

### Formen der Evaluierung experimenteller Politikgestaltung

Kaps, Petra

Veröffentlichungsversion / Published Version  
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:  
Verlag Barbara Budrich

#### Empfohlene Zitierung / Suggested Citation:

Kaps, P. (2020). Formen der Evaluierung experimenteller Politikgestaltung. *der moderne staat - dms: Zeitschrift für Public Policy, Recht und Management*, 13(1), 81-100. <https://doi.org/10.3224/dms.v13i1.09>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-SA Lizenz (Namensnennung-Weitergabe unter gleichen Bedingungen) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by-sa/4.0/deed.de>

#### Terms of use:

This document is made available under a CC BY-SA Licence (Attribution-ShareAlike). For more information see: <https://creativecommons.org/licenses/by-sa/4.0>

Petra Kaps

## Formen der Evaluierung experimenteller Politikgestaltung

### Zusammenfassung

Die Reformen der Arbeitsmarktpolitik der letzten 20 Jahre wurden von einer Reihe politischer Experimente begleitet, die jeweils anlassbezogen evaluiert wurden. An drei Beispielen – den Kombilohn-Experimenten, Experimenten zur Pauschalierung einmaliger Leistungen in der Sozialhilfe und am Wettbewerb zwischen Optionskommunen und Arbeitsagenturen in Arbeitsgemeinschaften mit den Kommunen – werden Evaluierungsarrangements als Kombinationen von Evaluierungsgegenständen, Evaluierungsaufträgen und Evaluierungspraktiken untersucht. Dabei wird nach Unterschieden und Gemeinsamkeiten zwischen diesen Arrangements, nach „varieties of evaluation“ (Schwab, 2017, S. 201), gefragt. Und es wird der Zusammenhang zwischen den institutionellen Rahmenbedingungen innerhalb des politisch-administrativen Systems und der Gestaltung dieser Evaluierungsarrangements untersucht. Dabei wird auch skizziert, welcher „politische Gehalt von Evaluierungsprozessen“ (ebd., S. 212) im Vergleich der Beispiele sichtbar wird.

*Schlagerworte:* Evaluationsforschung; politische Experimente; Arbeitsmarktpolitik; Verwaltungsforschung; Advokaten-Koalitionen

### Abstract

*Forms of evaluation in cases of experimental policy making*

The reform of labour market policies over the past 20 years has been accompanied by a series of political experiments. Each of these experiments has been evaluated. For three of these evaluation arrangements are studied as combinations of evaluation objects, assignments and practices. Cases are the evaluation of the wage subsidy experiments, of experiments with flat-rate payments for irregularly required benefits in social assistance and of the competition between two administrative models of jobcenters. Questions are asked about “varieties of evaluation” (Schwab, 2017, S. 201) and the connection between the institutional framework and the design of evaluation arrangements. In addition, “politics of evaluation” (Taylor & Balloch, 2005) visible in the comparison of the three examples are discussed.

*Keywords:* politics of evaluation; policy experiments; labour market policy; public administration; advocacy coalitions

## 1 Einleitung<sup>1</sup>

Die Reformen der Arbeitsmarktpolitik der letzten 20 Jahre wurden in einer Reihe von Themenfeldern von politischen Experimenten eingeleitet oder durch solche Experimente überhaupt erst ermöglicht. Diese Experimente wurden anlassbezogen evaluiert und zumindest formal als Instrument zur Evidenzbasierung von politischen Entscheidungen eingesetzt.

Am Beispiel von drei Evaluationen experimenteller Politikgestaltung unterschiedlicher instrumenteller und institutioneller Tragweite im Feld der Arbeitsmarktpolitik – der Evaluation verschiedener Kombilohn-Experimente (1998-2005), der Experimente zur Pauschalierung einmaliger Leistungen in der Sozialhilfe (1999-2004) und des Experiments zur Erprobung alternativer Modelle der Eingliederung von Arbeitsuchenden im Wettbewerb zwischen Optionskommunen und Arbeitsagenturen in Arbeitsgemeinschaften mit Kommunen (2005-2008) – werden hier Antworten auf die folgenden Fragen skizziert:

- Lassen sich für diese in ihrer institutionellen Reichweite unterscheidbaren Beispiele experimenteller Politik distinkte Evaluierungsarrangements identifizieren?
- Lässt sich dabei ein Zusammenhang zwischen den institutionellen Rahmenbedingungen und der Gestaltung von Evaluierungsarrangements erkennen?
- Welche politischen Deutungskämpfe wurden im Prozess der Evaluierung geführt?

Damit möchte der Artikel im Anschluss an Oliver Schwab (2017) Evaluierungsarrangements und deren Beeinflussung durch ihren politisch-administrativen Kontext systematisch untersuchen und damit einen Beitrag zum Verständnis von „varieties of evaluation“ (Schwab, 2017) leisten.

Evaluierungsarrangements werden hier verstanden als Kombinationen von Evaluierungsgegenständen, Evaluierungsaufträgen und Evaluierungspraktiken, die durch institutionell eingebettete Advokaten-Koalitionen (Sabatier, 1988) gestaltet werden. Es wird davon ausgegangen, dass diese Arrangements unterscheidbar sind und sich die Unterschiede zwischen ihnen, die hier als „varieties of evaluation“ verstanden werden, sowohl aus unterschiedlichen institutionellen Rahmenbedingungen als auch aus unterschiedlichen Interessen der sie prägenden Advokaten-Koalitionen erklären lassen.

## 2 Drei Beispiele experimenteller Arbeitsmarktpolitik und ihrer Evaluierung

### 2.1 Kombilohn-Experimente (1998 bis 2005)

In den 1990er Jahren hatte sich angesichts anhaltend hoher Arbeitslosigkeit eine intensive Debatte darum entwickelt, wie neue Beschäftigungsmöglichkeiten auch unterhalb tariflicher Entlohnung geschaffen werden könnten.

Mikroökonomisch orientierte Marktliberale argumentierten, die Zuverdienst-Regeln in der Sozial- und Arbeitslosenhilfe seien beschäftigungsbehindernd ausgestaltet, weshalb von diesen Leistungen abhängige Arbeitslose in einer Anreizfalle säßen (Boss, 1994; Siebert, 1996). Diese hindere die Menschen daran, existierende Niedriglohnbeschäftigung auszuüben. Deshalb müsse ein sogenannter Lohnabstand gewährleisten, dass auch mit gering entlohnter Tätigkeit mehr Einkommen erzielt werde als ausschließlich mit sozialen Sicherungsleistungen. Dieser Abstand könne entweder durch eine allgemeine Reduktion der Fürsorgeleistungen (Siebert, 1996) oder durch eine gezielte Subventionierung gering entlohnter Tätigkeiten (Spermann, 1996) erzielt werden.

Vertreter einer makroökonomischen Perspektive argumentierten hingegen, dass nicht ausreichend Beschäftigungsmöglichkeiten für gering qualifizierte Personen zur

Verfügung stünden. Deshalb sollten Löhne für Einfach Tätigkeiten flächendeckend subventioniert werden, damit die Lohnkosten für Unternehmen sinken und diese mehr einfache Arbeitsplätze anbieten. Dies sollte entweder über eine niedrigere Abgabenlast für gering entlohnte Beschäftigung (Scharpf, 1997) oder über zielgruppenorientierte Kombilöhne bzw. Lohnkostenzuschüsse erfolgen (Gerster & Deubel, 2001).<sup>2</sup>

Die CDU/CSU wollte im Wahlkampf 1994 „die Sozialhilfe bei denen, für die eine mögliche Erwerbstätigkeit zumutbar ist, ab[zu]senken und ein erzielttes Einkommen dafür nur teilweise auf die Sozialhilfe an[zu]rechnen“ (CDU-Bundesgeschäftsstelle, 1994, S. 17 f.). Die SPD hingegen wollte mit degressiv gestaffelten Lohnkostenzuschüssen an Arbeitgeber neue Beschäftigung für Langzeitarbeitslose am ersten Arbeitsmarkt schaffen, ohne die Tarifstrukturen oder die sozialen Sicherungssysteme zu verändern (SPD-Parteivorstand, 1994, S. 15). Im Jahr 1996 wurden mit dem „Gesetz zur Reform des Sozialhilferechts“ Lohnkostenzuschüsse an Arbeitgeber und Eingliederungszuschüsse an Arbeitnehmende als neue Instrumente in das Bundessozialhilfegesetz (BSHG) eingeführt, die aber zunächst kaum genutzt wurden.

Deshalb beschlossen die konservativ-liberalen Regierungsfractionen und die oppositionelle SPD 1998, eine Experimentierklausel in §18 Abs. 5 BSHG einzuführen, die den Sozialhilfeträgern erlaubte, „im Einzelfall oder durch eine Modellmaßnahme, (...) von den die Dauer oder die Höhe des [Eingliederungs-]Zuschusses begrenzenden Regelungen (...) abzuweichen“ (Deutscher Bundestag, 1998, S. 10). Die Ergebnisse dieser Experimente sollten durch eine Auswertung unter Beteiligung der Länder gesichert werden (ebd., S. 7). Diese Experimentierklausel hatte instrumentellen Charakter. Sie sollte mit der Öffnung der Regelungen zum Instrument des Eingliederungszuschusses in der Sozialhilfe neue Arbeitsanreize setzen.

Verschiedene Bundesländer entwickelten daraufhin eigene Kombilohn-Programme zur Förderung der Beschäftigung von Geringqualifizierten und Langzeitarbeitslosen, allerdings teilweise nicht oder nicht nur für Sozialhilfeempfangende, sondern auch für Langzeitarbeitslose, die Arbeitslosenhilfe erhielten und über die damaligen Arbeitsämter gefördert wurden. Auch einzelne Kommunen entwickelten eigene Fördermodelle über die Sozialhilfe. Bei aller Unterschiedlichkeit im Detail (Blancke, 2004, S. 112 ff.; Kaltenborn, 2001, S. 30 f.) setzten alle Experimente bei der Subventionierung von Niedriglohntätigkeiten an, einzelne alternativ dazu oder darüber hinaus auch an den Regeln zur Anrechnung von Erwerbseinkommen auf die Grundsicherungsleistungen.

Nach dem Regierungswechsel 1998 ließ das vom Bundeskanzleramt initiierte „Bündnis für Arbeit, Ausbildung und Wettbewerbsfähigkeit“ seine „Arbeitsgruppe Benchmarking“ parallel zu den laufenden Experimenten unter relativ großer Aufmerksamkeit einer breiten Fachöffentlichkeit eine Studie zu den Möglichkeiten der Förderung von Niedriglohnbeschäftigung erstellen, für die mehrere Forschungsinstitute mit Modellrechnungen Effekte einer allgemeinen Senkung der Lohnnebenkosten für gering qualifizierte Tätigkeiten schätzten (Bender & Rudolph, 1999; Schupp, Volz, Wagner & Zwiener, 1999). Die Ergebnisse dieser Simulationsrechnungen waren uneindeutig (Fels, Heinze, Pfarr & Streeck, 1999, S. 57 ff.).

