

Wie misst man wohlfahrtsstaatliche Institutionen?

Eine innovative Methode zur Messung qualitativer Daten

Patricia Frericks, Julia Höppner und Ralf Och

Beitrag zur Ad-hoc-Gruppe »Wohlfahrtsstaatlichkeit als Erklärungskonzept in der ländervergleichenden Analyse sozialer Ungleichheiten: Konzeption, Messung und Befunde«

Einleitung

Institutionen rahmen maßgeblich unsere soziale Ordnung. Sie beeinflussen sowohl das Verhalten von individuellen und kollektiven Akteuren als auch politische und gesellschaftliche Strukturen wie soziale Ungleichheit. Diese prominente Position von Institutionen macht es erforderlich, sie zu analysieren und zu verstehen, wie sie Verhalten und Strukturen rahmen. Eine fundierte Institutionenanalyse hängt dabei nicht nur von einer klaren theoretischen und empirischen Grundlage ab, sondern auch von einer angemessenen Methode, die der Gegenstand dieses Beitrags ist.

In der vergleichenden Sozialforschung gibt es zahlreiche qualitative und quantitative Methoden. Zur Messung und Analyse von Institutionen sind diese jedoch aus verschiedenen Gründen nur beschränkt geeignet. Qualitative Ansätze sind im Allgemeinen durch ihren narrativen Charakter auf die nuancierte Analyse von Institutionen bei kleinen Fallgrößen ausgerichtet. Quantitative Ansätze hingegen eignen sich in erster Linie zur Messung von Indikatoren mit metrischem Skalenniveau und können nicht-metrische Merkmale von Institutionen nur binär oder in groben Kategorien erfassen. Dieser Beitrag zielt darauf ab, dieses Defizit in den verfügbaren Methoden zur Messung von Institutionen zu reduzieren: Wir präsentieren im Folgenden eine innovative Methode für den systematischen Vergleich von Institutionen. Im Zentrum dieser Methode steht dabei die Quantifizierung qualitativer Daten, die auf eindeutig definierten Regeln zur Erfassung der Komplexität institutioneller Merkmale basiert.

Entwickelt haben wir diese Methode im Kontext der Analyse wohlfahrtsstaatlicher Institutionen (Anwendung in Frericks et al. 2016); sie kann unserer Einschätzung nach aber ohne weiteres auch für die Messung und den Vergleich von Institutionen anderer Politik- und Gesellschaftsbereiche verwendet werden.

Im nächsten Abschnitt des Beitrags werden die verschiedenen Methoden diskutiert, die bisher für den Vergleich von Institutionen genutzt worden sind. Danach präsentieren wir die von uns entwickelte Methode zur vergleichenden Analyse von Institutionen und schlagen einige statistische Maßzahlen zur Zusammenfassung der Daten vor, die insbesondere zur vergleichenden Analyse mehrerer Institutionen sinnvoll eingesetzt werden können. Der letzte Abschnitt des Beitrags fasst die gewonnenen Erkenntnisse zusammen.

Forschungsstand

Um Unterschiede und Entwicklungen von Institutionen erfassen zu können, ist eine geeignete Konzeptualisierung und Operationalisierung notwendig (Clasen, Clegg 2007: 166). Hierzu sind aus einer methodologischen Perspektive zwei Schritte erforderlich (Green-Pedersen 2007): Der erste beinhaltet eine präzise theoretische Definition des zu analysierenden Gegenstands, um zu klären, *was* überhaupt gemessen werden soll. Dies kann beispielsweise das Ausmaß sein, zu dem Rechte auf Sozialleistungen oder soziale Dienstleistungen universell ausgestaltet sind. Der zweite Schritt erfordert eine Operationalisierung des zu messenden Gegenstands, also die Klärung, *wie* genau gemessen werden soll.

In unserem Fall erfordert der erste Schritt, also die theoretische Definition dessen, *was* gemessen werden soll, eine klare Definition des Begriffs „Institution“. Institutionen können recht allgemein als „[...] the humanly devised constraints that structure political, economic and social interaction“ interpretiert werden (North 1991: 97). In der gegenwärtigen Institutionentheorie unterscheidet man zwischen drei grundlegenden Strömungen: dem soziologischen, dem Rational Choice und dem historischen Neo-Institutionalismus, die sich alle drei durch ein spezifisches Institutionenverständnis auszeichnen (Hall, Taylor 1996). Für die Zwecke unserer Methode definieren wir Institutionen in Übereinstimmung mit Hall (1986: 19) als „[...] formal rules, compliance procedures, and standard operating practices that structure the relationship between individuals in various units of the polity and economy“. Dabei folgen wir Streeck und Thelen (2005: 12) darin, Institutionen zu verstehen als „rules that can and need to be implemented and [...] if necessary be enforced by agents acting on behalf of the society as a whole“. Damit verwenden wir eine relativ enge Konzeptualisierung von Institutionen als formale Institutionen und schließen informelle Institutionen für den Einsatz unserer Methode zunächst einmal aus. Diese Definition von Institutionen ist hilfreich, da sie eine klare Abgrenzung gegenüber anderen Variablen oder Faktoren erlaubt und so eine empirisch eindeutige Operationalisierung ermöglicht. Anders ausgedrückt muss zwischen Institutionen selbst bzw. institutionellen Regeln auf der einen Seite und ihren Effekten oder ihrer ideellen Herkunft auf der anderen Seite unterschieden werden, um eine trennscharfe Analyse institutioneller Merkmale durchführen zu können (Thelen, Steinmo 1992: 5).

