

Vielfalt von Lebensläufen im sozialen Wandel

Operationalisierung von Destandardisierung

Okka Zimmermann

Einleitung

Vielfalt wird in soziologischer Theorie oft im Kontext von sozialem Wandel thematisiert, wobei für historische oder aktuelle Epochen zunehmende oder abnehmende Vielfalt postuliert wird. In Bezug auf Lebensläufe wird von historischen Phasen der Standardisierung und Destandardisierung wichtiger Übergänge gesprochen.

In der Theorie wird postuliert, dass Lebensläufe seit Ende der sechziger bzw. Anfang der siebziger Jahre einen zunehmend destandardisierten Charakter annehmen (Kohli 1985; Mayer 1990); wobei die empirische Forschung inzwischen ergeben hat, dass diese Destandardisierung nur graduell zu sein scheint und Lebensläufe weiterhin im historischen Vergleich ein hohes Maß an Standardisierung aufweisen (Kohli 2007; Mayer 2009). Im letzten Jahrzehnt wurden dafür vermehrt detaillierte empirische Belege auf Basis holistischer Lebenslaufbetrachtung erbracht (Aisenbrey, Fasang 2010). Dabei wird Vielfalt von Lebensläufen meist mittels Entropie der Verteilung von Zuständen zu verschiedenen Zeitpunkten oder Lebensläufen auf typische Verlaufsmuster gemessen.

Andere, seltener verwendete Verfahren messen Vielfalt auf der Basis eines direkten Vergleichs von Lebensläufen, ihr Schwerpunkt liegt auf der Bestimmung der durchschnittlichen Unterschiedlichkeit von Lebensläufen. Hier sollen diese Verfahren vorgestellt, kritisch auf der Basis der Theorie zu möglichen Dimensionen von Destandardisierung (zum Beispiel Prävalenz, Dauer, Alter, Verlauf) evaluiert und beispielhaft angewendet werden. Es werden Verfahren vorgestellt, die auf (a) OM-Distanzen (Optimal Matching) sowie (b) dem Vergleich von Subsequenzen beruhen, die für die Dimensionen Alters- (a) und Verlaufsstandardisierung bzw. Standardisierung der Reihenfolge (b) Verwendung finden. Daneben werden für die Dimensionen Standardisierung von Prävalenz, Dauer und Komplexität einfache, selbst konstruierte Maße vorgeschlagen. Die Analysemethoden werden auf ausgewählte Daten zu Fertilität und Partnerschaften des *Generations and Gender Survey* (GGS) angewendet.

Bevor ich zu den eigentlichen Inhalten komme, möchte ich zunächst die Ausgangssituation (1.) und Definitionen und Dimensionen von Destandardisierung beschreiben (2.). Danach werden Ausführungen zu Methoden der Se-

quenzdatenanalyse und beispielhafte Analysen mit dem GGS folgen (3.). Schließlich fasse ich die wichtigsten Punkte und Ergebnisse nochmals zusammen (4.).

Ausgangssituation

Die empirische Lebenslaufforschung konzentriert sich bisher überwiegend auf die Analyse von Übergängen unter Verwendung von Methoden der Ereignisdatenanalyse. Damit wird die Gesamtgestalt der Lebensläufe vernachlässigt, die Sackmann und Wingens (2001) mit *Verlauf* bezeichnen.¹ Eine Möglichkeit, diese zu berücksichtigen, bietet die Sequenzdatenanalyse. Methodisch verschiebt sich in der Sequenzdatenanalyse der Fokus von Übergängen zu Zuständen. Übergänge und Zustände können gesehen werden als zwei verschiedene Konzepte, den Lebenslauf zu beschreiben. Die berücksichtigte Information bleibt jeweils gleich, da jeder Übergang in einer Zustandsveränderung abgebildet wird und jeder Zustand durch die vorhergegangenen Übergänge bedingt ist.

Innerhalb der Sequenzdatenanalyse konzentrieren sich Forscher bisher auf die Untersuchung von Lebenslaufgruppen. Diese können durch Vorabdefinition von idealtypischen Lebensläufen und der darauffolgenden Zuordnung der empirisch auftretenden Lebensläufe entstehen. Sie können auch aus einer Clusteranalyse hervorgehen. In beiden Verfahren werden Lebensläufe jeweils paarweise verglichen und die Unterschiedlichkeit zwischen ihnen bestimmt. Diese Herangehensweise bietet viele Vorteile, es kann jedoch auch Kritik formuliert werden. Diese beruht auf dem Fakt, dass Lebensläufe empirisch oft eher ein Kontinuum bilden (Halpin 2010; Anyadike-Danes, McVicar 2010). Die Zuordnung zu den Gruppen scheint eher künstlich und stark von subjektiven Entscheidungen geprägt, wie dies auch bei anderen Anwendungen von Clusteranalysen der Fall ist.

