Teilnehmer besteht in der Verwendung bereits bestehender Datenbanken auf Einzelfirmenbasis. Informationen dieser Art werden von öffentlichen Einrichtungen gesammelt, diese sind aber aus Datenschutzgründen nur sehr selten in der Lage, solche Informationen für eine Evaluierung zu Verfügung zu stellen. Andere, kommerzielle Datenbanken unterliegen nicht so strengen Bestimmungen, da Unternehmen, die ihre Daten an solche Institutionen weitergeben, damit auch ihr Einverständnis erklären, dass diese Daten an Dritte übermittelt werden. Durch diese Art der Informationsbeschaffung kann es aber wieder zu einer statistischen Verzerrung, einem „selection bias“, kommen, d.h. die Verteilung der relevanten Merkmale ist bei Unternehmen, die diese Daten freiwillig an Dritte weitergeben, nicht notwendigerweise identisch mit der Verteilung in der Grundgesamtheit aller Unternehmen. Beispiel für solche kommerzielle Datenbanken sind das Performance Benchmarking Service (PBS) des Industrial Technology Institute (ITI) in den USA oder die Marketingdatenbank der Creditreform in Österreich.

Im PBS werden eine Reihe von Daten erhoben, die auch für Programmevaluierungen relevant sind. Firmen, die diese Daten an das ITI weitergeben, erhalten für eine Gebühr ein „benchmarking“, d.h. ihre Leistung hinsichtlich der angegebenen Kriterien wird mit den anderen Unternehmen in der Datenbank verglichen und bewertet. Da durch die MEP auch die Teilnahme am PBS-Benchmarking gefördert wird, enthält diese Datenbank Informationen über MEP-Teilnehmer sowie eine Kontrollgruppe von nichtgeförderten Unternehmen. Analysen aufgrund dieser Datenbestände wurden bereits durchgeführt, etwa durch Nexus Associates (1996) für die New Yorker MEP. Bei einem Vergleich zwischen MEP-Teilnehmern und anderen Unternehmen auf Basis der PBS-Daten muss allerdings sehr genau auf die Zusammensetzung der Kontrollgruppe geachtet werden, da sich in diesem Fall leicht ein „selection bias“ ergeben könnte.

Der angesprochene „selection bias“ führt bei einer ökonometrischen Analyse zu einer verzerrten Schätzung des Teilnahmeeffekts, der Wert des Koeffizienten γ wird also entweder

zu hoch oder zu niedrig ausgewiesen. Die Marketingdatenbank z.B. enthält nur Daten von Unternehmen, die bonitätsgeprüft sind, d.h. bei diesen Unternehmen wurde nie eine Geschäftsverbindung auf Kreditbasis abgelehnt. Die Teilnahme an einem Programm wird nicht von solchen Bedingungen abhängig gemacht, daher wäre eine Kontrollgruppe aus diesen Datenbeständen nur eingeschränkt verwendbar.

Eine weitere Datenquelle für Vergleichsstudien ist die Longitudinal Research Database (LRD) des US Census Bureau. In der LRD werden die Daten aus der fünfjährlich durchgeführten Vollerhebung „Census of Manufactures“ mit denen aus den jährlich stattfindenden „Annual Surveys of Manufactures“ verknüpft; der Umfang der an einem „Annual Surveys of Manufactures“ (ASM) beteiligten Unternehmen liegt etwa bei 15% der Gesamtpopulation. Unternehmen aus der Sachgüterproduktion mit mehr als 250 Beschäftigten werden in jedem ASM befragt, diese machen allerdings nicht mehr als 20% der ASM-Stichprobe aus. Aus den restlichen, kleineren Unternehmen wird eine nach Branchen- und Größenverteilung repräsentative Stichprobe ausgewählt; diese Stichprobe wird alle 5 Jahre neu gezogen. Da in der MEP eher kleinere Unternehmen gefördert werden, ist die Tauglichkeit der LRD für statistische Kontrollzwecke eingeschränkt, allerdings setzt sich die Stichprobe der Unternehmen mit weniger als 250 Beschäftigten aus etwa 40.000 Einzelbeobachtungen zusammen und bei dieser Anzahl sollte es schon möglich sein, zu signifikanten Ergebnissen zu kommen. Mit Hilfe der LRD können so unterschiedliche Variable verglichen werden wie die Entwicklung der Produktivitäten, Umsätze, Gewinne, Beschäftigung, Löhne oder die Exporte.

Da es sich bei der LRD um eine Einzelfirmendatenbank handelt, sind die enthaltenen Daten mit besonderer Vertraulichkeit zu behandeln. Selbst bei anonymisierten Einzelfirmendaten kann aufgrund relativ weniger Informationen (z.B.: Lage, Größe und Branche) oft schon auf das Unternehmen geschlossen werden. Um hier Missbrauch im Sinne des Datenschutz möglichst auszuschließen, werden diese Daten nicht freigegeben, sondern können nur im Center for Economic Studies (CES) des Census Bureau eingesehen werden, und zwar von speziell dafür vereidigten Wissenschaftlern, die allerdings nicht notwendigerweise Mitarbeiter des CES sein müssen.

1.3.6 Experimentelle Methoden

Die Vorteile eines experimentellen Evaluierungsdesigns sind vielversprechend: durch die Miteinbeziehung einer Kontrollgruppe ist eine quantifizierbare Beurteilung der Programmwirkung möglich und außerdem kann auch davon ausgegangen werden, dass sich bei der Analyse kein „selection bias“ ergibt. Bedauerlicherweise sind die grundlegenden Anforderungen an die verwendete Datenbasis so hoch, dass im Bereich der Programmevaluierung noch kaum jemand von diesen Methoden Gebrauch machen konnte.

Die hypothetische Vorgangsweise bei einer experimentellen Programmevaluierung ist einfach und gleichzeitig sehr effektiv. Das Evaluierungsteam müsste nur aus Grundgesamtheit aller Firmen eine Zufallsstichprobe von zukünftigen Programmteilnehmern sowie eine ebenfalls auf Zufall basierende Kontrollgruppe auswählen. Nachdem die auf diese Art bestimmten Teilnehmer das Programm durchlaufen haben, wird die in Frage

stehende Zielvariable des Programms, Y_{it}^p , zwischen den beiden Gruppen verglichen. Dabei wäre es schon völlig ausreichend, die Differenz der Mittelwerte $E(Y_{it}^p)$ aus beiden samples zu betrachten. Da sowohl die Teilnehmer als auch die Kontrollgruppe nach dem Zufallsprinzip ausgewählt wurden, ist statistisch sichergestellt, dass sich die Stichproben vor der Programmteilnahme nicht systematisch in irgendeiner Hinsicht unterscheiden. Alle Unterschiede in der Zielvariable können daher auf den Umstand der Programmteilnahme zurückgeführt werden.

Die konkrete Durchführung einer experimentellen Evaluierung ist leider mit großen Schwierigkeiten und hohem Aufwand verbunden, für die Analyse eines diffusionsorientierten Programms ist dieses Design nicht praktikabel. Durch die zufällige Auswahl der Teilnehmergruppe aus der Grundgesamtheit besteht eine relativ hohe Wahrscheinlichkeit, dass dabei Firmen ausgesucht werden, die nicht zur Zielgruppe des Programms zählen und die daher keinen oder nur einen geringen Zusatznutzen durch die angebotenen Förderungen erfahren. Eine Einschränkung der Grundgesamtheit auf Firmen, die potentielle Förderfälle des Programms wären, unterliegt einem trade off zwischen einem höheren Wirkungsgrad des Programms einerseits und der Abnahme der statistischen Aussagekraft dieses Evaluierungsdesigns auf der anderen Seite, da bei einer sehr genauen Vorselektion der zur Grundgesamtheit zählenden Firmen zwar die Programmeffekte tendenziell deutlicher aufgezeigt werden können, aber möglicherweise nur mehr so wenige Unternehmen zur Ziehung der Stichproben bleiben, dass keine signifikanten Aussagen mehr zulässig sind.

Eine weitere Schwierigkeit betrifft die Frage der Freiwilligkeit. Eine wichtige Voraussetzung, damit ein Förderprogramm auch seine vollen intendierten Effekte entfalten kann, ist die zustimmende Haltung möglichst aller direkt beteiligten Mitarbeiter des Unternehmens. Im Regelfall wird es vor einer Programmteilnahme im Unternehmen zu einer Abstimmung der relevanten Personen kommen. Bei einer zufälligen Zuteilung ist dies natürlich nicht durchführbar, und es besteht daher die Möglichkeit, dass wichtige Mitarbeiter der Teilnahme indifferent oder sogar ablehnend gegenüberstehen, wodurch sich keine umfassende Programmwirkung erreichen ließe.

Ein weiteres Argument gegen die Anwendbarkeit experimenteller Evaluierungsmethoden liegt wohl in dem damit verbundenen Aufwand. Das zu untersuchende Programm müsste in vollem Umfang für die fiktive Teilnehmergruppe durchgeführt werden, obwohl die oben angeführten Gründe gegen die Sinnhaftigkeit der Förderung von nach dem Zufallsprinzip ausgewählten Firmen sprechen. Aufgrund der hohen Kosten eines Technologieprogramms würde eine derartige Untersuchung jedes herkömmliche Evaluierungsbudget sprengen.

Anmerkung

Zum Abschluss sei noch auf ein Problem bei der Evaluierung von Förderprogrammen hingewiesen, das in allen bisher vorgestellten Wirkungsanalysen mittels Kontrollgruppe auftritt. Bei der Auswertung wird davon ausgegangen, dass alle Unterschiede zwischen den Teilnehmern und der Kontrollgruppe hinsichtlich der Zielvariable auf die beobachteten Merkmale, und darunter ganz besonders auf die Programmteilnahme, zurückzuführen sind. Dieser Ansatz vernachlässigt, dass die Entscheidung zur Teilnahme selbst bereits ein wesentlicher Unterschied zwischen diesen beiden Gruppen ist. Es besteht die Möglichkeit, dass sich die Teilnehmer von den Nicht-Teilnehmern systematisch unterscheiden, und zwar in solchen Merkmalen, die sich der Beobachtung entziehen. Programmteilnehmer könnten bspw. über besonders gute Kontakte zu Programmverantwortlichen verfügen, die betriebsinternen Entscheidungsträger könnten Förderungen oder grundsätzlich Innovationen gegenüber aufgeschlossener sein oder andere in einer quantitativen Analyse schwer zu überprüfende Unterschiede könnten eine Rolle spielen. Im Falle eines ökonometrischen Ansatzes würde sich die Schätzgleichung auf folgende Form erweitern:

$$Y_{it}^p = \alpha + \beta^b X_{it}^b + \beta^u X_{it}^u + \gamma Z_{it} + \varepsilon_{it}$$

wobei der Superskript b für „beobachtbar“ und der Superskript u für „unbeobachtbar“ steht. Die unbeobachtbaren Einflussfaktoren X_{it}^u sind von der Art wie oben angeführt und beeinflussen nicht nur die Zielvariable Y_{it}^p sondern auch die Teilnahmeentscheidung.

Mittels einer experimentellen Evaluierung könnte der Einfluss solcher unbeobachtbarer Merkmale ausgeblendet werden, allerdings ist fraglich, wie sinnvoll ein derartiges Vorgehen wäre, da bei der tatsächlichen Programmabwicklung genau diese Faktoren eine Rolle spielen.

1.4 Literatur

- Jarmin, Ronald S., „Using Matched Client and Census Data to Evaluate the Performance of MEP“, CES-Working Paper 95-7, (1995).
- Jarmin, Ronald S., „Manufacturing Extension and Productivity Dynamics“, CES-Working Paper 98-8, (1998).
- Jarmin, Ronald S., Jensen, J. Bradford, „Evaluating Government Technology Programmes: The Case of Manufacturing Extension“, in „Policy Evaluation in Innovation and Technology“, OECD Proceedings, (1997).
- Mowery, David, „The Practice of Technology Policy“ in „Handbook of the Economics of Innovation and Technological Change“, Hrsg. V. Paul Stoneman, (1995), S. 513-557.
- Shapira, Philip, Youtie, Jan, „Evaluating Technology Deployment at the State Level: Methods, Results and Insights from GMEA“ in „Policy Evaluation in Innovation and Technology“, OECD Proceedings, (1997).
- Shapira, Philip, Youtie, Jan, „Georgia Manufacturing Extension Alliance: Overview of the Evaluation Plan“ GMEA Evaluation Working Papers E9401, (1994).
- Youtie, Jan, Shapira, Philip, „Manufacturing Needs, Practices and Performance in Georgia: 1994-1998“ GMEA Evaluation Working Papers E9401, (1997).

2 EVALUIERUNG DES CIM-AKTIONSPROGRAMMS DER SCHWEIZER BUNDESREGIERUNG

*Christian Rammer
Austrian Research Centers Seibersdorf*

2.1 Einleitung

Staatliche Aktivitäten zur Förderung der Diffusion von neuen Technologien zielen in der Regel darauf ab,

- die **Anwendung** dieser Technologien in einer Volkswirtschaft zu **intensivieren** (d.h. die Zahl der Nutzer einer neuen Technologie über jenes Maß zu erhöhen, das sich auf der Basis von individuellen Entscheidungen unter den gegebenen Markt- und institutionellen Bedingungen ergibt);
- die **Anwendung** dieser Technologien in einer Volkswirtschaft zu **beschleunigen** (d.h. den Zeitpunkt der erstmaligen Anwendung einer neuen Technologie durch einen Nutzer vorzulegen gegenüber jenen Zeitpunkt, der sich auf der Basis von individuellen Entscheidungen unter den gegebenen Markt- und institutionellen Bedingungen ergibt);
- das **Anwendungsspektrum** dieser Technologien zu **erweitern** (d.h. neue Einsatzgebiete für eine neue Technologie ausfindig machen und zur Einsatzreife zu bringen).

Der Einsatz öffentlicher Förderung zur Diffusion neuer Technologien leitet sich aus der Annahme ab, dass die durch Marktstrukturen und institutionelle Rahmenbedingungen gegebenen Anreize für die potenziellen Nutzer neuer Technologien zu einem individuellen Adoptionsverhalten führen, das zu einer Technologiediffusion unter dem gesellschaftlich gewünschten Niveau führt – d.h. die erwarteten privaten Erträge liegen unter den gesellschaftlichen Erträgen bei einer breiten Diffusion (vgl. OECD 1997; Fritsch 1995). Ein niedriges Niveau der Technologiediffusion kann über verschiedene Wirkungskanäle zu einem ineffizienten Einsatz der Technologie insgesamt und zu Wohlfahrtsverlusten führen (etwa wenn Investitionen in die Qualifikation von Arbeitskräften, die für die Anwendung einer bestimmten Technologie benötigt wird, wegen eines zu geringen Verbreitungsgrads der Technologie unterbleiben; oder wenn Netzwerkökonomien in der Nutzung einer Technologie aufgrund eines geringen Verbreitungsgrads nicht lukriert werden können; oder wenn die geringe Nachfrage nach der neuen Technologie zur Einstellung weiterer Forschungs- und Entwicklungsarbeit bei den Technologieproduzenten führt).

Als Gründe für eine geringe Adoptionsneigung gegenüber neuen Technologien werden unter anderem angeführt:

- Informationsdefizite über die Möglichkeiten neuer Technologien,
- hohe Umstellungskosten bei Unsicherheiten über die künftige technologische Entwicklung,
- Defizite im Innovationsmanagement (Konzipierung und organisatorische Umsetzung der Innovation),
- mangelnde Absorptionskapazitäten für die Anwendung neuer Technologien (fehlende FuE-Kapazitäten, Mangel an qualifiziertem Personal),

- Akzeptanzprobleme in der Belegschaft,
- zu hohe Finanzierungskosten (bzw. Erwartung über künftig fallende Anschaffungskosten der Technologie),
- mangelnde Kompatibilität mit den bisher eingesetzten Technologien.

Staatliche Aktivitäten zur Förderung der Technologiediffusion versuchen in der Regel, bei einem oder mehreren der genannten Hemmfaktoren für die Technologieadoption anzusetzen. Typische Förderungsmaßnahmen sind – je nach Technologie und Anwendergruppe - die Bereitstellung von Informations- und Beratungsleistungen, die Einrichtung von Awarenessprogrammen (inklusive Öffentlichkeitsarbeit und Abbau von Akzeptanzbarrieren bei spezifischen Zielgruppen), die Qualifizierung von Beschäftigten, die (Mit-)Finanzierung des Technologieankaufs und der Technologieimplementierung sowie die Herstellung von Kontakten mit Technologieentwicklern (inklusive der finanziellen Unterstützung von unternehmensspezifischen Entwicklungsaktivitäten zur Einführung der Technologie).

Die Umsetzung dieser Förderungsmaßnahmen kann in unterschiedlichen institutionellen Arrangements vorgenommen werden. Eine in vielen entwickelten Industrieländern gängige Form ist die Einrichtung von Förderungsprogrammen für bestimmte Technologien, die deren Entwicklung und Verbreitung unterstützen sollen. Diese „Technologieprogramme“ haben in der Regel eine zeitlich begrenzte Laufzeit, eine definierte Zielgruppe und ein bestimmtes Spektrum an Förderungsmaßnahmen. Beispiele für solche Technologieprogramme finden sich unter anderem in Deutschland (vgl. Wengel et al. 1995; Becher und Kuhlmann 1995), der Schweiz (vgl. Arvanitis et al. 1998), Österreich (vgl. Polt et al. 1994), den USA (vgl. Jarmin und Jensen 1997; Shapira et al. 1996) und zahlreichen anderen OECD-Ländern (vgl. Arnold und Guy 1997).

Die Evaluierung von diffusionsorientierten Technologieprogrammen adressiert im Rahmen der drei zentralen Evaluierungsaufgaben folgende spezifische Fragen:

1. *Relevanz des Programms*: Liegt die dem Programmdesign zugrundeliegende Annahme eines Marktversagens tatsächlich in der vom Programm unterstellten Form vor? Ist die Zielgruppe des Programms gut gewählt? Treffen die im Programm gesetzten Maßnahmen die Hemmfaktoren der potenziellen Adoptoren? Lassen die durch das Programm intendierten Verhaltensänderungen eine Steigerung der sozialen Erträge erwarten? Sind von den durch das Programm angepeilten Veränderungen in den Anreizstrukturen negative (nicht intendierte) Effekte zu erwarten?
2. *Effizienz des Programms* (bzw. der Programmumsetzung): Wurde die Zielgruppe erreicht? In welchem Ausmaß nahm die Zielgruppe am Programm teil? Wurde das Programm ausreichend bekannt gemacht? Stehen die Aufwendungen für das Programmmanagement und die Programmumsetzung in einem vernünftigen Verhältnis zu den Programmaktivitäten und –zielen? Hätten die am Programm teilnehmenden Akteure die neuen Technologien auch ohne Förderungen eingeführt (Mitnahmeeffekte)?
3. *Effektivität des Programms*: Stimmen die Motive für die Programmteilnahme mit den angenommenen Hemmfaktoren der Technologieadoption überein? Haben Akteure, die am Programm teilgenommen haben, die Anwendung der Technologie ausgeweitet, intensiviert bzw. beschleunigt? Liegt die Anwendung, Intensivierung bzw.

Beschleunigung der Technologieadoption durch am Programm teilnehmende Akteure über dem über dem Durchschnitt einer vergleichbaren Kontrollgruppe?

Im folgenden wird ein Ansatz vorgestellt und diskutiert, auf den zwei der angeführten Fragen eine Antwort zu geben versucht: Im Rahmen einer mikroökonomischen Untersuchung des Adoptionsverhaltens der Akteursgruppe Unternehmen wird gleichzeitig überprüft, ob die Programmteilnahme einen positiven Effekt auf das Adoptionsverhalten (Ausweitung der Technologieadoption im Vergleich zur Kontrollgruppen nicht-teilnehmender Unternehmen) und ob die Programmteilnahme ein auslösender Faktor für diese Verhaltensänderung war.

2.2 Das CIM-Aktionsprogramm der Schweizer Bundesregierung

Das Aktionsprogramm „Computer-integrated Manufacturing“ (CIM-Aktionsprogramm) der Schweizer Bundesregierung, für dessen Evaluierung der hier vorgestellten Evaluierungsansatz entwickelt und angewandt wurde, ist als diffusionsorientiertes Förderungsprogramm konzipiert, dessen Ziel die Stärkung der Wettbewerbsfähigkeit der Wirtschaft durch die Unterstützung bei der Modernisierung der Produktionstechnologien war. Das Programm wurde von der „Kommission für Technologie und Innovation“ (KTI) abgewickelt. Die KTI ist die wichtigste Förderungsinstitution der Technologiepolitik auf Bundesebene in der Schweiz, über die sowohl die Projektförderungen als auch die Impuls- und Aktionsprogramme des Bundes größtenteils abgewickelt werden.

Das CIM-Aktionsprogramm kann an Hand der folgenden Merkmale charakterisiert werden (vgl. Dreher und Balthasar 1997; Arvanitis et al. 1998, 2f):

- *Zielgruppe*: Unternehmen des produzierenden Gewerbes, insbesondere Klein- und Mittelunternehmen und Unternehmen, die CIM noch nicht einsetzen
- *Förderungsinstrumente*: Informations- und Schulungsmaßnahmen für Unternehmen (Nachdiplomstudien, Weiterbildungskurse), Beratungsmaßnahmen (Unternehmensanalyse, Erstellung von Konzepten zur CIM-Einführung), finanzielle Unterstützung für technologische Entwicklungsprojekte (Einführung neuer Fertigungstechnologien)
- *Umsetzung des Programms*: Einrichtung von sieben regionalen CIM-Zentren, in denen Bildungs- und Beratungsmaßnahmen durchgeführt wurden und die teilweise sehr eng an technikorientierte Lehranstalten (insbesondere HTL) gekoppelt waren
- *Laufzeit*: 1990 bis 1996
- *Finanzierungsvolumen*: 102 Mio. SFR, davon 82 Mio. für CIM-Zentren
- *Zahl der teilnehmenden Unternehmen*: 116 geförderte CIM-Adoptoren (zusätzlich ca. 61.600 Schulungsteilnehmer und 3.700 Beratungen in den CIM-Zentren)

Von der konzeptionellen Seite war das Programm durch ein „ganzheitliches“ Herangehen an die Frage der CIM-Einführung gekennzeichnet. Neben technologischen Aspekten wurde auch nicht-technologischen Einflussgrößen wie der Arbeitsorganisation, der MitarbeiterInnenmotivation und –schulung sowie betriebswirtschaftlichen und organisatorischen Aspekten eine erhebliche Bedeutung für den Erfolg der CIM-Einführung beigemessen. Dies kommt auch durch die Kombination der drei Förderungsansätze Schulung, Beratung und Technologieentwicklung zum Ausdruck.

Das Programm hatte neben der Erhöhung der unternehmerischen Wettbewerbsfähigkeit als weiteres Ziel die Etablierung einer dezentralen Struktur des Wissens- und Technologietransfers in der Schweiz. Die im Rahmen des Aktionsprogramms eingerichteten regionalen CIM-Zentren sollten nach Ende des Programms auf Dauer selbständig funktionsfähig bleiben und ein Beratungs-, Schulungs- und technologisches Unterstützungsangebot darstellen, das Unternehmen bei der Einführung neuer Technologien zur Seite steht und Hemmnisse bei der Technologiediffusion abbauen hilft.

2.3 Der Evaluierungsansatz

Der Evaluierungsansatz der Forschungsgruppe des Instituts für Konjunkturforschung der ETH Zürich ist in einer Reihe von Veröffentlichungen dokumentiert (vgl. Arvanitis et al. 1998, Arvanitis, Hollenstein und Lenz 1999; Arvanitis und Hollenstein 1997, 1999). Ziel des Ansatzes ist es, für den konkreten Fall des Schweizer Programms zur Förderung der Anwendung von „Computer-Integrated Manufacturing“-Technologien (CIM) zu prüfen,

- ob die über das Programm geförderten Unternehmen rascher und umfassender CIM einsetzen als nicht-geförderte Unternehmen (*Zielerreichungsanalyse*) und
- welche Richtung ein etwaiger positiver Zusammenhang zwischen einer Förderung einerseits und einer überdurchschnittlich hohen Adoptionsneigung in Bezug auf CIM andererseits hat (*Kausalitätsanalyse*).

Der Evaluierungsansatz ist **mikroökonomischer Natur** und beobachtet die vom Programm anvisierte Verhaltensänderung innerhalb der Zielgruppe (mit der zu überprüfenden Hypothese, dass eine Programmteilnahme eine positive Verhaltensänderung, d.h. eine Erhöhung der Adoptionsneigung, bewirkt). Er ist als **Ex-post-Evaluierung** konzipiert und beobachtet die Verhaltensänderung zwischen den beiden Zeitpunkten vor Programmstart und nach Programmende. Die Datenbasis sind Informationen zu Unternehmen, die am Programm teilgenommen haben und solchen, die nicht teilgenommen haben (Kontrollgruppe). Die Analyse erfolgt als **Querschnittsanalyse** über alle erfassten Unternehmen.

Die formale Grundstruktur des Evaluierungsansatzes besteht aus zwei Verhaltensgleichungen:

- a. „**Adoptionsgleichung**“ zur Abbildung der Einflussfaktoren auf die Wahrscheinlichkeit, dass ein Unternehmen innerhalb des Beobachtungszeitraums (überhaupt, rascher und/oder stärker) CIM einsetzt:

$$A_i = a_1 + a_2 X_i + a_3 F_i + a_4 I_i \quad (1)$$

- b. „**Politikgleichung**“ zur Abbildung der Einflussfaktoren auf die Wahrscheinlichkeit, dass ein Unternehmen am Förderungsprogramm Teil nimmt:

$$P_i = b_1 + b_2 Y_i + b_3 A_i + b_4 I_i \quad (2)$$

wobei A die Adoptionsneigung in Bezug auf CIM im Beobachtungszeitraum (= Laufzeit des Programms) bezeichnet, \mathbf{X} ein Vektor, der jene Variable umfasst, die als wesentliche Einflussfaktoren auf die Adoptionsneigung eines Unternehmens angesehen werden, \mathbf{F} ein Vektor von Variablen, die Form und Umfang der Förderung abbilden, I das Ausmaß des CIM-Einsatzes zu Beginn des Programms, P die Teilnahme am Förderungsprogramm und i die Beobachtungseinheit (Unternehmen). a und b sind Parameter.

Die abhängige Variable A kann die Adoptionsneigung prinzipiell auf unterschiedlichen Messniveaus und mit Hilfe unterschiedlicher Indikatoren erfassen. In Bezug auf den Einsatz von CIM sind statische Indikatoren auf dichotomem (Anwendung ja/nein), ordinalem (Intensität des CIM-Einsatzes gering/mittel/stark) oder metrischem Messniveau (Zahl der eingesetzten CIM-Elemente, Anteil von mit Hilfe von CIM hergestellten Produkten am Umsatz) oder dynamische Indikatoren auf jedem der drei Messniveaus (Veränderung der Intensität des CIM-Einsatzes) denkbar. Darüber hinaus kann die Adoptionsneigung auch über Perzeptionsindikatoren (z.B. Einschätzung der Erreichung der Ziele, die mit dem Einsatz von CIM verfolgt wurden) und Performanzindikatoren (direkte, zuordenbare Effekte des CIM-Einsatzes auf die unternehmerische Wettbewerbsfähigkeit) erfasst werden.

Der Vektor \mathbf{X} repräsentiert Faktoren, die die Höhe der Adoptionsneigung eines Unternehmens beeinflussen. Diese Faktoren sollten alle Aspekte, die bei der Entscheidung eines Unternehmens, CIM zu adoptieren, eine Rolle spielen, abbilden. Hierzu kann auf theoretische Modelle zur Erklärung von Innovationsaktivitäten von Unternehmen im Allgemeinen bzw. auf die Einführung der jeweils betrachteten Technologie (hier: CIM) im Besonderen zurückgegriffen werden. Ziel sollte es sein, mit \mathbf{X} möglichst viele potenzielle Einflussfaktoren zu erfassen. Im Rahmen eines eklektischen Modells, das Argumente aus verschiedenen theoretischen Richtungen zusammenführt (vgl. Sarkar 1998; Karshenas und Stoneman 1995), können folgende Einflussfaktorengruppen unterschieden werden:

- *Technologiemarkt* (Bezugsmarkt): (erwarteter) Preis der Technologie, Unsicherheit über die technologische Entwicklung, Angebotsqualität, Anwendungsbereich der Technologie, Transparenz des Marktes, Informations- und Suchkosten, Struktur der Technologieanbieter
- *Absatzmarkt*: (erwartete) Nachfrageentwicklung, Art und Intensität des Wettbewerbs, Struktur und Bedürfnisse der Abnehmer
- *Faktormärkte*: Qualifikationsangebot am Arbeitsmarkt, Finanzierungsmöglichkeiten für Innovationen am Kapitalmarkt
- *Unternehmenscharakteristika*: Absorptionskapazität (wie Größe, FuE-Potenzial, Qualifikationsstruktur, Managementpotenzial), Merkmale des Produktionsprozesses (Produktspektrum, Fertigungsart), Unternehmensstrategie

Die „Politikvariable“ \mathbf{F} kann verschiedene Formen der öffentlichen Unterstützung und den Umfang der öffentlichen Förderung repräsentieren, im einfachsten Fall ist es eine dichotome Variable und entspricht P (Teilnahme am Programm ja/nein). Ein positives Vorzeichen und eine hinreichende statistische Signifikanz der Schätzwerte für den Parameter-Vektor α_3 zeigt an, dass der Einfluss der Programmteilnahme eine positive Wirkung auf die Adoptionsneigung hat.

Die Variable I bildet das Niveau des CIM-Einsatzes vor dem Start des Förderungsprogramms ab. Diese Variable soll den Effekt einer nach oben begrenzten Adoptionsneigung abbilden. Dieses unternehmensspezifische Sättigungsniveau des Technologieeinsatzes ergibt sich aus den Spezifika der Technologie und ihrer Anwendungsmöglichkeiten sowie aus einer Grenznutzenüberlegung (ab einem bestimmten Einsatzniveau übersteigen die Grenzkosten den Grenzertrag einer Ausweitung des Technologieeinsatzes). Der Effekt von I wird voraussichtlich nicht-linear sein. Unternehmen, die in eine bestimmte Technologie noch nicht eingestiegen sind, werden aufgrund von Einstiegsbarrieren (Unsicherheit, Kompatibilitätsprobleme, hohe Einstiegskosten durch technologische Unteilbarkeiten) eine geringere Adoptionsneigung haben als Unternehmen, die bereits erste Elemente einer neuen Technologie anwenden. Ab einer bestimmten Anwendungsintensität ist dann mit einer fallenden Adoptionsneigung bis zur Erreichung des Sättigungsniveaus zu rechnen.

Der Vektor \mathbf{Y} repräsentiert jene Faktoren, die die Entscheidung eines Unternehmens, an dem betrachteten Förderungsprogramm teilzunehmen, beeinflussen. Besonderes Augenmerk ist auf die Abbildung jener Aspekte zu legen, die mit einer Programmteilnahme direkt angesprochen werden (z.B. Finanzierungssituation, Qualifikationsstruktur der Beschäftigten, Managementkapazitäten, Informationsbeschaffungskapazitäten), da davon auszugehen ist, dass Unternehmen, die Defizite in diesen Bereichen aufweisen, eine hohe Bereitschaft zur Beteiligung am Programm haben (was auch programmintendiert ist). Die von \mathbf{Y} abgebildeten Faktoren sollten idealster nicht mit jenen ident sein, die Bestandteil des Vektors \mathbf{X} sind. Dies ist in der Realität schwer durchzuhalten, da insbesondere bestimmte Unternehmenscharakteristika sowohl die Neigung zur Programmteilnahme als auch die Adoptionsneigung wesentlich beeinflussen (z.B. Unternehmensgröße, Branchenzugehörigkeit, Standort).

Eine entscheidende Evaluierungsfrage ist jene nach der Richtung des Einflusses von P auf A : Führt die Förderung zu einer rascheren und intensiveren Anwendung neuer Technologien, oder sind Unternehmen, die neue Technologien rascher und intensiver als andere einführen, erfolgreicher bei der Lukrierung von öffentlichen Förderungen? Für letzteres spricht unter anderem auch der Selektionsmechanismus, der Förderungsprogrammen inhärent ist: Da die Effektivität des Programms (und damit der Erfolg der Programmadministration und der politisch Verantwortlichen) an der Erhöhung der Adoptionsneigung von Unternehmen überprüft wird, bestehen für die programmverantwortlichen Stellen große Anreize, angesichts knapper Förderungsmittel vor allem solche Unternehmen zu fördern, von denen erwartet werden kann, dass sie die Intensität der Technologieanwendung während der Programmlaufzeit deutlich erhöhen. Um die Kausalitätsbeziehung zwischen P auf A zu erfassen, werden (1) und (2) nicht nur separat, sondern auch simultan geschätzt, wobei in (1) an Stelle des Vektors \mathbf{F} die Politikvariable nur über P erfasst wird und in (2) die instrumentelle Variable I entfällt. Die durch die Vektoren \mathbf{X} und \mathbf{Y} repräsentierten Variablen werden so definiert, dass keine Variable in \mathbf{X} und \mathbf{Y} ident definiert ist:

$$A_i = \alpha_1 + \alpha_2 \mathbf{X}_i + \alpha_3 P_i + \alpha_4 I_i \quad (1a)$$

$$P_i = \beta_1 + \beta_2 \mathbf{Y}_i + \beta_3 A_i \quad (2a)$$

Ist bei einer simultanen Schätzung der beiden Gleichungen der Schätzwert für β_3 in (2a) positiv und statistisch signifikant, deutet dies eine Kausalität von der Adoptionsneigung zur Programmteilnahme an (unter der Annahme, dass der Schätzwert für α_3 in (1) positiv und statistisch signifikant ist). Ist der Schätzwert für β_3 nicht positiv bzw. statistisch nicht signifikant, deutet dies auf die umgekehrte Kausalität hin. Diese Vermutung wird gestärkt, wenn gleichzeitig der Schätzwert für α_3 in (1a) positiv und signifikant ist.

Die Aussagefähigkeit des Evaluierungsansatzes steht und fällt mit einer möglichst umfassenden Erklärung der Varianz von A durch den Vektor \mathbf{X} (sowie die instrumentelle Variable I) und der Programmteilnahme P durch \mathbf{Y} . Wird durch \mathbf{X} die Adoptionsneigung von Unternehmen unzureichend erklärt, besteht die Gefahr, dass nicht erfasste Einflussfaktoren mit der Politikvariablen korrelieren und sowohl schätztechnische Probleme verursachen (Nicht-Erfüllung der Unabhängigkeit des Fehlerterms) als auch eine Abdrängung dieser nicht beobachteten Einflüsse in die Politikvariable verursachen. Gleiches gilt für die Politikgleichung (2a) in Bezug auf den Vektor \mathbf{Y} und die unabhängige Variable A .

Zur Schätzung der Adoptionsgleichung (1) und der Politikgleichung (2) sowie zur simultanen Schätzung von (1a) und (2a) werden folgende Informationen benötigt:

- *Unternehmen, die am Programm teilgenommen haben*: vollständige Informationen zu den Variablen A , \mathbf{X} , \mathbf{Y} , \mathbf{F} , I und P für alle Unternehmen oder ein repräsentatives Sample
- *Unternehmen, die nicht am Programm teilgenommen haben*: vollständige Informationen zu den Variablen A , \mathbf{X} , \mathbf{Y} , \mathbf{F} und I für ein repräsentatives Sample der Grundgesamtheit aller Unternehmen (wobei \mathbf{F} andere Förderungen als die durch das zu evaluierenden Programm umfassen kann)

Dies bedeutet, dass für die Umsetzung des Evaluierungsansatzes eine umfangreiche Befragung einer Kontrollgruppe hinreichender Größe notwendig ist.

2.4 Erfahrungen mit der Anwendung des Evaluierungsansatzes in der Schweiz

Die Forschungsgruppe der ETH Zürich hat den oben vorgestellten Evaluierungsansatz im Zug der Wirkungsanalyse des Schweizer Programms zur Förderung des Einsatzes von computergestützten Fertigungstechnologien (CIM-Aktionsprogramm) angewandt. Dieses Programm lief von 1990 bis 1996 und hatte zum Ziel, den Einsatz von CIM durch die Bereitstellung von Informations- und Schulungsmaßnahmen, Innovationsberatung und die finanzielle Unterstützung von Entwicklungsvorhaben und Technologieimplementierungen zu stimulieren. Im Rahmen des Programms wurden nach Angaben der Förderungsstellen 116 Unternehmen unterstützt.

Die Erhebung der benötigten Unternehmensinformationen zur Schätzung von (1a) und (2a) erfolgte im letzten Jahr der Programmlaufzeit (1996) **im Zusammenhang mit der allgemeinen Schweizer Innovationserhebung**. Dadurch konnte für ein repräsentatives Sample des verarbeitenden Gewerbes der Schweiz (Wirtschaftsklassen 15 bis 36 der NACE-Gliederung) eine Vielzahl an Informationen zum Innovationsverhalten, zum

Technologieeinsatz und zu allgemeinen Unternehmenscharakteristika erfasst werden. Als Ergänzung zu dem Standardfragebogen wurde ein spezifischer Teil zum Einsatz von CIM und zum Erhalt von öffentlicher Unterstützung bei der Einführung bzw. Ausweitung des CIM-Einsatzes beigefügt. Erhebungseinheit der Befragung war das Unternehmen als rechtliche Einheit (und nicht der Betrieb als örtlich geschlossene Produktionseinheit). Aus der Grundgesamtheit wurde eine nach 17 Sektoren und 3 (sektorspezifischen) Größenklassen geschichtete Zufallsstichprobe gezogen (wobei alle Großunternehmen in die Stichprobe aufgenommen wurden). Dieses Sample wurde um jene Unternehmen ergänzt, die am Programm teilgenommen hatten, jedoch nicht in der Stichprobe vertreten waren.

Für die Anwendung des Evaluierungsansatzes standen Informationen zu 463 Unternehmen zur Verfügung, davon 96 (= 21 % der befragten Unternehmen), die an dem Förderungsprogramm teilgenommen hatten. Die Rücklaufquote betrug 34 %. Eine Non-response-Analyse wurde nicht durchgeführt, d.h. es kann nicht für eine eventuelle Verzerrung bei den antwortenden Unternehmen in Richtung adoptionsfreudige Unternehmen kontrolliert werden (die Erfahrung zeigt, dass eine solche Verzerrung bei Innovationsbefragungen die Regel ist).

Ein kritischer Punkt bei der Anwendung des Evaluierungsansatzes ist die sorgfältige Spezifikation der Modellvariablen. Hierbei sind die EvaluatorInnen schätztechnischen Restriktionen ausgesetzt: Die absolute Zahl der Modellvariablen ist durch die Größe des Samples (beobachtete Fälle) beschränkt, zur Konstruktion von Modellvariablen stehen nur jene Informationen zur Verfügung, für die zu allen befragten Unternehmen Werte vorliegen, und die Ausprägungen der abhängigen Variablen sollten eine möglichst gleiche Verteilung aufweisen (Mindestbesatz je Zelle). Das Messniveau der abhängigen Variablen beeinflusst die Wahl des angewandten statistischen Schätzverfahrens.

Für die abhängige Variable A in (1) und (1a) (Adoptionsneigung) wurden aus der Vielzahl an möglichen operationalen Spezifikationen drei ausgewählt:

- Veränderung der Intensität des CIM-Einsatzes: Zahl der eingesetzten CIM-Elemente 1996 minus Zahl der eingesetzten CIM-Elemente 1990, transformiert in eine Ordinalvariable mit den drei Ausprägungen „keine oder geringe Zunahme“ (um 0 oder 1 Element) „mittlere Zunahme“ (um 2 bis 4 Elemente) und „starke Zunahme“ (um mehr als 4 Elemente),
- erstmalige Anwendung eines oder mehrerer CIM-Elemente im Bereich „Planung“ (CAD, CAP, Rapid Prototyping) im Zeitraum 1990 bis 1996 (dichotome Variable),
- erstmalige Anwendung eines oder mehrerer CIM-Elemente im Bereich „Produktion“ (BDE, Handhabungssysteme, Industrieroboter, FMC, FMS, aber exklusive CNC-Maschinen), im Zeitraum 1990 bis 1996 (dichotome Variable).

Eine Schätzung von (1) und (1a) auf der Grundlage der beiden letzteren Variablen ist nur für ein Subsample der Unternehmen möglich, nämlich für jene, die vor 1990 noch keines dieser CIM-Elemente eingesetzt haben.

Die abhängige Variable P in (2) und (2a) (Programmteilnahme) wurde in Form von zwei alternativen Indikatoren abgebildet: einerseits als dichotome Variable für die

Programmteilnahme (ja/nein), andererseits als Impulsvariable (Einschätzung des Impulses durch die Förderung auf einer fünfstufigen Skala), die durch Zusammenfassung mehrerer Stufen ebenfalls dichotomisiert wurde (Förderimpuls hoch/niedrig). Beide Arten von Indikatoren wurden differenziert nach den drei Förderungsarten (Information/Schulung, Beratung, Einführung/Entwicklung) gemessen.

Der Vektor \mathbf{X} an unabhängigen Variablen in (1) und (1a) umfasst eine große Zahl an Indikatoren, die aus verschiedenen Ansätzen zur Erklärung des Innovationsverhaltens von Unternehmen abgeleitet wurden. Sie können zu sechs Variablengruppen zusammengefasst werden:

- Ziele des CIM-Einsatzes (sechs Leitziele, die aus einer Faktorenanalyse von 26 Einzelzielen, deren Relevanz die Unternehmen bewertet haben, gewonnen wurden)
- Hemmnisse für den CIM-Einsatz (sechs Hauptbarrieren, die aus einer Faktorenanalyse von 26 Einzelhemmnissen, deren Relevanz die Unternehmen bewertet haben, gewonnen wurden)
- Marktbedingungen (erwartete Nachfrageentwicklung, Intensität der Preis- und des nicht-preislichen Wettbewerbs, Konzentration am Absatzmarkt)
- Charakteristika der Produkte und des Produktionsprozesses (Art des Produktionsprogramms, Art des Fertigungsverfahrens)
- Kapazität zur Wissens- und Technologieabsorption (Qualifikationsstruktur der Beschäftigten, FuE-Kooperationen, Unternehmensgröße)
- Instrumentelle Variablen (Unternehmensalter, CIM-Intensität zu Programmstart).

Der Vektor \mathbf{Y} an unabhängigen Variablen in (2) und (2a) umfasst vier Messgrößen: die Unternehmensgröße (als absolute Zahl der Beschäftigten bzw. als Größenklassen), der Unternehmensstatus (hinsichtlich Zweigniederlassung und Auslandsbesitz), die Branchenzugehörigkeit (hinsichtlich der drei CIM-affinen Branchen Maschinenbau, Metallbearbeitung und Elektrotechnik sowie der CIM-unaffinen Branche Chemie), die frühere Erfahrung mit diffusionsorientierten Förderungsprogrammen (Erhalt von Förderungen durch die KTI vor dem CIM-Aktionsprogramm) sowie die Bedeutung von Finanzierungsproblemen als Hemmnis für den CIM-Einsatz.

Im Rahmen der Evaluierungsstudie der ETH Zürich wurde für Gleichung (1) eine Vielzahl an alternativen Modellen geschätzt, die sich durch die herangezogene Spezifikation der abhängigen Variablen, die berücksichtigten unabhängigen Variablen und die Abgrenzung der analysierten Unternehmensgruppe (gesamtes Sample, nur Klein- und Mittelunternehmen) voneinander unterscheiden (vgl. Arvanitis et al. 1998, 57ff; Arvanitis und Hollenstein 1999). Die Hauptergebnisse der Modellschätzungen in Bezug auf die gegenständliche Evaluierungsfragestellungen können folgend zusammengefasst werden:

- Die Neigung von Unternehmen, CIM-Technologien zu adoptieren bzw. den Einsatz von CIM-Technologien zu intensivieren, steigt mit der Größe des Unternehmens, der Fähigkeit und Ausmaß, externes Wissen zu absorbieren, mit ausgewählten Eigenschaften des Produktionsprozesses (insbesondere bei Fertigung in Großserien) und wenn das vorrangige Ziel des CIM-Einsatzes die Erhöhung der Profitabilität des Unternehmens ist.

- Für die Erklärung des Adoptionszeitpunktes (frühe, mittlere oder späte Adoption einer Technologie) sind vor allem grundsätzliche Hinderungsgründe (Akzeptanz- und Kompatibilitätsprobleme, Defizite im Wissensbereich und beim Personal) ausschlaggebend.
- Die Marktbedingungen (Nachfrage, Konkurrenzintensität, Marktkonzentration, Wettbewerbsform) beeinflussen hingegen die Adoptionsneigung nicht wesentlich.
- Die Ausweitung des CIM-Einsatzes (Zunahme der CIM-Intensität) ist abhängig vom Ausgangsniveau des CIM-Einsatzes: Je niedriger die Intensität des CIM-Einsatzes zum Beginn einer Beobachtungsperiode ist, desto größer ist die Ausweitung des CIM-Einsatzes bis zum Ende der Beobachtungsperiode (Bestätigung der Catch-up-These der Diffusion von Prozesstechnologien).
- Die Teilnahme am CIM-Aktionsprogramm erhöht die Adoptionsneigung der Unternehmen nur insignifikant. Dies ändert sich allerdings, wenn die Unternehmen in zwei Größengruppen geteilt werden (kleine/mittlere bis 200 Beschäftigte, große mit mehr als 200 Beschäftigten): Bei der Gruppe der Klein- und Mittelunternehmen (KMU) ist ein statistisch signifikanter positiver Einfluss auf die Programmteilnahme nachzuweisen, bei Großunternehmen hingegen nicht. Differenziert nach den drei Förderungsansätzen (Schulung, Beratung, Technologieentwicklung) ist der Einfluss bei Inanspruchnahme von Schulungs- und Beratungsangeboten bei KMU eindeutig signifikant, bei der Entwicklung nur in Bezug auf die Einführung von CIM-Elementen in der Planung und Produktion.

Gleichung (2) wurde für zwei verschiedene abhängige Variablen (Förderung ja/nein und Förderungsimpuls niedrig/hoch) geschätzt, die Ergebnisse unterscheiden sich dabei nur wenig. Die wichtigsten Einflussfaktoren auf die Teilnahme eines Unternehmens am CIM-Aktionsprogramm sind

- die Unternehmensgröße (sehr kleine und sehr große Unternehmen haben eine überdurchschnittliche Teilnahmeneigung)
- die Branche (die Zugehörigkeit zur Metallbearbeitung und zum Maschinenbau erhöht die Teilnahmewahrscheinlichkeit signifikant)
- sowie Finanzierungsprobleme und die bisherige Erfahrung mit Förderungen durch die KTI.
- Die Adoptionsneigung eines Unternehmens (gemessen als die Erhöhung der Intensität des CIM-Einsatzes während der Programmlaufzeit) hat dagegen für die Teilnahmewahrscheinlichkeit am CIM-Aktionsprogramm nur einen geringen Einfluss, der lediglich bei der Betrachtung eines reduzierten Unternehmenssamples (nur jene Unternehmen, die zum Beginn der Programms im Jahr 1990 noch keine CIM-Elemente im Fertigungsbereich eingesetzt hatten).

Für die simultane Schätzung von (1a) und (2a) wurden die Schätzungen von (1) und (2) dazu verwendet, jene unabhängigen Variablen auszuwählen, deren Einfluss auf die Adoptionsneigung bzw. die Programmteilnahmeneigung positiv und statistisch signifikant war. In der simultanen Schätzung wurde getestet, ob die Teilnahme am CIM-Aktionsprogramm einen signifikant positiven Einfluss auf die Adoptionsneigung und gleichzeitig die Adoptionsneigung keinen signifikanten Einfluss auf die Programmteilnahme ausüben. Die Schätzung wurde für vier Samples an Unternehmen durchgeführt: Erstens für alle in der Befragung erfassten Unternehmen, zweitens nur für Unternehmen, die 1990 noch

kein CIM-Element im Produktionsbereich eingesetzt hatten (da dies eine explizite Zielgruppe des Programms war), drittens nur KMU (weniger als 200 Beschäftigte, da dies ebenfalls eine explizite Zielgruppe des Programms war) und viertens KMU, die 1990 noch kein CIM-Element im Produktionsbereich eingesetzt hatten (als die engste Definition der Programm-Zielgruppe). Die Ergebnisse können folgend zusammengefasst werden:

- Die Adoptionsneigung hat durchgehend keinen signifikanten Einfluss auf die Teilnahmewahrscheinlichkeit am Programm.
- Die Programmteilnahme hat für die Samples der Unternehmen ohne CIM-Einsatz im Produktionsbereich (sowohl alle als auch nur KMU) einen statistisch signifikanten positiven Einfluss auf die Adoptionsneigung.
- Für die Gesamtheit aller befragten Unternehmen ist der Effekt der Programmteilnahme jedoch nicht signifikant.
- Für die Gruppe der KMU ist der Effekt der Programmteilnahme auf niedrigem Niveau statistisch signifikant.

Die Wirkungsanalyse des CIM-Aktionsprogramms der Schweizer Bundesregierung mit dem mikroökonomischen Analyseansatz des Instituts für Konjunkturforschung der ETH Zürich hat somit ergeben, dass die Förderung der Unternehmen im Rahmen des Programms zu einer signifikant rascheren und intensiveren Verbreitung von CIM beigetragen hat. Insbesondere die beiden Hauptzielgruppen des Programms – KMU und Unternehmen, die CIM noch nicht einsetzen – wurden durch das Programm in ihrem Verhalten signifikant im gewünschten Sinn beeinflusst. Die Förderung von Großunternehmen im Rahmen des Programms hat dagegen keinen Einfluss auf deren Adoptionsverhalten gehabt, hier liegt die Annahme nahe, dass es zu Mitnahmeeffekten der Programmförderung gekommen ist.

