

# Methoden der Längsschnittanalyse in den Wirtschafts- und Sozialwissenschaften

Philipp G. Sandner<sup>1</sup>, Andreas Größler<sup>2</sup>

In den Sozialwissenschaften werden Längsschnittstudien häufig angewendet, um sozialen und ökonomischen Wandel zu beschreiben und zu erklären. Das Potenzial von Längsschnittuntersuchungen liegt in der Analyse von Veränderungen, Entwicklungen, zeitlichen Prozessen oder Wandlungsvorgängen, wofür Daten für einen Zeitraum vorliegen müssen. Anwendungsmöglichkeiten sind auch in den Wirtschaftswissenschaften und der betriebswirtschaftlichen Forschung gegeben, allerdings wird diese Forschungsmethodik hier seltener angewendet. Dennoch besitzen Längsschnittanalysen ein außerordentliches Potenzial, weil sie anders als Querschnittsstudien wichtige Erkenntnisse über Dynamik, Veränderungen und Wandlungsprozesse liefern können und auch Aussagen über Kausalitäten zulassen (vgl. *Dierkes*, 1977, S. 113, *Baltes/Nesselrode*, 1979, S. 23–27, *Hakim*, 2000, S. 124 f.).

In diesem Beitrag soll der grundlegende Ansatz von Längsschnittuntersuchungen dargestellt werden. Die beiden wesentlichen Bestandteile einer Längsschnittstudie sind das **Design** zur **Erhebung** der Längsschnittdaten (Abschn. 2) und die **Methoden** zur **Auswertung** von Längsschnittdaten (Abschn. 3). Der bei der Erhebung gewonnene **Datenbestand** dient als Bindeglied zwischen diesen beiden Komponenten. Ziel dieser Arbeit ist, eine systematische Darstellung der vielfältigen Möglichkeiten, eine Längsschnittanalyse durchzuführen, vorzulegen.

## 1. Der Einsatz von Längsschnittstudien zur Analyse von Veränderungen

Querschnittsuntersuchungen „fotografieren“ die reale Wirklichkeit zu einem bestimmten Zeitpunkt. Damit können Zusammenhänge bereits beschrieben und aufgedeckt werden. Allerdings erlauben Längsschnittuntersuchungen durch Berücksichtigung der

---

<sup>1</sup> Institut für Innovationsforschung, Technologiemanagement und Entrepreneurship (INNO-tec), Ludwig-Maximilians-Universität München, Kaulbachstraße 45, D-80539 München. Telefon: +49-89-2180 5809, E-Mail: psandner@bwl.uni-muenchen.de.

<sup>2</sup> Nijmegen School of Management (NSM), Thomas van Aquinostraat 3, NL-6525 GD Nijmegen, E-Mail: a.grossler@fm.ru.nl.

Zeitkomponente weitergehende Zielsetzungen: Sie ermöglichen es, Zustände und Verhalten im Zeitablauf zu untersuchen und Wandlungsvorgänge zu erforschen. Erst wenn Daten über ein oder mehrere Untersuchungsobjekte für mindestens zwei Zeitpunkte vorliegen, kann von Längsschnittuntersuchungen gesprochen werden (vgl. *Coleman*, 1981, S. 7, *Menard*, 1991, S. 5). *Baltes/Nesselroade* (1979, S. 23–27) nennen als **grundlegende Einsatzmöglichkeiten** von Längsschnittstudien die Beobachtung bzw. Begründung von individuellem Wandel, die Analyse von Zusammenhängen in Verhaltensänderungen und schließlich den Vergleich verschiedener individueller Veränderungsprozesse zur Beobachtung bzw. Begründung von Unterschieden und Ähnlichkeiten.

Aufgrund des Einbezugs der Zeitkomponente eignen sich Längsschnittdesigns in hohem Maß zur **empirischen Überprüfung kausaler Hypothesen**. Längsschnittuntersuchungen vermögen auch diejenigen kausale Beziehungen zu analysieren, die – wenn Ursache und Wirkung zeitlich auseinander fallen – inhärente **Verzögerungen** aufweisen.