Daher wurde alternativ bereits die durchschnittliche Unterschiedlichkeit von Lebensläufen ermittelt, um Destandardisierung zu messen (Elzinga, Liefbroer 2007; Robette 2010; Fasang 2012). Die genannten Beispiele greifen allesamt auf den paarweisen Vergleich von Lebensläufen zurück, in dem kohortenweise alle Lebensläufe miteinander verglichen werden und das arithmetische Mittel der Ergebnisse dieser Vergleiche als Maß für den Grad der (De)Standardisierung

¹ In Anlehnung an die Verwendung der Sequenzdatenanalyse in englischsprachigen Forschungsarbeiten werde ich den gesamten Verlauf in einem festgelegten Lebenszeitraum als *Sequenz* bezeichnen. Dies scheint sinnvoll, um später die Dimension *Verlauf* davon unterscheiden zu können. Damit nutze ich eine von Sackmann und Wingens (2001) verschiedene Terminologie.

der Kohorten verwendet wird. Bei diesen Analysen ist meines Erachtens die Anbindung an die Theorie bisher eher schwach. Die ausgewählten sequenzdatenanalytischen Methoden werden nur in Ansätzen im Hinblick auf ihren Zusammenhang mit dem inzwischen vielschichtigen und mehrdeutigen Konzept der Destandardisierung diskutiert. Auch sind die vielen Möglichkeiten der deskriptiven Analyse bisher wenig kohortenvergleichend bzw. ergänzend genutzt worden. Daher möchte ich vorschlagen, unter Anbindung an die im Folgenden zu skizzierenden Dimensionen von Destandardisierung einfache Methoden der Sequenzdatenanalyse aus diesen beiden Bereichen systematisch auszuwählen und gemeinsam für eine mehrdimensionale Beschreibung des Wandels von Lebensläufen zu nutzen.

Destandardisierung

Die meisten Definitionen von Destandardisierung spiegeln die Komplexität und Vieldimensionalität dieses Konzeptes wieder. Oft wird Destandardisierung in Anlehnung an Brückner und Mayer (2005) wie folgt verstanden

»De-standardization would mean that life states, events and their sequences can become experiences which either characterize an increasingly smaller part of a population or occur at more dispersed ages and with more dispersed durations.« (Brückner, Mayer 2005: 32f.)

In diesem Konzept werden sowohl Zustände als auch Ereignisse sowie deren Abfolge erwähnt.² Dies ist in vielen anderen Darstellungen und Definitionen nicht der Fall, die sich stärker an Möglichkeiten empirischer Untersuchungen und daher am Übergangskonzept orientieren. Durch die Verschränkung der verschiedenen Aspekte ist die genannte Definition jedoch eine weniger brauchbare Ausgangsbasis für meine empirischen Analysen als Vorschläge, die eine Gliederung in Dimensionen vorschlagen. Dies geschieht meist in Vorbereitung einer empirischen Analyse (beispielsweise Konietzka 2010; Scherger 2007; Konietzka, Huinink 2003). Die dort gefundenen Dimensionen spiegeln die bisheri-

² Ereignisse sind dabei konzeptuell Übergängen ähnlich, jedoch nicht mit ihnen identisch. Das Übergangskonzept impliziert lebenslauftheoretisch meist eine Entwicklungsrichtung, beispielsweise beim Übergang ins Erwachsenenalter oder zur Elternschaft. Dabei wird der Ausdruck sowohl für einzelne Übergänge wie auch für Kombinationen genutzt. Das Konzept impliziert stärker als das des Ereignisses eine bewusst intendierte Handlung der Akteure. Ereignisse stoßen den Akteuren hingegen eher zu und sind außerdem weniger im Sinne einer Entwicklung gedacht. Demnach würde ein Unfall als Ereignis gewertet werden, eine Hochzeit als ein Übergang. Die Trennung der Konzepte ist jedoch nicht immer eindeutig möglich, wie beispielsweise bei einer Geburt oder einer Scheidung.

ge empirische Priorisierung des Übergangskonzeptes und werden daher von mir an die Perspektive der Sequenzdatenanalyse angepasst, also auf eine holistische Analyse der Lebensläufe unter Konzentration auf Zustände.

Ich unterscheide zwischen den *primären Dimensionen* Prävalenz, Dauer, Reihenfolge und Komplexität. Diese Dimensionen beziehen sich jeweils auf einen eindeutigen Aspekt und abstrahieren von einem relativ großen Anteil der Information. Die Indikatoren zur Untersuchung der Dimensionen gehören nicht zum klassischen Repertoire der Sequenzdatenanalyse. Aspekte jeweils mehrerer primärer Dimensionen werden in den sekundären Dimensionen Alters- und Verlaufsstandardisierung berücksichtigt, die mit klassischen Mitteln der Sequenzdatenanalyse (basierend auf paarweisem Vergleich von Lebensläufen) bestimmt werden. Damit wird ein größerer Anteil der in den Lebensläufen vorhandenen Information verwendet.

In der *ersten* Dimension *Prävalenz*_z wird das Auftreten von Elementen in Lebensläufen betrachtet. Lebensläufe können sich danach unterscheiden, ob ein Zustand auftritt oder nicht. Da der Fokus in dieser Untersuchung auf der ganzheitlichen Betrachtung des untersuchten Lebenslaufabschnittes liegt, sollen die *Kombinationen aller auftretenden Zustände* untersucht werden.

Die *zweite* Dimension betrachtet die Dauer der Episoden. Dies entspricht im gewissen Sinne der Kopplung bzw. Entkopplung von Ereignissen, da die Dauer einer Episode durch die einrahmenden Ereignisse bestimmt wird; eine Episode ist definiert als eine Reihe gleicher Zustände zwischen zwei Übergängen oder Ereignissen. In der hier angestrebten holistischen Perspektive auf ganze Lebenslaufabschnitte von Kohorten werden alle Lebenslaufabschnitte aller Befragten einer Kohorte als gleichberechtigt betrachtet.