2.5 Erfahrungen mit der Anwendung des Evaluierungsansatzes in Österreich

Ende 1999 wurde das Österreichische Forschungszentrum Seibersdorf (ARCS) vom BMWV beauftragt, eine Ex-post-Evaluierung des ITF-Programms „FlexCIM“ durchzuführen. Dieses Programm lief von 1992 bis 1998 und zielte auf die Unterstützung von Unternehmen bei der Einführung von computergestützten Fertigungstechnologien (CIM) ab. Die Förderung umfasste die Elemente Information & Schulung, Beratung sowie Technologieeinführung und setzte sich aus einer Konzeptphase und einer Durchführungsphase zusammen, wobei Unternehmen eine der beiden oder beide Phasen in Anspruch nehmen konnten. Förderungen standen sowohl für Technologieanwender als auch für Technologieentwickler (inklusive Softwareunternehmen) zur Verfügung.

Im Rahmen einer umfassenden Wirkungs- und Effizienzanalyse des Programms wurde vom ARCS ein Modul angeboten, das die Anwendung des Evaluierungsansatzes der ETH Zürich zum Inhalt hatte. Dieses Modul wird in Zusammenarbeit mit dem Institut für Konjunkturforschung der ETH Zürich, dem Institut für Technologie- und Regionalpolitik der Joanneum Research und dem Zentrum für Europäische Wirtschaftsforschung (ZEW) bearbeitet. Zum jetzigen Zeitpunkt (Mai 2000) liegen erste Erfahrungen zu diesem Evaluierungsansatz vor, über die kurz berichtet wird.

Die Erhebung der benötigten Unternehmensinformationen begann Ende 1999, d.h. ein Jahr nach dem Auslaufen des Förderungsprogramms, und lief bis April 2000. Das befragte Sample umfasste alle Unternehmen, die am Programm teilnahmen, sowie – als Kontrollgruppe – sämtliche Produktionsbetriebe mit mehr als 20 Beschäftigten (auf Basis der Betriebsdatenbank des ARCS). Erhebungseinheit war der Betrieb, d.h. bei Mehr-Betriebs-Unternehmen bezogen sich alle Fragen auf den einzelnen, örtlich geschlossenen Produktionsstandort.

Unter den Programmteilnehmern wurde eine Rücklaufquote von 47 % (= 84 Betriebe) erzielt⁴. Trotz mehrmaligem telefonischen Kontakts konnte über die Hälfte der programmteilnehmenden Betriebe nicht zu einer Beteiligung an der Befragung bewegt werden. Innerhalb der Kontrollgruppe betrug die Rücklaufquote 9 % (= 333 Betriebe). Von den Betrieben in der Kontrollgruppe waren Ende 1999 rund 65 % CIM-Anwender (= 217 Betriebe). Für die Kontrollgruppe wurde eine Non-response-Analyse auf Basis einer nach Sektoren und Größenklassen geschichteten Zufallsstichprobe (n = 163) durchgeführt, die zeigte, dass unter den nicht-antwortenden Betrieben der Einsatz von CIM niedriger ist (58 %) als unter den antwortenden Betrieben, die Differenz jedoch nur auf einem niedrigen Niveau statistisch signifikant ist.

Von den 417 Betrieben, die den Fragebogen beantwortet haben, werden für die Analyse im Rahmen des hier vorgestellten Evaluierungsansatzes nur jene 301 Betriebe (217 aus der Kontrollgruppe plus 84 Programmteilnehmer) herangezogen, die auch CIM einsetzen. Die Betriebe, die am FlexCIM-Programm teilgenommen haben, werden also gegen eine Gruppe von Unternehmen verglichen, die (zum Zeitpunkt der Befragung, also nach dem Ende der Programmlaufzeit) durchweg CIM anwenden (und damit von diesem Merkmal her der Gruppe der Programmteilnehmer entsprechen). Dies Vorgangsweise entspricht auch dem Analysedesign der ETH-Gruppe.

Zur Schätzung von (1a) und (2a) wurden im Wesentlichen die gleichen Variablen wie in 4. beschrieben herangezogen. Die Adoptionsneigung wird in drei alternativen operationalen Varianten gemessen (Intensität des CIM-Einsatzes 1998, Veränderung der Intensität des CIM-Einsatzes zwischen 1991 und 1998, Ausmaß der Vernetzung von CIM-Elementen 1998) operationalisiert. Die operationale Definition der einzelnen unabhängigen Variablen weicht teilweise – erhebungsbedingt – geringfügig von derjenigen der Schweizer Untersuchung ab. Insbesondere ist bei der österreichischen Untersuchung ein größeres Set an Unternehmenscharakteristika erfasst worden.

Zum Zeitpunkt der Berichtslegung liegen noch keine Ergebnisse zu den simultanen Schätzungen vor. Die separaten Schätzungen von Adoptions- und Politikgleichung zeigen ähnliche Resultate wie im Fall der Schweizer Evaluierung:

⁴ Unternehmen, die als Technologieentwickler bzw. –berater am Programm teilgenommen hatten (insbesondere Softwareentwickler) wurden nachträglich aus der Grundgesamtheit der 225 Programmteilnehmer ausgeschlossen, da sie selbst keine Anwender von CIM sind. Knapp 20 der Unternehmen, die am Programm teilgenommen hatten, konnten nicht befragt werden, da sie zum Zeitpunkt der Befragung nicht mehr existierten, wodurch die tatsächliche Grundgesamtheit der Produktionsunternehmen, die am FlexCIM-Programm teilnahmen, 180 beträgt.

- Im Modell zum Adoptionsverhalten ist der Effekt der Programmteilnahme für die Teilgruppe der KMU statistisch signifikant positiv, für Nicht-KMU ist der Effekt jedoch insignifikant.
- In der „Politikgleichung“ ist der Effekt der Adoptionsneigung auf die Wahrscheinlichkeit der Programmteilnahme für KMU insignifikant, für Nicht-KMU jedoch statistisch signifikant positiv.

2.6 Schlussfolgerungen

Der Ansatz des Instituts für Konjunkturforschung der ETH Zürich ist ein methodisch anspruchsvoller Weg zur Prüfung von Effektivität und Mitnahmeeffekten von Förderungsmaßnahmen, die auf die Verhaltensänderung von Akteuren abzielen. Der Einsatzbereich dieser Evaluierungsmethode ist nicht nur auf diffusionsorientierte Technologieprogramme beschränkt, sondern eignet sich für alle Förderungen, deren Ziel es ist, Aktivitäten von Akteuren in eine bestimmte Richtung zu lenken.

Die Evaluierungsmethode zeichnet sich durch folgende **Stärken** aus:

- Methodisch saubere und zuverlässige Schätzung der **direkten Wirkung des Programms** auf das Adoptionsverhalten der Zielgruppe
- Methodisch saubere und zuverlässige Schätzung von **Mitnahmeeffekten des Programms**, d.h. der Überprüfung der Kausalität zwischen Programmteilnahme eines Unternehmens und der intendierten Verhaltensänderung
- Quantifizierung des **relativen Beitrags des Programms** zur Veränderung des Adoptionsverhaltens von Unternehmen
- Quantifizierung der **Wirkung des Programms auf unterschiedliche Teilgruppen** innerhalb der Zielgruppe

Die **Defizite** des Evaluierungsansatzes liegen insbesondere in folgenden Bereichen:

- Die Konzentration auf den Aspekt der direkten Wirkungsanalyse unter (bewusster) Weglassung der Evaluierung der **Relevanz der Politikintervention** (d.h. ob diese Form der Intervention problemadäquat war) und der **Effizienz der Programmumsetzung** im Bereich des Programmmanagements.
- **Indirekte Wirkungen** des Programms auf Unternehmensebene (Wirkung auf die Wettbewerbsfähigkeit, auf das künftige Innovationspotenzial, auf Produktinnovationen, auf langfristige Veränderungen des Innovationsverhaltens im Bereich Innovationsorganisation, Innovationsmanagement und FuE-Kooperationen, auf die Unternehmensstrategie etc.) werden nicht untersucht.
- Die Analyse von **sektoralen und gesamtwirtschaftlichen Effekten** (Beitrag des Programms zur Diffusion einer bestimmten Technologie in der Wirtschaft insgesamt, Politikeffektivität auf gesamtgesellschaftlicher Ebene, gesamtwirtschaftliche Produktivitäts- und Beschäftigungseffekte, Veränderung der internationalen Wettbewerbsfähigkeit eines Sektors über Input-Output-Beziehungen etc.) ist nicht Gegenstand des Ansatzes.

Die praktische **Umsetzbarkeit** des Evaluierungsansatzes ist durch die hohen Anforderungen an Datenumfang und Datenqualität eingeschränkt. Es ist ein hoher Ressourcenaufwand

sowohl bei der evaluierenden Einrichtung als auch bei den Unternehmen notwendig, um die zwingend benötigten Informationen zu erheben. Neben der detaillierten Befragung der Programmteilnehmer stellt insbesondere die ebenso detaillierte Befragung einer Kontrollgruppe eine wesentliche Hürde dar, da es schwierig ist, Unternehmen aus der Kontrollgruppe für eine Befragung zu motivieren.

Die **Anwendungsmöglichkeiten** dieses Evaluierungsansatzes in der **österreichischen Technologiepolitik** sind vielfältig. Neben dem Einsatz zur Evaluierung von diffusionsorientierten Programmen (wie diese derzeit bei der Ex-Post-Evaluierung des FlexCIM-Programms geschieht und auch für andere Diffusionsprogramme in gleicher Weise vorstellbar ist) stellt die Evaluierungsmethode ein taugliches Instrument zur Überprüfung jedweder direkten Unternehmensförderung dar. Voraussetzung für die Anwendung der Methode ist jeweils, dass

- die Intervention auf eine beobachtbare Verhaltensänderung bzw. die Auslösung einer beobachtbaren Aktivität bzw. die Erzielung beobachtbarer Effekte bei einzelnen Akteuren abzielt (wobei angenommen wird, dass ohne die Förderung diese Verhaltensänderung bzw. Aktivität nicht eingetreten wäre) – diese Akteure können sowohl Unternehmen als auch Individuen oder Haushalte sein;
- diese Verhaltensänderung (bzw. diese Aktivität/dieser Effekt) auch bei Akteuren beobachtet werden kann, die keine Förderung erhalten haben (Existenz einer Kontrollgruppe);
- es möglich ist, das relevante Verhalten (bzw. die relevante Aktivität/den relevanten Effekt) über die Zeit hinweg sowohl bei den geförderten Akteuren als auch bei der Kontrollgruppe zu beobachten – als Mindestanforderung ist eine Beobachtung zum Zeitpunkt vor der Maßnahme und zum Zeitpunkt danach notwendig;
- es möglich ist, diese Verhaltensänderung (bzw. diese Aktivität/diesen Effekt) mit Hilfe eines operationalen Modells hinreichend gut zu erklären (d.h. die Einflussfaktoren der Zielgröße der Intervention zu kennen und beobachten zu können).

Für folgende technologiepolitischen Maßnahmenbereiche bzw. Förderungsansätze in Österreich scheint eine Anwendung der vorgestellten Evaluierungsmethode sinnvoll:

- Projektförderung des ERP-Fonds in den Technologieprogrammen (Wirkung der Förderungen auf Produkt- und Prozessinnovationen)
- Projektförderung des FFF (Wirkung der Förderungen auf FuE-Tätigkeit, wobei hier das Problem einer zu kleiner Kontrollgruppe auftreten könnte, da fast alle FuE-treibenden Unternehmen auch FFF-Förderungen erhalten; oder Wirkung der Förderungen auf Fund E-Kooperationen)
- Projektförderung der Bürges-Bank im Gründungsbereich (Wirkung auf die Überlebenswahrscheinlichkeit einer Unternehmensgründung)
- Projektförderung der Länder im Technologiebereich
- Qualifizierungsmaßnahmen des AMS im betrieblichen Bereich mit Technologiebezug

2.7 Literatur

- Arvanitis, S., L. Donzé, H. Hollenstein und S. Lenz (1998), Wirksamkeit der diffusionsorientierten Technologieförderung des Bundes. Eine Analyse anhand von Firmendaten, Bern (Reihe Strukturberichterstattung).
- Arvanitis, S. und H. Hollenstein (1997), Evaluating the promotion of advanced manufacturing technologies (AMT) by the Swiss government using micro-level survey data: some methodological considerations, in: OECD (Hrsg.), Policy Evaluation in Innovation and Technology. Towards Best Practices, Paris, 325-334.
- Arvanitis, S., H. Hollenstein und S. Lenz (1998), Are Swiss Governemnt Programmes fo Promotion of Advanced Manufacturing Technologies (AMT) Effecitive? An Economic Analysis Based on Micro-level Survey Data, Paper presented at the International Conference on „The Economic Evaluation of Technological Change“, Washington (15.-16.6.)
- Arvanitis, S. und H. Hollenstein (1999), The Determinants of the Adoption of Advanced Manufacturing Technology. An Empirical Analysis Based on Firm-level Data for Swiss Manufacturing, Zürich (mimeo).
- Becher, G. and S. Kuhlmann (eds.) (1995), Evaluation of Technology Policy Programmes in Germany, Dordrecht.
- Dreher, C. und A. Blathasar (1997), Evaluierung des Schweizer CIM-Aktionsprogramms 1990 bis 1996, Karlsruhe.
- Fritsch, M. (1995), The market, market failure, and the evaluation of technology-promoting programs, in: G. Becher and S. Kuhlmann (Hrsg.), Evaluation of Technology Policy Programmes in Germany, Dordrecht, 311-329.
- Karshenas, M. und P. Stoneman (1995), Technological diffusion, in: P. Stoneman (ed.), Handbook of the Economics of Innovation and Technological Change, Oxford.
- OECD (Hrsg.) (1997), Policy Evaluation in Innovation and Technology. Towards Best Practices, Paris.
- Papaconstantinou, G. und W. Polt (1997), Policy evaluation in innovation and technology: an overview, in: OECD (Hrsg.), Policy Evaluation in Innovation and Technology. Towards Best Practices, Paris, 9-14.
- Sarkar, J. (1998), Technological diffusion: alternative theories and historical evidence., Journal of Economic Surveys 12, 131-176.
- Shapira, P., J. Youtie und J.D. Roessner (1996), Current practices in the evaluation of US industrial modernization programmes, Research Policy 25, 215-232.

3 EVALUIERUNG DIREKTER FÖRDERUNG AM BEISPIEL DES „ADVANCED TECHNOLOGY PROGRAMS“ (ATP)

*Wolfgang Polt, Birgit Woitech, (unter Mitarbeit von Niko Gretzmacher)
Joanneum Research*

3.1 Geschichte und Organisation, Ziele und Ausrichtung

Das Advanced Technology Program (ATP) wurde 1990, unter der Bush-Administration, zur Unterstützung US-amerikanischer Unternehmen in der Entwicklung und Kommerzialisierung von Technologien mit hohem Neuigkeitsgrad und großem Risiko („Leap-frog technologies“) eingerichtet. Von zentraler Bedeutung ist dabei der vorwettbewerbliche Charakter dieses Programms, welcher einerseits dazu beitragen soll, die Spillovers zu maximieren und andererseits die Gefahr der negativen Auswirkungen auf die Produktmärkte (Anbieter-Konzentration) minimieren. Die gesetzliche Grundlage des Programms bildet der „Omnibus Trade and Competitiveness Act (the Trade Act)“ aus 1988, welche 1992 durch den „American Technology Preeminence Act“ wichtige Veränderungen erfahren hat und der ATP unter anderem verpflichtete, einen umfassenden Statusbericht über Ergebnisse und Erfolg bis spätestens 1996 sowohl an den Kongress als auch an den Präsidenten einzureichen. Mit der Abwicklung wurde das NIST (National Institute of Standards and Technology) beauftragt.

Ursprünglich konzipiert als ein sehr kleines Programm, mit einem Budget von 10 Millionen US\$, wurde es unter der Regierung Clinton substantiell ausgeweitet (145,4 Mio. US\$ im Jahr 2001). Dennoch ist der Stellenwert von ATP innerhalb der US-amerikanischen Technologiepolitik vom Volumen her gering: Die kumulierten Förderungen ergeben bis zum Jahr 2000 eine Summe von 1,64 Mrd. US\$ - eine Summe, die anderen Programmen, wie zum Beispiel DARPA (Defense Advanced Research Projects Agency), jährlich zur Verfügung steht (ca. 2 Mrd. US\$). Da ATP jedoch als experimentelles Programm mit konzeptionellen Neuerungen - im Vergleich zu anderen Technologienprogrammen - gilt, war und ist es politisch umstritten. Vor allem nach den Wahlen 1994, mit einer republikanischen Mehrheit im Kongress, wurde der Rechtfertigungs- und Legitimierungsdruck sehr stark. Dies führte dazu, dass, obwohl mit vergleichsweise geringem Budget ausgestattet, ATP zu den „most rigorously studied, open, and closely scrutinized program of any era, in any country“ (C. Hill, 1999, S. 41) gehört. Die im Zuge dieser Evaluierungsstudien entwickelten und angewandten qualitativen und quantitativen Methoden stellen ein sehr gutes Ausgangsmaterial für die Bewertung der Möglichkeiten und Grenzen von quantitativen Ansätzen dar.

Zentrale Aufgabe von ATP ist es, im Rahmen von Wettbewerben anteilmäßig eine begrenzte Summe an Fördermitteln an einzelne US-amerikanische Unternehmen oder Forschungsgemeinschaften (research joint venture, research consortia) zu vergeben, um diese bei der Entwicklung neuer Technologien zu unterstützen, die entweder von den teilnehmenden Unternehmen selbst und/oder anderen eingesetzt und vermarktet werden können.

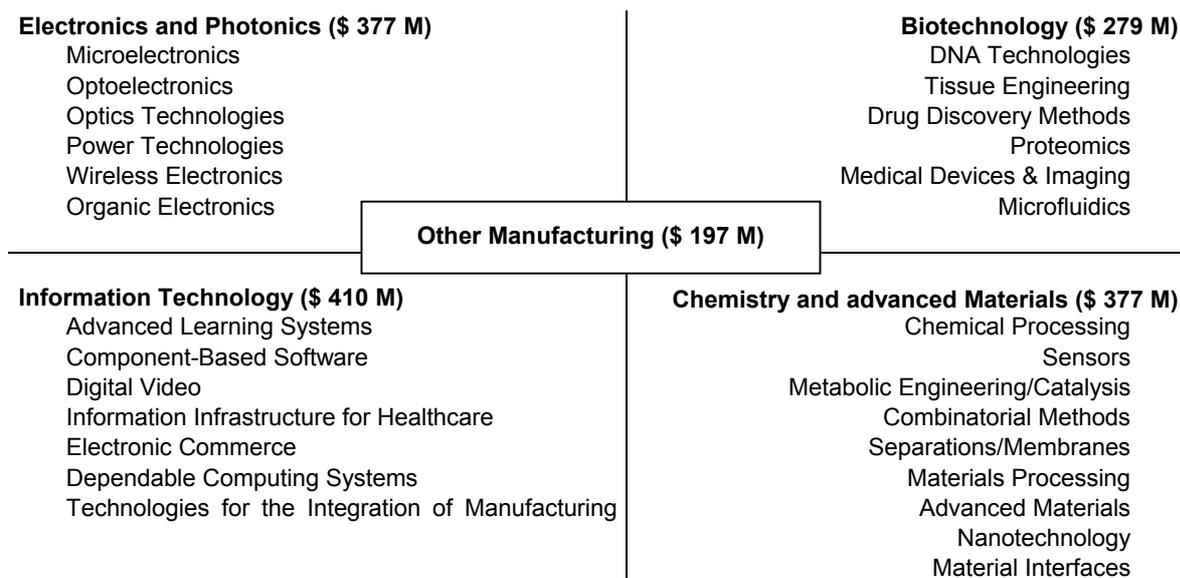
Demgemäß lassen sich die Ziele von ATP wie folgt bestimmen:

- Förderung der Entwicklung von Technologien mit hohem Risiko und Neuigkeitsgrad („leap-frog technologies“)
- Förderung von Technologien mit breiten Anwendungsmöglichkeiten und großem Diffusionspotential („generic technologies“)
- Förderung der Kooperation zwischen Unternehmen und Forschungseinrichtungen
- Beschleunigung der Kommerzialisierung von technologischen Innovationen
- Stärkung der Wettbewerbsfähigkeit der US-amerikanischen Wirtschaft
- Erzielung hoher ‚sozialer‘ Erträge

Seit dem Start von ATP im Jahr 1990 bis September 2001 wurden 581 Projekte mit insgesamt 1.250 TeilnehmerInnen und einer annähernd gleich großen Summe an SubauftragnehmerInnen kofinanziert. Etwas mehr als zwei Drittel der Projekte sind Einzelanträge (396), etwa ein Drittel wird von Joint Ventures (185) getragen. Von wissenschaftlicher Seite waren 160 Universitäten und mehr als 20 nationale Forschungseinrichtungen beteiligt. Über die Hälfte der Projekte (61%) wird von KMU's geleitet. Die Gesamtprojektsumme beträgt 3,6 Mrd. US\$, und ist zu gleichen Teilen auf das ATP und die US-amerikanische Industrie verteilt (ATP 1,8 Mrd. US\$, Industrie 1,8 Mrd. US\$).

Der Schwerpunkt von ATP liegt auf (zivilen) Technologien, die das Potential haben, zu beträchtlichen Produktivitätszuwächsen und Wettbewerbsverbesserungen der Unternehmen beizutragen, KonsumentInnen mit neuen, besseren und billigeren Gütern und Dienstleistungen zu versorgen und die Beschäftigung von hochqualifizierten Arbeitskräften in den USA zu erhöhen (Ruegg, 1999). ATP klassifiziert die geförderten Projekte nach fünf Technologiefeldern je nachdem in welchem Bereich die technologische Innovation entwickelt wird. Die Projekte verteilen sich wie folgt auf die Technologiefelder:

Tabelle 4: Finanzierung nach Technologiefeldern, 1990-2000

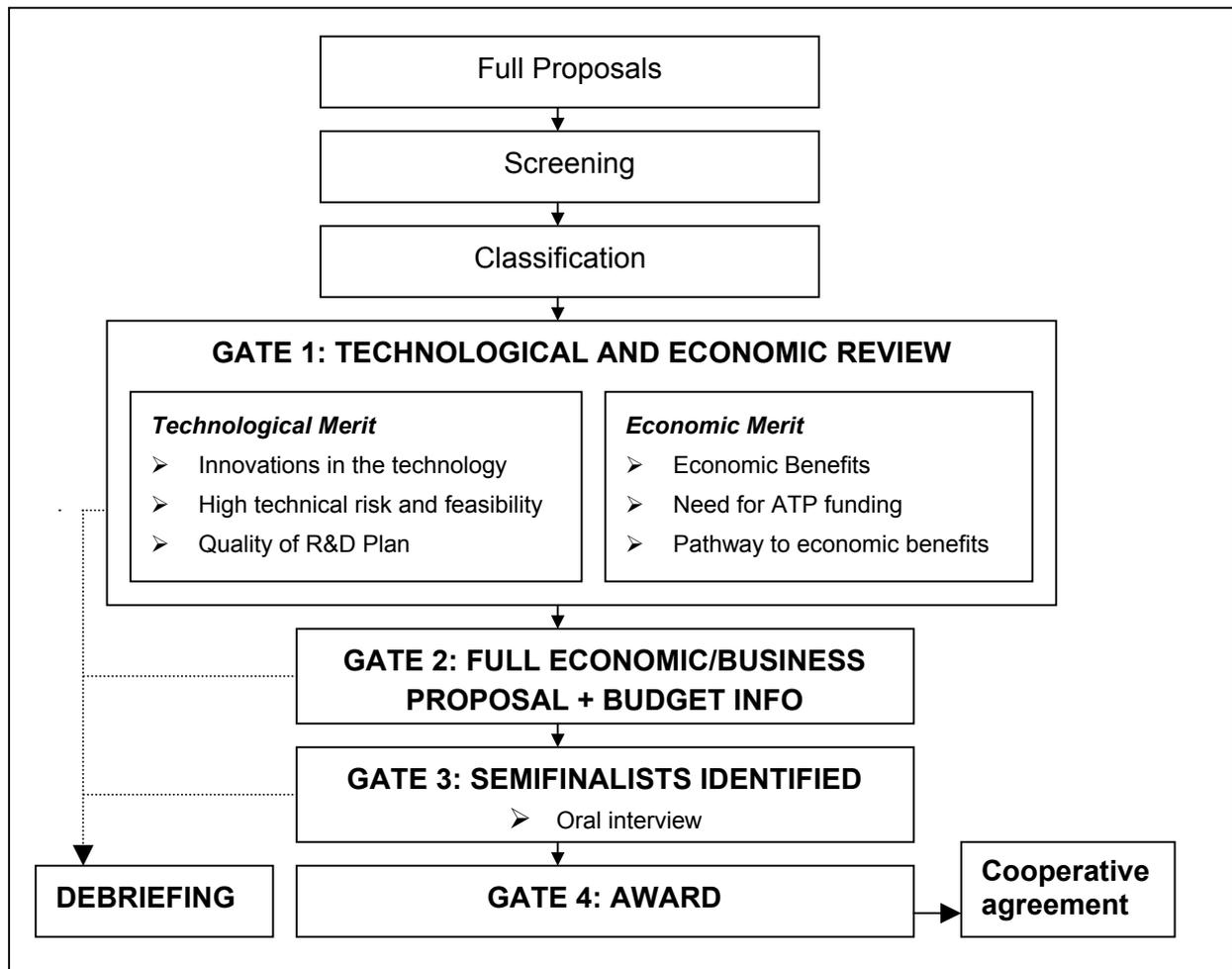


Für die mehrjährigen Förderungen von ATP können sich die Unternehmen entweder individuell oder als Forschungsgemeinschaft (research joint ventures) bewerben. Joint Ventures müssen aus mindestens zwei gewinnorientierten Unternehmen bestehen, die eigenständige Forschung betreiben und die anfallenden Kosten anteilig untereinander aufteilen. Sehr häufig werden die Projekte, unabhängig von der Ausgestaltung der Einreichung, in Kooperation mit Universitäten und/oder außeruniversitären Forschungseinrichtungen durchgeführt.

Die Förderungen unterliegen einer zeitlichen Begrenzung, sind allerdings längerfristig angelegt. Sie liegt für Einzelunternehmen bei 3 Jahren, für Joint Ventures bei 5 Jahren. Zusätzlich gilt: Bei Anträgen von einzelnen Unternehmen müssen diese die gesamten indirekten Kosten tragen. Von ATP kofinanziert werden nur die direkten (Forschungs)Kosten bis zu einer Höhe von 2 Mio. US\$. Sehr große Unternehmen müssen mehr als 60% der Projektkosten selbst tragen. Bei Joint Ventures gibt es zwar keine monetäre Obergrenze, jedoch müssen mehr als 50% der Gesamtkosten von der Forschungsgemeinschaft übernommen werden. Erfindungen oder Patente, die aus ATP geförderten Projekten entstehen, sind Eigentum der (gewinnorientierten) Unternehmen. Universitäten, non-profit Organisationen oder staatliche Einrichtungen haben keinerlei Anspruch, jedoch können die Lizenzgebühren zwischen den Partnerinstitutionen aufgeteilt werden.

Die Auswahl der von ATP kofinanzierten Projekte erfolgt in einem mehrstufigen Selektionsprozess, welcher gleichzeitig auch den ersten Schritt des Evaluierungsmodells von ATP darstellt. Mittels eines Peer-review Verfahrens werden die Projektvorschläge hinsichtlich ihres technologischen und ökonomischen Inhaltes bewertet. Jeder Vorschlag wird – ex-ante - anhand seines Innovationspotentials, des technischen Risikos, seiner Machbarkeit, der Qualität des FuE-Planes, der Notwendigkeit für ATP Kofinanzierung und seines ökonomischen Nutzens, sowohl des privaten als auch des sozialen evaluiert. Abbildung 2 zeigt in einer schematischen Darstellung die einzelnen Stufen des Auswahlprozesses:

Abbildung 2: Projektselektionsprozess



Das Auswahlgremium umfasst neben MitarbeiterInnen von ATP auch unabhängige ExpertInnen aus den Bereichen Technik, Wirtschaft und Wissenschaft. Technologische und ökonomische Selektionskriterien werden zu gleichen Teilen gewichtet, da für ATP neben der Entwicklung neuer, mit hohem Risiko behafteter Technologien die Kommerzialisierung der daraus entstehenden Produkte und/oder Verfahren einen ebenso wichtigen Stellenwert einnimmt. Ziel des Selektionsprozesses ist es, jene Projekte auszuwählen, deren zu erwartender sozialer Nutzen deutlich höher ist als die privaten Erträge aus der Investition (Maximierung der Spillovers) und die ohne Kofinanzierung durch ATP überhaupt nicht durchgeführt (hohe Additionalität), jedenfalls aber nicht innerhalb eines bestimmten Zeitrahmens fertiggestellt hätten werden können.

Bereits an dieser kurzen Darstellung der Inhalte und des Auswahlprozesses lässt sich feststellen, das ATP ein komplexes Programm ist, welches versucht, technologische und ökonomische Komponenten in der Selektion der Projekte bestmöglich zu verbinden. Ein Evaluierungsprogramm muss zum einen diesen Aspekten Rechnung tragen und zum anderen auch die zeitliche Dimension der zu erwartenden Rückflüsse berücksichtigen. Die wirtschaftliche Verwertung von neuen Technologien oder Verfahrenstechniken ist ein längerfristiger Prozess, der sich häufig überhaupt erst (lange) nach Projektende zeigt. Die besondere Herausforderung des Evaluierungsmodells von ATP ist es daher, (a)

Selektionskriterien, die soziale Erträge ex-ante adäquat berücksichtigen, (b) kurzfristige Indikatoren über Projektfortschritt und Wirkung mit (c) einer Abschätzung der langfristigen volkswirtschaftlichen Effekte zu kombinieren. Die Ausgestaltung dieses Evaluierungssystem soll nun im folgenden Abschnitt dargestellt werden.

3.2 Das Evaluierungsdesign von ATP

Wie bereits ausgeführt, ist ATP aufgrund des hohen politischen Drucks das am häufigsten und auch am strengsten evaluierte Programm. Schon vor Inkrafttreten des „Government Performance and Results Act“ (GPRA) 1993, der eine eingehende Bewertung aller öffentlichen Programme und der Leistungen von öffentlichen Institutionen vorsieht, war es daher notwendig, eine entsprechende Bewertung des Programms vorzunehmen. So erschien 1993 der erste Bericht, eine Befragung über die (Aus)Wirkungen der ersten 11 ATP-Projekte auf die teilnehmenden Unternehmen (Solomon Associates, 1993).

In der Ausarbeitung eines adäquaten Evaluierungssystems war von Anfang an klar, dass ATP „a unique measurement challenge“ (Kammer, 1999, S. 31) sein würde. Denn so wie ATP ein „experiment in technology policy“ (Spender, 1999, S. 83) darstellt, so sehr ist es auch „an experiment in evaluations research“ (Spender, 1999). Für die Entwicklung einer verbindlichen Evaluierungsstrategie, welche auch Richtlinien für die Erstellung von Evaluierungen des ATP enthält, wurde das Economic Assessment Office (EAO) eingerichtet, das darüber hinaus noch eine Vielzahl von weiteren Aufgaben erfüllt. Unterstützt wurde es dabei von externen ForscherInnen (z. B. des NBER) und ExpertInnen aus dem Bereich der Evaluation (Edwin Mansfield, Adam Jaffe etc.). Ziel war, die Programmevaluierung vollständig in die Dynamik der Programmstruktur bzw. des Managements zu integrieren, um dadurch eine ständige Verbesserung und Weiterentwicklung gewährleisten zu können.

Um sowohl die Nachfrage nach kurzfristigen Programmsergebnissen befriedigen, aber auch die langfristigen Effekte messen und bewerten zu können, wurde eine Evaluierungsstrategie entworfen, die verschiedenste methodische Komponenten umfasst und die auch den unterschiedlichen zeitlichen Dimensionen der Rückflüsse Rechnung tragen soll. Zu den wichtigsten Elementen gehören:

- (1) Deskriptive Datenerfassung von BewerberInnen, Projekten, TeilnehmerInnen und Technologien
- (2) Messung des Projektfortschritts mittels Befragungen und dem „Business Reporting System“
- (3) Real-Time Monitoring durch ATP und Unternehmensberichte
- (4) Statusberichte zu abgeschlossenen Projekten
- (5) Mikro- und makroökonomische vertiefende Fallstudien zu den Impacts von einzelnen Projekten oder Projekt-Clustern (private und soziale Erträge, Additionalität)
- (6) Studien zu speziellen Aspekten (z. B. Beteiligung vom KMU, Rolle von Universitäten)
- (7) Laufende Weiterentwicklung von Methoden und Ansätzen zur Evaluierung langfristiger Programme oder Projekte
- (8) Vergleichende Studien mit anderen Programmen

Zentraler Bestandteil des Evaluierungsprozesses, welcher nicht nur für das laufende Monitoring und Projektmanagement sondern auch für die (kurz- und langfristige) Impact-Messung von Bedeutung ist, ist die von ATP eingerichtete Datenbank, in der die wichtigsten Daten aus den Projektanträgen erfasst sind. Diese Datenbank besteht aus der „Awards Database“ (TeilnehmerInnen, Projekte, Technologien), der „Applicants Database“ (BewerberInnen) sowie dem „Business Reporting System (BRS)“. Das BRS besteht aus fünf Hauptkomponenten, welche sich durch den Projektfortschritt bestimmen:

- (1) Zu *Projektbeginn* wird ein Bericht erstellt, in dem die Unternehmen die geplanten Anwendungsbereiche ihrer Technologien sowie ihre Vermarktungsstrategien identifizieren.
- (2) Die während der Projektzeit abgefassten *Jahresberichte* beinhalten zum einen bereits kurzfristig realisierte ökonomische Effekte, wie zum Beispiel frühzeitige Verkaufserlöse, eine Verkürzung des Entwicklungszyklus. Zum anderen umfassen sie Auswirkungen des Projektfortschrittes auf die Vermarktungsstrategie, wie zum Beispiel die Erschließung neuer Anwendungsfelder oder Entwicklungsrichtungen, Änderungen des technologischen Entwicklungszieles oder den Abbruch des Projektes.
- (3) Darüber hinaus werden in von den Unternehmen verfassten *Quartalsberichten* ein oder zwei wesentliche wirtschaftliche Entwicklungen im Rahmen des ATP-Projektes dargestellt.
- (4) Zu Ende eines Projektes erfolgt ein *Abschlussbericht*. Dieser beinhaltet einerseits eine Aufstellung über die Zielerreichung in wirtschaftlicher und technologischer Hinsicht sowie Hindernisse und Hemmnisse für eine erfolgreiche Vermarktung. Andererseits werden spezifische Geschäftsstrategien für die nächsten fünf Jahre entworfen und zu erwartende Spillover-Effekte des Projektes skizziert.
- (5) Da ein großer Teil der Erträge häufig erst nach Projektabschluss realisierbar ist, erfolgt ein systematischer Follow-Up bis sechs Jahre nach Projektende. In sogenannten *Post-Project Reports* müssen die Unternehmen insgesamt dreimal (alle 2 Jahre) über den Vermarktungsprozess der Technologie und damit verbundene Impacts für das Unternehmen selbst oder andere Auskunft geben. Darüber hinaus werden vom EAO (Economic Assessment Office) eine Reihe von Interviews mit der technischen und kaufmännischen Leitung von kofinanzierten Projekten sowie mit Vertretern jener Unternehmen und/oder Organisationen, die zwar keine Förderung erhalten haben, welche aber die von ATP-unterstützte Technologie einsetzen, durchgeführt.

Basierend auf dem „Business Reporting System“ werden laufend Evaluierungen – seien sie nun zur Bewertung des Projektfortschrittes oder zur Messung von Impacts bei Projektende – durchgeführt. Diese sind zunächst einmal ausschließlich auf die kurzfristig realisierbaren, direkten Effekte fokussiert. So werden beispielsweise im Rahmen des Monitorings folgende Indikatoren (auf Unternehmensebene) erhoben:

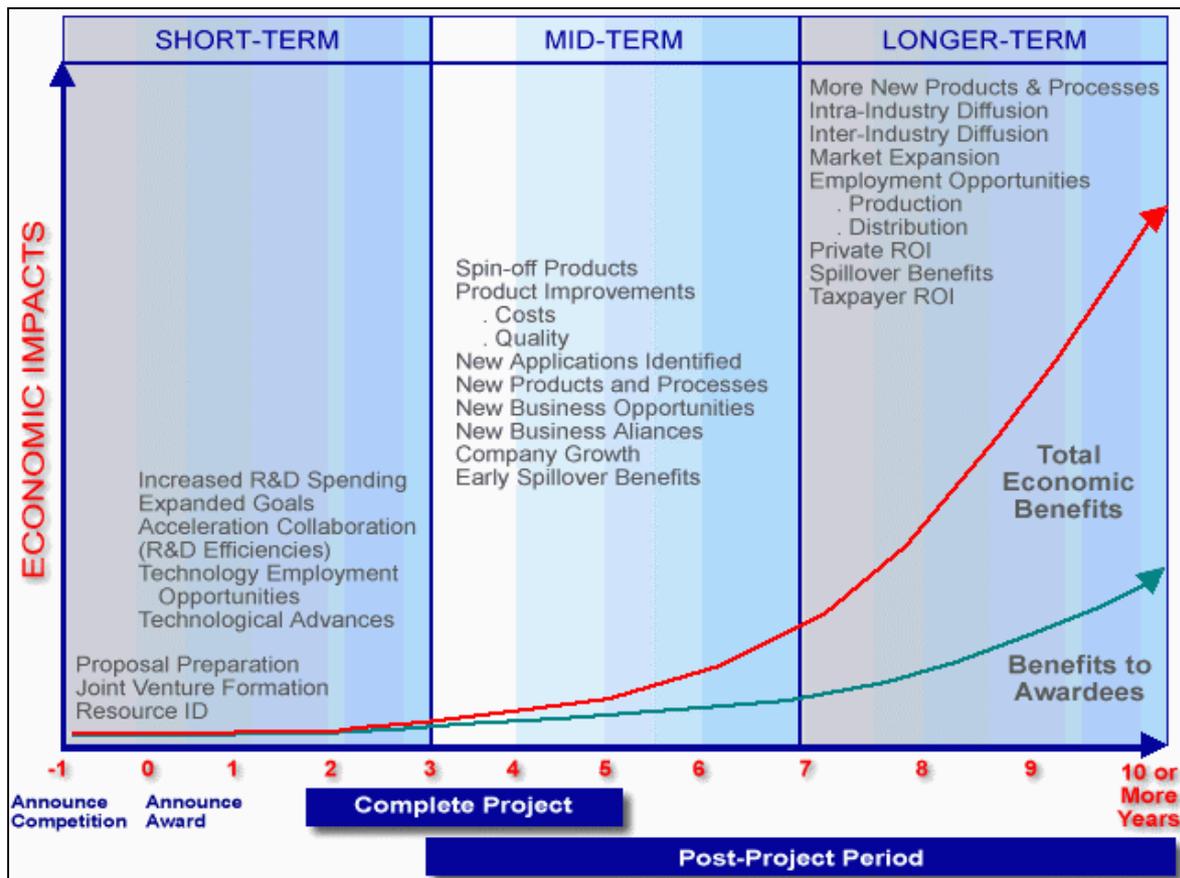
- Entwicklung von Technologien mit verschiedensten Anwendungsmöglichkeiten (Zuordnung der Technologien zu den jeweiligen Anwendungsgebieten)
- Erreichung quantitativer kaufmännischer Ziele (Verbesserung der wirtschaftlichen Leistungsfähigkeit, Kostensenkung, Reduzierung der Produkteinführungszeit)

- Strategien für eine Vermarktung der von ATP kofinanzierten Technologien (Lizenzen, Verwertung des Verfahrens im Unternehmen, Produktion von Gütern oder Dienstleistungen)
- Realisierung von Erträgen (Wahrscheinlichkeit des Eintritts von Rückflüssen)
- Beschleunigung des Entwicklungszyklus (in Jahren)
- Stimulierung von FuE-Investitionen durch die Industrie
- Veränderung in der Beschaffenheit von FuE in den Unternehmen
- Stimulierung von Kooperationen
- Erzielte Wirkungen aus der Zusammenarbeit mit anderen Organisationen oder Unternehmen im Rahmen des ATP-Projektes
- Kosten der Zusammenarbeit (Verzögerung in der Forschungs- oder Markteinführungsphase, erhöhter Koordinationsaufwand)

Diese begleitenden Evaluierungen stellen nur einen Teilabschnitt des Evaluierungsplans von ATP dar, denn dieser bestimmt sich an der Gesamtheit der ATP zugrunde gelegten, in der Regel langfristigen, Dimensionen – Beschleunigung wissenschaftlichen und technologischen Fortschritts, Erzielung privater Erträge aus Kommerzialisierung, Erreichen von gesamtwirtschaftlichen Spillover-Effekten. Die Evaluierung eines ATP-Projektes oder einer Projektgruppe umfasst daher die Abschätzung verschiedenster, diesen Programmzielen zugeordneter Impacts. Dazu zählen beispielsweise, der Umfang bzw. die Größe des Beitrages, den das ATP-Projekt zur wissenschaftlichen bzw. technologischen Erweiterung der Wissensbasis zu leisten vermag, das Diffusionspotential, das Ausmaß der Kooperation bzw. der gemeinschaftlichen Forschung im Rahmen von Joint-Ventures, der Grad der Verkürzung des Produkteinführungs- und Kommerzialisierungsprozesses, Additionalität, Umfang der Spillover-Effekte. Im Rahmen der Evaluierung müssen diese sehr breit gefassten Kategorien durch exakte Indikatoren zur Messung der jeweiligen Impacts konkretisiert werden (Vergleich Abbildung 3).

Jedoch unterscheiden sich diese Wirkungen sehr stark bezüglich ihres zeitlichen Auftretens. Eine Erhöhung von Ausgaben für FuE, die Verbesserung gemeinschaftlicher Forschung und damit auch ihrer Effizienz oder die Erreichung der technologischen Ziele zeigen sich bereits relativ frühzeitig, entweder noch während der Projektlaufzeit oder jedenfalls kurz nach Beendigung. Hingegen sind beispielsweise Diffusionsprozesse, die Erhöhung von Marktanteilen oder gesamtwirtschaftliche Spillovers deutlich längerfristiger in ihrem Realisierungsgrad. Diese unterschiedlichen Zeitpfade des Anfalls der ökonomischen Effekte lassen sich anhand folgender Abbildung darstellen:

Abbildung 3: Zeitpfad der erwarteten ökonomischen Impacts

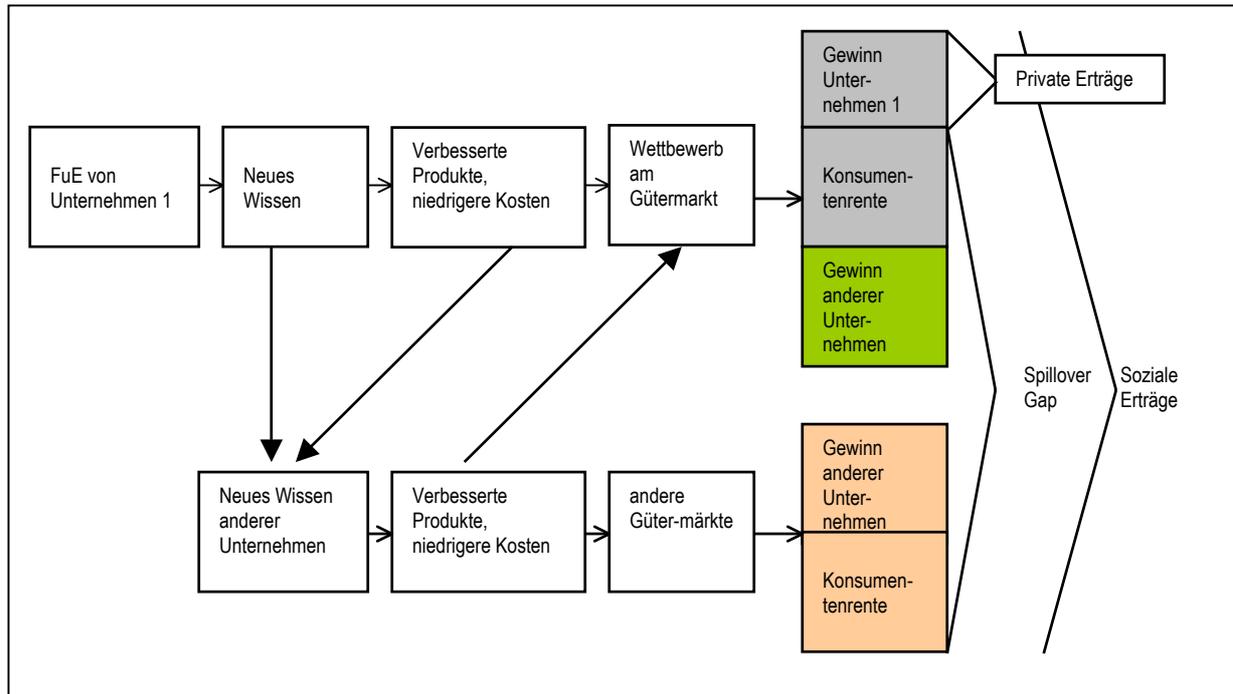


Quelle: Ruegg, 1999, S. 75

Der Erfolg von ATP bemisst sich daher wesentlich auch an diesem langfristigen ökonomischen Nutzen, der entweder direkt über die marktbezogenen Erträge oder indirekt durch Diffusion von Wissen oder institutionelle Effekte anfällt. Beide können Ursache von beträchtlichen Spillovers sein: „Market spillovers“ entstehen aus der Veränderung der Konsumentenrente und sind daher den direkten Effekten zuzuordnen. „Knowledge und network spillovers“ resultieren beispielsweise aus der Wissensdiffusion zu anderen MarktteilnehmerInnen und erzeugen damit indirekt wiederum (gesamtwirtschaftlichen) Nutzen.

Als Spillover bezeichnet man dabei positive externe Effekte, die sich aus der Differenz von privaten und sozialen Erträgen ergeben und die somit nicht unmittelbar dem Unternehmen, das die eigentliche Forschungs- und Entwicklungsarbeit leistet, zugute kommen. Wie bereits erwähnt, gibt es im Rahmen von FuE-Projekten verschiedenste Mechanismen aus denen Spillover-Effekte entstehen können. Abbildung 4 gibt einen Überblick über den Wirkungszusammenhang.

Abbildung 4: Private und soziale Erträge: Market Spillovers, Knowledge Spillovers, Interaktion von Market und Knowledge Spillover



Quelle: Jaffe, 1998, S. 13

Markt-Spillovers ergeben sich aus der erfolgreichen Vermarktung der neuen Technologien, welche entweder in einer Erhöhung der Produktqualität oder einer Verringerung der Produktionskosten resultiert. Der Wettbewerb am Gütermarkt führt dazu, dass der ökonomische Nutzen nicht zur Gänze als Unternehmensprofit (privater Ertrag) lukriert werden kann, sondern dass sich ein Teil auch in einer Erhöhung der Konsumentenrente niederschlägt (Spillover).

Wissens-Spillovers entstehen durch Wissenstransfer (durch Patente, wissenschaftliche Arbeiten, abgeworbene Arbeitskräfte etc.) von einem innovierenden Unternehmen zu einem anderen. Das andere Unternehmen wird dadurch in die Lage versetzt, seinerseits eine Produktverbesserung bzw. Absenkung der Produktionskosten zu erreichen. Der Spillover-Effekt hat sich somit um den Gewinn des anderen Unternehmen und die zusätzliche Konsumentenrente erhöht.

Realistischerweise werden auch Konkurrenzunternehmen von diesem Wissenstransfer profitieren. Durch den daraus resultierenden stärkeren Wettbewerb am Gütermarkt, werden die Gewinne des innovierenden Unternehmens zugunsten des Konkurrenzunternehmens geschmälert (Interaktion zwischen Markt- und Wissens-Spillover). Damit wird zwar der sozialen Nutzen ebenso wie der Spillover-Effekt größer, dieser geht jedoch zu Lasten des Unternehmens, das die eigentliche Forschungs- und Entwicklungsarbeit geleistet hat. Das bedeutet nicht nur, dass ein Teil der Erträge aus Forschung und Entwicklung anderen zufließt, sondern auch, dass die Gewinne aus der Vermarktung der Innovation durch den gestiegenen Wettbewerbsdruck geschmälert werden.

Für ATP ist die Analyse dieser Spillover-Effekte und ihrer Entstehungs- bzw. Wirkungsmechanismen aus mehreren Gründen von zentraler Bedeutung: Die politische Legitimierung von ATP beruht auf der Erwartung, dass das Programm einen Beitrag zur Korrektur des bestehenden Marktversagens (Forschungsaufwand geringer als gesellschaftlich optimal) zu leisten vermag; die sozialen Erträge aus der Existenz von ATP also höher sind, als bei seinem Fehlen. Deren Umfang bestimmt sich wiederum zum überwiegenden Teil durch die positiven Externalitäten (also Spillovers). Es ist daher bereits im Projektauswahlprozess entscheidend, jene Projekte zu identifizieren, die eine möglichst hohe Wahrscheinlichkeit für diese positiven externen Effekte haben. Denn eine Fokussierung auf die ausschließliche Maximierung des privaten Nutzens (anstelle der Spillovers) würde neben dem Risiko der Verdrängung nicht geförderter (aber rentabler) privater Investitionen, auch die Gefahr von Mitnahmeeffekten in sich bergen. Eine ex-ante Abschätzung der zu erwartenden Spillovers – also noch vor dem eigentlichen Projektbeginn – ist daher ein wichtiges Element des Peer-review Verfahrens.

Da also der politische Auftrag von ATP vor allem in der Erzeugung von Spillover-Effekten und der Erzielung hoher sozialer Erträge liegt, muss auch die Evaluierungsstruktur auf die Messung dieser Impacts ausgerichtet sein. Dementsprechend erfolgt die Messung des langfristigen (ökonomischen) Erfolgs von ATP entlang dreier Fragestellungen (siehe Ruegg, 2000):

- (1) Ist das Projektportfolio derart gestreut, dass es einen möglichst hohen sozialen Nutzen stiftet?
- (2) Ist ein wesentlicher Teil dieses sozialen Nutzens direkt ATP zurechenbar? Also: wie hoch ist die Additionalität von ATP?
- (3) Sind die sozialen Nettoerträge deutlich größer als die Summe der privaten Erträge? Wurden also hohe Spillover-Effekte erzielt?

