

**Aus dem Institut für Betriebswirtschaft, Agrarstruktur
und ländliche Räume**

**Bernhard Forstner
Regina Grajewski
Peter Mehl**

**Evaluation von Politikmaßnahmen zur Förderung
ländlicher Räume : Theorie und Praxis**

Manuskript, zu finden in www.fal.de

Vortrag gehalten am 2.10.2002 auf der GEWISOLA-Tagung in
Halle

**Braunschweig
Bundesforschungsanstalt für Landwirtschaft (FAL)
2002**

Also available at:

http://www.bal.fal.de/download/Gewisola_2002_endg_Stand21_11.pdf

Evaluation von Politikmaßnahmen zur Förderung ländlicher Räume: Theorie und Praxis

von

B. FORSTNER, R. GRAJEWSKI, P. MEHL^{*}

1 Einleitung

Seit der Reform der Strukturfonds der Europäischen Union (EU) im Jahr 1988 sind Evaluationen zunehmend zum festen Bestandteil des Politikzyklus in diesem Politikbereich geworden. Mit Beginn der Förderperiode 2000 bis 2006 hat die EU die diesbezüglichen Anforderungen verschärft und systematisiert. Alle EU-Mitgliedstaaten sind nunmehr verpflichtet, die im Rahmen der EU-Verordnung über die Förderung der Entwicklung des ländlichen Raums (VO (EG) Nr. 1257/1999) durchgeführten Fördermaßnahmen zu begleiten (Monitoring) und zu bewerten (Evaluation).

Die zunehmende Bedeutung von Evaluationen im Bereich der Politik zur Entwicklung der ländlichen Räume trägt dem wachsenden politischen und fiskalischen Stellenwert dieses Politikbereichs als der zweiten Säule der gemeinsamen Agrarpolitik (GAP) Rechnung. Die gesteigerte Verbindlichkeit ist gewissermaßen die Kehrseite des erweiterten Spielraums, den die o. g. Verordnung den Mitgliedstaaten bei der Ausgestaltung ihrer Politik im Unterschied zu den Vorgängerregelungen eröffnet. Überdies werden die Programme der Bundesländer in hohem Maße durch EU-Gelder und Bundesmittel aus der Gemeinschaftsaufgabe "Verbesserung der Agrarstruktur und des Küstenschutzes" und nur in geringem Maße durch eigene Landesmittel finanziert. In den besonders intensiv geförderten Ziel-1-Gebieten beträgt der maximale Kofinanzierungssatz von Bund und EU 90 % der öffentlichen Mittel. Es erscheint zweifelhaft, ob dieser geringe Grad an fiskalischer Äquivalenz das Interesse der Programmgestalter an einem effektiven und effizienten Mitteleinsatz hinreichend absichert. Auch deshalb ist die Aufwertung der Phase der Evaluation im Policy-Zyklus und die feste Verankerung von Begleitung und Bewertung in der Politik der zweiten Säule der GAP zu begrüßen.

Vor diesem Hintergrund widmet sich der folgende Beitrag der Frage nach der praktischen Durchführung der Evaluationen im Bereich der Politik zur Förderung der Entwicklung ländlicher Räume. Dabei werden auf der Grundlage eigener Evaluationen und von Berichten anderer Evaluatoren die bisherigen Erfahrungen im Hinblick auf dabei aufgetre-

^{*} Bernhard Forstner, Regina Grajewski, Peter Mehl, Institut für Betriebswirtschaft, Agrarstruktur und ländliche Räume, Bundesforschungsanstalt für Landwirtschaft (FAL), Bundesallee 50, 38116 Braunschweig; e-mail: bernhard.forstner@fal.de.

ne Probleme dargestellt, auf ihre Ursachen hin analysiert und Vorschläge für eine verbesserte Bewertungspraxis vorgelegt. Hierzu werden zunächst in Kapitel 2 die Vorgaben der EU für die aktuelle Förderperiode und die von Seiten der EU-Kommission damit verbundenen Ziele umrissen. Kapitel 3 analysiert den Umstand, dass Evaluationen in der EU aus dem komplexen Zusammenwirken heterogener Stakeholder in einem Mehrebenensystem erwachsen und untersucht deren mögliche Auswirkungen auf die Evaluationspraxis. In Kapitel 4 werden dann die Erfahrungen (bezüglich Methodik, Datengrundlage, Ergebnisse) mit den bisher durchgeführten Bewertungen vor dem Hintergrund der in Kapitel 2 dargestellten Ansprüche an Evaluationen und den in Kapitel 3 formulierten besonderen Rahmenbedingungen eines Mehrebenensystems analysiert. Kapitel 5 zieht einen Vergleich mit der Erfolgskontrolle im Rahmen der Gemeinschaftsaufgabe "Verbesserung der Agrarstruktur und des Küstenschutzes" (GAK) vor der Agenda 2000 und bietet so die Gelegenheit, die oben dargestellten Bemühungen innerhalb der EU um Fortschritte in Richtung auf eine systematischere und regelmäßigerere Erfolgskontrolle einzuordnen. Vorschläge zur Verbesserung der Evaluationspraxis, die sich im Rahmen des bestehenden Systems bewegen, beschließen den Beitrag.

2 Evaluation der EU-Strukturpolitik

Seit der Reform der Strukturfonds im Jahr 1988 ist die Bewertung von Programmen und Maßnahmen in den Strukturfondsverordnungen vorgesehen. Gab es in der ersten Programmphase 1988 bis 1993 nur vereinzelte exemplarische Evaluationen, so wurde mit der zweiten Programmphase 1994 bis 1999 der Stellenwert von Bewertungen im EU-Kontext systematisch ausgebaut (vgl. ESER, 2001, S. 327). Dies hängt zusammen mit einer Initiative der EU-Kommission zur Förderung eines effizienten Finanzmanagements.¹ Bewertungsaktivitäten erfassen inzwischen alle Politikbereiche, die im Rahmen der sogenannten Zweiten Säule, d. h. der ehemals flankierenden Maßnahmen und der agrarstrukturellen Maßnahmen, umgesetzt werden. In der aktuellen Förderperiode 2000 bis 2006 wurden die Bewertungsaktivitäten gegenüber der Vorgängerperiode noch einmal deutlich ausgeweitet. Der formale Stellenwert der Bewertungen im Implementationsprozess wurde weiter gestärkt und die inhaltlichen Anforderungen wurden konkretisiert.

Die EU-Kommission verfolgt mit der Bewertung von Förderprogrammen und Fördermaßnahmen im Wesentlichen die folgenden drei Ziele (vgl. NAGARAJAN und VANHEUKELEN, 2002):

- Verbesserung des Managements der Programme,
- Rechenschaftspflicht gegenüber politischen Entscheidungsträgern und der Öffentlichkeit,
- Unterstützung bei der Budgetierung von Haushaltsmitteln.

¹ Kurz: SEM 2000 (Sound and Efficient Management).

Mit diesen Zielen bewegt sich die Evaluation von EU-Programmen im Spannungsfeld zwischen Rechenschaft nach außen gegenüber den Stakeholdern (summative Evaluation - Leistungsmessung) und der Suche nach Verbesserungsmöglichkeiten innerhalb der Organisation (formative Evaluation - Lernmedium) (vgl. ESER, 2001, S. 335). Korrespondierend mit dem Politikzyklus sind in der laufenden Programmperiode drei Evaluationstypen vorgesehen (siehe Tabelle 1): (a) Eine Ex-ante-Bewertung, die vor allem eine Abschätzung der relevanten Wirkungsindikatoren beinhaltet und als Grundlage für die Programmerrstellung und für nachfolgende Evaluationen dient, (b) eine Halbzeitbewertung, die mit ihrem überwiegend formativen Charakter besonders auf eine Unterstützung des Programmmanagements abzielen soll und (c) eine Ex-post-Bewertung, die in erster Linie summativ ist und Auskunft über Programm- oder Maßnahmenwirkungen der abgeschlossenen Förderperiode geben soll (vgl. TAYLOR et al., 2001, S. 341-343).²

Während die Ex-ante-Bewertung von der zuständigen Verwaltungsbehörde unter fakultativer Hinzuziehung von Experten durchgeführt wurde, ist bei der Halbzeit- und Ex-post-Bewertung zwingend eine externe Bewertung durch unabhängige Sachverständige vorgeschrieben. Die Bewerter sind bei der Erstellung von Bewertungen keineswegs frei. Die EU-Kommission hat in Zusammenarbeit mit den Mitgliedstaaten Mindeststandards für Bewertungen erarbeitet, von deren Einhaltung sie auch die Mitfinanzierung der Bewertungsstudien abhängig macht. Diese inhaltlichen Anforderungen für den Bereich der ländlichen Entwicklungsprogramme gemäß VO (EG) Nr. 1257/99 wurden in spezifischen Bewertungsleitfäden und in einem Set gemeinsamer Bewertungsfragen samt Kriterien und Indikatoren konkretisiert (KOM 1999, 2000 und 2002). Zur Unterstützung der Evaluation von sozio-ökonomischen Programmen gab die EU-Kommission des weiteren eine Reihe von Bewertungshandbüchern (MEANS collection; COM, 1999) heraus.

² In der laufenden Förderperiode 2000 bis 2006 wird die Halbzeitbewertung im Rahmen der Strukturförderung gemäß VO (EG) Nr. 1260/99 (beispielsweise der Ziel-1-Programme) auch dazu genutzt, um in geringem Umfang eine leistungsbezogene Zuweisung auf die besonders erfolgreichen Förderprogramme vorzunehmen. Hier nimmt die Halbzeitbewertung auch einen summativen Charakter an.

Tabelle 1: Wesentliche Aufgaben und Inhalte der Bewertungsstufen des Programmplanungszeitraums 2000 bis 2006

| Bewertung | Zeitpunkt | Inhaltliche Schwerpunkte | Verantwortlichkeit |
|--|-------------------------------------|---|--|
| Ex-ante-Bewertung | Mit Abgabe des EPLR/OP* (1999/2000) | <ul style="list-style-type: none"> – Voraussetzung für spätere (formative und summative) Bewertungen – Grundlage zur Ausarbeitung der Pläne (EPLR/OP*) – Stärken-Schwächen-Analyse – Kohärenz und Relevanz – Beschreibung der erwarteten wirtschaftlichen, ökologischen und sozialen Wirkungen – Quantifizierung der Ziele – Festlegung von Indikatoren zur Überprüfung der Zielerreichung | Für die Ausarbeitung des EPLR/OP* zuständige Behörde (i.d.R. Landesverwaltung); |
| Halbzeitbewertung | Spätestens Ende 2003 | <ul style="list-style-type: none"> – weitgehend formativer Charakter – Beurteilung der Verwendung der Finanzmittel und der Qualität der Begleitung und Durchführung – ggf. erforderl. Programmverbesserungen begründen – Darstellung erster Ergebnisse | Für die Abwicklung der EPLR zuständige Behörde (Landesverwaltung) in Zusammenarbeit mit der EU-Kommission; <u>Ziel-1-Gebiet</u> : ebenso, aber in Zusammenarbeit mit dem Mitgliedstaat |
| Up-date der Halbzeitbewertung (optional) | Spätestens Ende 2005 | Ebenso | Ebenso |
| Ex-post-Bewertung | Spätestens Ende 2008 | <ul style="list-style-type: none"> – weitgehend summativer Charakter – Darstellung und systematische Analyse der Programmwirkungen | Ebenso; <u>Ziel-1-Gebiet</u> : EU-Kommission in Zusammenarbeit mit dem Mitgliedstaat und der Verwaltungsbehörde |

* Entwicklungsplan Ländlicher Raum/Operationelles Programm.

Quelle: Eigene Darstellung.