Die durchschnittliche Länge der Episoden nimmt ab, da die Anzahl der Episoden zunimmt. Die Anzahl der Episoden ist ein einfaches Maß für die Komplexität von Lebensläufen³, die als *dritte* Dimension berücksichtigt werden soll. Die Anzahl der Episoden eines Lebenslaufabschnittes entspricht der Anzahl der Ereignisse plus eins, wobei gleichzeitig stattfindende Ereignisse als ein Ereignis gezählt werden.

Die *vierte*, etwas weniger abstrakte Dimension Reihenfolge wird durch die Prävalenz beeinflusst. Gleiche Reihenfolgen können nur entstehen, wenn gleiche Elemente im Lebenslauf auftreten. Zusätzlich wird jedoch die temporale Ordnung dieser Elemente, also ein Aspekt der zeitlichen Struktur in die Betrachtung einbezogen (ohne Berücksichtigung der Dauer der Episoden).

Die *sekundären Dimensionen* Alters- und Verlaufsstandardisierung beziehen jeweils Informationen mehrerer primärer Dimensionen ein und spiegeln damit

³ Elzinga und Liefbroer (2007) berücksichtigen für ihre Maßzahl der Komplexität zusätzlich die Varianz in der Länge der Episoden. Da diese hier bereits als Aspekt der Dauer einzeln dargestellt wird, soll sie in Bezug auf die Komplexität nicht berücksichtigt werden.

mehrere Aspekte des Destandardisierungskonzeptes. In der Dimension Alter wird untersucht, wie ähnlich oder unterschiedlich die Lebensläufe von Kohortenmitgliedern zu bestimmten Alterszeitpunkten, also beispielsweise im dritten Quartal des 15. Lebensjahres sind. In der Dimension Verlauf wird untersucht, wie ähnlich sich Lebensläufe in Bezug auf die Reihenfolge der auftretenden Zustände sind, wobei die Dauer der Episoden berücksichtigt wird. Beide sekundäre Dimensionen setzen sich mehr oder weniger direkt aus Aspekten der primären Dimensionen zusammen; Veränderungen in den primären Dimensionen beeinflussen daher die sekundären Dimensionen.

Methoden der Sequenzdatenanalyse und beispielhafte Analysen

Die in der Soziologie als erste und am meisten verwendete Methode fortgeschrittener Sequenzdatenanalyse ist die Optimal Matching Analyse (OMA, zum Beispiel Abbott 1990, 1995; Abbott, Tsay 2000). Sie wird von Abbott in Zusammenhang mit der Entwicklung seines *narrativem Positivismus* gebracht; einem neuen theoretischen Paradigma, in dem er die Beschreibung von sozialen Phänomenen (zunächst) als vorrangig gegenüber der Erklärung derselben sieht (Abbott 1992). Ursprünglich stammt die OMA aus der Informatik und fand dort vor allen Dingen in der Bioinformatik breite Anwendung bei dem Vergleich und der Gruppierung von DNA-Sequenzen (Lesnard 2006, 2008). Mit dieser Methode wird die Ähnlichkeit von Sequenzen durch die Anzahl der Operationen bestimmt, die zur Überführung eines Lebenslaufes in einen anderen benötigt werden. Dabei können Einfügungen, Löschungen oder Substitutionen von Elementen vorgenommen werden. Einfügungen können theoretisch vernachlässigt werden, da eine Einfügung in einer Sequenz einer Löschung in der jeweils verglichenen Sequenz entspricht. Die Operationen werden mit verschiedenen Kosten versehen. Diese drücken aus, für wie relevant ein Unterschied zwischen zwei Sequenzen gehalten wird. Das Verhältnis der Kosten zueinander ist besonders bedeutsam und es werden in der Regel verschiedene Kombinationen von Kosten ausprobiert, um sinnvolle Gruppierungen von Lebensläufen zu erhalten (Lesnard 2006, 2008).

In den Sozialwissenschaften wurde diese Methode übernommen, um Lebensläufe zu vergleichen. Diese werden als eine Folge von Zuständen dargestellt. Die Methode wird schwerpunktmäßig eingesetzt, um Lebensläufe zu gruppieren und anschließend besser beschreiben zu können (Anyadike-Danes, McVicar 2010). Ähnlich wie in Anwendungen der Bioinformatik werden meist verschiedene Kostenkombinationen ausprobiert und anhand der Ergebnisse der Clusteranalyse evaluiert. Bei neueren Anwendungen dieser Ähnlichkeitsmaße

zur Beschreibung von Destandardisierung im Kohortenvergleich (wie sie hier angestrebt wird) ist jedoch eine theoretische Begründung der Kostenstruktur notwendig, die auf der Bedeutung im Bezug auf die Unterschiedlichkeit von Lebensläufen beruht und eine Interpretierbarkeit der Ergebnisse ermöglicht. Da dieses Unterfangen sich als äußerst schwierig erweist (Wu 2000; Levine 2000), sollen beide Operationen des Optimal Matching (Substitutionen, Löschungen) getrennt voneinander betrachtet werden.

Werden nur Substitutionen vorgenommen, wird die Stellung der Zustände in der Sequenz gegenüber ihrer Reihenfolge priorisiert (Lesnard 2006, 2008). Damit können diese Verfahren der Dimension Alter zugerechnet werden. Das einfachste Verfahren ist die sogenannte Hamming-Distanz (Hamming 1950; 1980; Bookstein et al. 2002; Kruskal 1983), die für jeweils zwei Lebensläufe die Anzahl der Positionen zählt, in denen sich Lebensläufe unterscheiden. In Bezug auf zeitlich parallel verlaufene Lebensläufe kann die Hamming-Distanz als ein Maß der kalendarischen Gleichzeitigkeit verstanden werden – in Bezug auf Lebenslaufabschnitte, die anhand des chronologischen Alters bestimmt werden, dagegen als ein Maß der altersbezogenen Parallelität der Lebensläufe.