3.3 Evaluierung der ökonomischen Auswirkungen von F&E-Programmen

In der ökonomischen Theorie ist es unbestritten, dass private Forschungsprojekte förderungswürdig sind, wenn es sich um ein öffentliches Gut handelt, Externalitäten in der Produktion oder im Konsum existierten, ein hohes Risiko und Ungewissheit über den Forschungsausgang besteht oder asymmetrische und unvollständige Information vorliegt (Vgl. etwa Stiglitz/Wallsten 2000). Treffen diese Bedingungen zu, so ist zu erwarten, dass der soziale Ertrag von staatlich geförderten Programmen höher als der private Nutzen ist. Selbst bei Vorliegen dieser „Marktversagemomente“ besteht jedoch Bedarf an Evaluierungsstudien, die die Größenordnungen dieser sozialen Erträge quantitativ bestimmen können.

Georghiou /Roessner (2000) stellen dazu drei mögliche Evaluierungsmethoden dar:

- (1) *Output-Analyse*: Evaluierung der Anzahl an neuen Patenten, Wert und Einkünfte aus Lizenzen, etc. Während bei der Output-Analyse an direkt messbaren Resultaten angesetzt wird – und diese daher vorwiegend eine Frage der richtigen Auswahl von Indikatoren ist, ist die Identifikation von ökonomischen Effekten komplizierter.

- (2) *makroökonomische Analyse*: Die Auswirkung der technologische Innovation auf das BIP wird anhand einer angepasste Produktionsfunktion quantifiziert.
- (3) *mikroökonomische Analyse*: Der soziale Nutzen des Projekts wird über den Zuwachs an Konsumenten- bzw. Produzentenrente gemessen.

Der Ansatz der makroökonomischen Analyse beruht auf der Annahme eines Zusammenhangs zwischen FuE-Ausgaben und Produktivität bzw. anderen ökonomischen Outputvariablen. Georghiou/Roessner (2000) weisen auf einige Probleme dieses Ansatzes hin:

- Die Technologie-Variable in der Produktionsfunktion wird nur durch eine grobe Annäherung (meist. FuE-Ausgaben) bestimmt.
- Die Methode weist Unzulänglichkeit im Umgang mit Externalitäten auf.
- Eine Differenzierung zwischen den Auswirkungen von Grundlagenforschung und angewandter Forschung ist schwer möglich, wäre aber für eine Politikevaluation dringend notwendig.
- Die makroökonomische Analyse kann den Zusammenhang zwischen Forschungsausgaben und Wirtschaftswachstum darstellen. Öffentliche Investitionsprojekte können aber, neben dem Einfluss auf das Wirtschaftswachstum, auch einen Fokus auf andere Zielvariable haben.
- Modelle mit Produktionsfunktionen untersuchen ökonomische Effekte der Veränderung der Produktionseffizienz, aber nicht Auswirkungen durch die Entwicklung neuer Produkte.

Im Falle der mikroökonomischen Analyse wird der aggregierte soziale Nutzen den FuE-Ausgaben gegenübergestellt. Georghiou/Roessner (2000) betonen, dass zu zeigen gilt, dass die aggregierte Ertragsrate positiv und der soziale Ertrag höher als der private Nutzen sein muss. Die methodologischen Probleme dieser Methode lassen sich wie folgt zusammenfassen:

- (1) Es fehlt oft der Nachweis eines kausalen Zusammenhangs zwischen dem sozialen bzw. privaten Nutzen und den F&E-Ausgaben.
- (2) Eine Veränderung der Konsumenten- und Produzentenrente kann nur dann nachgewiesen werden, wenn klar bestimmte Angebots- und Nachfragefunktionen existieren.
- (3) Studien über soziale Ertragsraten beruhen auf individuellen Fällen und sind schwer zu generalisieren.
- (4) Langzeiteffekte, die durch den Forschungsoutput generiert werden, können nicht mittels der kurzfristigen Analyse berücksichtigt werden.
- (5) Über den mikroökonomischen Ansatz können mögliche Spill-over Effekte nicht abgeschätzt werden.

Eine weitere vielversprechende Evaluierungsmethode nach Hertzfeld (1992) ist die auf Erhebungsdaten basierende Fallstudie. Die wesentlichen Nachteile einer Fallstudie manifestieren sich in der Objektivität und Expertise des Gutachters, den limitierten quantitativen Informationen sowie dem Kostenfaktor bei einer Vielzahl an Projekten.

Im Folgenden sollen kurz einige Beispiele für einschlägige Ansätze im Rahmen der Evaluierung des ATP angeführt werden:

- 1996 wurde von der CONSAD Research Corporation ein 53 Sektoren umfassendes *regionales Input-Output-Modell* zur Abschätzung der langfristigen Effekte von verbesserten Fertigungsmethoden im Karosseriebau entwickelt. Dabei handelte es sich um ein Projekt, das als Joint Venture („Auto Body Consortium“) - als Zusammenarbeit von Chrysler und General Motors mit kleinen und mittleren Unternehmen der Autozulieferindustrie und den Universitäten von Michigan und Wayne State – konzipiert war.
- 1997 wurde das Research Triangle Institute (RTI) mit der Evaluierung von medizinischen Innovationsprojekten beauftragt. Dazu mussten *sieben einheitliche Fallstudien* erstellt werden, die im Rahmen eines homogenen Evaluierungssystems einander gegenübergestellt wurden. Eine weitere ausführliche Fallstudie (Austin/Macauley, 2000) behandelt zwei Projekte aus dem Bereich der digitalen Speichersysteme (Digital Data Storing). Diese beinhaltet eine umfassende Bewertung des Wohlfahrtsgewinnes der KonsumentInnen. Eine detaillierte Darstellung dieser Studien findet sich im nächsten Abschnitt.
- In einer Studie von Vonortas (1999) wurde erstmals auch ein *Kontrollgruppenansatz* zur Evaluierung der von ATP kofinanzierten Joint Ventures verwendet. Dazu wurde eine Kontrollgruppe bestehend aus nicht subventionierten Joint Ventures auf Basis von Daten des Department of Justice entwickelt. Insgesamt waren Informationen zu 102 ATP Joint Ventures und 510 „non ATP-joint ventures“ verfügbar. Die Fragestellungen umfassten unter anderem Unterschiede in den Technologiefeldern und dem Unternehmenswachstum, der Zusammensetzung der Joint Ventures und dem Kooperationsverhalten innerhalb bzw. auch außerhalb dieser, der Art der Spillover-Effekte etc. Zusätzlich zu der ökonomischen Analyse wurden Fallstudien zu den ATP-Joint Ventures durchgeführt. In einer laufenden Arbeit (zur Beschreibung des Ansatzes siehe Zucker/Darby, 1999) soll ebenfalls ein Kontrollgruppenansatz zur Analyse der Unterschiede zwischen TeilnehmerInnen und Nicht-TeilnehmerInnen eingesetzt werden. Darüber hinaus sieht die Studie auch einen Vergleich der ATP Unternehmen vor, während und nach ihrer Programmteilnahme vor.
- *TeilnehmerInnenbefragungen* werden einerseits laufend, im Rahmen des „Business Reporting System“ durchgeführt (Powell, 1996). Ferner werden periodisch Statusberichte über abgeschlossene Projekte erstellt (siehe Powell 1997, Long 1999, Ruegg 2000).

Im folgenden werden zunächst die Ergebnisse von TeilnehmerInnenbefragungen dargestellt, um einen Überblick über die, mit Standardmethoden abgeschätzten, Programmwirkungen zu geben. In der Folge werden dann Evaluationen in Form von Kosten-Nutzen-Analysen ausführlicher dargestellt. Diese Methode wurde deshalb ausgewählt, weil das ATP in diesem Bereich Ansätze in der Programmevaluierung erprobt hat, die nicht oder kaum in anderen Programmen zu finden sind..

Der nachfolgende Abschnitt fasst Ergebnisse der Wirkungsanalysen zweier fertiggestellter Berichte – dem Progress Report von Powell (1997) und dem Status Report von Long (1999)

– kurz zusammen, um den ökonomischen und technologischen Erfolg von ATP zu illustrieren. Darüber hinaus werden erste Ergebnisse der Analyse von fünfzig abgeschlossenen Projekte (Ruegg, 2000) vorgestellt.

3.4 Evaluation auf der Basis von TeilnehmerInnenbefragungen

In der von *Powell 1997* durchgeführten Evaluierung (basierend auf den Daten des BRS, ergänzt um Aussagen von ProjektteilnehmerInnen) sind alle Unternehmen und Organisationen erfasst, die im Zeitraum 1993 bis 1995 von ATP Zuschüsse erhalten haben. Der Bericht bietet eine Momentaufnahme der Projektfortschritte sowie geplanter Maßnahmen zur Erreichung der technologischen und ökonomischen Ziele. Er umfasst Informationen zu bereits abgeschlossenen Forschungsarbeiten, eine Untersuchung über den Stand der Tätigkeiten im Rahmen der Vermarktungsstrategie sowie eine Analyse anderer Wirkungen der ATP-Projekte, wie zum Beispiel der Stimulierung von Forschungsk Kooperationen, einer Erhöhung privater FuE-Ausgaben, den Umfang von Patenten oder Lizenzen etc. Die zentralen Ergebnisse dieses Berichtes lassen sich wie folgt zusammenfassen:

- In 35% der Fälle führt die Anwendung der entwickelten Technologie zu absoluten Weltneuheiten (sog. „new-to-the-world solutions“).
- Für 86% der Unternehmen hat sich die Forschungs- und Entwicklungsphase deutlich verkürzt.
- Die jeweils entwickelten Technologien bzw. Verfahren sind vielseitig einsetzbar. Pro Projekt konnten mehr als vier mögliche Anwendungsgebiete identifiziert werden.
- Mehr als 100 neue Produkte, Verfahren oder Dienstleistungen befinden sich bereits in der Kommerzialisierungsphase.
- In 85% der Fälle hatte ATP einen positiven Einfluss auf die Stimulierung von Kooperation zwischen Unternehmen und Forschungseinrichtungen.

In dem von *Long 1999* erstellten Statusbericht wurden die ersten 38 abgeschlossenen Projekte von ATP einer Bewertung hinsichtlich zentraler ökonomischer und technologischer Fragestellungen unterzogen. Unter diesen 38 Projekten waren 34 Einzelfirmenprojekte und 4 Joint Ventures (vgl. Abbildung 5, Abbildung 6) mit einem gesamten Projektvolumen von US\$ 130 Mio. Die durchschnittlichen Projektkosten für Einzelfirmenprojekte betragen US\$ 2.9 Mio. Bei Research Joint Ventures sind diese wesentlich höher und machen im Durchschnitt US\$ 7.9 Mio. aus.

ATP übernimmt bei Einzelfirmenprojekten zwar nur die direkten Kosten bis zu einer Höhe von maximal US\$ 2 Mio., jedoch sind diese bei kleinen Unternehmen (hier weniger als 500 Beschäftigte) meist höher als die indirekten Kosten. Es kann also angenommen werden, dass ATP bei vielen dieser Projekte mehr als die Hälfte der Projektkosten finanziert hat. Der Beitrag von ATP betrug insgesamt US\$ 64.6 Mio., was etwas weniger ist als die Hälfte der gesamten Projektsumme. In 19 Projekten betrug der Anteil von ATP mehr als 50% der Gesamtkosten. Dabei handelte es sich ausschließlich um Einzelfirmenprojekte. Bei den Research Joint Ventures lag die ATP Kofinanzierung immer unter 50% der Projektgesamtkosten.

Abbildung 5: Projektvolumina der ATP-Projekte von 1990 – 1997 - Einzelfirmenprojekte

Mio US\$	ATP Anteil an der Finanzierung		Gesamtprojektkosten	
	Zahl der Projekte	In % aller Projekte	Zahl der Projekte	In % von allen Projekten
<1	5	15	2	6
>1, <2	29	85	5	15
>2, <3			14	41
>3, <4			5	15
>4, <5			5	15
>5, <6			3	9
Gesamt	34	100	34	100
Kosten	gesamt	98.4 Mio US\$	durchschnittlich	2.9 Mio US\$

Abbildung 6: Projektvolumina der ATP-Projekte von 1990 – 1997 – Research Joint Venture Projekte

Mio US\$	ATP Anteil an der Finanzierung		Gesamtprojektkosten	
	Zahl der Projekte	In % aller Projekte	Zahl der Projekte	In % von allen Projekten
<5	3	75	1	25
>5, <10	1	25	1	25
>10, <15			2	50
Gesamt	4		4	
Kosten	gesamt	31.9 Mio US\$	durchschnittlich	7.9 Mio US\$

Abbildung 7 zeigt den wissenschaftlichen (Anzahl der Publikationen) und technologischen (Anzahl der erteilte Patente) Output der Projekte: Hinsichtlich des technologischen Output lässt sich feststellen, dass 15 der Projekte bereits über ein oder mehrere Patente verfügen, in Summe erreichen diese Projekte 50 erfolgreiche Patentanmeldungen. Aus 12 Projekten wurden zwar Patente angemeldet, diese jedoch noch nicht erteilt. Davon sind wiederum 8 Projekte zur Gruppe mit bereits erteilten Patenten zuzurechnen. Daraus abgeleitet ergibt sich, dass in 19 der insgesamt 38 untersuchten Projekte Aktivitäten zur Entwicklung patentreifer Technologien eine zentrale Rolle einnahmen.

Das im Rahmen der ATP Projekte entstandene Wissen kann durch Publikation in wissenschaftlichen Journals einer breiteren Öffentlichkeit zugänglich gemacht werden. Zumindest 16 der Projekte konnten ihre Forschungsergebnisse in renommierten Fachjournalen platzieren. Fünf davon haben mehr als 10 Publikationen veröffentlicht.

Abbildung 7: Wissenschaftlicher und technologischer Output

Publikationen			Patente		
Anzahl	Zahl der Projekte	In % von allen Projekten	Anzahl	Zahl der Projekte	In % von allen Projekten
0 oder k.A.	22	58	0 oder k.A.	23	61
1-5	9	24	1	3	8
6-10	2	5	2	5	13
11-20	4	11	3	2	5
Mehr als 20	1	3	4	2	5
			5 oder mehr	3	8

Ersichtlich ist, dass in den Frühphasen der Projektevaluierung der wissenschaftliche Output dominiert. Dies war sowohl vom Zeitpunkt der Messung als auch vom vorwettbewerblichen Charakter des Programms her zu erwarten.

Die nachfolgende Tabelle fasst die Ergebnisse einer Befragung der Projektleiter aus jedem Unternehmen zur Additionalität von ATP zusammen. Konkret wurde an sie die Frage gestellt, ob das Projekt auch ohne finanzielle Zuschuss von ATP durchgeführt worden wäre, sei es nun im gleichen Projektzeitraum oder mit einer zeitlichen Verzögerung. Dabei zeigt sich eine hohe Additionalität von ATP: *66% der Projekte wären ohne ATP Kofinanzierung nicht durchgeführt worden*. Keines der Projekte hätte im gleichen Zeitraum fertiggestellt werden können, die durchschnittliche zeitliche Verzögerung würde 2 Jahre betragen. Darüber hinaus gaben 13 (von 32) Unternehmen an, dass die Zuerkennung von ATP-Mitteln die Aufnahme zusätzlicher Kapitalmittel erleichtert hat. In 23 Fällen trug ATP zur Förderung von Kooperationen bei.

Abbildung 8: Additionalität

Projekt wäre ohne ATP durchgeführt worden	Zahl der Projekte	In % von allen Projekten
Ja, aber mit einer Verzögerung von ... Monaten	11	34
18	4	
21	3	
24	3	
60	1	
Wäre nicht durchgeführt worden	21	66
Gesamt	32	

Eine aktuelle, noch laufende, Evaluierungsstudie (Ruegg, 2000) umfasst die ersten fünfzig abgeschlossenen ATP-Projekte. Sie basiert, ebenso wie der Bericht von Powell (1997) auf den Daten des BRS sowie Interviews mit Projektmanagern und unabhängigen ExpertInnen. Erste Ergebnisse daraus sollen hier kurz dargestellt werden (Ruegg, 2000, S. 121):

- 72% der Unternehmen haben ihre Forschungsarbeiten abgeschlossen.
- 52% der beteiligten Unternehmen haben ihre Forschungsergebnisse veröffentlicht. In Summe wurden zumindest 180 Publikationen gezählt.
- In über 50% der Projekte sind erfolgreiche Patentanmeldungen zu beobachten.
- Mehr als 60% der Projekte konnten die entwickelten Technologien in marktfähige Produkte umsetzen, sei es nun direkt oder als Input in der Herstellung neuer Güter.

- 52% der Projekte verfügen somit sowohl über wissenschaftlichen Output als auch über kommerziellen Erfolg.

TeilnehmerInnenbefragungen über Additionalität und Impacts haben – wie schon erwähnt - allerdings deutliche Grenzen in ihrer Aussagekraft: Zum einen ist ein positiver Antwort-Bias zu erwarten, zum anderen überblicken die Programmteilnehmer naturgemäß nur einen Teil der Effekte – nämlich die direkt bei ihnen anfallenden privaten Kosten. Im Folgenden wird mit der Kosten-Nutzen-Analyse eine ökonomische Methode zur Abschätzung von sozialen Erträgen dargestellt, die diese Beschränkungen aufheben will. Diese soll anhand von Fallbeispielen illustriert werden.

3.5 Evaluation mittels Kosten-Nutzen-Analysen (KNA)

Die KNA dient dem Vergleich von Alternativen und der Evaluierung von öffentlich geförderten Investitionsprojekten. Sie unterscheidet sich von der betriebswirtschaftlichen Investitionsrechnung insofern, als sie neben den betrieblichen Erträgen (private returns) auch die volkswirtschaftliche Effekte (social returns) eines Investitionsprojektes berücksichtigt. Durch die Monetarisierung von Kosten und Nutzen eines Projektes soll ein einheitlicher Bewertungsmaßstab geschaffen werden. Dabei berücksichtigt die KNA auch das Zeitprofil der Kosten und Nutzen, da insbesondere bei größeren oder langfristig angelegten Projekten Kosten und Nutzen in unterschiedlichen Zeiträumen anfallen. Die angestrebte Maximierung des Gegenwartswerts aller Nutzen vermindert um den aller Kosten, soll der Rationalisierung der Verwendung finanzieller Mittel im öffentlichen Sektor dienen.

Man unterscheidet in der KNA zwischen einer ex-post und einer ex-ante Analyse der Auswirkungen von Forschungsprogrammen. Die ex-ante Analyse birgt das Problem der Schätzung der zukünftigen Kosten und des erwarteten Nutzens, wobei über erstere deutlich sicherere Vermutungen angestellt werden können als über letztere. Bei der ex-post Analyse wiederum besteht die Ungewissheit über zeitverzögerte Auswirkungen (sehr spät anfallende, positive Effekte werden i.d.R. nicht berücksichtigt). Ebenfalls können bei der Evaluierung von abgeschlossenen Projekten Schwierigkeiten bei der Zurechnung und Ursächlichkeit der Kosten und des Nutzens auftreten.

Wegen der unterschiedlichen Realisierungszeitpunkte müssen alle verfügbaren Daten über Kosten- und Nutzenfaktoren gesammelt, vollständig monetarisiert und abdiskontiert werden, um eine Vergleichbarkeit einzelner Projekte herzustellen. Die Definition der Diskontrate (Kapitalnutzungskosten, Inflation, Opportunitätskosten, etc.) ist durch die verwendete Methode bestimmt. Tabelle 5 gibt einen Überblick über mögliche Quellen des Nutzens und den Kosten der Betroffenen:⁵

⁵ siehe Polt, W. (1998)

Tabelle 5: Beispiele für im Rahmen eines Projektes anfallende Kosten und Nutzen

	Förderungsnehmer	Förderungsgeber	Gesellschaft
Kosten	<ul style="list-style-type: none"> • Opportunitätskosten durch Projektteilnahme • Direkte Teilnahmekosten • Verlust von Förderungen aus anderen Programmen 	<ul style="list-style-type: none"> • Steuergelder • Projekt- und Verwaltungskosten 	<ul style="list-style-type: none"> • Opportunitätskosten (Einkommen von anderen potentiellen Förderungsprogrammen) • Kosten des Förderungsprogramms
Nutzen	<ul style="list-style-type: none"> • Steigerung des Nettogewinns • Zusatznutzen durch Transfers • nicht-monetäre Leistungen 	<ul style="list-style-type: none"> • Steuereinnahmen • Absenken anderer Förderungen 	<ul style="list-style-type: none"> • Erhöhung des Volkseinkommens (direkter Effekt) • Spill-over Effekte (indirekter Effekt)

Für die Monetarisierung von Projektkosten und -nutzen stehen drei Standardmethoden zur Verfügung:

Der *Kapitalwert (NPV – Net Present Value)* beschreibt die Differenz zwischen dem erwarteten zukünftigen sozialen Nutzen (N_t) und den gegenwärtigen Kosten (K_t). Ein Investitionsprojekt ist nur dann volkswirtschaftlich profitabel, wenn die Differenz zwischen dem abdiskontierten Nutzen und den abdiskontierten Kosten positiv ist. Als Diskontsatz (r) wird der reale erwartete Zinssatz verwendet.

$$NPV = \sum_{t=0}^n \frac{N_t - K_t}{(1+r)^t}$$

Die *Methode des internen Zinssatzes (IRR – Internal Rate of Return)* ermöglicht die Berechnung einer (über Projekte vergleichbaren) Kapitalrendite, indem der Kapitalwert der Investition gleich Null gesetzt und der Diskontsatz r explizit dargestellt wird. Ist die Kapitalrendite größer dem erwarteten realen Zinssatz, so kann die Investition positiv bewertet werden.

$$IRR = \sum_{t=0}^n \frac{N_t - K_t}{(1+r)^t} = 0$$

Zusätzlich kann das *Kosten-Nutzen-Verhältnis*, als normierte Kennzahl, berechnet werden.

$$\frac{N}{K} = \frac{\sum_{t=0}^n \frac{N_t}{(1+r)^t}}{\sum_{t=0}^n \frac{K_t}{(1+r)^t}}$$

Die Bewertung von Kosten ist dann wenig problematisch, wenn Märkte hinreichend funktionieren und Marktpreise zur Anwendung gelangen können. Hingegen ist die Bestimmung von Preisen für auftretende externe Effekte deutlich schwieriger. Sehr häufig kommt es daher zu einer Unterschätzung der Kosten durch die unvollständige

Berücksichtigung von negativen externen Effekten. Der Nutzen eines Projektes wird prinzipiell an der Zahlungsbereitschaft für die Inanspruchnahme des Projektoutputs gemessen. Verzerrungen entstehen meist dadurch, dass Angebots- und Nachfragekurven falsch ermittelt und anstelle der Zahlungsbereitschaften betriebswirtschaftlich erzielbare Einnahmen zu Grunde gelegt werden.

Darüber hinaus ist auch die Wahl der Diskontrate ein entscheidender Aspekt, vor allem dann, wenn unterschiedliche Projekte mit verschiedenen zeitlichen Ausgabenprofilen verglichen werden. Welcher Diskontfaktor gewählt wird, hängt einerseits von der Zeitpräferenz (werden Gegenwartskosten oder Zukunftsnutzen höher eingeschätzt) ab, andererseits aber auch von der politischen bzw. persönlichen Zielsetzung der Evaluatoren. Eine Entscheidungshilfe kann dabei die Berechnung des internen Zinsfußes darstellen, da dieser zeigt, wie niedrig die soziale Diskontrate jedenfalls anzusetzen ist, damit das Projekt einen positiven Kapitalwert bringt. Um die Unsicherheit bezüglich der Diskontrate zu minimieren, sollten Sensitivitätsanalysen eingesetzt werden, welche die Robustheit der Berechnungen gegenüber Änderungen zentraler Annahmen überprüfen.

Obwohl die KNA eine Vielzahl an methodischen Schwächen und Einschränkungen aufweist, so ist sie doch für die Bewertung technologiepolitischer Programme von Bedeutung, wenn es sich – wie im Falle von ATP – um die Abschätzung von gesellschaftlichem Nutzen handelt. Wie bereits ausgeführt, ist es Ziel des Programms, den sozialen Ertrag der getätigten Forschungsprojekte zu erhöhen, weshalb auch entsprechende Evaluierungsmethoden entwickelt wurden. In den hier diskutierten Fallstudien gelangte die oben diskutierte Kosten-Nutzen-Analyse, welche mit Hilfe von Fallstudien durchgeführt wurde, zur Anwendung. Sieben Projekte aus dem Bereich der medizinischen Innovation, sowie zwei Projekte zum Thema digitale Speichersysteme wurden im Detail analysiert.

3.5.1 Fallstudie I: Medizinische Innovationsprojekte

Das Center for Economic Research des Research Triangle Institute (RTI) wurde im Rahmen der ATP-Forschungsauftrages damit betraut, ein einheitliches Evaluierungssystem für ein breites Feld von Technologieprojekten zu entwickeln und diese Methode in sieben Fallstudien über ATP-Projekte anzuwenden. Folgende sieben Projekte wurden aus dem Technologiefeld Medizinische Innovationsprojekte untersucht:

Tabelle 6: Überblick- zu untersuchende Projekte

ATP Project Title ⁶	Duration	Funding Level
In-Depth Case Studies		
Stem Cell Expansion	2 years	\$ 1,220,000
Biopolymers for Tissue Repair	3 years	\$ 1,999,000
Living Implantable Microreactors	3 years	\$ 4,263,000
Proliferated Human Islets	3 years	\$ 2,000,000
Brief Case Studies		
Biomaterials for Clinical Prostheses	3 years	\$ 1,999,000
Gene Therapy Applications	3 years	\$ 1,996,000
Universal Donor Organs	3 years	\$ 1,999,000

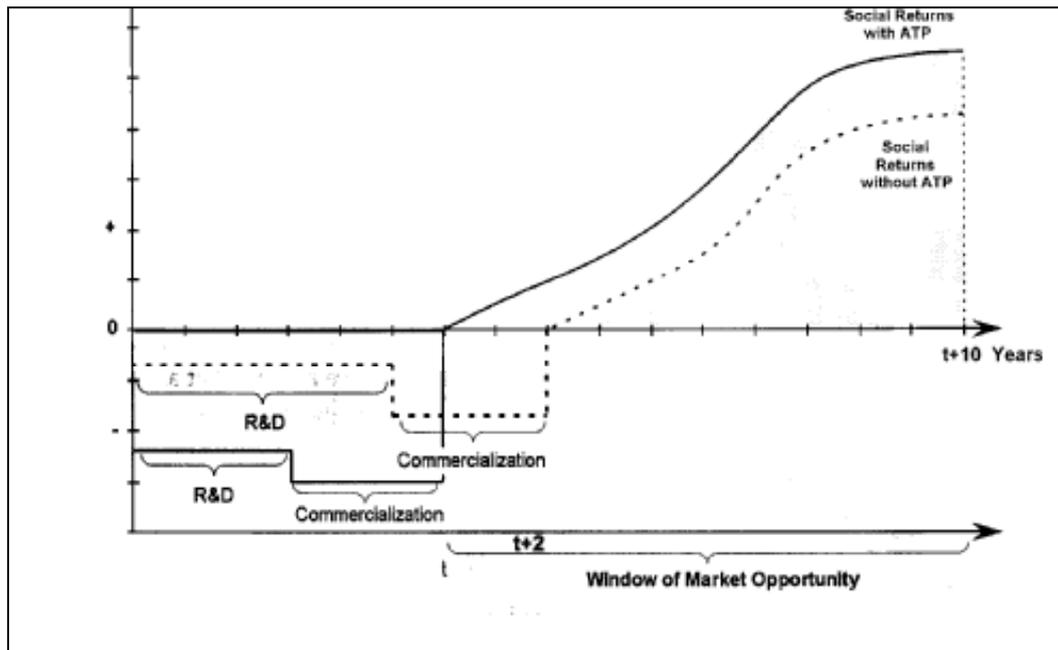
Quelle: Martin, Sheila A. et al. (1998), S. 1-13.

Der Schwerpunkt der Fallstudien lag auf der Untersuchung des sozialen Nutzens der neuen medizinischen Anwendungen. Zeit, Budget und Datenrestriktionen verlangten nach einigen Einschränkungen: Pro Projekt wurden nur die Auswirkungen einer Anwendung untersucht, auch wenn es eine Vielzahl an Outputs gab. Zusätzlich wurde der Zeithorizont auf eine zehnjährige Verwendung der Innovation beschränkt, bevor eine neue technologische Entwicklung eintritt. Die Schätzung der Effekte kann folglich als sehr konservativ betrachtet werden und liegt am unteren Ende der möglichen Bandbreite.

Nachfolgende Abbildung (Abbildung 9) zeigt die dieser Fallstudie zugrundeliegenden Annahmen über den Projektverlauf. Sie beschreibt einerseits die zeitliche Entwicklung einer technischen Innovation bis zur Marktreife und andererseits den Nettoertrag während jeder Periode. Zu Beginn des Forschungsprojektes, der FuE-Phase, liegt der Fokus auf der Lösung des technischen Problems. Wird diese Phase erfolgreich abgeschlossen, folgt die Kommerzialisierung der Innovation, indem in Verkauf, Marketing und Produktion investiert wird. Erst nachdem das Produkt die Marktreife erlangt hat, beginnt die letzte Phase, in welcher der sozialen Erlös lukriert werden kann. Unterstellt wird eine 10-jährige Nutzungsdauer der Innovation bevor sie durch eine neue Technologie ersetzt wird. Das Advanced Technology Program versucht diesen Zyklus auf mehreren Ebenen zu beeinflussen.

⁶ für vollständige Titel der Projekte siehe Martin, Sheila A. et al. (1998), p 1-13

Abbildung 9: Hypothetischer Projektverlauf



Quelle: Martin, Sheila A. et al. (1998), pp 2-8.

Mögliche Kanäle der Einflussnahme auf den Projektverlauf durch eine ATP-Subvention sind:

- Eine Beschleunigung der Produkteinführung durch die subventionsbedingte Verkürzung der FuE-Zeit. Griliches/Pakes/Hall (1987) zeigen, dass die Wahrscheinlichkeit eines Forschungserfolges eine Funktion der FuE-Ausgaben eines bestimmten Jahres ist.
- Eine Verminderung des Risikos durch erhöhte Finanzierung und folglich höhere Chancen eines Forschungserfolgs.
- Eine Erweiterung des Einsatzgebietes des Forschungsoutputs durch eine Verbreiterung des Forschungsvorhaben.

Der zusätzliche soziale Ertrag der Innovation wurde nun insofern gemessen, als der Projektverlauf sowohl mit als auch ohne ATP-Finanzierung simuliert wurde. Die Differenz hängt vom Einfluss der drei oben genannten Wirkungskanäle ab. Kann die FuE-Phase durch eine ATP-Finanzierung maßgeblich verkürzt werden, so endet die Kommerzialisierungsphase (Produktentwicklung) früher und folglich verlängert sich die unterstellte Produktionsphase. Der durch Unternehmensbefragung eruierte Einfluss der drei Wirkungskanäle ist in nachfolgender Tabelle dargestellt.

Tabelle 7: Einfluss der möglichen Wirkungskanäle von ATP

ATP Project Title	Project Acceleration (years)	Increase in the Probability of Success (percent)	Widening of Technology Applications
Stem Cell Expansion	1 to 2	9%	None reported
Biopolymers for Tissue Repair	At least 10	171%	Significant but not quantified
Living Implantable Microreactors	2	11%	None reported
Proliferated Human Islets	3 to 5	2%	None reported
Biomaterials for Clinical Prostheses	2	1%	None reported
Gene Therapy Applications	2	20%	Some effect reported but not quantified
Universal Donor Organs	1 to 2	16%	None reported

Quelle: Martin, Sheila A. et al. (1998), pp 1-23

Die entwickelten Methoden ermöglichen den soziale Nutzen der Projekte, in Form von Kapitalwert (NPV) und internen Zinssatz (IRR), darzustellen. Nachfolgende Tabelle 8 zeigt, dass Projekte mit ATP-Finanzierung einen Kapitalwert von \$47 Millionen bis \$17.7 Milliarden aufweisen und einen internen Zinssatz, der in der Spannweite von 21-148% liegt. Der gesellschaftliche Nutzen pro Projekt variiert aufgrund der Anzahl der behandelten Patienten, der Auswirkungen auf die Gesundheitskosten, sowie der Wahrscheinlichkeit eines Forschungserfolges. Der Gesamteinfluss der ATP-Förderungen wird für die sieben Projekte mit einem NPV von \$34.3 Milliarden und einer IRR von 116% beziffert.

Tabelle 8: Net present value und internal rate of return

ATP Project Title	Project Time Horizon	NPV (1996 Mio. US\$)	IRR (%)
Stem Cell Expansion	1992 to 2009	\$ 47	21%
Biopolymers for Tissue Repair	1994 to 2009	\$ 98	51%
Living Implantable Microreactors	1994 to 2009	\$ 17,750	148%
Proliferated Human Islets	1995 to 2008	\$ 1,297	34%
Biomaterials for Clinical Prostheses	1993 to 2010	\$ 15,058	128%
Gene Therapy Applications	1995 to 2011	\$ 945	111%
Universal Donor Organs	1992 to 2011	\$ 783	92%
Composite	1992 to 2011	\$ 34,258	116%

Anmerkung: Die Composite Berechnung des NPV setzt sich aus der Summe des erwarteten Ertrages und der Kosten pro Jahr zusammen (unterschiedliche Projektdauer).

Quelle: Martin, Sheila A. et al. (1998), pp 1-21

Abschließend können die Werte dem privaten Ertrag der Forschungsprojekte gegenübergestellt werden. Aus einem Vergleich von Tabelle 8 und Tabelle 9 ist deutlich

ersichtlich, dass der kumulierte gesellschaftliche Nutzen der Innovationen um ein Vielfaches höher ist als jener des einzelnen Unternehmens.

Tabelle 9: Composite Private Returns

Composite Private Returns	NPV (1996\$ million)	IRR (%)
Project returns (composite)	\$1,564	12%
Increment attributable to ATP	\$914	13%

Quelle: Martin, Sheila A. et al. (1998), pp 1-23

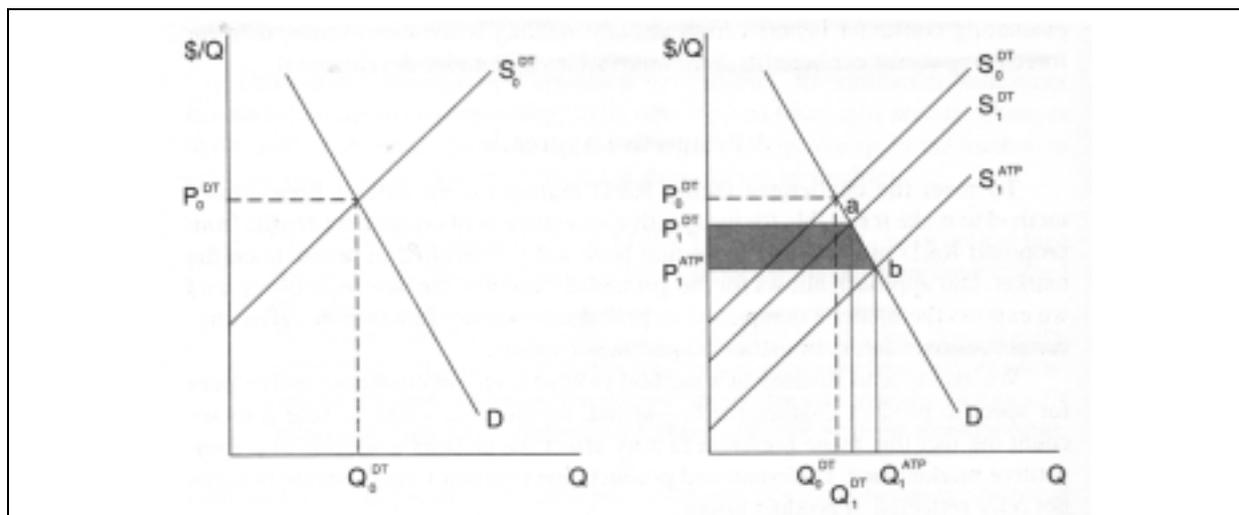
Während mit diesem Ansatz plausibel gezeigt werden kann, dass die betrachteten ATP-Projekte hohe soziale Erträge hatten, und diese die privaten Erträge auch substanziell übersteigen, dürfte er nur beschränkt zur Auswahl unter verschiedenen Projekten taugen, da die einzelnen Schätzungen mit großen Unsicherheiten behaftet sind. Allerdings lassen sich für einzelne besonders erfolgversprechende Projekte Größenordnungsunterschiede feststellen (klinische Prothesen, Mikroreaktoren).

3.5.2 Fallstudie II: Digitale Speichersysteme (Digital Data Storage - DDS)⁷:

Die zweite Gruppe von Fallstudien behandelt zwei Projekte aus dem Bereich der digitalen Speichersysteme. Mit diesen neuen Technologien wird eine verbesserte Lese- und Schreibgeschwindigkeit sowie eine höhere Speicherkapazität angestrebt. Die Projektkosten betragen mit \$21 Millionen (eingereichte ATP-Finanzierung \$10.4 Millionen) für den „magnetic tape recorder“ bzw. \$2.9 Millionen (eingereichte ATP-Finanzierung von \$2 Millionen) für das „optical tape drive“.

Zur Abschätzung dieser Effekte wurde eine neue Methode zur Abschätzung des Konsumentennutzens angewandt. (siehe Austin/Macauley 2000) Normalerweise wird der KonsumentInnennutzen aus einer Verschiebung der Angebotskurven ermittelt.

Abbildung 10: Derived demand for new technologies: Illustration of net surplus change



Quelle: Austin, David; Macauley, Molly (2001), p. 242

⁷ siehe Austin and Macauley (2000)

Die neue Technologie (S_1^{ATP}) erlaubt die Befriedigung des gleichen Konsumentennutzens zu niedrigeren Kosten. Wenn man in Rechnung stellt, dass auch die existierende Technologie weiterentwickelt wird (ausgedrückt durch die Verschiebung der Angebotskurve von S_0^{DT} nach S_1^{DT}), dann ist der Zusatznutzen für die Konsumenten aus der neuen Technologie gleich der schattierten Fläche.

Schwierig wird diese Abschätzung dann, wenn sich ökonometrisch keine Nachfragekurve (D) schätzen lässt (z. B. weil die Effekte der Produktinnovation auf die Qualität und die Reaktion der Konsumenten schwer quantitativ abzuschätzen sind).

Die Autoren schlagen für einen solchen Fall vor, eine von Bresnahan (1986) entwickelte Methode zu verwenden. Dieser Ansatz verwendet eine Abschätzung der Veränderung der Kosten für die Konsumenten in Folge der Innovation. Diese wird hier als Funktion der KonsumentInnennachfrage, der Marktadoptionrate und den KonsumentInnenpräferenzen hinsichtlich einer Verbesserung der Leistungsfähigkeit von DDS konstruiert. Für die Berechnung werden zwei Szenarien konstruiert und in ihren Wirkungen verglichen: Das sog. „Baseline-Szenario“ unter Beibehaltung der alten Technologie (DT) sowie das „Alternativ-Szenario“ der Einführung der technologischen Innovation.

Für die Berechnung wird der sogenannte *Törnquist-Kosten-Index* verwendet, wobei die digitalen Speichersysteme als intermediäre Güter betrachtet werden. Der Törnqvist-Index ist das (gleich) gewichtete geometrische Mittel aus den zwei Standard-Indizes mit denen üblicherweise Preisveränderung gemessen wird: aus (a) dem *Laspeyres-Index*⁸, der hier zur Messung der Höhe der Kompensation verwendet wird, die man den Konsumenten zahlen müsste, damit sie auf die Innovation verzichten („willingness to accept“) und (b) dem *Paasche-Index*⁹, den man zur Messung der Zahlungsbereitschaft der Konsumenten für die Innovation heranziehen kann. („willingness to pay“)

Die Präferenzen der KonsumentInnen werden bezüglich der Variablen Speicherkapazität (CAP), Transferrate (TR) und Datenzugriffszeit (FAT) abgefragt. Es wird angenommen, dass die KonsumentInnen höhere Kapazitäten und Transferraten sowie niedrigere Zugriffsraten präferieren. Veränderungen in den jeweiligen nominalen Stückpreisen repräsentieren dabei die jeweiligen KonsumentInnenpräferenzen. Um die qualitätsangepassten Preise der existierenden Technologie sowie der Innovation zu berechnen, wird der Schattenwert¹⁰ jeder

⁸ Beim Laspeyres-Index wird der Warenkorb $q_0^1, q_0^2, \dots, q_0^n$ der Basiszeit 0 auf die Berichtszeit 1 übertragen:

$$\left(\sum_{i=1}^n p_1^i \cdot q_0^i \right) : \left(\sum_{i=1}^n p_0^i \cdot q_0^i \right)$$

Der Divisor gibt somit reale, der Dividend fiktive Umsatzwerte wieder.

⁹ Der Paasches-Index ist durch Übertragung des Warenkorbs vom Zeitpunkt 1 ($q_1^1, q_1^2, \dots, q_1^n$) auf die Basiszeit 0 gekennzeichnet:

$$\left(\sum_{i=1}^n p_1^i \cdot q_1^i \right) : \left(\sum_{i=1}^n p_0^i \cdot q_1^i \right)$$

Der Dividend ist also eine reale, der Divisor ein fiktive Umsatzgröße.

¹⁰ „Schattenwerte“ existieren nicht als reale Marktpreise, sondern sind errechnete Größen, die die „wirklichen“ sozialen Kosten und Erträge widerspiegeln sollen.

Variablen (β_{CAP} , β_{TR} , β_{FAT}) mittels hedonistischer Regression¹¹ berechnet und zu den Verkaufspreisen der Produkte („Regalpreise“) in Beziehung gesetzt.

$$W^{DT} = p^{DT} + \beta_{CAP}(\Delta CAP) \cdot (\Delta CAP > 0) + \beta_{TR}(\Delta TR) \cdot (\Delta TR > 0) - \beta_{FAT}(\Delta FAT) \cdot (\Delta FAT < 0)$$

und

$$W^I = p^I + \beta_{CAP}(\Delta CAP) \cdot (\Delta CAP < 0) + \beta_{TR}(\Delta TR) \cdot (\Delta TR < 0) - \beta_{FAT}(\Delta FAT) \cdot (\Delta FAT > 0)$$

wobei W^{DT} und W^I die qualitätsangepassten Preise der existierenden Technologie (defender technology) bzw. der Innovation darstellen. p ist der erwartete (nominelle) Regalpreis.

Die Veränderung der „Kosten der Dienstleistung“ für die KonsumentInnen wird mittels dem Törnqvist-Kosten-Index geschätzt: (Service Cost) durch die Innovation.

$$C^{*DT} = \frac{E^*(u^{DT}, P^{DT}, W^{DT})}{E^*(u^{DT}, P^I, W^A)} \quad \text{und} \quad C^{*I} = \frac{E^*(u^I, P^{DT}, W^{DT})}{E^*(u^I, P^I, W^A)},$$

wobei $E(\bullet)$ die Ausgabenfunktion der Konsumenten, u das jeweilige Nutzenniveau und P die Preise von anderen Gütern angeben. Da angenommen wird, dass die Innovation nicht sofort übernommen wird, wird ein durchschnittlicher Preis W^A unter Einbeziehung der Adoptionsrate berechnet. C^{*I} entspricht dabei dem Paasche-Index („willingness to pay“) und bezeichnet die Kosten für das Erreichen des optimalen Nutzenniveaus im Innovationsszenario. C^{*DT} ist demgemäß der Laspeyres-Index („willingness to accept“) und entspricht den Kosten für die Erzielung des optimalen Nutzenniveaus im Ausgangsszenario, also ohne Einführung der neuen Technologie.

Unter der Annahme, dass die Ausgabenfunktion der Konsumenten als Translog-Funktion darstellbar ist (Caves, et al., 1982), reduziert sich der Törnqvist-Index auf

$$\frac{1}{2} \ln(C^{*DT} \times C^{*I}) = \left(\frac{1}{2} (s^{DT} + s^I) \left(\frac{W^{DT}}{W^A} \right) \right),$$

wobei s^{DT} und s^I jenen Teil der DDS-Ausgaben repräsentieren, die im Falle der existierenden bzw. innovativen Technologie vom KonsumentInnen aufgewendet werden muss. Dieser Index wird von den Autoren für die Jahre 2000-2005 prognostiziert. Basis für die KonsumentInnenausgaben und die Ausgaben für DDS (Digital-Data-Storage) sind vorhandene Daten aus der Vergangenheit, für die Veränderungsrate der Preise von DDS werden Annahmen getroffen.

Dieser Ansatz wurde nur zur Abschätzung der Effekte von zwei DDS-(Digital-Data-Storage) Technologien verwendet, nämlich dem linearen Scannen von Magnetbändern und einer Weiterentwicklung von Lese/Schreibtechnologien bei optischen Bändern. Da das zugrundeliegende Modell eine große Anzahl von Schätz-Parametern enthält

¹¹ Detaillierte Ausführungen dazu, siehe Austin, David Macauley (2001), p.23 ff

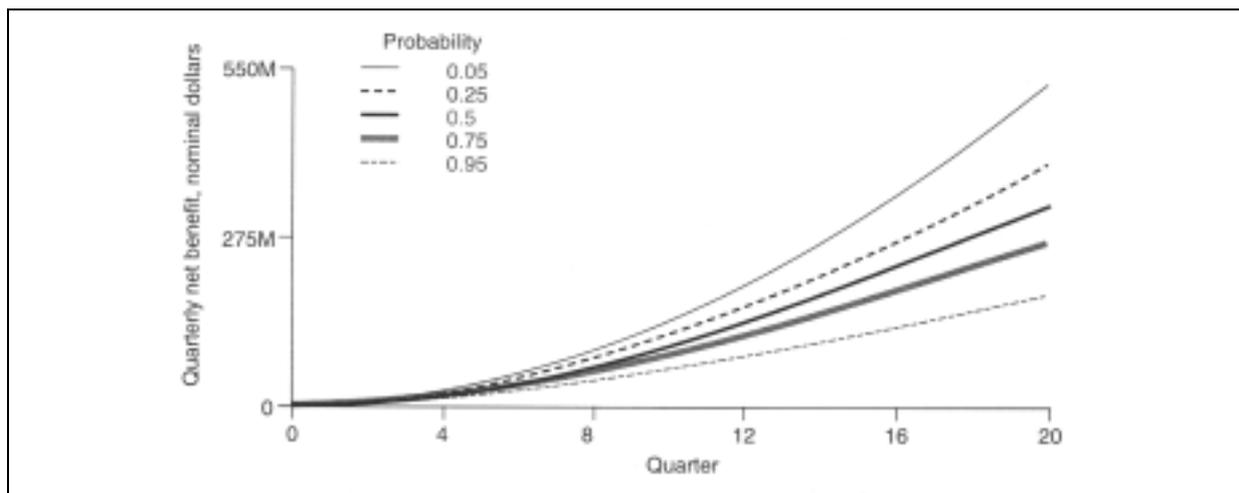
(Verbesserungen durch die neue Technologie, gleichzeitige Verbesserung der alten Technologie, Marktbedingungen, Bandbreiten der Konsumenteneinschätzungen etc.) wurden Wahrscheinlichkeitsbandbreiten für die zu erwartende Zusatznutzen der neuen Technologie verrechnet (s. Tabelle 10)

Tabelle 10: DDS Innovation - Nettogegenwerte des Wohlfahrtsgewinnes der KonsumentInnen (über fünf Jahre, in Milliarden US\$, 2000)

Wahrscheinlichkeits-Percentile	Linear Scanning	Optical Tape
5 th	1.25	1.05
25 th	1.79	1.30
Median	2.16	1.45
75 th	2.53	1.62
95 th	3.17	1.88

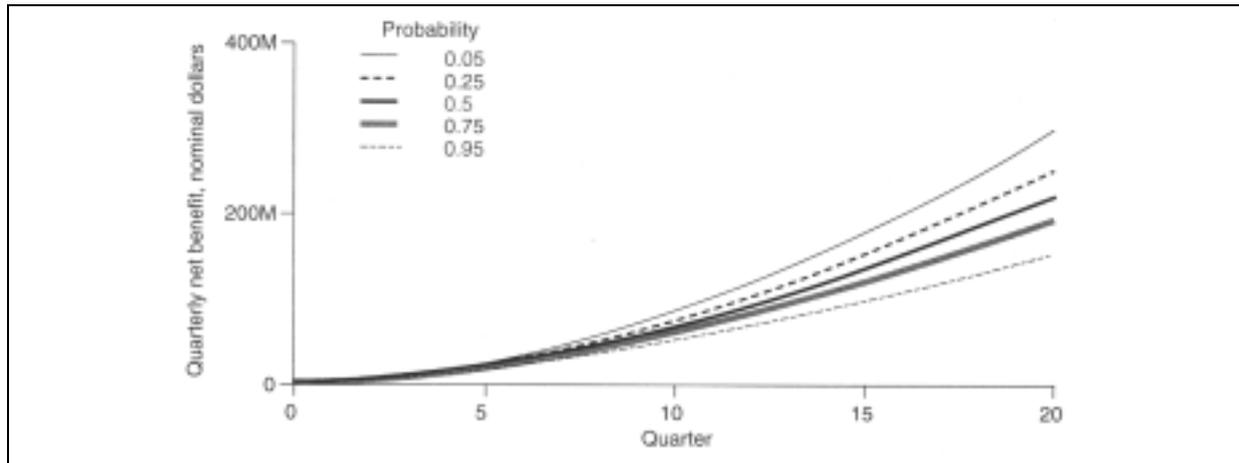
Quelle: Austin und Macauley (2000). D.h. dass mit einer 95% Wahrscheinlichkeit mindestens 1,25 bzw. 1,05 Mrd. US\$ lukriert werden und mit einer 5% Wahrscheinlichkeit 3,17 bzw. 1,88 Mrd. US\$.

Abbildung 11 Consumer benefit given successful introduction: Linear scanning innovation vs. defending products



Quelle: Austin, David; Macauley, Molly (2001), p. 246

Abbildung 12 Consumer benefit given successful introduction: Optical tape innovation vs. defending products



Quelle: Austin, David; Macauley, Molly (2001), p. 246

Diese Resultate hängen natürlich entscheidend von den Annahmen über Geschwindigkeitszuwachs, Technologieadaption, Preisveränderungen, Schattenwert und Entwicklung der Marktgröße ab, wobei gerade die Größe des Marktes und die Adaptionrate die größten Einflussfaktoren in der Berechnung der Konsumentenrendite waren.