Die verschiedenen sozialen, politischen oder ökonomischen Institutionen bestehen aus multi-dimensionalen Regelungen, die sich als Rechte und Pflichten verstehen lassen. Institutionelle Analysen beziehen sich entsprechend auf das (Rechts-)Subjekt und die Bedingungen, von denen die Anwendung des Rechts abhängt (Hall, Soskice 2001; Streeck, Thelen 2005; Campbell 2011; Frericks 2013). Ausgehend von dieser theoriebasierten Definition lassen sich die zentralen Analysedimensionen eines institutionellen Vergleichs bestimmen. So führt zum Beispiel Pierson aus, dass drei Formen von Wandel in den Rechten und Pflichten wohlfahrtsstaatlicher Institutionen vorliegen müssen, damit man von Wohlfahrtsstaatsrückbau (retrenchment) sprechen kann: „(1) significant increase in the reliance on means-testing; (2) major transfers of responsibility to the private sector and (3) dramatic changes in benefit and eligibility rules that signal a qualitative reform of a particular program“ (Pierson 1996: 157).

Im Anschluss an die Definition von Institutionen ist zu bestimmen, *wie* Institutionen zu messen sind, da eine theoretische Definition für eine eindeutige empirische Operationalisierung nicht ausreicht (Green-Pedersen 2007: 16). Und tatsächlich gibt es eine Vielzahl an qualitativen und quantitativen Ansätzen zur Operationalisierung und Messung von Institutionen.

Ein häufig genutztes Instrument zur Analyse und zum Vergleich von Institutionen sind Typologien (Ebbinghaus 2011). Die entwickelten Typologien können Weberschen Idealtypen entsprechen, und

nutzen damit klar definierte Begrifflichkeiten, um die Realität analytisch eindeutig zu erfassen (Weber 1985). Sie stellen damit ein heuristisches Instrument für die soziologische Analyse und Systematisierung empirisch-historischer Realitäten dar. Wenn wir Institutionen bzw. die uns an ihnen interessierenden Merkmale als Idealtypen konzeptualisieren, steht uns damit ein klarer, theoretisch bestimmter Fixpunkt zu ihrer Messung zur Verfügung. Anhand dessen lässt sich die Distanz eines untersuchten empirischen Falls zu diesem Fixpunkt und zu den anderen Fällen eindeutig erfassen (Weber, Winckelmann 1985: 213). Im Folgenden präsentieren wir die wichtigsten auf Idealtypen basierenden Methoden zur Messung von Institutionen.

Die vielleicht einfachste Art, um Institutionen zu messen und zu vergleichen, ist die Durchführung qualitativer Fallstudien, ein Ansatz, der häufig mit einer Idealtypenanalyse kombiniert wird (Hall, Soskice 2001). Das übliche methodische Vorgehen besteht hierbei darin, zunächst den Institutionenbegriff im Kontext der Forschungsfrage zu definieren, dann die Indikatoren für die zu untersuchenden Institutionen zu operationalisieren und abschließend die Unterschiede zwischen den untersuchten Ländern oder Jahren zu beschreiben. Der Vorteil dieses Vorgehens liegt dabei in der detaillierten Erfassung und dem systematischen Vergleich von Institutionen (Bennett, Elman 2006). Der narrative Charakter von Fallstudien bringt jedoch auch einige bedeutende Nachteile mit sich. Einer besteht darin, dass ein Vergleich von mehr als ein paar Fällen kaum umzusetzen ist, weil die detaillierten und umfangreichen Beschreibungen die Vergleichbarkeit stark erschweren. Um dies zu kompensieren, werden oft Kategorien mit recht groben Differenzierungen gebildet. In der Wohlfahrtsstaatsanalyse wird beispielsweise häufig zwischen einer niedrigen, mittleren und hohen Ausprägung eines bestimmten Merkmals unterschieden (Pfau-Effinger et al. 2009), und im Varieties of Capitalism (VoC)-Ansatz (Hall, Soskice 2001) zwischen zwei Idealtypen von politischen Ökonomien, der liberalen und der koordinierten Marktwirtschaft. Der Vorteil des VoC-Ansatzes besteht darin, dass er den Vergleich einer Vielzahl von Institutionen ermöglicht. Er wurde jedoch aufgrund seiner binären Herangehensweise vielfach kritisiert (Kogut 2010). Der Hauptkritikpunkt an qualitativen Datenanalysen liegt kurz gefasst darin, dass sie entweder durch sehr präzise Beschreibungen einen genauen Vergleich für eine nur geringe Fallzahl erlauben oder aber durch die Bildung sehr breiter Kategorien empirische Varianz nicht hinreichend erfassen. Dies kann auch dann der Fall sein, wenn der Forschungsgegenstand idealtypisch definiert wird (Alber 1996; Green-Pedersen 2007: 19), da eine konsistente Messweise über verschiedene Untersuchungseinheiten hinweg selten generiert wird (Scruggs 2007: 135). Der Fallstudien-Ansatz ist also nützlich, um qualitative Unterschiede zwischen Institutionen aufzudecken. Für den systematischen Vergleich mehrerer Institutionen oder Fälle fehlt ihm jedoch eine eindeutige, nuancierte und konsistente Messweise.