Die Hamming-Distanz kann verfeinert werden, indem bei nicht übereinstimmenden Zuständen der Grad der Unterschiedlichkeit berücksichtigt wird. Dies kann zum Beispiel durch die City Block Distanz geschehen, die für metrische Zustände die Differenz als Maß der Unterschiedlichkeit verwendet. Da Zustände in den Sozialwissenschaften jedoch selten metrisch sind, findet die City Block Distanz eben so wenig Anwendung und wird auch hier vernachlässigt. Der Grad der Unterschiedlichkeit zwischen zwei Zuständen kann weiterhin theoretisch festgelegt werden. Dies würde der Überführung des nominalen oder ordinalen Meßniveaus in ein metrisches implizieren und wird daher aus methodischen Gründen abgelehnt.

Deswegen wurde vorgeschlagen, die Unterschiedlichkeit auf Basis von Übergangsraten zu bestimmen (Rohwer, Pötter 1999; angewendet in Widmer, Ritschard 2009; Anyadike-Danes, McVicar 2010). Hier stellt sich zunächst das Problem der Symmetrie. Substitutionskosten müssen symmetrisch sein, damit die Überführung des Lebenslaufes A in B die gleichen Kosten ergibt wie die Überführung des Lebenslaufes B in A und damit die jeweils bestimmten Kosten unabhängig von der Vergleichsrichtung sind. Daher werden in der Praxis die Übergangsraten zwischen zwei Zuständen in beide Richtungen ermittelt und schließlich gemittelt. Es regt sich jedoch auch gegen diesen Ansatz (meines Erachtens berechtigte) Kritik, da Übergänge als Aspekte der temporalen Ordnung verstanden werden und Substitutionen als abstrakte Operation auf eine abstrahierte Darstellung von Lebensläufen zu verstehen sind (Halpin 2010). Es lassen sich zudem leicht Beispiele finden, in denen diese Übertragung irreführend erscheint. So wird der Übergang von der Ausbildung in die Berufstätigkeit

von nahezu allen Mitgliedern westlicher Gesellschaften durchlaufen; es ist daher zu erwarten, dass die Übergangsrate hoch ist. Dennoch sind die Zustände *in Ausbildung* und *berufstätig* als sehr unterschiedlich zu betrachten.

Die aufgeführten Argumente gelten in ähnlicher Form auch für Verfahren, bei denen die Übergangsraten zeitpunktspezifisch bestimmt werden, wie das Dynamic Hamming (Lesnard 2006, 2008). Diese Konzepte bestimmen Substitutionskosten anhand der Anzahl der Übergänge, die im Bereich eines zu substituierenden Zustandes stattfinden. Damit wird abgebildet, dass in Übergangsphasen zwischen zwei Zuständen die Unterschiedlichkeit zwischen diesen als weniger bedeutsam empfunden wird. Obwohl der Ansatz der Übergangsphasen recht einleuchtend und sinnvoll erscheint; bleiben oben genannten Zweifel an übergangsbasierten Substitutionskosten meines Erachtens bestehen. Aufgrund dieser Zweifel an den Verfahren mit variierenden Substitutionskosten werde ich nur die Hamming-Distanz zur Messung der Standardisierung der Dimension Alter nutzen.

Werden nur Löschungen zur Überführung der Lebensläufe ineinander verwendet, dann werden vorhandene gleiche Zustandsfolgen, die zu unterschiedlichen Zeitpunkten in zwei Lebensläufen zu finden sind, in Übereinstimmung gebracht. Unter Fokussierung auf die Reihenfolge werden dabei die Zeitpunkte von Zuständen verändert; die Zeit wird durch die Verschiebung der Zustände gestaucht (*time warping*, Lesnard 2006, 2008). Elzinga (2003, 2005, 2009; auch Elzinga, Liefboer 2007) verwendet den Begriff der Subsequenzen für eine Sequenz, die durch Streichung beliebiger Teile aus einer anderen Sequenz hervorgeht. Diese Betrachtung ermöglicht die Entwicklung einer Reihe von Maßzahlen zur Messung von Verlaufsstandardisierung.⁴ Die wichtigsten Maßzahlen aus einer holistischen Perspektive auf einen Lebenslaufabschnitt sind die längste gemeinsame Subsequenz und der Anteil der gemeinsamen Subsequenzen. Erstere fokussiert auf die temporale Ordnung des gesamten Lebenslaufes und wurde als Ergebnis von Elementlöschungen bereits in anderen Kontexten verwendet; letztere berücksichtigt stärker Ähnlichkeiten in Substrukturen bzw. Lebenslaufabschnitten, indem auch kleinere Strukturen aus unterschiedlichen Lebenslaufabschnitten miteinander verglichen werden. Beide Maßzahlen können ohne Dauer, wodurch sich ein Maß für die Dimension Reihenfolge ergeben würde, und mit Dauer, wodurch sich ein Maß für die Dimension Verlauf ergeben würde, berechnet werden.