Trotz – nach Angabe der Autoren – konservativen Schätzungen über die Erträge, weisen also beide, im Rahmen des ATP entwickelten – Technologien substantielle soziale Erträge auf, die die Förderung rechtfertigen. Der Kosten-Index-Ansatz konnte dabei als zusätzliches Instrument zur Abschätzung dieser sozialen Erträge nachgewiesen werden. Seine Stärken liegen in der Möglichkeit Konsumentennutzen auch dort anschätzen zu können, wo keine empirisch erhobenen Nachfragekurven existieren. Weiters kann durch die simultane Berücksichtigung einer Vielzahl von Variablen Unsicherheit modelliert werden und die Sensitivität der Resultate gegenüber der Veränderung einzelner Parameter dargestellt werden.

3.5.3 Möglichkeiten und Grenzen der KNA

Die KNA stellt durch die systematische, objektive Bewertung von Kosten und Nutzen und der expliziten Berücksichtigung von Annahmen ein wichtiges Instrument der Wirkungsanalyse dar. 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. Die KNA wird bereits in sehr vielen Bereichen als Instrument der Evaluierung eingesetzt, am häufigsten in der Umwelt- und Arbeitsmarktpolitik. Dennoch weist sie in ihrem methodischen Konzept einige Schwächen auf, die dementsprechend Auswirkungen auf Einsatzmöglichkeiten und Ergebnisse haben:

Bei Unsicherheit in bezug auf das Eintreten bzw. den zeitlichen Anfall von Erträgen (z. B. bei Innovationsprojekten) ist die KNA nur ein unzureichendes Instrument für eine ex-ante Evaluierung. Die monetäre Bewertung des zu erwartenden Nutzens würde beispielsweise

schon durch die Wahl eines falschen Zeithorizonts (zu kurz oder zu lang) zu verzerrten Ergebnissen führen.

Bei der ex-post Analyse wiederum besteht das Problem der Zurechnung und Ursächlichkeit von Kosten und Nutzen. Die Frage, welcher Anteil der – kosten- und nutzenstiftenden – technologischen Veränderung auf das Programm zurückzuführen ist, wird umso schwieriger beantwortbar, je mehr Zeit vergangen ist. Gleichzeitig ist dies jedoch auch die Voraussetzung dafür, dass ökonomische Auswirkungen überhaupt realisierbar sind.

Eine weitere Schwierigkeit besteht in der Bewertung intangibler Effekte. Beispiele dafür wären die Erhöhung der Adoptionsfähigkeit, das Fortschreiten organisatorischer Innovationen, die Förderung technologischer Lernprozesse etc. Da für diese intangiblen Effekte keine Marktpreise existieren und auch Opportunitätskosten nur schwer konstruierbar sind, würde eine KNA für jene Programme zu einer systematischen Fehleinschätzung gelangen, deren Zielsetzung sich an diesen Effekten orientiert.

Darüber hinaus ist die KNA eine Analysemethode, die in ihrer klassischen Form auf die Bewertung von wenigen und größeren Projekten zugeschnitten ist. Sie erscheint zwar prinzipiell machbar, aber weniger praktikabel für die Evaluierung einer größeren Zahl von Kleinprojekten, da hier die Kosten der Datenbeschaffung sowie der Ertragsabschätzung sehr hoch wären.

Insgesamt kann das ATP als Programm eingestuft werden, das

- (a) ein sehr ausgebautes Monitoring und Berichtswesen hat, was die Datenlage für Evaluierungen sehr verbessert.
- (b) Mit einigen Evaluierungsansätzen experimentiert und innovative Ansätze entwickelt
- (c) Insbesondere bezüglich der Anwendung von Kosten-Nutzen-Analysen weiter ist als vergleichbare europäische Programme.

3.6 Anwendbarkeit des Evaluierungsansatzes in Österreich

Im vorigen Abschnitt wurden die prinzipiellen Möglichkeiten und Grenzen der KNA beschrieben. Umgelegt auf die Anwendungsmöglichkeit von KNA zur Evaluierung österreichischer Technologie- und innovationspolitischer Programme und Initiativen bedeutet das:

- (1) KNA ist eine wichtige Methode zur Abschätzung von sozialen Erträgen und der weiteren ökonomischen Effekte von Programmen. In Österreich ist sie bisher im Rahmen der Bewertung von Technologie- und Innovationspolitik nur vereinzelt zum Einsatz gekommen (vgl. etwa Leo, 1991) und sollte verstärkt angewendet werden.
- (2) Aufgrund der methodischen Grenzen sollte die KNA in der Regel nicht alleiniger Untersuchungsansatz sein, sondern zusammen mit anderen Methoden zum Einsatz kommen (etwa verschiedene Ansätze der Additionalitätsmessung).
- (3) Aufgrund der praktischen Grenzen der KNA (Schwierigkeiten und Kosten der Datenbeschaffung etc.) wird sie wohl am ehesten bei größeren Projekten geringer Zahl, bei denen der soziale Ertrag relativ leicht identifiziert und mit nicht zu großen Einschränkungen auch monetarisierbar ist zur Anwendung kommen können.

Jedenfalls übertragbar aus dem ATP scheinen aber Monitoring und Berichtswesen sowie der innovative Einsatz mehrerer Evaluierungsmethoden.

Diese Randbedingungen für die Anwendung lassen sich z.B. bei missionsorientierten Programmen mit explizit gesellschaftlichen Zielsetzungen (neben der im engeren Sinne technologiepolitischen) finden. Dazu zählen in Österreich etwa die Programme des BMVIT im Bereich ‚Verkehrstechnik‘ („Intelligente Infrastruktur“, „MOVE“ etc.), im Bereich ‚Nachhaltig Wirtschaften‘ sowie die weltraumbezogenen Aktivitäten. Auch Programme in den Bereichen ‚Energie, Gesundheit, umweltrelevanter FuE und Innovation‘ wären erfolversprechende Anwendungsfelder. Allerdings ist – wie schon oben angemerkt – die Liste der Programme, die mittels KNA analysiert werden könnten, keinesfalls taxativ. Die Beschränkungen des Anwendungsbereiches ergeben sich in Österreich vor allem durch den – in den meisten Fällen – gegenüber den geringen Programmvolumina vergleichsweise hohen Kosten für diesen Evaluierungsansatz.

Jedenfalls übertragbar scheint aber insgesamt der strukturierte und systematische Evaluierungsprozess mit dem die einzelnen Projekte des ATP auf ihre sozialen Erträge überprüft werden.

3.7 Literatur

- Austin, D.; Macauley, M. (2000): Estimating Future Consumer Benefits from ATP-Funded Innovation: The Case of Digital Data Storage. www.atp.nist.gov/eao/abstracts-eao.htm.
- Austin, David; Macauley, Molly (2001): Estimating Future Benefits from ATP Funding of Digital Data Storage. In: *The Advanced Technology Program: Assessing Outcomes*. National Academy Press, Washington D.C., 239 – 248.
- Bresnahan, T. (1986): Measuring the Spillovers from Technical Advance: Mainframe Computers in Financial Services”, in: *American Economic Review*, 76(4):742-755.
- Feldmann, Maryann (2001): Assessing the ATP: Halo Effects and Added Value. In: *The Advanced Technology Program: Assessing Outcomes*. National Academy Press, Washington D.C., 131 – 135.
- Feldmann, Maryann, Kelley, Kelley Maryellen R., (2001): Leveraging Research and Development: The Impact of the Advanced Technology Program; In: *The Advanced Technology Program: Assessing Outcomes*. National Academy Press, Washington D.C., 189 – 210.
- Feller, Irwin (2001): Perspectives on Program Evaluation. In: *The Advanced Technology Program: Assessing Outcomes*. National Academy Press, Washington D.C., 123 – 130.
- Fritz, O.; Grasenick, K.; Gruber, M.; Hartmann, C.; Novakovic, M.; Polt, W.; Sturn, D.; Woitech, B.; Zinöcker, K. (2001): EvinA – Evaluierung innovativen Aktionen in der Technologie-, Struktur- und Arbeitsmarktpolitik: Entwicklung von Methoden, Indikatoren und “Good Practice”. Projektbericht gefördert durch das Bundesministerium für Wissenschaft und Verkehr. Wien – Graz.
- Georghiou, L.; Roessner, D. (2000): Evaluating technology programs: tools and methods. In: *Research Policy* 29, 657 – 678. Elsevier. Amsterdam.
- Hertzfeld, H. (1992): Measuring the returns to space research and development. In: Hertzfeld, H.; Greenberg, J. Eds., *Space Economics*. American Institute of Astronautic, Washington D.C.
- Hill, C. (1999): An Overview of the Program’s History and Objectives. In: *The Advanced Technology Program – Challenges and Opportunities*, National Academy Press, Washington D.C, 38-42.
- Jaffe, A. B. (1998): The Importance of “Spillovers” in the Policy Mission of the Advanced Technology Program. In: *The Journal of Technology Transfer*, Vol. 23, Summer 1998, 11-21.
- Kammer, R. (1999): Opening Remarks. In: *The Advanced Technology Program – Challenges and Opportunities*, National Academy Press, Washington D.C, 31-37.
- Link, Albert N. (2001): Enhanced R&D Efficiency in an ATP-funded Joint Venture. In: *The Advanced Technology Program: Assessing Outcomes*. National Academy Press, Washington D.C., 223 – 238.
- Long, W. (1999): Performance of Completed Projects: Status Report Number 1, NIST Special Publication 950-1.
- Martin, S. A.; Winfield, D. L.; Kenyon, A. E.; Farris, J. R.; Bala, M. V.; Bingham, T. H. (1998): A Framework for Estimating the National Economic Benefits of ATP Funding of Medical Technologies – Preliminary Applications to Tissue Engineering Projects Funded from 1990 to 1996. Final report. Research Triangle Institute, Center for Economics Research.

- Polt, W (1998): "Evaluation and Impact Assessment" – Advanced Training Course in Science & Technology Policies and Programmes. Paper prepared for presentation at Universidad Carlos III. Madrid.
- Powell, J. (1996): The ATP's Business Reporting System: A Tool for Economic Evaluation, Paper presented at Conference on Comparative Analysis of Enterprise Data, Helsinki, Finland. <http://www.atp.nist.gov/eao/helsinki.htm>.
- Powell, J. (1997): Development, Commercialization, and Diffusion of Enabling Technologies: Progress Report for Projects Funded 1993-1995, U.S. Department of Commerce.
- Ruegg, R. (1998): The Advanced Technology Program, Its Evaluation Plan, and Progress in Implementation. In: *The Journal of Technology Transfer*, Vol. 23, Summer 1998, 5-11.
- Ruegg, R. (1999): Assessment of ATP. In: *The Advanced Technology Program – Challenges and Opportunities*, National Academy Press, Washington D.C., 71-81.
- Ruegg, R. (2000): Taking a Step Back: An Early Results Overview of Fifty ATP Awards. In: *The Advanced Technology Program: Assessing Outcomes*, National Academy of Sciences.
- Solomon Associates (1993): *The Advanced Technology Program, An Assessment of Short-Term Impacts: First Competition Participants*.
- Spender, J. C.: The ATP and Program Management. In: *The Advanced Technology Program – Challenges and Opportunities*, National Academy Press, Washington D.C., 114-115.
- Stiglitz, Joseph E.; Wallstein, Scott J. (2000): *Public-Private Technology Partnerships: Promises and Pitfalls*. In: *Public-Private Policy Partnerships*, The MIT Press, Massachusetts.
- Vonortas, N. (1999): *Business Diversification Through Research Joint Ventures: Advanced Technology Program*. Gaithersburg, MD: National Institute of Standards and Technology.
- Wessner, C. W. (Ed.) (2001): *The Advanced Technology Program: Assessing Outcomes*, National Academy Press, Washington D.C.
- Zucker, L.; Darby, M. (1999): Performance Measures as Indicators of ATP Effects on Long-term Business Success. In: *The Advanced Technology Program – Challenges and Opportunities*, National Academy Press, Washington D.C, 81-82.

4 METHODEN DER EVALUATION ÖKONOMISCHER EFFEKTE VON TECHNOLOGIEPROGRAMMEN: DIE B.E.T.A METHODE ZUR BEWERTUNG VON ESA PROGRAMMEN

*Dorothea Sturn
Joanneum Research*

4.1 Einleitung

Bei den ESA Programmen handelt es sich um den Prototyp **missionsorientierter Programme**: Vergleichsweise viel öffentliches Geld fließt in ein möglichst klar umrissenes gesellschaftliches Ziel, welches der Markt nicht oder nur unzureichend erreichen kann. Die Programme sind auf genau spezifizierte und eng definierte Endprodukte (Satelliten, Launcher) fokussiert und benötigen große kritische Massen, es partizipieren eher große, forschungsintensive Firmen und Forschungseinrichtungen.

Da die eigentliche Mission (nationale Sicherheit, nationales Prestige) oft unklar oder auch in der öffentlichen Diskussion umstritten ist, wurden in der Vergangenheit unter zunehmendem Legitimationsdruck für die hohen öffentlichen Ausgaben vermehrt Untersuchungen zur Quantifizierung des "ökonomischen Nutzens" oder auch "Zusatznutzens" durchgeführt. Dieses "Zusatznutzenkonzept" stammt historisch aus der Diskussion um die Budgets für Verteidigung. Da der Primärnutzen dieser Ausgaben (im Prinzip das öffentliche Gut nationale Sicherheit) nur begrenzt in der Lage ist, stark wachsende Ausgaben öffentlich zu legitimieren, konzentrierte sich die Beurteilung der Sinnhaftigkeit der öffentlichen Ausgaben auf den ökonomischen Nutzen als "Zusatznutzen".

Dieser Zusatznutzen, in der Raumfahrt in der Regel "**Spin-off**" genannt, wird definiert als "Transfer von F&E Ergebnissen, die für einen ursprünglich nicht geplanten Anwendungszweck industriell genutzt werden". Neben Produktinnovationen, die direkte Folge der in Raumfahrt investierten F&E sind (Teflonpfannenargument), handelt es sich bei Spin-offs auch um verbesserte Organisations- und Managementtechniken, um Qualifikationsverbesserungen etc. Rentieren sich die Raumfahrtausgaben schon bereits für diese Spin-offs, dann - so das Argument -

4.2 ESA: Daten und Ziele

Die ESA (European Space Agency) wurde 1975 mit der Motivation gegründet, europäische Weltraumanstrengungen zu koordinieren und effizienter zu gestalten.

4.2.1 Ziele der ESA

Die ESA selbst gibt folgenden **Zielkatalog** an (vgl. ESA 1984, Leo 1991 sowie <http://www.esrin.esa.it/index.html>):

1. **Förderung der Kooperation** europäischer Staaten auf den Gebieten Weltraumforschung, Weltraumtechnologie und deren Anwendung für wissenschaftliche und operationale Raumfahrtapplikationen (ausschließlich friedliche Zwecke) durch:

- langfristiges europäisches Weltraumprogramm
 - Koordination nationaler Programme mit dem europäischen
 - Industriepolitische Maßnahmen
2. Erleichterung des **Austauschs von wissenschaftlichen und technischen Informationen** über Weltraumforschung und Weltraumtechnologie (wissenschaftliche Ergebnisse sollen öffentlich zugänglich gemacht werden, Erfindungen und Patente der ESA allen Mitgliedsstaaten gebührenfrei zur Verfügung gestellt werden)
3. **Industriepolitische Zielsetzungen:**
- Steigerung der Wettbewerbsfähigkeit der europäischen Industrie
 - Hohe Kosteneffizienz (cost effective) der Forschung und Entwicklung sowie der Anwendung
 - Sicherung der Beteiligung der Industrie der Mitgliedsstaaten gemäß ihren Beiträgen

4.2.2 Finanzierung der ESA Programme

Zur Verfolgung dieser Zielsetzungen bedient sich die ESA einer Programmorganisation und Finanzierung, die Ansätze einer **variablen Geometrie** zeigt: Alle Mitgliedsstaaten partizipieren in den Aktivitäten zur Weltraumforschung sowie in einen gemeinsamen Set an Programmen (mandatory programmes). Darüber hinaus können die Mitgliedsländer ihr Partizipationsniveaus bei den sogenannten "optionalen" Programmen wählen.

Die Beiträge zum "Pflichtprogramm" werden entsprechend der ökonomischen Stärke des Landes gewichtet, die Beiträge zu den optionalen Programmen ergeben sich aufgrund der – gewählten - Teilnahmeintensitäten in den einzelnen Programmen.

Das Gesamtbudget der ESA betrug 1998 2592,6 Mio ECU (davon 2089,1 Mio als Beiträge der Mitgliedsländer), Österreichs Beitrag beläuft sich mit insgesamt 30,3 Mio ECU auf etwas 1,4% der Gesamtbeträge, davon sind 12,8 Mio ECU Pflichtbeitrag.

Die ESA – und das ist mit eine der industriepolitischen Zielsetzungen – garantiert den Mitgliedsländern einen weitgehenden **Rückfluss ihrer Beiträge** in Form von Aufträgen an die nationalen Industrien und Forschungseinrichtungen.

4.2.3 Bewertung des Zielsystems

Bewertet man das ESA Zielsystem entlang den Kategorien von "good practice" für den Zusammenhang von Zielsysteme und Evaluation (vgl. die im Rahmen dieses Projektes vorgestellten arbeiten im rahmen des FTSP EvinA), so kann ein potentieller Evaluator diesen Zielen nur sehr kritisch gegenüber treten:

Die Ziele sind vage, allgemein, unspezifiziert und nicht einmal in Ansätzen konkretisiert oder operationalisiert. Sätzen wie die "Steigerung der Wettbewerbsfähigkeit" bleiben leere Worthülsen, solange dies nicht genauer spezifiziert wird. Es wird hier nicht einmal die leiseste qualitative These über den Wirkungszusammenhang zwischen den ESA Programmen und der Wettbewerbsfähigkeit der europäischen Industrie formuliert.

Entsprechend verwundert auch der Widerspruch zwischen angekündigter industriepolitischer Zielsetzung und den rein technischen Kriterien, nach denen Programme und Projekte bewertet und ausgewählt werden.

Eine konsistente Evaluation, die auch überprüft, ob und in welchem Ausmaß gesetzte Ziele erreicht wurden, ist angesichts dessen nahezu unmöglich. Nichtsdestotrotz konzentrieren sich Evaluationsstudien von Raumfahrtaktivitäten von Ex-post-Evaluationen und dort auf die Abschätzung bzw. Quantifizierung der ökonomischen Aspekte.

Im folgenden wird ein kurzer Überblick über die verschiedenen methodischen Ansätze zur Messung ökonomischer Auswirkungen von Raumfahrtprogrammen gegeben. Im Zentrum steht die Ermittlung der Spin-offs nach der Methodologie des BETA Instituts der Universität Straßburg, die im Auftrag der ESA dreimal für Kosten-Nutzen-Analysen angewandt wurde.

4.3 Überblick über Methoden der Evaluierung von Raumfahrtaktivitäten

Die Methoden können grob wie folgt unterschieden werden (für einen Überblick vgl. Hertzfeld 1992 und Leo 1991):

- Fallstudien
- Ökonometrische Modelle bzw. Growth Accounting Ansätze
- Vorher – Nachher - Vergleiche: Kosten-Nutzen-Analysen bzw. Methoden zur Ermittlung des gesamtwirtschaftlichen Nutzens

4.3.1 Fallstudien

Fallstudien sind aufwendig und daher relativ teuer. Darüber hinaus gelangen sie in der Regel nur zu wenig quantitativen Aussagen (eine Ausnahme ist Scientific Consulting 1989) und Verallgemeinerungen der qualitativen Aussagen sind nicht immer möglich. Dennoch können sie sehr verdienstvoll sein bei der Klärung von Begriffen und der Katalogisierung bzw. Systematisierung von Wirkungen. Im Kontext der Evaluation von Raumfahrtprogrammen wurden mit Hilfe von Fallstudien die angewandten Technologien sehr genau analysiert und mögliche Bereiche für Spin-off identifiziert (z.B. Welles – Waterman 1964, Scientific Consulting 1989 – für einen Überblick vgl. Leo 1991). Eine zweite, gerade im Kontext der Analyse ökonomischer Wirkungen, wichtige Funktion von Fallstudien kann die sein, Basisdaten für die spätere Modellierung von Wirkungen zu erfassen. So schreibt Hertzfeld resümierend in seinem Methodenüberblick:

... the most promising type of economic study for measuring returns to space R&D is the documentation of actual cases based on direct survey information. Case studies provide relatively clear examples of the returns, both theoretical and actual, to space R&D investments. A well structured series of case studies coupled with good theoretical modelling aimed at integrating the data and incorporating those data with more general economic measures of benefits may well be the way to establish reliable and "believable" economic measures of the returns to space R&D (Hertzfeld 1992, S. 168)

Bei neueren Studien, die in verschiedenen europäischen Ländern auf nationaler Ebene durchgeführt wurden, um künftige Raumfahrtstrategien auszuarbeiten, handelt es sich auch im wesentlichen um Fallstudien, allerdings steht die Strategieentwicklung gegenüber der Bewertung von Wirkungen vergangener Aktivitäten deutlich im Vordergrund. Es sind damit sehr formativ angelegte Evaluationen, die meist auf Interviews und Workshops basieren¹. Tichy und Posch (1999) kritisieren in ihrem Überblick vor allem, dass die meisten dieser Studien von Raumfahrtexperten erstellt wurden,

"die ihrem Fachgebiet zwangsläufig positiv gegenüberstehen..... es wird nicht einmal versucht zu begründen, oder gar abzuschätzen, wieweit staatliche Mittel in der Raumfahrt besser angelegt sind als in alternativen Forschungs- bzw. Technologiefeldern" (S. 7).

4.3.2 Ökonometrische Modelle

Ökonometrische Modelle sind in der Regel makroökonomische Produktionsfunktionsschätzungen, bei denen versucht wird, die Beziehung zwischen Forschungsausgaben (bzw. Forschungsausgaben im Rahmen des untersuchten Programms) und outputbezogenen Größen (in der Regel Produktivität) zu quantifizieren. Solche Modelle machen nur dann Sinn, wenn die Ausgaben so groß sind, dass sich deutliche ökonomische Wirkungen zeigen können. Im Kontext der Wirkungsevaluation von Weltraumausgaben finden sich ökonometrische Modelle daher fast ausschließlich in den USA. Den Stärken solcher Modelle (Quantifizierung von Primär – und Sekundärnutzen, quantitatives Maß für die Stärke des Zusammenhangs zwischen Programm und intendierter Wirkung, makroökonomische Perspektive und damit Erfassung von Wirkungen auf die gesamte Volkswirtschaft) stehen folgende Schwächen gegenüber² (vgl. Georghiou und Roessner 2000):

- In der Produktionsfunktion geht als Input die Daten über R&D Ausgaben ein, ohne dass der Effekt dieser Ausgaben auf das Produktionssystem als solches hinreichend geklärt ist
- Externalitäten als Resultat von Forschungsaktivitäten können nicht hinreichend abgebildet werden
- Die Mission des untersuchten Forschungsprogramms beinhaltet in der Regel andere bzw. weitreichende Ziele als die Stimulierung ökonomischen Wachstums. Dies lässt sich nicht abbilden
- Die jeweiligen Beiträge von Grundlagenforschung und angewandter Forschung können nicht differenziert werden
- Die ökonomischen Wirkungen inkrementaler Veränderungen der Effizienz können abgebildet werden, nicht aber die Wirkungen völlig neuer Produkte oder radikaler Veränderungen in der Technologie

¹ Eine Ausnahme stellt die norwegische Studie (Amundsen und Skovholt 1994) dar, die eine quantitative Ermittlung des Spin-off Faktors versucht. Vgl. hierzu kritisch Tichy und Posch 1999)

² Diese Schwächen gelten für alle Arten von Produktionsfunktionsanalysen, nicht nur für ihren Einsatz bei der Evaluation der Wirkung von Weltraumausgaben.

4.3.3 Vorher-Nachher-Vergleiche

Ein prinzipielles methodologisches Problem, welches mit Vorher-Nachher-Vergleichen einhergeht ist, dass die eigentlich relevante Fragestellung hinsichtlich der Wirkungen von Interventionen nicht die ist, wie sich die Situation nach einer Intervention von einer hypothetischen Situation ohne Intervention unterscheidet. Im Vorher-Nachher-Vergleich kann die Wirkung der Intervention nur schlecht von den Wirkungen anderer Einflüsse separiert werden. Der Grund, weshalb diese Vergleiche dennoch vergleichsweise attraktiv sind ist, dass die Modellierung eines Referenzpfades (also die hypothetische Situation ohne die Intervention) häufig auf methodische Schwierigkeiten stößt (vgl. die Methoden zur Konstruktion von Kontrollgruppen im Abschnitt zu diffusionsorientierten Programmen).

Insgesamt stellen Vorher-Nachher-Vergleiche die umfangreichste Methodengruppe zur Evaluierung von Raumfahrtprogrammen dar. Wenn auch in ihrer konkreten Ausgestaltung unterschiedlich, so versuchen doch alle Ansätze den "Nutzen" der Raumfahrt zu messen. Dabei kommen meist nur Sekundärnutzenkonzepte zur Anwendung, selten wird auch der Primärnutzen untersucht (dies tun beispielsweise Teubal und Steinmüller 1982).

Eine der wichtigsten Studien in diesem Zusammenhang sind die Untersuchungen des Sekundärnutzens von ESA Programmen des B.E.T.A. (Bureau d'Économie Théorique et Appliquée Strassbourg). Diese Untersuchung wurde bereits dreimal im Auftrag der ESA durchgeführt und stellt die wichtigste europäische Untersuchung zu diesem Thema dar. Es handelt sich im Prinzip um eine modifizierte **Kosten Nutzen Analyse**: Dem Zusatznutzen, also den Spin-offs der ESA werden Kosten in Höhe der an die Unternehmer vergebenen Aufträge gegenübergestellt.

Methode und Ergebnisse werden im folgenden genauer dargestellt:

4.4 Der B.E.T.A. Ansatz

4.4.1 Ziele und Umfang der Evaluation

Die **Ziele** der drei von B.E.T.A. durchgeführten Evaluationen betreffend der Ermittlung der ökonomischen, indirekten Effekte (Spin-off Effekte) der ESA_Programme sind wie folgt definiert:

- Quantitativ = Effektivität eines Programms mittels Abschätzung der **minimalen** indirekten wirtschaftlichen Effekte der ESA Kontrakte
- Qualitativ und präskriptiv = ökonomische, wissenschaftliche und organisatorische Wirkungen der Programme auf die Programmteiligen, genauere Beschreibung des Spin-off Phänomens

Es handelt sich bei dieser Evaluation um eine reine Wirkungsanalyse, es sind keine anderen Evaluationselemente enthalten.

4.4.2 Methode

Motivation

Die hier angewendete Methode basiert auf direkten Interviews von B.E.T.A. Mitarbeitern mit den ESA-Vertragspartnern. **Es handelt sich hiermit um eine mikroökonomische Analyse indirekter Effekte (Spin-offs) basierend auf Primärerhebungen.** Die Anzahl der geführten Interviews ist dabei so groß, dass die Ergebnisse als repräsentativ für alle ESA Vertragspartner angesehen werden kann.

Der Grund für diese recht aufwendige Art der Datenerhebung ist, dass komparativ statistische Analysen nicht anwendbar sind, um die Diffusion des in den ESA Programmen erlangten Wissens in andere Bereiche abzubilden: Es gibt schlicht keine Daten, welche diese Wissensströme abbilden.

Beschreibung

Die angesprochenen indirekten Effekte, welche B.E.T.A. adressiert, umfassen:

"The different learning processes undergone by firms during their work for ESA, affecting them in many varied ways (widening of scientific, technical and "organisational" knowledge, innovation in products and procedures, new links with new external organisations, etc.) and applied to activities other than ESA contracts (space or non-space related activities)" (Cohendet, 1997, S.201)

Zur Ermittlung und Quantifizierung dieser indirekten Effekte wird zunächst die folgende Typologie entworfen:

Tabelle 11

<p>Technologische Effekte</p> <p>Derivate von ESA Produkten</p> <p>Neue Produkte</p> <p>Verbesserung bestehender Produkte</p> <p>Diversifikation</p>	<p>Kommerzielle Effekte</p> <p>Internationale Kooperation</p> <p>Neue Vertriebsnetzwerke</p> <p>ESA als Marketing Referenz</p>
<p>Effekte auf Organisation und Methoden</p> <p>Qualitätskontrolle</p> <p>Projektmanagement</p> <p>Produktionstechnik</p>	<p>Effekte, den Faktor Arbeit betreffend</p> <p>Formation einer kritischen Masse aus Spezialisten</p> <p>Verbesserung der Qualifikation und Fähigkeiten</p>

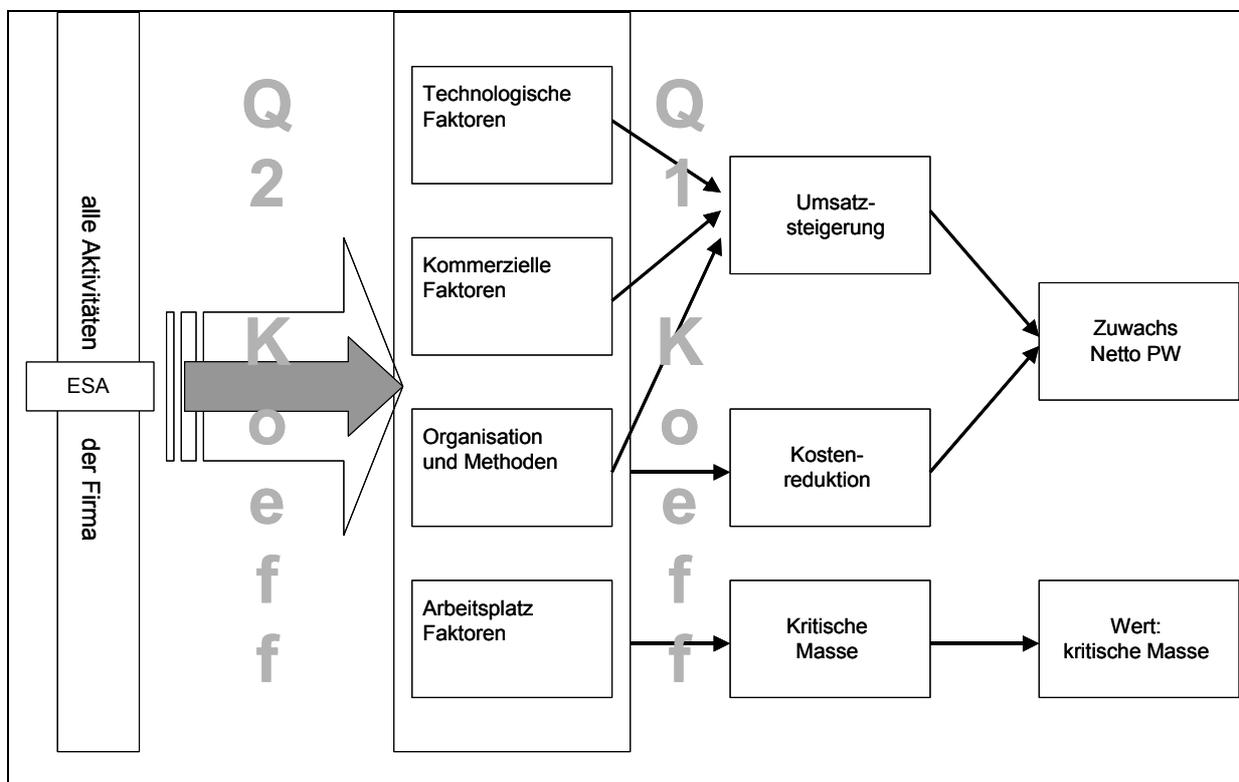
Die Effekte werden wie folgt **quantifiziert**:

Der Nutzen der ESA wird gebildet als die Differenz zwischen der tatsächlichen Struktur der Unternehmen und einer hypothetischen Situation (ohne ESA). Damit wird ein ambitionierter Schritt weg von einem schlichten Vorher-Nachher-Vergleich hin zu einem **echten "Mit**

(hypothetisch) Ohne-Vergleich" unternommen. Allerdings gehen in den Vergleich nur die Vertragspartner selbst sowie ihre direkten Lieferanten ein, andere Wirkungen auf die Volkswirtschaft bleiben vernachlässigt (vgl. die Kritik am Ansatz unten).

Wie bereits erwähnt, wird der Nutzen aus **Interviews mit dem Management der Unternehmen** ermittelt. B.E.T.A. geht davon aus, dass das Management den Unterschied zur hypothetischen Situation kennt und quantifizieren kann. Als Nutzenindikator dient der Zuwachs am Netto-Produktionswert, der sich aus Umsatzsteigerungen und reduzierten Kosten zusammensetzt. Die befragten Personen geben damit pro Wert zwei Koeffizienten an: erstens die Wirkung des jeweiligen Effekts auf den Umsatz (Q1) und zweitens der Anteil des Effektes, welcher auf die ESA zurückzuführen ist (Q2). Die folgende Abbildung zeigt das Prinzip der Quantifizierung:

Abbildung 13 Prinzip der Quantifizierung



Implementierungsschritte

1. Festlegen der **Typologie** der indirekten Effekte
2. Bestimmung der **Koeffizienten Q1 und Q2 durch Befragungen**
 Q1 gibt den Einfluss des jeweiligen Faktors (Technologie = Q1T, Kommerziell = Q1C, Organisation = Q1O) auf den Umsatz an
 Q2 gibt den Einfluss der ESA auf den jeweiligen Faktor an (Technologie = Q2T, Kommerziell = Q2C, Organisation = Q2O)
3. **Berechnung der Wirkung** der drei Effekte (Technologischer Effekt, Kommerzieller Effekt, Organisation und Methoden) auf den Nettoproduktionswert (**NPW**)

Technologischer Effekt: Umsatz x Quote des NPW x Q1T x Q2T

Kommerzieller Effekt: Umsatz x Quote des NPW x Q1C x Q2C

Organisatorischer Effekt: Umsatz x Quote des NPW x Q1O x Q2O

(z.B. Der Umsatz des Unternehmens beträgt 100 Mio öS, die NPW Quote 60%. Die Manager schätzen, dass die technologischen Faktoren 10% des Umsatzes induzieren und dass die ESA Verträge mindestens zwei Drittel der technologischen Faktoren bewirken. Damit kann der Effekt der technologischen Faktoren auf den NPW berechnet werden:

$100 \text{ Mio öS} \times 0,6 \times 0,1 \times 0,66 = 4 \text{ Mio öS.}$)

4. Berechnung der Effekte bei **den direkten Lieferanten** der Vertragspartner

Es gehen hier lediglich die Lieferanten ein, die Produkte fertigen, welche einen der indirekten Effekte betreffen. Die entsprechenden Koeffizienten werden nicht bei den Lieferanten selbst, sondern bei den Vertragspartnern erhoben.

5. Quantifizierung der **Kostenreduktion**

Hier werden die Kostenreduktionen erhoben, die aufgrund verbesserter Organisation und Arbeits- bzw. Produktionsmethoden anfallen

6. Quantifizierung der **Kritischen Masse**

Die Manager schätzen die kritische Masse: Sie geben die für die kritische Masse notwendige Anzahl an Personen an und schätzen den Anteil der kritischen Masse, die durch ESA Verträge geschaffen oder gehalten wird. Diese Ziffer wird mit den durchschnittlichen Personalkosten multipliziert

Wert der kritischen Masse = Personenanzahl x Anteil ESA Verträge x Personalkosten

(z.B. Geben die Manager 30 Personen als kritische Masse an, von denen 20% durch ESA Verträge gehalten oder geschaffen werden. Die Kosten für einen durchschnittlichen Ingenieur beträgt im Jahr öS 840.000. Der Wert der kritischen Masse ergibt sich dann als $30 \times 0,2 \times 840.000 = 7,56 \text{ Mio öS}$)

7. Aggregation der 4 verschiedenen Nutzenkategorien zu einem Gesamtnutzen, Ermittlung eines ökonomischen Spin-off Koeffizienten (**Multiplikator**)

Der Gesamtnutzen wird durch schlichte Addition der 4 Nutzenkategorien ermittelt. Als Kosten werden dem die an die Unternehmen vergebenen Aufträge gegenübergestellt. Der Quotient aus Gesamtnutzen und Gesamtkosten ergibt den "Multiplikator", der anzeigt, wie viel indirekt ökonomischer Nutzen mindestens aus einer Geldeinheit ESA-Vertrag resultiert.

(z.B. Ein Multiplikator von 3 gibt an, dass aus jeder Geldeinheit, die von der ESA an die Industrie gezahlt wird, ein indirekter ökonomischer Nutzen von mindestens 3 Geldeinheiten resultiert.)

4.4.3 Datenerfordernis und Aufwand

Datenerfordernis und **Aufwand der Primärerhebung** dieser Methode ist enorm: Bei der im Jahr 1980 durchgeführten Evaluation wurden 128 Firmen interviewt, im Jahr 1988 waren es 67. Es ist davon auszugehen, dass pro Firma mindestens 5 Interviews geführt werden müssen, da ja die Effekte auf die einzelnen Abteilungen der Firmen ermittelt werden. Bei

jedem der Interviews müssen alle Effekte abgefragt werden, also mindestens 2 Mal 11 Koeffizienten, hinzu kommen qualitative Fragen.

Die aus der Primärerhebung gewonnenen Daten werden wahrscheinlich (es ist dies nicht dokumentiert) mit den Inhalten der ESA Verträge auf Kompatibilität überprüft. Es ist anzunehmen, dass es mehrerer Runden an Nachfragen benötigt, bis ein vollständiges und kompatibles Set an Koeffizienten gewonnen werden kann.

Sekundärdaten benötigt die angewandte Methode überhaupt keine.

4.5 Ergebnisse der Evaluation

Die Studien berechnen enorm hohe Multiplikatoren in der Größenordnung von 3 (vgl. nachfolgende Tabelle).

Tabelle 12

	ESA 1980	ESA 1988	Canada 1989
Period covered	64-82	77-91	79-93
Number of firms in the panel	128	67	10
total indirect effects	7551 (MAU 86)	12680 (MAU 86)	256 (MAU 89)
among ESA contractors	6023 (MAU 86)	9214 (MAU 86)	189 (MAU 86)
Ratio effects/contracts	2,9	3,2	3,5
Indirect effects outside space sector	50%	21,1%	24,4%
Indirect effects on exports	28,2% (out of ESA Member States)	12,8%	66,4%
Nature of the effects (% of contractors' effects)			
Technological	25	32	40
Commercial	27	8	18
Org.&Methods	19	6	18
Work factor	29	54	24

Quelle: Cohendet 1997

Auffallend ist auch die sehr unterschiedliche Aufteilung zwischen den Nutzenkategorien der Untersuchungen im Jahr 1980 und 1988. Während die erste Untersuchung eine relativ ausgeglichene Verteilung der Nutzen auf die einzelnen Kategorien zeigt, ist bei der Untersuchung 1988 eine Konzentration auf die Kategorien "Technologischer Nutzen" und "Nutzen für den Faktor Arbeit" festzustellen.

In einer Aufgliederung nach Ländern ergeben sich auch große Verschiebungen: Für Schweden beispielsweise errechnet sich im Jahr 1980 ein Multiplikator von 0,8, für 1988 einer von 9,8. Einen sehr hohen Multiplikator, nämlich 5,7 weisen auch die drei Länder Irland, Norwegen und Österreich auf, die 1988 erstmals mit untersucht wurden.

Disaggregierte Zahlen für Österreich finden sich in Leo (1991): Demnach wurden in die Erhebungen ORS, Schrack, Elin und die Universität Wien einbezogen. Interessanterweise hat in Österreich - im Gegensatz zu den meisten anderen Ländern - der Nutzen in Bezug auf den Faktor Arbeit eine nur untergeordnete Bedeutung. Der Multiplikator beträgt insgesamt für Österreich lediglich 1,14, das ist der schlechteste Wert von allen untersuchten Ländern für 1980.

Multiplikatoren dieser Größenordnung finden sich auch in anderen Studien (beispielsweise der NASA Beschäftigungsmultiplikator von 2,8, der norwegische Multiplikator von 2,8), allerdings gibt es eine Reihe methodischer Einwände gegenüber diesen Arbeiten und darüber hinaus steht ein solch hoher Multiplikator im Widerspruch zu empirischen Evidenzen hinsichtlich der kommerziellen Umsätze im Raumfahrtbereich (vgl. hierzu genauer Tichy und Posch 1999)

Auch die zu Beginn zitierten Fallstudien lassen Spin-offs in eher bescheidenerem Ausmaß vermuten: Die Weltraumwirtschaft ist mit den übrigen Bereichen der Wirtschaft nur sehr mangelhaft verflochten - Technologietransfer zwischen Raumfahrtunternehmen zu Nicht-Raumfahrtunternehmen ist eher selten. Hinzu kommt, dass aufgrund der hohen Sicherheitsanforderungen in der Raumfahrt nicht immer die neusten Technologien eingesetzt werden (vgl. z.B. Scientific Consulting 1989).

4.6 Stärken und Schwächen der Methode

4.6.1 Stärken der Methode

Die Methode hat die folgenden Stärken aufzuweisen:

- Durch die Erfassung und **Befragung unterschiedlicher Funktionsbereiche und Funktionsträger** innerhalb der betroffenen Firmen wird Einblick gewährt in die umfassenden firmeninternen und funktionsübergreifenden Auswirkungen technologischer Änderungen. So werden Effekte nicht nur im engen technologischen Sinn erhoben, sondern auch die Wirkungen auf andere Funktionsbereiche der Firma (Organisation, Qualifikation etc.) . Auch umgeht man mit der – zugegebenermaßen aufwendigen – Interviewtechnik mit den jeweiligen Abteilungsleitern in den Firmen das Problem, dass Fragebögen in der Regel von Ferialpraktikanten ausgefüllt werden bzw. dass man völlig andere Antworten erhält, je nach dem ob den Bogen nun zufällig eine Person aus der F&E Abteilung oder eine Person aus der Marketing Abteilung ausfüllt.
- Die hier vorgestellte **Mikroerhebung** scheint darüber hinaus die **einzige Möglichkeit**, solide Daten in Bezug auf messbare ökonomische Effekte zu sein. Die Ergebnisse haben natürlich den Nachteil **subjektiv gefärbt** von genau denjenigen Personen und Einrichtungen zu sein, die letztlich das größte Eigeninteresse an Erhaltung oder Ausbau

der Weltraumprogramme haben. Doch angesichts dessen, dass keine anderen Personen über dieses spezifische Know-how verfügen, scheint kein Weg daran vorbei zu gehen, diese Subjektivität in Kauf zu nehmen. B.E.T.A. versucht, dieses Problem zu mindern, indem sie „**minimale**“ **indirekte Effekte** abbilden. Dies heißt, aus der Bandbreite der Quantifizierungen, welche die Manager angeben, geht jeweils nur der niedrigste Wert in die Kalkulation ein.

Dem stehen allerdings erhebliche Schwächen gegenüber:

4.6.2 Schwächen bei der konkreten Anwendung der Methode

- Die Ermittlung des **Nutzens für den Faktor Arbeit ist mit methodischen Mängeln verbunden**. Da diese Kategorie einen großen Teil des Gesamtnutzens darstellt (in der Untersuchung 1988 sind es 41%), führen diese Mängel zu einer starken Überschätzung des Nutzenkoeffizientens (vgl. hierzu ausführlich Leo 1991, S. 48ff):
- Der Nutzen für den Faktor Arbeit soweit er den Beitrag der ESA zum Erhalt einer „kritischen Masse“ betrifft, wird ermittelt, indem die geschätzte Zahl an MitarbeiterInnen, welche die kritische Masse darstellen mit dem durchschnittlichen Personalaufwand multipliziert wird. Dies stellt eine Bestandsgröße dar, die nicht einfach zu den übrigen Nutzenkategorien, die alle als Stromgrößen ausgewiesen sind, hinzugerechnet werden kann.
- Der Beitrag der ESA zum Erhalt der kritischen Masse stellt eher einen direkten Nutzen der ESA, keinen indirekten dar. Er berechnet einen Inputfaktor dessen Wirkungen auf Outputgrößen bereits in den anderen Nutzenkategorien erfasst ist. Die Einführung einer eigenen Kategorie führt zu Doppelzählungen.
- Die Variablen der anderen Nutzenkategorien messen eher die Wirkung der Entwicklung nationaler und internationaler Raumfahrtmärkte auf **einzelne Unternehmen als den volkswirtschaftlichen Nutzen** der ESA. Für staatliche Allokationsentscheidungen ist diese Methode daher wenig geeignet.
- Als Maßstab für die **Kosten** verwendet B.E.T.A. die an die untersuchten Unternehmen vergebenen Aufträge. Damit bleibt ein großer Teil der Kosten (Kosten der Gesamtorganisation und Verwaltung, etc.) unberücksichtigt.
- Die Kosten und Nutzen fallen zu jeweils unterschiedlichen Zeitpunkten an. Um hier vergleichen zu können, müssten alle Größen auf **Gegenwartswerte abdiskontiert** werden.

4.6.3 Prinzipielle Schwächen bzw. Probleme des Ansatzes

- Es gibt keinen Vergleich mit anderen Programmen bzw. **alternative Ausgaben der öffentlichen Mittel** werden nicht untersucht. Dies ist zugegebenermaßen auch sehr schwer und aufwendig, würde sich aber bei großen, missionsorientierten Programmen durchaus anbieten, wenn die Ergebnisse tatsächlich Grundlage für staatliche Allokationsentscheidungen sein sollen. Doch auch in kleinerem Ausmaß sind Untersuchungen dieser Art wenig als Basis für strategische Entscheidungen geeignet: Es wird nicht differenziert, für welchen Zweck die Mittel eingesetzt wurden, und es können

daher auch keine Aussagen gemacht werden, in welcher Weise Mittel auf unterschiedlich nutzenstiftende Bereiche der Raumfahrt alloziert werden sollen – auf den ersten Blick scheinen hier sehr heterogene Wirkungsmuster vorzuliegen (Telekommunikation hat sicherlich andere ökonomische Wirkungen als die Mondlandung).

- Die Kalkulation des ökonomischen Nutzens und überhaupt die Definition von Spin-offs vermitteln den Eindruck, **die Raumfahrt sei der einzige Innovationsimpuls** für die involvierte Firma – und möglicherweise auch für den Rest der Ökonomie. Abgesehen davon, dass aufgrund hoher Sicherheitsbestimmungen in der Raumfahrt nur in den seltensten Fällen Spitzentechnologien entwickelt werden, spiegelt dies ein falsches **lineares Innovationsmodell** wider. Einer der Autoren der B.E.T.A. Studien kommentiert selbst die damit verbundenen Fehlschlüsse:

In fact, experience show that the space industry is where technologies developed elsewhere are assembled and improved. This is why we should perhaps consider space-sector spinoffs in terms of their complementarity and interactivity with other sectors rather than in terms of impacts justifying space programmes. (Cohendet 1997, S.218)

- Der Untersuchung liegt nicht wirklich ein **volkswirtschaftliches Nutzenkonzept** zu Grunde, erhoben wird eigentlich nur der Nutzen der direkten Vertragspartner sowie der ihrer direkten Lieferanten. Damit werden sowohl alle intangiblen Kosten- und Nutzenkategorien aber auch alle Wirkungen, die an anderen Stellen der Volkswirtschaft auftreten, vernachlässigt. Im Kontext technologischer Impulse handelt es sich hierbei im wesentlichen um die folgenden zwei Effekte:

Erstens jene Effekte die aufgrund von Input-Output Beziehungen der Firmen weniger beim Vertragspartner und seinen direkten Lieferanten selbst auftreten, sondern bei anderen vor- und nachgelagerten Elementen der Wertschöpfungskette und zwar sowohl im gleichen Sektor als auch sektorübergreifend. Diese Effekte, welche im Fall von „embodied technology“ auftreten, sind im Prinzip in Preisen und Mengen fassbar. Zweitens die Wirkungen des generellen Wissenszuwachses, welches sich die ESA Vertragspartner nicht exklusiv aneignen können und welches in die gesamte Volkswirtschaft diffundiert. Diese Effekte „disembodied technology“ schlagen sich nur unvollständig in Preisen und Mengen wider und sind daher schwieriger fassbar.

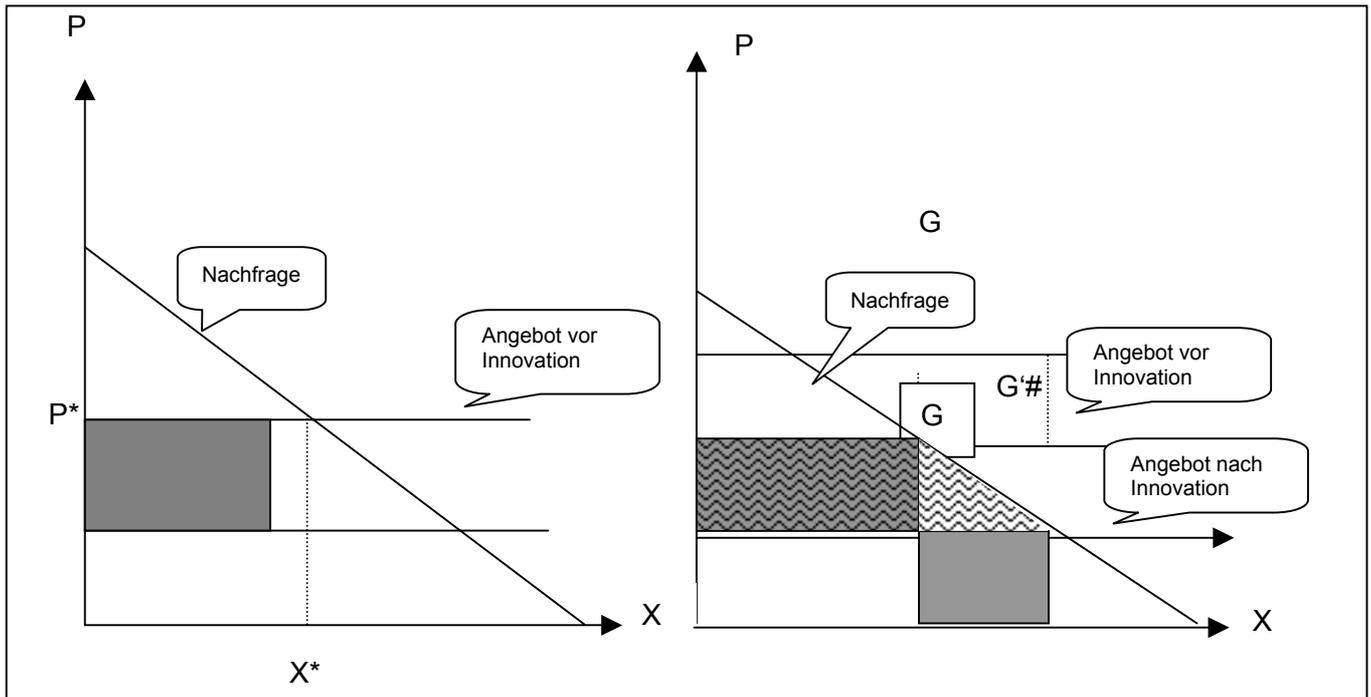
Diese Wirkungen können für unterschiedliche Sektoren und unterschiedliche Nachfragekomponenten positive wie auch negative Auswirkungen haben.