3 Besonderheiten von Evaluationen im Mehrebenensystem der Europäischen Union und ihre Folgen

Nach der Typologie von MAYNTZ und SCHARPF (1973) zählt Evaluation zu den Informationsverarbeitungsprozessen innerhalb des Policy-Making im Unterschied zu Konflikt-Konsens-Prozessen. Da sich die Aussagen von Evaluationsforschung häufig auf Zielerreichungsgrad und Wirkungen einer Politik beziehen, sind ihre Aussagen von potenzieller

Virulenz für Entscheidungen über die Weiterführung, Novellierung oder Beendigung einer politischen Maßnahme. Allerdings kann Informationsverarbeitung (wie Analyse, Prognose, Planung und Evaluation) vom politischen Prozess nur in dem Maße genutzt werden, wie die Konflikt- oder Konsensprozesse dies erlauben oder fordern. Für die Evaluation bedeutet dies, dass vom politischen Umfeld erhebliche Rückwirkungen auf die Bedingungen einer Evaluation und auf das Verhalten der involvierten Akteure zu erwarten sind. Dies gilt insbesondere dann, wenn die Betroffenen und Beteiligten einer Evaluation (Stakeholder) annehmen, dass Evaluationsergebnisse auf das weitere Programm-schicksal Einfluss nehmen.

Akteure, Institutionen und Entscheidungsprozeduren innerhalb der EU unterscheiden sich von denen in föderativen oder unitarischen Nationalstaaten. Die Vermutung scheint daher naheliegend, dass die spezifischen Bedingungen der Politikformulierung, Implementation und deren Evaluation im Bereich der Politik der EU zur Förderung der Entwicklung ländlicher Räume zur Erklärung der bisherigen Evaluationspraxis beitragen können. Im Folgenden werden daher Besonderheiten von Evaluationen im Mehrebenensystem der Europäischen Union dargestellt und einige Hypothesen über deren potenzielle Auswirkungen auf die Evaluationspraxis formuliert.

3.1 Besonderheiten von Evaluationen im Mehrebenensystem der Europäischen Union

Im politik- oder verwaltungswissenschaftlichen Bereich ist Evaluation der Teil des Politikzyklus, der auf die Phasen Problemdefinition, Agenda-Setting, Politikformulierung, Implementation folgt und in Termination bzw. Weiterführung oder Neuformulierung mündet (vgl. z. B. SCHUBERT, 1991). Evaluation befasst sich vorrangig mit dem Zielerreichungsgrad, den Wirkungen und der Effizienz von Politik und den sie bedingenden Faktoren und ist der Politikformulierung und -umsetzung nachgelagert.

Ein charakteristisches Merkmal von Evaluation im Rahmen der EU-Strukturpolitik besteht darin, dass diese nicht erst nach der Implementation der Programme, sondern quasi begleitend im Politikzyklus einsetzt: zunächst bei der Ex-ante-Evaluation als integraler Bestandteil der Programmentwicklung, dann bei der Halbzeitbewertung nach zwei- bis dreijähriger Laufzeit des Programms (bzw. als „up date“ zwei Jahre später) und schließlich als Ex-post-Evaluation nach Ablauf des Programmzeitraums in 2006. Daher verläuft die Evaluation im Rahmen der EU-Strukturförderung parallel zur Ausfüllung der Rahmenvorgaben, die zuvor auf europäischer Ebene im Hinblick auf Finanzausstattung, Mittelverteilung der Strukturfonds und Förderinhalte festgelegt wurden, durch die entsprechenden Programme der Bundesländer.

Besonderheiten sind auch in Bezug auf die Stakeholder festzuhalten. Diese sind zu identifizieren, damit ihre Interessen berücksichtigt werden können (WIDMER, 1996, S.15).³

³ "In undertaking their work, evaluators usually find themselves confronted with individuals and groups who hold competing and sometimes combative views on the appropriateness of either the programme

Gängige Unterteilungen unterscheiden beim Kreis der Stakeholder von Evaluationen a) die Entscheidungsträger und Programmgestalter, b) Programmfinanzierende, c) Evaluationsfinanzierende, d) Zielgruppen des Programms, e) Programmverantwortliche, f) Evaluatoren, g) deren Mitbewerber und schließlich h) die Evaluationsgemeinde. Eine Untersuchung von Stakeholdern umfasst die jeweiligen Interessenlagen und Handlungsrechte der Akteursgruppen in Bezug auf das zu evaluierende Programm und dessen Evaluation sowie deren Beziehungs- und Interaktionsmuster untereinander.

Bezogen auf die EU-Strukturpolitik können Besonderheiten im Vergleich zu Evaluationen in einem unitarischen politischen System im Besonderen für die unter a) bis c) genannten Stakeholder unterstellt werden.⁴ Zentrales Kennzeichen der Stakeholder im Rahmen der EU-Strukturpolitik für ländliche Räume und deren Evaluation ist der Umstand, dass Formulierung und Umsetzung dieser Politik ebenso wie ihre Bewertung und Neuformulierung bzw. Terminierung Interaktionen in einem Mehrebenensystem darstellen (MARKS, 1993, S. 392). Entscheidungsfindung, Finanzierung, Implementation, Evaluation und Neuformulierung der EU-Strukturpolitik für ländliche Räume werden im Zusammenwirken von Akteuren auf supranationaler, mitgliedstaatlicher, regionaler und lokaler Ebene entschieden. Politik in einem Mehrebenensystem ist komplexer, intransparenter und aufgrund der Verantwortlichkeit mehrerer Ebenen schlechter zurechenbar (SUMMA, 1998, S. 3). Worin unterscheiden sich nun Stakeholder einer Evaluation im Rahmen der multi-level-Governance europäischer Strukturpolitik von Stakeholdern im Rahmen eines unitarischen Nationalstaats?

Auch bei Evaluationen im Rahmen eines unitarischen Nationalstaats kann nicht von homogenen Interessen der Akteure ausgegangen werden; hier kommen insbesondere ressortbezogene Interessenunterschiede zum Tragen. Im Mehrebenensystem kommen allerdings zu sektoralen noch ebenenbezogene Interessenunterschiede und Konfliktlinien hinzu:

- Sektorbezogene Interessenunterschiede: Bei Akteuren auf den gleichen politischen Ebenen kann grob unterschieden werden zwischen allgemeinen Akteuren mit Interessen an effizienter und effektiver Mittelverwendung (etwa Haushaltsbehörde) und Akteuren mit politikbereichsbezogenen Interessen. Bei Letzteren dürften zum Interesse an effizienter und effektiver Mittelverwendung auch sektorbezogene Interessen am Fortbestand der von ihnen entwickelten Maßnahmen und den hierfür bereitgestellten Mitteln kommen.⁵
- Ebenenbezogene Interessenunterschiede: Im Mehrebenensystem kann diese sektorbezogene Konfliktlinie überlagert werden durch Interessenunterschiede zwischen den Ebenen. Insbesondere bei Mischfinanzierungssystemen kann das oben angenommene Interesse der allgemein-politischen Akteure an effizienter und effektiver Mittelver-

or its evaluation, and whose interest will be affected by the outcome. To conduct their work effectively and contribute to the resolution of the issues at hand, evaluators must understand their relationship with the stakeholders involved as well as the relationships between stakeholders." (ROSSI und FREEMANN, 1993, S. 110).

⁴ Parallelen bestehen hingegen zur Politikverflechtung im Rahmen des deutschen kooperativen Föderalismus. (vgl. z. B. SCHARPF, 1985).

⁵ SUMMA (1998, S. 4) bestätigt diese Zusammenhänge auch für die EU-Kommission.

wendung überlagert werden vom Interesse, das der eigenen politischen Ebene zugeordnete Finanzvolumen zu erhalten oder gar auszuweiten. Diese Verschiebung der Interessenposition dürfte mit der Höhe des eigenen Kofinanzierungsanteils negativ korrelieren. Bezogen auf die zweite Säule der gemeinsamen Agrarpolitik können folgende Interessenprofile der beteiligten Ebenen festgestellt werden: Das Handeln der Bundesländer zielt darauf ab, landesbezogene Präferenzen und Ausgestaltungsanliegen unter gleichzeitiger maximaler Realisierung der Kofinanzierungsmöglichkeiten von Seiten der EU und des Bundes verwirklichen zu können (vgl. MEHL und PLANKL, 2002).⁶ Der Bund hat ein Interesse an Kohärenz und Konsistenz des Förderangebots innerhalb des Mitgliedstaates, an der Wahrung zentralstaatlicher Steuerungsanliegen sowie ein Interesse an einer Verbesserung der Zahlerposition durch Einsparungen oder einer Maximierung des eigenen Anteils an den Fördergeldern der EU. Dagegen kann der EU-Ebene ein Interesse an Rechenschaft nach außen über summative Evaluationen unterstellt werden (vgl. ESER, 2000, S. 328).

3.2 Vermutete Auswirkungen auf die Evaluationspraxis

Diese Besonderheiten im Bereich der Politik der EU zur Förderung der Entwicklung ländlicher Räume lassen vermuten, dass die mit den Evaluationen verbundenen Ziele und Erwartungen nur schwer umzusetzen sein werden. Diese Auffassung wird in der Literatur gestützt.

- (1) Bei den Evaluationen im Rahmen der EU-Politik zur Förderung der Entwicklung ländlicher Räume handelt es sich um Auftragsevaluationen. MANN (2000) beleuchtet auf der Grundlage ökonomischer Theorien der Bürokratie die Motive von Programmverantwortlichen, ihre Programme evaluieren zu lassen. Er unterscheidet dabei Eigenevaluation von Evaluationen, die von Seiten Dritter veranlasst werden, wie etwa die obligatorischen Evaluationen der Programme im Rahmen der EU-Strukturpolitik. Wird eine Evaluation von den Programmverantwortlichen selbst angestoßen, sei davon auszugehen, dass diese sich von einer Evaluation Nutzen für den Fortbestand des Programms versprechen. Etwa, indem die Evaluation die Bedeutung des mit dem Programm zu lösenden Problems und dessen Eignung, dieses tatsächlich auch zu tun, betone. Auch bei einer obligatorischen Evaluation würden die Programmverantwortlichen versuchen, die Evaluation in eine Richtung zu beeinflussen, die ihrer Interessenlage entspreche. Die entsprechenden Möglichkeiten hingen stark von den genauen Umständen ab. Generell könnten drei Möglichkeiten angenommen werden: a) die Auswahl des Evaluators; b) eine selektive Informationspolitik gegenüber dem Evaluator und c) die Option, Druck auf den Evaluator auszuüben. Letztes könne bedeuten, dass Folgeverträge für Auftragnehmer von für den Auftraggeber zufriedenstellenden Ergebnissen der laufenden Evaluation abhängig gemacht würden. Sowohl bei fakultativen als auch bei obligatorischen Evaluationen

⁶ Zum Pochen auf ein eigenständiges Länderprofil gehört auch die Kritik der Länder an den mit der Agenda 2000 ihnen auferlegten Monitoring- und Berichtspflichten (vgl. etwa die Beiträge von ALTMANN und WILSTACKE in FAA, 2000).

sei demnach damit zu rechnen, dass das Resultat von Auftragsevaluationen zugunsten der Sichtweisen der Programmverantwortlichen verzerrt werde.⁷

- (2) Bezogen auf die Agrarumweltprogramme der EU und die Förderphase 1994 bis 1999 bestätigen AHRENS et al. (2000, S. 11) die Sichtweise MANNs. Sie verweisen auf die starke Stellung der Bundesländer als Hauptnutznießer der EU-Kofinanzierung und als zugleich verantwortliche Ebene für die Evaluationen. Nach Artikel 16 der VO (EG) Nr. 746/96 haben die Mitgliedstaaten für die Bewertung der Agrarumweltprogramme zu sorgen und die Kommission über die Ergebnisse der Bewertung zu unterrichten. Üblich sei die Praxis, Evaluationsberichte durch den Hauptgeldgeber bei unabhängigen externen Gutachtern zu vergeben. Hier würden dagegen die Hauptempfänger des Mittelzuflusses beauftragt, eine Evaluation zu veranlassen und über deren Ergebnisse zu berichten. Sie seien somit quasi aufgefordert, sich selbst zu evaluieren, wobei nur bei positivem Ergebnis das Weiterfließen der Mittel aus Brüssel in Aussicht stünde. Dies müsse "wie eine indirekte Aufforderung zur Ablieferung eines besonders positiven Berichts wirken" (ebenda, S. 111). Ein regionaler externer Bewerter müsse zudem annehmen, dass alle anderen Mitgliedstaaten bzw. Regionen aufgrund dieser geschilderten Zusammenhänge besonders positive Berichte abliefern. Auf ihm laste ein Erwartungsdruck, die Region nicht im Stich zu lassen (ebenda, S. 112):
- (3) Schließlich ist zu berücksichtigen, dass es sich bei der Zusammenarbeit in der EU im Wesentlichen um die vertraglich fixierte Kooperation von Nationalstaaten handelt. Übereinstimmend wird im Völkerrecht wie bei der Theorie der internationalen Beziehungen angenommen, dass gesetzliche Regelungen in einem Nationalstaat ein höheres Maß an Folgebereitschaft aufweisen als internationale Vereinbarungen (vgl. WOLF, 1999). Nationalstaatliches Recht sei zwangsläufig deshalb wirksamer als internationales Recht, weil hinter ihm die hierarchische Durchsetzungsgewalt des Nationalstaates stehe (Kelsen, 1966). Wenn nationale Regelungen einen höheren Grad an Regelbefolgung (Compliance) aufweisen als internationale Abkommen oder Verträge, dann liegt in Bezug auf die Evaluation der Strukturpolitik die Vermutung nahe, dass die EU-Kommission als Principal weniger stark in der Lage sein wird, die Befolgung seiner Vorgaben zu garantieren, als dies in einem nationalstaatlichen Kontext der Fall wäre.