Werden über einen Zeitraum mehrere Veränderungen in Folge betrachtet, können **Sequenzen** Gegenstand einer Untersuchung sein, was die Analyse und den Vergleich von Entwicklungsmustern erlaubt (vgl. *Menard*, 1991, S. 15). Beispielsweise haben *Abbott/Hrycak* (1990) auf diese Weise Karrierepfade analysiert und miteinander verglichen.

[Abbildung 1]

Wie *Abb. 1* verdeutlicht, liegt Längsschnittstudien ein mehrstufiger Prozess zugrunde. Zunächst muss ein Untersuchungsdesign spezifiziert werden, aus dem ein Datenbestand hervorgeht, der wiederum die Basis für verschiedene Analyseverfahren bildet. Der Datenbestand kann als Kopplungspunkt angesehen werden, weil der Bestand dieser Daten das vorgelagerte Erhebungsdesign mit der nachgelagerten Datenanalyse verbindet. Die konzeptionelle Trennung von Design und Daten ist erforderlich, weil verschiedene Designs die gleiche Datenbeschaffenheit erzeugen können, z.B. Paneldesign und retrospektives Design (vgl. *Campbell*, 1988, S. 47).

## 2. Längsschnittdesigns zur Erhebung von Longitudinaldaten

Zur Erzeugung von Längsschnittdaten können verschiedene Längsschnittdesigns verwendet werden. Konzeptionelle Unterschiede zwischen diesen Designs stellt *Tab. 1* anhand wesentlicher Merkmale dar.

[Tabelle 1]

Die verschiedenen Längsschnittsdesigns, die in *Tab. 1* vergleichend dargestellt wurden, hängen mit verschiedenen Datenqualitäten und Analyseverfahren zusammen. Diese Zusammenhänge werden schematisch in *Abb. 2* dargestellt und werden nachfolgend zusammen mit ihren charakteristischen Besonderheiten beschrieben.

[Abbildung 2]

Eine **simultane Querschnittsstudie** (engl. simultaneous cross-sectional study) liegt vor, wenn eine Querschnittserhebung zu einem einzigen Zeitpunkt durchgeführt wird, die so gewonnene Stichprobe aber nach verschiedenen Altersklassen geschichtet werden kann. Durch vergleichende Analyse der verschiedenen Altersgruppen können Veränderungen längsschnittanalog untersucht werden (vgl. *Helmreich, 1977, S. 12*). Vorteilhaft ist dieses Design, weil eine Erhebungswelle bereits ausreicht, um Untersuchungseinheiten verschiedener Altersklassen zu betrachten. Es handelt sich hierbei um keine Längsschnittstudie im eigentlichen Sinne.

**Trendstudien** (engl. repeated cross-sectional study) erfassen mit dem wiederholten Einsatz desselben Erhebungsinstruments verschiedene Untersuchungsobjekte. Damit handelt es sich um eine Reihe zeitlicher Querschnitte, weil – anders als bei Panelstudien – zu jeder Erhebungswelle eine neue Stichprobe eingesetzt wird (vgl. *Galtung, 1973, S. 85*). Deshalb kann nur die Nettoveränderung (engl. net change) auf aggregierter Ebene betrachtet werden, was z.B. durch Vergleich mehrerer Mittelwerte im Zeitablauf möglich ist (vgl. *Firebaugh, 1997, S. 3*). Wie diese aggregierte Veränderung auf individueller Ebene zustande kommt, ist der Analyse nicht zugänglich (vgl. *Helmreich, 1977, S. 11*). Trendstudien gestatten also den Vergleich von Zuständen zu verschiedenen Zeitpunkten. Der zeitliche und organisatorische Aufwand ist im Vergleich zur nachfolgend beschriebenen Paneluntersuchung geringer, weil bei jeder Erhebungswelle neue Untersuchungseinheiten befragt werden können und sich auch die Datenauswertung einfacher gestaltet (vgl. *Ruspini, 2002, S. 28*). Kausalzusammenhänge oder Entwicklungsmuster können hingegen nur ansatzweise untersucht werden (vgl. *Menard, 1991, S. 27*).