Die Kosten für Löschungen werden meist konstant gehalten; Vorschläge mit variierenden Kosten für Löschungen wurden bisher nur selten verwendet und wenig diskutiert und werden daher hier nur kurz thematisiert; ähnlich wie die Bedeutung von Indelkosten insgesamt theoretisch vernachlässigt wird (Hol-

⁴ Vgl. zusammenfassend zu den Möglichkeiten der Analyse und Variationen der Maßzahlen Elzinga 2009.

lister 2009: 244). Es sind drei Varianten vorstellbar: Die Kosten können variiert werden in Abhängigkeit von (a) den Zuständen, (b) der Länge einer Episode (Halpin 2010; Rohwer, Pötter 1999) und (c) den umgebenden Zuständen (*localized index*, Hollister 2009: 247). Es treten bei diesen Varianten ähnliche Schwierigkeiten bei der theoretischen Begründung der variierenden Kosten auf, wie es bei Substitutionen der Fall ist; daher werden diese Varianten hier nicht weiter verfolgt.

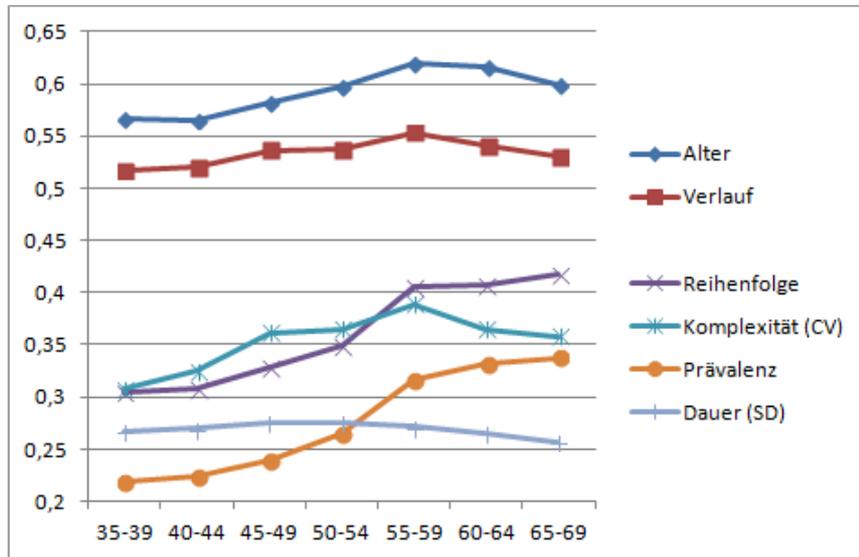
Ausgewählte Analysemethoden und beispielhafte Analysen mit dem GGS

In den nachfolgenden beispielhaften Analysen werden die ausgewählten Methoden für französische Frauen, die zwischen 1935 und 1969 geboren sind, angewandt. Insgesamt gingen die Daten von 3.368 Befragten aus dem GGS in die Analyse ein⁵; bei quartalsweiser Zustandsdefinition wurde der Altersabschnitt 15–35 Jahre untersucht. Es werden die Dimensionen Fertilität und Partnerschaft betrachtet, die Zweidimensionalität wurde in der Zustandsdefinition berücksichtigt⁶. Für jede Dimension wurden drei Zustände definiert, in der Kombination ergeben sich neun Zustände. In der Dimension der Fertilität wurde unterschieden, ob der Befragte mit einem, keinem bzw. zwei oder mehr Kindern in einem Haushalt lebt. Da es selten vorkommt, dass die Befragten vor dem 35. Lebensjahr mehr als zwei Kinder bekommen, wurde die Kinderzahl nicht weiter differenziert. In der Dimension der Partnerschaft wurde zwischen Alleinleben, nichtehelicher Lebensgemeinschaft und ehelicher Lebensgemeinschaft unterschieden. *Living-Apart-Together*-Beziehungen wurden nicht berücksichtigt, da diese schwierig zu definieren und daher ebenso schwierig eindeutig zu erfassen sind. Insgesamt liegt der Fokus damit auf Beziehungen in der selbst gegründeten Kernfamilie im gemeinsamen Haushalt, andere Beziehungen (zum Beispiel zu eigenen Eltern, Stiefkindern oder eigenen Kindern außerhalb des Haushaltes) wurden nicht berücksichtigt. Diese Beschränkung wurde eingeführt, da Beziehungen mit gemeinsamer Haushaltsführung als am intensivsten und wichtigsten angesehen werden.

5 Einige Befragte mussten ausgeschlossen werden, da ihre retrospektiven Angaben an entscheidender Stelle unvollständig oder widersprüchlich waren.

6 Die wichtigste alternative Vorgehensweise wäre die getrennte Berechnung und anschließende lineare Kombination der Unterschiedlichkeit in den beiden Dimensionen Fertilität und Partnerschaft.

Abb. 1: Entwicklung in den Dimensionen von Standardisierung (normierte, kohortenweise berechnete Koeffizienten).



Quelle: GGS, 1. Welle (2004), eigene Berechnungen, Geburtskohorten 1935–69, Altersabschnitt 15–35 Jahre, französische Frauen, Dimensionen Fertilität und Partnerschaft.