Beispielsweise können durch die ESA Aufträge Verdrängungseffekte auftreten, die bei anderen Firmen negative Auswirkungen auf Umsatz und Beschäftigte haben.

Ein zentraler Punkt in diesem Zusammenhang ist, dass die Ergebnisse der Mikroerhebungen (Befragung der einzelnen Unternehmen hinsichtlich der Wirkungen auf Umsatz bzw. auf Nettoproduktionswert) ja auch Ausdruck bestimmter Konstellationen des Unternehmensumfeldes sind. Je nachdem, wie vor- und nachgelagerte Bereiche, wie konkurrierende Unternehmen von der technologischen Änderung betroffen sind und auch je nachdem, in welcher Weise das Wissen im betreffenden Sektor – und darüber hinaus in die gesamte Ökonomie – diffundiert, sehen auch **die Effekte auf die untersuchten Unternehmen ganz unterschiedlich** aus. Das heißt die mit der B.E.T.A. Methode erhobenen Effekte sind hoch abhängig davon, was sich im Unternehmensumfeld tut. Ein

einfaches Beispiel: **Die Effekte affizieren nicht unbedingt den Umsatz des Unternehmens.** Wenn das Wissen um eine neue Technologie (nehmen wir im einfachsten Fall an, eine verbesserte Produktionstechnik) sofort und kostenlos von allen Anbietern im betreffenden Markt angewendet werden kann, so sinkt bei entsprechend elastischer Nachfrage der Preis. Diese Preissenkung kann dazu führen, dass trotz einer mengenmäßigen Outputsteigerung der Umsatz des Unternehmens sinkt.

Abbildung 14



Die Innovation wirkt hier so, dass sich die Grenzkostenkurve (= Durchschnittskostenkurve) des innovierenden Unternehmens ausgehend von der alten Gleichgewichtssituation P^*X^* nach unten verschiebt.

Der erste Teil von der Abbildung zeigt die Situation bei der Innovator in der Lage ist, das technologische Wissen – wie auch immer – **geheim zu halten**. Alle anderen Unternehmen produzieren weiterhin auf der alten Marktgrenzkostenkurve (= Angebotskurve), der Innovator erwirtschaftet einen Gewinn (Produzentenrente) in Höhe der markierten Fläche, an Preis und Umsatz ändert sich nichts, nach der B.E.T.A. Methode erhoben, wird die Firma eine entsprechende Kostensenkung als ökonomischen Nutzen angeben. Es ist dies eine kurzfristige Situation, bei der davon ausgegangen wird, dass der Innovator seine Kapazitäten nicht sofort ausweiten kann, längerfristig wird er – wenn die Geheimhaltung weiter möglich ist – den gesamten Markt als Monopolist beherrschen und der Preis wird von ihm in Abhängigkeit davon, wie schwer der Markteintritt für neue Firmen ist, festgelegt werden.

Der zweite Teil der Abbildung zeigt die Situation bei der **alle Firmen im Markt sofort auch über das technologische Wissen des Innovators verfügen**: In dem Fall setzt sich die neue, niedrigere Grenzkostenkurve als neue Marktangebotskurve durch, der Preis sinkt und

für die Produzenten fällt keine Rente an. Ob die neue Situation mit Umsatzsteigerungen für die Firmen einhergeht oder nicht, hängt von der Elastizität der Nachfrage ab - in dem hier dargestellten Fall, zeigt ein Vergleich der beiden grau unterlegten Flächen, dass der Umsatz insgesamt sinkt. Der nach der B.E.T.A. Methode erhobene Nutzen wäre in diesem Fall wahrscheinlich recht klein und vor allem deutlich geringer als im ersten Fall: Die befragte Firma wird zwar Kostensenkungen zu verzeichnen haben, nicht aber positive Gewinne und auch keine Umsatzsteigerung. Aus der Perspektive der Maximierung eines gesamtwirtschaftlichen Nutzens hingegen ist diese Situation – ceteris paribus – überlegen: Der Zuwachs an Konsumentenrente (Wellenmuster) übersteigt deutlich die Produzentenrente der Geheimhaltungssituation.

4.7 Weiterentwicklungsmöglichkeiten und Anwendungsgebiete

Gemäß der Analyse der Schwächen der Methode in Abschnitt 4.6 sind folgende Weiterentwicklungsmöglichkeiten denkbar:

- Korrektur um die in Abschnitt 4.6.2 genannten methodischen Unzulänglichkeiten
- Erweiterung des Ansatzes um eine volkswirtschaftliche Dimension, bei der Effekte in allen Sektoren und bei allen Komponenten der Endnachfrage lokalisiert werden. Dafür bietet es sich an, die B.E.T.A. Methode mit anderen Methoden der Ermittlung von Spillovers zu kombinieren wie multisektorale ökonomische Schätzungen, Input-Output-Analysen und Fallstudien.
- Vergleich der ermittelten Kosten und Nutzen mit alternativen Verwendungsmöglichkeiten öffentlicher Mittel

Selbst wenn der Ansatz in dieser Weise weiterentwickelt wird – Spin-off und Spillover Phänomene sind derart komplex, dass eine adäquate Abbildung derselben noch einer Menge weiterer Forschungsarbeiten bedarf. So auch Cohendet selbst:

„Many research projects need to be carried out in order to develop an analytical framework able to take into account the variety of spinoff phenomena and the complexity of the channels through which they have an impact on economic activity and, on this basis, design methodologies that can provide accurate measurements.“
(Cohendet, 1997, S. 218)

4.7.1 Einsatzmöglichkeiten der Methode

Aufgrund des enormen Aufwands, der mit dieser Evaluationsmethode verbunden ist, lohnt sich der Einsatz nur bei **sehr großen Programmen**. Dort wiederum ist die Methode vor allem für **missionsorientierte** Programme die eine gesellschaftspolitische Zielsetzung verfolgen, interessant. Im Kontext solcher Programme macht die hier verfolgte Fragestellung des „**Sekundärnutzens**“ Sinn, allerdings sind – wie oben gezeigt – die Grenzen zwischen Primär- und Sekundärnutzen häufig nicht ganz trennscharf zu ziehen. Insgesamt sind ähnliche Methoden sehr verbreitet bei der Ermittlung des ökonomischen Nutzens von Rüstungs- und Weltraumausgaben – sowohl auf nationaler als auch auf internationaler Ebene (vgl. Leo 1991 sowie Tichy und Posch 1999). Neben dem besonderen Interesse an Spin-offs in diesem Bereich bietet sich die Methode hier besonders an, da es

vergleichsweise wenige Firmen und Forschungseinrichtungen gibt, die an den Programmen teilnehmen.

4.7.2 Einsatzmöglichkeiten bei österreichischen Programmen

Daraus folgt eine nur **eingeschränkte Anwendungsmöglichkeit** der Methode auf österreichische Programme: Diese sind in der Regel zu klein und wenig fokussiert. Bei den neuen Schwerpunktprogrammen (Impulsprogramm „Nachhaltig Wirtschaften“ oder „Move“) kann die Frage nach dem Sekundärnutzen ein möglicherweise interessantes Tool einer Evaluation sein, doch eine umfassende Erfassung dieser Effekte nach der vorgestellten B.E.T.A. Methode wird den Aufwand nicht wert sein.

Im Zusammenhang mit **Raumfahrtaktivitäten** gibt es eine österreichische Studie, welche die B.E.T.A. Methode eins zu eins übernimmt: Lothaller (1997) berechnet „indirekte wirtschaftliche und sozioökonomische Auswirkungen des ESA-Engagements“ (S.17) exakt wie in oben vorgestellter Weise. Tichy und Posch (1999) kritisieren diese unhinterfragte Übernahme und äußern sich auf Grundlage eines intensiven Literaturstudiums sowie eigenen Befragungen der Firmen und Forschungseinrichtungen skeptisch hoher Spin-off Erwartungen – sowohl international als auch insbesondere für Österreich. Allerdings schlagen sie eine Weltraum-Strategie für Österreich vor, die sich um genau dies konzentriert: „letztes Ziel bleibt immer die Erzielung von Synergien und Spin-offs“ (S.50): Die Förderwerber müssen dem gemäß ex-ante ihre Zielvorgaben in Bezug auf Synergien und/oder Spin-offs bekannt geben, die Zielerreichung wird regelmäßig überprüft und eine Nicht-Erreichung hat deutliche Konsequenzen (Rückzahlung oder Ausschluss aus dem Förderprogramm). Wenn diese vorgeschlagene Strategie in Österreich künftig verfolgt wird, so muss auf Seiten der Evaluatoren Know-how zur problemadäquaten Ermittlung von Spin-offs aufgebaut werden. Die hier vorgestellte Methode mit allen ihren Mängeln und Verbesserungsmöglichkeiten kann hier als Ausgangspunkt dienen.

4.8 Literatur

- Amundsen, E.F., Skovholt, J. (1994), Evaluation of the industrial spin-off effects of Norwegian participation in ESA collaboration: 1985-1993, Oslo: Norsk Romcenter (Norwegian Space Center).
- B.E.T.A. (1988), Study of the Economic Effects on European Space Expenditure, Results (Vol.1) and Report on Investigation Theory and Methodology (Vol. 2), Reports for the European Space Agency, ESA Contract No. 7062/87/F/RD/(SC).
- Chapman, R.L. et al. (1989), An Exploration of Benefits from NASA "Spin-off", Chapman Research Group Inc., Little Colorado, USA.
- Cohendet, P. (1997), Evaluating the Industrial Indirect Effects of Technology Programmes: The Case of the European Space Agency (ESA) Programmes, in: OECD Proceedings, Policy Evaluation in Innovation and Technology, Towards Best Practices, Paris.
- Damianisch, W., (1984), Beziehungen Österreich zur ESA, Auswirkungen auf die Wirtschaft, Bundesministerium für Wissenschaft und Forschung, Wien.

- Deutsches Institut für Wirtschaftsforschung, Wochenbericht 43/97, Deutsche und europäische Raumfahrtindustrie - fit für die Zukunft? Berlin. Dykewicz, P., 1998, The Global Satellite Market-Emerging Trends and Future Prospects, Paris.
- ESA (1984), Europe Two Decades in Space, ESA SP-1060, Noorwijk.
- ESA (1987), Europe Long Term Space Plan 1987-2000, Paris.
- Georghiou, L. and Roessner, D. (2000), Evaluating technology programs: tools and methods, in: Research Policy 29 (2000), S. 657-678.
- Hertzfeld, H. (1992), Measuring the returns to space research and development, in: Hertzfeld, H. and Greenberg, J. (Eds.), Space Economics, American Institute of Astronautics, Washington.
- Kaufmann, A. (1997), Industriepolitik der ESA in EU-Relation und technologiepolitische Auswirkungen des ESA-Engagements auf die österreichische Volkswirtschaft, Band 2, Forschungsauftrag des Bundesministeriums für Wissenschaft und Verkehr, Wien: Industrewissenschaftliches Institut.
- Leo, H., (1991), Die österreichische Weltraumindustrie. Ausmaß und Perspektiven, Studie des Österreichischen Instituts für Wirtschaftsforschung im Auftrag des Bundesministeriums für Wissenschaft und Forschung, Wien: Wifo.
- Lothaller, W. (1997), Industriepolitik der ESA in EU-Relation und technologiepolitische Auswirkungen des ESA-Engagements auf die österreichische Volkswirtschaft, Band 1, Forschungsauftrag des Bundesministeriums für Wissenschaft und Verkehr, Wien: Austrian Space Agency.
- Mathematica Inc. (1977), A Cost Benefit Analysis of Selected Technology Utilisation Office Programs, NASA, Washington.
- Schnee, J.E. (1977), Space programme impacts revisited, in: California Management Review 20(1).
- Scientific Consulting (1989), Spin-offs der Raumfahrt, ihre Auswirkungen auf Firmenstrategie und Märkte, Köln.
- Teubal, M., Steinmueller, E. (1982), Government Policy, Innovation and Economic Growth, in: Research Policy 1982 (11).
- Welles, J.G., Waterman, R.H, (1964), Space Technology: Pay-off From Spin-off, in: Harvard Business Review.

5 EVALUIERUNG VON NIST-LAB PROGRAMMEN

Anton Geyer

Austrian Research Centers Seibersdorf

5.1 Organisation und Tätigkeit des NIST-National Institute of Standards and Technology

Das NIST-National Institute of Standards and Technology ist eine Einrichtung innerhalb des US-amerikanischen Wirtschaftsministeriums (Department of Commerce). Die Aufgabe des NIST ist es, das Wachstum der amerikanischen Wirtschaft durch die Entwicklung und Anwendung von neuen Technologien zu stimulieren, sowie durch die Entwicklung von Mess-, Test- und Prüfverfahren zu fördern. Das NIST wurde vom amerikanischen Kongress im Jahr 1901 als National Bureau of Standards gegründet und arbeitet seit 1988 unter dem Namen NIST-National Institute of Standards and Technology.

5.1.1 Organisatorische Gliederung

Das NIST ist in vier Programmbereiche gegliedert. Neben den Measurement and Standards Laboratories (MSL) betreibt das NIST die Programmbereiche Advanced Technology Program (ATP), Manufacturing Extension Partnership (MEP) und das National Quality Program (Baldrige National Quality Award). Die Aufgabe der MSL ist es, die technischen und organisatorischen Voraussetzungen für eine hochwertige Mess-, Test- und Prüfinfrastruktur in den Vereinigten Staaten sicherzustellen. Dazu stellen die NIST-Laboratorien der Industrie Referenzdaten sowie Mess-, Test- und Prüf-Kapazitäten zur Verfügung.

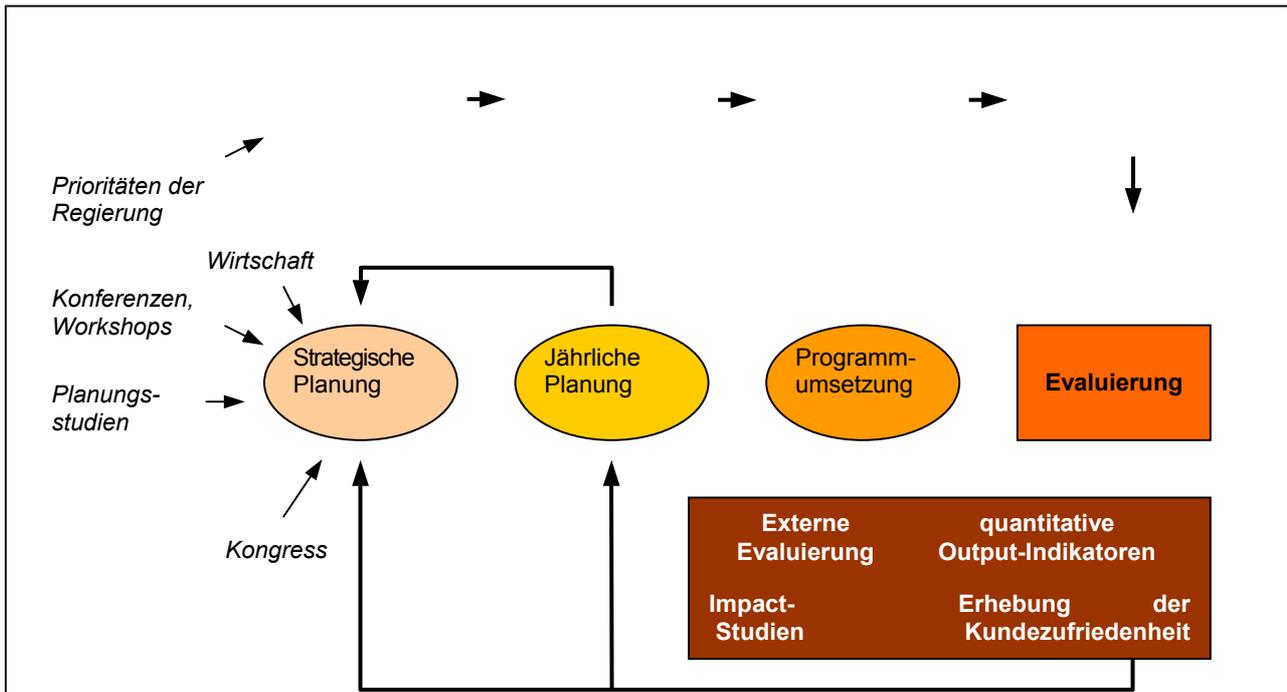
5.1.2 Programmplanungs- und Evaluierungssystem

Der Programmplanungs- und Evaluierungsprozess am NIST umfasst einen Kreislauf mit vier Kernelemente:

- **Strategische Planung:** Hier fließen unter anderem die Vorstellungen und Vorgaben des Kongresses ein. Weiters wird die strategische Planung von den politischen Prioritäten der amerikanischen Regierung beeinflusst. Innerhalb des NIST tragen strategische Planungsstudien, Planungskonferenzen und Workshops dazu bei, die langfristige Strategie festzulegen und kontinuierlich weiterzuentwickeln. Eine enge Kooperation mit der amerikanischen Industrie unterstützt diesen Prozess. Über eine Feedback-Schleife fließen die Ergebnisse von Evaluierungen und die Erfahrungen aus dem jährlichen Programmplanungsverfahren in die Strategieentwicklung mit ein.
- **Jährliche Programmplanung:** Das jährliche Arbeitsprogramm wird auf Grundlage des strategischen Planungsprozesses erarbeitet. Die Ergebnisse von Evaluierungen der vorangegangenen Arbeitsprogramme werden für die laufende Programmentwicklung genutzt.
- **Programmumsetzung und -administration:** Die operative Umsetzung der Programme wird durch ein konsequentes Programmmonitoring begleitet.

- **Evaluierung:** Die Leistungsevaluierung der Aktivitäten am NIST umfasst die vier Elemente:
 - (a) externe ExpertInnenbegutachtung (Peer review)
 - (b) quantitative Output-Indikatoren
 - (c) ökonomische Impact-Studien
 - (d) Kundenzufriedenheitserhebungen

Abbildung 15 Programmplanungs- und Evaluierungsprozess NIST



Quelle: Planungs- und Evaluierungsprozess des NIST (Doremus, 1998)

Die *externe Evaluierung* des NIST erfolgt durch das *Board on Assessment* des NRC-*National Research Council*. Das Board wurde 1959 gegründet und setzt sich aus rund 150 VertreterInnen der Industrie, der Wissenschaft und der öffentlichen Verwaltung zusammen. Das NRC-Board ist für die Nominierung der Arbeitsgruppenmitglieder in den Evaluierungsteams verantwortlich. Das NIST berät das Board bei der Auswahl geeigneter Personen. Die Letztverantwortung für die Zusammensetzung der Arbeitsgruppen (z.B. hinsichtlich der Vermeidung von Interessenkonflikten) und die Qualität der Evaluierungen liegt jedoch beim NRC-Board.

Das NRC-Board legt jährlich einen Bericht über die Qualität und Relevanz der sieben großen NIST-Laboratory Programme vor (NAS, 2000). Dazu ist für jedes Laboratorium eine eigene Arbeitsgruppe eingerichtet (sowie zusätzlich in bestimmten Bereichen Unterarbeitsgruppen), die eine Peer-review Evaluierung der einzelnen Laboratorien durchführt.¹² Das NRC-Board informiert die NIST-Leitung, die Programmmanager und das „Visiting Committee on Advanced

¹²Die sieben großen NIST-Laboratorien sind: (1) Electronics and Electrical Engineering Laboratory, (2) Manufacturing Engineering Laboratory, (3) Chemical Science and Technology Laboratory, (4) Physics Laboratory, (5) Materials Science and Engineering Laboratory, (6) Building and Fire Research Laboratory, (7) Information Technology Laboratory.

Technology“ (NIST-Beirat) über die Ergebnisse der Evaluierungen. Der Bericht wird jeweils im Oktober des dem Berichtszeitraums folgenden Jahres veröffentlicht.

Der Evaluierungsbericht des NRC-Boards konzentriert sich vor allem auf die technische Qualität der NIST-Tätigkeit. Vier Punkte werden besonders berücksichtigt:

- Die erbrachten technischen Leistungen der Laboratoriumsprogramme im Vergleich zum state-of-the-art im jeweiligen Fachgebiet
- Das Ausmaß, in dem die Laboratoriumsprogramme mit der Mission der Laboratorien übereinstimmen
- Die Effektivität der Programmumsetzung und der Verbreitung der Ergebnisse
- Die Angemessenheit der vorhandenen Einrichtungen, das heißt der technischen und personellen Ausstattung der Laboratorien, sofern sie die Qualität der technischen Ergebnisse am NIST beeinflussen.

Die NIST-Laboratorien erarbeiten eine Stellungnahme zu den Evaluierungsberichten, die jeweils im darauf folgenden Jahr vom NRC-Board im Evaluierungsbericht kommentiert wird.

Neben dem Peer-review werden sind Erhebungen der Kundenzufriedenheit und Output-Indikatoren zwei weitere Elemente des NIST-Evaluierungssystems für Laboratorien. Als *Output-Indikatoren* für NIST-Laboratorien, die ebenfalls ein Evaluierungselement im NIST-System sind, werden unter anderem die folgenden quantitativen Maße herangezogen:

- Anzahl der verfügbaren NIST-Referenzdatenbanken
- Anzahl der verfügbaren NIST-Referenzmaterialien
- Anzahl der NVLAP (National Voluntary Laboratory Accreditation Program) Akkreditierungen
- Anzahl der gegenseitigen NVLAP Anerkennungsvereinbarungen
- Anzahl der eingereichten Patente und der vergebenen Lizenzen
- Anzahl der Mitgliedschaften und Leitungen in nationalen Normungsausschüssen
- Anzahl der Mitgliedschaften und Leitungen in internationalen Normungsausschüssen
- Anzahl der technischen Publikationen
- Anzahl der Teilnahmen an Konferenzen und Workshops
- Anzahl der Abfragen über den zentralen NIST-Server
- Anzahl der durchgeführten Kalibrierungen und Prüfungen
- Anzahl der verkauften Referenzmaterial-Einheiten
- Anzahl der vertriebenen Referenzdatenbanken

5.2 Ökonomische Impact-Studien für NIST-Forschungsprogramme

Impact-Studien stellen das vierte Element der regelmäßigen Evaluierungen von NIST-Laboratoriumsprogrammen dar. Das hier beschriebene quantitative Verfahren hat sich zu einer Standardvorgehensweise bei der Evaluierung von NIST-Laboratory Programmen

entwickelt (Link und Scott, 1998). Die Ergebnisse bereits durchgeführten Impact-Studien, sowie Hintergrundinformationen zur verwendeten Methodik können über das Internet abgerufen werden (<http://www.nist.gov/director/planning/strategicplanning.htm>). In der aktuellen Liste (April 2000) finden sich 19 abgeschlossene Evaluierungen sowie vier laufende Studien (<http://www.nist.gov/director/planning/studies.htm>).

5.2.1 Ziel der Impact-Studien

Das Ziel der Impact-Studien ist es, die Art und das Ausmaß der ökonomischen Auswirkungen von NIST-Programmen zu bestimmen. Die Herausforderung ist dabei, quantifizierbare und überprüfbare Kosten-Nutzen Indikatoren zu entwickeln, die robust und aussagekräftig die Auswirkungen von Forschungsprojekten über lange Zeitperioden erfassen.

5.2.2 NIST-Methodik zur Impact-Evaluierung

Die am NIST verwendete Methodik zur Ermittlung der ökonomischen Effekte von Laboratorienprogrammen geht von konventionellen Methoden der Wirtschafts- und Finanzanalyse aus und adaptiert diese für die Beurteilung von NIST-finanzierten Forschungs- und Entwicklungsprogrammen. Dabei wird untersucht, welchen Einfluss die Arbeit des NIST, bzw. die vom NIST zur Verfügung gestellten Infratechnologien, auf die Produktivität von Unternehmen, auf die Effizienz industrieller Forschungs- und Entwicklungsarbeit, auf Transaktionskosten und auf damit zusammenhängende Faktoren an den verschiedenen Stellen der Produktionskette haben.

Zusätzlich führt das NIST auch eine qualitative Abschätzung der Programmimpacts durch. Dabei wird unter anderem erhoben, auf welche Art und Weise die NIST-Laboratoriumsprogramme Investitionsentscheidungen der Industrie (z.B. hinsichtlich Forschungs- und Entwicklungsvorhaben, Produkt- und Qualitätsstrategien, Produktzyklen etc.) beeinflussen.

Die Durchführung von NIST-Impact Studien folgt einem vierstufigen Prozess (Link and Scott, 1998):

- Schritt 1: Ermittlung (Quantifizierung) der aufgewendeten Mittel in jenen NIST-Einrichtungen, die für die Entwicklung und Umsetzung des zu evaluierenden Programms verantwortlich waren.
- Schritt 2: Ermittlung der mit den aufgewendeten finanziellen Ressourcen erzielten Ergebnisse (outputs).
- Schritt 3: Identifizierung jener Gruppen, die von den Programme bzw. den Ergebnissen der Programme profitiert haben.
- Schritt 4: Ermittlung (Quantifizierung) des Nutzens (impacts) bei den profitierenden Gruppen.

Für die Bestimmung der ökonomischen Impacts beauftragt das NIST multidisziplinäre Teams von Ingenieuren, Industrieanalysten, Ökonomen, Finanzanalysten und UmfrageexpertInnen. Die Evaluierungsteams legen jene Aspekte fest, die in der Analyse untersucht werden sollen.

Auf Basis der zur Verfügung stehenden Daten – und unter Berücksichtigung der Datenqualität und -verlässlichkeit – werden Art und Ausmaß der ökonomischen Impacts abgeschätzt. Weiters legen die Projektteams den Beurteilungszeitraum für die Untersuchung fest.

Ökonometrisches Modelle für die Bestimmung des Nutzens von Forschungsaktivitäten

In der ökonomischen Fachliteratur wird zur Bestimmung des ökonomischen Nutzens von Forschung und Entwicklung häufig der Griliches/Mansfield-Ansatz herangezogen. Mit Hilfe dieses Modells können spillover-Effekte von Forschungs- und Entwicklungsleistungen, wie Produktverbesserungen, Prozessverbesserungen und die Schaffung neuer Märkte bestimmt werden. Das Griliches/Mansfield-Modell nimmt an, dass Forschung und Entwicklung zu einer Verringerung der Herstellungskosten eines Produktes führen. Zu jedem Zeitpunkt existiert für den Markt des Produktes eine Nachfrage- und eine Angebotskurve. Technologische Innovationen als Folge von Forschungs- und Entwicklungstätigkeit verringern die Herstellungskosten des Produktes und senken daher die Angebotskurve. Damit verringert sich der Preis des Produktes und führt so zu einer zusätzlichen Rente für die KonsumentInnen. Diese zusätzliche Konsumentenrente errechnet sich als die Differenz zwischen den beiden Preisen multipliziert mit der verkauften Stückzahl im jeweiligen Beobachtungszeitraum. Auch die (zusätzliche) Produzentenrente kann mit diesem Modell bestimmt werden. Hier wird die Differenz zwischen dem erzielten Preis und den Grenzkosten abzüglich der Fixkosten über die Gesamtzahl der verkauften Produkte aufsummiert. Das Griliches/Mansfield Modell lässt sich gut auf kostenreduzierende Prozessinnovationen anwenden, es muss jedoch bei Produktinnovationen (in diesen Fällen existiert keine Ausgangsangebots- und -nachfragekurve) entsprechend adaptiert werden.

In beiden Fällen können die aufsummierten zusätzlichen Renten der KonsumentInnen und ProduzentInnen als „sozialer Nutzen“ (*Return on Investment*) von Forschung und Entwicklung verstanden werden. Der Griliches/Mansfield-Ansatz gibt somit Antwort auf die Frage: Wie hoch ist der soziale Nutzen einer Investition in Forschung und Entwicklung, ausgedrückt als Summe des Nutzen (der zusätzlichen Renten) für KonsumentInnen und ProduzentInnen?

Das NIST interessiert allerdings üblicherweise eine andere Frage bei der Bestimmung der ökonomischen Impacts seiner Forschungstätigkeit: *Welchen Betrag hätte der private Sektor investieren müssen, (unter der Annahme, es wären keine NIST-Mittel eingesetzt worden) um die selben Ergebnisse zu erzielen?* Der Vorteil dieses Modells gegenüber dem Griliches/Mansfield-Ansatz bei der Evaluierung eines öffentlich finanzierten F&E-Programmes ist, dass damit eine Aussage gemacht werden kann, ob die öffentlich getätigte Investition tatsächlich effizienter gewesen ist, als eine hypothetisch angenommene private Investition.

Berücksichtigte Indikatoren für die Impact-Abschätzung

Für die Bestimmung der Programmimpacts werden sowohl output- als auch outcome-Indikatoren berücksichtigt. Zu den berücksichtigten *outputs* zählen beispielsweise die Anzahl der mit dem Programmen in Zusammenhang stehenden entwickelten

- Normen,
- Referenzmaterialien und Kalibrierungsverfahren,
- Referenzdaten und -datenbanken,
- Verbreitung von NIST-Normen,
- Mess- und Prüfverfahren,
- Qualitätssicherungsverfahren und
- Simulationsmodelle.

Die Kosten für die Erstellung dieser outputs fließen in die Berechnung des ökonomischen impacts ein. Unter den berücksichtigten *outcomes-Maßen* für die Programme finden sich Indikatoren wie

- Auswirkungen auf FuE-Entscheidungen der Industrie,
- Auswirkungen auf den Marktzugang,
- Verringerung von Industriezykluszeiten,
- Erhöhung der Produktivität von industriellen Prozessen und FuE-Aktivitäten,
- Verbesserte Qualität von Produkten und Dienstleistungen,
- Erhöhung der Zuverlässigkeit von Produkten und IKT-Systemen,
- Verringerung von Transaktionskosten,
- Verbesserung der Interoperabilität von IKT-gestützten Systemen

Diese Aspekte beschreiben den Nutzen der Programme.

5.2.3 Quantitative Beurteilungsmaße

Üblicherweise werden für NIST Impact-Studien drei verschiedene quantitative Maße für Impact-Studien verwendet:

Net Present Value (NPV, Nettogegenwartswert)

Dieser Wert ist ein absolutes Maß für den kumulierten Wert (bzw. für den erwarteten Wert) eines Forschungsprogramms bezogen auf einen bestimmten Zeitpunkt. Dieser Wert wird durch die zeitliche Aufwertung bzw. Abwertung der über den Beobachtungszeitraum entstehenden Kosten und Nutzen des Programms errechnet. Der NPV ist die Differenz zwischen Nettonutzen und Nettokosten des gesamten Programms. Wie bei vielen solchen Maßen ist das Ergebnis abhängig vom zugrundegelegten Zinssatz mit dem Kosten und Nutzen auf- bzw. abgewertet werden. Die Höhe des Zinssatzes wird in Abhängigkeit von Risiko, Inflation und der Zeitpräferenz von Kapital festgelegt.

$$NPV = [(B_0 - C_0) / (1 + i)^0] + \dots + [(B_n - C_n) / (1 + i)^n]$$

Social Rate of Return (SRR)

Der SRR entspricht per Definition dem Wert jenes Zinssatzes i , bei dem der Nettogegenwartswert des Nettonutzens eines Forschungsprogramms, vom Zeitpunkt des Programmbeginns bis zum festgelegten Ende des Beobachtungszeitraumes gleich null ist. Der Nettonutzen errechnet sich für jeden Zeitabschnitt aus der Differenz von Gesamtnutzen und Gesamtkosten.

$$NPV = 0 = [(B_0 - C_0) / (1 + i)^0] + \dots + [(B_n - C_n) / (1 + i)^n]$$

Der so berechnete Wert für SRR kann nun mit einem angenommenen Wert r verglichen werden, der die Opportunitätskosten (bzw. den dafür anzunehmenden Zinssatz) der finanzierenden öffentlichen Einrichtung widerspiegelt. Ist der Wert für SRR höher als der Zinssatz für die Opportunitätskosten r , war die Investition von einem ex-post Standpunkt betrachtet gesellschaftlich sinnvoll.

Da der Wert für i maßgeblich vom zeitlichen Auftreten der Kosten und Nutzen abhängt, sollte nach Link und Scott (1998) ein Vergleich verschiedener Programme auf Basis des SRR vermieden werden. Auch sollte eine einfache Analogie des SRR mit einem Zinssatz im herkömmlichen Sinn – beispielsweise für eine Bankeinlage – vermieden werden. Bei einer Bankeinlage tritt ein Zinseszinsseffekt auf, der in dieser Form für Forschungsinvestitionen nicht sinnvoll angenommen werden kann. Insbesondere treten die Kosten bei Forschungsprogrammen nicht ausschließlich am Beginn des Beobachtungszeitraumes auf, und der Nutzen wird nicht erst vollständig am Ende des Zeitraums lukriert.

Um diesem Problem der zeitlichen Veränderungen von Kosten und Nutzen zu begegnen wird häufig auch ein *Implied Rate of Return* (bzw. auch *Adjusted Internal Rate of Return*) berechnet. Dieser Berechnung liegt die Annahme zugrunde, dass alle Kosten am Beginn des Programms entstehen, der Nutzen aber erst am Ende des Beobachtungszeitraumes realisiert wird. Für die Berechnung werden dazu die Kosten auf den Zeitpunkt $t = 0$ bezogen, die Nutzen auf den Zeitpunkt $t = n$ (dies entspricht einem hypothetischen Reinvestieren des Nutzens zum Zeitpunkt $t = i$). Der Implied Rate of Return x bestimmt sich in der Folge als

$$x = -I + (TVB / PVC)^{1/n}$$

wobei TVB der Wert des Nutzens zum Ende des Beobachtungsraumes ist, und PVC der Nettogegenwartswert der Kosten:

$$TVB = \sum_{t=0 \text{ bis } t=n} B_t (1 + r)^{n-t}$$

$$PVC = \sum_{t=0 \text{ bis } t=n} C_t / (1 + r)^t$$

Kosten-Nutzen-Verhältnis (KNV)

Dieser Wert wird aus den Schätzungen der Nettokosten und des Nettonutzes eines Programms errechnet. Die entsprechenden Werte zum Zeitpunkt t_i werden dabei auf den Referenzzeitpunkt $t = 0$ bezogen.

$$KNV = B / C = [\sum_{t=0 \text{ bis } t=n} B_t / (1 + r)^t] / [\sum_{t=0 \text{ bis } t=n} C_t / (1 + r)^t]$$

Wenn der Zahlenwert des KNV größer als 1 ist, kann davon ausgegangen werden, dass das Programm erfolgreich war. Aus dem Kosten-Nutzen-Verhältnis kann auch einfach der Nettogegenwartswert des Programms ($NPV = B - C$) bestimmt werden.

Im folgenden Beispiel soll die Durchführung einer NIST-Impact-Evaluierung genauer illustriert werden.

5.3 Abschätzung der ökonomischen Impacts am Beispiel des NIST-Forschungsprogramms „Alternative Kältemittel“

5.3.1 Hintergrund des NIST Forschungsprogramms „Alternative Kältemittel“

Seit den frühen 80er Jahren wurden am NIST Chemical Science and Technology Laboratory (CSTL) Untersuchungen über die chemischen und physikalischen Eigenschaften von Ersatzstoffen für Fluorchlorkohlenwasserstoffe (FCKW) durchgeführt. Im Jahr 1998 wurde eine umfassende Evaluierung des Programms veröffentlicht, die entsprechend der oben beschriebenen NIST-Evaluierungsmethode den wirtschaftlichen des Programms für die amerikanische Wirtschaft näher untersuchte (Shedlick et al., 1998).

Bis in die 80er Jahre wurden FCKW in großen Mengen als Kältemittel in Klimaanlage, Kühlanlagen, Kühlschränken und Wärmepumpen etc. verwendet. Aufgrund der ozonschichtschädigenden Eigenschaften dieser Stoffe wurde im Rahmen eines internationalen Abkommens (Montreal-Protokoll 1987) ein akkordierter Ausstieg aus der Produktion und Verwendung von FCKW beschlossen. Als Konsequenz dieser völkerrechtlichen Vereinbarung mussten für die betroffenen industriellen Anwendungen technisch und wirtschaftlich geeignete, und weniger umweltschädigende Ersatzstoffe gefunden werden.

Um den im Montreal-Protokoll vorgesehenen Zeitplan für das Verbot von FCKW einhalten zu können war es notwendig, Untersuchungen über die Eigenschaften von potentiell geeigneten alternativen Kältemitteln durchzuführen. In den USA übernahm das NIST dabei die Rolle, grundlegende technische Anforderungen an alternative Kältemittel festzulegen und die physikalischen und chemischen Eigenschaften von geeigneten Substituten zu bestimmen. Dieses Forschungsprogramm sollte es der amerikanischen Industrie erlauben, zeitgerecht und mit vertretbarem Aufwand auf FCKW-freie Kältemittel umzusteigen.

Das Hauptprodukt des NIST-Forschungsprogrammes „Alternative Kältemittel“ ist das Computerprogramm REFPROP. Diese Software erlaubt es den Herstellern und Verwendern von Kältemitteln, die physikalisch-chemischen Eigenschaften von Kältemitteln und ihren Mischungen zu simulieren. Diese Technik erleichtert die Auswahl einer optimalen Alternative zu FCKW für die jeweilige Anwendung. Dadurch werden besonders energieeffiziente Lösungen für die kältetechnischen Anwendungen ermöglicht.

5.3.2 Modell für die Impact Abschätzung

Das Modell für die Kosten-Nutzen-Analyse des Forschungsprogramms „Alternative Kältemittel“ umfasst die Bestimmung der wesentlichen technischen Ergebnisse des Programms (outputs), die Entwicklung eines Ansatzes für die ökonomische Bewertung der Programmauswirkungen (outcomes) und damit zusammenhängend, die Festlegung des relevanten AdressatInnenkreises für die empirische Erhebung.

Ergebnisse des NIST-Programmes „Alternative Kältemittel“

Über die Laufzeit des Programms wurden fünf wesentliche technische Ergebnisse realisiert:

- Zahlreiche Publikationen über Messergebnisse und Rechenmodelle über die thermodynamischen Eigenschaften von FCKW-Ersatzstoffen, bzw. von alternativen Kältemittelmischungen wurden veröffentlicht. In den Jahren 1986 bis 1996 veröffentlichten MitarbeiterInnen des NIST zusätzlich über 140 wissenschaftliche Arbeiten mit Ergebnissen des alternativen Kältemittel-Programms.
- Eine umfassende Überarbeitung der Zahlenwerke und Graphiken im American Society of Heating, Refrigerating, and Air-Conditioning (ASHRAE) Handbook of Fundamentals wurde durchgeführt. Dieses Handbuch dient weltweit über 50.000 IngenieurInnen als Referenzwerk im Kältemittelbereich.
- NIST-MitarbeiterInnen beteiligten sich, und leiteten internationale Gremien, die zur Ziel hatten, den Stand der Technik im Sektor Kältemittel zusammenzufassen und eine internationale Norm über die Eigenschaften von Kältemitteln auszuarbeiten.
- Eine umfassende Datensammlung über die Eigenschaften alternativer Kältemittel wurde im CRC Press Verlag veröffentlicht (Referenzpublikationen für PhysikerInnen und ChemikerInnen).
- Das Softwareprogramm REFPROP, das die Eigenschaften von Mischungen mit bis zu fünf Komponenten aus einem Set von insgesamt 41 möglichen reinen Kältemitteln berechnet, wurde entwickelt. Dieses Programm ist heute ein Standardtool für die Kältemittelindustrie. Andere Forschungsinstitute übernahmen in der Folge die REFPROP-Datenbank als Basis für eigene Simulationsprogramme. Seit dem Erscheinen des Programms im Jahr 1990 wurde die Software über 500 mal verkauft.

Erfassung des ökonomischen Programmnutzens

Um die ökonomischen Effekte des Programms zu ermitteln, wurden die NIST-Forschungsausgaben dem Programmnutzen gegenübergestellt. Die Schätzung des Nutzens erfolgt auf Grundlage der Hypothese, dass die direkten ökonomischen Effekte des Programms mit jenen zusätzlichen Kosten abgeschätzt werden können, die die Industrie für die Erreichung der gleichen technischen Ergebnisse hätte aufwenden müssen, unter der Voraussetzung, das NIST-Programm wäre nicht durchgeführt worden. Anders ausgedrückt werden im Modell jene hypothetischen zusätzlichen Kosten für die Industrie bestimmt, die durch die Existenz des NIST Programms, bzw. durch die Verfügbarkeit von Programmerngebnissen und Softwareprodukten vermieden werden konnten.

Diese Hypothesemodell, als *counterfactual experiment* bezeichnet, wurde für die Evaluierung des Alternativen Kältemittel-Programms auch deshalb benutzt, weil in diesem Fall keine

geeignete Kontrollgruppe für die Abschätzung der ökonomischen Auswirkungen der Maßnahmen zur Verfügung stand. Die Industrie musste sich auf die vom NIST geschaffene Datenbasis verlassen, da der Industrie zum Zeitpunkt der Programmdurchführung keine alternativen Daten zur Verfügung hatte.

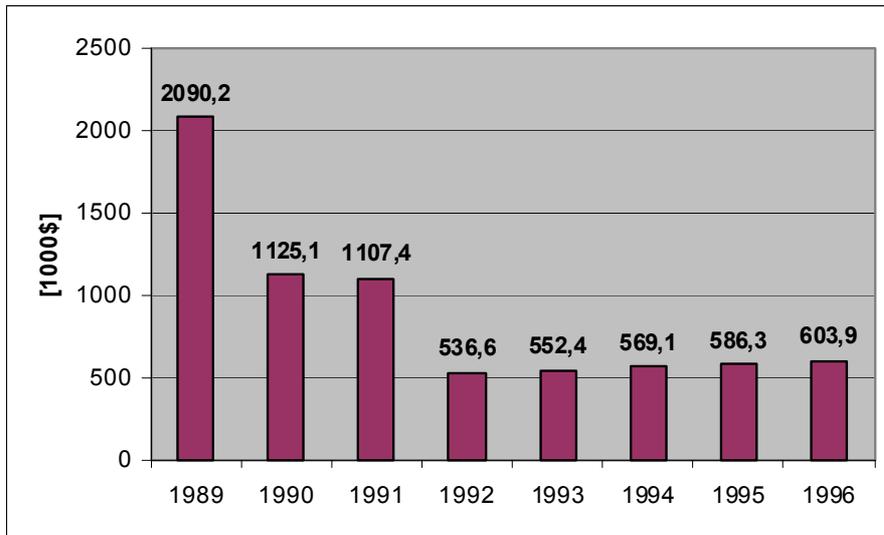
Für die Sammlung der Informationen über die ökonomischen Effekte wurden semi-strukturierte Telefoninterviews mit führenden Anwenderunternehmen der Software durchgeführt. Die erste Zielgruppe in der Untersuchung waren fünf große amerikanische Hersteller von FCKW-freien Kältemitteln. Diese Gruppe von Unternehmen besaßen zum Zeitpunkt der Untersuchung einen Marktanteil von ca. 90% (Umsatzanteil). Die sechs wichtigsten amerikanischen Verbraucher von Kältemitteln (ca. 70% Beschäftigtenanteil) bildeten die zweite Zielgruppe für die Untersuchung. Da bei dieser Stichprobe keine Repräsentativität hinsichtlich des Nutzens für die gesamte Industrie angenommen werden kann, wurde auf eine Hochrechnung des Programmnutzens auf die gesamte Kältemittelerzeugende und -verarbeitende Industrie verzichtet.

5.3.3 Durchführung der empirischen Erhebung

Ermittlung des Programmnutzens

Für die Hersteller und die Anwender von alternativen Kältemitteln wurden jeweils Interviewleitfäden entwickelt. In den Gesprächen wurde unter anderem nach der Einschätzung der Industrie hinsichtlich der Auswirkungen des Montrealprotokolls gefragt sowie die internen Forschungsaktivitäten und die Nutzung des NIST REFPROP-Programmes im jeweiligen Unternehmen besprochen. Jedes Unternehmen wurde bei den Interviews gefragt: *„Angenommen, das NIST hätte kein Kältemittel-Programm durchgeführt, wie hoch würden Sie Ihren zusätzlichen Forschungsaufwand in Personenjahren annehmen, um das derzeitige vorhandene Ausmaß an Wissen und Fertigkeiten mit eigenen Anstrengungen erreicht zu haben? Wie hätten sich diese zusätzlichen FuE-Personenjahre auf die vergangenen Jahren verteilt?“* Weitere Fragen beschäftigten sich mit den Kosten für ein FuE-Personenjahr und mit den Aufwendungen für die hypothetische zusätzliche Forschungsausstattung.

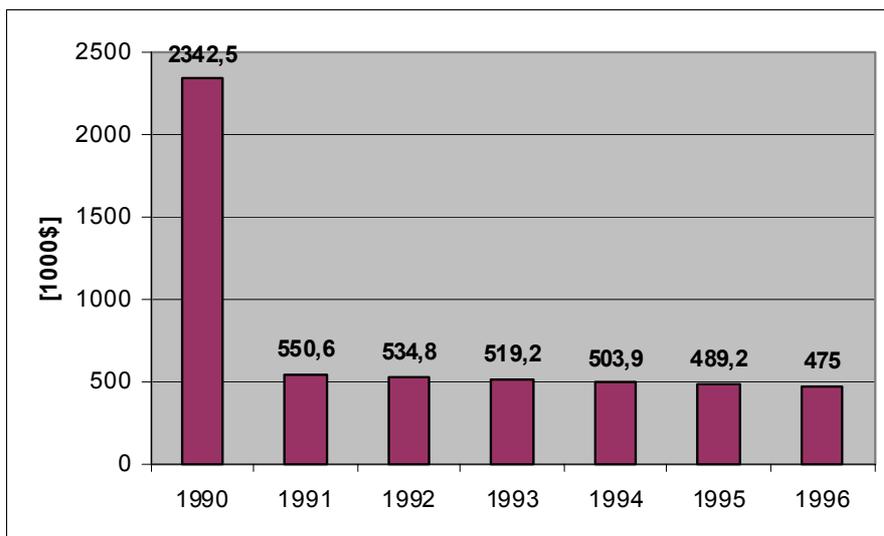
Abbildung 16



Quelle: Shedlick et al., 1998

Die Verwender von Kältemitteln wurden ebenfalls befragt, welcher zusätzliche Aufwand notwendig gewesen wäre, um die gleichen Ergebnisse ohne NIST-Forschungsprogramm zu erreichen. Hier zeigte sich, dass die Unternehmen mit dem Forschungsaktivitäten des NIST wenig vertraut waren, obwohl auch sie die NIST-Software REFPROP intensiv nutzten. Darüber hinaus verstanden die Verwender ihre Aktivitäten zum FCKW-Ersatz weniger als „Forschung und Entwicklung“ und eher als „Komponentenentwicklung“ (Shedlick et al., 1998, p.21).

Abbildung 17

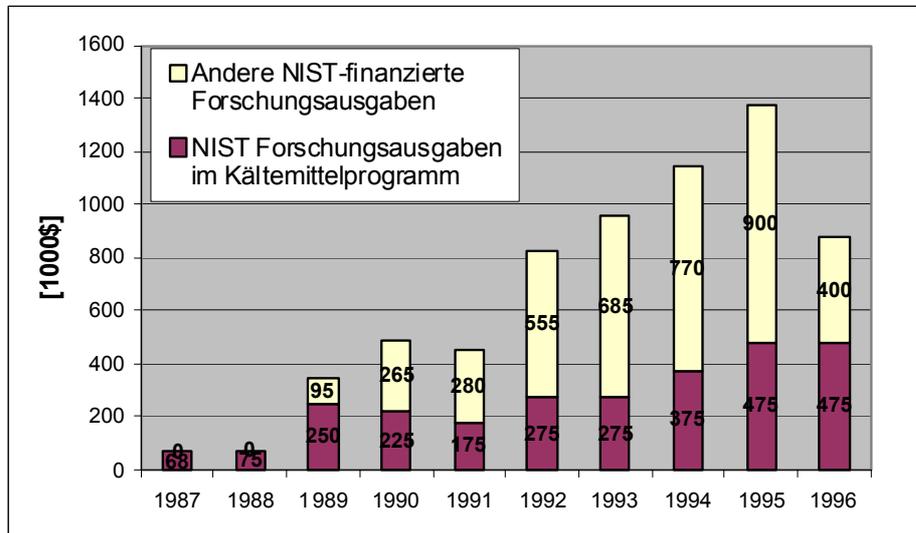


Quelle: Shedlick et al., 1998

Aufwendungen des NIST

Den Nutzen der Hersteller und der Verwender von alternativen Kältemitteln werden die jährlichen Ausgaben des NIST für das Kältemittelprogramm gegenüber gestellt:

Abbildung 18



Quelle: Shedlick et al., 1998

Für die Berechnung des Nettogesamtnutzens des Nist-Programms wurden die Aufwendungen der Jahre 1994 bis 1996 nicht berücksichtigt, da die entsprechenden Mittel noch nicht in ein Produkt eingeflossen waren, von dem die Industrie hätte profitieren können. Der alternative Ansatz wäre gewesen, diese Forschungsaufwendungen doch einzurechnen, gleichzeitig aber auch den erwarteten Nutzen der Industrie für die Jahre 1997 bis 1999 zu berücksichtigen.