[Abbildung 3]

**Panelstudien** (engl. panel study) betrachten nicht nur verschiedene Zustände, sondern können darüber hinaus die teils gegenläufigen Bewegungen zwischen den Zuständen analysieren (vgl. *Dierkes, 1977, S. 127*). Dies ist möglich, weil eine Panelstudie dieselben Einheiten mit dem gleichen Erhebungsinstrument wiederholt untersucht und die in mehreren Wellen erhobenen Daten miteinander verknüpft (siehe *Abb. 3*). Dies erlaubt Aussagen über die Merkmalskonstanz und -veränderung einzelner Einheiten und lässt

so eine detaillierte Betrachtung von Veränderungsmustern zu. Auch für die Untersuchung kausaler Beziehungen bietet sich ein Paneldesign an (vgl. *Finkel*, 1995, S. 1). Als Panel wird die im Zeitverlauf mehrmals untersuchte gleiche Gruppe von Untersuchungsobjekten bezeichnet (vgl. *Lazarsfeld/Rosenberg/Thielens*, 1972, S. 253). Weil die Panelmitglieder über einen gewissen Zeitraum hinweg begleitet werden, kann es zum Ausfall von Untersuchungseinheiten kommen, was ein besonderes Problem von Panelstudien darstellt (vgl. *Dierkes*, 1977, S. 129). Besondere Analysemöglichkeiten ergeben sich durch Verknüpfung der Paneldaten mit weiteren Datenquellen (vgl. *Ruspini*, 2002, S. 38 f.).

Bei der **Kohortenuntersuchung** (engl. cohort study) besteht hinsichtlich der teilnehmenden Untersuchungseinheiten eine Restriktion (vgl. *Hakim*, 2000, S. 115): Untersuchungsobjekte bilden dann eine Kohorte, wenn bei ihnen zum gleichen Zeitpunkt ein bestimmtes Ereignis in ihrer Entwicklung eingetreten ist, z.B. Existenzbeginn, Börsengang, externer Schock (vgl. *Ryder*, 1968, S. 546). So bilden alle Personen, die im Jahr 1960 geboren sind, eine Kohorte, weil das Ereignis der Geburt in diesem Jahr diese Personengruppe eint. Durch Bezug auf dieses gemeinsame Ereignis kann die Entwicklung als besondere Form der Veränderung betrachtet werden.

Das **retrospektive Design** (engl. retrospective longitudinal study) ermöglicht nur Quasi-Längsschnittdaten, indem Untersuchungsobjekte zu vergangenen Entwicklungen und Ereignissen befragt werden (vgl. *Hakim*, 2000, S. 120). Weil bereits mit einer einzigen Erhebungswelle Längsschnittdaten für vergangene Zeiträume betrachtet werden können, werden die entstehenden Daten auch als Pseudo-Paneldaten bezeichnet. Basierend auf der Erinnerung werden hier rückblickend Daten aus der Vergangenheit erhoben. Nachteilig an diesem Design ist, dass die Qualität der Daten vom Erinnerungs- und Einschätzungsvermögen der untersuchten Einheiten abhängt. Weiterhin können kausale Prozesse nur unzureichend geprüft werden. Dafür bietet das retrospektive Design den Vorteil, über die Abfrage von Ereignisketten lückenlose Lebensverläufe nachzuzeichnen.

Bei der **Einzelfallbetrachtung** ergeben sich die Daten aus der „Betrachtung einer einzelnen Untersuchungseinheit“ (*Petermann*, 1989, S. 3). Dies können z.B. Personen, Unternehmen, ganze Branchen oder Gesellschaften sein. Serielle Daten, wie sie bei der Einzelfallbetrachtung anfallen, lassen sich mit Techniken der Zeitreihenanalyse untersuchen (vgl. *Kratochwill*, 1978, S. 1 f.). Schwerpunkt dieser Analyse ist das Erkennen von Trends und saisonalen Effekten, aber auch von Diskontinuitäten (vgl. *Winker*, 1997,

S. 50): Die Wertesequenz für eine Zeitreihenuntersuchung umfasst eine große Anzahl von Werten. Eine solche Wertereihe bezieht sich auf eine oder sehr wenige Beobachtungsobjekte.