Das Bündnis für Arbeit, Ausbildung und Wettbewerbsfähigkeit beschloss daraufhin, verschiedene Kombilohn-Modelle auf Bundesebene zu testen (Andres, 2000, S. 44 f.). Unter den vorhandenen Konzepten fiel die Wahl auf das Mainzer Modell und das SGI-Modell des Saarlandes – nach Ansicht von Susanne Blancke aufgrund der besonders aktiven und frühzeitigen Lobbyarbeit der jeweiligen Landesregierungen beim

Bundeskanzleramt (Blancke, 2004, S. 136 ff.).<sup>3</sup> Unter dem Namen „Chancen und Anreize zur Aufnahme sozialversicherungspflichtiger Tätigkeiten (CAST)“ wurden diese beiden Modelle ab Mitte 2000 in mehreren Regionen erprobt und parallel die Experimentierbedingungen in den beiden laufenden Landes-Modellprojekten im Sinne des CAST-Programmes verändert. Auch die Evaluationen mussten sich diesen Veränderungen anpassen.

Nahezu alle dezentralen Kombilohn-Experimente wurden zwar dezentral evaluiert, wegen des dezentralen Evaluierungsarrangements entstanden aber trotz vergleichbarer Evaluationsgegenstände weder ein einheitliches methodisches Evaluationsdesign noch einheitliche Evaluationskriterien. Die Mehrzahl der Evaluationen nutzte mikroökonomische Analysen mit einem Kontrollgruppendesign oder Vorher-Nachher-Vergleiche, die teilweise mit einer Nutzerbefragung oder einer Implementationsanalyse ergänzt wurden. Dabei waren die Evaluatoren davon abhängig, welche Daten die Arbeitsämter oder Sozialämter zur Verfügung stellen und welche sie selbst generieren konnten.

Auch wenn am Ende unter dem Dach der zentralen Evaluierung des Bundesmodellprojekts die Evaluationsergebnisse aus den dezentralen Projekten zusammengestellt wurden (Kaltenborn, 2001), kann von einem einheitlichen, an systematisch vergleichbaren Kriterien orientierten Evaluationsdesign für die verschiedenen erprobten Ansätze keine Rede sein. Eher könnte man argumentieren, dass die Belastbarkeit der Evaluationsergebnisse im folgenden politischen Prozess auch aufgrund der eingesetzten Evaluationsmethodik in unterscheidbarem Maß anerkannt wurde. Im Unterschied zu den dezentralen Evaluationen wurde das Bundesprogramm CAST zentral von einem Konsortium aus dem Forschungsinstitut der Bundesagentur für Arbeit (IAB) und anderen Forschungs- und Befragungsinstituten evaluiert. Dieses Konsortium konnte, neben eigens durchgeführten Nutzer- und Betriebsbefragungen auf zentral in der Bundesagentur für Arbeit (BA) gehaltene bzw. speziell für das CAST-Programm generierte Prozessdatensätze zugreifen und deshalb zur Evaluation des zentralen Bundesprogramms einen komplexen Methodenmix einsetzen (Kaltenborn, Krug, Rudolph, Weinkopf & Wiedemann, 2005, S. 13).

Die umfangreichen Evaluationen aller Landesprogramme und des Bundesprogramms ergaben nur selten positive Netto-Effekte der Förderung von Kombilöhnen für die Integration in den Arbeitsmarkt, teilweise wurden auch deutlich negative Effekte ermittelt. In allen Experimenten zeigte sich, dass fehlende Arbeitsanreize bzw. die „Faulheit der Arbeitslosen“ nicht die zentrale Ursache hoher Arbeitslosigkeit darstellten. Stattdessen offenbarten sich andere Kernprobleme: Fehlende Arbeitsplätze im Bereich von Einfach Tätigkeiten, mangelnde Beschäftigungsfähigkeit eines Teils der Arbeitslosen und der hohe vermittelnde Aufwand, um Angebot und Nachfrage im Bereich von Einfach Tätigkeiten zusammenzubringen. Für diese Probleme boten Kombilöhne allein keine Lösung.

Das SGI-Modell wurde wegen geringer Nutzungszahlen vorzeitig abgebrochen, das Mainzer Modell wurde ebenfalls vorzeitig abgebrochen (Kaltenborn, Krug, Rudolph, Weinkopf & Wiedemann, 2005, S. 12), denn die Neuregulierung des Niedriglohnssektors war im Jahr 2002 zur Verhandlungsmasse um das „Zweite Gesetz für moderne Dienstleistungen am Arbeitsmarkt“ geworden. Damit war auch die umfangreiche und methodisch anspruchsvolle Evaluierung des zentralen Bundesprogramms, obwohl umfangreiche Daten zur Verfügung standen, durch Diskontinuität der Experimentier-

bedingungen erschwert. Die Evaluationsergebnisse spielten wegen paralleler politischer Aktivitäten zur Einführung von Mini- und Midi-Jobs keine Rolle für den folgenden politischen Entscheidungsprozess.

Sowohl im politischen Prozess der Entwicklung der verschiedenen Kombilohn-Experimente als auch bei der Auswahl der im Bundesmodellprojekt zu erprobenden Varianten zeigten sich verschiedene Advokaten-Koalitionen (Sabatier, 1988), in denen sich Akteure mit jeweils geteilten *belief systems* unter dem Schirm innovativen Experimentierens zusammenfanden: unter anderem Landespolitiker, die eigene Modelle entwickelten und Wirtschaftswissenschaftler, die der Landespolitik eigene Modelle andienten und diese dann für die Länder evaluierten. Sie lassen sich in eine mikroökonomisch und eine makroökonomisch orientierte Advokaten-Koalition unterscheiden. Während die ersten argumentierten, fehlende Arbeitsanreize seien das Kernproblem für die Reintegration von Langzeitarbeitslosen und entsprechend einen Fokus auf verbesserte Arbeitsanreize legten, argumentierten die Vertreter der zweiten Position eher, es mangle an Beschäftigungsmöglichkeiten im Niedriglohnbereich oder an passenden institutionellen Rahmenbedingungen. Entsprechend entwickelten sie eher Kombilohn-Konzepte, die auf das Angebot von Einfacharbeitsplätzen oder deren Rahmenbedingungen Einfluss nehmen sollten.

In den politischen Deutungskämpfen beurteilte jede Advokaten-Koalition ihre eigenen Projekte als erfolgreich oder begründete geringe Erfolge mit fehlerhafter Umsetzung ihrer Vorschläge durch die Arbeits- oder Sozialverwaltung. Immerhin aber drangen über die Zeit die zentralen Erkenntnisse der Evaluierungen der verschiedenen Kombilohn-Experimente als Wissensbestände in beide Advokaten-Koalitionen und in die breite Fach-Community vor, so dass in den letzten Jahren eine weitere Verstärkung von Arbeitsanreizen nicht mehr als zentraler Lösungsansatz zur Bekämpfung von Langzeitarbeitslosigkeit gilt. Das breite Spektrum der erprobten Kombilohn-Konzepte hat dazu sicher beigetragen. Auffällig ist aber zugleich, dass eine zentrale Erkenntnis der Kombilohn-Experimente über die Zeit verschüttet wurde und erst in der wesentlich späteren Evaluierung des Bundes-ESF-Programms zur Bekämpfung der Langzeitarbeitslosigkeit wiederentdeckt werden musste: Dass die Erschließung und Sicherung von Arbeitsplätzen für Langzeitarbeitslose eine aktive Gestaltung der konkreten individuellen Arbeitsbedingungen im Betrieb durch bewerberorientierte Vermittlung und durch ein begleitendes Coaching benötigt (Boockmann et al., 2019).

## 2.2 Experimente zur Pauschalierung der Sozialhilfe (1998 bis 2004)

In den 1980er und 1990er Jahren wurden diverse Vorschläge zur fundamentalen Änderung des Grundsicherungssystems vorgelegt (Kaltenborn, 2000; Lessenich, 2009). Alle Vorschläge, die auf Einführung eines bedingungslosen Grundeinkommens oder einer negativen Einkommensteuer zielten und mit einer Ausgabensteigerung in den öffentlichen Haushalten verbunden waren, hatten aus haushaltspolitischen Gründen keine Umsetzungschance. Angesichts der Kostenentwicklung und einer ausbleibenden Neugestaltung der Regelsätze der Sozialhilfe entwickelte sich parallel zu diesen Forderungen eine Debatte darum, wie innerhalb der bestehenden Sozialhilfe durch eine Pauschalierung bestimmter einmaliger Leistungen über Verwaltungsvereinfachung Kosten gesenkt werden könnten.

Der Landkreistag Baden-Württemberg forderte, wie mehrere konservativ regierte Bundesländer, schließlich, alle „laufenden Leistungen der Hilfe zum Lebensunterhalt (Regelsatz, Mehrbedarf, Kosten der Unterkunft, Kosten der Heizung, Nebenkosten der Unterkunft und die einmaligen Beihilfen für Bekleidung, Hausrat, persönliche Anlässe usw.) zu einem Gesamtbetrag zusammenzufassen, der in der Regel monatlich ausbezahlt wird. Mit diesen monatlichen Gesamtpauschalen“, so die Argumentation, „sollten die Sozialhilfeempfängerhaushalte auskommen, so wie jeder Arbeiter, Angestellte oder Beamte mit dem ihm monatlich nach Abzug der Steuern und Sozialversicherungsabgaben verbleibenden Betrag den Lebensunterhalt für sich und seine Familie zu bestreiten hat“ (Klinger, 2000, S. 127 f.). Die Vertreter dieser Advokaten-Koalition argumentierten eher paternalistisch, haushaltspolitisch und leistungsorientiert im Sinne konservativer fürsorglicher Ansätze.