Ökonomische Auswertung der Erhebung

Aus den oben angeführten Daten errechnen sich die folgenden Werte:

- Internal (Social) Rate of Return = 433%
- Implied Rate of Return = 20,7%
- Kosten-Nutzen-Verhältnis = 1 : 3,9

5.4 Anwendungsgebiete, Diskussion und Folgerungen

5.4.1 Erfahrungen des NIST mit dem Impact-Modell

Bei der Schätzung der ökonomischen Impacts eines Nist-Programms wird ein relativ konservativer Ansatz gewählt. Wie das Beispiel zeigt, wurden ausschließlich Effekte der ersten Ebene zur Bestimmung des Nutzens herangezogen. Der ausgewiesene Wert für den gesellschaftlichen Nutzen (SRR) kann daher als untere Grenze des tatsächlichen Nutzens verstanden werden. Nicht mit berücksichtigt wurden beispielsweise die vermiedenen

Transaktionskosten der Industrie durch das Programm. Wären keine verlässlichen Daten über die Eigenschaften alternativer Kältemittel verfügbar gewesen, hätten die verarbeitenden Unternehmen auf weniger umfassende, weniger verlässliche und weniger homogene Herstellerdaten zurückgreifen müssen. Die Kosten für die Überprüfung der Daten durch die Verwender wären beträchtlich gewesen. Auch ist anzunehmen, dass die Industrie durch eigene Anstrengungen nie ein vergleichbares Niveau an standardisierten Daten hätte erarbeiten können.

Ein weiterer Faktor, der nicht in die Berechnung eingeflossen ist, hängt mit Verbesserungen der Energieeffizienz von Kühlgeräten und Kälteanlagen zusammen, die ebenfalls eine mittelbare Auswirkung des Nist-Programms sind. Ohne Durchführung des NIST Forschungsprogramms wäre anzunehmen, dass in der ersten Phase nur suboptimale Lösungen für den Einsatz alternativer Kältemittel realisiert worden wären. Eine mangelnde Optimierung des Kältemittels hätte beträchtliche Konsequenzen für den Energieverbrauch der Anlagen bei einer geforderten konstanten Kühl- bzw. Heizleistung.

5.4.2 Nutzen des Modells für das NIST

Der Nutzen der Evaluierungen für das NIST kann auf drei verschiedenen Ebenen gesehen werden. Zum Ersten trägt das Evaluierungsverfahren maßgeblich dazu bei, Art und Ausmaß der ökonomischen Effekte von öffentlichen Forschungsprogrammen besser zu verstehen. Zweitens liefern die Evaluierungen neue Einsichten über die Innovationsdynamik in den betroffenen Industrien und Märkten. Ein dritter wesentlicher Punkt ist die Nutzung der Evaluierungsergebnisse im Politikkreislauf und im Budgeterstellungprozess des NIST. Hier ist besonders auf die gesetzlichen Verpflichtungen von Bundesbehörden durch den GPRA-Government Performance and Results Act hinzuweisen (Doremus, 1999). Schließlich trägt das Evaluierungsverfahren auch dazu bei, das Programmdesign und das Programmmanagement am NIST insgesamt zu verbessern.

5.4.3 Grenzen des Verfahrens zur Impact-Abschätzung

Einige einschränkende Bemerkungen hinsichtlich der Brauchbarkeit des geschilderten Impact-Evaluierungsverfahrens sollen nicht vergessen werden. Die Robustheit der Ergebnisse hängt maßgeblich von der Verfügbarkeit und der Qualität der zugrunde liegenden Daten ab. In zahlreichen Studien bieten diese Aspekte Anlass für eine kritische Auseinandersetzung. Viele direkte und indirekte Effekte von Forschungsprogrammen können seriöserweise nur äußerst grob geschätzt werden, bzw. sie entziehen sich überhaupt einer empirischen Bestimmung.

Die Werte für die quantitativen Maßzahlen hängen zusätzlich vom Zeitpunkt ab, zu dem die Kosten und Nutzen der Programme auftreten. Es kann daher durchaus vorkommen, dass für zwei Programme, die hinsichtlich ihrer Nettogesamtkosten und des Nettogesamtnutzens übereinstimmen, unterschiedliche SRR erhalten werden. Dies beschränkt die Anwendung des Modells für einen Vergleich der Performance verschiedener Programme.

Diese Grenzen der Impact-Abschätzung sind auch der Grund, warum am NIST die quantitative Impact-Evaluierung nur *ein* Element innerhalb des Evaluierungssystems für die

Laboratorien darstellt. Um den Erfolg und die Auswirkungen der Programme möglichst umfassend bewerten zu können, werden Peer Reviews, Output-Indikatoren, Erhebungen der Kundenzufriedenheit und die quantitative ökonomische Bestimmung der Programmimpacts gleichermaßen berücksichtigt. Erst die Zusammenschau der Ergebnisse aller vier Ansätze bildet ein tragfähiges Fundament für institutionelles Lernen und für die Weiterentwicklung der Laboratoriumsprogramme am NIST (Doremus, 1998).

5.4.4 Anwendbarkeit des Modells in Österreich

Eine Anwendung des NIST-Modells zur Evaluierung von Laboratoriumsprogrammen wäre in Österreich vor allem bei der Institutionenevaluierung von anwendungsorientierten, außeruniversitären Forschungseinrichtungen möglich. Klar abgrenzbare Programme, bzw. große durchgeführte Forschungsvorhaben des Österreichischen Forschungszentrums Seibersdorf, von Joanneum Research oder von Arsenal Research, bei denen die Ergebnisse der Forschungen einem breiteren NutzerInnenkreis zur Verfügung standen, könnten durch die hier vorgestellte Methode evaluiert werden.

5.5 Literatur

- Doremus, Paul N. (1998): Performance Evaluation at the NIST Measurement and Standards Laboratories. Powerpoint-presentation given at the Workshop on Methods to Evaluate Research Outcomes. Committee on Science, Engineering, and Public Policy, National Academy of Sciences, June 8, 1998.
- Doremus, Paul, N. (1999): GPRA Performance Reporting: Views from an S&T Bureau. Powerpoint-presentation given at the Fall Symposium of the American Association for Budget and Program Analysis, 15 November 1999.
- Link, Albert N. (1996): Economic Impact Assessments: Guidelines for Conducting and Interpreting Assessment Studies (<http://www.nist.gov/director/prog-ofc/rept-961.htm>), Zugriff: 24. Jänner 2000.
- Link, Albert N. and Scott, John T. (1997): Evaluating Technology-based Public Institutions: Lessons from the National Institute of Standards and Technology. In: OECD (1997) (ed.): Policy Evaluation in Innovation and Technology, Towards Best Practices (Paris).
- Link, Albert N. and Scott, John T. (1998): Public Accountability: Evaluating Technology-Based Institutions (Kluwer, Norwell).
- NAS (2000): Board on Assessment of National Institutes of Standards and Technology (NIST) Programs (online: <http://www4.nationalacademies.org/cpsma/nist.nsf>), Zugriff: 2. Juni 2000.
- NIST (2000): Outputs and Outcomes of NIST Laboratory Research (online: <http://www.nist.gov/director/planning/studies.htm>), 17. April 2000.
- NIST (2000): Strategic Planning and Economic Analysis at NIST (online: <http://www.nist.gov/director/planning/strategicplanning.htm>), 17. April 2000.
- OECD (1997) (ed.): Policy Evaluation in Innovation and Technology, Towards Best Practices (Paris).
- Shedlick Matthew T., Link, Albert N. and Scott John T. (1998): Economic Assessment of the NIST Alternative Refrigerants Research Program. NIST-Planning Report 98-1 (Gaithersburg).
- Tassey, Gregory (1999): Lessons learned about the methodology of economic impact studies: the NIST experience. Evaluation and Program Planning 22 (1999) 113-119.
- Tassey, Gregory (1999): Assessing the economic impacts of government R&D programs. Presentation to Technology Transfer Society, May 20, 1999.

6 EIN MODELL ZUR ABSCHÄTZUNG DER KOSTENEFFEKTIVITÄT: DIE EVALUATION VON INNOVATIONSFÖRDERPROGRAMMEN DES BRITISCHEN DEPARTMENT OF TRADE (DTI)

*Anton Geyer
Austrian Research Centers Seibersdorf*

6.1 Einleitung und Hintergrund

Das Department of Trade and Industry (DTI) ist in Großbritannien verantwortlich für die Umsetzung der Innovationspolitik der britischen Regierung. Ziel der innovationsfördernden Maßnahmen des Ministeriums ist es, neue Ideen und neue Technologien erfolgreich zu verwerten, um damit die Wettbewerbsfähigkeit der britischen Industrie zu erhöhen. Dieses Ziel soll durch eine Reihe von Programmlinien verwirklicht werden, die auf die jeweiligen Bedürfnisse der Industrie abgestimmt sind.

Die Förderinitiativen des DTI lassen sich in zwei Kategorien einteilen. Einerseits werden Förderprogramme durchgeführt, die Firmen Zuschüsse für die Entwicklung innovativer Technologien gewähren (Technologieentwicklungsprogramme). Andererseits existieren Programme, deren hauptsächliches Ziel es ist, den Transfer von Technologien, bzw. von technischen Fähigkeiten und Best-practise Methoden zu fördern. Die geförderten Initiativen können sowohl im technischen als auch im Managementbereich der Unternehmen ansetzen (Technologietransfer- bzw. Diffusionsprogramme).

Großbritannien ist ein Land, in dem in vielen Bereichen Grundlagenforschung auf höchstem internationalen Niveau durchgeführt wird, gleichzeitig aber die industrielle Verwertung neuer, innovativer Technologien vergleichsweise schlecht funktioniert. Um diesem Problem wirksam zu begegnen, kam es bereits 1993 zu einer Neuorientierung der britischen Innovationspolitik, die der Förderung von Technologietransfer- und Technologiediffusionsmaßnahmen im Vergleich zur Technologieentwicklung eine höhere Bedeutung zumisst. Konkret beabsichtigte das Ministerium, den prozentuellen Anteil an Technologiediffusionsförderungen am gesamten Innovationsbudget von 23% im Haushaltsjahr 1992-93 auf mehr als 50% im Haushaltsjahr 1996-97 zu erhöhen.

Im Rahmen einer größeren Verwaltungsreform kam es in den Jahren 1993 und 1994 am DTI zu einer Reorganisation der Administration der Innovationsprogramme. Insbesondere wurden die bis dahin bestehenden Regionalstellen des DTI mit den Büros anderer Ministerien zusammengelegt. Gleichzeitig wurde damit begonnen, bestimmte Aufgaben zur Betreuung von Firmen an ein Netzwerk von rund 200 regionalen Anlaufstellen zu übertragen.

Die Phase der Neustrukturierung des DTI wurde vom NAO-National Audit Office (britischer Rechnungshof) dazu genutzt, **insgesamt acht Technologieentwicklungs- und Technologiediffusionsprogramme des DTI gemeinsam zu evaluieren**. Das Evaluierungsprogramm des NAO umfasste vier Bereiche (NAO, 1995, p.16):

- Eine telefonische Unternehmensbefragung wurde durchgeführt, in der der Bekanntheitsgrad der Programme und die Einschätzung der Wirtschaft hinsichtlich der Qualität der Durchführung durch das DTI erhoben wurde. In der Umfrage wurden auch die Projektergebnisse bei den teilnehmenden Firmen erhoben.
- Die relative Kosteneffektivität (relative cost-effectiveness) der Programme wurde bestimmt. Zu diesem Zweck wurde ein analytisches Modell entwickelt und für die verschiedenen Programme Leistungsindikatoren erarbeitet, gewichtet und aggregiert. Dieses Modell erlaubt es, die einzelnen Programme sowohl hinsichtlich einzelner, als auch hinsichtlich aller Indikatoren gemeinsam untereinander zu vergleichen. Die Methode liefert jedoch keine Absolutwerte für die Kosteneffektivität, sondern ermöglicht Aussagen über das relative Verhältnis zwischen verschiedenen Programmen.
- Ein internationaler Vergleich von Innovationsförderprogrammen wurde durchgeführt. Das NAO versandte Fragebögen in insgesamt elf Staaten, besuchte die mit Evaluierung befassten Abteilungen bei der Europäischen Kommission und führte Interviews mit VertreterInnen von Regierungsbehörden in anderen Ländern.
- Zusätzlich führte das NAO zahlreiche Interviews mit DTI-Programmmanagern und EvaluatorInnen. Insgesamt 48 Projekte aus allen Programmlinien wurden genauer untersucht.

Das relative Kosteneffektivitäts-Modell wurde mit Unterstützung einer privaten Beratungsfirma und von ExpertInnen der London School of Economics entwickelt (NAO, 1995, p.51).

Eine Dokumentation über die Ergebnisse der NAO-Evaluierung wurde im August 1995 unter dem Titel „The Department of Trade and Industry's Support for Innovation“ veröffentlicht (NAO, 1995). Die Darstellung der Cost-Effectiveness-Methode in diesem Berichtes bezieht sich überwiegend auf diese Publikation. Einen Überblick über die aktuellen DTI-Förderprogramme (<http://www.dti.gov.uk/support/index.htm>) und Informationen über die dazu durchgeführten Evaluierungen (<http://www.dti.gov.uk/tese/au.html>) geben die Internetseiten des DTI.

6.2 Das NAO-Modell zur Abschätzung der relativen Kosteneffektivität (relative cost-effectiveness)

Als Grundlage für die Anwendung der Cost-effectiveness-Methode auf die acht DTI-Programmlinien standen die Daten, Informationen und Auswertungen zur Verfügung, die im Rahmen des DTI-Planungs- und Evaluierungskreislaufes erhoben und erarbeitet werden. Das unter dem Akronym ROAME bekannte Planungs- und Evaluierungsverfahren sieht eine systematische Begründung und Zielfestlegung in den Programmen, ein verbindliches Projektbegutachtungs- und Monitoringverfahren, sowie entsprechende Evaluierungs- und Lernstrukturen vor (vgl. dazu Hills und Dale, 1995, Guy, 1998).

Für die Erarbeitung des relativen Kosteneffektivitätsmodells wurde im Jahr 1994 eine gemeinsame Arbeitsgruppe des DTI und des NAO eingerichtet. Diese Gruppe erarbeitete ein Modell, mit dessen Hilfe der Erfolg der Programme hinsichtlich der Stimulierung von Innovationen, sowie hinsichtlich der Steigerung der Wettbewerbsfähigkeit bewertet werden kann. Die Durchführung der Untersuchung umfasste die folgenden Arbeitsschritte:

- Identifizierung geeigneter Kriterien, die als Grundlage für ein Indikatorensystem zur Beurteilung der relativen Kosteneffektivität von DTI-Programmen dienen können.
- Prüfung und Sicherstellung, dass die festgelegten Kriterien und die darauf aufbauenden Indikatoren tatsächlich ein brauchbares Maß für die Kosteneffektivität der Programme sind.
- Zusammenfassung der identifizierten Kriterien und Indikatoren in ein Modell zur vergleichenden Bewertung der Kosteneffektivität.
- Anwendung des Modells auf insgesamt acht Förderprogramme des DTI.

6.2.1 Kriterien und Indikatoren für die Ermittlung der Programmergebnisse

Der im NAO-Modell herangezogene Ansatz basiert auf einer Adaption der Entscheidungstheorie für die Programmevaluierung (NAO, 1995). Die Methode erlaubt auch den Vergleich von Programmen, die sich hinsichtlich einzelner Evaluationskriterien bzw. bestimmter Ergebnisindikatoren unterscheiden.

Dem NAO-Modell liegen zwei übergeordnete Programmziele (das sind die zentralen Kriterien für die Beurteilung der Programme) zugrunde. Einerseits sollen die mit den Programmen verbundenen *Kosten minimiert*, andererseits ihr *Nutzen* (im Sinne der Stimulierung innovativer Entwicklungen und gesteigerter Wettbewerbsfähigkeit) *maximiert* werden. Beide Ziele gemeinsam bilden die Grundlage für die Beurteilung der Kosteneffektivität für jedes einzelne Programm. Aus den beiden Hauptzielen werden sieben verschiedene Beurteilungskriterien abgeleitet, die ihrerseits durch 19 einzelne Indikatoren erfasst und gemessen werden können.

6.2.2 Bestimmung der Kosteneffektivität verschiedener Programme

Im ersten Schritt werden die verfügbaren Daten aus der Programmplanung, dem Programmmonitoring und der Evaluierung dazu genutzt, den einzelnen Indikatoren entsprechende Werte zuzuordnen. In einem zweiten Schritt werden diese Werte auf eine sogenannte *Präferenzskala* übertragen. Mit diesem Begriff wird eine relative Skala bezeichnet, auf der ein Wert von „100“ das höchste Präferenzniveau für einen bestimmten Indikator anzeigt. Ein Präferenzwert von „0“ weist ein Programm aus, das hinsichtlich eines bestimmten Indikators die niedrigste Präferenz *relativ zu anderen Programmen* hat. Der Wert „0“ bedeutet also nicht, dass das beurteilte Programm überhaupt nicht kosteneffektiv ist, sondern nur, dass es im Vergleich zu anderen Programmen und in bezug auf einen bestimmten Indikator die geringste Präferenz aufweist.

Die einzelnen Programme werden – nach ihrer Kosteneffektivität gereiht – auf den Präferenzskalen der verschiedenen Indikatoren eingeordnet. Ein Programm, das hinsichtlich eines bestimmten Indikators genau zwischen der Kosteneffektivität des besten und des am schlechtesten abschneidenden Programms liegt, bekommt also in diesem Modell den Präferenzwert „50“ zugewiesen.

Tabelle 13

Kosten und Nutzen	Ergebniskriterien	Indikatoren	Beschreibung der Indikatoren
Innovation und Wettbewerbsfähigkeit	Verwertung der Ergebnisse	LICS ISS'D	Anzahl der vergebenen Lizenzen
		COM'NT TO EXPLT	Ausmaß der Aktivitäten bei der Produktverwertung
		PRODUCTS	Anzahl der markteingeführten Produkte nach 36 Monaten
	Forschung und Entwicklung	COM'NT R&D	Ausmaß des FuE Engagements
		PROJECTS	Anzahl abgeschlossene Projekte
	Erwerb von Know-how	CONT COLLB	Anzahl kooperierende Firmen nach Projektende
		TECH ACQUISITION	Anzahl bedeutender technischer Verbesserungen
		BUS PRACT	Anzahl bedeutender organisatorischer Verbesserungen
		ICS EXECD	Anzahl erworbenen Lizenzen
		COLL AGMTS	Anzahl der Organisationen mit Kooperationsverträgen
	Zugang zum Programm	CONTACTS	Anzahl der Informationskontakte
		CONTRACTS	Anzahl der Einrichtungen in den geförderten Anträgen
	Bekanntheit des Programms	APPLICANTS	Anzahl der Projektanträge
		PARTICIPANTS	Anzahl der Programmteilnehmer pro Jahr
Kosten	Nicht-monetäre Kosten	FAILED APPNS	Anzahl der abgelehnten Anträge
		PROCTIME	Durchschnittliche Bearbeitungszeit der Anträge
		OUT BODIES	Anzahl der involvierten externen Einrichtungen
	Monetäre Kosten	PROGRAM	Kosten des Programms
		RUNNING	Kosten der Programmadministration

Quelle: NAO, 1995, p.53

Im nächsten Schritt werden die verschiedenen Präferenzskalen gewichtet, um sie untereinander vergleichbar zu machen. Dem als am wichtigsten eingeschätzten Indikator wird ein relatives Gewicht von 100% zugeordnet. Für die anderen Indikatoren werden Gewichte festgelegt, die zwischen 0 und 100% liegen können. Dieser Schritt wird jeweils getrennt für die Kosten- und die Nutzenindikatoren durchgeführt. Im letzten Schritt werden die gewichteten Präferenzwerte für die Kosten- und die Nutzenseite der einzelnen Programms addiert.¹³

¹³ Die relative Gewichtung der einzelnen Indikatoren zueinander beeinflusst natürlich die erhaltenen Resultate bei der Evaluierung.

Die erhaltenen Werte für die relative Kosteneffektivität der Programme können nun miteinander verglichen werden. Auch ist es möglich, die Ergebnisse hinsichtlich einzelner Indikatoren in den jeweiligen Programmen zu verfolgen.

6.2.3 Untersuchte DTI-Programme

Das Evaluierungsmodell wurde für die Beurteilung der folgenden acht DTI-Programme herangezogen:

Tabelle 14

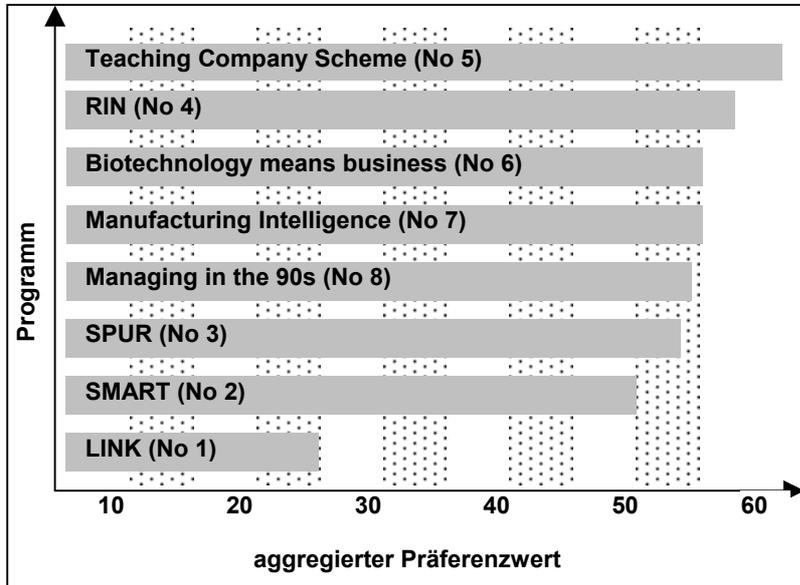
	Programmbeschreibung	Durchführung	Budget 91-94
Technologieentwicklungsprogramme			
LINK	Ressortübergreifendes Programm (seit 1986) unter Einbeziehung von neun Ministerien und vier Research Councils, das industriell relevante Forschungsk Kooperationen zwischen Wirtschaft und Universitäten in bestimmten Technologiebereichen fördert. Das LINK Programm steht allen Unternehmen offen.	DTI-Zentrale	49,8 m£
SMART	Beihilfenprogramm für Kleinunternehmen (< 50 Beschäftigte) auf Basis eines jährlichen Wettbewerbs. Das Programm zielt darauf ab, Innovationen anzuregen und vermarktbar e Produkte zu entwickeln. Die Beihilfen werden in zwei Phasen vergeben: In der ersten Phase kann eine Machbarkeitsstudie gefördert werden, in der zweiten Phase wird die Prototypenentwicklung unterstützt.	DTI-Zentrale, Regionalstelle	45,9 m£
SPUR	Beihilfenprogramm für Klein- und Mittelbetriebe (< 500 Beschäftigte) mit dem die Entwicklung von technologisch avancierten neuen Produkten und Prozessen unterstützt wird.	Regionalstelle	14,8 m£
RIN	Beihilfenprogramm für Kleinbetriebe (< 50 Beschäftigte) in bestimmten Regionen, mit dem die Entwicklung kommerziell erfolgreicher, innovativer Produkte unterstützt wird.	Regionalstelle, Umweltbehörde	20,0 m£
Technologietransfer-Programme			
Biotechnology Means Business	Das Programm diente der Verbreitung von Informationen über die wirtschaftliche Bedeutung von Biotechnologien für potentielle industrielle Anwender. Das Programm bestand aus zwei sich überschneidenden Phasen (Laufzeit 1990 bis 1993, bzw. ab 1992).	DTI-Zentrale	0,8 m£
Manufacturing Intelligence	Das Programm förderte die Anwendung von Software-Technologien mit dem Ziel, Probleme in den Bereichen Produktionstechnologien und -management, Produktdesign, Marketing, Vertrieb und Kundendienst zu lösen. Das Programm thematisierte die Bedeutung von Software-Technologien für die Steigerung der Wettbewerbsfähigkeit (Laufzeit: 1990 bis 1993).	DTI-Zentrale	0,9 m£
Teaching Company Scheme	Ressortübergreifendes Programm (seit 1975), das im Auftrag des DTI sowie neun anderer öffentlicher Förderungsgeber durch ausgelagerte Teaching-Company-Büros administriert wird. Das Programm unterstützt den Transfer von Wissen und Fertigkeiten von den Universitäten in die Wirtschaft mittels der Förderung von Anstellungen von Universitätsabsolventen in Unternehmen für einen Zeitraum von zwei bis drei Jahren.	Teaching-Company-Abteilung	21,1 m£
Vermittlung von Best-Practice Methoden			
Managing in the 90s	Das Programm förderte Best-practise Management in vier Kernbereichen: Produktion, Einkauf, Qualität, Design. Das Programm startete 1989. Zum Zeitpunkt der NAO Evaluierung lagen die Schwerpunkte beim Innovations- und Change-Management.	DTI-Zentrale, Regionalstelle	6,2 m£

Quelle: NAO, 1995, p.14

6.2.4 Evaluierungsergebnisse

Nach der Bestimmung der relativen Kosteneffektivität der einzelnen Programme ergab sich ein Bild, welches das TCS-Programm vor dem RIN-Programm an die erste Stelle reihte, während das LINK-Programm abgeschlagen an der letzten Stelle lag.

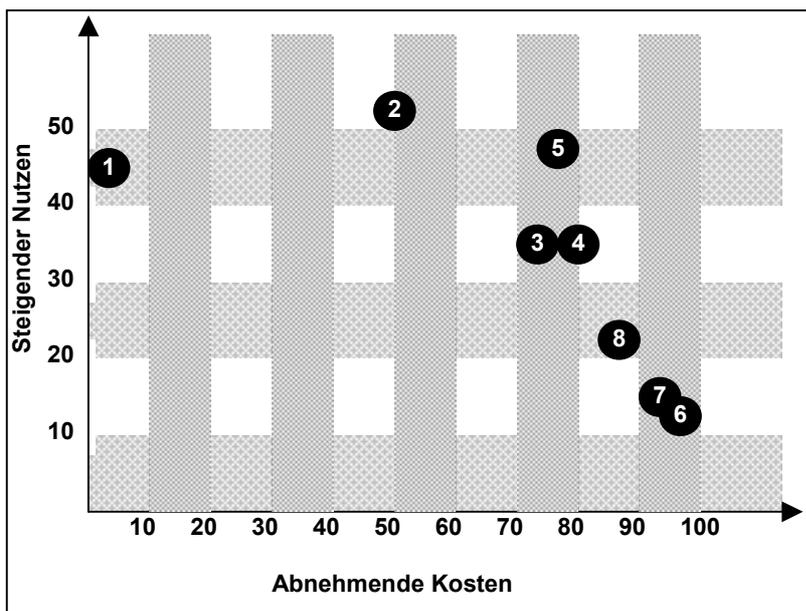
Abbildung 19



Relative Kosteneffektivität von DTI-Programmen, Quelle: NAO, 1995, p.36

Um ein besseres Bild über die relative Kosteneffektivität der einzelnen Programme zu erhalten, werden in der folgenden Abbildung die Nutzen- (Innovationen und gesteigerte Wettbewerbsfähigkeit) und Kostenkriterien (monetäre und nicht-monetäre) getrennt betrachtet und in einer zweidimensionalen Matrix dargestellt.

Abbildung 20



Relative Kosteneffektivität von DTI-Programmen, Quelle: NAO, 1995, p. 36

6.2.5 Diskussion des NAO-Modells

Um die Evaluierungsmethode auf Basis der relativen Kosteneffektivität zu überprüfen, hat das NAO bereits bei der Durchführung seiner Untersuchung

- (1) die Auswirkungen von Veränderungen bei der Gewichtung von Kosten- und Nutzenindikatoren auf die Ergebnisse der vergleichenden Evaluierung bestimmt,
- (2) die Haupteinflussfaktoren, die das Ranking der Programme besonders beeinflussen, ermittelt und daraus
- (3) mögliche Ansätze für die Verbesserung der Programme hinsichtlich bestimmter Indikatoren diskutiert.

Auswirkungen von Veränderungen der Gewichtung von Kosten und Nutzen

Die oben angeführten Ergebnisse der relativen Kosteneffektivität der Programme basieren auf einer Gleichgewichtung der beiden Hauptbeurteilungskriterien Kosten und Nutzen (jeweils 50% Beitrag zum Gesamtergebnis). Wird die relative Gewichtung von Kosten und Nutzen im Modell verändert, ändert sich auch die Reihung der verschiedenen Programme. Wenn den Kosten eine höhere Gewichtung zugewiesen wird, verbessern sich die Technologietransfer- und Best Practise-Programme relativ zu den Beihilfenprogrammen. Nur im Falle von LINK wird das Ranking von der Gewichtung der Kosten- und Nutzenseite nicht beeinflusst. Die folgende Tabelle illustriert die Auswirkungen von Veränderungen der Gewichtung von Kosten und Nutzen auf die relative Kosteneffektivität der Programme.

Tabelle 15

Programm	Gewichtungsanteil der Kostenfaktoren		
	25%	50%	75%
Teaching Company Scheme	55 (1)	61 (1)	68 (5)
RIN	47 (3)	57 (2)	69 (4)
Biotechnology Means Business	34 (7)	55 (3)	79 (1)
Manufacturing Intelligence	35 (6)	55 (3)	77 (2)
Managing in the 90s	40 (5)	54 (5)	71 (3)
SPUR	44 (4)	53 (6)	64 (6)
SMART	52 (2)	50 (7)	49 (7)
LINK	34 (7)	25 (8)	15 (8)

Quelle: NAO 1995, p.38

Haupteinflussfaktoren auf die Evaluierungsergebnisse

Durch eine genauere Analyse des Modells lässt sich zeigen, dass die Verwendung von nur 12 der 19 Indikatoren zu einem Ergebnis führt, das zu 96% jenem entspricht, welches man erhält, wenn alle 19 Indikatoren zur Berechnung herangezogen werden. Das Teaching

Company Scheme Programm erreicht vor allem deswegen das höchste Ranking, weil es über alle zwölf Kern-Indikatoren hinweg gute Ergebnisse liefert. Insbesondere sind die relativen Programmkosten gering, die Bearbeitungszeit der Anträge ist kurz und das Programm fördert Kooperationen. Im Gegensatz dazu erzielt das LINK Programm vergleichsweise schlechte Ergebnisse, weil es trotz eines auf die Förderung von Kooperationen ausgerichteten Programmdesigns hinsichtlich vieler anderer Indikatoren gegenüber anderen Programmen zurückfällt. Insbesondere ist die Bearbeitungszeit für Projektanträge relativ lang, die Kosten für die Administration des Programms sind hoch und die Verwertung der Projektergebnisse ist im Vergleich zu anderen Programmen bescheiden. Die nachfolgende Tabelle zeigt den Beitrag der 12 wichtigsten Modellindikatoren zum Ranking der verschiedenen Programme. Die Indikatoren sind entsprechend ihrer Bedeutung im Modell absteigend angeordnet.

Abbildung 21



Beitrag der zwölf wichtigsten Indikatoren zum Ergebnis : Quelle: NAO, 1995, p.39

6.3 Erfahrungen und Anwendungsmöglichkeiten

Aus der obigen Tabelle versuchte das NAO Vorschläge abzuleiten, wie Verbesserungen bei der relativen Kosteneffektivität von DTI-Programmen erreicht werden können. Dabei ist zu berücksichtigen, dass die Ergebnisse hinsichtlich einzelner Indikatoren zu einem guten Teil eine direkte Folge des Designs der Programme sind. Beispielsweise lässt sich der vergleichsweise geringe Präferenzwert des Programms „Managing in the 90s“ bei projekt- und marktbezogenen Indikatoren darauf zurückführen, dass unter diesem Programm keine Projekte gefördert wurden (die aber einen wesentlichen Beitrag zu hohen Präferenzwerten bei dieser Indikatorengruppe leisten). Gleichfalls sind die geringen Präferenzwerte des LINK-Programms bei marktbezogenen Indikatoren - zumindest teilweise - Folge einer relativ starken Orientierung des Programms auf vorkompetitive Forschung, bei der das Stadium der Markteinführung von neuen Produkten noch relativ fern ist.

Trotz der nicht immer vergleichbaren Rahmenbedingungen zog das NAO folgende Schlussfolgerungen aus der vergleichenden Evaluierung der Programme:

- Die große Gewichtung der Indikatoren Programmkosten und Administrationskosten streichen die Bedeutung einer Minimierung von Mitnahmeeffekten bei Technologieentwicklungsprogrammen heraus. Gleichzeitig ist auf die Effizienz der Projektadministration besonders zu achten.
- Innerhalb der Gruppe der Innovations- und Wettbewerbsindikatoren (Nutzen) besitzen die produkt- und projektbezogenen Indikatoren ein höheres Gewicht, als jene, die den Informationstransfer in den Programmen berücksichtigen.
- Diffusionsprogramme erzielen dann ein gutes Ergebnis, wenn die Inhalte und die Zielgruppen so definiert sind, dass die Wahrscheinlichkeit von relevanten Änderungen hinsichtlich von Managementpraktiken oder der Verwendung von Technologien hoch ist. Der im Teaching Company Scheme-Programm verfolgte Ansatz, der Technologieentwicklungs- und -diffusionselemente miteinander verbindet, schien dafür besonders geeignet zu sein.
- Die Ergebnisse zeigen, dass für das SPUR- und das LINK-Programm Verbesserungen im Rating durch eine Konzentration auf effizienteres Projektmanagement und eine bessere wirtschaftliche Verwertung der Projektergebnisse durchaus möglich wären. Einen wesentlichen Beitrag zur Erreichung dieses Zieles könnte die Verbesserung der ex-ante Evaluierungsverfahren der eingereichten Projektanträge leisten, mit der die Bearbeitungszeit der Anträge entsprechend beeinflusst werden können.

Weiters wurde dem DTI vorgeschlagen, das vorgestellte Evaluierungsmodell der vergleichenden Kosteneffektivität in die Evaluierungsstrukturen am DTI einzubinden und weiterzuentwickeln. Die dabei gemachten Erfahrungen sollten auch in den Prozess der Budgeterstellung und der Mittelallokation einbezogen werden.

6.3.1 Mögliche Anwendungsgebiete des Modells in Österreich

Die Erfahrungen mit dem Modell in Großbritannien zeigen, dass auch bei quantitativen und formalisierten Evaluierungsverfahren (inter-) subjektive Einschätzungen und Bewertungen einen starken Einfluss auf die Ergebnisse haben. Die vorgestellte Methode ist dann gut

anwendbar, wenn zwischen den Beteiligten Einigkeit über die relative Gewichtung der verschiedenen Kosten- und Nutzenindikatoren besteht. Die Erfahrungen mit der Evaluierung der DTI-Programme zeigt weiters, dass die Methode ihre Grenzen bezüglich der Vergleichbarkeit verschiedener Programme hat, wenn sich die Ausrichtung und Zielsetzung, sowie die eingesetzten Instrumente in den Programmen allzu stark unterscheiden. Ein Vergleich der Kosteneffektivität von Programmen, die unmittelbar auf Produktentwicklung abzielen, und solchen die beispielsweise den Informationstransfer als Hauptaufgabe verfolgen, wird notwendigerweise zu Ergebnissen führen, die nur beschränkt aussagekräftig sind. Die NAO hat in der Diskussion der Ergebnisse zur DTI-Evaluierung auf dieses Problem auch hingewiesen (NAO, 1995).

Trotz dieser Einschränkungen wäre das NAO-Kosteneffektivitätsmodell auch für österreichische Programme ein interessanter Ansatz. Um den oben beschriebenen Beschränkungen des Modells Rechnung zu tragen, sollte die NAO-Methode jedoch nur auf Programme ähnlichen Typs angewandt werden. Für Österreich würde sich hier beispielsweise eine Anwendung auf die in den vergangenen Jahren durchgeführten Technologiediffusionsprogramme anbieten. Voraussetzung dafür wäre allerdings, dass in ausreichendem Ausmaß Daten über das Programmmanagement und die Programmresultate aus dem Monitoring und/oder aus bereits durchgeführten Evaluierungen vorliegen. Ohne umfassende Monitoring- und Evaluierungsdaten ist eine Ermittlung der im Modell verwendeten Indikatoren, und damit die Abschätzung der vergleichenden Kosteneffektivität der Programme kaum möglich.

6.4 Literatur

Department of Trade and Industry (DTI) (2000): DTI Business Support (online-Dokument: <http://www.dti.gov.uk/support/index.htm>), Zugriff am 13. März 2000.

Department of Trade and Industry (DTI) (2000): The Assessment Unit (online-Dokument: <http://www.dti.gov.uk/support/index.htm>), Zugriff am 13. März 2000.

Guy, Ken (1998): Strategic Options for the Evaluation of the R&D Programmes of the European Union. Final Report prepared for STOA (Brighton).

Hills, Philip; Dale, Alison (1995): Research and technology evaluation in the United Kingdom. Research Evaluation 5, 35-44.

National Audit Office (1995): The Department of Trade and Industry's Support for Innovation. Report by the Comptroller and Auditor General. HC 715 Session 1994-95 (HMSO, London).

7 COMEVAL EVALUIERUNGSTOOLKIT

Anton Geyer

Austrian Research Centers Seibersdorf

7.1 COMEVAL Entwicklungsgeschichte

Die Europäische Kommission versuchte Mitte der 90er Jahre mit der Entwicklung des Evaluierungstoolkits COMEVAL eine allgemeine Evaluierungsmethodik für F&E-Programme zu erarbeiten. In einem ersten Schritt wurden durch die Europäische Kommission Aufträge für die Entwicklung zweier Testversionen für den COMEVAL Toolkit vergeben. Diese Projekte bildeten die Basis für eine gemeinsame COMEVAL-Software Version, die vor allem zur Nutzung durch die Forschungsgeneraldirektion gedacht war.

Der COMEVAL-Toolkit wurden zwar in den vergangenen Jahren innerhalb der Kommission kontinuierlich weiterentwickelt (EUSTAT, 1998), jedoch nie einer breiteren Öffentlichkeit zugänglich gemacht (Georghiou, 2000). Die in der Folge beschriebenen Inhalte und Charakteristika des COMEVAL-Toolkits beziehen sich daher ausschließlich auf eine der Testversionen des Toolkits vom September 1996 (Clifton et al., 1996).

7.1.1 Ziele des COMEVAL-Toolkits

COMEVAL sollte ursprünglich im Fünften Forschungsrahmenprogramm das Programmmanagement dabei unterstützen, mit Hilfe von bestimmten Indikatoren und Deskriptoren standardisierte Projektauswertungen zu erstellen. Die innovativen Effekte der Projekte waren dabei von zentralem Interesse. Der Nutzen einer allgemein anwendbaren Methodik zum Monitoring von Forschungsergebnissen in den EU Forschungsprogrammen kann auf verschiedenen Ebenen gesehen werden:

- Auf der Ebene des Rahmenprogramms wird die Transparenz und die politische Verantwortlichkeit im Hinblick auf die übergeordneten politischen Ziele des Fünften Rahmenprogramms verbessert.
- Auf Programmebene kann eine allgemeine Evaluierungsmethodik dazu beitragen, innovative Ergebnisse und Trends von Projekten zu clustern, und die vorliegenden Informationen auch über die Programmgrenzen hinweg zu vergleichen. Dadurch würden die später durchzuführenden Evaluierungen unterstützt, und zusätzliche detaillierte analytische Auswertungen (z.B. makroökonomische Effekte) erstellt werden können.
- Die Verantwortlichen für bestimmte Projekte wären in der Lage ihr Portfolio im Hinblick auf die übergeordneten Zielen des Rahmenprogramms zu kontrollieren, und die Übereinstimmung von Projektzielen und Programmzielen regelmäßig zu überprüfen.
- Die Projektmanager könnten COMEVAL dazu nutzen, im Sinne eines umfassenden Projekt-Life-Cycle-Managements in verschiedenen Stadien des Projektzykluses die erwarteten und erreichten innovativen Effekte zu vergleichen.
- COMEVAL würde eine Möglichkeit bieten, mit den ProjektleiterInnen spezifische Anforderungen für die Projektberichterstattung zu vereinbaren, um damit die Rückmeldungen an die Kommission effizienter zu strukturieren.

Zum Stand März 1998 umfasste COMEVAL eine Software und eine "Gemeinsame Terminologie" für die Projektevaluation. COMEVAL war nicht als eine starre Methode gedacht, sondern vielmehr als ein flexibler Rahmen, der an die Anforderungen in den verschiedenen Programmen angepasst werden sollte. Zusätzlich zur Software wurde ein Handbuch erstellt, das die Nutzung der Software erleichtern und unterstützen sollte (Clifton et al., 1996).

7.1.2 Beschreibung der COMEVAL-Methode

Die COMEVAL-Methode erlaubt die Beschreibung und die graphische Darstellung der Ziele und Ergebnisse von Forschungsprojekten und ermöglicht das Nachvollziehen des „Programmprofils“ über den gesamten Projektzyklus, beginnend vom Projektstart, bis zum Ende der Implementierungsphase.

Die in COMEVAL verwendeten Indikatoren (bzw. Deskriptoren) zielen in erster Linie auf eine Charakterisierung der Projekte ab. Sie bezwecken nicht unbedingt, die Projekte hinsichtlich der Ziele und der erreichten Ergebnisse zu reihen. So ist beispielsweise ein Projekt das eine „radikal neue Technologie“ entwickelt nicht von vornherein *besser*, sondern nur *verschieden* von einem Projekt, das eine „wesentliche Verbesserung einer bestehenden Technologie“ als Projektziel oder Ergebnis aufweist. In dieser Hinsicht sind die meisten COMEVAL-Indikatoren mit einem Multiple-Choice-System vergleichbar, und nur in untergeordnetem Maße mit einem reihenden Bewertungsverfahren. Das heißt aber nicht, dass die BenutzerInnen der COMEVAL-Software bloß eine „Ankreuzübung“ durchführen. Sie sind aufgefordert, zu jeder ihrer Angaben eine kurze qualitative Erläuterung zu geben.

Die COMEVAL-Methodik deckt alle Phasen des Projektzykluses abdecken und soll

- eine ex-ante Bewertung von Projekten für die Projektauswahl ermöglichen;
- eine Überprüfung (Monitoring) der erzielten Ergebnisse gegenüber den Projektvorgaben, den vereinbarten Zeitplänen und den im Qualitätsplan festgelegten Qualitätskriterien ermöglichen, wobei gegebenenfalls darauf aufbauend Maßnahmen zur Neuausrichtung des Projektes gesetzt werden können;
- eine ex-post-Bewertung im Hinblick auf die Effektivität des Projektes auf Grundlage eines Vergleichs der erzielten Ergebnisse mit den Zielsetzungen, Zielvorgaben und den zu erbringenden Leistungen, sowie hinsichtlich der Effekte durch die Verwertung der Ergebnisse nach dem Projektende erlauben,;
- Lernprozesse für die Zukunft unterstützen.

Der umfassende Evaluierungsansatz von COMEVAL sollte insbesondere dazu dienen, die verschiedenen administrativen Ebenen bei der Evaluierung (Projekt, Programm, Rahmenprogramm) besser miteinander zu verknüpfen und alle relevanten AkteurInnen in den Evaluierungsprozess mit einzubeziehen: die Programmmanager stellen Evaluierungsanforderungen an COMEVAL; die Projektbetreuer nutzen den Toolkit für ihr Projektportfolio; die Auftragnehmer erstellen ihre Berichte entsprechend den COMEVAL-Vorgaben; die EvaluatorInnen nutzen die Ergebnisse für ihre Studien (EC, 1998).

Die COMEVAL-Methode ist so ausgerichtet, dass subjektive Bewertungen der EvaluatorInnen möglichst beschränkt, und damit die Vergleichbarkeit zwischen den Bewertungen verschiedener Projekte ermöglicht wird. Dies erscheint umso wichtiger zu sein, als die EU-Projektevaluierungen in einem äußerst heterogenen Umfeld erstellt werden, in dem verschiedene Programme, verschiedene EvaluatorInnen, unterschiedliche Arten von Projekten in verschiedenen Projektphasen aufeinander treffen. Um die Aussagekraft und die Vergleichbarkeit der Ergebnisse von COMEVAL-Auswertungen zu verbessern, wurde in der COMEVAL-Software zu jedem Indikator eine besondere Erläuterung erstellt, in der die Verwendung des Indikators anhand von Beispielen aus Projektevaluierungen genau beschrieben wird. Um die Erfassung der notwendigen Projektinformationen rationell zu gestalten, war daran gedacht, für jedes spezifische Programm und jede Key Action eine programmspezifische Version der „gemeinsamen Terminologie“ zu erstellen, mit deren Hilfe im Zuge der Vertragsverhandlungen mit den Projekteinreichern besondere Anforderungen an die Berichterlegung vereinbart werden sollten.

Die COMEVAL-Software fasst auf der Projektebene unterschiedliche Bewertungen zusammen und ermöglicht eine graphische Darstellung von Projektprofilen bzw. von Projektradars. Diese summative Form der COMEVAL-Indikatoren erleichtert eine statistische Auswertung der Projektergebnisse und macht eine weitere Aggregation auf höherer Ebene (z.B. Projektcluster oder auf Programmebene) möglich.

7.1.3 Entwicklungsphasen und Test des COMEVAL-Toolkits

Die Koordinierung der Entwicklung von COMEVAL wurde 1995 durch die damalige DGXIII/D wahrgenommen, die dabei durch eine Steuerungsgruppe unterstützt wurde. Dieser Arbeitsgruppe gehörten zusätzlich Direktoren von betroffenen Abteilungen in der DG III, der DG XII und des IPTS, sowie fünf Vertreter der Europäischen Industrie an. Um den Unterschieden in den verschiedenen Forschungsprogrammen gerecht zu werden, entschied die Steuerungsgruppe in der ersten Projektphase zwei verschiedene Versionen von COMEVAL ausarbeiten zu lassen, die sich auf verschiedene Forschungsbereiche konzentrieren sollten.

Da die Erstfassungen der beiden Testversionen jeweils spezifische Stärken und Schwächen aufwiesen, entschied die Steuerungsgruppe beide Versionen einem vergleichenden Test in der Kommission zu unterziehen. In der Abteilung DG XIII/D wurde im September 1996 eine Auswahl von 25 Forschungsprojekten mit COMEVAL evaluiert. Auf Grundlage der gemachten Erfahrungen wurde im Dezember 1996 damit begonnen, die beiden Testversionen umfassend zu vereinfachen und miteinander zu verschmelzen.

Diese Aufgabe wurde von der DG XIII/D übernommen. Im April 1997 wurde das Resultat (das COMEVAL Software-Tool) intern versendet und für eine zweite Testrunde eingesetzt. Die Rückmeldungen aus dem Test wurden in das COMEVAL Indikatorensystem eingearbeitet. Gleichzeitig wurde der Toolkit an die politischen Vorgaben und Auswahlkriterien für Projekte im Fünften Rahmenprogramm angepasst. Die zweite Testphase der Software führte zu einigen wesentlichen Vereinfachungen und Verbesserungen:

- Die Anzahl der Indikatoren wurde deutlich reduziert und besser auf die politischen Ziele des Fünften Rahmenprogramms abgestimmt
- Auch nicht-industrielle Forschungsprogramme mit hohem Charakter eines öffentlichen Gutes konnten bewertet werden
- Die Bewertungsskalen wurden geändert damit auch potentielle negative Effekte der Projekte erfasst werden können
- Ein Kern von Indikatoren erlaubte eine ad-hoc-Bewertung der Projekte auch ohne zusätzliche Datensammlung
- Die Software selbst wurde verbessert

Die vorläufige „Endversion“ der COMEVAL Software und des COMEVAL Handbuches wurde im Jänner 1998 fertiggestellt und konnte kommissionsintern verwendet werden.

7.1.4 Komplementarität von COMEVAL mit anderen Ansätzen und Hilfsmitteln

Für die Entwicklung von COMEVAL wurde eine große Anzahl von Vorarbeiten innerhalb der Kommission mit berücksichtigt, wie etwa bereits existierende Evaluierungsleitfäden, Studien und Handbücher. Drei damals bereits bestehende Initiativen der beteiligten Kommissionsstellen wiesen besonders starke Berührungspunkte zu COMEVAL auf:

Aufgrund der Ausrichtung von COMEVAL auf das laufende Monitoring von Ergebnissen und Effekten von Forschungsprojekten, war eine Abstimmung mit dem *Programmevaluierungsverfahren* der DG XII/A von besonderem Interesse. Das 1994 eingeführte Verfahren besteht im Grunde aus einer Anzahl von Kernindikatoren, die für alle Programme gleich sind. Die Direktion beteiligte sich aktiv an der Erstellung von COMEVAL mit der Zielsetzung, die Ergebnisse für die Weiterentwicklung des eigenen Verfahrens heranzuziehen.

Für die Abteilungen DGXII/C und DGXII/F war die Entwicklung der COMEVAL-Methodik ein Anlass, das eigene Projektmonitoringsystem weiter zu entwickeln. Dieses Monitoringsystem hat zum Ziel, die Qualität des Programmmanagements mit Hilfe von Projektqualitätsindikatoren (PQI) zu verbessern. Dieses Indikatorensystem wird von den ProjektbetreuerInnen in der Kommission dazu genutzt, die Qualität der Projekte in den verschiedenen Stadien zu beurteilen und zu verfolgen. Während der zweiten COMEVAL-Testphase entwickelte die DGXII eine überarbeitete Version der PQI, die auch Aspekte hinsichtlich der Effizienz und Qualität von Projekten zur Diskussion brachte, die ursprünglich in COMEVAL keine Berücksichtigung fanden.

Die COMEVAL-Methode und das PQI-System berücksichtigen vielfach die selben Faktoren, jedoch aus einer jeweils unterschiedlichen Perspektive. Die COMEVAL-Indikatoren bestanden in erster Linie aus Multiple Choice-Deskriptoren zur Charakterisierung von Projekthinhalten. Die Frage, die mit COMEVAL beantwortet wird ist: „Worauf zielt das Projekt ab und was sind bzw. waren die Ergebnisse?“. Die Projektqualitätsindikatoren sind stärker quantitativ orientiert und bewerten die Qualität des Projektes. Sie beantworten die Frage: „Wie hoch ist die Qualität des Projektes in bezug auf seinen methodischen Ansatz, Projektplan, Kooperationen und seine Durchführung etc.“? Die zweite Testphase bestätigte

die Ähnlichkeit der beiden Ansätze, und sowohl DG XII und DG XIII überarbeiteten in der Folge ihre Indikatorensysteme.