Wurde durch die Datenerhebung ein Datenbestand generiert, so kann eine Vielzahl von Methoden zur Datenanalyse eingesetzt werden. Je nach Datenbestand und je nach Ausrichtung der Forschungsfrage eignen sich verschiedene Analysemethoden, von denen nachfolgend die wichtigsten Verfahren vorgestellt werden sollen.

### 3. Modelle zur Analyse von Längsschnittdaten

Daten, die mittels Längsschnittdesigns erhoben wurden, können – bevor sie in den Analyseprozess eingehen – auf verschiedene Weise transformiert werden. Die Umformung des Datenbestands lässt damit eine **Verbreiterung des Spektrums möglicher Analyseverfahren** zu.

Die Analyse von Längsschnittdaten und das eingesetzte Analyseverfahren sind von der Forschungsfrage abhängig: Je nachdem, ob Ereignisse untersucht, kausale Beziehungen analysiert oder Entwicklungsmuster aufgedeckt werden sollen, sind bestimmte Analyseansätze zu wählen (vgl. *Goldstein*, 1979, S. 3 f.). *Menard* (1991, S. 49 ff.) unterscheidet **drei Analyserichtungen**, denen verschiedene statistische Modelle gerecht werden können. Erstens können zeitliche Entwicklungsmuster etwa mit Markov-Modellen, log-linearen Modellen und Modellen der Ereignisdatenanalyse beschrieben werden. Zweitens können zeitliche Reihenfolgen beschrieben werden, zu deren Untersuchung z.B. die Sequenzanalyse eingesetzt werden kann, um ganzheitliche Trajektorien zu betrachten. Drittens können kausale Zusammenhänge analysiert werden, denn der Verlauf der Zeit gibt den Messwerten eine eindeutige Reihenfolge der Geschehnisse vor (vgl. *Finkel*, 1995, S. 22 und 87). So können die Einflussfaktoren für Veränderungen und Entwicklungen erörtert werden.

Veränderungen können mittels **diskreter oder kontinuierlicher Variablen** gemessen werden, was die Wahl des Analyseverfahrens beeinflusst. *Abb. 4* bietet eine Zusammenstellung einiger Verfahren zur Längsschnittdatenanalyse. Die aufgeführten Methoden werden nun detailliert vorgestellt, wobei nicht nur das grundlegende Konzept der jeweiligen Methode, sondern auch auf die Einsatzmöglichkeiten eingegangen wird. Erst ein Überblick über die zur Verfügung stehenden Methoden ermöglicht den angemessenen Einsatz einzelner Verfahren.

[Abbildung 4]

Die **Zeitreihenanalyse** (engl. time series analysis) hat nicht nur das Ziel, dynamische Phänomene mit Hilfe von zeitlichen Datenreihen zu identifizieren und zu beschreiben, sondern auch zukünftige Werte auf Basis vergangener Daten zu prognostizieren (vgl. *Hartung/Elpelt/Klösener*, 1998, S. 637). Hierzu wird angenommen, dass die Höhe eines Wertes durch eine bestimmte Systematik und einen zufälligen Fehlerterm (engl. random noise, white noise) bestimmt wird (vgl. *Dierkes*, 1977, S. 153, *Menard*, 1991, S. 51). Basierend hierauf versuchen statistische Verfahren eine Trennung dieser Bestandteile (vgl. *Freitag*, 2003, S. 27 f.). Vorhersagen sind dann durch Fortschreiben des systematischen, aber latenten Wertes möglich.

Es können statistische Methoden eingesetzt werden, um latente Muster hinter den tatsächlichen Zeitreihen zu identifizieren. Diese systematischen Muster können durch Trends, saisonale Effekte oder eine Kombination aus beidem beschrieben werden (vgl. *Ruspini*, 2002, S. 108). Ein Trend stellt ein systematisches Verhalten – ob linear oder nicht – dar, das sich nicht wiederholt und damit eine Grundrichtung vorgibt. Saisonale Effekte basieren auf einer periodisch, also in bestimmten Zeitabständen, wiederkehrenden Systematik. Ein weiterer Analyseschwerpunkt ist die Autokorrelation. Hierbei wird angenommen, dass die Höhe des konkreten Wertes einer Zeitreihe nicht willkürlich ist, sondern unmittelbar von den vorherigen Werten der Zeitreihe abhängt.