Die von SPD und Bündnis 90/Die Grünen geführte Bundesregierung verabredete im Koalitionsvertrag 1998, „für einen Modellversuch zur Pauschalierung von Sozialhilfeleistungen (...) unter Beachtung des Prinzips der Bedarfsdeckung und der Freiwilligkeit die gesetzlichen Grundlagen“ zu schaffen und parallel dazu „ein Konzept für eine bedarfsorientierte soziale Grundsicherung [zu] entwickeln, das schrittweise eingeführt werden“ sollte (SPD & Bündnis 90/Die Grünen, 1998, S. 30). Teile beider Parteien wie auch Wohlfahrtsverbände und Sozialrechtler standen der Pauschalierung aus soziallibertärer Haltung skeptisch gegenüber, weil sie darüber eine schleichende Leistungskürzung befürchteten, auch wenn ein Potential für größere Handlungsautonomie der Leistungsberechtigten gesehen wurde, wenn nicht mehr für jede Kleinigkeit ein zusätzlicher Antrag zu stellen wäre.

Nach einigem Hin und Her über das Wie einigten sich Regierung und Opposition, eine entsprechende Experimentierklausel in § 101a Bundessozialhilfegesetz einzuführen. Beobachter argumentierten, das damalige Scheitern, eine Regelsatzverordnung für das BSHG zu erlassen, sei der Auslöser für diesen Weg gewesen (Armborst, 2003, S. 1152). Die soziallibertäre Advokaten-Koalition wurde mit dem Argument zur Zustimmung bewegt, dass eine Pauschalierung der Sozialhilfe die Handlungsautonomie der Leistungsberechtigten erhöhen würde und die eingesparten Verwaltungskosten in eine bessere Ausstiegsberatung investiert werden könnten.

Mit dieser Experimentierklausel sollte „zur Weiterentwicklung der Sozialhilfe [...] die Pauschalierung weiterer Leistungen [...] erprobt werden“. Die Länder sollten die Kommunen als Träger der Sozialhilfe dazu per „Rechtsverordnung ermächtigen, in Modellvorhaben solche Leistungen der Sozialhilfe pauschaliert zu erbringen, für die Beträge nicht schon durch [das BSHG] festgesetzt oder auf Grund [des BSHG] festzusetzen sind. Die Pauschalbeträge“ waren dabei „für einen bestimmten Bedarf festzusetzen und“ mussten „dem Grundsatz der Bedarfsdeckung gerecht werden. Die Modellvorhaben“ sollten so ausgewertet werden, dass „sie eine bundesweite Bewertung zulassen“. Dazu hatten die Kommunen, die zuständigen Landesministerien und das Bundesministerium für Arbeit und Soziales (BMAS) zusammenzuarbeiten (Deutscher Bundestag, 1999, S. 5).

Die Experimentierklausel hatte materiellen Charakter und war, anders als die Kombilohn-Experimente, ausschließlich auf die Sozialhilfe konzentriert. Sie sollte mit der Pauschalierung von Teilen des materiellen Existenzminimums einerseits den Aufwand einer „Massenverwaltung“ (Armborst, 2003, S. 1151) kleinteiliger einmaliger Leistungen reduzieren und je nach politischer Position andererseits das „Einzelan-

spruchsdenken“ (Klinger, 2000, S. 127) der Leistungsberechtigten zurückdrängen oder zu „mehr Autonomie für die Sozialhilfeempfänger“ (Deutscher Bundestag, 1999, S. 7) beitragen.

Zwölf Bundesländer erließen Verordnungen zur Umsetzung solcher Pauschalierungsexperimente (Deutscher Bundestag, 2001, S. 1). Diese überließen die Ausgestaltung der Pauschalierung und die Festlegung des Teilnehmerkreises weitgehend den Kommunen und regelten selbst nicht Näheres zur konkreten Ausgestaltung der Modellprojekte. In einigen Ländern sollten über die lokalen Voraussetzungen Einvernehmen mit dem zuständigen Landesministerium hergestellt bzw. lokale Beiräte in die Ausgestaltung der Experimente einbezogen werden (GV NRW 2000, S. 250, Gesetzblatt Baden-Württemberg 2000, Nr. 9, S. 433 ff. und Hessisches GVBl. I, S. 528).

Die Landesverordnungen eröffneten einen weiten lokalen Handlungsspielraum, den rund 40 Kommunen bundesweit sehr unterschiedlich nutzten (Felder, 2003). Die kommunalen Experimente wurden von teils heftigen, allerdings auf eine kleine Fach-Community begrenzten, sozialpolitischen und juristischen Debatten um die Rechtmäßigkeit der lokal getesteten Pauschalen, die Pflicht zur Teilnahme für die Leistungsberechtigten, die Informationspflichten der Kommunen und grundsätzlich um den Bedarfsdeckungsgrundsatz begleitet (Rothkegel, 2002; Maaß, 2001, S. 118 ff.). Mehrere lokale Experimente landeten vor Gericht. Das Bundesverwaltungsgericht hob Ende 2004 ein Urteil des Bayrischen Verwaltungsgerichts zur Pauschalierung von Bekleidungsbeihilfen in München auf. Die Stadt hatte „den Betroffenen den Text der Ausführungsbestimmungen (...) nicht bekannt gegeben“. Dem Rechtsstaatsprinzip sei aber „nur dann Genüge getan, wenn der Betroffene unmittelbar Kenntnis von den Bestimmungen selbst nehmen kann. Nur dann kann er diese auf ihre Rechtmäßigkeit und Anwendbarkeit überprüfen und sich des Inhalts der durch sie für ihn begründeten Rechte und Pflichten vergewissern“, urteilte das Bundesverwaltungsgericht (2004, S. 12).

Nur zwei der zwölf Länder mit Landesverordnung ließen die Experimente in ihrem Land evaluieren (Nothbaum, Lübker & Kämper, 2004; Univation e.V. & Dr. Beywl & Associates, 2005). Vereinzelt wurden daneben lokale Evaluationen durchgeführt. Für diese dezentralen Evaluierungen ist nur teilweise rekonstruierbar, mit welchen Methoden und durch wen sie durchgeführt wurden. Die bundesweite Meta-Evaluation, mit der das BMAS Mummert Consulting GmbH beauftragt hatte, wurde nie veröffentlicht.

Die dezentralen Evaluationen hatten alle mit der Schwierigkeit zu kämpfen, dass für die Sozialhilfe keine einheitlichen Forschungsdaten verfügbar waren. Die Kommunen nutzten unterschiedliche Prozesssoftware für die Administration der Sozialhilfe, wenn sie denn Ende der 1990er Jahre überhaupt schon auf digitale Aktenhaltung umgestellt hatten. Je nach Struktur der Software mussten die Evaluatoren deshalb die Daten für ihre Untersuchungen weitgehend selbst generieren oder mit jeder Kommune einzeln vereinbaren, welche Daten für die Evaluation nutzbar gemacht werden konnten. Unter anderem deshalb wurden hier keine mikroökonomischen Kontrollgruppen-Designs eingesetzt. Zudem fehlt es, vor allem wegen der Nichtveröffentlichung der Meta-Evaluation, an einem Überblick darüber, für welche Teilleistungen der Sozialhilfe Pauschalierungen in welcher Höhe erprobt wurden, welche Teilnahmebedingungen an den Pauschalierungsexperimenten herrschten und mit welchen Methoden die einzelnen Evaluationen gearbeitet haben.

Das zersplitterte dezentrale Evaluierungsarrangement führte dazu, dass nicht alle Experimente evaluiert wurden und dass weder ein einheitliches Evaluationsdesign noch



einheitliche Evaluationskriterien entwickelt wurden. Die vom Bund beauftragte Meta-Evaluation setzte, anders als die zentrale Evaluation des CAST-Programms (und soweit das überhaupt rekonstruierbar ist), lediglich auf dezentralen Evaluierungen auf und führte nicht zu einem einheitlichen, an systematisch vergleichbaren Kriterien orientierten Evaluationsdesign. Die Frage, ob eine Pauschalierung bestimmter Sozialhilfeleistungen die Handlungsautonomie der Leistungsberechtigten stärkte, wurde nicht untersucht. Der Aufwand für die Antragsbearbeitung wurde nach Angaben aus einer Befragung in den Sozialämtern zwar reduziert, die Sachbearbeitenden beschrieben aber, dass sie dennoch nicht mehr Zeit hatten, die Leistungsberechtigten zu beraten (Felder, 2003, S. 9). Dies war aber ein zentrales Argument für die Pauschalierungs-Experimente gewesen. Die verfügbaren Informationen aus der Meta-Evaluation (Mummert Consulting, 2003; Felder, 2003) lassen nicht erkennen, ob die Zufriedenheits- und Machbarkeits-einschätzungen der Nutzer und der Sachbearbeitenden in einem Zusammenhang mit der Art und Höhe der lokalen Pauschalen in den 40 Kommunen standen, sie wurden lediglich über alle experimentierenden Kommunen aggregiert. Damit berücksichtigte die Meta-Evaluation die Grundidee der experimentellen Politgestaltung im Evaluationsdesign nicht, so dass ihre Ergebnisse auch wegen methodischer Schwächen wenig Aussagekraft hatten.

Trotz der fehlenden Transparenz über die Evaluationspraxis und die Evaluationsergebnisse wurde in den politischen Deutungskämpfen schon während des Evaluierungszeitraums mit ausgewählten Ergebnissen argumentiert. Soweit es sich aus der Analyse dieser Deutungskämpfe rekonstruieren lässt, hatten die Evaluationen ergeben, dass grundsätzlich die Pauschalierung von Leistungen der Sozialhilfe machbar sei. Allerdings lässt sich nicht erkennen, ob das für alle Arten erprobter Pauschalierungen gleichermaßen galt – was aber angesichts der kritischen gerichtlichen Einschätzungen zu den verfassungsrechtlichen Grenzen derartiger Experimente mit dem Existenzsicherungsniveau notwendig gewesen wäre. So wurde zwar berichtet, dass insgesamt etwa die Hälfte der teilnehmenden Leistungsberechtigten mit der lokalen Pauschalierungsregel zufrieden waren, aber unklar blieb, ob es einen Zusammenhang zwischen Zustimmung und Art oder Ausmaß der Pauschalierung oder veränderten Verwaltungsprozessen gab. Unter den Menschen mit Kindern und denen, die schon länger mit Leistungsbezug lebten, lag die Zufriedenheit zudem deutlich niedriger. Und immerhin 55 Prozent aller Befragten hatten angegeben, dass sie aus den Pauschalen in keinem Monat Rücklagen bilden konnten, um bei Bedarf notwendige Anschaffungen zu tätigen, was aber ein wesentlicher Sinn der Pauschalierung war. Zudem hatten die Kommunen keine Kosten gespart, einige von ihnen mit den Pauschalen sogar Mehrkosten verursacht (Mummert Consulting, 2003; Felder, 2003).