4 Bewertungsaktivitäten in der Praxis

Im folgenden Abschnitt werden Problembereiche dargestellt, die bei den derzeit laufenden Halbzeitbewertungen und auch schon bei vorhergehenden Bewertungen aufgetreten

⁷ MANN (2000) testet diese Hypothese am Beispiel von Evaluationen im Bereich der EU-Strukturpolitik in Deutschland zwischen 1990 und 1999 und kommt zu dem Ergebnis, dass in Auftrag gegebene Evaluationen im Durchschnitt ein deutlich besseres Ergebnis erbringen als Evaluationen ohne Auftrag. Seine Empfehlung: Die Evaluation politischer Programme sollte in andere Hände gelegt werden als das Programm selbst, z. B. in die der Rechnungshöfe.

sind. Diese basieren auf Erfahrungen aus selbst durchgeführten Evaluationsprojekten⁸ sowie auf Berichten anderer Evaluatoren (z.B. TOEPEL, 2000; RIEDEL, 1997; HEINTEL, 2002). Da nicht davon ausgegangen werden kann, dass alle bei bisherigen Evaluationen beobachteten Probleme auf das komplexe Zusammenwirken heterogener Stakeholder in einem Mehrebenensystem zurückzuführen sein werden, wird nach drei Problemkategorien differenziert.

- (1) Probleme, die generell bei Evaluationen auftauchen, unabhängig von Politikfeld, Akteuren und Spielregeln. Hierzu zählen die Schwierigkeit, kausale Beziehungen zwischen einer Maßnahme und den beobachteten Effekten zu überprüfen, eine oftmals unzureichende Datengrundlage sowie die nur selten gegebene Vergleichbarkeit und Kumulierbarkeit von Evaluationsergebnissen (vgl. TOEPEL, 2000, S. 400).
- (2) Probleme, die den Eigenarten des zu evaluierenden Politikfeldes zuzuschreiben sind, also etwa auch in einem unitarischen Nationalstaat erwartet werden könnten. Dazu zählen beispielsweise der in der Agrarstrukturpolitik häufige Einsatz von Maßnahmenbündeln sowie deren meistens erst mittel- bis langfristig feststellbare Wirkungen.
- (3) Probleme, die auf die in Kapitel 3 umrissenen spezifischen politics- und polity-bezogenen Rahmenbedingungen der EU-Politik zur Förderung der ländlichen Entwicklung zurückzuführen sind. Dazu zählen zum einen a) konfligierende Interessen der Stakeholder und zweitens b) die den jeweiligen Akteuren zugewiesenen Aufgaben und Handlungsrechte bei der Evaluation.

Im Folgenden werden ausgewählte technische und methodische Aspekte (Datenerfassung und -bereitstellung, Festlegung von Zielen und Indikatoren, Wirkungsmessung), das Verhältnis zwischen Auftraggeber und Evaluator, die finanziellen und personellen Rahmenbedingungen einer Evaluation sowie Fragen der Offenlegung und Klarheit von Ergebnissen beschrieben. Diskutiert werden diese Aspekte einer Evaluation vor dem Hintergrund der eben beschriebenen drei Problemkategorien.

⁸ Das Institut für Betriebswirtschaft, Agrarstruktur und ländliche Räume (BAL) der Bundesforschungsanstalt für Landwirtschaft (FAL) verfügt über mehrjährige Evaluationserfahrung im Bereich der EU-Strukturpolitik. Dies bezieht sich sowohl auf methodische Arbeiten (z.B. PLANKL und SCHRADER, 1991; TISSEN, 1997) wie auch die Durchführung einer ganzen Reihe von Bewertungsstudien im Auftrag der EU-Kommission, des Bundes und der Länder. Derzeit ist das BAL in die Halbzeitbewertung der Programme zur Entwicklung des ländlichen Raums in Deutschland gemäß VO (EG) Nr. 1257/99 eingebunden. Beispielsweise werden die Programme der Länder Schleswig-Holstein, Niedersachsen, Hessen, Nordrhein-Westfalen, Bremen und Hamburg bewertet. Darüber hinaus wird zentral die Halbzeitbewertung einzelner Maßnahmen (Agrarinvestitionsförderungsprogramm, Ausgleichszulage) für alle Länder in Deutschland durchgeführt.

4.1 Ausgewählte Aspekte der Durchführung von Bewertungen im EU-Kontext

4.1.1 Defizite bei der Datenerfassung und -bereitstellung

Die Erfassung von Förderdaten stellt eine wesentliche Voraussetzung für die Darstellung des Vollzugs von Maßnahmen dar, schafft aber auch die Basis für die Kontrolle und Bewertung der Zielerreichung, der Wirkungen und Effizienz durchgeführter Fördermaßnahmen. Hierfür hat die EU-Kommission Vorschriften erlassen. Beispielsweise müssen die Mitgliedstaaten der EU-Kommission jährlich einen Lagebericht vorlegen, in dem sie u. a. anhand spezifischer materieller und finanzieller Indikatoren (Monitoringdaten), die vorher im Konsens festgelegt wurden, den Durchführungsstand des Förderprogramms dokumentieren. Hierfür wurde von der EU-Kommission zwar ein Monitoringsystem installiert, dieses liefert aber – entgegen den gesetzten Ansprüchen – kaum Anknüpfungspunkte für die Evaluation (vgl. GRAJEWSKI und KOCH, 2002).

Insbesondere die Bewertungsfragen der EU-Kommission und die damit verbundenen Kriterien und Indikatoren machen eine umfangreiche Datensammlung auf der Ebene der Fördermittelempfänger (Individualdatensätze) oder Projekte (Projektdatensätze) erforderlich. Die EU-Kommission verweist in diesem Kontext darauf, dass zur Beantwortung der Bewertungsfragen aus Kostengründen möglichst auf Sekundärdaten zurückgegriffen und nur ergänzend selbst Daten erhoben werden sollten (KOM 1999, S. 21).

Ergebnis und Empfehlung aller bislang durchgeführten Bewertungen im EU-Kontext war, ein konsistentes, rechnergestütztes Datenerfassungssystem als Grundlage für die Programmsteuerung und –evaluation aufzubauen. Dies beruht auch auf der Erfahrung, dass bisher allein die Darstellung des Vollzugs der Maßnahmen sehr zeitaufwändig war und für die eigentliche Aufgabe der Bewertung, nämlich die Beurteilung der Zielerreichung, der Wirkungen und Effizienz, häufig die Zeit fehlte.

Bislang hat sich in dieser Hinsicht jedoch wenig geändert. In der Praxis der Halbzeitbewertung der laufenden ländlichen Entwicklungsprogramme ergeben sich im Zusammenhang mit der Datenerfassung und -bereitstellung für die Bewertung weiterhin erhebliche Probleme. Relevante Sekundärdaten liegen häufig nur auf Papier vor. Falls Einzelfalldaten in digitalisierter Form zur Verfügung gestellt werden, mangelt es ihnen teilweise an einer einheitlichen Struktur, so dass diese nur mit großem Zeitaufwand in verarbeitbare Datenformate übergeführt werden können. Professionell erstellte Datenbanken stellen eher eine Ausnahme dar. Die in der Praxis erfassten Einzelfalldaten sind häufig recht lückenhaft und im Variablenumfang begrenzt; folglich werden zeit- und kostenträchtige Nacherhebungen notwendig.

Bei der Suche nach den Gründen für die bestehenden Datenprobleme stößt man zum einen auf die Besonderheiten des Politikfeldes und zum anderen auf die bestehenden Spielregeln. Die verschiedenen, im Bereich der Politik zur Förderung des ländlichen Raums, relevanten politischen Ebenen führen gemeinsam Maßnahmen durch, flankieren diese Maßnahmen aber auch mit eigenen, rein national finanzierten Förderprogrammen. Diese Vielfalt bringt Monitoringsysteme mit sich, die untereinander bislang nur unzureichend

abgestimmt sind. Zu nennen sind hier beispielsweise die Agrarstrukturberichterstattung im Rahmen der GAK und das EU-Monitoringsystem.

Bei gleichzeitig existierender Geld- und Personalknappheit führt dies bei dem auf der untersten Ebene zuständigen Verwaltungspersonal zu viel Unverständnis angesichts der zu erfassenden Datenflut und der geringen Datenkonsistenz. Mangelnde Motivation und Kooperation bei der zusätzlich notwendigen Datenerfassung jenseits der Begleitsysteme sind die natürliche Folge. Die Verwaltung selbst nutzt die umfangreichen Datensammlungen bislang nur unzureichend für die Programmsteuerung, da diese nicht in der Form vorliegen, die den Ansprüchen einer professionellen Programmsteuerung genügen würden.

4.1.2 Unzureichende Ex-ante-Bewertung

Bei EU-kofinanzierten Programmplanungsansätzen ist generell eine Ex-ante-Bewertung zu erstellen (siehe Tabelle 1). Die EU-Kommission beschreibt die Ex-ante-Bewertung als einen interaktiven Prozess, bei dem gegebenenfalls zu Rate gezogene Experten unabhängig von den Verwaltungsbehörden Politik- oder Programmfragen beurteilen und Empfehlungen aussprechen können (KOM, 1999 S. 25). Ein konstruktiver Dialog zwischen Bewertern und Verwaltungsbehörden soll zur Qualitätsverbesserung der Entwicklungspläne beitragen.

Insbesondere die Erfahrungen der Programmerstellung der Förderperiode 2000 bis 2006 zeigen, dass die Qualität der Ex-ante-Bewertungen in sehr vielen Fällen unter dem erheblichen Termindruck bei der Erstellung der Entwicklungspläne litt. Mitverursacht wurde dies dadurch, dass die relevanten Unterlagen zur Durchführung einer Ex-ante-Bewertung den Programmverantwortlichen von der EU-Kommission teilweise erst im Laufe der Programmerstellung an die Hand gegeben wurden. Die Entwicklungspläne enthalten i. d. R. keine Hinweise auf konkrete Ergebnisse der Ex-ante-Bewertung, da umfangreichere Analysen meistens nicht durchgeführt wurden bzw. werden konnten. Die Programmplanungsdokumente waren durch den Dialog zwischen Mitgliedstaat und EU-Kommission ständigen Änderungen unterworfen, so dass es schwierig bis unmöglich für einen Evaluator war, eine geeignete Grundlage für die Ex-ante-Bewertung zu fixieren. Dies ist auch Ausdruck einer fehlenden Konkretisierung der Anforderungen an eine interaktive, prozessorientierte Zusammenarbeit zwischen Evaluator und programmerstellender Behörde. Steht das formative Element einer Evaluation im Vordergrund, d. h. die Initiierung von Lernprozessen und die kontinuierliche Verbesserung der Programmplanungsdokumente, müsste zum einen mehr Zeit für die Programmerstellung eingeräumt werden und zum anderen müsste sich der Evaluator stärker als Moderator verstehen.