Dieses Problem spielt bei der quantitativen Messung von Institutionen eine deutlich geringere Rolle. Diesen Ansatz wenden insbesondere die zahlreichen Studien zur Wohlfahrtsstaatsgenerosität an, deren bekanntestes Beispiel die Analyse Esping-Andersens ist (1990), dessen Typologie verschiedentlich weiterentwickelt wurde (siehe für einen Überblick Ferragina, Seeleib-Kaiser 2011). Bei den quantitativen Ansätzen werden in der Regel zahlenbasierte Merkmale von Institutionen wie Lohnersatzraten als Indikatoren verwendet. Dieses Vorgehen erlaubt somit statistische Analysen für große Fallzahlen; es birgt jedoch zwei gravierende methodologische Schwächen. Zum einen werden die Grenzen zwischen institutionellen Regulierungen und deren Folgen häufig verwischt. Zum anderen werden institutionelle Merkmale, die sich nicht in metrische Skalen transformieren lassen, wie Anspruchsvoraussetzungen, als kategoriale Variablen kodiert, die die Komplexität nicht-metrischer Variablen nicht abbilden können und somit, ähnlich wie bei qualitativen Fallstudien, zu einer Unschärfe in der Messung von Institutionen führen.

Methodologisch basieren auch viele für die Wohlfahrtsstaatsanalyse entwickelte Datensätze, wie das Social Citizenship Indicator Programme (SCIP) oder das Comparative Welfare Entitlements Dataset (CWED), auf der quantitativen Operationalisierung von Institutionen. Sie bieten den Vorteil, dass sie einen einfachen Zugang zu relevanten Daten für eine große Zahl an Ländern bereitstellen (Grünewald 2014). Sie haben aber auch drei zentrale Nachteile. Erstens können natürlich nicht alle möglichen Institutionen oder institutionellen Merkmale durch die Daten abgedeckt werden, sodass Daten für die jeweils interessierenden Merkmale gegebenenfalls nicht zur Verfügung stehen. Zweitens werden institutionelle Regulierungen, die sich nicht als metrische Variable erfassen lassen, in der Regel als kategoriale Variablen kodiert, was den oben geschilderten Nachteil der nur recht groben Differenzierung mit sich bringt. Drittens entspricht die Kodierung der Variablen einer spezifischen, von den Entwickler/innen des Datensatzes zugrunde gelegten Definition der institutionellen Regulierungen, die mit den Idealtypen der eigenen Untersuchung eventuell nicht übereinstimmt. Dies kann zu theoretischer Ungenauigkeit und – je nach Operationalisierung – zu Unterschieden in der Messung führen (siehe Ferrarini et al. 2013) und zeigt wiederum, dass sich Unterschiede in der genauen Bestimmung der Institutionen auf die Messung auswirken können.

Ein vollkommen anders ausgestalteter Ansatz zur quantitativen Messung von Institutionen ist *fuzzy set* QCA (Qualitative Comparative Analysis). QCA wurde ursprünglich entwickelt, um methodische Probleme bei der Analyse von mittleren Fallzahlen zu überwinden (Ragin 2000: 23). Es ist eine ausgefeilte Methode zur Identifizierung von so genannten notwendigen und hinreichenden Bedingungen sowie der idealtypischen Kombinationen dieser Bedingungen für die Erklärung eines beobachteten Phänomens. Obwohl QCA für Kausalanalysen entwickelt wurde, kann es wie bei Kvist (2007) auch für deskriptive Zwecke verwendet werden. Zwei Eigenschaften von QCA relativieren jedoch seine Eignung für die deskriptive Messung von Institutionen. Erstens bietet QCA keine Richtlinie, anhand derer die Werte der Indikatoren berechnet werden können. Damit wird dem/der Forscher/in kein Anhaltspunkt dafür gegeben, wie sich qualitative Regelungen in Zahlwerte transformieren lassen (Wagemann 2008: 245). Diese methodische Schwäche schuldet sich dem Fokus von QCA auf *Kausalanalysen* und den damit im Vordergrund stehenden notwendigen und hinreichenden Bedingungen. Auch ein zweites Charakteristikum von QCA ist für unsere Zwecke unvorteilhaft: Die Fälle sind mengentheoretisch als Mitglieder von Sets konzipiert. Dadurch wird der mittlere Wert jeder Skala zur zentralen Bruchstelle, an der der Wert weder dem einen noch dem anderen Set zugeordnet werden kann, was maximale Unschärfe bedeutet (Kogut 2010: 165f.). Eine solche Konzeptualisierung macht nur dann Sinn, wenn die Unterscheidung zwischen zwei Gruppen von Institutionen das Ziel der Analyse ist (Mitglied eines Sets oder nicht). Sollte dies nicht der Fall sein, ist eine Konzeptualisierung angemessener, die eine kontinuierliche Zuordnung der Werte auf einer kontinuierlichen Skala erlaubt.