Während die Kombilohn-Experimente zumindest teilweise über die Arbeitsverwaltung administriert wurden, waren die Pauschalierungs-Experimente ausschließlich in Sozialämtern angesiedelt. Deshalb standen den Evaluationen unterschiedlich gute Daten zur Verfügung, weshalb unterschiedliche Evaluationsmethoden eingesetzt werden konnten bzw. mussten. Im Bereich der kommunalen Sozialverwaltung waren Evaluationen bis dahin zudem, mit Ausnahmen im Bereich kommunaler Beschäftigungspolitik (Kaps, 2006), noch weniger etabliert als in der Arbeitsverwaltung.

Die negativen Ergebnisse zur für die Mehrheit der befragten Sozialhilfehaushalte nicht ausreichenden Höhe der Pauschalen wurden aus dem parallel zu den Pauschalierungs-Experimenten begonnenen Prozess zur Neubestimmung der Regelsätze weit-

gehend herausgehalten. Die Vertreter der Bundes-, Landes- und Kommunalverwaltung argumentierten nun vor allem haushaltspolitisch, die Pauschalierung sei machbar und würde Verwaltungskosten sparen, ohne den Leistungsberechtigten materiell zu schaden.<sup>4</sup> Dass die Leistungsberechtigten aus den Pauschalen keine Ansparungen für größere Investitionen tätigen konnten, hätte hingegen zu einer Vergrößerung der Pauschalen und damit zu einem Anstieg der Kosten der Grundsicherung führen müssen, wie die Vertreter der nun deutlich verkleinerten soziallibertären Advokaten-Koalition forderten (Spindler, 2004; Martens, 2006).

Da die Evaluationsergebnisse nicht veröffentlicht wurden, konnte die letztere Advokaten-Koalition nicht öffentlichkeitswirksam darauf zurückgreifen. Zudem fanden diese Ergebnisse ohne Veröffentlichung keinen Eingang in den Wissensbestand der Fach-Community. Für den Auftraggeber der Meta-Evaluierung war es wegen des stark dezentralen Evaluationsarrangements und der gleichgerichteten haushaltspolitischen Interessen der Verwaltung aller Ebenen leicht, die Ergebnisse nur so weit zu veröffentlichen, wie es in die damals aktuelle Strategie zur Abschaffung der Arbeitslosenhilfe passte.

An diesem Beispiel lässt sich zeigen, dass, wie von Schwab (2017, S. 214) vermutet, eine „Evaluierung in Mehrebenenkonstellationen“, die „stark von Steuerungsinteressen [des Auftraggebers] motiviert ist“, „Auswirkungen auf die Ausgestaltung der Evaluierungsarrangements“ hat.

Zugleich wurde im Vergleich zum Kombilohn-Beispiel deutlich, dass eine breite Evaluations-Community und eine Institutionalisierung der Evaluationsfunktion (im ersten Beispiel beim IAB und beim Bündnis für Arbeit) einem politischen Interesse eines Auftraggebers, die Ergebnisse von Evaluationen intransparent zu halten und in politischen Deutungskämpfen selektiv für die eigenen Ziele zu nutzen, entgegenwirken kann. Hingegen gewinnt ein zentraler Auftraggeber in dezentralen Evaluierungsarrangements ohne institutionalisierte Evaluationsfunktion und entsprechende Community einen hohen Handlungsspielraum, die Ergebnisse von Evaluationen selektiv für eigene Steuerungsinteressen zu nutzen.

### 2.3 Experimente zum Wettbewerb zwischen zwei Modellen öffentlicher Aufgabenerbringung (2005 bis 2008)

Auch die Modernisierung der Leistungsverwaltung der Sozialhilfe und der Arbeitslosenhilfe wurde damals breit diskutiert. Dabei wurde als Ursache für den steigenden Anteil von Langzeitarbeitslosen und Sozialhilfeempfängern der so genannte Drehtüreffekt zwischen dem Einsatz der arbeitsmarktpolitischen Förderinstrumente von Sozialhilfe- und Arbeitsverwaltung verantwortlich gemacht, die Ineffizienz der doppelten Verwaltungsstrukturen für bedürftige Arbeitslose kritisiert und die unterschiedlichen Regeln von Arbeitslosen- und Sozialhilfe zum Fordern und Fördern zum zentralen institutionellen Hindernis für wirksame Arbeitsmarktintegrationen erklärt (Sell, 1999; Hoffmann, 2000).

Schon seit den 1980er Jahren ging es darum, die Zusammenarbeit zwischen Sozial- und Arbeitsämtern zu verbessern, um das gegliederte System kompatibler zu gestalten und die Fördermöglichkeiten beider Systeme für alle Leistungsberechtigten zugänglich zu machen. Dazu hatten die Bundesanstalt für Arbeit und die Kommunalen Spitzen-

verbände mehrfach Vereinbarungen zur Verbesserung der Verwaltungskooperation geschlossen. Seit 1993 hatten beide Seiten nach § 12 Arbeitsförderungsgesetz (AFG) zusammenzuwirken, aber die praktische Umsetzung zog sich hin.

In der Debatte um das „Gesetz zur Reform der Sozialhilfe“ 1995 wurden dann drei Advokaten-Koalitionen zu diesem Thema sichtbar: BA, Deutscher Städtetag, SPD, Bündnis 90/Die Grünen sowie die meisten Gewerkschaften wollten die Leistungsverwaltungen durch interne Verwaltungsmodernisierung verbessern und deren Wirksamkeit durch verbesserte Kooperation erhöhen. Der Deutsche Landkreistags, Deutsche Städte- und Gemeindebund sowie die CDU/CSU wollten die Aufgaben der BA für Langzeitarbeitslose kommunalisieren, weil die Kommunen „näher am Menschen“ seien als die lokalen Arbeitsämter. Arbeitgeberverbände, FDP und marktliberale Ökonomen wollten die Arbeitsvermittlung privatisieren und damit langfristig der BA allein die Aufgabe der Zahlung der materiellen Versicherungsleistungen zuschreiben oder sie gleich ganz auflösen und ihre sonstigen Aufgaben kommunalisieren.

Nach dem Regierungswechsel 1998 dynamisierte sich diese Debatte. Treiber waren die nun oppositionelle CDU und die Bertelsmann-Stiftung. Die CDU forderte sowohl eine bessere Zusammenarbeit beider Seiten als auch die Kommunalisierung der Aufgaben (Fink, 1998). Die Bertelsmann-Stiftung initiierte mehrere Projekte zur Verwaltungsmodernisierung in Kommunen nach dem Neuen Steuerungsmodell (Banner, 2008), unter anderem das Projekt „Beschäftigungsförderung in Kommunen (BiK)“, in dem neue Formen der Zusammenarbeit von Sozial- und Arbeitsämtern erprobt wurden (Empter & Frick, 1999), und das „Netzwerk Kommunen der Zukunft“, in dem Kommunen verschiedene Aspekte des Neuen Steuerungsmodells erprobten (Bertelsmann-Stiftung, Hans-Böckler-Stiftung & KGSt, 2002). Daneben tagte unter dem Dach der Bertelsmann-Stiftung ein Arbeitskreis „Reform der Arbeitslosenhilfe und Sozialhilfe“, in dem Vertreter der Exekutiven von Bund und Ländern mit Vertretern der BA, der kommunalen Spitzenverbände, von Gewerkschaften und Arbeitgeberverbänden, ausgewählten Praktikern der kommunalen Beschäftigungsförderung und Experten der Bertelsmann-Stiftung und der *con\_sens* GmbH zusammentrafen, um eine Reform von Arbeitslosen- und Sozialhilfe vorzubereiten (Schmachtenberg, 2003; Bertelsmann-Stiftung, 2002; Pautz, 2008).

Die Bundesregierung förderte dann zunächst unter großer Aufmerksamkeit der Fachöffentlichkeit Modellprojekte zur Verbesserung der Zusammenarbeit von Arbeits- und Sozialämtern (MoZArT-Projekte, Hess, Schröder, Smid & Reis, 2004). Roland Koch (CDU), damals Hessischer Ministerpräsident, dynamisierte die Debatte, indem er – inspiriert von Erfahrungen des Hessischen Partnerlandes Wisconsin in den USA mit dem *workfare*-Programm „Wisconsin works“ – lautstark die „Zusammenlegung“ von Arbeitslosen- und Sozialhilfe und die Kommunalisierung der dann neuen Grundsicherungsleistung forderte. Die CDU übernahm diese Position und forderte ein kommunal administriertes „Sozialgeld“ analog der Sozialhilfe (CDU-Bundesgeschäftsstelle, 2001, S. 115). Im Zuge des so genannten Vermittlungsskandals der BA wurden dann im Bundestagswahlkampf 2002 fundamentale Veränderungen des Grundsicherungssystems gefordert. Während die CDU das Grundsicherungssystem kommunalisieren wollte und darin von marktliberalen Ökonomen unterstützt wurde (Sinn, Holzner, Meister, Ochel & Werding, 2002), wollte die FDP die BA privatisieren (FDP-Bundesgeschäftsstelle, 2002, S. 6 f.).

Nun veröffentlichte der Arbeitskreis „Reform der Arbeitslosenhilfe und Sozialhilfe“ der Bertelsmann-Stiftung ein Positionspapier zur Neugestaltung des Grundsiche-

rungssystems. „Kernstück der Reform von Arbeitslosen- und Sozialhilfe“, so formulierte dieses Papier, müsse „die Schaffung eines einheitlichen Systems aktivierender und passiver Leistungen für die Zielgruppe der erwerbsfähigen Erwerbslosen sein. Dieses einheitliche, steuerfinanzierte System aktivierender und passiver Leistungen soll zudem sicherstellen, dass die Leistungsgewährung ‚aus einer Hand‘ erfolgt und nicht in verschiedene behördliche Strukturen bzw. Zuständigkeiten aufgeteilt ist“ (Bertelsmann-Stiftung, 2002, S. 4). In diesem Papier wurde erstmals die enge Advokaten-Koalition zwischen verantwortlichen Experten der Ministerialverwaltungen von Bund und Ländern, den Kommunalen Spitzenverbänden und Beratungsunternehmen sichtbar, die später eine Reihe von Aufträgen in diesem Bereich generierten. Der politische Einfluss der hier vernetzten Akteure wird unter anderem daran erkennbar, dass sie mehrere Papiere in die Hartz-Kommission einspeisten (Hartmann, 2002; Kruse, 2002).