Ein weiterer Anknüpfungspunkt zu bereits bestehenden Aktivitäten bei der Entwicklung von COMEVAL war die PROVIEW-Initiative im IT-Programm, das durch die DGIII/F entwickelt wurde. Auf der Ebene der Programme hatte PROVIEW eine ähnliche Zielsetzung wie COMEVAL auf der Rahmenprogrammebene. Die ESPRIT ProgrammbetreuerInnen nutzen PROVIEW, um den Projekten jeweils eine Reihe von Deskriptoren zuzuordnen. Dieses Deskriptorensystem umfasste die Bereiche (a) Industriesektor, (b) sozialer und ökonomischer Nutzen, (c) strategische Ziele für die Informationsgesellschaft, (d) Projektgruppen und Handlungsansätze sowie (e) Art des Projektantrages. Hinsichtlich des sozio-ökonomischen Nutzens und der strategischen Projektziele erlaubte PROVIEW auch ein Rating verschiedener Projekte. PROVIEW setzte dieses Rating zu einigen Grundinformationen des Projektes in Beziehung (Nummer des Call, Budget, etc.) und erlaubte es, die Projekte entlang der Deskriptoren zu clustern.

Zusammen mit den oben erwähnten Kommissions-Instrumenten sollte COMEVAL zu einem nutzerInnenfreundlichen Managementtool für die Durchführung des Fünften Forschungsrahmenprogramms werden. Die Kommissionsarbeitsgruppe zur Projektevaluation und zum Monitoring wollte COMEVAL als einen Startpunkt für die Entwicklung eines integrierten Ansatzes für das Projekt-Life-Cycle-Management im Rahmenprogramm nutzen.

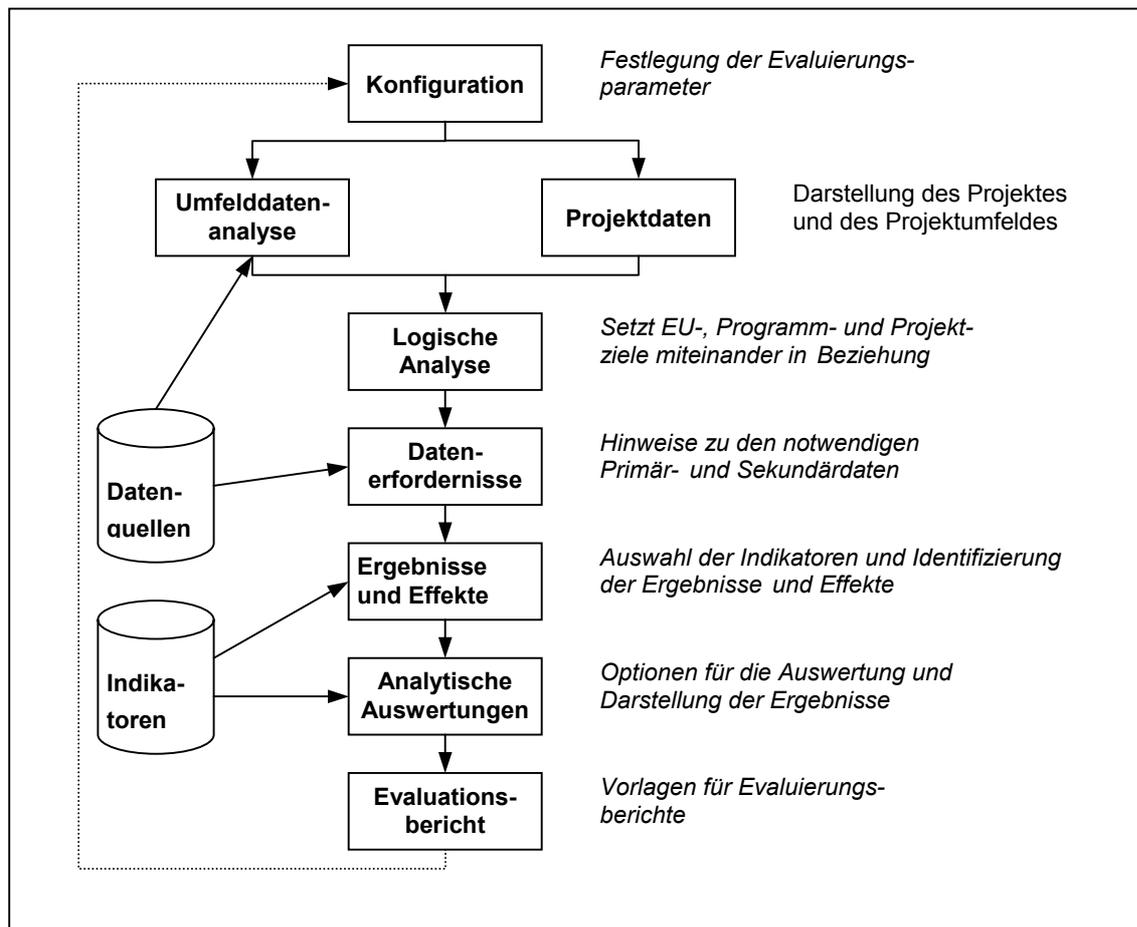
7.2 Module, Inhalte und Bewertungsverfahren des COMEVAL-Toolkits

Die hier vorgestellten Module des COMEVAL-Toolkits stützen sich auf eine der beiden von der Europäischen Kommission in Auftrag gegebenen Testversionen (Clifton et al., 1996). Der hier vorgestellte Toolkit wurde insbesondere für die Anwendung in den EU Forschungsprogrammen zu Informations- und Kommunikationstechnologien, (nicht nukleare) Energieforschung, Industrielle Technologien, Umwelttechnologien und Verkehrstechnologien entwickelt. Da es sich bei dem Toolkit um ein adaptionsfähiges Expertensystem handelt, wäre aber auch eine Anwendung auf andere Programme durchaus möglich.

7.2.1 Module des Toolkits

Der COMEVAL-Toolkit besteht aus acht verschiedenen Modulen, die von den NutzerInnen für ihre jeweiligen Bedürfnisse entsprechend konfiguriert werden können. Zusätzlich sind zwei Supportmodule vorgesehen (Indikatorendatenbank, Datenquellenmodul).

Abbildung 22



Quelle: Clifton et al., 1996

Die Nutzung der verschiedenen Module ist abhängig vom Ziel der Evaluierung, bzw. vom Ziel des Projektmonitorings. Vier hauptsächliche Anwendungsgebiete lassen sich dabei unterscheiden:

- Beurteilungsorientierte Nutzung: Hier sollen Aussagen über die Effektivität von Projekten getroffen werden, wie z.B.: „Welche Art von Projekten war hinsichtlich ihrer innovativen Ergebnisse besonders effektiv?“
- Managementorientierte Nutzung: In diesem Fall dient die Anwendung von COMEVAL dazu, den Verlauf von Projekten zu verfolgen und gegebenenfalls Interventionen zu setzen: „Ist das Projekt im Plan in bezug auf die Erreichung der angestrebten Ziele?“
- An spezifischen Zielen orientierte Nutzung: COMEVAL kann dazu verwendet werden, für eine Gruppe von Projekten aggregierte Ergebnisse bzw. Effekte zu erfassen und zu beurteilen, wie beispielsweise Antworten auf: „Welcher Anteil an den Projekten zielt auf die Schaffung eines neuen Marktes in der EU ab?“

- Auswirkungsorientierte Nutzung: Fragen, die hier beantwortet werden sollen zielen auf die Effekte der Projekte ab: „Wie viele Arbeitsplätze konnten als direkte Folge des Projektes geschaffen werden?“

In der Folge soll vor allem auf die im Toolkit verwendete Methode zur Beschreibung und Darstellung (Quantifizierung) der Ergebnisse und Effekte von FuE-Projekten Programme näher beschrieben werden.

7.2.2 Darstellung von Projektergebnissen und Effekten im COMEVAL-Toolkit

Im Modul „Ergebnisse und Effekte“ werden die Resultate und Wirkungen des Projektes auf Basis eines hierarchischen Indikatorensystems identifiziert. Das Indikatorensystem ist in der Indikatorendatenbank festgelegt und wird dort genau beschrieben. Die wesentlichsten Ergebnisse und Effekte können graphisch in Abhängigkeit von den Projektphasen (entlang der Zeitachse) dargestellt werden.

Das Modul unterscheidet auf der obersten Hierarchieebene zwischen sieben verschiedenen Ergebniskategorien (outputs):

- | | | | |
|---|--------------------|---|-----------------------------|
| 1 | Zwischenergebnisse | 5 | Normen |
| 2 | Produkte | 6 | Wissen und Fertigkeiten |
| 3 | Prozesse | 7 | Verbreitung von Ergebnissen |
| 4 | Dienstleistungen | | |

Die unten angeführte Tabelle illustriert für die Ergebniskategorie „Produkte“ das verwendete Indikatorensystem.

Tabelle 16

Ebene 1		Ebene 2		Ebene 3		Indikator
Produkte	<input type="checkbox"/>	Neue Produkte	<input type="checkbox"/>	Weltweit	<input type="checkbox"/>	Ja/nein (Anzahl)
				In Europa	<input type="checkbox"/>	Ja/nein (Anzahl)
				Im Wirtschaftszweig	<input type="checkbox"/>	Ja/nein (Anzahl)
				Im Unternehmen	<input type="checkbox"/>	Ja/nein (Anzahl)
		Verbesserte Produkte	<input type="checkbox"/>	Verbesserte Qualität	<input type="checkbox"/>	Ja/nein (% Steigerung)
				Verbesserte Leistung	<input type="checkbox"/>	Ja/nein (% Steigerung)
				Verb. Funktionalität	<input type="checkbox"/>	Ja/nein (Welche?)
				Verringerte Kosten	<input type="checkbox"/>	(% Kostenersparnis)

Quelle: Clifton et al., 1996

Bei den Effektkategorien für die Indikatoren unterscheidet das COMEVAL-Verfahren auf der obersten Hierarchieebene zwischen zehn unterschiedlichen Ausprägungen:

- | | |
|---------------------------------|---|
| 1 Wettbewerbsfähigkeit | 6 Europäische Kohäsion |
| 2 Beschäftigungseffekte | 7 Infrastruktur |
| 3 Organisatorische Innovationen | 8 Rationelle Energieerzeugung und Nutzung |
| 4 Lebensqualität | 9 Industrielle Entwicklung |
| 5 Umweltschutz | 10 Regulierung und Politik |

Für den Bereich der organisatorischen Innovationen veranschaulicht die folgende Tabelle das dreigliedrige Indikatorensystem:

Tabelle 17

Ebene 1		Ebene 2		Ebene 3		Indikator
Organisation	<input type="checkbox"/>	Schaffung eines neuen Unternehmens	<input type="checkbox"/>	Mit Projektpartnern	<input type="checkbox"/>	Ja / Nein (Warum?)
				Durch einen Partner im Konsortium	<input type="checkbox"/>	Ja / Nein (Warum?)
				Mit neuen Partnern	<input type="checkbox"/>	Ja / Nein (Warum?)
		Joint venture	<input type="checkbox"/>	Mit Projektpartnern	<input type="checkbox"/>	Ja / Nein (Warum?)
				Mit neuen Partnern	<input type="checkbox"/>	Ja / Nein (Warum?)
		Neue Netzwerke und Kontakte (Technik)	<input type="checkbox"/>	Mit Projektpartnern	<input type="checkbox"/>	Ja / Nein (Art?)
				Mit neuen Partnern	<input type="checkbox"/>	Ja / Nein (Art?)
		Neue Netzwerke und Kontakte (Markt)	<input type="checkbox"/>	Mit Projektpartnern	<input type="checkbox"/>	Ja / Nein (Art?)
				Mit neuen Partnern	<input type="checkbox"/>	Ja / Nein (Art?)
		Gesteigerte Fähigkeit Wissen aufzunehmen	<input type="checkbox"/>	Investitionen in Technologien	<input type="checkbox"/>	Höhe der Ausgaben
				Erwerb von Lizenzen	<input type="checkbox"/>	Höhe der Ausgaben
		Verbesserung der Kernkompetenzen	<input type="checkbox"/>		<input type="checkbox"/>	Anzahl der Aufträge in neuen Gebieten
		Zusätzliche FuE	<input type="checkbox"/>	EU Programme	<input type="checkbox"/>	Höhe der Ausgaben
				Eureka	<input type="checkbox"/>	Höhe der Ausgaben
				Nationale Programme	<input type="checkbox"/>	Höhe der Ausgaben
				Eigen FuE, vorkomp.	<input type="checkbox"/>	Höhe der Ausgaben
				Eigen FuE, kompetitiv	<input type="checkbox"/>	Höhe der Ausgaben
				Forschung in verwandten Gebieten	<input type="checkbox"/>	Höhe der Ausgaben
				Forschung in nicht verwandten Gebieten	<input type="checkbox"/>	Höhe der Ausgaben
		Strategische Neuausrichtung	<input type="checkbox"/>	Technologische Strategie	<input type="checkbox"/>	Veränderung des Kompetenzprofils
				Marktstrategie	<input type="checkbox"/>	Veränderung des Marktprofils
		Firmenreorganisation zur Ergebnisverwertung	<input type="checkbox"/>			Ja/nein (warum?)
		Bessere Profilierung des Unternehmens		In bereits bestehenden Sektoren	<input type="checkbox"/>	Ja/nein
				In neuen Sektoren	<input type="checkbox"/>	Ja/nein (in welchen?)

Quelle: Clifton et al., 1996

Das Modul „Analytische Auswertungen“ bietet die Möglichkeit, Analysen auf Grundlage der empirischen Daten und/oder hypothetischer Annahmen durchzuführen. Es kann geprüft werden, inwieweit die Ziele der Projekte erreicht wurden. Verschiedene Projekte können untereinander verglichen werden. Auch eine Aggregation von Daten zu Projektclustern ist hier möglich.

Die Beurteilung der Projektergebnisse auf einer Bewertungsskala ist ebenfalls ein Teil des Moduls „Analytische Auswertungen“. Die quantitative Einstufung erfolgt in COMEVAL in einer äußerst vereinfachten Form. Auf einer Skala kann der/die EvaluatorIn in der zusammenfassenden Projektbewertung beurteilen, inwieweit das Projekt bestimmten übergeordneten Zielen entsprochen hat. Dieses halbquantitative Ranking erfolgt in einer stark zusammenfassenden Form. Es stehen den EvaluatorInnen nur sechs verschiedene Beurteilungsdimensionen zur Verfügung:

- | | | | |
|---|-------------|---|---------------|
| 1 | Forschung | 4 | Infrastruktur |
| 2 | Technologie | 5 | Umwelt |
| 3 | Wirtschaft | 6 | Gesellschaft |

Im Konfigurationsmodul werden diesen Beurteilungsdimensionen bestimmte Datenerfordernisse sowie relevante Ergebnis- und Effektindikatoren zugeordnet. Beurteilungsdimensionen, die für ein bestimmtes Projekt nicht relevant sind, können bereits im Konfigurationsmodul ausgeblendet werden. Die Beurteilungsskala hat die Ausprägungen

- | | | | |
|----|------------------|---|--------------------|
| -1 | negativer Effekt | 2 | mäßiger Effekt |
| 0 | kein Effekt | 3 | großer Effekt |
| 1 | geringer Effekt | 4 | sehr großer Effekt |

Die Darstellung der Ergebnisse im Evaluierungsbericht erfolgt in Form einer Tabelle, oder alternativ dazu in Form eines Projektbewertungsradars.

Tabelle 18

	Selbstbewertung	Evaluator 1	Evaluator 2	Anmerkungen
Forschung				
Technologie				
Wirtschaft				
Infrastruktur				
Umwelt				
Gesellschaft				
Gesamt				

Quelle: Clifton et al., 1996

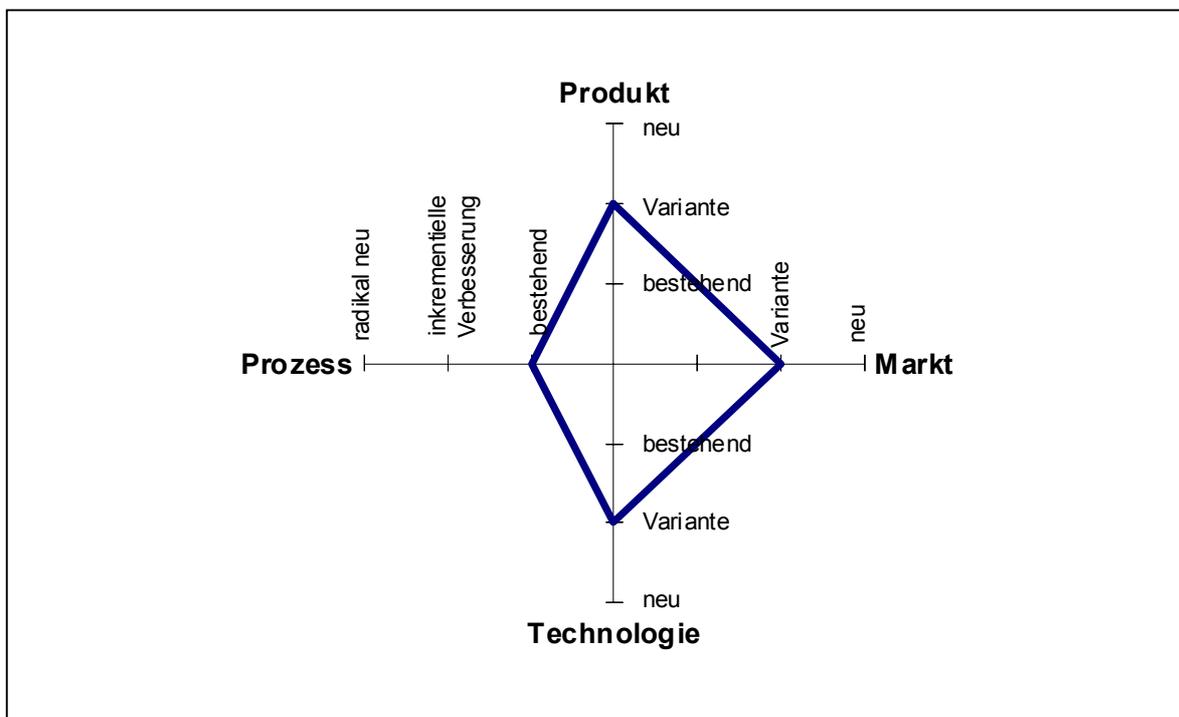
7.2.3 Der Evaluierungsbericht

Der Standard-Evaluierungsbericht ist in sechs Teile gegliedert:

- Grundinformationen zum Projekt
- Projektannahmen
- Projektziele und Zielerreichung
- Projektergebnisse und Effekte
- Projektbewertung
- Gesammelte Erfahrungen

Der erste Teil des Evaluierungsberichtes umfasst die wesentlichen Grundinformationen und beinhaltet eine graphische Darstellung der Positionierung des Projektes hinsichtlich innovativer Aspekte.

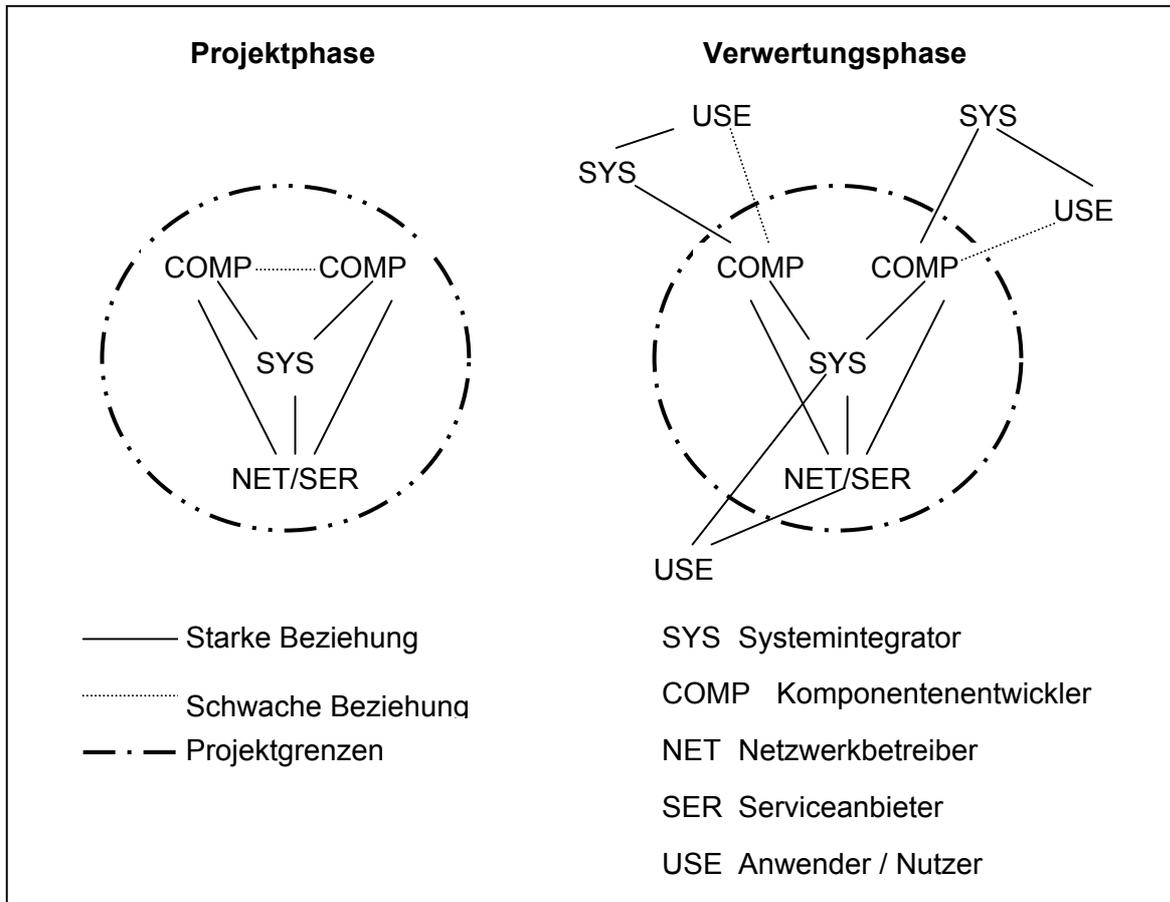
Abbildung 23



Strategische Positionierung eines Projektes hinsichtlich Technologie und Innovationsaspekten (Quelle: Clifton et al., 1996)

Zusätzlich kann mit diesem Modul das Projekt Netzwerk in verschiedenen Projektstadien dargestellt werden.

Abbildung 24



Beschreibung des Projektnetzwerks nach COMEVAL (Quelle: Clifton et al., 1996)

Teil zwei des Berichtes stellt die dem Projekt zugrundeliegenden Annahmen dar und beschreibt die Auswirkungen potentieller Änderungen dieser Annahmen auf das Projekt. Teil drei befasst sich mit den technologischen, wirtschaftlichen und sonstigen Projektzielen, dokumentiert Änderungen der Ziele während der Projektlaufzeit und gibt Aufschluss darüber, inwieweit die Projektziele tatsächlich erreicht wurden. Teil vier beschreibt die Projektergebnisse und -effekte und ordnet sie den einzelnen Projektphasen zu (jeweils aggregiert auf der ersten und zweiten Ebene der Ergebnis- und Effektindikatoren, siehe Beispiel unten).

Abbildung 25

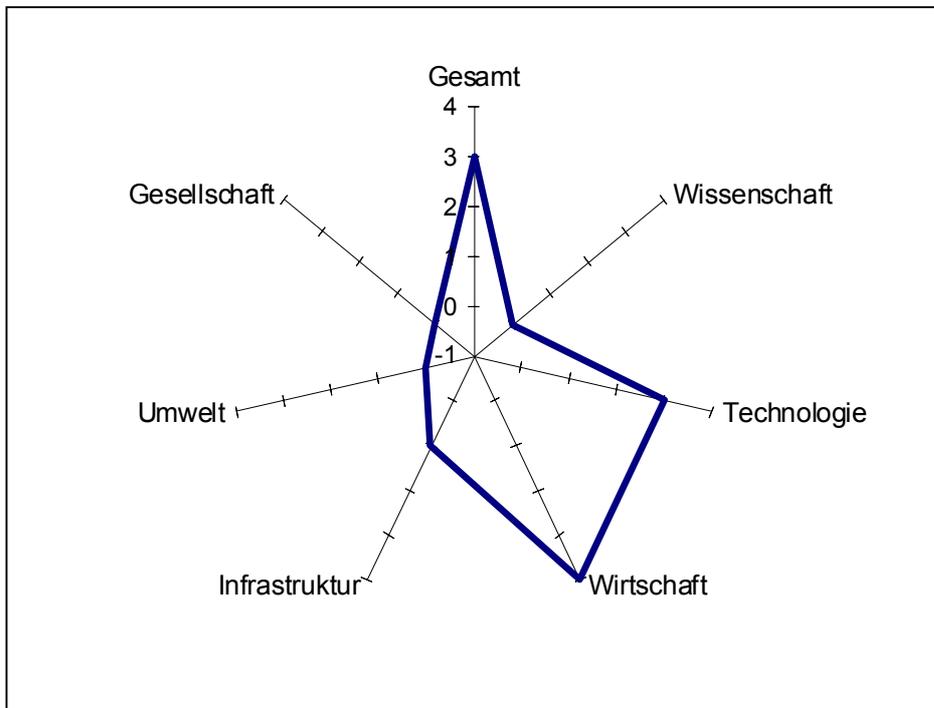
Was sind die Ergebnisse des Projektes und wann wurden diese erreicht?		
Ergebnisse Ebene 1	Ergebnisse Ebene 2	Zeitverlauf $t_{\text{Beginn}} \rightarrow t_{\text{Ende}} \rightarrow t_{\text{Evaluation}}$
Zwischenergebnisse	Prototyp	❖ ■
	Geistiges Eigentum	❖ ■
	Demonstration	❖ ■
Produkte	Verbessertes Produkt	❖ ■
Normen	De jure Norm	❖ ■
Was sind die Ergebnisse des Projektes und wann wurden diese erreicht?		
Ergebnisse Ebene 1	Ergebnisse Ebene 2	Zeitverlauf $t_{\text{Beginn}} \rightarrow t_{\text{Ende}} \rightarrow t_{\text{Evaluation}}$
Wettbewerbsfähigkeit	Erhöhter Umsatz	❖ ■
	Lizenzgebühren	❖ ■
Beschäftigung	Jobs geschaffen	■ ❖
Infrastruktur	Telekom	❖ ■

❖ (geplant) ■ (erreicht) ! (unerwartetes Ergebnis)

Ergebnis- und Auswirkungsdatenblatt des COMEVAL-Berichtes
(Quelle: Clifton et al., 1996)

Im darauf folgenden Abschnitt des Berichtes befindet sich das bereits oben angeführte zusammenfassende Projektbewertungsblatt, das auch in Form eines Radardiagramms dargestellt werden kann.

Abbildung 26



COMEVAL-Projektbewertungsradar (Clifton et al., 1996)

Teil sechs dokumentiert die gesammelten Erfahrungen im Projekt, die in der Zukunft für andere Projekte und Programme genutzt werden sollen.

7.3 Erfahrungen und mögliche Anwendungsgebiete

7.3.1 Anwendung des COMEVAL-Toolkits in der Europäischen Kommission

Die COMEVAL-Methode wurde ursprünglich dafür entwickelt, die Ergebnisse und Effekte von Forschungsprojekten im Fünften Rahmenprogramm zu bewerten. Die Implementierung des Toolkits ging jedoch nie über die Pilotphase hinaus. Laut Informationen aus der Programmevaluierungsgruppe in der Generaldirektion Forschung ist auch nicht daran gedacht, COMEVAL als geschlossene Methode für das Monitoring im Fünften Rahmenprogramm zu verwenden (Dumont, 2000). Nach Einschätzung eines Kommissionsmitarbeiters der maßgeblich an der Entwicklung des Toolkits beteiligt war, hat die Generaldirektion Forschung jedoch zentrale Teile von COMEVAL in das Projektmonitoringverfahren des Rahmenprogramms integriert (Löwe, 2000).

Für das Programmmonitoring von FuE-Programmen wurde in der Kommission das System der Projektqualitätsindikatoren beibehalten. Für das Fünfte Rahmenprogramm sieht dieses Modell insgesamt 15 Indikatoren vor, die in fünf Gruppen zusammengefasst sind: Wissenschaftlich-technologische Qualität und Innovation; Zusätzlicher Wert für die

Gemeinschaft und Beitrag zur Gemeinschaftspolitik; Beitrag zu den gesellschaftlichen Zielen der Gemeinschaft; Wirtschaftliche Entwicklung und Perspektiven für Wissenschaft und Technologie; Ressourcen, Partnerschaften und Management. Dieses Indikatorensystem wird heute von der Projektantragsphase (ex-ante Evaluierung) über das Monitoring bis zur ex-post Programmevaluierung verwendet. Besonders hinsichtlich der zusammenfassenden Darstellung der Projekteffekte hat das PQI-System zahlreiche Berührungspunkte zum COMEVAL-Beurteilungsschema.

7.3.2 Anwendungsmöglichkeiten in Österreich

Der COMEVAL-Toolkit ist weniger eine quantitative Methode der Evaluierung, als vielmehr ein ExpertInnensystem für die Projektverfolgung und das Projektmonitoring. Aufgrund der bei der Europäischen Kommission gemachten Erfahrungen (Probleme bezüglich Komplexität und Zeitaufwand) wäre vor der Anwendung eines derartigen Toolkits jedenfalls eine Vereinfachung und Anpassung an die spezifischen Anwendungsanforderungen notwendig. Ein Zeitbedarf für eine Standardevaluierung von rund zweieinhalb Tagen je Projekt (Clifton et al., 1996) erscheint in der Praxis für viele Programme unrealistisch hoch zu sein, insbesondere wenn angenommen wird, dass der Evaluierungstoolkit während der Projektlaufzeit mehrmals zur Anwendung kommen soll.

Die Entwicklung eines einfachen COMEVAL-ähnlichen Evaluierungs- und Monitoringtoolkits in Österreich wäre vor allem für eine Anwendung beim Projektmonitoring des Forschungsförderungsfonds für die gewerbliche Wirtschaft denkbar. Hier würden die Struktur des Toolkits die Verfolgung von Projekten über die Laufzeit entscheidend verbessern helfen und damit mehr Transparenz hinsichtlich der Ergebnisse und Wirkungen der geförderten Projekte schaffen. Die dabei gemachten Erfahrungen könnten für die Begleitung anderer Projekte sinnvoll genutzt werden.

7.4 Literatur

Dumont, Yves (2000): e-mail-Korrespondenz, 3. Februar 2000.

Löwe, Peter (2000): e-mail-Korrespondenz, 4. Februar 2000.

Georghiou, Luke (2000): e-mail-Korrespondenz, 1. Februar 2000.

Clifton, Trevor, London, Lee, Cunnigham, Paul, Georghiou, Luke and Howells, Jeremy (1996): Common Methodology for the Evaluation of RTD Project Results: Toolkit User Guide. Unveröffentlichter Bericht im Auftrag der Europäischen Kommission, Generaldirektion XIII/D.

EC-European Commission (1998): The „Common Methodology for the Evaluation of RTD Results“ (COMEVAL). Internes, unveröffentlichtes Dokument, März 1998.

EUSTAT (1998): Europäischer Plan für Forschungstätigkeit im Bereich der amtliches Statistik (EPROS). Fünftes FTE-Rahmenprogramm 1999-2003, Entwurf (Luxemburg).

TEIL B

8 EVALUIERUNG STEUERLICHER ANREIZE FÜR FORSCHUNG UND ENTWICKLUNG

Gernot Hutschenreiter
WIFO

8.1 Steuerliche Anreize für F&E

8.2 Vor- und Nachteile steuerlicher Anreize

Die entwickelten Volkswirtschaften setzten bei der Förderung von Forschung und Entwicklung sowohl direkte als auch indirekte (steuerliche) Förderungsinstrumente ein. Als potentielle Vorteile der indirekten gegenüber der direkten F&E-Förderung (siehe etwa *Leyden – Link, 1992, Harhoff, 1994, Link, 1996, OECD, 1998, Hall – van Reenen, 2000*) werden vor allem die folgenden genannt:

- Die indirekte F&E-Förderung ist durch ein hohes Maß an "Neutralität" in Bezug auf unternehmerische Allokationsentscheidungen gekennzeichnet. Dies betrifft vor allem den Inhalt und Charakter der F&E-Projekte, aber auch die Kombination der Produktionsfaktoren. Eine generelle steuerliche F&E-Förderung ist im Prinzip auch neutral gegenüber der sektoralen Zugehörigkeit der Unternehmen. "Im Prinzip" deshalb, weil die F&E-Intensität der Unternehmen nach ihrer Branchenzugehörigkeit beträchtlich und systematisch variiert.
- Steuerliche Förderungen weisen eine relativ niedrige Zutrittschwelle auf, was insbesondere für kleine und mittelgroße Unternehmen (KMU) von Vorteil sein sollte.
- Die Förderungsbedingungen sind vergleichsweise transparent.
- Indirekte Förderungen sind für die Unternehmen planbar. Sie unterscheiden sich diesbezüglich von der direkten (diskretionären) Projektförderung, es sei denn letztere ist nicht selektiv.
- Die Administrations- und "Compliance"-Kosten können vergleichsweise gering gehalten werden. Das heißt es entsteht kein großer Aufwand für die Programmplanung, das Programmmanagement etc. und die Unternehmen werden weniger beansprucht.
- Indirekte Förderungen verursachen in weit geringerem Ausmaß "Rent seeking"-Aktivitäten als gezielte (diskretionäre) direkte Förderungsprogrammen (siehe etwa *Cohen – Noll, 1991*).

Es ist festzuhalten, dass diese Vorteile nur potentiell gegeben sind. Ob sie auch tatsächlich realisiert werden, hängt von der konkreten Ausgestaltung des Förderungsinstrumentes und seiner administrativen Handhabung ab. Die Neutralität der Förderung ist der "harte Kern" der ökonomischen Argumente für die steuerliche Förderung von F&E.

Den Vorteilen der indirekten F&E-Förderung werden häufig die folgenden potentiellen Nachteile gegenüber Instrumenten der direkten Förderung gegenübergestellt:

- Mitnahmeeffekte bzw. ein hohes Maß an Unsicherheit über die tatsächlichen Anreizwirkungen.

- Steuerliche Anreize für private F&E begünstigen von ihrer Konstruktion her eher F&E-Aktivitäten mit hohen privaten Erträgen und nicht notwendigerweise solche mit hohen sozialen Erträgen.
- "Inequity", d. h. Ungleichheit in der Beanspruchung der Begünstigung. In der Praxis ist – aus verschiedenen Gründen – ein hohes Maß an Heterogenität der effektiven "F&E-Preise" festzustellen (*Hall – van Reenen*, 2000).
- Die Anreize indirekter F&E-Förderung variieren über den Produktlebenszyklus. Dieser Umstand widerspricht zwar dem Postulat der Neutralität der Förderung, muss aber nicht notwendigerweise von Nachteil sein.
- Der durch die indirekte Förderung induzierte Ausfall von Steuereinnahmen führt ceteris paribus zu höheren Steuersätzen bzw. in der Praxis auch zu einer die Ressourcenallokation verzerrenden kompensatorischen Besteuerung.
- Steuerliche Anreize unterlaufen die Budgetkontrolle. Direkte Förderungsprogramme werden in der Regel mit fixen Mitteln dotiert. Steuerausfälle werden weniger beachtet als in den öffentlichen Haushalten ausgewiesene direkte Förderungen.

Eine ausführlichere Abwägung der zitierten Argumente für und wider steuerliche F&E-Anreize findet sich in *Leyden – Link* (1992), *Link* (1996), *OECD* (1998), *Hall – van Reenen* (2000).

Gestützt auf ökonometrische Analysen auf der makroökonomischen Ebene gelangt die *OECD* (1998, S. 178) zu der Feststellung, dass sowohl direkte Förderungen als auch Steuerbegünstigungen eine signifikante, jedoch eher geringe positive Wirkung auf private F&E-Investitionen ausüben. Ein Unterschied zwischen den beiden Förderungstypen wird eher in den Zielen und Anreizmechanismen gesehen: "Tax concessions encourage applied research with sufficient private return. Direct subsidies are granted to projects selected by government, taking a longer-term view on their social return."

Aus den genannten Vor- und Nachteilen steuerlicher Anreize für F&E lassen sich grundsätzliche Schlussfolgerungen hinsichtlich des Designs derartiger Instrumente ziehen. Mithilfe indirekter F&E-Förderungsinstrumente lassen sich – wie auch die Praxis in den OECD-Ländern (*OECD*, 1998) zeigt – durchaus spezifischere Ziele als die allgemeine Stimulierung von F&E-Aktivitäten in der Wirtschaft verfolgen (wie z.B. die Stimulierung von technologieorientierten Start-up-Unternehmen). Allerdings sollten die Ziele, die mit einem steuerlichen Anreizinstrument verfolgt werden, aus grundsätzlichen Überlegungen heraus nicht zu spezifisch sein. Aus ökonomischer Sicht besteht der wichtigste Grund dafür darin, dass eine zu spezifische Auslegung die potentiellen Vorteile steuerlicher Instrumente der F&E-Förderung (Neutralität hinsichtlich unternehmerische Allokationsentscheidungen, niedrige Zutrittsschwelle, geringe Administrationskosten etc.) unterläuft. Zielgerichtete technologiepolitische Interventionen sollten prinzipiell eher der direkten Förderung vorbehalten bleiben.

Über diese theoretischen Überlegungen hinaus hat der Grad der "Spezifität" und "Selektivität" in der EU auch beihilfenrechtliche Implikationen. Die *Europäische Kommission* (1998) hat klargestellt dass "Maßnahmen, die ein Ziel der allgemeinen Wirtschaftspolitik verfolgen, indem die mit bestimmten Produktionskosten verbundene Steuerbelastung reduziert wird (z.B. Forschung und Entwicklung, Umweltschutz, Ausbildung, Beschäftigung)", keine Beihilfen im Sinn des Artikel 92 EG-Vertrag sind, vorausgesetzt sie gelten gleichermaßen für alle Unternehmen und Produktionszweige im Gebiet eines

Mitgliedsstaats¹⁴). Spezifische oder selektive steuerliche Maßnahmen hingegen unterliegen in vollem Umfang den beihilfenrechtlichen Bestimmungen der EU.

8.3 Steuerliche Anreize für F&E: Ein internationaler Überblick

Eine ganze Reihe von OECD-Ländern verfügt über Instrumente der steuerlichen Förderung von Forschung und Entwicklung. Diese wurden in der überwiegenden Mehrzahl seit Beginn der achtziger Jahre eingeführt. Dabei kommen folgende Instrumente zum Einsatz (siehe *OECD*, 1998, S. 168):

- Volle Geltendmachung der laufenden F&E-Aufwendungen im Jahr ihrer Verausgabung. Dies stellt insofern eine Begünstigung dar als F&E-Aufwendungen eigentlich Investitionen darstellen; das F&E-Kapital generiert über einen längeren Zeitraum hinweg Einkommen.
- Vorzeitige Abschreibung von Maschinen und Ausrüstungen für F&E sowie, in manchen Fällen, von Gebäuden.
- Zusätzlich Freibeträge, die es den Begünstigten erlauben, mehr als 100 Prozent der F&E-Ausgaben von ihrem steuerbaren Einkommen abzuziehen.
- Steuergutschriften ("R&D tax credits"), die es den Unternehmen gestatten, einen gewissen Prozentsatz ihrer F&E-Aufwendungen unmittelbar von ihrer Steuerschuld abzuziehen. Steuergutschriften werden in zwei Formen gewährt: Bei "volumenbasierten" ("volume-based") tax credits sind die gesamten F&E-Aufwendungen förderbar, bei inkrementellen ("incremental") tax credits nur die zusätzlich getätigten, also die marginalen F&E-Aufwendungen.

Eine Übersicht der *OECD* (1998, S. 168ff) fasst die im Jahr 1996 herrschende Praxis der steuerlichen F&E-Förderung in 25 OECD-Ländern zusammen:

- Die Steuersysteme aller OECD-Länder (mit Ausnahme Neuseelands) sehen vor, dass laufende F&E-Aufwendungen noch im selben Jahr geltend gemacht werden, in denen sie getätigt werden. Sieben Länder wenden diese Begünstigung auch auf andere als laufende F&E-Aufwendungen an.
- Zwei Länder, nämlich Australien und Österreich räumen Freibeträge im Ausmaß von 125 Prozent (bis 1996 150%) im Fall der australischen "R&D Tax Concession" bzw. (vor dem Steuerreformgesetz 2000) bis zu 112% bzw. bis zu 118% Prozent im Fall des österreichischen Forschungsfreibetrages (nach der Steuerreform 2000: bis zu 125% bzw. bis zu 135% für zusätzliche Forschungsaufwendungen) ein.
- In acht Ländern existieren Steuergutschriften für F&E: Volumenbasierte Steuergutschriften (tax credits) in Italien, Kanada und den Niederlanden, inkrementelle Steuergutschriften in Frankreich und in den USA. In Japan, Korea und Spanien existieren beide Formen nebeneinander.

¹⁴) Eine Maßnahme ist als Beihilfe im Sinne des Artikel 92 EG-Vertrag anzusehen, wenn sie sämtliche der folgenden Voraussetzungen erfüllt (*Europäische Kommission*, 1998):

1. Die Maßnahme muß dem Begünstigten einen Vorteil verschaffen, durch den seine normalerweise zu tragenden Belastungen vermindert werden.
2. Der Vorteil muß vom Staat oder aus staatlichen Mitteln gewährt werden.
3. Die betreffende Maßnahme muß den Wettbewerb und den Handel zwischen Mitgliedsstaaten beeinträchtigen.
4. Die betreffende Maßnahme muß spezifisch oder selektiv sein, also "bestimmte Unternehmen und Produktionszweige" begünstigen.

- Sechs Länder sehen besondere steuerliche Begünstigungen für Forschung durch kleinere Unternehmen vor.

Einen etwas über den Kreis der OECD-Länder hinausreichenden Überblick (einschließlich Brasilien, China, Indien, Singapur, Südafrika und Taiwan) bieten *Hall – van Reenen* (2000). Unterschiede zwischen den Ländern bestehen u.a. hinsichtlich:

- Der Definition der begünstigten F&E-Aufwendungen.
- Des fiskalischen Status der Steuervorteile.
- Der Plafondierung der steuerlichen Begünstigung.
- Der Behandlung von Firmen, die Verluste machen.
- Der Behandlung von kleinen und mittelgroßen Unternehmen (KMU).

Es ist festzuhalten, dass die seit Inkrafttreten der Steuerreform 2000 in Österreich praktizierte Form der steuerlichen Begünstigung von F&E in ihrer konkreten Form international einzigartig ist.

8.4 Evaluierungen steuerlicher Anreize für F&E

8.4.1 Evaluierungskriterien

In der Wirkungsanalyse von F&E-Förderungen kommen vor allem die folgenden Evaluationskriterien zur Anwendung:

"Additionalität": In diesem Zusammenhang wird versucht abzuschätzen, in welchem Umfang eine Reduktion der F&E-Kosten zusätzliche F&E-Ausgaben der Unternehmen induziert. Technisch gesprochen hängt die Antwort auf diese Frage von der zu schätzenden Preiselastizität der Nachfrage nach F&E-Leistungen ab.

Maximierung des "sozialen Nutzens": Streng genommen ist das Vorliegen von "Additionalität" keine Rechtfertigung für eine staatliche Intervention in Form von Förderungen. Die (neo-) "klassische" theoretische Begründung für die Förderung privater F&E-Investitionen besteht darin, dass die gesamtwirtschaftlichen ("sozialen") Erträge von Investitionen von F&E die privat aneignbaren Erträge übersteigen (siehe etwa *Hutschenreiter*, 1995). Optimalerweise sollte daher jene Förderungshöhe bestimmt werden, die die Kluft zwischen der privaten und sozialen Ertragsrate von F&E-Ausgaben schließt.

Ein Überblick über Evaluationskriterien sowie kritischer Aspekte des Designs von F&E-Förderungsinstrumenten findet sich in *OECD* (1998, S. 174ff). Siehe auch *Lhuillery* (1996).

8.4.2 Quantitative Methoden der Evaluierung

In der Praxis ist eine hinreichend exakte Beobachtung der Ertragsverläufe zur Bestimmung der optimalen Förderungshöhe nicht möglich. Die einschlägigen ökonometrischen Studien beschränken sich daher häufig auf eine Art reduzierte Kosten-Nutzen-Analyse, wobei die geschätzten, durch die indirekte Förderung zusätzlich induzierten F&E-Ausgaben den damit verbundenen Steuerausfällen gegenübergestellt werden. Diese Gegenüberstellung ist zwar intuitiv einleuchtend und wohl auch für die Förderungspolitik eine sinnvolle erste Annäherung an das Problem, aus theoretischer Sicht ist sie jedoch nicht zufriedenstellend. Darüber

hinaus gibt es auch vollständigere Kosten-Nutzen-Analysen, die jedoch aufwendig sind und weniger häufig angewandt werden (Beispiel Australien).

8.4.3 Empirische Evidenz

Die umfangreichste empirische Evidenz zur steuerlichen F&E-Förderung liegt für den "R&D Tax Credit" in den USA vor. Dieser wurde 1981 eingeführt ("Economic Recovery Tax Act") und war zunächst bis Ende 1985 befristet. In der zweiten Hälfte der achtziger Jahre und in den neunziger Jahren wurde die Laufzeit mehrfach verlängert, wobei auch Änderungen in den Modalitäten vorgenommen wurden. Siehe dazu den Überblick bei Hall (1992). Der "R&D Tax Credit" wurde ermittelt, indem die F&E-Ausgaben in der jeweils gültigen Definition, die einen bestimmten Basiswert überschreiten, mit dem gültigen Satz ("credit rate") multipliziert wurde. Der sich daraus ergebende Betrag wurde von der Körperschaftssteuer ("corporate income tax") abgezogen. Es bestand eine dreijährige Rücktrags- und 15jährige Vortragsmöglichkeit, falls im laufenden Jahr kein steuerbares Einkommen vorhanden war.

Ein bemerkenswertes Faktum stellt der Umstand dar, dass – obwohl sich der gesetzlich fixierte ("statutarische") Satz für den "R&D Tax Credit" zumeist in der Höhe von 20-25% bewegte – der ermittelte effektive Satz sehr viel niedriger lag, nämlich bei rund 4% oder etwas darüber. Dies hatte eine Reihe von Ursachen: Erstens erhöhten zusätzliche F&E-Ausgaben im laufenden Jahr die Basis zur Berechnung der steuerbegünstigten zusätzlichen F&E-Ausgaben in den drei Folgejahren. Zweitens gab es eine Obergrenze für die zusätzlichen F&E-Ausgaben. Darüber hinaus gibt es komplexe Wechselwirkungen zwischen dem Steuersystem der USA und anderen nationalen Steuersystemen.

Frühe Studien (bis etwa zur Mitte der achtziger Jahre) beurteilten die Anreizwirkung indirekter F&E-Förderungen eher skeptisch. Typischerweise bewegt sich die in diesen Studien geschätzte Preiselastizität von F&E im Intervall -0,3 bis -0,4. Dies besagt, dass eine Reduktion der F&E-Kosten um ein Prozent eine Erhöhung der F&E-Ausgaben im Ausmaß von 0,3-0,4% zur Folge hat (siehe etwa den Überblick bei Cordes, 1989, Hall, 1992, Harhoff, 1994, Hall – van Reenen, 2000).

Untersuchungen aus den neunziger Jahren gelangen zu deutlich höheren Schätzungen für die Preiselastizität von F&E, darunter auch die bekannte Studie von Hall (1992). Die Untersuchung von Hall ist ein technisch anspruchsvoller Versuch, die Wirkungen eines Instruments der steuerlichen F&E-Förderung, des "R&D Tax Credit" in den USA, auf die F&E-Ausgaben mit ökonometrischen Methoden zu quantifizieren. Den analytischen Rahmen für die Analyse von Hall bildet ein dynamisches Investitionsmodell mit Anpassungskosten. Diese Anpassungskosten werden als proportional zur Intensität der F&E-Investitionen angenommen. Diese werden minimiert, wenn die F&E-Investitionen den Ersatzinvestitionen entsprechen.

Die empirische Grundlage der Untersuchung besteht aus einem Sample von Unternehmen des Manufacturing-Bereichs (rund 1.000 pro Jahr) im Zeitraum 1980-1991. Die darin enthaltenen Unternehmen tätigen etwa 85% der F&E-Ausgaben der Industrie. Ein – durch Rückgriff auf ältere Schätzungen bzw. per Annahmen behelfsmäßig gelöstes – Problem der Studie besteht darin, dass weder der Anteil der steuerlich qualifizierten F&E-Ausgaben an den gesamten F&E-Ausgaben der Unternehmen bekannt noch Information über den exakten steuerlichen Status der Unternehmen verfügbar war.

Auf Basis ökonometrischer Schätzungen kommt Hall zu dem Ergebnis, dass die Preiselastizität der F&E-Ausgaben in der Nähe von 1 liegt. Diese – im Vergleich zu früheren

Studien – relativ starke Reaktion der F&E-Ausgaben auf Preisänderungen impliziert, dass die durch den "R&D Tax Credit" zusätzlich induzierten F&E-Ausgaben höher sind als die damit verbundenen Einnahmefälle. Weiters legen die Ergebnisse von Hall den Schluss nahe, dass die (allerdings kaum präzise schätzbaren) langfristigen Effekte des "R&D Tax Credit" relativ groß sind. Aus der verzögerten Anpassung an Preisänderungen kann umgekehrt die Empfehlung abgeleitet werden, mit kurzfristigen Adjustierungen bzw. kurzfristigen Laufzeiten der steuerlichen Förderung – im Gegensatz zur tatsächlichen Geschichte des "R&D Tax Credit" in den USA – vorsichtig umgegangen werden sollte.