Im Genehmigungsverfahren selbst hat die EU-Kommission lediglich formal das Vorhandensein einer Ex-ante-Bewertung geprüft; inhaltliche Anforderungen an eine Ex-ante-Bewertung waren dagegen nachrangig.

Im Bewertungsleitfaden der EU-KOMMISSION (1999, Anhang III) sind die einzelnen Elemente einer Ex-ante-Bewertung vorgegeben. Vergleicht man die hier enthaltenen inhaltlichen Anforderungen mit den Erfahrungen der laufenden Halbzeitbewertung, so treten

deutliche Defizite zu Tage. So fehlen beispielsweise häufig in den Programmplanungsdokumenten konkrete Zielangaben für die wesentlichen Wirkungsindikatoren. Eine Quantifizierung wird zumeist nur bei Input- und Outputindikatoren vorgenommen. Dies liegt auch an dem geringen Interesse sowohl der politischen Entscheidungsebene als auch der Verwaltungsebene an Zielformulierungen, die ein wichtiger Bestandteil einer jeden Ex-ante-Bewertung sind. Eine prospektive Festlegung auf ein Punktziel wird von Seiten der Verwaltungsbehörde meistens auch deshalb gern vermieden, weil eine deutliche Zielverfehlung später angreifbar macht (vgl. SCHUBERT, 2001, S. 317 ff.). Diese Vorbehalte der Verwaltungsbehörden der Länder im Rahmen einer formativen Evaluation abzubauen, ist aufgrund der Zeitknappheit und der unklar definierten Position des Evaluators nur begrenzt gelungen.

Ziele und Inhalte der Ex-ante-Bewertung sind in den Dokumenten der EU-Kommission weitgehend klar definiert sind. Es handelt sich somit um ein Vollzugsdefizit, das sowohl aus den zeitlichen Engpässen bei der Programmerstellung (Programmverantwortliche) und Programmgenehmigung (EU-Kommission) als auch aus Vorbehalten der Programmverantwortlichen gegenüber strategischen und zielbezogenen Festlegungen resultiert. Der von der EU-Kommission im Zusammenhang mit der Ex-ante-Bewertung eingeführte Begriff des interaktiven Prozesses eröffnet der praktischen Umsetzung der Ex-ante-Bewertung einen erheblichen Spielraum, der wenig zur qualitativen Verbesserung der Ex-ante-Bewertung beigetragen hat.

4.1.3 Methodische Probleme bei der Wirkungsmessung

Die EU-Kommission geht in ihren methodischen Leitfäden davon aus, dass Netto- und nicht Bruttowirkungen der EU-finanzierten Programme zu erfassen sind. Das heißt, die Wirkungen sind zu isolieren und es sind Mitnahme-, Substitutions- und Verdrängungseffekte abzuschätzen. Da es sich um einen Programmplanungsansatz und nicht um isolierte Einzelmaßnahmen handelt, ist zudem der Mehrwert dieses Ansatzes darzustellen. Bei der Wirkungsmessung ist auch die Langfristigkeit von Strukturmaßnahmen zu berücksichtigen. Je länger der Zeitraum zwischen Umsetzung der Maßnahmen und dem Eintritt von Wirkungen ist, desto stärker schlagen Überlagerungseffekte zu Buche.

Isolierung der Programm-/Maßnahmenwirkungen

Agrarstruktur- und Agrarumweltmaßnahmen, die im Rahmen der ländlichen Entwicklungsprogramme angeboten werden, zeichnen sich dadurch aus, dass sie im Rahmen unterschiedlicher Mischfinanzierungssysteme durchgeführt werden. Im Rahmen der Verordnung (EG) Nr. 1257/99 werden EU-Mittel mit Bundes-, Landes- und kommunalen Mitteln kombiniert. Das gleiche Maßnahmenspektrum der ländlichen Entwicklungsprogramme gemäß VO (EG) Nr. 1257/99 kann aber auch rein national finanziert durchgeführt werden (sog. Artikel 52-Maßnahmen). Daneben gibt es allein aus den Landeshaushalten finanzierte Staatsbeihilfen, die sich aber in ihren Zielsetzungen teilweise nur wenig von den ländlichen Entwicklungsmaßnahmen unterscheiden. Den Beitrag einzelner Maßnahmen und Akteure in einer Evaluation zu isolieren wird dadurch schwierig, wenn nicht ganz unmöglich.

Netto- versus Bruttowirkungen

Bei den ländlichen Entwicklungsprogrammen gemäß der VO (EG) 1257/99 handelt es sich - bezogen auf das Finanzvolumen - nur um einen kleinen Politikbereich im ländlichen Raum⁹, so dass andere Politiken bzw. exogene Rahmenbedingungen Effekte auf hochaggregierte Größen wie Einkommen oder Beschäftigung überlagern (siehe unten). Makroorientierte Wirkungsanalysen (z. B. mit Hilfe von Modellen¹⁰) sind demnach nur ein sehr beschränkt geeignetes methodisches Instrument. Im Vordergrund stehen stattdessen Wirkungsanalysen mit Hilfe eines Bündels von Einzelindikatoren. Solche Analysen stützen sich methodisch v. a. auf Befragungen und Fallstudien. Auf Basis von Einzelindikatoren können jedoch Hinweise auf Mitnahme-, Verlagerungs- und Verdrängungseffekte nur in Einzelfällen gewonnen werden; dies gilt ebenso für Multiplikatoreffekte. Demnach muss von dem hohen theoretischen Anspruch einer Erfassung von Nettowirkungen abgewichen werden; abgebildet werden können zumeist nur Bruttowirkungen (TOEPEL, 2000). Dies gilt umso mehr, als es für eine Vielzahl von Maßnahmen schwierig, wenn nicht sogar unmöglich ist, geeignete Vergleichsgruppen oder -regionen zu finden.

Referenzgruppen

In den Leitlinien der Kommission wird als Technik zur Messung von Wirkungen von Förderprogrammen oder -maßnahmen auf den Vergleich geförderter Gruppen (Unternehmen, Dörfer, Regionen etc.) mit einer Referenzgruppe (= Null- oder Vergleichsgruppe ohne Förderung) hingewiesen. In der Evaluationspraxis zeigt sich jedoch, dass Referenzgruppen häufig nicht existieren oder nur mit erheblichen Einschränkungen zusammengestellt werden können. Ein Beispiel aus dem Bereich der einzelbetrieblichen Investitionsförderung in der Landwirtschaft möge dies illustrieren: Im Milchviehbereich, in dem ein Großteil der Investitionsfördermittel eingesetzt wird, wird fast jede größere Stallbaumaßnahme öffentlich gefördert. Nicht geförderte Maßnahmen scheidern als Vergleichsobjekte aus, da es sich in diesen Fällen meistens um Betriebe handelt, die aufgrund ihrer Einkommenssituation (Prosperitätsgrenze) oder völlig unzureichender ökonomischer Leistungen von der Förderung ausgeschlossen werden und daher nicht mit den geförderten Betrieben vergleichbar sind. Ähnliche Probleme sind aus den Bereichen Ausgleichszulage, Agrarumweltmaßnahmen und Marktstrukturverbesserung bekannt.

Das Problem fehlender Referenzgruppen ist vor allem der Spezifik des Politikbereichs zuzuordnen, in dem bei vielen Maßnahmen (-bereichen) nicht selektiv gefördert wird.

Langfristigkeit der Strukturförderung

⁹ So fließen beispielsweise 10 % der EAGFL-Garantie-Mittel in die ländliche Entwicklung. Die übrigen 90 % werden im Rahmen der Gemeinsamen Agrarpolitik für Flächen- und Tierprämien, Exporterstattungen usw. aufgewendet.

¹⁰ Makroorientierte Ansätze wurden beispielsweise in den Evaluationen des irischen Ziel-1-Programms (ECONOMIC AND SOCIAL RESEARCH INSTITUTE, 1997) und der neuen Bundesländer (RIEDEL et al., 1997) eingesetzt.

Ein wesentliches Kennzeichen der Strukturförderung ist ihre mittel- bis langfristige Ausrichtung. Die Wirkungen sind i. d. R. nicht kurzfristig erkennbar, sondern manifestieren sich mitunter erst nach mehreren Jahren. Folglich treten Schwierigkeiten auf, wenn – wie von der EU-Kommission in den Verordnungen festgelegt – Strukturfördermaßnahmen einer Programmperiode (z. B. 1994 bis 1999, 2000 bis 2006) bereits innerhalb oder unmittelbar zum Ende dieser Periode bewertet werden sollen.

Am Beispiel der Einzelbetrieblichen Investitionsförderung wird deutlich, welche Probleme hiermit verbunden sind. Die ersten Jahre nach Abschluss einer umfangreichen baulichen Investition, die i. d. R. vom Beginn bis zum Abschluss ein bis zwei Jahre in Anspruch nimmt, sollten aufgrund der mit der Investition verbundenen Umstellungs- und Anpassungsprobleme keinesfalls in die Erfassung der Förderwirkungen einbezogen werden. Dies sieht auch die EU-Kommission so, wenn sie in ihren Bewertungsleitlinien vorschreibt, nur die Wirkungen von Förderfällen zu bewerten, deren Investitionsabschluss mindestens zwei Jahre zurückliegt (KOM, 2000). Zusätzliche Schwierigkeiten erwachsen daraus, dass die Jahresabschlüsse der geförderten Betriebe, aus denen die wirtschaftliche und teilweise auch die strukturelle Entwicklung ersichtlich ist, erst mit einer zeitlichen Verzögerung von rund einem Jahr zur Rechnungsperiode verfügbar sind. Zudem sind für eine belastbare Aussage zur wirtschaftlichen Entwicklung eines Betriebe mindestens drei aufeinanderfolgende Jahresabschlüsse¹¹ notwendig. Folglich sollten bei der Zwischen- und Endbewertung der Förderperiode 2000 bis 2006 die Wirkungen der im Förderzeitraum 1994 bis 1999 geförderten Betriebe analysiert werden, um belastbare Ergebnisse zu erzielen.

Die bestehenden Probleme resultieren vor allem aus den Vorschriften zur Halbzeit- und Ex-post-Bewertung, die im Wesentlichen eine Bewertung der im Rahmen der Förderperiode durchgeführten Maßnahmen fordern. Für die Bewertung von Strukturfördermaßnahmen sind diese Vorschriften nicht sachdienlich. Diese Vorschriften werden von den Akteuren mitgestaltet, die teilweise in erster Linie schnelle Erfolge von Fördermaßnahmen darstellen wollen.

Programmplanungsansatz

Aufgrund des Programmplanungsansatzes nimmt die Beurteilung des Zusammenwirkens von Maßnahmenbündeln und der in diesem Zusammenhang beabsichtigten Synergieeffekte an Bedeutung zu. Gerade die Erreichung solcher (positiver) Synergieeffekte ist der eigentliche Grund dieses Ansatzes (COM, 1999, S. 73). Untersucht werden müssen nicht nur interne Synergien (innerhalb eines Programms), sondern auch externe Synergien (zwischen verschiedenen Programmen). Für die in den Programmen der VO (EG) Nr. 1257/99 umgesetzten Maßnahmen müssen methodische Ansätze zur Analyse von Synergiewirkungen (weiter-) entwickelt werden, da sich die bisher durchgeführten Evaluationen im Rahmen des Ziels 5a und der Agrarumweltmaßnahmen bislang fast ausschließlich auf die Maßnahmenwirkungen konzentriert haben, ohne Interdependenzen mit

¹¹ Bei stark fluktuierenden Märkten (z. B. für Schweine) werden sogar fünf Jahresabschlüsse empfohlen.

anderen Maßnahmen zu berücksichtigen.¹² Nur in den Ziel-5b- und Ziel-1-Evaluationen wurde versucht, Synergiewirkungen mit Hilfe eines qualitativen Ansatzes darzustellen (z.B. FÄHRMANN et al., 2001, RIEDEL et al., 1997, S. 396 ff.).