Zusammenfassend lässt sich sagen, dass es eine Vielzahl an Methoden zur Messung von Institutionen gibt, die sowohl qualitative als auch quantitative Operationalisierungen heranziehen. Dabei sind qualitative Ansätze von geringen Fallzahlen geprägt und quantitative Ansätze von Indikatoren mit metrischem Skalenniveau. Die Kombination quantitativer und qualitativer Methoden in Form von QCA wiederum bietet sich für unsere Zwecke nicht an, da sie weder die Vergabe der numerischen Werte für qualitative Indikatoren klärt, noch als Gradmesser auf einer kontinuierlichen Skala verwendet werden kann. Unter den bestehenden Methoden zur vergleichenden Analyse von Institutionen fehlt es also bislang an einer, die eine nuancierte Messung von Institutionen erlaubt, für eine große Fallzahl geeignet ist und unterschiedliche Skalenniveaus von Indikatoren berücksichtigt. Wir präsentieren im Folgenden die Methode, die wir explizit für diese drei Anforderungen entwickelt haben.

Methode für die vergleichende Messung von Institutionen (CMI)

In diesem Abschnitt präsentieren wir die von uns entwickelte Methode für die vergleichende Messung von Institutionen (Comparative Measurement of Institutions, CMI). Diese Methode basiert auf theoretisch bestimmten Idealtypen, die als Maßstab für die Messung und den systematischen Vergleich von Institutionen dienen. Im Folgenden stellen wir die einzelnen Schritte zur Anwendung dieser Methode vor.

Bestimmung der Idealtypen

Die Definition eines Idealtyps hängt von der theoretischen Perspektive der Studie ab. Die in unserer Methode notwendigen Idealtypen können als Webersche Idealtypen verstanden werden, d.h. sie sind nicht „Ziel, sondern [...] Mittel“ zur Analyse der Wirklichkeit und werden durch die „gedankliche Steigerung bestimmter Elemente der Wirklichkeit gewonnen“, ohne in ihrer „volle[n] begriffliche[n] Reinheit“ empirisch vorkommen zu müssen (Weber 1985: 190-194). Idealtypen Weberscher Art dienen somit als Referenzpunkte für die Messung von Institutionen und die Distanz zwischen den untersuchten Fällen und gewährleisten so die Vergleichbarkeit von verschiedenen Fällen oder über einen längeren Zeitraum hinweg (Aspalter 2011: 736). Sie müssen für unsere Methode so gefasst werden, dass sie eine Möglichkeitsskala mit zwei Enden bilden (Achsenausprägung zwischen voller und keiner Übereinstimmung mit dem Idealtyp). Die empirischen Fälle lassen sich zwischen diesen beiden Extremen einordnen, wobei die Idealtypen als Endpunkte einer *kontinuierlichen Skala* zu verstehen sind. Je nach Untersuchungsgegenstand können Idealtypen als Endpunkte mehrerer aufeinander bezogener und theoretisch hergeleiteter *separater Skalen* gefasst werden (siehe nächster Abschnitt). Durch diese Offenheit bei der Interpretation der Skalen ist unsere Methode deutlich ergebnisoffener als die Idealtypen-Ansätze des VoC oder der Wohlfahrtsregimetypologien, bei denen eine bestimmte Kombination von Dimensionsmerkmalen jeweils einem Idealtyp zugeordnet wird. Da unsere Methode mit kontinuierlichen Skalen arbeitet und keine Zuordnung zu einem Set erzwingt, ist sie darüber hinaus auch flexibler und weniger willkürlich als QCA.