[Abbildung 5]

**Regressions-, Pfad- und Strukturgleichungsmodelle** können dem explikativen Erkenntnisinteresse gerecht werden, da sie die Analyse kausaler Verknüpfungen durch Betrachtung von Ursache und Wirkung erlauben. Die Unterscheidung von Ursache und Wirkung wird dabei von Längsschnittdaten erleichtert, weil durch die Messzeitpunkte eine eindeutige Reihenfolge auf der Zeitachse gegeben ist (vgl. *Engel/Reinecke*, 1994, S. 8). Damit eignen sich Längsschnittdaten für die in diesem Abschnitt erläuterten Modelle besonders (vgl. *Rogosa*, 1979, S. 264 f.), weil sie Daten verschiedener Zeitpunkte simultan analysieren und die bestehenden intertemporalen Abhängigkeiten innerhalb eines Modells untersuchen. Dies erfordert mehrere zeitversetzte Messungen einer Variablen. Grundlage dieser Modellgruppe ist die Regressionsanalyse. **Regressionsmodelle** (engl. regression models) formulieren eine Beziehung zwischen mehreren Variablen und können damit die Wirkung einer oder mehrerer unabhängiger Variablen auf eine abhängige abbilden (vgl. *Markus*, 1979, S. 46 und 48).

Bei **Pfadmodellen** (engl. path model) dagegen handelt es sich um miteinander verbundene multiple Regressionsmodelle, die komplexere Zusammenhangsstrukturen zwischen Variablen zu untersuchen vermögen. Pfadmodelle erlauben dabei die Modellie-

rung mehrstufiger Prozesse und können Quer- und Längsschnitteffekte gleichzeitig berechnen (vgl. *Engel/Reinecke*, 1994, S. 22). Häufig können die in den Pfadmodellen berücksichtigten Variablen aber nicht direkt gemessen werden.

**Strukturgleichungsmodelle** (engl. structural equation model) ermöglichen die Konzeptualisierung von latenten Variablen, die sich einer direkten Beobachtung entziehen und nur indirekt über beobachtbare Indikatorvariablen gemessen werden können (vgl. *Urban*, 2004, S. 17, *Bollen*, 1989, S. 18). Implizit werden also Messfehler für diese hypothetischen Konstrukte berücksichtigt. Die beobachteten Variablen werden dabei auch als Indikatoren, die hypothetischen Konstrukte als latente Variablen bezeichnet. Strukturgleichungsmodelle bilden eine Synthese aus Pfadanalyse und Faktorenanalyse, wobei die Faktorenanalyse prinzipiell eine Bündelung von Variablen vornimmt und so latente Konstrukte konzeptualisiert. Die Besonderheit von Strukturgleichungsmodellen liegt im Allgemeinen in der Möglichkeit, ein komplexes Geflecht von intertemporalen kausalen Beziehungen zu analysieren. Der Hauptzweck von Strukturgleichungsmodellen bildet das Überprüfen statistischer Hypothesen über die Zusammenhänge mehrerer Variablen (vgl. *Rogosa*, 1979, S. 266 f.). Diese Hypothesen werden auf Basis einer Theorie formuliert und ergeben ein Modell, das mit der Struktur in der Realität möglichst genau übereinstimmen sollte und mittels Gleichungssystemen dargestellt wird. Intertemporale Relationen können abgebildet werden, indem das Modell nicht nur die strukturellen Beziehungen eines Zeitpunkts wie bei einer Querschnittsanalyse enthält, sondern Variablen verschiedener Zeitpunkte innerhalb eines Modells integriert (vgl. *Kline*, 1998, S. 261).

*Abb. 5a* zeigt ein Beispiel für ein Strukturgleichungsmodell, das zwei Messzeitpunkte derselben Variablen enthält. Dabei wird ersichtlich, wie jeweils zwei Indikatoren eine latente Variable messen. Diese steht in zeitpunktübergreifenden Beziehungen zu anderen latenten Variablen. Strukturgleichungsmodelle werden auch als Kausalmodelle (engl. causal models) oder Kovarianzstrukturanalysen (engl. covariance structure analysis) bezeichnet.