In dem Positionspapier wurde diskutiert, ein einheitliches Grundsicherungssystem in bundesunmittelbarer Trägerschaft durch die BA, in kommunaler Selbstverwaltung oder in Bundesauftragsverwaltung durch die Kommunen umzusetzen. Sichtbar wurde, dass alle drei Organisationsmodelle zur Leistungsgewährung „aus einer Hand“ massive Verschiebungen der Aufgaben- und Finanzverteilung zwischen Bund, Ländern und Kommunen bedeuteten, die Mehrheiten in Bundestag und Bundesrat oder einer Veränderung des Grundgesetzes bedurften. Lösungsvorschläge für diese institutionelle Herausforderung hatte der Arbeitskreis allerdings nicht entwickelt (Bertelsmann-Stiftung, 2002).

Weder die Hartz-Kommission noch die Gemeindefinanzreform-Kommission legten einen konsensfähigen Organisationsvorschlag zur „Zusammenführung“ von Arbeitslosen- und Sozialhilfe vor (Kaps, 2008). Immerhin wurden in den Verhandlungen der letzteren die mit den verbliebenen Organisationsvarianten – Zentralisierung bei der BA in Bundesauftragsverwaltung oder Kommunalisierung in Selbstverwaltung – verbundenen finanziellen Folgen als der zentrale haushaltspolitische Streitpunkt sichtbar (AG Arbeitslosenhilfe/Sozialhilfe der Kommission zur Reform der Gemeindefinanzen, 2003).

Die rot-grüne Bundesregierung entschied sich in ihrem Entwurf für das „Vierte Gesetz für moderne Dienstleistungen am Arbeitsmarkt“ für die Zentralisierungsvariante mit der BA als Leistungsträgerin (Deutscher Bundestag, 2003). Die CDU/CSU-Opposition hielt mit dem Entwurf für ein „Existenzgrundlagen-Gesetz“ dagegen, das die Kommunalisierung der Aufgaben vorsah (Bundesrat 2003).

Letztlich einigten sich Bund und Länder im Vermittlungsausschuss auf einen Vorschlag des Deutschen Städtetags und des Deutschen Städte- und Gemeindebundes zu einer geteilten Leistungsträgerschaft als Ausweg aus dem haushaltspolitischen Verhandlungspatt (Wienand, 2003). Die Kommunen sollten einen Teil der Leistungen des neuen SGB II in eigene Trägerschaft übernehmen. Roland Koch forderte im Gegenzug für Kommunen ein Optionsrecht auf eigenverantwortliche Trägerschaft. „Nachts um halb Drei“ wurde schließlich „ein nicht ausformulierter Kompromissvorschlag von den Beteiligten zu dem vermeintlichen Königsweg erhoben“. Genau das aber hätten „eigenem Bekunden nach alle Beteiligten vor Beginn des Vermittlungsverfahrens unbedingt vermeiden wollen“ (Henneke, 2004, S. 7).

Der folgende Entwurf zum „Kommunalen Optionsgesetz“ sah vor, die optierenden Kommunen über eine Organleihe mit den Leistungen zu beauftragen und sie damit formal der Bundesagentur für Arbeit zuzuordnen. Dagegen argumentierten alle Kom-

munalen Spitzenverbände, die kommunale Selbstverwaltung würde gefährdet. Sie kündigten an, unter dieser Bedingung würden kaum Kommunen eine solche Option ziehen. Damit hätte die BA die Hauptlast der Einführung des SGB II tragen müssen, die das zu diesem Zeitpunkt aber nicht wollte, um den Prozess der Verwaltungsreform in den Arbeitsagenturen nicht zu gefährden (Deutscher Bundestag, 2004, S. 10 ff.).

Im Vermittlungsausschuss einigten sich beide Seiten schließlich darauf, eine begrenzte Zahl von Jobcentern in kommunaler Trägerschaft aller Leistungen des SGB II für einen begrenzten Zeitraum zuzulassen, wobei „die Erprobung (...) insbesondere auf alternative Modelle der Eingliederung von Arbeitsuchenden im Wettbewerb zu den Eingliederungsmaßnahmen der Agenturen für Arbeit ausgerichtet“ werden sollte (§ 6a SGB II i. d. F. von 2005). Letztlich wurde so aus einem Konflikt zwischen rot-grüner Bundesregierung und den konservativ regierten Ländern einerseits und zwischen Landkreisen und kreisfreien Städten andererseits um die Trägerschaft und die Finanzierung der neuen Grundsicherungsleistung ein „konstitutionelles Experiment“ zwischen zwei Organisationsmodellen öffentlicher Leistungsverwaltung, das „existenzielle Eigeninteressen der beteiligten Akteure auf allen Ebenen“ berührte (Kaps, 2009, S. 612 f.).

Die nach § 6b SGB II zugelassenen kommunalen Träger (sog. Optionskommunen) und die nach § 44b SGB II gegründeten Arbeitsgemeinschaften aus Arbeitsagentur und kommunalem Träger (ARGEn) begannen Anfang 2005 zu arbeiten. Im Sommer 2005 startete die in § 6c SGB II (a. F.) verankerte und groß angelegte zentrale Evaluierung der beiden Modelle der Aufgabenwahrnehmung im Auftrag des BMAS, an der in vier Konsortien insgesamt zwölf Forschungsinstitute beteiligt waren. Das BMAS ließ sich dabei von einem weiteren Forschungsinstitut administrativ unterstützen und wissenschaftlich beraten (Institut für Sozialforschung und Gesellschaftspolitik, 2006).

Die Evaluierung des Wettbewerbs zwischen den beiden Modellen der Leistungsträgerschaft war von Beginn an von politischen Deutungskämpfen begleitet, die deutlich vom Wettbewerb der Modelle geprägt waren. Der Deutsche Landkreistag (DLT) beauftragte im Sommer 2005 im Vorgriff auf die Ergebnisse der gesetzlichen Evaluierung nach § 6c SGB II ein anderes Institut mit einer eigenen Evaluationsstudie für seine Mitgliedskommunen, mit der er auf Vorrat Argumente sammelte, um sich bei Bedarf gegen die Ergebnisse der vom BMAS in Abstimmung mit den Ländern eingeleiteten gesetzlichen Evaluation stellen zu können (Hesse, Götz, Rüdiger & Schubert, 2008). Und er beauftragte die Bertelsmann-Stiftung und später die *con\_sens* GmbH damit, ein Benchmarking für die Optionskommunen zu entwickeln, mit dem er immer wieder an die Öffentlichkeit ging und für das Optionsmodell warb (Bertelsmann-Stiftung, 2007; Hartmann, Ahlrichs & Böttcher, 2008). Die Optionskommunen stellten zudem wiederholt die Kompetenz und Neutralität der BA in Frage, die gesetzlich mit der gesamten Statistik zum SGB II beauftragt ist, weil sie Nachteile im Wettbewerb der Modelle fürchteten, wenn die BA als Konkurrent die Hoheit für die Datengenerierung habe. Zur Umsetzung des Statistik-Auftrags der BA wurde deshalb ein Expertenkreis Statistik gebildet, in dem BA, BMAS, zuständige oberste Landesbehörden, kommunale Spitzenverbände, Statistisches Bundesamt und die Statistischen Landesämter vertreten waren und in dem zentrale Verhandlungen um Indikatoren und Kennzahlensets zum Controlling und Benchmarking im SGB II geführt wurden.

Die BA hielt sich mit Öffentlichkeitsarbeit zu den Arbeitsgemeinschaften eher zurück, wirkte allerdings nach innen in deren Handlungsspielräume deutlich steuernd hinein. Deshalb bildeten sich Landesarbeitsgemeinschaften der ARGE-Geschäftsführer

und ein „Bundesnetzwerk Arge SGB II“, um beim Bund die Interessen der ARGEn als eigenständige Grundsicherungsträger (auch gegen die Position der BA-Zentrale) wahrzunehmen und so ihre Chancen im Wettbewerb mit den Optionskommunen zu wahren (Bundesnetzwerk Arge SGB II, 2008).

Alle bis Mitte 2008 laufenden Evaluationsstudien zum Wettbewerb zwischen Arbeitsgemeinschaften und Optionskommunen wurden veröffentlicht. Die gesetzliche Evaluierung wurde durch mehrere Fachtagungen begleitet, in denen die Evaluierenden ihre Zwischenergebnisse vor Optionskommunen und Arbeitsgemeinschaften, Ländervertretern, kommunalen Spitzenverbänden und der BA zur Diskussion stellten. Die Evaluation ergab, dass die Unterschiede in der Umsetzung des SGB II zwischen Arbeitsgemeinschaften und Optionskommunen nicht gravierend waren. Insgesamt hatten Leistungsberichtigte in ARGEn eine um rund vier Prozentpunkte höhere Wahrscheinlichkeit, innerhalb von zwölf Monaten nicht mehr auf SGB II-Leistungen angewiesen zu sein als in Optionskommunen. Bei der Nachhaltigkeit der Integrationen schnitten die ARGEn um rund drei Prozentpunkte besser ab als die Optionskommunen (ZEW, IAQ & TNS emnid, 2008, S. 132). Die ARGEn waren etwas erfolgreicher, Männer in Arbeit zu vermitteln (ebd., S. 137 ff.), die Optionskommunen hingegen waren etwas erfolgreicher, Frauen in Arbeit zu vermitteln (ebd., S. 211 ff.). Eine eindeutige Erklärung für diese geringfügigen Unterschiede fand die Evaluation nicht. Die umfangreichen Informationen aus der zentralen Evaluation flossen nur teilweise in die Wissensbestände der Fach-Community ein.

Als das Bundesverfassungsgericht Ende 2007 die Arbeitsgemeinschaften für nicht vereinbar mit dem Grundgesetz erklärte und bis Ende Dezember 2010 eine grundlegende Neuorganisation der Grundsicherung für Arbeitsuchende forderte, änderte dies die institutionellen Rahmenbedingungen des Modellwettbewerbes grundlegend. Das Gericht sah zwei Lösungsmöglichkeiten (Bundesverfassungsgericht 2007, Abs. 172): Entweder der Bund führt das Gesetz nach Artikel 87 Grundgesetz mit einer eigenständigen Bundesbehörde aus<sup>5</sup> oder der Gesamtvollzug wird nach Artikel 83 Grundgesetz Länderangelegenheit. Eine geteilte Leistungsträgerschaft war demnach nicht möglich. Während die Evaluierung noch lief, entwickelten BMAS und BA das Modell eines „Kooperativen Jobcenters“, eine Bund-Länder-Arbeitsgruppe entwickelte ein Modell für „Zentren für Arbeit & Soziales“ (Kaps, 2008, S. 414).

Der zentrale Ergebnisbericht wurde zwar veröffentlicht, der Bundestag beschäftigte sich aber nicht damit und die Evaluierungsergebnisse fanden bei der Neugestaltung der Organisationsmodelle kaum Berücksichtigung – auch nicht die der Parallel-Evaluierung im Auftrag des DLT. Wesentlicher wurde nun, juristisch wasserdichte Lösungen für die Organisation der Jobcenter zu entwickeln, die auch den Normen des Grundgesetzes entsprachen.