Dimensionen

Den nächsten Schritt bildet die Bestimmung der für die Untersuchung relevanten und theoretisch abgeleiteten Dimensionen. Diese helfen, Unterschiede und Gemeinsamkeiten zwischen den einzelnen Untersuchungsobjekten zu analysieren, die durch bestimmte Kombinationen von Dimensionswerten gekennzeichnet sind (Kluge 2000). Aus diesem Grund müssen die Dimensionen sorgfältig bestimmt werden, was ein vertieftes Wissen über den Forschungsgegenstand erfordert (Flick 2009: 135). Die Dimensionen entsprechen den kontinuierlichen Skalen, die sich zwischen den idealtypischen Eckpunkten erstrecken. Die Anzahl der idealtypischen Eckpunkte beträgt 2^k , wobei k die Zahl der Dimensionen ist (Lazarsfeld 1937). Abhängig von der Anzahl der Dimensionen ist das Möglichkeitsspektrum unterschiedlich ausgeprägt und kann einen zwei-, drei- oder mehrdimensionalen Merkmalsraum bilden (Schnell et al. 2005: 168). Die einzelnen Dimensionswerte, die auf Basis des unten beschriebenen Verfahrens bestimmt werden, lassen sich als Koordinaten im Merkmalsraum interpretieren (zwei Werte für eine Koordinate im zweidimensionalen Raum, drei Werte für eine Koordinate im dreidimensionalen Raum usw.).

Indikatoren

Nachdem die Dimensionen theoretisch bestimmt wurden, muss deren Messung nun operationalisiert werden. Um die Ausprägung des untersuchten Phänomens auf den Dimensionen erfassen zu können, werden zunächst die relevanten Indikatoren identifiziert. Die Indikatorwerte werden dann nach dem untenstehenden Verfahren gemessen und später zu den Dimensionswerten aggregiert (siehe Aggregation). Die Auswahl der Indikatoren ist ein entscheidender Schritt für jede empirische Analyse. An erster Stelle sollten die Indikatoren valide sein, damit sie das messen, was sie messen sollen (Figueredo et al. 2013: 349). Zudem sollten die gewählten Indikatoren über alle Fälle und untersuchten Zeitpunkte vergleichbar sein (siehe Mahoney 2004: 95f.). In unserer Methode sind Indikatoren *alle* institutionellen Regelungen, die die Ausprägung eines Dimensionswertes beeinflussen. Die Identifizierung der jeweils relevanten Indikatoren basiert dabei auf den empirischen Daten.

Messung

Im Rahmen der Wertevergabe wird bestimmt, in welchem Ausmaß ein Indikator einem Idealtyp entspricht. Unserem methodischen Ansatz folgend können die Indikatoren Werte zwischen 0 (keinerlei Übereinstimmung mit dem Idealtyp, eventuell als eigener gegenüberliegender Idealtyp gefasst) und 1 (volle Übereinstimmung) auf einer kontinuierlichen Skala annehmen. Zudem können Indikatoren, wie oben beschrieben, auf einer metrischen oder ordinalen Skala gemessen werden. Wenn sich institutionelle Regelungen beispielsweise in Form von Geldbeträgen messen lassen, sind die Werte metrisch und recht einfach zu bestimmen. Wenn sich institutionelle Regelungen hingegen nur qualitativ erfassen lassen, sind die Indikatorwerte ordinal. Geschlechterspezifische Anreize zur familialen Kinderbetreuung (Javornik 2012) oder politische Freiheitsrechte (Fedderke, Garlick 2011) sind zwei Beispiele für ordinale Skalen. Die Bestimmung ordinaler Indikatorwerte ist dabei komplexer als die metrischer Indikatoren, da zunächst geeignete und vergleichbare Kriterien zu deren Bestimmung entwickelt werden müssen. Diese sollten in Form eines Kodierleitfadens eine konsistente Wertevergabe über die untersuchten Fälle hinweg gewährleisten.

Gewichtung

Nach einer sorgsam Bestimmung des Indikatorwerts muss berücksichtigt werden, dass zusätzliche Bedingungen, die in institutionellen Regelungen festgelegt sind, diesen Wert gemessen am Idealtyp relativieren können. So kann, um auf ein Beispiel aus der Wohlfahrtsstaatsforschung zurückzugreifen, der Anspruch auf Kindergeld davon abhängen, ob das Kind im gleichen Haushalt wohnt wie der/die Antragsteller/in, wodurch das Kriterium der Familienzugehörigkeit relativiert wird, da es für den Anspruch nicht allein ausschlaggebend ist (Frericks et al. 2016). Die Indikatorwerte müssen in solchen Fällen auf Basis eines systematischen Verfahrens relativiert, das heißt reduziert werden. Die Literatur zur Indexbildung befasst sich intensiv mit der Verwendung von Gewichten. Einige Autor/innen gehen davon aus, dass explizite Gewichtungen der relativen Bedeutung der einzelnen Indikatoren am besten Rechnung tragen. Andere Forscher/innen sind der Ansicht, dass dies die Ergebnisse verzerrt und eine implizite Gewichtung im Rahmen der Wertevergabe geeigneter sei (Booyen 2002: 126ff.). Für unsere Methode empfehlen auch wir eine implizite Gewichtung, die noch vor der Aggregation vorgenommen wird. Bei der Bestimmung des Gewichts sollte berücksichtigt werden, dass die Indikatorwerte umso kleiner ausfallen, je niedriger das Gewicht ist. Dieser Effekt zeigt sich insbesondere dann, wenn man mehrere Bedingungen berücksichtigt, also mehrfach gewichtet werden muss. Das für die Messung geeignete Gewicht hängt somit in mehrfacher Hinsicht vom untersuchten Gegenstand ab, und wir

empfehlen, die Berechnung der Indikatorwerte mit unterschiedlichen Gewichten durchzuführen, um die Robustheit der Gewichtung zu überprüfen.