**Latente Wachstumskurvenmodelle** (engl. latent growth curve models) können auf individueller Basis Veränderungsvorgänge analysieren, um bestimmte Muster in einzelnen Veränderungspfaden aufzudecken. Mit diesen Modellen lassen sich differenzierte Entwicklungen verschiedener Objekte berücksichtigen und somit individuelle Entwicklungen wie z.B. Alterungs- oder Lernprozesse beschreiben (vgl. *Urban*, 2004, S. 19 f.). Besonderes Analysepotenzial haben diese Modelle, wenn sie mit Strukturgleichungs-

modellen verknüpft werden. Dann besteht die Möglichkeit, Veränderungsmuster in der zeitlichen Entwicklung latenter Konstrukte zu betrachten.

Mit **log-linearen Modellen** werden typischerweise Assoziations- und Interaktionsmuster zwischen kategorialen Daten untersucht, wenn zunächst keine kausalen Beziehungen angenommen werden (vgl. *Agresti*, 1984, S. 47). Dazu wird eine Kreuztabelle mehrerer kategorialer Variablen eingesetzt. „Das allgemeine Ziel der log-linearen Analyse ist [...] darin zu sehen, die Zelhäufigkeiten [...] als Funktion genau der Variablen zu studieren, welche die (Mobilitäts-)Tabelle aufspannen.“ (*Engel/Meyer*, 1996, S. 140) Bei der Häufigkeitenschätzung werden dann Zeilen-, Spalteneffekte und deren Kombinationen berücksichtigt (vgl. *Hagenaars*, 1993, S. 11). Die zur Klassifikation verwendeten Merkmale spannen die Kreuztabelle auf und sind somit die unabhängigen Variablen, die in die Schätzung der Zelhäufigkeit als der abhängigen Variablen eingehen (vgl. *Glogg/Shockey*, 1988, S. 364). Bei **Logit-, Probit- und logistischen Modellen** handelt es sich um Modifikationen von log-linearen Modellen. Diese Modelle können für die Längsschnittanalyse eingesetzt werden, indem in einer Kreuztabelle verschiedene Variablen ihren Messzeitpunkten gegenübergestellt werden, was *Abb. 5b* veranschaulicht. Die Zeit selbst kann unabhängige Variable sein.

Anders als log-lineare Modelle stellen **Markov-Modelle** Daten verschiedener Zeitpunkte nicht als Vergleich „nebeneinander“ dar, sondern analysieren einen Satz individueller Wechsellvorgänge simultan, indem sie in einer einzigen Kreuztabelle Variablenausprägungen verschiedener Zeitpunkte einander gegenüberstellen (vgl. *Menard*, 1991, S. 52, *Ruspini*, 2002, S. 118). Sie wurden speziell für die Längsschnittdatenanalyse entwickelt und können intertemporale Muster individueller Veränderungsprozesse abbilden, da sie die Bewegungen zwischen verschiedenen Zuständen betrachten. Im einfachsten Fall wird eine solche Matrix für kategoriale Daten zu diskreten Zeitpunkten generiert, die dann die Übergangswahrscheinlichkeiten für mögliche Zustandswechsel enthält (vgl. *Plewis*, 1985, S. 147, *Markus*, 1979, S. 8). Bei einer bivariaten Markov-Matrix, wie sie in *Abb. 5c* skizziert ist, sind Informationen über die Merkmalskonstanz in den diagonalen Zellen zu finden, während die anderen Zellen Informationen über Zustandsveränderungen enthalten.

Bei der **Ereignisdatenanalyse** (engl. event history analysis), die auch Lebensdauer-, Überlebens-, Überlebensdauer- oder Zeitdaueranalyse (engl. survival analysis, analysis of duration data) genannt wird, werden prinzipiell keine Zustände, sondern die Zeitpunkte von Ereignissen betrachtet (vgl. *Allison*, 1990, S. 1). Prinzipiell stellen Ereignis-



se die Veränderungen von Zuständen dar und gewährleisten so einen höheren Informationsgehalt als die Betrachtung von Zuständen, da keine diskret sondern kontinuierlich aufgezeichnete Daten vorliegen (vgl. *Hakim*, 2000, S. 112). Bei der Ereignisdatenanalyse handelt es sich um eine Familie statistischer Verfahren zur Untersuchung von Lebensverläufen (vgl. *Yamaguchi*, 1991, S. 1). Ebenso wie Personen können auch Unternehmen, Produkte oder Märkte solche Lebensverläufe aufweisen. Untersuchungsobjekte wie z.B. Produkte oder Märkte durchlaufen ihre spezifischen Lebenswege als Trajektorien, die durch aneinander gekettete Phasen gekennzeichnet sind (siehe *Abb. 5d*).