Für die Untersuchung der Prävalenz durch paarweisen Vergleich ist keine Maßzahl bekannt. Daher wird eine neue einfache Maßzahl eingeführt. In paarweisem Vergleich wird dabei ermittelt, in Bezug auf das Auftreten von wie vielen Zuständen sich die Lebensläufe unterscheiden. Das theoretische Maximum von neun Unterschieden (welches zur Normierung verwendet wird) wäre erreicht, wenn in einem Lebenslauf beispielsweise die Zustände 1, 2, 3, 4 und 5 auftreten und in einem anderen die Zustände 6, 7, 8 und 9. Ein weiteres Beispiel: Die Lebensläufe A mit den Elementen 1, 2, 5, 6 und B mit den Elementen 1, 4, 5, 8, 9 unterscheiden sich in Bezug auf das Auftreten von sechs Elementen (2, 3, 4, 6, 8, 9); die Unterschiedlichkeit beträgt also $6/9$.

Abbildung 1 zeigt in der Dimension Prävalenz kontinuierliche Destandardisierung; die Entwicklung beschleunigte sich bis zur Kohorte 1955–59 und verlor in den jüngeren Kohorten wieder an Stärke. Demnach stieg die Variabilität der Lebenslaufzustände, die die Befragten zwischen 15 und 35 Jahren erleben, besonders deutlich bis zur Kohorte 1955–59, danach waren schwächere Zunahmen zu verzeichnen.

Eine ähnliche Entwicklung zeigt sich in der Dimension der Reihenfolge, in der im paarweisen Vergleich die längste gemeinsame Subsequenz zwischen den Lebensläufen jeder Kohorte ermittelt und daraus abschließend die durchschnitt-

liche Unterschiedlichkeit bestimmt wird. Dabei bleibt die Dauer der Episoden unberücksichtigt; jede Episode wird beim Vergleich nur durch ein Symbol repräsentiert. Die Destandardisierung ist hier ebenso wie bei der Prävalenz in der Kohorte 1955–59 besonders deutlich und schwächt sich hinterher ab.

Tab. 1: Übersicht betrachteter Dimensionen

Dimensionen Destandardisierung		Betrachtete Unterschiedlichkeit
Primär	Prävalenz	Auftretende Zustände
	Dauer	Dauer der Episoden
	Komplexität	Anzahl der Episoden
	Reihenfolge	Abfolge der Zustände
Sekundär	Alter	Anteil der ungleichen Zustände zu Alterszeitpunkten
	Verlauf	Längste gemeinsame Subsequenz

Bei der Dimension Dauer wird die Streuung der durchschnittlichen Episodendauer der Lebensläufe der Kohorten ermittelt, gemessen durch die Standardabweichung⁷. Diese vergrößert sich zwischen den älteren Kohorten und verringert sich anschließend in den drei jüngsten Kohorten wieder. Damit findet in den jüngeren Kohorten *Restandardisierung* statt, ein Begriff mit dem in jüngerer Zeit die Umkehrung von Destandardisierungsprozessen beschrieben wird. Die Restandardisierung in den jüngeren Kohorten fällt deutlicher aus als die Destandardisierung in den älteren Kohorten, die ebenso wie in der Prävalenz zwischen den jüngeren und älteren Kohorten etwas geringer ausfällt als zwischen den mittleren.

Wie oben beschrieben, hängt die Dauer der Episoden mit der Anzahl der Episoden zusammen. Daher ist die dritte Dimension (Streuung der Komplexität) in gewisser Weise komplementär zur Dauer. Als einfaches Komplexitätsmaß wird die Anzahl der Episoden verwendet. Um die sich verändernde durchschnittliche Anzahl der Episoden zu berücksichtigen, wird als Streuungsmaß der Varianzkoeffizient (Standardabweichung durch arithmetisches Mittel) verwendet (Hartung et al. 2005: 47). Dieser zeigt prinzipiell Destandardisierung bis zur Kohorte 1955–59 und im Folgenden Restandardisierung.

Die beiden sekundären Dimensionen Alter und Verlauf zeigen eine einander ähnliche Entwicklung wie die Komplexität mit einem Wendepunkt bei der Ko-

⁷ Dabei wurde die erste und letzte Episode, die an jeweils einer Seite durch die Grenzen des untersuchten Lebenslaufzeitraums bestimmt werden, mit einbezogen. Zur Kontrolle des Einflusses der künstlich beschnittenen Episoden wurde die Maßzahl alternativ ohne die erste und letzte Episode berechnet, wobei sich sehr ähnliche Ergebnisse ergaben.

horte 1955–59 und nachfolgender Restandardisierung. Diese beiden Dimensionen fassen Aspekte der vorhergegangenen Dimensionen zusammen und sind daher die aussagekräftigsten Maßzahlen für die Gesamtentwicklung. In der Dimension Alter sinkt dabei zunächst der Anteil der gleichen Zustände, die Befragte in einem bestimmten Alter erleben, um später wieder anzusteigen. Gleichzeitig verkürzt sich zunächst die durchschnittliche längste gemeinsame Subsequenz zwischen den Befragten, also der Zeitanteil der Lebensläufe, die sich gleich entwickeln bzw. aus gleichen Zustandsfolgen bestehen. In den jüngeren Kohorten nimmt auch dieser wieder zu. Der hier nicht abgebildete Anteil der gemeinsamen Subsequenzen entwickelte sich sehr ähnlich. Die sekundären Dimensionen scheinen die Kombination der anderen Dimensionen zu reflektieren, in denen die Kohorte 1955–59 ebenfalls meist einen Wendepunkt markiert.