Über einen längeren Zeitraum schien eine Einigung nicht möglich. Erst nach einem Regierungswechsel und weiteren zähen Verhandlungen zwischen den Advokaten-Koalitionen der Kommunalisierungsbefürworter und der Zentralisierungsbefürworter wurde mit dem „Gesetz zur Weiterentwicklung der Organisation der Grundsicherung für Arbeitsuchende“ im August 2010 die heutige Organisation der Grundsicherung für Erwerbssuchende installiert (Deutscher Bundestag 2010). Die Ergebnisse der Evaluierung spielten dabei keine Rolle.

Im Vergleich dieser Evaluierung mit den zuvor erörterten Beispielen sticht hervor, dass hier der Evaluationsauftrag zentralisiert war, mit erheblichen Ressourcen und der Beauftragung einer breiten Fach-Community einherging und von Beginn an unter brei-

ter Beobachtung durch eine große Zahl interessierter Akteure stand, die immer wieder auch ihre Positionen in die Medienöffentlichkeit trugen.

### 3 Vergleich der Evaluierungsarrangements und ihrer Rahmenbedingungen

#### 3.1 Vergleich zentraler institutionellen Rahmenbedingungen der Evaluierungen

Vergleicht man die institutionellen Rahmenbedingungen der Evaluierungen dieser drei Beispiele experimenteller Politikgestaltung, so lassen sich einige grundlegende Unterschiede konstatieren.

Während die beiden ersten Beispiele im Bereich der kommunalen Selbstverwaltung angesiedelt waren und durch Landesrecht konkretisiert wurden, war das dritte Beispiel auf der Ebene des Bundes angesiedelt. In allen drei Beispielen sah das Bundesgesetz einen Evaluierungsauftrag vor. Je stärker dezentralisiert der Evaluierungsgegenstand angesiedelt war, desto dezentraler war auch der gesetzliche Evaluierungsauftrag gestaltet. Im ersten Beispiel sollten die kommunal erprobten Ansätze „unter Beteiligung des [jeweiligen] Landes“ ausgewertet werden. Man könnte hier von einem zweidimensionalen *bottom-up*-Auftrag sprechen. Im zweiten Beispiel sollten die Ergebnisse unter Zusammenarbeit von Kommunen, Ländern und dem zuständigen Bundesministerium so ausgewertet werden, dass eine bundesweite Bewertung möglich würde (dreidimensionaler *bottom-up*-Auftrag). Im dritten Beispiel hatte das zuständige Bundesministerium den Auftrag, die Umsetzung zu untersuchen und „die Länder bei der Entwicklung der Untersuchungsansätze und der Auswertung der Untersuchung zu beteiligen“ (*top-down*-Auftrag). Zudem war nur in diesem Beispiel auch eine Berichtspflicht an Bundestag und Bundesrat gesetzlich festgeschrieben. Mit den in Abschnitt 2 vorgestellten Ergebnissen würde Schwabs These gestärkt, dass Mehrebenenkonstellationen die Abstimmung von Evaluierungsaktivitäten und die Entwicklung eines einheitlichen konzeptionellen Ansatzes erschweren (Schwab, 2017, S. 214).

Auch der Grad der „Institutionalisierung der Evaluierungsfunktion“ (Konzendorf, 2009) im Politikfeld hatte einen relevanten Einfluss auf das Evaluationsarrangement. Während im Bereich der Arbeitsmarktpolitik (Beispiele eins und drei) das IAB einen gesetzlichen Auftrag zu Arbeitsmarkt- bzw. Wirkungsforschung hat, gibt es keinen vergleichbaren Auftrag im Bereich der Sozialhilfepolitik, zu dem das zweite Beispiel zählt. Im BSHG galten zwar Regeln zur Generierung einer bundeseinheitlichen Sozialhilfestatistik, aber eine Regel zu wissenschaftlicher Begleitforschung existierte bis 2002 nicht. Erst dann wurde eine Rechtsgrundlage für den Datentransfer zu Forschungszwecken im Auftrag des Bundessozialministeriums eingeführt, eine institutionalisierte Evaluierungsfunktion war aber dennoch nicht gegeben. Dies änderte sich erst 2005 und nur für die Grundsicherung für Arbeitsuchende. In § 55 SGB II ist nun das IAB mit der regelmäßigen Wirkungsforschung beauftragt. Schwab (2017, S. 214) argumentiert, dass „eine institutionelle Verankerung [...] die Rolle von Evaluation [stärkt]“. Die von ihm angeführte „Gefahr einer Ritualisierung von Evaluierungsprozessen“ ließ sich in den hier aufgeführten Beispielen experimenteller Politik nicht er-

kennen. Inwiefern sie für die regelmäßige Wirkungsforschung des IAB besteht, das als Teil der BA von dieser nicht unabhängig ist, wäre gesondert zu untersuchen.

### 3.2 Vergleich der Evaluierungsarrangements

Hatten diese Unterschiede in den institutionellen Rahmenbedingungen Auswirkungen auf die Evaluierungsarrangements?

Im Vergleich zeigt sich, dass der Evaluierungsgegenstand im Sinne von „Eigenschaften der politischen Programme“ (Schwab, 2017, S. 215) die Rolle der Evaluierungen mitbestimmte. Das erste Beispiel griff Fragen der finanziellen und regulativen Ausgestaltung eines Instruments der Arbeitsförderung auf und kann als typisch für Programm-Evaluationen verstanden werden. Grundlegende Legitimationsfragen stellten sich dabei nicht, ein Scheitern der Kombilohn-Modelle hätte die Teilnehmenden nicht existenziell gefährdet. Das zweite Beispiel war mit materiellen Fragen der Bemessung der Existenzsicherung beschäftigt und berührte das grundgesetzlich geschützte Recht auf menschenwürdige Existenz. Derart existenzielle Gegenstände erzeugen ein hohes Interesse der von der Regelung Betroffenen, das allerdings nur dann sichtbar wird, wenn diese sich kollektiv für ihre Interessen einsetzen (können). Im zweiten Beispiel konnten die Betroffenen ihre durchaus berechtigten Interessen punktuell individuell, aber nicht kollektiv organisieren. Anders im dritten Beispiel, in dem es für Kommunen und die BA darum ging, wer die Grundsicherung als Leistungsträger zugeschrieben bekommt und wer diese Aufgabe „verlieren“ würde. Hier waren die Kommunalen Spitzenverbände als kollektive Akteure und die BA als interessierte Partei und zugleich mit der Generierung der notwendigen Daten beauftragter Umsetzungsakteur in die Ausgestaltung und Begleitung der Evaluierung einbezogen. Darüber hinaus organisierten sich auch die Jobcenter beider Umsetzungsmodelle über eigene Vertretungen und beteiligten sich aktiv an den Diskursen über die Evaluation. Hier zeigt sich die in der Evaluationsforschung thematisierte Bedeutung der Machtposition der Stakeholder bei der Gestaltung der Evaluierungsarrangements. „„Meaning‘ in evaluation research is not negotiated within a simple plurality of competing stakeholder interests. Different stakeholder views and interests are expressed from positions of more or less power“, argumentieren beispielsweise David Taylor und Susan Balloch (2005, p. 2).

Die Evaluierungen, die auf Sozialhilfe-Daten angewiesen waren, hatten einen schwierigeren Zugang zu Primärdaten als die Evaluierung im Bereich der Arbeitslosenversicherung und des SGB II, weil die Kommunen die Sozialhilfedaten dezentral und mit verschiedenen Softwaresystemen generierten und die Evaluierungen wesentlich auf zusätzliche Befragungen angewiesen waren. Dies führte dazu, dass eher Vorher-Nachher-Vergleiche als Kontrollgruppen-Vergleiche angestellt wurden. Ein Teil der Evaluierungen zu Kombilohn-Modellen konnte aber auf Daten der BA zugreifen (Kaltenborn, Krug, Rudolph, Weinkopf & Wiedemann, 2005, S. 13, 45). Deshalb konnten hier multivariate Analysen zum Kontrollgruppen-Vergleich auf der individuellen Ebene der Teilnehmenden eingesetzt werden. Die Evaluierung zum SGB II konnte auch auf zentral gehaltene Daten der BA zum SGB II zugreifen, weshalb ambitionierte wissenschaftliche Analysemethoden – neben multivariaten Analysen zum Kontrollgruppen-Vergleich auf der individuellen Ebene auch solche auf der Ebene der Jobcenter – eingesetzt werden konnten.



## 4 Fazit

Mit der vorliegenden Analyse konnten verschiedene Thesen Oliver Schwabs zur Rolle von Evaluierung in politisch-administrativen Systemen (Schwab, 2017) getestet werden. In der Gesamtbetrachtung der drei hier diskutierten Evaluierungen zeigt sich ein deutlicher Zusammenhang zwischen institutionellen Rahmenbedingungen und der Gestaltung der Evaluierungsarrangements. Je dezentralisierter der Evaluationsauftrag und je weniger institutionalisiert die Evaluationsfunktion war, desto weniger ambitioniert waren die eingesetzten Evaluierungsmethoden, desto weniger waren die Evaluationsergebnisse miteinander vergleichbar und desto unwahrscheinlicher war ihre vollständige Veröffentlichung. Zugleich hatten auch die Evaluationsgegenstände und die Verfügbarkeit von Datenquellen einen Einfluss auf die Evaluierungsarrangements.

Diese „varieties of evaluation“ (Schwab, 2017) erscheinen zudem davon abhängig, wie stark Akteure der jeweiligen Advokaten-Koalitionen politisches Interesse an der Deutung und der Veröffentlichung von Ergebnissen hatten. Die drei untersuchten Fälle unterlagen im Evaluierungsprozess unterschiedlich starken politischen Deutungskämpfen, die abhängig waren von der institutionellen Ebene und Umstrittenheit des Gegenstandes sowie dem Organisationsgrad und der Machtposition der handelnden Akteure.

## Anmerkungen

- 1 Aus Gründen der besseren Lesbarkeit wird im Folgenden ausschließlich die männliche Sprachform verwendet. Selbstverständlich sind alle Geschlechter gleichermaßen gemeint.
- 2 Zur grundlegenden Kritik an der Debatte um das Lohnabstandsgebot siehe Gerhard Bosch (1999).
- 3 Die Landesregierungen versprachen sich dabei vor allem finanzielle Unterstützung vom Bund, die dieser im Rahmen des Sonderprogramms CAST auch gewährte.
- 4 Damit passt das Beispiel zu den haushaltspolitisch motivierten Evaluierungsarrangements der zweiten Welle, wie Hellmut Wollmann (2004) herausgearbeitet hat.
- 5 Art. 87 Abs. 2 Grundgesetz sieht eigenständige Bundeseinrichtungen als Körperschaften öffentlichen Rechts für die Sozialversicherungen, aber nicht für Fürsorgesysteme, vor. Nach Art. 87 Abs. 3 Grundgesetz kann der Bund durch Bundesgesetz selbständige Bundesoberbehörden oder bundesunmittelbare Körperschaften errichten.