8.5 Best Practices der Evaluierung

Eine Reihe von OECD-Ländern haben bereits seit längerer Zeit Erfahrung mit steuerlichen Anreizen zur Stimulierung von F&E. Im Lauf der Zeit haben sie Erkenntnisse über die Wirksamkeit dieser Instrumente gesammelt. Die in diesem Zusammenhang vorgenommenen Evaluierungsanstrengungen der Länder unterscheiden sich in Umfang und Tiefe. Australien etwa blickt auf mehr als ein Jahrzehnt intensiver Evaluierungen in Zusammenhang mit der australischen "R&D Tax Concession" zurück, an denen mehrere Forschungsinstitutionen beteiligt waren. Australien hat das steuerliche Anreizinstrumentarium im Licht dieser Erfahrungen reformiert. Die niederländische Evaluierung hingegen kann als "Pilotprojekt" zur Bewertung der Effektivität eines relativ jungen Anreizinstrumentes betrachtet werden.

8.5.1 Australien: Kosten-Nutzen-Analyse

In Australien gab es umfassende und nachhaltige Bemühungen, die Wirkung der steuerlichen Anreize für F&E zu bewerten. Dieser Evaluierungsschwerpunkt ist nicht zufällig. Im Jahr 1985/86 führte Australien eine der großzügigsten indirekten Förderungen im OECD-Raum ein, die "R&D Tax Concession". Ursprünglich gestattete die Tax Concession Firmen, eine Reihe F&E-bezogener Ausgaben im Umfang von bis zu 150% geltend zu machen. Dieser Satz wurde später (gestützt auf Erkenntnisse der Evaluierung) auf 125% reduziert. Weitere Modifikationen betrafen die Zulassung von Firmen-Syndikaten im Jahr 1987 (und deren Abschaffung im Jahr 1996). Nach *Lattimore* (1997) wurde 1994/95 die allgemeine (150%) Tax Concession von ungefähr 3.500 Unternehmen genutzt, die ungefähr \$ 550 Millionen an F&E-Förderungen beanspruchten. Das Syndikats-Programm umfasste lediglich 103 Nutzer, die Subventionen erreichten jedoch ein Volumen von \$ 255 Millionen. Im Vergleich zu diesen beiden Programmen waren alle anderen öffentlichen F&E-Förderungsaktionen in Australien verschwindend klein.

Als Hauptziel der "R&D Tax Concession" wurde die Schaffung von Anreizen für höhere F&E-Ausgaben im Unternehmenssektor postuliert, der Zielkatalog ist jedoch weiter ausdifferenziert, aber klar nachvollziehbar. Auf diese Ziele nehmen die Terms of Reference der Evaluierung durch das Australian Bureau of Industry Economics explizit Bezug (siehe Textkasten).

In den neunziger Jahren wurden mehrere unabhängige Evaluierungen der Tax Concession vorgenommen. Dabei wurden qualitative und quantitative Indikatoren ihrer Wirkung und des privaten und sozial Nutzens induzierter F&E verwendet. Von besonderem Interesse ist, dass dabei auch die soziale Kosten-Nutzen-Analyse als konzeptioneller Rahmen verwendet wurde. In diesem Rahmen wurde der soziale Netto-Nutzen der Förderung definiert als Summe des induzierten privaten Nutzens plus der nicht durch den privaten Investor

angeeigneten Spillovers minus marginale "excess burden of taxation", Abfluss an ausländische Shareholders, staatliche Administrationskosten, Compliance-Kosten der Unternehmen und die Kosten des "Rent seeking" (Lobbying für zusätzliche Begünstigungen oder zur Verhinderung der Erosion bestehender Begünstigungen).

Die in Australien durchgeführten Evaluationsstudien bieten wertvolle Information hinsichtlich der Komponenten des sozialen Netto-Nutzens. Eines der wichtigsten Elemente in diesem Zusammenhang ist die empirische Forschung über induzierte F&E-Investitionen, d.h. der Sensitivität der F&E-Ausgaben auf preisliche Anreize (Kostensenkung durch die steuerliche Begünstigung). Als Maß der Induzierung wird dabei der Anteil der politikinduzierten laufenden F&E-Aktivität an der gesamten laufenden F&E-Aktivität verwendet. Die australischen Studien verwenden sowohl ökonometrische Methoden als auch subjektive Einschätzungen der Wirkung der R&D Tax Concession (BIE, 1993, Lattimore, 1997).

Tabelle 19

<p>Terms of Reference der Evaluierung</p> <ol style="list-style-type: none">1. Assess the effectiveness of the 150 per cent R&D tax concession against its stated objectives: To make Australian companies more innovative and internationally competitive through:<ol style="list-style-type: none">a. increasing companies' investment in R&D;b. Encouraging better use of Australia's existing research infrastructure;c. improving conditions for the commercialisation of new process and product technologies developed by Australian companies; andc. developing a greater capacity for the adoption of foreign technology.2. Analyse the intended and unintended impacts of the concession on Australian industry;3. Examine the overall relevance of program objectives within the context of industry, science and technology policy objectives;4. Assess the role of the concession in relation to other government programs aimed at promoting the contribution of R&D to create internationally competitive industries; and5. Recommend whether the tax concession should continue, and if so, suggest changes to improve its effectiveness and efficiency.

Quelle: BIE (1993).

Nach Lattimore (1997) war die geschätzte "social rate of return" der 125% Tax Concession deutlich positiv. Interessanterweise übersteigt sie die rate of return der 150% Concession. Dies ist darauf zurückzuführen, dass der "bang for a buck" (d.h. das Verhältnis der Induktion zur Subvention) im Fall der 125% Tax Concession höher ist. Dies weist darauf hin, dass die marginale Induktion mit dem Subventionsgrad abnimmt. Die Ergebnisse der Kosten-Nutzen-Analyse weisen auch auf Schwächen der Tax Concession hin, z.B. auf eine insgesamt relativ geringe Induktion zusätzlicher F&E-Investitionen. Die Klarheit des analytischen Rahmens

erleichtert die Ableitung von Optionen für das Re-Design der steuerlichen Anreize. Insbesondere zielen diese Vorschläge auf eine Erhöhung der induzierten F&E ab. Die konzeptionelle Klarheit und der empirische Gehalt machen die australischen Untersuchungen besonders attraktiv und können daher als wertvolle Referenz für künftige Arbeiten in Österreich dienen.

8.5.2 Niederlande: Ein pragmatischer Mix

In den Niederlanden wurde 1994 die "Wage Tax and Social Insurance Allowance Act (WVA/S&O)" eingeführt. Das niederländische Schema sieht eine Reduktion der Lohnsteuer und Sozialversicherungsbeiträge, die vom Arbeitgeber zu tragen sind, vor.

Für die ersten NLG 150.000 an F&E-Personalkosten können 40 Prozent geltend gemacht werden, für die diesen Betrag überschreitenden hinausgehenden F&E-Personalkosten 17,5 Prozent. Darüber hinaus gibt es eine Plafondierung des Absetzbetrages je Einheit bei 15 Mill. NLG. KMU werden durch eine degressive Gestaltung der steuerlichen Begünstigung bevorzugt. Die Grenze der ersten Stufe (NLG 150.000) erscheint relativ niedrig gegriffen, die Obergrenze von NLG 15 Mill. hingegen vergleichsweise hoch. Projektanträge werden – vorab – durch Senter, eine Agentur des Ministeriums für wirtschaftliche Angelegenheiten (und nicht ex post durch die Steuerbehörden) evaluiert.

Wichtiger noch ist der Umstand, dass die Konstruktion des niederländischen Modells – die genannten Prozentsätze der F&E-Personalkosten werden von den Lohnnebenkosten abgezogen – den steuerlichen Anreiz gewinnunabhängig macht. Dies bedeutet, dass die Probleme von Start-up-Unternehmen oder auch der prozyklischen Wirkung vermieden werden. Überdies lässt diese flexible Konstruktion auch die Einbeziehung von nicht-gewinnorientierten Forschungseinrichtungen in den Kreis der Begünstigten zu (wie dies in den Niederlanden tatsächlich der Fall ist).

Die Evaluation (siehe *Dorsman, 1997* und *Dorsman, 1998*) befasst sich mit drei separaten Fragestellungen:

- die Wirkung der Steuerbegünstigung auf die F&E-Anstrengungen,
- die Wirkung der F&E-Anstrengungen auf die Innovationsfähigkeit und schließlich
- die Wirkung der Innovationsfähigkeit auf die wirtschaftliche Performance.

Der erste Wirkungszusammenhang betrifft die Effekte des untersuchten Förderungsinstruments auf verschiedene F&E-Inputindikatoren (F&E-Intensität, die Zahl der zusätzlichen F&E-Beschäftigten und die Zunahme der F&E-Personalkosten in Relation zu den F&E-Ausgaben insgesamt). Dabei wurde, im Rahmen ökonometrischer Analysen, die Gruppe der Teilnehmer mit einer Gruppe von Nicht-Teilnehmern verglichen. Auch wurde durch Unterteilung der Gruppe der Teilnehmer zusätzliche Ergebnisse gewonnen. Die übrigen beiden Wirkungszusammenhänge betreffen Effekte "höherer Ordnung".

8.6 Die indirekte F&E-Förderung in Österreich

8.6.1 Der Forschungsfreibetrag vor der Steuerreform 2000

Das österreichische Steuersystem kennt seit längerem eine ganze Reihe steuerlicher Begünstigungen für forschungs- bzw. innovationspolitisch relevante Tatbestände¹⁵). Das bei weitem bedeutendste Instrument der indirekten F&E-Förderung in Österreich ist jedoch der Forschungsfreibetrag. Die nachfolgenden Ausführungen beziehen sich ausschließlich auf diesen.

Gemäß § 4 Abs. 4 EStG steht – zusätzlich zu der sofortigen Geltendmachung als Betriebsausgaben – für "Aufwendungen zur Entwicklung oder Verbesserung volkswirtschaftlich wertvoller Erfindungen" ein Forschungsfreibetrag zu. Davon ausgenommen sind "Verwaltungs- und Vertriebskosten sowie Aufwendungen für Wirtschaftsgüter des Anlagevermögens" (Abschreibungen).

"Der volkswirtschaftliche Wert der angestrebten oder abgeschlossenen Erfindung ist durch eine Bescheinigung des Bundesministers für wirtschaftliche Angelegenheiten nachzuweisen. Die Bescheinigung ist nicht erforderlich, wenn die Erfindung bereits patentrechtlich geschützt ist."

Der Forschungsfreibetrag betrug vor der Reform grundsätzlich bis zu 12% der F&E-Aufwendungen. Ein erhöhter Forschungsfreibetrag (bis zu 18%) wurde anerkannt, "wenn die Erfindungen nicht anderen Personen zur wesentlichen Verwertung überlassen werden." "Eine Verwertung durch Dritte lag dann vor, "wenn Nutzungsrechte (Lizenzen) eingeräumt werden; wesentlich ist die Verwendung durch Dritte dann, wenn die Einnahmen (Nettolizenzeneinnahmen; auf den gesamten Betrieb bezogen, und nicht auf die einzelne Erfindung) aus der Verwertung durch Dritte mehr als 25% der Forschungsaufwendungen dieses Wirtschaftsjahres betragen" (*Doralt*, 1997, S. 268). Doralt vertrat in diesem Zusammenhang die Auffassung: "Das Verhältnis der Lizenzeneinnahmen aus einer bestehenden Erfindung zu den laufenden Forschungsaufwendungen für neue Erfindungen ergibt allerdings keinen Sinn"; die Abgrenzung von Eigennutzung und Fremdnutzung sei "insgesamt unbrauchbar und sachlich nicht gerechtfertigt".

Der Forschungsfreibetrag von bis zu 12% stand der Auftragsforschung zu. Der Auftraggeber hingegen konnte den Forschungsfreibetrag nicht geltend machen, wohl aber den Investitionsfreibetrag (siehe *Doralt*, 1997, S. 267). In der Praxis spielte dieser Satz laut vorliegenden Informationen eine geringe Rolle.

Zusammenfassend war der Forschungsfreibetrag vor der Reform 2000 durch folgende wesentliche Merkmale charakterisiert:

- 1) Begünstigung von "volkswirtschaftlich wertvollen Erfindungen" bzw. bei Vorhandensein eines Patentschutzes,
- 2) "volumensbasierte" Förderung (d.h. Begünstigung der gesamten laufenden F&E-Ausgaben),
- 3) differenzierte Begünstigung nach dem Kriterium der Eigen- oder Fremdverwertung.

¹⁵) Siehe dazu den Überblick von *Taucher* (1994), der in bezug auf die quantifizierten Einnahmehausfälle jedoch irreführend ist. Eine Darstellung neueren Datums ist *Schneider* (2000).

Die direkte Wirkung des Forschungsfreibetrages auf das Budget besteht in dem durch den Forschungsfreibetrag induzierten Einnahmenausfall¹⁶⁾. Die Förderungsberichte (bis zum Förderungsbericht 1997) des Bundesministeriums für Finanzen bezifferten die geschätzten Einnahmenausfälle aufgrund § 4 Abs. 4 EStG (Forschungsfreibetrag, Spendenbegünstigung für Wissenschaft und Forschung) seit fast einem Jahrzehnt mit jeweils insgesamt 450 Mill. S (davon Bund: 400 Mill. S) per annum. Die Zuverlässigkeit dieser Schätzung war fraglich, die Konstanz der jährlichen Einnahmenausfälle in den neunziger Jahren erscheint jedenfalls a priori unplausibel. Als Indiz sei hier angeführt, dass die durch den Unternehmenssektor finanzierten nominellen F&E-Ausgaben zwischen 1990 und 1998 gemäß den – vermutlich die tatsächliche Dynamik eher unterschätzenden (siehe *Leo*, 1999) – Angaben des ÖSTAT um 48% gestiegen sind. Im Förderungsbericht 1998 wurden diese Schätzungen nach oben revidiert, und zwar auf insgesamt 600 Mill. S (1996), 800 Mill. S (1997) und 1.000 Mill. S (1998). Der Bund ist nach dieser Schätzung mit 540 Mill. S im Jahr 1996 und 700 Mill. S in den beiden Folgejahren betroffen.

Eine zuverlässige Schätzung der Steuerausfälle in der Vergangenheit sowie der zusätzlichen Budgetbelastung ("tax expenditure") durch die Ausweitung der indirekten F&E-Förderung lag zum Beschluss der Steuerreform 2000 nicht vor.

8.6.2 Vorschläge zur Reform der indirekten F&E-Förderung

In den neunziger Jahren wurden verschiedene Vorschläge unterbreitet, die auf einen Ausbau bzw. eine Reform der indirekten F&E-Förderung in Österreich abzielten (siehe etwa *Expertenentwurf*, 1996 oder *Hochleitner – Schmidt*, 1997). Der *Beirat für Wirtschafts- und Sozialfragen* (1994) hielt eine F&E-Personalkostenförderung für überlegenswert (ohne sich darauf festzulegen, ob diese als direkte oder indirekte Förderung konzipiert werden sollte). Seitens der Wirtschaft wurde eine Erhöhung des Forschungsfreibetrages auf 30% gefordert. Den maßgeblichen Hintergrund für diese Vorschläge bildet das für eine hochentwickelte Volkswirtschaft seit langem als unbefriedigend empfundene Niveau der F&E-Ausgaben, insbesondere des Unternehmenssektors (siehe etwa *Hutschenreiter et al.*, 1998). In der jüngsten Zeit verdichtet sich zwar die Evidenz, dass die amtliche F&E-Statistik Niveau bzw. Dynamik der F&E-Ausgaben des Unternehmenssektors unterschätzt (siehe *Leo*, 1999), ohne dass dadurch das generelle Ziel einer Intensivierung der Industrieforschung in Frage gestellt wäre.

Mit dem Bericht der Steuerreformkommission vom November 1998 kam die Option des Ausbaus der indirekten F&E-Förderung erneut auf die wirtschaftspolitische Tagesordnung (wobei sich die Kommission bezüglich einer Erhöhung der indirekten oder der direkten Förderung indifferent zeigte)¹⁷⁾.

Anlässlich der Regierungsklausur in Bad Aussee (21. bis 22. Jänner 1999) wurde eine Anhebung der österreichischen F&E-Quote bis 2005 auf 2,5% (1999 laut revidierter amtlicher

¹⁶⁾ Der Gesamteffekt ist natürlich wesentlich komplexer, da die F&E-Förderung die mittel- und langfristig erzielbaren Einkommen und damit das Steueraufkommen beeinflusst. Darüber hinaus gibt es Wechselwirkungen im Steuersystem sowie – durch die Allokationsentscheidungen multinationaler Unternehmen – komplexe Interaktionen zwischen dem nationalen Steuersystem und den Steuersystemen anderer Länder (siehe *Hall*, 1992, *Hines*, 1994).

¹⁷⁾ Die Steuerreformkommission sah folgende "Möglichkeiten, die Forschungsförderung attraktiver zu gestalten: Anhebung des Forschungsfreibetrags auf 30%, was einen Einnahmenausfall in Höhe von 500 bis 600 Mio S auslösen würde und aus EU-rechtlicher Sicht unbedenklich erscheint. Direkte Förderungen aus dem Budget etwa in Höhe des durch den Forschungsfreibetrag entstehenden Steuerausfalls, bei gleichzeitigem Auslaufen des Forschungsfreibetrags." Die Kommission ging offensichtlich von der inzwischen nach oben revidierten Schätzung des Einnahmenausfalls des Bundesministeriums für Finanzen aus.

Statistik 1,63%) als Ziel formuliert. Der Ergebnisbericht der Klausur stellt fest, dass es hierzu erforderlich sein werde, "die Unternehmen zu mehr Einsatz im Bereich F&E zu stimulieren, aber auch die aus dem Bereich der öffentlichen Hände stammenden Mittel zu erhöhen". In diesem Zusammenhang wird unter anderem eine "signifikante Erhöhung des Forschungsfreibetrages" angekündigt.

Mit dem Steuerreformgesetz 2000 wurde der Forschungsfreibetrag schließlich neu gestaltet.

8.6.3 Die Steuerreform 2000

Die Reform des Forschungsfreibetrags im Rahmen der Steuerreform 2000 umfasst die folgenden Kernpunkte:

- 1) Der Forschungsfreibetrag wird auf generell bis zu 25% angehoben.
- 2) "Zusätzliche" (über eine – als gleitender Dreijahresdurchschnitt berechnete – Baseline hinausgehende) F&E-Ausgaben lukrieren einen Forschungsfreibetrag von bis zu 35%.
- 3) Die Differenzierung der Begünstigung nach dem Kriterium der Eigen- oder Fremdverwertung wird aufgehoben.

Zur Bewertung der mit dem Steuerreformgesetz 2000 durchgeführten Reform lässt sich folgendes feststellen:

In der Praxis spielte der Forschungsfreibetrag von bis zu 12% vor Reform nur eine geringe Rolle. Er konnte daher schon aus diesem pragmatischen Grund abgeschafft werden. Darüber hinaus gab es wie erwähnt begründete Zweifel an der Sinnhaftigkeit der Abgrenzung zwischen Eigennutzung und Fremdnutzung (*Doralt*, 1997, S. 268).

In der Diktion der OECD stellt der neue österreichische Forschungsfreibetrag eine Kombination aus einer "volumensbasierten" und "inkrementellen" Förderung dar. Darüber hinaus wurde, entgegen der internationalen Praxis der letzten Jahre, ein gleitender Durchschnitt zur Bestimmung der Baseline gewählt. Die Abschätzung der Konsequenzen dieser in der konkreten Form international einzigartige Kombination wäre Gegenstand einer Design-Evaluierung.

8.7 Anwendung der internationalen Evaluierungserfahrungen auf Österreich

Der Forschungsfreibetrag wurde bisher keinerlei Evaluierung unterzogen. Es besteht sogar erhebliche Unsicherheit über grundlegende Fakten, etwa bezüglich der Höhe des induzierten Steuerausfalls und die Struktur der Inanspruchnahme.

Da aber die Wirkung steuerlicher Anreize für F&E in hohem Maße von der konkreten Ausgestaltung der Förderung, ihrer Durchführung und von den jeweiligen nationalen Rahmenbedingungen (z.B. Umfang und Qualität der direkten F&E-Förderung) beeinflusst wird, kommt Evaluierungen gerade in diesem Bereich eine besondere Bedeutung zu. Aus dieser Feststellung folgt aber auch, dass Evaluierungsergebnisse aus anderen Ländern nicht ohne weiteres übertragbar sind.

Dass es bisher zu keiner Evaluierung indirekter Instrumente der Forschungsförderung in Österreich kam, hängt wohl auch damit zusammen, dass Steuerbegünstigungen ("tax expenditures") generell weit weniger im Blickpunkt der Öffentlichkeit stehen als direkte Förderungen, obwohl sie aus finanzwissenschaftlicher Sicht – grob gesprochen – äquivalent

sind¹⁸⁾). Dennoch mag es erstaunen, dass die Ausweitung der durchaus nicht unbedeutenden indirekten Förderung im Rahmen der Steuerreform 2000 ohne vorbereitende Evaluierung durchgeführt wurde.

Dass die in Österreich gewährten steuerlichen Anreize für F&E auch im internationalen Vergleich nicht unbedeutend sind, zeigt eine Untersuchung der OECD. Die OECD hat ein Länder-Ranking in Bezug auf die Generosität der indirekten F&E-Förderung auf Basis des sogenannten B-Index¹⁹⁾ vorgenommen. Danach lag Österreich – vor der Steuerreform – im oberen Mittelfeld der OECD-Länder. Mit der Reform des Forschungsfreibetrages im Rahmen der Steuerreformgesetzes 2000 rückt Österreich zweifellos in das internationale Spitzenfeld auf.

Es ist daher unbedingt zu empfehlen, im Sinne der Sicherung der Effizienz und Effektivität des Forschungsfreibetrags der Reform einen verbindlichen Evaluierungsplan folgen zu lassen. In den Niederlanden etwa wurde unmittelbar nach Inkrafttreten der 1994 eingeführten steuerlichen Förderung von F&E-Personalkosten mit Evaluierungsarbeiten begonnen. Bemerkenswert sind wie gezeigt auch die umfangreichen und nachhaltigen Evaluierungsanstrengungen in Zusammenhang mit der australischen "R&D Tax Concession".

Das analytische Instrumentarium für eine Evaluierung indirekter F&E-Förderungsinstrumente ist vorhanden²⁰⁾. Erleichtert wird eine effektive Evaluierung durch möglichst operationale Ziele sowie die laufende Erfassung relevanter Daten im Monitoring/Auditing-Prozeß.

8.7.1 Die Ziele der steuerlichen F&E-Förderung in Österreich

Es ist aus mehreren Gründen erforderlich, die Ziele festzumachen, die durch die indirekte F&E-Förderung erreicht werden sollen. Die Definition der Ziele beeinflusst das Design des Förderungsinstruments und dessen Positionierung in der Förderungslandschaft. Schließlich ist nur auf Grundlage einigermaßen präzise definierter Ziele eine effektive Evaluierung möglich.

- Als allgemeinstes Ziel eines Instruments der steuerlichen F&E-Förderung kann die generelle Stimulierung unternehmerischer Forschung und Entwicklung definiert werden, die sich in einer Erhöhung der Forschungsausgaben oder –quoten niederschlägt.

Diese generelle Zielsetzung kann mit der Erwartung verbunden werden, damit das Innovationspotential der Unternehmen mitsamt der Fähigkeit zur Übernahme von extern produziertem Wissen und so letztlich ihre Wettbewerbsfähigkeit zu erhöhen.

Darüber hinaus können mit der steuerlichen Begünstigung von F&E weitere Neben- oder Unterziele verfolgt werden, im österreichischen Kontext zum Beispiel die folgenden:

- Sicherung und Stärkung des Forschungsstandorts.

¹⁸⁾ Diese Einschätzung stimmt mit der Beobachtung überein, daß die Effektivität und Effizienz auch von umfangmäßig bedeutenderen Instrumenten der indirekten Förderung, wie etwa des Investitionsfreibetrages – der Förderungsbericht 1998 des Bundesministeriums für Finanzen beziffert den durch den Investitionsfreibetrag induzierten Steuerausfall in den Jahren 1996-1998 mit 8 Mrd. S (1996) bzw. 5,5 Mrd. S (1997 und 1998) – nur phasenweise im Blickpunkt der wirtschaftspolitischen Diskussion stehen. Zuletzt wurde die Investitionsförderung in ihrer Gesamtheit in den achtziger Jahren eingehenden Untersuchungen, einschließlich ökonomischer Analysen, unterzogen (siehe etwa Szopo – Aiginger – Lehner, 1985). Für den Bereich der Forschungs- und Technologieförderung haben Fritz – Hutschenreiter – Sturn (1997) festgestellt, daß in Österreich Evaluierungen typischerweise bei volumensmäßig eher unbedeutenden direkten Förderungsprogrammen stattfanden.

¹⁹⁾ Siehe OECD (1998). Nähere Details finden sich bei Warda (1996).

²⁰⁾ Einen Überblick über eine Reihe empirischer Studien bieten Cordes (1989), Harhoff (1994), OECD (1995A) und zuletzt OECD (1997, 1998), Hall – van Reenen (2000).

- Stabilisierung der F&E-Aktivitäten in Unternehmen, die sporadisch F&E betreiben.
- Erweiterung des Kreises der Forschung und Entwicklung betreibenden Unternehmen.

Die zuletzt genannten beiden Zielsetzungen betreffen insbesondere (jedoch nicht ausschließlich) KMU, im zweiten Fall auch technologieorientierte Start-up-Unternehmen. Eine (steuerliche) Förderung der F&E-Aktivitäten dieses Adressatenkreises ist von besonderer Bedeutung in Ländern mit schwach entwickelten Venture-Capital-Märkten. Aber auch in Ländern mit hochentwickelten Venture-Capital-Märkten kommt nur ein relativ kleiner Teil der technologieorientierten Neugründungen in den Genuss von Risikokapitalfinanzierung (Lerner, 1999).

Aus der Gestaltung des Forschungsfreibetrages vor der Steuerreform 2000 war keine Intention erkennbar, KMU oder Neugründungen in besonderem Maße zu fördern. In der Praxis dürfte die Inanspruchnahme des Forschungsfreibetrages in stärkerem Maße auf größere Unternehmen konzentriert gewesen sein als dies in Hinblick auf die einem indirekten Förderungsinstrument zugeordnete Funktion (niedrige Zutrittschranken, breite Streuung des Kreises der Begünstigten etc.) gerechtfertigt erscheint.

Die Regierungsklausur in Bad Aussee Anfang 1999, nannte in Zusammenhang mit der im Rahmen der Steuerreform 2000 realisierten Erhöhung und Neugestaltung des Forschungsfreibetrags explizit die Stimulierung von F&E-Ausgaben im Unternehmenssektor als Ziel.

Mit der Einführung eines besonderen Anreizes für zusätzliche F&E-Ausgaben im Rahmen der Steuerreform 2000 dürfte darüber hinaus die Absicht verbunden sein, F&E in neu F&E betreibenden Unternehmen, darunter technologieorientierte junge Unternehmen, speziell zu begünstigen.

Im Vorfeld der Reform wurde darauf hingewiesen, dass gerade kleinere Unternehmen und technologieorientierte Start-ups in besonderem Maße von Marktversagen (z. B. asymmetrische Information) betroffen sind und die spezifischen Bedürfnisse dieser Zielgruppe deutlich stärker zu berücksichtigen sind. Eine solche Orientierung steht in Einklang mit den von der *OECD* (1998) erarbeiteten Best Practices.

Generell ist festzustellen, dass die Ziele von Instrumenten der steuerlichen F&E-Förderung in anderen Ländern expliziter und umfassender formuliert wurden (Beispiel Australien).

8.7.2 Design/Operationalisierung

Internationale Forschungsarbeiten belegen, dass das Design indirekter Förderungsinstrumente gravierende Auswirkungen auf die davon ausgehenden Anreizwirkungen haben kann. Gestützt auf die in den letzten zwei Jahrzehnten gesammelten Erfahrungen gelangt die *OECD* (1998, S. 180) zu der Schlussfolgerung: "The detailed design of tax incentive schemes often makes the difference between efficient and wasteful fiscal support to R&D."

Es liegt mittlerweile eine Vielzahl von Studien vor, die sich mit dem Design und den Effekten steuerlicher Förderungsinstrumente in einer Reihe von Ländern beschäftigen. Diese Erfahrungen wurden in den letzten Jahren vor allem im Rahmen der *OECD* systematisiert und zu "Best Practices" kondensiert (siehe *OECD*, 1996, 1997, 1998).

Ausgehend von den Zielen ist die Frage zu stellen, ob das Design, d.h. die konkrete Ausgestaltung des Förderungsinstruments der Erreichung dieser Ziele angemessen ist.

Gewinnabhängige Anreize (wie der österreichische Forschungsfreibetrag) vor und nach Reform benachteiligen technologieorientierte Unternehmensneugründungen insofern als diese in der Startphase in der Regel keine Gewinne erzielen. Darüber hinaus wirkt er prozyklisch, was unter dem Gesichtspunkt der Verstetigung der F&E-Tätigkeit über den Konjunkturzyklus nicht wünschenswert ist. Steuergutschriften für F&E wurden daher zuweilen mit einer "Negativsteuer-Komponente" versehen (Umwandlung des nicht beanspruchbaren Teils des Absetzbetrags in einen Zuschuss). Alternativ zur Negativsteuer gibt es Vortrags- oder Nachtragsmöglichkeiten. Dies wirkt zwar der Zyklizität entgegen, die Liquiditätsprobleme von technologieorientierten Unternehmen in der Start-up-Phase werden dadurch jedoch nicht gemildert.

Es besteht die begründete Vermutung, dass bei der Erhöhung und Umgestaltung des Forschungsfreibetrags im Rahmen der Steuerreform 2000 die Möglichkeiten zur Steigerung der Effektivität und Effizienz der steuerlichen F&E-Förderung durch eine Reform des Instruments bzw. der administrativen Abwicklung nicht hinreichend genutzt wurden.

Einbettung in das System der F&E-Förderung

Die Ausweitung der indirekten F&E-Förderung hat Rückwirkungen auf die bestehenden Instrumente und Institutionen der F&E-Förderung. Umgekehrt kann davon ausgegangen werden, dass die Anreizwirkungen einer steuerlichen Förderung umso geringer sind, je ausgebauter die direkte F&E-Förderung ist (OECD, 1998). Daraus ergibt sich ein Bedarf, die Förderungsinstrumente aufeinander abzustimmen.

Es wurde darauf hingewiesen, dass dies in besonderem Maße, wenngleich nicht ausschließlich, den FFF betrifft, der in mancher Hinsicht Funktionen erfüllt, die zumindest aus theoretischer Sicht eher der indirekten F&E-Förderung zugeordnet werden (geringe Spezifität, relativ leichter Zugang etc.). Ein Ausbau der indirekten F&E-Förderung entlastet die bestehenden Förderungsinstitutionen und bietet ihnen mehr (finanziellen) Spielraum in Kernbereichen der direkten Förderung. Dies könnte sich sowohl in höherer "Spezifität" als auch – bei konstanter Allokation der direkten Förderungsmittel – in höheren Förderintensitäten niederschlagen.

Darüber hinaus könnte eine erweiterte, leicht zugängliche indirekte F&E-Förderung – insbesondere im Bereich der KMU und Neugründungen – komplementär auch Segmente bedienen, die bisher (etwa aufgrund von Mindestanforderungen an die Projektgröße und –darstellung) nur in geringem Maß Zugang zu Förderungen hatten.

Begünstigung der gesamten oder der zusätzlichen F&E-Ausgaben

Wie oben erwähnt kommen international zwei Modelle der steuerlichen Förderung von F&E zum Einsatz: Im volumenbasierten Modell werden die gesamten F&E-Ausgaben (in der jeweiligen Spezifikation) begünstigt, im zweiten Modell nur die zusätzlichen ("inkrementellen") F&E-Ausgaben. Aus ökonomischer Sicht gibt es Argumente für beide Formen.

Eine inkrementelle Förderung ist prinzipiell die kosteneffektivere Form der Förderung. Voraussetzung für die Bestimmung der "zusätzlichen" F&E-Ausgaben ist die Festlegung eines Referenzwerts ("baseline"). Die zusätzlichen Ausgaben sind als Differenz zwischen den aktuellen und den Referenzausgaben definiert. Die Definition der "baseline" ist allerdings tückisch, eine Fehlspezifikation kann die Wirksamkeit eines indirekten Förderungsinstruments drastisch vermindern. Dies war etwa beim "R&D Tax Credit" in den

USA der Fall (siehe *Hall*, 1992), wo die "baseline" in der Vergangenheit als Durchschnitt der F&E-Ausgaben der vorangegangenen drei Jahre definiert wurde. In der Zwischenzeit sind die USA auf eine fixe Baseline übergegangen (siehe *Whang*, 1998).

Die inkrementelle Form der Förderung hat gegenüber dem volumensbasierten Modell den Nachteil, dass Unternehmen mit demselben Niveau an – aus ökonomischer Sicht förderungswürdigen – F&E-Aktivitäten unterschiedlich behandelt werden. Überdies ist die inkrementelle Form der Förderung manipulationsanfälliger.

In Österreich wurde mit der Steuerreform 2000 eine Mischform aus volumensbasierter und inkrementeller Förderung gewählt.

Definition der förderungswürdigen F&E-Ausgaben

Die überwiegende Mehrzahl der OECD-Länder begünstigt sowohl laufende F&E-Ausgaben als auch entsprechende tangible Investitionen. Österreich und die USA beziehen ihren Forschungsfreibetrag bzw. "R&D Tax Credit" ausschließlich auf die laufenden F&E-Ausgaben. Im Fall Österreichs ist dies in Zusammenhang mit der Existenz des Investitionsfreibetrags zu sehen.

Das sowohl innovative als auch praktikable steuerliche Anreizinstrument der Niederlande ("Wage Tax and Social Insurance Allowance Act" - WVA/S&O) setzt wie erwähnt an den F&E-Personalkosten (entspricht in etwa den Bruttolöhnen des F&E-Personals, exklusive Sozialversicherungsbeiträge der Arbeitgeber) an²¹). Eine Besonderheit des niederländischen Modells besteht darin, dass es eine Reduktion der arbeitskostenbezogenen Steuerbelastung und der Sozialversicherungsbeiträge der Unternehmen vorsieht. Dieses Modell reduziert also die Lohnnebenkosten und ist unabhängig von den laufenden Gewinnen. Eine an den Lohnkosten für das (gesamte oder zusätzliche) F&E-Personal ansetzende Förderung ist vom Ansatz her transparent, leichter administrierbar als alternative Varianten und hat auch hohe Signalwirkung. Ein weiterer Vorteil besteht in der politischen Akzeptanz²²). Gewinnunabhängige steuerliche Anreize für F&E – wie etwa in den Niederlanden – haben positive Effekte insbesondere auf technologieorientierte Start-ups. Eine Konstruktion wie im niederländischen Modell ist auch konsistent mit kurz- und längerfristigen beschäftigungspolitischen Zielsetzungen.

²¹) Außer in den Niederlanden gibt es im OECD-Raum ein an den Personalkosten ansetzendes Instrument der indirekten F&E-Förderung derzeit nur in der kanadischen Provinz Quebec.

²²) Dies äußert sich etwa in dem Umstand, daß sich Überlegungen zu einer F&E-Personalkostenförderung in der Wirtschaftstandort-Studie des *Beirats für Wirtschafts- und Sozialfragen* (1994) finden. Die F&E-Ausgaben der Industrie bestehen de facto zu einem großen Teil aus Personalkosten. Laut F&E-Erhebung 1993 der Wirtschaftskammer Österreich entfielen auf Personalkosten 62%, auf Sachausgaben 30% und auf Investitionen 8% der gesamten betriebsinternen F&E-Ausgaben der österreichischen Industrie. Sofern – wie in Österreich – F&E-bezogene Aufwendungen für Wirtschaftsgüter des Anlagevermögens nicht begünstigt werden, würde ein Übergang auf eine Förderung von Personalkosten daher die begünstigten F&E-Ausgaben um rund ein Drittel vermindern. Eine solche Reduktion könnte natürlich durch einen höheren Satz kompensiert werden.

Zusätzliche Kriterien ("volkswirtschaftlicher Wert" und Patentschutz)

Wie dargestellt bedarf es zur Inanspruchnahme des Forschungsfreibetrages auch nach der Reform des Nachweises des volkswirtschaftlichen Werts der betreffenden Erfindung. Dieser Nachweis entfällt, wenn ein Patent vorliegt.

Der volkswirtschaftliche Wert einer Erfindung ist nur durch anspruchsvolle ökonomische Forschungsarbeiten abzuschätzen, und auch das – selbst ex post – nur annähernd. Wäre es einfach, die Differenz zwischen privaten und sozialen Erträgen einer Innovation zu ermitteln, so könnte Marktversagen treffsicher durch eine optimale *direkte* Subvention behoben werden. Als ein in der Verwaltungspraxis administrierbares Konzept dürfte sich der "volkswirtschaftliche Wert" kaum eignen. Es wurden auch verfassungsrechtliche Bedenken gegen die Voraussetzung des volkswirtschaftlichen Werts "mangels Determiniertheit" geäußert (siehe *Doralt*, 1997, S. 266).

Die Äquivalenz von "volkswirtschaftlichem Wert" und dem Faktum einer Patentanmeldung ist fragwürdig. So zeigt die wirtschaftswissenschaftliche Literatur, dass die Patentneigung der Branchen beträchtlich und systematisch variiert und überdies der ökonomische Wert von Patenten enorm streut.

Es gibt gute Gründe, den Nachweis des volkswirtschaftlichen Werts bzw. die Patentierung als Voraussetzung für die Gewährung der Förderung fallen zu lassen und sich ausschließlich auf den Nachweis des (entsprechend definierten) F&E-Aufwands zu beschränken. Daraus würde jedoch ein gewisser Bedarf zur Neugestaltung des Auditing resultieren.

8.7.3 Wirkungsanalyse

Bei der Bewertung der ökonomischen Wirkungen eines Instruments der indirekten F&E-Förderung wird häufig die Frage nach den dadurch bewirkten zusätzlichen unternehmerischen F&E-Ausgaben (d.h. nach der sogenannten "Additionalität") gestellt. Obgleich die "Additionalität" wie erwähnt keineswegs das einzige, und aus ökonomisch-theoretischer Perspektive nicht einmal das entscheidende Evaluierungskriterium darstellt, ist diese Frage natürlich von praktischer Relevanz und ist auch Ausgangspunkt für umfassendere Wirkungsanalysen.

Im Zusammenhang mit den forschungs- und technologiepolitischen Zielen Österreichs kann gefragt werden, in welchem Umfang erwartet werden kann, dass eine Erhöhung der steuerlichen F&E-Förderung zur Erreichung einer Forschungsquote von 2,5% bis zum Jahr 2005 beiträgt. Der Forschungsfreibetrag vor der Reform 2000 war rein volumensbasiert und vermutlich ist die Wirkung des Freibetragsmodells nach Reform nicht sehr weit von der eines volumensbasierten Ansatzes entfernt. Hinsichtlich volumenbasierter Steueranreize gelangt die *OECD* (1998, S. 177) zusammenfassend zu der Schlussfolgerung, dass diese wenig mehr F&E-Ausgaben induzieren als der Steuerausfall bei den öffentlichen Haushalten ausmacht. Unter Zugrundelegung von Resultaten der internationalen Forschung kann man – bei vorsichtiger Beurteilung des Beitrags, den die steuerlichen F&E-Förderung zur Erreichung einer Forschungsquote von 2,5% leisten kann – die Hypothese aufstellen, dass diese größenordnungsmäßig in etwa dem Steuereinnahmenschwund entspricht. Eine Prüfung dieser Hypothese ist nur durch eine eigene Wirkungsanalyse des Forschungsfreibetrags möglich.

Eine künftige Wirkungsanalyse des Forschungsfreibetrags sollte im Kern zumindest eine (ökonometrische) Schätzung der durch den Forschungsfreibetrag bei den Unternehmen induzierten F&E-Ausgaben umfassen. Eine solche Schätzung ist sowohl für die Durchführung einer umfassenderen Kosten-Nutzen-Analyse erforderlich (siehe das Beispiel Australien) als auch Voraussetzung für die Quantifizierung von "Effekten höherer Ordnung" wie die Folgewirkungen einer Erhöhung der F&E-Ausgaben auf die Innovationsfähigkeit und in weiterer Folge auf die wirtschaftliche Leistungsfähigkeit (siehe das Beispiel Niederlande).

8.7.4 Datenerfordernisse

Für Zwecke der Evaluierung, insbesondere der Wirkungsanalyse, bedarf es quantitativer Informationen. Die quantitative Evaluierung wird wesentlich erleichtert durch eine laufende Erfassung der F&E-Ausgaben nicht nur der steuerlich begünstigten F&E-Vorhaben, sondern der gesamten F&E-Ausgaben aller begünstigten Unternehmen.

Prinzipiell ist zu empfehlen, eine Definition der F&E-Aufwendungen zu wählen, die den im Frascati-Manual (OECD, 1994) festgelegten Standards der OECD entspricht. Dies hätte als wünschenswerten Nebeneffekt die Verfügbarkeit rezenter, standardisierter F&E-Daten für den Unternehmenssektor, die nicht nur für Zwecke der Evaluierung der F&E-Förderung, sondern auch als Element eines Management-Informationssystem der Technologiepolitik von großem Nutzen sein könnte. Derzeit ist die ihrer Zielsetzung nach zukunftsorientierte Technologiepolitik paradoxerweise oft gezwungen, sich auf fatal veraltete Daten als Informationsquelle zu stützen.

8.8 Schlussfolgerungen

Die indirekte F&E-Förderung ist keine unbedeutende Komponente der F&E-Förderung in Österreich. Laut Förderungsbericht 1998 induzierten im Jahr 1998 (also vor der Erhöhung des Forschungsfreibetrags im Rahmen der Steuerreform 2000) der Forschungsfreibetrag und die Spendenbegünstigungen für Wissenschaft und Forschung einen Steuerausfall in Höhe von einer Milliarde Schilling. Auch international dürfte Österreich hinsichtlich der Generosität der steuerlichen F&E-Förderung im Spitzenfeld liegen.

Dennoch wurde der Forschungsfreibetrag niemals einer Evaluierung unterzogen. Vielmehr besteht Unsicherheit über grundlegende Fakten wie die Höhe des tatsächlichen Steuerausfalls oder die Struktur der Inanspruchnahme. Über die Höhe der induzierten zusätzlichen F&E-Ausgaben liegt keinerlei Information vor.

Quantitative Evaluierungen indirekter F&E-Förderungsinstrumente wurden in einer Reihe von Ländern erprobt und dienen auch in der Praxis als Entscheidungsgrundlage für Reformen. Als Best-Practice-Beispiele können etwa die umfangreichen Evaluierungsanstrengungen in Australien sowie die Realisierung eines Evaluierungsplans unmittelbar nach Einführung eines innovativen Instruments der steuerlichen F&E-Förderung in den Niederlanden gelten. Ebenso liegen umfangreiche und durchaus praxisrelevante internationale Erkenntnisse und Erfahrungen bezüglich des Designs steuerlicher Anreizinstrumente für F&E vor.

Angesichts der Bedeutung des Forschungsfreibetrags ist zu empfehlen, das Design und die Wirkung dieses Instruments einer Evaluierung zu unterziehen. Dies ist nicht zuletzt zur Gewährleistung der Effizienz und Effektivität der indirekten F&E-Förderung geboten. Die Wirkungsanalyse sollte zumindest eine (ökonometrische) Schätzung der durch den Forschungsfreibetrag bei den Unternehmen induzierten F&E-Ausgaben beinhalten. Eine

solche Schätzung ist sowohl für die Durchführung einer umfassenderen Kosten-Nutzen-Analyse erforderlich (siehe das Beispiel Australien) als auch Voraussetzung für die Quantifizierung von "Effekten höherer Ordnung" wie die Folgewirkungen einer Erhöhung der F&E-Ausgaben auf die Innovationsfähigkeit und in weiterer Folge auf die wirtschaftliche Leistungsfähigkeit (siehe das Beispiel Niederlande).

Die quantitative Wirkungsanalyse wird wesentlich erleichtert durch eine laufende Erfassung relevanter Daten, insbesondere der F&E-Ausgaben nicht nur der steuerlich begünstigten F&E-Vorhaben, sondern der gesamten F&E-Ausgaben aller begünstigten Unternehmen (bevorzugt in der geltenden OECD-Definition). Eine solche Erfassung von F&E-Daten wäre darüber hinaus ein wertvolles Element eines aktuellen Informationssystems der Forschungs- und Innovationspolitik.

8.9 Literatur

- Beirat für Wirtschafts- und Sozialfragen, Wirtschaftsstandort Österreich, Wien, 1994.
- "BIE": Australian Bureau of Industry Economics, R&D, Innovation and Competitiveness. An Evaluation of the Research and Development Tax Concession, Research Report 50, Australian Government Publishing Service, Canberra, 1993.
- Cohen, L., Noll, R., The Technology Pork Barrel, Brookings Institution, Washington, D.C., 1991.
- Cordes, J.J., "Tax Incentives and R&D Spending: A Review of the Evidence", Research Policy, 1989, (18), 119-133.
- Doralt, W., Einkommensteuergesetz. Kommentar Teil I bis III, §§ 1-32, 3. Auflage, WUV-Universitätsverlag, Wien, 1997.
- Dorsman, M., "Evaluation of Industrial R&D Support in the Netherlands: The Wage Tax and Social Insurance Allowances Act/R&D Allowance", in OECD (1997), 135-145.
- Dorsman, M., Evaluation of the Wage Tax and Social Security Premiums Allowances Act, Research and Development Component, Formerly the WBSO, Ministry of Economic Affairs, Den Haag, 1998.
- Europäische Kommission, Mitteilung der Kommission über die Anwendung der Vorschriften über staatliche Beihilfen auf Maßnahmen im Bereich der direkten Unternehmensbesteuerung, 11. November 1998.
- "Expertenentwurf": WIFO, Forschungszentrum Seibersdorf, Joanneum Research, Technologiepolitisches Konzept 1996 der Bundesregierung. Expertenentwurf, Wien, 1996.
- Fritz, O., Hutschenreiter, G., Sturn, D., Evaluierung von FTE-Programmen, Joanneum Research, Graz, 1997.
- Hall, B. H. , "R&D Tax Policy During the Eighties: Success or Failure?", NBER Working Paper, 1992, (4240).
- Hall, B., van Reenen, J., "How Effective are Fiscal Incentives for R&D? A Review of the Evidence", Research Policy, 2000, 29(4-5), 449-469.
- Harhoff, D., Zur steuerlichen Behandlung von Forschungs- und Entwicklungsaufwendungen. Eine internationale Bestandsaufnahme, ZEW Dokumentation, 1994, (94-02).
- Hines, J.R., "Taxes, Technology Transfer, and the R&D Activities of Multinational Firms", NBER Working Paper, 1994, (4932).
- Hochleitner, A., Schmidt, A., Forschung und Wettbewerb. Technologieoffensive für das 21. Jahrhundert, Bericht an die Bundesregierung, Wien, 1997.
- Hutschenreiter, G., "Intersektorale und internationale 'F&E-Spillovers'. Externe Effekte von Forschung und Entwicklung", WIFO-Monatsberichte, 1995, 68(6), 419-427.
- Hutschenreiter, G., Knoll, N., Paier, M., Ohler, F., Österreichischer Technologiebericht 1997, tip-Programm, Wien, 1998.

- Joint Committee on Taxation, Description and Analysis of Tax Provisions Expiring in 1992, U.S. Government Printing Office, Washington, D.C.
- Klette, T.J., Moen, J., Griliches, Z., "Do Subsidies to Commercial R&D Reduce Market Failures? Microeconomic Evaluation Studies", *Research Policy*, 2000, 29(4-5), 471-495.
- Lattimore, R., "Research and Development Fiscal Incentives in Australia: Impacts and Policy Lessons", in OECD (1997), 91-134.
- Leo, H., "Österreichs Innovations- und Forschungsleistung im internationalen Vergleich", *WIFO-Monatsberichte*, 1999, 72(6), 435-443.
- Lerner, J., "The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program", *Journal of Business*, 1999, 72(3), 285-318.
- Leyden, D.P., Link, A.N., *Government's Role in Innovation*, Kluwer Academic Publishers, Dordrecht, 1992.
- Lhuillery, S., "Problems Involved in Designing and Implementing R&D Tax Incentive Schemes", in OECD (1996), 38-57.
- Link, A., "Fiscal Measures to Promote R&D and Innovation – Trends and Issues", in OECD (1996), 23-33.
- Mohnen, P., "R&D Externalities and Productivity Growth", *STI Review*, 1996, (18), 19-66.
- OECD, *Frascati Manual 1993. Proposed Standard Practice for Surveys of Research and Experimental Development*, Paris, 1994.
- OECD (1995A), *Impacts of National Technology Programmes*, Paris, 1995.
- OECD (1995B), *National Systems for Financing Innovations*, Paris, 1995.
- OECD, *Fiscal Measures to Promote R&D and Innovation*, OECD/GD(96)165, Paris, 1996.
- OECD, *Policy Evaluation in Innovation and Technology. Towards Best Practices*, Paris, 1997.
- OECD, *Technology, Productivity and Job Creation. Best Policy Practices*, Paris, 1998.
- Office of Technology Assessment (OTA), *The Effectiveness of Research and Experimentation Tax Credits*, Congress of the United States, Washington, D.C., 1995.
- Schneider, H.W., *Steuerliche Begünstigung von Forschung und Entwicklung*, Linde Verlag, Wien, 2000.
- Szopo, P., Aiginger, K., Lehner, G., *Investitionsförderung in Österreich*, Österreichisches Institut für Wirtschaftsforschung, Wien, 1985.
- Taucher, "Prüfung des Einsatzes des Steuerrechts für die indirekte Förderung von F&E nach dem EU-Beitritt Österreichs", *Finanz-Journal*, 1994, (12), 274-280.
- Warda, J., "Measuring the Value of R&D Tax Provisions", in OECD (1996), 8-22.
- Whang, K.J., "Fixing the Research Credit", *Issues in Science and Technology Online*, 1998, (Winter), <http://bob.nap.edu/issues/15.2/whang.htm>.

TEIL C