Neben den Synergiewirkungen sind für das Gesamtprogramm thematisch ausgerichtete maßnahmenübergreifende Querschnittsfragen zu beantworten, die u. a. die Lebensbedingungen und die Struktur der Bevölkerung im ländlichen Raum, die Beschäftigung und die Einkommen aus der landwirtschaftlichen und nichtlandwirtschaftlichen Tätigkeit, die Wettbewerbsfähigkeit land- und forstwirtschaftlicher Grunderzeugnisse und Umweltaspekte betreffen. Allerdings kommen mit diesen Querschnittsfragen z. T. neue Zielkategorien in die maßnahmenbezogenen Fragestellungen. So wird beispielsweise im Kapitel "Agrarumwelt" der Schwerpunkt auf die Frage nach den Ressourcenschutzwirkungen gelegt. Beschäftigungs- und Einkommenswirkungen spielen im kapitelbezogenen Ziel-Wirkungssystem keine Rolle. Mit jeder neuer Zielkategorie müssen aber neue Wirkungsbeziehungen geprüft werden. Durch die möglichen Interdependenzen wird der Analyseaufwand überproportional erhöht. Die Gewichtung der Ergebnisse wird schwierig, da die Querschnittziele nicht konsequent in das kapitelbezogene Zielsystem einbezogen sind. Wie gewichtet man beispielsweise eine Agrarumweltmaßnahme, die zwar ihre ökologischen Zielsetzungen erreicht, aber negative Beschäftigungswirkungen erzielt?

Die dargestellten methodischen Schwierigkeiten der Wirkungsmessung und -zuordnung sind wiederholt in den durchgeführten Evaluationsstudien und auch während der MEANS-Tagungen diskutiert worden. Es handelt sich weitgehend, zumindest bezüglich der Unterscheidung von Brutto- und Nettowirkungen, um Probleme allgemeiner Natur, mit denen sich das Untersuchungsdesign einer jeden Evaluation auseinandersetzen muss.

Die Leitfäden zur Evaluation der neuen Programmplanungsperiode stellen allerdings – ohne Beachtung der Diskussion um die methodischen Grenzen der Evaluation - z. T. recht unrealistische Erwartungen an die Ergebnisse der Evaluation. Während bis 1999 die Strukturfondsverordnungen und Leitlinien zur Evaluation flexibel angelegt waren und hinsichtlich der konkreten methodischen Regelungen einen weiten Spielraum ließen (TOEPEL, 2000, S. 401), hat die EU-Kommission seit 2000 ein enges Korsett vorgegeben. Dieses ist mit seinem stark auf Quantifizierung ausgerichteten und auf die Daten des Begleitsystems abstellenden Ansatz der Komplexität und Vielfalt der Strukturfondsprogramme nicht immer angemessen (siehe Kapitel 4.1.4).

¹² Die EU-Kommission unterschied bis 1999 im Rahmen ihrer Strukturpolitik regional ausgerichtete Programme, mit denen bestimmte Regionen verstärkt gefördert wurden, und regionsunabhängige Programme, mit denen Maßnahmen horizontal, d. h. unabhängig von einer regionalen Schwerpunktsetzung, angeboten wurden. Zu den regional ausgerichteten Programmen gehörte die sog. Ziel-1- und Ziel-5b-Förderung, die in wirtschaftlich rückständigen Gebieten eine besondere Wirtschaftsförderung ermöglichte. Zu den horizontalen Programmen gehört die sog. Ziel-5a-Förderung, in deren Rahmen beispielsweise einzelbetriebliche Fördermaßnahmen im Agrarbereich, wie z. B. die Ausgleichszulage und die einzelbetriebliche Investitionsförderung, durchgeführt werden konnten.

4.1.4 Überbetonung der quantitativen Analyse

Die von der EU-Kommission im Zusammenhang mit den Bewertungsfragen gewählten Kriterien und Indikatoren besitzen eine Tendenz, die Evaluation schwerpunktmäßig quantitativ auszurichten. Durch die Konzentration auf quantifizierbare Indikatoren besteht jedoch die Gefahr, dass zwar formal eine Vergleichbarkeit zwischen verschiedenen Evaluationsberichten hergestellt wird. Bei genauerer Analyse und Einbindung in die komplexen Wirkungszusammenhänge kann allerdings die Aussagefähigkeit und auch Vergleichbarkeit nicht zwingend hergestellt werden. Es besteht also die Gefahr, dass Scheinwahrheiten produziert werden (TOEPEL, 2000, S. 403). Gerade regionalpolitische Fördermaßnahmen oder auch Agrarumweltmaßnahmen erzeugen multiple Effekte, die häufig nicht quantifiziert werden können oder wegen der Zeitnähe der Evaluation (zunächst) nur qualitativ zu erfassen sind. Eine rein quantitative Wirkungsmessung würde die Breite und die Aussagekraft der Evaluation somit erheblich einschränken.

Die Tendenz der Bewertungen zu quantitativen Analysen ist eine Folge der von der EU-Kommission zur Begleitung und Bewertung verfassten Leitfäden, also des formellen Rahmens, den die EU-Kommission als Mindeststandard für jede von ihr kofinanzierte Bewertung vorgegeben hat. Dies hängt unmittelbar mit dem Anspruch der EU-Kommission als einem der wichtigsten Stakeholder zusammen, auf der Grundlage der Bewertungsberichte relativ einfach Synopsen erstellen zu können, um ihrer Berichtspflicht gegenüber dem EU-Rechnungshof und dem Europäischen Parlament zu genügen. Dem Prinzip „value for money“ folgend liefern quantitative Ergebnisse vielen interessierten Akteuren eine vordergründig klare Auskunft. Die meist sehr viel schwieriger aggregierbaren und zu interpretierenden Ergebnisse qualitativer Analysen finden dagegen weniger Interesse. Die Herstellung von Vergleichbarkeit kann damit auf Kosten des Informationsgehalts gehen.

4.1.5 Evaluation als externe Verpflichtung

Die für die Durchführung der Entwicklungsprogramme verantwortlichen Länderverwaltungen (Auftraggeber) zeigen für die Bewertungen ihrer Programme bislang ein eher geringes Eigeninteresse. „Evaluation ist eine externe Verpflichtung (Teil der Spielregel, um Förderquellen zu erschließen) mit geringer innerer Emotionalität. [...] Es besteht bei allen Beteiligten ein relativ geringes Interesse, sowohl am Prozess der Evaluation als auch am Ergebnis. Wichtig allein ist, sie regelkonform durchzuführen“ (HEINTEL, 2002). Gleiches gilt auch für den Bund, der im Zusammenhang mit der Bewertung der Fördermaßnahmen teilweise eine Koordinierungsfunktion einnimmt. Die Vorschläge der EU-Kommission zur Bewertung der ländlichen Entwicklungsmaßnahmen wurden mit den Mitgliedstaaten abgestimmt und im Ausschuss für Agrarstrukturen und ländliche Entwicklung (STAR-Ausschuss) verabschiedet. Es gab verschiedene Positionspapiere von Seiten des BMVEL und der relevanten Bund-Länder-Arbeitsgemeinschaften zu den Bewertungsansätzen der EU-Kommission, die v. a. eine Flexibilisierung und eine Reduzierung der Bewertungserfordernisse beinhalteten. Damit konnten sich Bund und Länder jedoch nur in wenigen Punkten durchsetzen. Der eigentliche Fragenkanon und die Struktur der Bewertung (KOM, 1999) blieben als Mindestkatalog bestehen. Die Möglichkeit, diesen Mindestkatalog durch eigene Fragestellungen an die spezifischen Programmzielsetzungen der Län-

der anzupassen, hat kaum ein Bundesland genutzt. Der Leistungskatalog, der der Mehrzahl der Evaluationen zugrundeliegt, deckt sich somit mit den Vorgaben der EU-Kommission. Die Abarbeitung dieses Pflichtprogramms sichert die Kofinanzierung der Bewertung und schafft die erforderliche Basis für mögliche Programmänderungen zur Halbzeit. Die Chancen, auch eigene Zielsetzungen überprüfen zu lassen, bleiben so ungenutzt.

Das geringe Eigeninteresse der Auftraggeber ist v. a. auf die Eigenheiten des Politikbereichs zurückzuführen. Der Bereich der Agrarpolitik war bislang deutlich von der Markt- und Preispolitik dominiert. Eine Programmierung und Bewertung von markt- und preispolitischen Maßnahmen wurde - verglichen mit der langen Tradition, die diese Interventionen haben - erst sehr spät vorgenommen und auch immer durch die EU-Ebene selbst veranlasst, die diesen Politikbereich zu 100 % finanziert. Aus diesem Bereich sind demnach keine Impulse zu erwarten, die Bewertungsaktivitäten, auch auf Ebene der Mitgliedstaaten, unterstützen.

Der Europäische Ausrichtungs- und Garantiefonds für Landwirtschaft (EAGFL) - Abteilung Ausrichtung hingegen ist in den allgemeinen Bewertungszyklus der EU-Strukturpolitik einbezogen, allein durch seine Verknüpfung mit den regional ausgerichteten Ziel-Politiken. Weniger verankert waren die Evaluationsaktivitäten im Bereich der Ziel-5a-Politik, obwohl gerade diese Instrumente z. T. schon lange bestehen. Die Agrarumweltmaßnahmen, seit 1994 als flankierende Maßnahmen in Deutschland angeboten, waren nur einer Ex-post-Bewertung zu unterziehen. Das heißt, dass wesentliche Teile der Agrarpolitik der Gemeinschaft nicht systematisch bewertet wurden. Über eine längere Evaluationserfahrung verfügen nur die in die allgemeine Strukturpolitik der EU eingebundenen Bereiche des EAGFL. Der geringe Stellenwert von Evaluation im Bereich des EAGFL auf EU-Ebene setzt sich auf der Ebene der Mitgliedstaaten fort.

Die verantwortlichen Verwaltungen in den Mitgliedstaaten, die nun obligatorisch Bewertungen durchführen, sehen diese eher als ein Kontrollinstrument seitens der EU-Kommission (siehe Kapitel 3 - ebenenbezogene Interessenunterschiede). Beide Ebenen (die Generaldirektion Landwirtschaft der EU-Kommission und die Landwirtschaftsressorts der Länder) verfolgen aber auch ein gemeinsames sektorbezogenes Interesse. Angesichts von knappen Haushaltsmitteln wollen beide Ebenen als Auftraggeber der Evaluation den Fortbestand der von ihnen entwickelten Maßnahmen und den hierfür bereitgestellten Mitteln sichern. Beide Interessenkonstellationen (Kontrollfunktion und Sicherung der Haushaltsmittel) schaffen kein günstiges Klima für Evaluationen, die sich als Lerninstrument verstehen (vgl. HOLZINGER, 2001, S. 36).