## Literatur

- AG Arbeitslosenhilfe/Sozialhilfe der Kommission zur Reform der Gemeindefinanzen (2003). *Bericht der Arbeitsgruppe „Arbeitslosenhilfe/Sozialhilfe“ der Kommission zur Reform der Gemeindefinanzen*.
- Andres, Gerd (2000). Über die Chancen, die Potenziale eines aktivierenden Sozialstaats in praktische Politik zu verwandeln. In Erika Mezger & Klaus W. West (Hrsg.) *Aktivierender Sozialstaat und politisches Handeln* (S. 41-49), Marburg: Schüren.
- Armborst, Christian (2003). § 101a Experimentierklausel. In Christian Armborst et al. (Hrsg.). *Bundessozialhilfegesetz. Lehr- und Praxiskommentar (LPK-BSHG) mit einer Kommentierung zum Asylbewerberleistungsgesetz* (S. 1149-1157). Baden-Baden: Nomos.
- Banner, Gerhard (2008). Logik des Scheiterns oder Scheitern an der Logik? Kommentar zu Lars Holtkamp ‚Das Scheitern des Neuen Steuerungsmodells‘. *dms – der moderne staat*, (1) 2, 447-455.
- Bender, Stefan & Rudolph, Helmut (1999). *Kosten eines gestaffelten Zuschusses zu den Sozialversicherungsbeiträgen. Simulation des Zuschußbedarfs auf der Basis des Jahreszeitraummaterials der Beschäftigtenstatistik*. IAB-Werkstattbericht 08/1999. Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung.

- Bertelsmann-Stiftung (2002). *Positionspapier: Eckpunkte einer Reform von Arbeitslosen- und Sozialhilfe*. Gütersloh: Bertelsmann-Stiftung.
- Bertelsmann-Stiftung (2007). *Jahresbericht 2006 für das SGB II-Benchmarking der Optionskommunen*. Gütersloh: Bertelsmann-Stiftung.
- Bertelsmann-Stiftung, Hans-Böckler-Stiftung & KGSt (Hrsg.) (2002). *Netzwerk Kommunen der Zukunft. Produkte der Netzwerkarbeit. Bd. 14. Kommunen und lokale Beschäftigungsförderung. Bausteine guter Praxis*. Frankfurt am Main: VAS.
- Blancke, Susanne (2004). *Politikinnovationen im Schatten des Bundes*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Boockmann, Bernhard, Brändle, Tobias, Fuchs, Philipp, Klee, Günther, Kugler, Philipp, Laub, Natalie, Sippli, Khira & Verbeek, Hans (2019). *Evaluation des ESF-Bundesprogramms zur Eingliederung langzeitarbeitsloser Leistungsberechtigter nach dem SGB II auf dem allgemeinen Arbeitsmarkt*. Zwischenbericht 31. März 2019.
- Bosch, Gerhard (1999). Niedriglöhne oder Innovation: Überlegungen zur Zukunft der Erwerbsarbeit. *WSI-Mitteilungen*, (52) 12, 861-869.
- Boss, Alfred (1994). *Explizite und implizite Besteuerung geringer Arbeitseinkommen - Aspekte der Armutsfalle in der Bundesrepublik Deutschland*. Kieler Arbeitspapier Nr. 643. Kiel: Institut für Weltwirtschaft.
- Bundesverfassungsgericht (2007). Urteil des Zweiten Senats vom 20. Dezember 2007 – 2 BvR 2433/04 - Abs. (1 - 228).
- Bundesverwaltungsgericht (2004). Urteil des 5. Senats vom 25. November 2004 – BVerwG 5 CN 1.03.
- Bundesnetzwerk Arge SGB II (2008). *Stellungnahme des Bundesnetzwerks Arge zur Geschäftsabweisung 13 / 2008 der Bundesagentur für Arbeit (BA)*. Verfügbar unter: [http://www.lag-arbeitsnrw.de/de/aktuelles/pdf-sgb/2008-04-25\\_Bundesnetzwerk\\_ARGE\\_SGB\\_II\\_Stellungnahme.pdf](http://www.lag-arbeitsnrw.de/de/aktuelles/pdf-sgb/2008-04-25_Bundesnetzwerk_ARGE_SGB_II_Stellungnahme.pdf) [12. Dezember 2019].
- Bundesrat (2003). *Gesetzesantrag der Länder Hessen, Baden-Württemberg, Bayern, Sachsen. Entwurf eines Gesetzes zur Sicherung der Existenzgrundlagen (Existenzgrundlagengesetz – EGG)*, Drucksache 654-03.
- CDU-Bundesgeschäftsstelle (1994). *Wir sichern Deutschlands Zukunft*. Regierungsprogramm von CDU und CSU.
- CDU-Bundesgeschäftsstelle (2001). *Neue Soziale Marktwirtschaft*. Diskussionspapier der CDU Deutschlands, 27. August 2001, Berlin.
- Deutscher Bundestag (1998). *Beschlußempfehlung und Bericht des Ausschusses für Gesundheit (14. Ausschuß) zu dem Gesetzentwurf der Bundesregierung – Drucksachen 13/10422, 13/10868 – Entwurf eines Ersten Gesetzes zur Änderung des Medizinproduktegesetzes (1. MPG-ÄndG)*, Drucksache 13/11021.
- Deutscher Bundestag (1999). *Beschlußempfehlung und Bericht des Ausschusses für Arbeit und Sozialordnung (11. Ausschuß) zu dem Gesetzentwurf der Bundesregierung – Drucksache 14/389 – Entwurf eines Siebten Gesetzes zur Änderung des Bundessozialhilfegesetzes*. Drucksache 14/820.
- Deutscher Bundestag (2001). *Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Pia Maier, Dr. Klaus Grehn, Dr. Ruth Fuchs, weiterer Abgeordneter und der Fraktion der PDS – Drucksache 14/7224 – Pauschalierung der Sozialhilfe*, Drucksache 14/7426.
- Deutscher Bundestag (2003). *Gesetzentwurf der Fraktionen SPD und BÜNDNIS 90/DIE GRÜNEN. Entwurf eines Vierten Gesetzes für moderne Dienstleistungen am Arbeitsmarkt*. Drucksache 15/1516.
- Deutscher Bundestag (2004). *Materialien für die öffentliche Anhörung von Sachverständigen am 26. April 2004 in Berlin zum Entwurf eines Gesetzes zur optionalen Trägerschaft nach dem Zweiten Buch Sozialgesetzbuch (Kommunales Optionsgesetz) – BT-Drs. 15/2816*. Zusammenstellung der schriftlichen Stellungnahmen. Ausschussdrucksache 15(9)1112.
- Deutscher Bundestag (2010). *Beschlußempfehlung und Bericht des Ausschusses für Arbeit und Soziales (11. Ausschuss) a) zu dem Gesetzentwurf der Bundesregierung – Drucksache 17/1940, 17/2057 – Entwurf eines Gesetzes zur Weiterentwicklung der Organisation der Grundsicherung für Arbeitsuchende b) zu dem Gesetzentwurf der Fraktionen der CDU/CSU, SPD und FDP –*

- Drucksache 17/1555 – Entwurf eines Gesetzes zur Weiterentwicklung der Organisation der Grundsicherung für Arbeitsuchende. Drucksache 17/2188.
- Empter, Stefan & Frick, Frank (Hrsg.) (1999). *Beschäftigungsorientierte Sozialpolitik in Kommunen. Strategien zur Integration von Sozialhilfeempfängern in das Erwerbsleben*. Gütersloh: Verlag Bertelsmann Stiftung.
- FDP-Bundesgeschäftsstelle (2002). *Bürgerprogramm 2002. Programm der FDP zur Bundestagswahl 2002 beschlossen auf dem 53. Ord. Bundesparteitag vom 10. bis 12. Mai 2002 in Mannheim*. Berlin: FDP-Bundesgeschäftsstelle.
- Felder, Bernd (2003). *Stellungnahme zur Anhörung „Entwurf eines Gesetzes zur Einordnung des Sozialhilferechts in das Sozialgesetzbuch“* – Bundestags-Drucksache 15/1514. (13) Ausschuss für Gesundheit und Soziale Sicherung. Ausschussdrucksache 0293. Berlin: Deutscher Bundestag.
- Fels, Gerhard, Heinze, Rolf G., Pfarr, Heide & Streeck, Wolfgang (1999). *Bericht der Wissenschaftlergruppe der Arbeitsgruppe Benchmarking über Möglichkeiten zur Verbesserung der Beschäftigungschancen gering qualifizierter Arbeitnehmer*. Verfügbar unter: [http://www.mpifg.de/people/ws/downloads/Bericht\\_Benchmarking-Gruppe.pdf](http://www.mpifg.de/people/ws/downloads/Bericht_Benchmarking-Gruppe.pdf) [6. April 2020].
- Fink, Ulf (1998). *Arbeit für Alle. Neue Initiativen zur Beschäftigungsförderung*. Bonn: Bouvier.
- Gerster, Florian & Deubel, Ingolf (2001). Das Mainzer Modell für Beschäftigung und Familienförderung. In Florian Gerster (Hrsg.). *Arbeit muss sich lohnen!* (S. 59-77). Baden-Baden: Nomos.
- Hartmann, Helmut (2002). *Organisationsmodell der Arbeitsförderung*. Revidierter Entwurf, erstellt für die Unterarbeitsgruppe „Reform der Arbeitslosen- und Sozialhilfe“ der Bertelsmann Stiftung, 25. Februar 2002. Hamburg: con\_sens GmbH.
- Hartmann, Helmut, Ahlrichs, Jan & Böttcher, Beate (2008). *Die 69 Optionskommunen der Bundesrepublik Deutschland. Benchmarking der Optionskommunen – Vergleichsring IV – Entwicklung und Tendenzen im SGB II. Berichtsjahr 2007*. Hamburg: con\_sens GmbH.
- Henneke, Hans-Günter (2004). Die Saga um „Hartz IV“: zur Reform der sozialen Sicherungssysteme und ihrer Umsetzung. *Zeitschrift für Staats- und Europawissenschaften*, (2) 4, 548-579.
- Hess, Doris, Schröder, Helmut, Smid, Meno & Reis, Claus (2004). MoZArT. *Neue Strukturen für Jobs. Abschlussbericht der wissenschaftlichen Begleitforschung*. Mit BMWA und den Ländern abgestimmte Endfassung. Bonn: infas Institut für angewandte Sozialwissenschaft.
- Hesse, Joachim Jens, Götz, Alexander, Rüdiger, Ronald & Schubert, Simon (2008). *Evaluation der Aufgabenträgerschaft nach dem SGB II 2005 – 2008. Abschlussbericht*. Untersuchung im Auftrag des Deutschen Landkreistages. Band 80 der Veröffentlichungen des Vereins für Geschichte der Deutschen Landkreise e.V., Berlin: Deutscher Landkreistag.
- Hoffmann, Susanne (2000). Hilfe durch Systemreform? *Bundesarbeitsblatt*, (51) 5, 5-10.
- Institut für Sozialforschung und Gesellschaftspolitik (2006). *Jahresbericht 2006 an das Bundesministerium für Arbeit und Soziales. Evaluation der Experimentierklausel nach § 6c SGB II – Vergleichende Evaluation des arbeitsmarktpolitischen Erfolgs der Modelle der Aufgabenwahrnehmung „zugelassene kommunale Träger“ und „Arbeitsgemeinschaften“*. Dienstleistungsauftrag: Administrative Unterstützung und wissenschaftliche Beratung, Köln: Institut für Sozialforschung und Gesellschaftspolitik GmbH.
- Kaltenborn, Bruno (2000). Reformkonzepte für die Sozialhilfe und ihre Konsequenzen für Fiskus und Arbeitsangebot. Mikroökonomische Analyse mit dem Sozio-ökonomischen Panel 1986 bis 1996 und dem Simulationsmodell SimTrans. *Mitteilungen aus der Arbeitsmarkt und Berufsforschung*, 33, 68-79.
- Kaltenborn, Bruno (2001). Kombilöhne in Deutschland – Eine systematische Übersicht. *IAB-Werkstattbericht Nr. 14*. Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung.
- Kaltenborn, Bruno, Gerhard, Rudolph, Helmut, Weinkopf, Claudia & Wiedemann, Eberhard (2005). *Evaluierung der arbeitsmarktpolitischen Sonderprogramme CAST und Mainzer Modell*. BMAS Forschungsbericht 552. Berlin: Bundesministerium für Wirtschaft und Arbeit.
- Kaps, Petra (2006). *Arbeitsmarktintegration oder Haushaltskonsolidierung. Interessen, Zielkonflikte und Ergebnisse kommunaler Beschäftigungspolitik*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kaps, Petra (2008). Die Grundsicherung in der Politikverflechtungsfalle? *dms – der moderne staat*, (1) 2, 399-422.