Aggregation

Abschließend müssen die gewichteten Indikatorwerte zum Dimensionswert aggregiert werden. Welcher Aggregationsmethode der Vorzug gegeben wird, hängt von den jeweiligen Merkmalen der Dimensionen ab. Bei unserer idealtypischen Vorgehensweise sind die Stärke der Ausprägung eines Indikators und die Anzahl der Indikatoren ausschlaggebend dafür, wie hoch der Dimensionswert insgesamt ausfällt. Dem trägt die Addition Rechnung.

Bei einzelnen Indikatoren kann es vorkommen, dass es Subindikatoren gibt, sich also verschiedene institutionelle Regelungen auf den gleichen Gegenstand beziehen. Hier ist es möglich, dass sich die Bedingungen für diese Regelungen gegenseitig ausschließen und eine Addition entsprechend logisch inkorrekt wäre (wenn beispielsweise die Bedingung für einen Anspruch die Scheidung bzw. der Tod des/der Partners/in ist). Den Indikatorwert berechnen wir in diesen Fällen als das arithmetische Mittel der Subindikatoren.

Zur genaueren Messung der Dimensionswerte sollten die einzelnen Indikatorwerte auf zwei Nachkommastellen erfasst werden. Bei der Aggregation der Indikatorwerte hingegen kann es für die Interpretation hilfreich sein, die Dimensionswerte auf eine Nachkommastelle zu runden. Dadurch ergibt sich für die Dimensionen eine elfstufige Skala (0,0 bis 1,0).

Reliabilität

Damit die auf diese Weise gemessenen Werte tatsächlich zuverlässige Ergebnisse liefern, ist es von Bedeutung, dass die Wertevergabe, Gewichtung und Aggregation reliabel sind. Um die Reliabilität zu verbessern, empfehlen wir das Verfahren der Intercoder-Reliabilität für das gesamte Messverfahren. Dies bedeutet, dass die Richtlinien für die Wertevergabe so klar und eindeutig sind, dass verschiedene Coder zu den gleichen oder sehr ähnlichen Messungen der Indikator- und aggregierten Dimensionswerte kommen. Falls dies nicht der Fall ist, sollten die Richtlinien überarbeitet und klarer gefasst werden (Schreier 2012: 166ff.).

Unsere Methode bietet eine gute Ausgangsbasis für eine hohe Reliabilität, da die Dimensionen auf theoretisch bestimmten Idealtypen und den damit einhergehenden Skalen mit Endpunkten beruhen, wodurch der Spielraum für eine alternative Wertevergabe für Indikatorwerte stark eingeschränkt ist. Zweitens erfordert die Methode eine umfassende Erschließung der Daten, da die aggregierten Dimensionswerte *alle* relevanten institutionellen Regelungen beinhalten sollten, was wiederum die Gefahr abweichender Dimensionswerte minimiert. Nichtsdestotrotz ist es ein grundlegendes Merkmal qualitativer Forschung, dass Interpretationsspielräume bestehen bleiben. Das vorgeschlagene Verfahren gewährleistet jedoch, dass die es anwendenden Untersuchungen nicht zu substantiell abweichenden Ergebnissen führen können.

Verwendung der Dimensionswerte für weitere Analysen

Die auf Basis unserer Methode erzielten Dimensionswerte können für weiterführende statistische Analysen verwendet werden, wenn deren Anwendung im Forschungskontext sinnvoll erscheint. So ist es beispielsweise möglich, verschiedene Lage- und Streuungsmaße zu berechnen, um durchschnittliche Merkmalsausprägungen und Unterschiede zwischen den untersuchten Fällen in verschiedenen Analysebereichen abzubilden. Außer für solche deskriptiven Verfahren können die gewon-

nenen Dimensionswerte auch für induktive multivariate Verfahren eingesetzt werden, wenn es das Ziel ist, die erfasste institutionelle Ausgestaltung als abhängige oder unabhängige Variable zu verwenden.

Zusammenfassung

In diesem Beitrag haben wir eine innovative Methode für den systematischen Vergleich von Institutionen (CMI) vorgestellt. Wir haben diese Methode im Rahmen der vergleichenden Wohlfahrtsstaatsanalyse entwickelt, halten sie jedoch auch für ein geeignetes Messinstrument zur Analyse von Institutionen anderer Gesellschafts- und Politikbereiche.