Evaluation eher als externe Verpflichtung zu sehen, hat auch Auswirkungen auf das Verhältnis zwischen Evaluatoren und Auftraggebern. Eigentlich soll die Halbzeitbewertung nicht nur ein Rechenschaftsbericht gegenüber Dritten (EU-Kommission, EU-Rechnungshof, EU-Parlament) sein, sondern v. a. formativen Charakter haben. "Die *Halbzeitbewertung* bietet die Möglichkeit, gegebenenfalls Korrekturen am Programm vorzunehmen und die Durchführung zu verbessern. Außerdem kann sie den Behörden, die für das Programm zuständig sind, wertvolle Informationen liefern" (KOM, 1999, S. 6). Gerade in der geforderten Umsetzungsorientierung liegt ein Defizit bislang durchgeführter Evaluationen. Umsetzbare Empfehlungen können aber nur in einem Dialog mit den

verantwortlichen Länderverwaltungen entwickelt werden, da externe Evaluatoren für wesentliche Teile ihres Wissens grundsätzlich auf Insider und deren Kooperationsbereitschaft angewiesen sind (vgl. SCHUBERT, 2001, S. 317). Kooperationsbereitschaft und die Offenheit für Lernprozesse sind jedoch aufgrund der Interessenlage und der vielfach bestehenden Personalknappheit der Länderverwaltungen verhalten. Länderverwaltungen und der Bund möchten die im Rahmen dieser Bewertungen gewonnenen Ergebnisse vorwiegend als Legitimation ihrer eigenen Politik nutzen und "Misserfolge" nur dann nach Brüssel melden, wenn sich diese mit den Einschätzungen der jeweiligen Länderministerien decken. „Werden Lerneffekte dennoch angestrebt, so tritt hier häufig eine Vermischung zwischen Lernen und Kontrolle auf, was ‚sanktionsfreies Lernen‘ von vorneherein unmöglich macht" (Heintel, 2002).

4.1.6 Knappe Mittelausstattung und kurze Bearbeitungszeit

Angesichts der zu erledigenden Aufgaben sind finanzielle Mittel und der zur Verfügung stehende Bearbeitungszeitraum für die Evaluation, betrachtet man die derzeit laufenden Halbzeitbewertungen der ländlichen Entwicklungsprogramme, teilweise als gering zu bezeichnen. Verzögerungen bei der Auftragserteilung¹³, auch aufgrund der z. T. sehr spät erfolgten Programmgenehmigungen, führen zu einer recht knappen Bearbeitungszeit. Beides schlägt sich dann i. d. R. in der gewählten Untersuchungsmethodik und in der Qualität der Bewertungsberichte nieder. Für einen - allgemein empfohlenen - Methodenmix aus eingehender quantitativer und qualitativer Analyse auf breiter und einzelfallbezogener Ebene reichen die vorhandenen Ressourcen häufig nicht aus. Folglich entscheidet man sich für die Methode, die „machbar“ ist. Um zu allen Bewertungsfragen Stellung beziehen zu können, werden Expertenbefragungen und -interviews durchgeführt. Bisweilen werden anstelle der aufwändigen Analyse der Einzelfalldaten (s. Punkt 4.2.1) einige wenige Fallbeispiele dargestellt und darauf aufbauend generelle Schlussfolgerungen gezogen, die jedoch häufig nur wenig belastbar sind.

Neben den genannten Umsetzungsdefiziten im Zusammenhang mit der Programmgenehmigung und den zeitaufwändigen Vorschriften zur Vergabe von Bewertungen tragen sicherlich auch die Vorstellungen der bewertenden Stellen dazu bei, dass die im Rahmen der Bewertungen angesetzten Mittel und Bearbeitungszeiten als knapp empfunden werden. Da es sich bei den im Zusammenhang mit der Förderung des ländlichen Raums durchzuführenden Bewertungen nicht um Grundlagenforschung oder angewandte Forschung handelt, müssen wissenschaftlich ausgerichtete Institutionen bei der Durchführung von Evaluationen ihren Arbeitsstil an die Rahmenbedingungen anpassen.¹⁴ Insbesondere

¹³ Die Verzögerungen bei der Auftragserteilung resultieren teilweise auch aus den Schwierigkeiten der Verwaltungsbehörden, im Rahmen von Ausschreibungsverfahren Leistungsbeschreibungen für Evaluationen zu formulieren und die eingehenden Angebote formell und fachlich zu prüfen. Bei den zentralen Evaluationen resultiert die Verzögerung bei der Auftragserteilung v. a. auf dem schwierigen Abstimmungsprozess zwischen Bund und Ländern.

¹⁴ Zur Abgrenzung der Evaluationsforschung von der Grundlagen- oder angewandten Forschung siehe BORTZ und DÖRING (1995, S. 96 ff.).

umfangreiche Literaturrecherchen und komplexe Modellkalkulationen müssen aufgrund der gegebenen finanziellen und personellen Restriktionen unterbleiben.

4.1.7 Keine Offenlegung der Ergebnisse und fehlende Klarheit des Berichtes

Ein wesentlicher Standard für Evaluationen ist die Offenlegung der Ergebnisse für die relevanten Stakeholder einschließlich der interessierten Öffentlichkeit (Sanders, 1999, S. 139 ff.). Hier sind erhebliche Defizite zu verzeichnen, da in der Vergangenheit nur wenige Evaluationsstudien veröffentlicht wurden. Die Veröffentlichung von Evaluationsergebnissen war bisher nicht klar geregelt. Die Veröffentlichung von Ergebnissen ist aber wichtig, um den Vorwurf der Parteilichkeit, der Evaluationen oft anhängt, zu entkräften. Nur nach Publikation der Befunde können andere die Resultate und Schlussfolgerungen einer Evaluation auf ihre Stichhaltigkeit und Trägfähigkeit hin kritisch prüfen. „Die ‚Scientific Community‘ kann diese Kontrollfunktion aber nur dann wirkungsvoll übernehmen, wenn Untersuchungsberichte nachvollziehbar und vollständig abgefasst sind und öffentlich zugänglich gemacht werden.“ (BORTZ und DÖRING, 1995, S. 98).

Für die neue Förderperiode 2000 bis 2006 legt die EU-Kommission fest, dass die Bewertungsergebnisse der Öffentlichkeit - entweder in vollständiger Fassung oder in einer Zusammenfassung - zugänglich gemacht werden sollten, sobald der Auftraggeber der Bewertung den entsprechenden Bericht angenommen hat. Die EU-Kommission selbst will zumindest die Zusammenfassungen der Berichte im Internet veröffentlichen (KOM, 1999, S. 27). Gegenüber den Vorgängerperioden stellt das Veröffentlichungsgebot eine Verbesserung dar. Allerdings sollte u. E. noch konkreter festgelegt werden, in welcher Form und zu welchem Zeitpunkt eine Veröffentlichung - auch der Einzelberichte der Länder - zu erfolgen hat.

Die Veröffentlichung der Ergebnisse ist ein wesentlicher Standard, um die Wirkung einer Evaluation zu erhöhen. Die "Klarheit des Berichts" (SANDERS, 1999, S. 73 ff.) stellt ebenfalls einen wichtigen Evaluationsstandard dar. "Ein guter Evaluator sollte - auch wenn dies zugebenermaßen nicht ganz einfach ist - in der Lage sein, seine Ergebnisse so aufzubereiten, dass sie auch für ein weniger fachkundiges Publikum nachvollziehbar sind" (BORTZ und DÖRING, 1995, S. 97). Die bisherigen Evaluationsberichte entsprachen nur begrenzt dem Standard "Klarheit des Berichts" und waren wenig auf die Bedürfnisse der unterschiedlichen Adressaten (Auftraggeber, Beteiligte und Betroffene) zugeschnitten. Evaluationen haben zwar wissenschaftlichen Standards zu folgen, in der Ergebnisdarstellung ist die Wissenschaft aber nur ein Stakeholder unter vielen, so dass die Berichtsformen sich entsprechend anpassen müssen. Dass Evaluationsergebnisse bisher nur ein geringes Umsetzungspotenzial beinhalteten, ist allerdings stärker dem Charakter der Evaluation als externer Verpflichtung geschuldet (vgl. HEINTEL, 2002) als der Form, in der Ergebnisse bislang präsentiert wurden.

4.2 Zusammenfassung

Dargestellt wurden wesentliche Aspekte, bei denen in bisher durchgeführten und derzeit laufenden Evaluationen Probleme auftreten. Diese Probleme wurden vor dem Hintergrund diskutiert, ob und inwieweit

- es sich um generelle Probleme im Rahmen von Evaluationen handelt;
- sie mit den Eigenarten des zu evaluierenden Politikfeldes zusammenhängen
- oder den spezifischen Bedingungen für Entscheidungsfindung, Finanzierung und Implementation im Rahmen der EU-Strukturpolitik sowie den unterschiedlichen Interessen der Stakeholder geschuldet sind (siehe Übersicht 1).

Generelle Probleme im methodischen Design von Evaluationen wirken sich auch auf die Durchführung von Evaluationen von EU-kofinanzierten Programmen aus. Diese beziehen sich v. a. auf die Differenzierung von Brutto- und Nettowirkungen, während die anderen Problembereiche der Wirkungsmessung z. T. "hausgemacht" sind, d. h., sich auch auf die unrealistischen Vorgaben der EU-Kommission zurückführen lassen.

Die Eigenheiten des Politikbereichs haben v. a. Auswirkungen auf die methodischen Probleme der Wirkungsmessung. Durch den Einsatz von Maßnahmebündeln wird beispielsweise eine effiziente, für die Evaluation nutzbare Datenerfassung erschwert.

Konfligierende Interessen der Stakeholder und die den jeweiligen Akteuren zugewiesenen Aufgaben und Handlungsrechte wirken sich weniger auf die methodischen Probleme der Evaluation aus. Vielmehr beeinflussen sie die Rahmenbedingungen der Evaluation, das Zusammenwirken zwischen Auftraggeber und Evaluator sowie die Frage der Wirkung der Evaluationsergebnisse. Gerade aufgrund dieser spezifischen politics- und polity-bezogenen Rahmenbedingungen können formative Evaluationsansätze nur schwer realisiert werden.

Übersicht 1: Problembereiche der Durchführung von Evaluationen von Maßnahmen und Programmen zur Entwicklung ländlicher Räume

| Aspekte der Durchführung von Evaluationen von EU-kofinanzierten Programmen | Probleme lassen sich zurückführen auf | | |
|--|---------------------------------------|---|---|
| | Generelle Probleme von Evaluationen | Eigenarten des zu evaluierenden Politikfeldes | spezifische politics- und polity-bezogene Rahmenbedingungen |
| Defizite bei Datenerfassung und Datenbereitstellung | nein | teilweise | teilweise |
| Unzureichende Ex-ante-Bewertung | nein | nein | teilweise |
| Methodische Probleme bei der Wirkungsmessung | | | |
| - Isolierung der Programm-/Maßnahmenwirkungen | teilweise | ja | nein |
| - Netto- versus Bruttowirkungen | ja | teilweise | nein |
| - Referenzgruppen | teilweise | ja | teilweise |
| - Langfristigkeit der Strukturförderung | teilweise | teilweise | teilweise |
| - Programmplanungsansatz | teilweise | nein | ja |
| Überbetonung der quantitativen Analyse | nein | nein | ja |
| Evaluation als externe Verpflichtung | nein | teilweise | ja |
| Knappe Mittelausstattung und kurze Bearbeitungszeit | nein | nein | teilweise |
| Keine Offenlegung der Ergebnisse und fehlende Klarheit des Berichts | Teilweise | nein | teilweise |

Quelle: Eigene Darstellung

5 Bewertung und Empfehlungen

5.1 Bewertung

Im vorangehenden Kapitel wurden zahlreiche Schwierigkeiten der praktischen Evaluationsarbeit dargestellt. Teilweise könnte dabei der Eindruck entstehen, als sei die Evaluation von Fördermaßnahmen in den vergangenen Jahren kaum vorangekommen. Ein Vergleich mit der Erfolgskontrolle im Rahmen der Gemeinschaftsaufgabe "Verbesserung der Agrarstruktur und des Küstenschutzes" (GAK) vor der Agenda 2000 bietet die Gelegenheit, die oben dargestellten Bemühungen innerhalb der EU um Fortschritte in Richtung auf eine systematischere und regelmäßigeren Erfolgskontrolle einzuordnen. Hierfür bietet

sich der Bereich des Agrarinvestitionsförderungsprogramms (AFP) als finanziell bedeutendste Einzelmaßnahme¹⁵ der GAK an.