Der jeweilige Status während dieser aneinander gereihten Phasen ergibt die abhängige Variable (vgl. *Petersen*, 1990, S. 260), die somit im kontinuierlichen Zeitverlauf durch diskrete Zustände bzw. kategoriale Daten repräsentiert wird (vgl. *Blossfeld/Hamerle/Mayer*, 1986, S. 18). Die abhängige Variable wie z.B. eine Phase im Produktlebenszyklus wird dabei von einer Reihe unabhängiger kovariierender Variablen begleitet, z.B. Marktvolumen, Absatzvolumen, Preiselastizität. Ziel des Analysevorgangs stellt es, von diesen unabhängigen begleitenden Variablen, auf den Übergang von einem Status in einen anderen zu schließen (vgl. *Petersen*, 1990, S. 260). Sofern Wahrscheinlichkeiten für die Eintrittszeitpunkte von Zuständen berechnet werden können, wird die Prognose von Statusübergängen und damit die Vorhersage von Ereignissen und Phasenzeitdauern möglich (vgl. *Stovel/Bolan*, 2004, S. 561 f.). Beispielsweise können so die Lebensdauer von technischen Geräten oder der Zeitpunkt von Unternehmensinsolvenzen bzw. -fusionen untersucht und geschätzt werden.

Bei der **Sequenzanalyse** (engl. sequence analysis) wird nicht der spezifische Eintritt von Ereignissen, sondern die komplette Sequenz als serielle Abfolge aufeinander folgender Ereignisse analysiert (vgl. *Abbott*, 1995, S. 206 f.). In die Analyse fließen nicht nur die Eintrittszeitpunkte und Zeitdauern aufeinander folgender Ereignisse, sondern auch die Reihenfolgen der Ereignisse und die Position innerhalb der Ereignissequenzen ein, was einem hohen Informationsgehalt entspricht (vgl. *Coleman*, 1981, S. 185). So können Muster in Ereignissequenzen entdeckt und die synchrone Abfolge mit anderen Untersuchungsobjekten abgeglichen werden. Damit erlaubt die Sequenzanalyse die ganzheitliche Betrachtung der Trajektorien von Untersuchungsobjekten (vgl. *Stovel/Bolan*, 2004, S. 562). Insbesondere eignet sich die Sequenzanalyse zur Untersuchung von Systemen oder Organisationen (vgl. *Stovel/Bolan*, 2004, S. 562).

Das Prinzip der Sequenzanalyse verkörpert, wie in *Abb. 5e* dargestellt, die „Formulierung“ spezifischer Buchstabenfolgen oder Zeichenketten, die den Ereignisketten der

Trajektorien entsprechen (vgl. *Abbott*, 1995, S. 211). Wenn jedem Ereignis ein Buchstabe zugewiesen wird, entsteht aus einer Ereignissequenz eine Buchstabenfolge (vgl. *Frederiksen/Rotondo*, 1979, S. 113). Bei der Analyse mehrerer Untersuchungsobjekte können anschließend Distanzmaße zwischen verschiedenen Buchstabenfolgen errechnet werden, die die Ähnlichkeit verschiedener Sequenzen angeben (vgl. *Ruspini*, 2002, S. 131, *Taris*, 2000, S. 128 f.). Damit lässt sich etwa eine Klassifikation verschiedener Sequenzen erreichen. Sequenzanalysen haben den Vorteil, dass sie anstatt Momentaufnahmen Muster ganzheitlicher Lebensverläufe untersuchen können (z.B. Entscheidungsprozesse in Unternehmen, Entwicklungsprofile, Produktionsprozesse, Karrierepfade, Marktstrukturen und Konsumentenverhalten; vgl. *Abbott/Hrycak*, 1990, S. 99 f.).