Bei der Entwicklung unserer Methode haben wir die grundlegenden Fragen danach gestellt, *was* gemessen wird und *wie*. Der Gegenstand unserer Methode sind formale Institutionen, da diese eine recht unproblematische und konsistente Messung von Institutionen als Variablen erlauben. Wir gehen jedoch davon aus, dass abhängig vom theoretischen und empirischen Gegenstand auch informelle Institutionen mit unserer Methode erfasst werden können.

In Bezug auf die Frage, *wie* gemessen wird, haben wir eine Reihe an Methoden diskutiert, die zur Messung von Institutionen verwendet werden. Hierbei haben wir uns auf solche konzentriert, die mit Idealtypen arbeiten, um empirische Fälle anhand von theoretisch klar bestimmten Fixpunkten zu messen. Unter diesen Ansätzen eignen sich Fallstudien aufgrund ihres narrativen Charakters zur nuancierten Analyse von Institutionen bei kleinen Fallzahlen. Quantitative Ansätze wiederum erlauben zwar die Analyse von größeren Fallzahlen, sie weisen jedoch eine unpräzise Operationalisierung nicht-metrischer Variablen auf. Es sind aber gerade diese Variablen, die in der Institutionenanalyse von großer Bedeutung sind. Auch die Kombination von qualitativen und quantitativen Methoden in der QCA-Methode ist für die Messung von Institutionen nicht geeignet, da sie keine Richtlinien für die Bestimmung von Indikator- und Dimensionswerten bereitstellt und, noch problematischer, Messwerte als Mitglieder verschiedener Sets konzipiert, während wir für unser Analyseziel Werte auf einer kontinuierlichen Skala benötigen.

Wir haben die hier vorgestellte Methode entwickelt, um diese methodologischen Probleme zu reduzieren und einen systematischen Vergleich von Institutionen zu ermöglichen. Der Ansatz basiert auf der Quantifizierung qualitativer Daten und verwendet Webersche Idealtypen als Maßstab für den institutionellen Vergleich. Unser Ansatz vermeidet die oben beschriebenen Probleme etablierter Methoden durch die auf Idealtypen beruhende Herangehensweise und eine differenzierte und eindeutige Wertevergabe. Dadurch erlaubt unser Ansatz eine nuancierte Messung von metrischen sowie ordinalskalierten Daten. Die aggregierten Daten können für eine Vielzahl weiterer statistischer Analysen verwendet werden. Die von uns entwickelte Methode zur Messung von Institutionen schließt damit eine Lücke in den sozialwissenschaftlichen Methoden. Sie bietet sich zur Messung von Institutionen verschiedener Art an und kann zudem als ein methodischer Baustein in etablierten Analysen verwendet werden.

Literaturverzeichnis

- Alber, J. 1996: Selectivity, universalism, and the politics of welfare retrenchment in Germany and the United States, Paper for the 92nd Annual Meeting of the American Political Science Association, August 1996, San Francisco.
- Aspalter, C. 2011: The development of ideal-typical welfare regime theory. *International Social Work*, Vol. 54, No. 6, 735–750.
- Bennett, A., Elman, C. 2006: Qualitative research: Recent developments in case study methods. *Annual Review of Political Science*, No. 9, 455–476.
- Booyesen, F. 2002: An overview and evaluation of composite indices of development. *Social Indicators Research*, Vol. 59, 115–151.
- Campbell, J. 2011: Institutional reproduction and change. In G. Morgan, J. Campbell, C. Crouch, O.K. Pedersen and R. Whiteley (Hg.), *The Oxford handbook of comparative institutional analysis*. Oxford: Oxford Univ. Press, 87–115.
- Clasen, J., Clegg, D. 2007: Levels and levers of conditionality: measuring change within welfare states. In J. Clasen, N. Siegel (Hg.), *Investigating welfare state change. The 'dependent variable problem' in comparative analysis*. Cheltenham Northampton: Edward Elgar, 166–197.
- Ebbinghaus, B. 2011: Case selection in medium-N comparative welfare state analysis. *Qualitative & Multi-Method Research*, Vol. 9, No. 2, 15–20.
- Esping-Andersen, G. 1990: *The three worlds of welfare capitalism*. Cambridge: Polity Press.
- Fedderke, J. Garlick, J. 2011: Measuring institutions: Indicators of political and property rights in Malawi. *Social Indicators Research*, Vol. 106, 491–521.
- Ferragina, E., Seeleib-Kaiser, M. 2011: Welfare regime debate: past, present, futures? *Policy & Politics*, Vol. 39, No. 4, 583–611.
- Ferrarini, T., Nelson, K., Korpi, W., Palme, J. 2013: Social citizenship rights and social insurance replacement rate validity: pitfalls and possibilities. *Journal of European Public Policy*, Vol. 20, No. 9, 1251–1266.
- Figueredo, A., Olderbak, S., Schlomer, G., Garcia, R., Abril Wolf, P. 2013: Program evaluation: Principles, procedures, and practices. In T. D. Little (Hg.), *The Oxford handbook of quantitative methods in psychology*, Vol. 1. New York: Oxford University Press, 332–360.
- Flick, U. 2009: *An introduction to qualitative research*, Fourth Edition. London: SAGE Publications.
- Frericks, P. 2013: Strengthening market principles in welfare institutions: How hybrid pension systems impact on social-risk spreading. *Journal of Social Policy*, Vol. 42, No. 4, 665–683.
- Frericks, P., Höppner, J., Och, R. 2016: Institutional individualisation? The family in European social security institutions. *Journal of Social Policy*, Vol. 45, No. 4, 747–764.
- Green-Pedersen, C. 2007: More than data questions and methodological issues: Theoretical conceptualization and the dependent variable 'problem' in the study of welfare reform. In J. Clasen, N. Siegel (Hg.), *Investigating welfare state change. The 'Dependent Variable Problem' in comparative analysis*. Cheltenham Northampton: Edward Elgar Publishing Ltd., 13–23.
- Grünewald, A. 2014: Social security around the world: A review of datasets. *ZeS-Arbeitspapier 03/2014*, Universität Bremen.
- Hall, P. 1986: *Governing the economy. The politics of state intervention in Britain and France*. New York: Oxford University Press.
- Hall, P., Soskice, D. (Hg.) 2001: *Varieties of capitalism: The institutional foundations of comparative advantage*. Oxford, New York: Oxford University Press.