Seit Ende der siebziger Jahre wurden von jedem Förderfall im Bereich der einzelbetrieblichen Investitionsförderung (EFP) zahlreiche Daten der Betriebsentwicklungsplanung und der Auflagenbuchführung gesammelt, um eine systematische und laufende Erfolgskontrolle durchzuführen und eine fundierte Informationsgrundlage für die künftige Maßnahmenplanung im ohnehin schwierigen Konsensbildungsprozess zwischen Bund und Ländern zu erhalten (WILLER, 1977, S. 199). Das Ergebnis der Anstrengungen ist jedoch rückblickend enttäuschend. Die Auswertungen, veröffentlicht in zahlreichen Agrarstrukturberichten (BMELF, versch. Jgg.), gleichen mehr einem oberflächlichem Verwendungsnachweis der eingesetzten Mittel als der ursprünglich angestrebten sozio-ökonomischen Erfolgskontrolle. Erst im Jahr 1997 wurde, nachdem eine Studie von STRIEWE et al. (1996) zu einem sehr negativen Urteil hinsichtlich der Effizienz der einzelbetrieblichen Investitionsförderung in der Landwirtschaft kam, vom Bund eine umfassendere Studie zur Untersuchung der Wirtschaftlichkeit dieser von Bund und Länder prioritär behandelten Fördermaßnahme an die FAL vergeben.

Trotz des zuletzt genannten Bewertungsauftrags kann festgestellt werden, dass die Erfolgskontrollen im Bereich der GAK weit hinter den durch die EU-Kommission ausgelösten Bemühungen um die Evaluation der gemeinschaftlich mitfinanzierten Förderprogramme und -maßnahmen zurückblieben. Dies kann darauf zurückgeführt werden, dass die Verfahrensvorgaben der EU sowohl bei der Programmplanung und -genehmigung wie auch in Bezug auf die Berichtspflicht der Länder über die Durchführung der Programme sehr viel rigider gehalten sind, als dies bei einer reinen Bund-Länder-Kooperation im Rahmen der GAK der Fall war und ist. Während die Länder im Bereich der EU-Politik zur Förderung der Entwicklung ländlicher Räume detaillierte Vorgaben zur Finanzkontrolle, zu den Berichtspflichten und zur Erstellung von Bewertungen beachten müssen, heißt es im Gemeinschaftsaufgabengesetz lediglich: „Die Durchführung des Rahmenplans ist Sache der Länder. Die Landesregierungen unterrichten die Bundesregierung und den Bundesrat auf Verlangen über die Durchführung des Rahmenplanes und der Gemeinschaftsaufgabe“.¹⁶

5.2 Empfehlungen

Obwohl aus dem Mehrebenensystem der EU sicherlich Restriktionen für Evaluationen erwachsen, stellt der Vergleich mit Erfolgskontrollen im Rahmen der GAK diese in ein günstigeres Bild. Die Mehrzahl der in Kapitel 4 dargestellten Probleme sind nicht der

¹⁵ Die Ausgaben des Bundes und der Länder für die einzelbetriebliche Investitionsförderung im Rahmen der GAK - dem zentralen gemeinsamen Planungs- und Finanzierungsinstitut von Bund und Ländern für die Agrarstrukturpolitik - beliefen sich allein im früheren Bundesgebiet in den vergangenen 10 Jahren auf durchschnittlich rund 215 Mio. Euro pro Jahr (BMELF, versch. Jgg.).

¹⁶ § 9 des Gesetzes über die Gemeinschaftsaufgabe „Verbesserung der Agrarstruktur und des Küstenschutzes“. WOLF (1999) bestätigt diese Aussagen für den Bereich der regionalen Wirtschaftsförderung.

spezifischen Akteurskonfiguration geschuldet; häufig kann ihnen durch systemimmanente Innovationen begegnet werden. Einige Vorschläge hierzu werden im Folgenden vorgestellt:

- **Datenerfassung und -bereitstellung:** Die Datenerfassung sollte künftig nach einem abgestimmten einheitlichen System erfolgen. Dies beinhaltet sowohl einen inhaltlichen als auch einen formalen Standard. Darüber hinaus sollten die Länder im Einklang mit ihren individuellen Verhältnissen zusätzliche Daten und Informationen erfassen, die für die Evaluation wichtig sind, aber im Nachhinein nur durch hohen Aufwand zu erheben sind. Um die Daten effizient nutzen zu können, ist es notwendig, die Datensammlung auf die Erfordernisse der unterschiedlichen Akteure (z. B. seitens der Bewilligungsstelle, der Evaluation, der Zahlstelle) auszurichten. Aus Akzeptanzgründen ist sicherzustellen, dass keine unnötigen Daten gesammelt werden.
- **Ex-ante-Bewertung:** Es sollte künftig sichergestellt werden, dass die Entwicklungsprogramme auf einer Ex-ante-Bewertung beruhen, die allgemeinen Bewertungsstandards entspricht. Dies beinhaltet u. a. ein Indikatorensystem, das mit den Zielen des Programms und mit den Bewertungsfragen der EU-Kommission für die Halbzeit- und Ex-post-Bewertung abgestimmt ist. Dies ist bereits vor der Genehmigung der Programme einzufordern und später - zum Zweck der Qualitätssicherung - z. B. durch eine Evaluationssynthese zu dokumentieren.
- **Inhaltliche und methodische Ausrichtung der Bewertung:** Künftig sollten neben den, vor allem quantitativ fassbaren Indikatoren verstärkt qualitative Indikatoren in die Bewertung einfließen, um die inhaltliche Basis und die Aussagekraft der Bewertung zu verbessern. Die gegenwärtig vor allem kurz- bis mittelfristig ausgerichteten Wirkungsanalysen sollten verbindlich um langfristige Analysen ergänzt werden, um die in ihrer Natur eher langfristigen strukturellen Wirkungen besser ermitteln zu können.
- **Abstimmungsprozess zwischen Evaluatoren:** Bislang findet auf Ebene der Evaluatoren kaum ein Diskussionsprozess sowohl bezüglich des methodischen Vorgehens als auch bezüglich der Evaluationsstandards statt. Da alle Evaluatoren sich mit vergleichbaren methodischen Problemen auseinandersetzen müssen, sollte auch der EU-Kommission an der Organisation eines verstärkten Diskussionsprozesses gelegen sein.¹⁷
- **Wettbewerb der Evaluatoren:** Evaluationen sollten künftig grundsätzlich in einer freien Ausschreibung vergeben werden. Dabei müssen die Anforderungen („terms of reference“) klar und ausführlich formuliert sein, damit ein fairer Wettbewerb zwischen den Anbietern entstehen kann.
- **Mittelausstattung und Bearbeitungszeit:** Künftig sollte durch eine frühzeitige Ausschreibung der Bewertungsaufträge dafür gesorgt werden, dass ausreichend Bearbeitungszeit für die Evaluatoren zur Verfügung steht. Der Termin der Auftragsvergabe sollte bereits im Entwicklungsprogramm festgehalten werden.

¹⁷ Auf nationaler Ebene wurde im Rahmen der Deutschen Gesellschaft für Evaluation (DeGEval) im März 2002 mit dem Arbeitskreis „Strukturpolitik“ ein Forum geschaffen, das allen mit der Evaluation im Bereich der Strukturpolitik befassten Personen und Institutionen die Möglichkeit gibt, sich über methodische und inhaltliche Fragen im Zusammenhang mit Evaluationen auseinanderzusetzen.

- **Nützlichkeit der Evaluation:** Durch die stärkere Einbindung der sog. Stakeholder bei der Konzeption und Durchführung der Evaluation (Stärkung der formativen Evaluation) kann die Akzeptanz und Kooperation der Informationsträger deutlich gesteigert werden. Voraussetzung hierfür ist eine umfassende Information der Beteiligten über das Ziel und die Konsequenzen (z. B. Zeitaufwand) der laufenden Einbindung der wesentlichen Akteure in den Evaluationsprozess.
- **Transparenz und Qualitätssicherung:** Um die Qualität der (Ex-ante-, Halbzeit- und Ex-post-) Bewertungen zu erhöhen, sollten alle Bewertungen zeitnah zur Abgabe der Berichte veröffentlicht werden. Eine Meta-Evaluation mit schwerpunktmäßiger Ausrichtung auf die Einhaltung der formalen Anforderungen und deren baldige Veröffentlichung könnte zusätzlich zu mehr Transparenz und Qualität der Bewertungen beitragen. Die Veröffentlichungen sollten, um eine möglichst große Resonanz zu erreichen, inhaltlich und formal stärker auf die jeweiligen Adressaten zugeschnitten werden. Umfangreiche wissenschaftliche Abhandlungen sind sicherlich hierfür nicht geeignet. Für die programmverantwortlichen Akteure müssen praxistaugliche Handlungsempfehlungen formuliert werden.

Um die Nützlichkeit der Evaluation zu verbessern, ist der formative Charakter der Evaluation zu verstärken. Gegenwärtig ist jedoch teilweise eine gegenläufige Entwicklung zu beobachten, indem beispielsweise die Zuweisung von zusätzlichen Mitteln im Rahmen der operationellen Programme der Ziel-1-Regionen an die Ergebnisse der Halbzeitbewertung geknüpft werden. Derlei Anreiz- oder Sanktionsmechanismen könnten im Hinblick auf die Aussagekraft der Bewertungsergebnisse auch kontraproduktive Wirkungen entfalten, wenn die Folgen einer negativen Bewertung gravierend sind. In diesem Fall besteht ein Anreiz für die Programmverantwortlichen, Gefälligkeitsgutachten zu initiieren. Auf diese Weise würde der Bewertungszyklus als wertvoller Lernprozess für Politik und Verwaltung geradezu pervertiert. Es bleibt abzuwarten, ob und inwieweit eine direktere Verbindung zwischen Evaluationsergebnissen und daraus resultierenden Politikentscheidungen auf die Evaluationsbedingungen negativ zurückwirkt (vgl. SUMMA, 1998, S. 3).

6 Zusammenfassung

Die zunehmende Bedeutung von Evaluationen im Bereich der Politik zur Entwicklung der ländlichen Räume trägt dem wachsenden politischen und fiskalischen Stellenwert dieses Politikbereichs als der zweiten Säule der gemeinsamen Agrarpolitik Rechnung. Seit der Reform der Strukturfonds der Europäischen Union (EU) im Jahr 1988 sind Bewertungen zunehmend zu einem festen Bestandteil des Politikzyklus in diesem Politikbereich geworden. Mit Beginn der Förderperiode 2000 bis 2006 hat die EU die diesbezüglichen Anforderungen verschärft und systematisiert. Alle Mitgliedstaaten in der Europäischen Union sind verpflichtet, die im Rahmen der EU-Verordnung über die Förderung der Entwicklung des ländlichen Raums (VO (EG) Nr. 1257/1999) durchgeführten Fördermaßnahmen zu begleiten (Monitoring) und zu bewerten (Evaluation). Insofern gewinnt auch die Frage nach der praktischen Durchführung der Evaluationen in diesem Politikbereich an Relevanz. Die EU-Kommission will mit der Bewertung von Förderprogrammen das Management der Programme verbessern, ihrer Rechenschaftspflicht gegenüber politischen Entscheidungsträgern und der Öffentlichkeit nachkommen sowie Unterstützung bei

der Budgetierung von Haushaltsmitteln erhalten. Allerdings bewegen sich diese Evaluationen in einem komplexen Umfeld, in dem Stakeholder mit heterogenen Interessen in einem Mehrebenensystem zusammenwirken. Der Beitrag stellt auf der Grundlage eigener Evaluationen und Berichte anderer Evaluatoren die bisherigen Erfahrungen im Hinblick auf dabei aufgetretene Probleme dar, analysiert diese auf ihre Ursachen hin und legt Vorschläge für eine verbesserte Bewertungspraxis vor.