Die Ähnlichkeit der Entwicklungen könnte darauf hindeuten, dass die Kombination von zwei Indeloperationen wie eine Substitution verwendet wird (*pseudo-substitutions*, Hollister 2009: 258). Damit würde die Unähnlichkeit zwischen den Lebensläufen auch in dieser Dimension hauptsächlich durch die altersbezogene Parallelität von Lebensläufen bestimmt. Indeloperationen, die nicht kombiniert als Pseudo-Substitutionen verwendet werden, scheinen dazu beizutragen, dass die Kosten für die Überführung etwas niedriger sind. Es scheint jedoch zwischen den beiden Dimensionen Alter und Verlauf keine grundsätzlichen Unterschiede in der Entwicklung zu geben, weswegen gegebenenfalls zu überdenken wäre, ob die parallele Betrachtung durch eine andere Kombination betrachteter Dimensionen ersetzt werden kann.

Zusammenfassung

Durch die Verwendung von Sequenzdatenanalyse ist durch die Betrachtung von Zustandsreihen ein holistischer Zugang zu Lebensläufen möglich. Der Fokus dieser Darstellung lag auf abstrahierten Eigenschaften von Kohorten, statt (wie in vielen anderen Untersuchungen) auf Gruppen von Lebensläufen. Destandardisierung wurde theoretisch als multidimensionales Konzept gefasst und einfache Maßzahlen der Sequenzdatenanalyse zur Operationalisierung der Dimensionen verwendet.

In Anlehnung an mehrdimensionale Konzepte von Destandardisierung wurden systematisch einfache Methoden der Sequenzdatenanalyse ausgewählt und passende Betrachtungsweisen von einzelnen Sequenzeigenschaften ergänzt. Es wurden beispielhafte Analysen für den Lebensverlauf zwischen 15 und 35 Jahren von zwischen 1935 und 1969 geborenen Französisinnen durchgeführt.

In den sekundären, mehrere Aspekte reflektierenden Dimensionen Alter (Hamming-Distanz) und Verlauf (längste gemeinsame Subsequenz) zeigte sich eine einheitliche Entwicklung: Nach einer Phase der Destandardisierung folgte eine Phase der Restandardisierung nach der Kohorte 1955–59. Es erscheint daher plausibel, dass diese Entwicklung eine Reflexion der Entwicklungen in den primären Dimensionen ist, vielleicht sogar als eine Kombination dieser Effekte verstanden werden kann: Auch in allen diesen primären Dimensionen zeigte sich nach der Kohorte 1955–59 entweder Restandardisierung (Dimensionen Dauer, Komplexität) oder abnehmende Stärke von Destandardisierung (Dimensionen Prävalenz, Reihenfolge). Besonders stark fiel die Restandardisierung in Bezug auf die durchschnittliche Episodendauer aus; es scheint daher plausibel, dass die Restandardisierung in den abstrakten Dimensionen Alter und Verlauf zu einem großen Teil durch eine abnehmende Unterschiedlichkeit in der Dauer der Episoden verursacht wird. Es wären weitere Untersuchungen nötig, um die angenommenen Einflüsse zu verifizieren.

Die gewählte Herangehensweise bietet Vorteile über die bisherigen Anwendungen der Sequenzdatenanalyse zur Messung von Destandardisierung. Neben der stärkeren Anbindung an die Theorie wird vermieden, dass bei der Methodenwahl eine unnötige (frühe) Festlegung auf eine Methode zur Bestimmung der Unterschiedlichkeit von Sequenzen erfolgen muss. Es können gleichartige und auch widersprüchliche Entwicklungen sowie mögliche Wechselwirkungen identifiziert werden. Bei dem Vergleich verschiedener Länder oder auch Geschlechter könnten detaillierte Unterschiede in den Dimensionen ermittelt und interpretiert werden.

Es konnte bestätigt werden, was schon für auf Lebenslaufgruppierung ausgerichtete Anwendungen der Sequenzdatenanalyse gezeigt haben⁸: Variationen des Optimal Matchings (hier zwei stark vereinfachte Versionen: nur Substitutionen bzw. nur Löschungen) führen zu vergleichbaren Ergebnissen. Dies könnte darauf verweisen, dass in der Dimension Verlauf Einfügungen und Löschungen als *pseudo-substitutions* (Hollister 2009: 258; vgl. oben) verwendet werden; damit würde die Unterscheidung zwischen diesen Dimensionen empirisch weniger wichtig. Die Ergänzung durch andere Maßzahlen, die spezifische Aspekte von Lebensläufen vergleichen (Prävalenz, Reihenfolge, Komplexität, Dauer) ist daher eine umso wichtigere Ergänzung, um Prozesse der Destandardisierung zu beschreiben und zu verstehen.

⁸ Siehe Robette, Bry 2012; Hollister 2009; Robette, Thibaut 2008; Aisenbrey, Fasang 2010; Abbott, Hrycak 1990.