- Hall, P., Taylor, R. 1996: Political Science and the Three New Institutionalisms. *Political Studies*, Vol. 44, No. 5, 936–957.
- Javornik, J. 2012: Rethinking comparative childcare policy analysis: Example of Central and Eastern Europe. Working paper 10/2012. Department of Sociology, Umeå University.
- Kluge, S. 2000: Empirically grounded construction of types and typologies in qualitative social research. *Forum: Qualitative Social Research*, Vol. 1, No. 1.
- Kogut, B. 2010: Qualitative comparative analysis of social science data. In G. Morgan, J. L. Campbell, C. Crouch, O. K. Pedersen and R. Whiteley (Hg.), *The Oxford handbook of comparative institutional analysis*. Oxford: Oxford Univ. Press, 139–179.
- Kvist, J. 2007: Fuzzy set ideal type analysis. *Journal of Business Research*, Vol. 60, No. 5, 474–481.
- Lazarsfeld, P. 1937: Some remarks on the typological procedures in social research. *Zeitschrift für Sozialforschung*, Heft 6, 119–139.
- Mahoney, J. 2004: Comparative historical methodology. *Annual Review of Sociology*, Vol. 30, 81–101.
- North, D. 1991: Institutions. *Journal of Economic Perspectives*, Vol. 5, No. 1, 97–112.
- Pfau-Effinger, B., Flaquer, L., Jensen, P. (Hg.) 2009: Formal and informal work. The hidden work regime in Europe. New York: Routledge (Routledge Advances in Sociology, 46).
- Pierson, P. 1996: The new politics of the welfare state. *World Politics*, Vol. 48, No. 2, 143–179.
- Ragin, C. 2000: *Fuzzy-set social science*. Chicago/London: The University of Chicago Press.
- Schnell, R., Hill, P., Esser, E. 2005: *Methoden der empirischen Sozialforschung*. München: Oldenbourg Verlag.
- Schreier, M. 2012. *Qualitative content analysis in practice*. London/Thousand Oaks/New Delhi/Singapore: SAGE.
- Scruggs, L. 2007: Welfare state generosity across space and time. In J. Clasen, N. Siegel (Hg.), *Investigating welfare state change. The 'Dependent Variable Problem' in comparative analysis*. Cheltenham Northampton: Edward Elgar Publishing Ltd., 133–166.
- Streeck, W., Thelen, K. 2005: Introduction: Institutional change in advanced political economies. In W. Streeck, K. Thelen (Hg.), *Beyond continuity – Institutional change in advanced political economies*. Oxford: Oxford University Press, 3–39.
- Thelen, K., Steinmo, S. 1992: Historical institutionalism in comparative politics. In S. Steinmo, K. Thelen, F. Longstreth (Hg.), *Structuring politics. Historical institutionalism in comparative analysis*. Cambridge/New York: Cambridge University Press, 1–32.
- Wagemann, C. 2008: Qualitative Comparative Analysis und Policy-Forschung. In F. Jannig, K. Toens (Hg.), *Die Zukunft der Policy-Forschung: Theorien, Methoden, Anwendungen*. Wiesbaden: VS Verlag, 242–258.
- Weber, M. 1985: Die ‚Objektivität‘ sozialwissenschaftlicher und sozialpolitischer Erkenntnis. In M. Weber, J. Winckelmann (Hg.), *Gesammelte Aufsätze zur Wissenschaftslehre*. Tübingen: J.C.B. Mohr, 146–214.
- Weber, M., Winckelmann, J. (Hg.) 1985: *Gesammelte Aufsätze zur Wissenschaftslehre*. Tübingen: J.C.B. Mohr.