Die Erfahrungen mit den bisher durchgeführten Bewertungen werden vor dem Hintergrund der Ansprüche an Evaluationen und den besonderen Rahmenbedingungen eines Mehrebenensystems analysiert. Dabei werden ausgewählte technische und methodische Aspekte, das Verhältnis zwischen Auftraggeber und Evaluator, die finanziellen und personellen Rahmenbedingungen einer Evaluation sowie Fragen der Offenlegung und Klarheit von Ergebnissen beschrieben. Diskutiert werden diese Aspekte einer Evaluation vor dem Hintergrund von drei Problemkategorien: (a) Probleme, die generell bei Evaluationen auftauchen, (b) Probleme, die den Eigenarten des zu evaluierenden Politikfeldes zuzuschreiben sind, sowie (c) Probleme, die auf die spezifischen politics- und polity-bezogenen Rahmenbedingungen der EU-Politik zur Förderung der ländlichen Entwicklung zurückzuführen sind.

Generelle Probleme im methodischen Design von Evaluationen zeigen sich v. a. im Hinblick auf die Differenzierung von Brutto- und Nettowirkungen, während die anderen Problembereiche der Wirkungsmessung z. T. "hausgemacht" sind. Die Eigenheiten des Politikbereichs haben v. a. Auswirkungen auf die methodischen Probleme der Wirkungsmessung. Durch den Einsatz von Maßnahmebündeln wird etwa eine effiziente und für die Evaluation nutzbare Datenerfassung erschwert. Konfligierende Interessen der Stakeholder und die den jeweiligen Akteuren zugewiesenen Aufgaben und Handlungsrechte beeinflussen das Zusammenwirken zwischen Auftraggeber und Evaluator sowie die Frage der Wirkung und Verwendung der Evaluationsergebnisse. Gerade aufgrund dieser spezifischen politics- und polity-bezogenen Rahmenbedingungen können formative, d. h., vor allem auf die Verbesserung des Programmmanagements ausgerichtete Evaluationsansätze nur schwer realisiert werden. Obwohl aus dem Mehrebenensystem der EU sicherlich Restriktionen für Evaluationen erwachsen, stellt der Vergleich mit Erfolgskontrollen im Rahmen der Gemeinschaftsaufgabe "Verbesserung der Agrarstruktur und des Küstenschutzes" (GAK) vor der Agenda 2000 diese in ein günstigeres Bild. Die Mehrzahl der dargestellten Probleme sind nicht der spezifischen Akteurskonfiguration geschuldet; häufig kann ihnen durch systemimmanente Innovationen begegnet werden. Im Beitrag werden Vorschläge hierzu gemacht.

Literatur

- AHRENS, H.; LIPPERT, Ch.; RITTERSHOFER, M. (2000): Überlegungen zu Umwelt- und Einkommenswirkungen von Agrarumweltprogrammen nach VO (EWG) Nr. 2078/92 in der Landwirtschaft. *Agrarwirtschaft* 49, H. 2, S. 99-115
- BERGSCHMIDT, A.; PLANKL, R. (1999): Evaluation der Agrarumweltprogramme gemäß der Beschlüsse der Agenda 2000. *Berichte über Landwirtschaft* 77, H. 4, S. 570-590

- BMELF (versch. Jgg.) (Bundesministerium für Ernährung, Landwirtschaft und Forsten): Die Verbesserung der Agrarstruktur in der Bundesrepublik Deutschland.
- BORTZ, J.; DÖRING, N. (1995): Forschungsmethoden und Evaluation für Sozialwissenschaftler. Berlin.
- ECONOMIC AND SOCIAL RESEARCH INSTITUTE (1997): EU Structural Funds in Ireland. A Mid-Term Evaluation of the CSF 1994-99. Dublin.
- ESER, T. W. (2001): Evaluation und Qualitätsmanagement - Anforderungen und Konsequenzen für die EU-Strukturpolitik. Informationen zur Raumentwicklung, Heft 6/7, S. 327-339
- FÄHRMANN, B. et al. (2001): Ex-post-Bewertung des saarländischen Ziel 5b-Programms 1994-1999, Endbericht. Institut für Betriebswirtschaft, Agrarstruktur und ländliche Räume in Zusammenarbeit mit dem Institut für praxisorientierte Forschung und Bildung e. V., Braunschweig.
- FORSCHUNGSGESELLSCHAFT FÜR AGRARPOLITIK UND AGRARSOZIOLOGIE e.V., Bonn (FAA) (2001): Politik zur Entwicklung ländlicher Räume - wie geht es weiter? Verhandlungen zur öffentlichen Arbeitstagung am 9.11.2000 in Bonn-Röttgen. Bonn.
- FORSTNER, B. (2000): Untersuchung zur Wirtschaftlichkeit der einzelbetrieblichen Investitionsförderung in Deutschland. In: Alvensleben, R. von; Koester, U.; Langbehn, C.: Wettbewerbsfähigkeit und Unternehmertum in der Land- und Ernährungswirtschaft. Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e. V., Band 36. Münster-Hiltrup, S. 151-158
- GRAJEWSKI, R.; KOCH, B. (2002): Welche Anforderungen gibt es an das Monitoring im Rahmen der ländlichen Entwicklungsprogramme gemäß VO (EG) Nr. 1257/99 aus Sicht von Evaluatoren? Schriftliche Fassung eines Vortrags auf einem Workshop der Deutschen Gesellschaft für Evaluation bezüglich der Anforderungen an ein Monitoringsystem als Grundlage der Evaluation ländlicher Entwicklungsprogramme (http://www.bal.fal.de/download/graj_Vortrag.pdf).
- HEINTEL, M. (2002) Probleme der Evaluation von Regionalmanagements - oder: Ist Erfolg messbar? In: Sedlacek, P. (Hrsg.): Evaluation in der Stadt- und Regionalentwicklung. Opladen (erscheint im November 2002).
- HOLZINGER, E. (2001): Evaluation ist zum Lernen da. Raum 44, Österreichische Zeitschrift für Raumplanung und Regionalpolitik, S. 36-37 (download als pdf-file unter: http://eur-mas.iff.ac.at/menu_r/download/evaluation_holzinger.pdf)
- KELSEN, H. (1966): Principles of International Law. New York, 2. Auflage.
- EUROPEAN COMMISSION (COM) (1999): Means Collection: „Evaluating socio-economic programmes“ (volume 1-6).
- EUROPÄISCHE KOMMISSION (KOM) (1999): Bewertung von Programmen zur Entwicklung des ländlichen Raums im Zeitraum 2000-2006 mit Unterstützung des Europäi-

schen Ausrichtungs- und Garantiefonds für die Landwirtschaft. Leitfaden (Dokument VI/8865/99).

- EUROPÄISCHE KOMMISSION (KOM) (2000): Gemeinsame Bewertungsfragen mit Kriterien und Indikatoren zur Bewertung von Programmen zur Entwicklung des ländlichen Raums, die von 2000 bis 2006 durchgeführt und durch den EAGFL gefördert werden (Dokument VI/12004/00 Endg.)
- EUROPÄISCHE KOMMISSION (KOM) (2002): Leitfaden für die Halbzeitbewertung der Programme zur Entwicklung des ländlichen Raums 2000-2006 mit Fördermitteln des EAGFL (Dokument VI/43517/02).
- MANN, S. (2000): The Demand for Evaluation from a Public Choice Perspektive. In: Evaluation im Spannungsfeld von Wissenschaft und Politik. Deutsches Institut für Wirtschaftsforschung. Vierteljahreshefte zur Wirtschaftsforschung 69, H. 3, S. 371-378
- MARKS, G. (1993): Structural Policy and Multi Level Governance in the European Community. In: Cafruny, A. W. and Rosenthal, G. W. (Ed.): The State of the European Community. Boulder, S. 388-422
- MAYNTZ, R. ; SCHARPF, F.W. (1973): Planung als politischer Prozess. Aufsätze zur Theorie der planenden Demokratie. Frankfurt/Main.
- MEHL, P.; PLANKL, R. (2002): Regionale Agrarpolitik in Deutschland - Handlungsspielräume in einem verflochtenen Politikfeld. In: Europäisches Zentrum für Föderalismus-Forschung: Jahrbuch des Föderalismus. Föderalismus, Subsidiarität und Regionen in Europa. Baden-Baden: Nomos, S. 199-210
- NAGARAJAN, N; VANHEUKELEN, M. (2002): Evaluation der ausgabenwirksamen Programme der EU. Ein Leitfaden zur Halbzeit- und Ex-post-Evaluation. Generaldirektion XIX - Haushalt. <http://europa.eu.int/comm/budget/evaluation/guide>.
- PLANKL, R.; SCHRADER, H. (1991): Monitoring and evaluation of objective 5b regions. Final report, prepared by Centre for European Agricultural Studies Wye College, Centre d'Estudis de planificacio and Institut für Strukturforchung, Bundesforschungsanstalt für Landwirtschaft (unveröffentlicht).
- RAT (1999) (Europäischer Rat): Verordnung (EG) Nr. 1257/1999 des Rates vom 17.05.1999 über die Förderung der Entwicklung des ländlichen Raums durch den EAGFL und zur Änderung bzw. Aufhebung bestimmter Verordnungen.
- RIEDEL, J. (1997): Die Europäischen Strukturfornds in den neuen Bundesländern. Zwischenbewertung des Mitteleinsatzes von 1994 bis 1996. ifo dresden studien 12. Dresden.
- ROSSI, P. H.; FREEMANN, H. E. (1993): Evaluation. A Systemic Approach. 5. Edition. London.
- SANDERS, J. R. (Hrsg.) (1999): Handbuch der Evaluationsstandards. Opladen.

- SCHARPF, F. W. Die Politikverflechtungsfalle: Europäische Integration und deutscher Föderalismus im Vergleich. Politische Vierteljahresschrift 26 (1985), H. 4, S. 323-356
- SCHLAGHECK, H. (1994): Auswirkungen der Kompetenzverteilung im kooperativen Föderalismus Deutschlands und in der EG auf die Gestaltung der Agrarstrukturpolitik. In: Hagedorn, K.; Isermeyer, F.; Rost, D.; Weber, A. (Hrsg.): Gesellschaftliche Forderungen an die Landwirtschaft. Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V., Bd. 30. Münster-Hiltrup, 1994, S. 89-99
- SCHUBERT, A. (2001): Das Evaluationskonzept der Strukturfonds und die deutsche Verwaltung - ein nicht spannungsfreies Verhältnis. Informationen zur Raumentwicklung, Heft 6/7, S. 317-319
- SCHUBERT, K. (1991): Politikfeldanalyse. Opladen.
- STRIEWE, L.; LOY, J.-P.; KOESTER, U. (1996): Analyse und Beurteilung der einzelbetrieblichen Investitionsförderung in Schleswig-Holstein. Agrarwirtschaft 45, Heft 12, S. 423-434
- SUMMA, H. (1998): The Role of Evaluation in European Policies. 1. Jahrestagung der Deutschen Gesellschaft für Evaluation: "Innovation durch Evaluation" am 24. und 25. September 1998 in Köln. <http://degeval.de/koeln1998/summa.html>.
- TAYLOR, S.; BACHTLER, J.; POLVERARI, L. (2001): Structural Fund evaluation as a programme management tool: comparative assessment and reflections on Germany. Informationen zur Raumentwicklung, Heft 5/6, S. 341-357
- TISSEN, G. (1997): Ein Ansatz für die Zwischenbewertung der Ziel 5b-Politik. Agrarwirtschaft 46, Heft 6, S. 225-232
- TOEPEL, K. (2000): Evaluation in der Regionalpolitik. Vierteljahresschriften zur Wirtschaftsforschung, 69, Heft 3, S. 395-405
- VOELZKOW, H. (1999): Europäische Regionalpolitik zwischen Brüssel, Bonn und den Bundesländern. In: Derlien, H.-U. und Murswieck, A. (Hrsg.): Der Politikzyklus zwischen Bonn und Brüssel. Opladen, S. 105-120
- WOLF, D. (1999): Regelungsbefolgung auf verschiedenen politischen Ebenen. Wird nationales Recht wirklich besser verfolgt als internationales Recht? In: van Deth, J.; König, T. (Hrsg.): Europäische Politikwissenschaft: Ein Blick in die Werkstatt, S. 281-305
- WILLER, H. (1977): Agrarpolitische Planung und politisch-administrative Praxis. Berichte über Landwirtschaft, Band 55, Heft 2, S. 177-213