Literatur

- Abbott, A. 1990: ›Conceptions of Time and Events in Social Science Methods. *Historical Methods and Research*, 23. Jg., 140–150.
- Abbott, A. 1992: From Causes to Events – Notes on Narrative Positivism. *Sociological Methods & Research*, 20. Jg., 428–455.
- Abbott, A. 1995: Sequence Analysis – New Methods for Old Ideas. *Annual Review of Sociology*, 21. Jg., 93–113.
- Abbott, A., Hrycak, A. 1990: Measuring Resemblance in Sequence Data: An Optimal Matching Analysis of Musicians' Careers. *American Journal of Sociology*, 96. Jg., 144–185.
- Abbott, A., Tsay, A. 2000: Sequence Analysis and Optimal Matching. *Sociological Methods & Research*, 29. Jg., 3–33.
- Aisenbrey, S., Fasang, A.E. 2010: New life for old ideas: The ›Second Wave‹ of Sequence Analysis. Bringing the ›Course‹ Back into the Life Course. *Sociological Methods & Research*, 38. Jg., 420–462.
- Anyadike-Danes, M., McVicar, D. 2010: My Brilliant Career – Characterizing the Early Labor Market Trajectories of British Women From Generation X. *Sociological Methods & Research*, 38. Jg., 482–512.
- Bookstein, A., Kulyukin, V.A., Raita, T. 2002: Generalized Hamming Distance. *Information Retrieval*, 5. Jg., 353–375.
- Brückner, H., Mayer, K.U. 2005: The De-Standardization of the Life Course: What It Might Mean and If it Means Anything Whether It Actually Took Place. *Advances in Life Course Research*, 9. Jg., 27–54.
- Elzinga, C.H. 2003: Sequence Similarity – A Non-Aligning Technique. *Sociological Methods & Research*, 31. Jg., 3–29.
- Elzinga, C.H. 2005: Combinatorial Representations of Token Sequences. *Journal of Classification*, 22. Jg., 87–118.
- Elzinga, C.H. 2009: CHESA 3.1 User Manual. Vrije Universiteit Amsterdam, Faculty of Social Sciences, Dept. of Social Science Research methods, <http://home.fsw.vu.nl/ch.elzinga/CHESA%203.1%20pack.zip> (letzter Aufruf 14.01.2013).
- Elzinga, C.H., Liefbroer, A.C. 2007: De-Standardization of Family-Life Trajectories of Young Adults: A Cross-national Comparison Using Sequence Analysis. *European Journal of Population*, 23. Jg., 225–250.
- Fasang, A. 2012: Institutional Change and Family Formation. The Reunification of East and West Germany in 1989. CIQLE Working Paper 2012-01. New Haven, CT: Yale University, Center for Research on Inequalities and the Life Course.
- Halpin, B. 2010: Optimal Matching Analysis and Life-Course Data: The Importance of Duration. *Sociological Methods & Research*, 38. Jg., 365–388.
- Hamming, R.W. 1950: Error-detecting and error-correcting codes. *Bell System Technical Journal*, 29. Jg., 147–160.
- Hamming, R.W. 1980: *Coding and Information Theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Hartung, J., Elpelt, B., Klösener, K.-H. 2005: *Statistik: Lehr- und Handbuch der angewandten Statistik*. München: Oldenbourg.
- Hollister, M. 2009: Is Optimal Matching Suboptimal? *Sociological Methods & Research*, 38. Jg., 235–264.

- Kohli, M. 2007: The institutionalization of the life course: Looking back to look ahead. *Research in Human Development*, 4. Jg., 253–271.
- Konietzka, D. 2010: *Zeiten des Übergangs. Sozialer Wandel des Übergangs in das Erwachsenenalter*. Wiesbaden: VS.
- Konietzka, D., Huinink, J. 2003: Die De-Standardisierung einer Statuspassage? Zum Wandel des Auszugs aus dem Elternhaus und des Übergangs in das Erwachsenenalter in Westdeutschland. *Soziale Welt*, 54. Jg., 285–311.
- Kruskal, J.B. 1983: An overview of sequence comparison. In D. Sankoff, J.B. Kruskal, *Time warps, string edits, macromolecules: the theory, practice of sequence comparison*. Addison: Wesley, 1–44.
- Lesnard, L. 2006: Optimal Matching, *Social Sciences. Série des Documents de Travail du CREST*, 2006-1, 1–25.
- Lesnard, L. 2008: Off-scheduling within dual-earner couples: an unequal, negative externality for family time. *American Journal of Sociology*, 114. Jg., 447–490.
- Levine, J.H. 2000: But What Have You Done for Us Lately? Commentary on Abbott and Tsay. *Sociological Methods & Research*, 29. Jg., 34–40.
- Mayer, K.U. 2009: New Directions in Life Course Research. *Annual Review of Sociology*, 35. Jg., 413–433.
- Robette, N. 2010: The diversity of pathways to adulthood in France: evidence from a holistic approach. *Advances in Life Course Research*, 15. Jg., 89–96.
- Robette, N., Bry, X. 2012: Harpoon or Bait? A Comparison of Various Metrics in Fishing for Sequence Patterns. *Bulletin of Sociological Methodology*, 30. Jg., 5–24.
- Robette, N., Thibault, N. 2008: Comparing Qualitative Harmonic Analysis and Optimal Matching. An Exploratory Study of Occupational Trajectories. *Population-E*, 63. Jg., 533–556.
- Rohwer, G., Pötter, U. 1999: *TDA User's Manual*. Bochum: Manuskript Ruhr-Universität.
- Sackmann, R., Wingens, M. 2001: *Strukturen des Lebenslaufs. Übergang – Sequenz – Verlauf*. Weinheim: Juventa.
- Scherger, S. 2007: *Destandardisierung, Differenzierung, Individualisierung. Westdeutsche Lebensläufe im Wandel*. Wiesbaden: VS.
- Widmer, E.D., Ritschard, G. 2009: The de-standardization of the life course: Are men and women equal? *Advances in Life Course Research*, 14. Jg., 28–39.
- Wu, L.L. 2000: Some comments on sequence analysis, optimal matching methods in sociology: Review, prospect. *Sociological Methods & Research*, 29. Jg., 41–64.