- Kaps, Petra (2009). Konstitutionelles Experiment als Governanceform. In Sebastian Botzem, Jeanette Hofmann & Sigrid Quack (Hrsg.). *Governance als Prozess. Koordinationsformen im Wandel* (S. 607-633), Baden-Baden: Nomos.
- Klinger, Roland (2000). Pauschalierung der Sozialhilfe wird in den Stadt- und Landkreisen verwirklicht. Reformvorschlag des Landkreistags Baden-Württemberg setzt sich durch. *Landkreisnachrichten Baden-Württemberg*, (39) 3, 127-131.
- Konzendorf, Götz (2009). Institutionelle Einbettung der Evaluationsfunktion in Politik und Verwaltung in Deutschland. In Thomas Widmer, Wolfgang Beywl & Carlo Fabian (Hrsg.) *Evaluation. Ein systematisches Handbuch* (S. 27-39), Wiesbaden: VS Verlag.
- Kruse, Jürgen (2002). *Reform von Arbeitslosenhilfe und Sozialhilfe. Modell einer Zusammenführung beider Systeme*. Präsentation von Ergebnissen einer Arbeitsgruppe unter Moderation der Bertelsmann-Stiftung. Köln, 7. März 2002.
- Lessenich, Stephan (2009). *Das Grundeinkommen in der gesellschaftspolitischen Debatte*. WISO Diskurs. Bonn: Friedrich-Ebert-Stiftung.
- Maaß, Volker (2001). *Experimentierklauseln für die Verwaltung und ihre verfassungsrechtlichen Grenzen: zugleich ein Beitrag zu § 7a BerlHG*. Berlin: Duncker & Humblot.
- Martens, Rudolf (2006). Neue Regelsatzberechnung 2006: Zu den Vorschlägen des Paritätischen Wohlfahrtsverbandes und der Bundesregierung. *Soziale Sicherheit*, (55) 6, 182-194.
- Mummert Consulting (2003). Stand der bundesweiten Evaluation. Zwischenergebnisse von Mummert Consulting. *Informationen, Termine und Notizen zu Modellvorhaben der Kommunen zur Pauschalierung der Sozialhilfe. PASO Brief 7 (9-13)*, Köln: Univation e.V..
- Nothbaum, Norbert, Lübker, Susanne & Kämper, Andreas (2004). *Die Konsequenzen einer pauschalieren Sozialhilfe für die HilfeempfängerInnen*. Zusammenfassung, Bielefeld: GOE Gesellschaft für Organisation und Entscheidung.
- Pautz, Hartwig (2008). Think-tanks in Germany: the Bertelsmann Foundation's role in labour market reform. *Zeitschrift für Politikberatung*, (1) 3/4, 437-456.
- Rothkegel, Ralf (2002). Der rechtliche Rahmen für die Pauschalierung von Sozialhilfeleistungen – insbesondere zur Experimentierklausel des 101 a BSHG. *ZFSH/SGB* (41) 10/11: 585-592, 657-668.
- Sabatier, Paul A. (1988). An Advocacy Coalition Framework of Policy Change and the Role of Policy Oriented Learning Therein. *Policy Sciences*, (21) 1, 129-168.
- Scharpf, Fritz W. (1997). Beschäftigungsfreundlich und sozial: ein Widerspruch? *Die Mitbestimmung*, (43) 4, 36-39.
- Schmachtenberg, Rolf (2003). *Zusammenlegung von Arbeitslosenhilfe und Sozialhilfe*. Vortrag auf der Fachtagung der nationalen Armutskonferenz (NAK) „Aktueller Stand der Arbeitsmarkt- und Sozialhilfereform“, Berlin, 14. Mai 2003.
- Schupp, Jürgen, Volz, Joachim, Wagner, Gert G. & Zwiener, Rudolf (1999). Zuschüsse zu den Sozialversicherungsbeiträgen im Niedriglohnbereich: wenig zielgerichtet und teuer. *DIW-Wochenbericht*, (66) 27, 499-509.
- Schwab, Oliver (2017). Varieties of Evaluation? Zur Rolle von Evaluierung in politisch-administrativen Systemen. Literaturbasierte Standortbestimmung und Ausblick. In Sabine Kuhlmann & Oliver Schwab (Hrsg.). *Starke Kommunen – wirksame Verwaltung. Fortschritte und Fallstricke der internationalen Verwaltungs- und Kommunalforschung* (S. 199-225). Wiesbaden: VS Verlag.
- Sell, Stefan (1999). Zur Schnittstellenproblematik zwischen Arbeitslosenversicherung und Sozialhilfe. Bestandsaufnahme der Risikoverlagerungen und Plädoyer für eine funktionsgerechte institutionelle und instrumentelle Modernisierung. *Sozialer Fortschritt*, (48) 1-2, 24-33.
- Siebert, Horst (1996). *Hundert Punkte für mehr Beschäftigung*. Kieler Diskussionsbeiträge 264. Kiel: Institut für Weltwirtschaft.
- Sinn, Hans-Werner, Holzner, Christian, Meister, Wolfgang, Ochel, Wolfgang & Werding, Martin (2002). Aktivierende Sozialhilfe. Ein Weg zu mehr Beschäftigung und Wachstum. *Ifo-Schnelldienst*, (55) 9, 7-14.
- SPD-Parteivorstand (1994). *Reformen für Deutschland*. Das Regierungsprogramm der SPD. Bonn.

- SPD & Bündnis 90/Die Grünen (1998). *Aufbruch und Erneuerung – Deutschland Weg ins 21. Jahrhundert*. Koalitionsvertrag zwischen der Sozialdemokratischen Partei Deutschlands und BÜNDNIS 90/DIE GRÜNEN.
- Spermann, Alexander (1996). Das „Einstiegsgehd“ für Langzeitarbeitslose. *Wirtschaftsdienst*, (76) 5, 240-246.
- Spindler, Helga (2004). Die neue Regelsatzverordnung – Das Existenzminimum stirbt in Prozentschritten. *Info also*, (22) 4, 147-166.
- Taylor, David & Balloch, Susan (2005). *The Politics of Evaluation. Participation and Policy Implementation*. Bristol: The Policy Press.
- Univation. Institut für Evaluation und wissenschaftliche Weiterbildung e.V. & Dr. Beywl & Associates GmbH (2005). *Pauschalisierte Leistungsgewährung in sozialen Sicherungssystemen – Erfahrungen und Konsequenzen*. Abschlussbericht der wissenschaftlichen Begleitung zu den Modellvorhaben zur Pauschalierung von Sozialhilfe in NRW. Köln.
- Wienand, Manfred (2003). Verzahnung von Arbeitslosen- und Sozialhilfe: Das Kooperationsmodell des Deutschen Städtetages. *Der Städtetag*, (56) 12, 9-10.
- Wollmann, Helmut (2004). Zur (Nicht-)Verwendung von Evaluationsergebnissen in Politik und Verwaltung. Eine vernachlässigte Frage der Evaluationsforschung. In Sabine Kropp & Sabine Kuhlmann (Hrsg.). *Wissen und Expertise in Politik und Verwaltung*, dms Sonderheft 1 (S. 87-102). Opladen: Barbara Budrich.
- ZEW, IAQ & TNS emnid (2008). *Evaluation der Experimentierklausel nach § 6c SGB II - Vergleichende Evaluation des arbeitsmarktpolitischen Erfolgs der Modelle der Aufgabenwahrnehmung „Zugelassener kommunaler Träger“ und „Arbeitsgemeinschaft“*. Untersuchungsfeld 3: „Wirktungs- und Effizienzanalyse“. Abschlussbericht. Mannheim, Gelsenkirchen, Bielefeld.

*Anschrift der Autorin:*

Petra Kaps, ZEP – Zentrum für Evaluation und Politikberatung, Mansteinstraße 8, 10783 Berlin, E-Mail: kaps@zep-partner.de.