#### 4. Fazit

Der Beitrag zeigt die **breite Palette möglicher Designs** und **Analysemethoden** von Längsschnittstudien auf und grenzt die verschiedenen Optionen voneinander ab. Anders als Querschnittstudien untersuchen Längsschnittstudien Entwicklungen und Veränderungen im Zeitablauf. Auch die Untersuchung von systematischen Veränderungsmustern ist möglich. Letztlich bieten Längsschnittdesigns durch ihre inhärente zeitliche Ordnung der Daten eine ideale Basis, um kausale Zusammenhänge zu untersuchen.

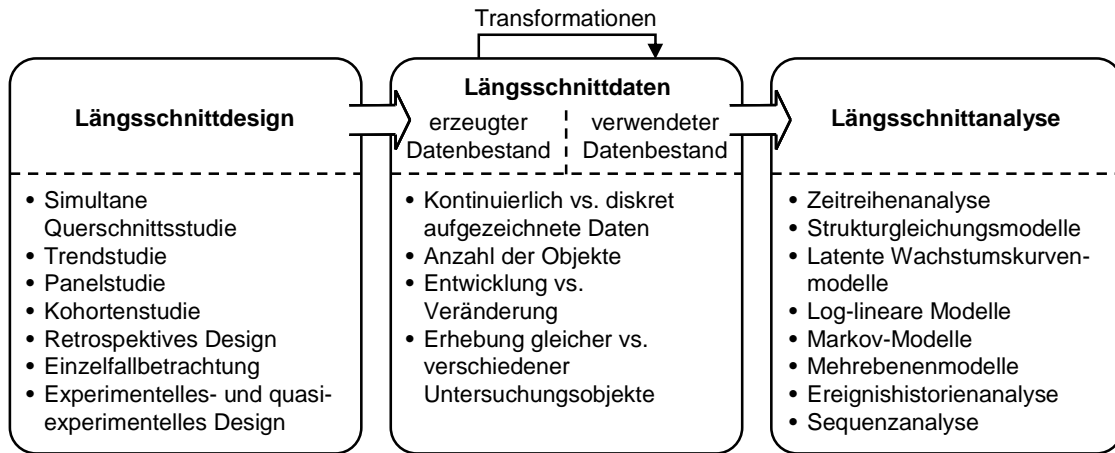
Zunächst bestimmt die spezifische Konfiguration der vorliegenden oder zu erhebenden Längsschnittdaten den **Einsatz verschiedener Analyseverfahren**. Weiterhin ist aber vor allem das Erkenntnisinteresse ein maßgeblicher Faktor, da die Untersuchung von Veränderungsmustern andere Analysemethoden erfordert als z.B. die Analyse von Kausalitäten.

Durch die **Berücksichtigung der Dimension Zeit** verbreitert sich die Menge an denkbaren Forschungsfragen. Damit verbunden ist auch eine höhere Komplexität der Durchführung einer Studie. Erst das Wissen über Längsschnittdesigns und Analyseverfahren und die Fähigkeit einzelne Konzepte voneinander abzugrenzen, ermöglicht es, aus dem breiten Inventar an verfügbaren Instrumenten die für den jeweiligen Untersuchungszweck am besten geeignete Methodik auszuwählen.

		Simultane Querschnitte	Trendstudie	Panelstudie	Kohortenstudie	Einzelfallbetrachtung	Retrospektives Design	Experiment, Quasi-Experiment
Anzahl Erhebungswellen	eine möglich	•					•	
	mehr als 2		•	•	•	•		•
Datenaufzeichnung	diskret	•	•	•	•	•		•
	kontinuierlich						•	
Objektanzahl	ein Objekt					•		
	mehrere	•	•	•	•		•	•
Zeitbezug der erhobenen Daten	aktuell	•	•	•	•	•		•
	rückblickend						•	
Objektauswahl im Zeitablauf	die gleichen			•	•	•	•	•
	verschiedene	•	•		•			
Beobachtungsschwerpunkt	Veränderung		•	•		•		
	Entwicklung	•			•			
	Ereignisse						•	•

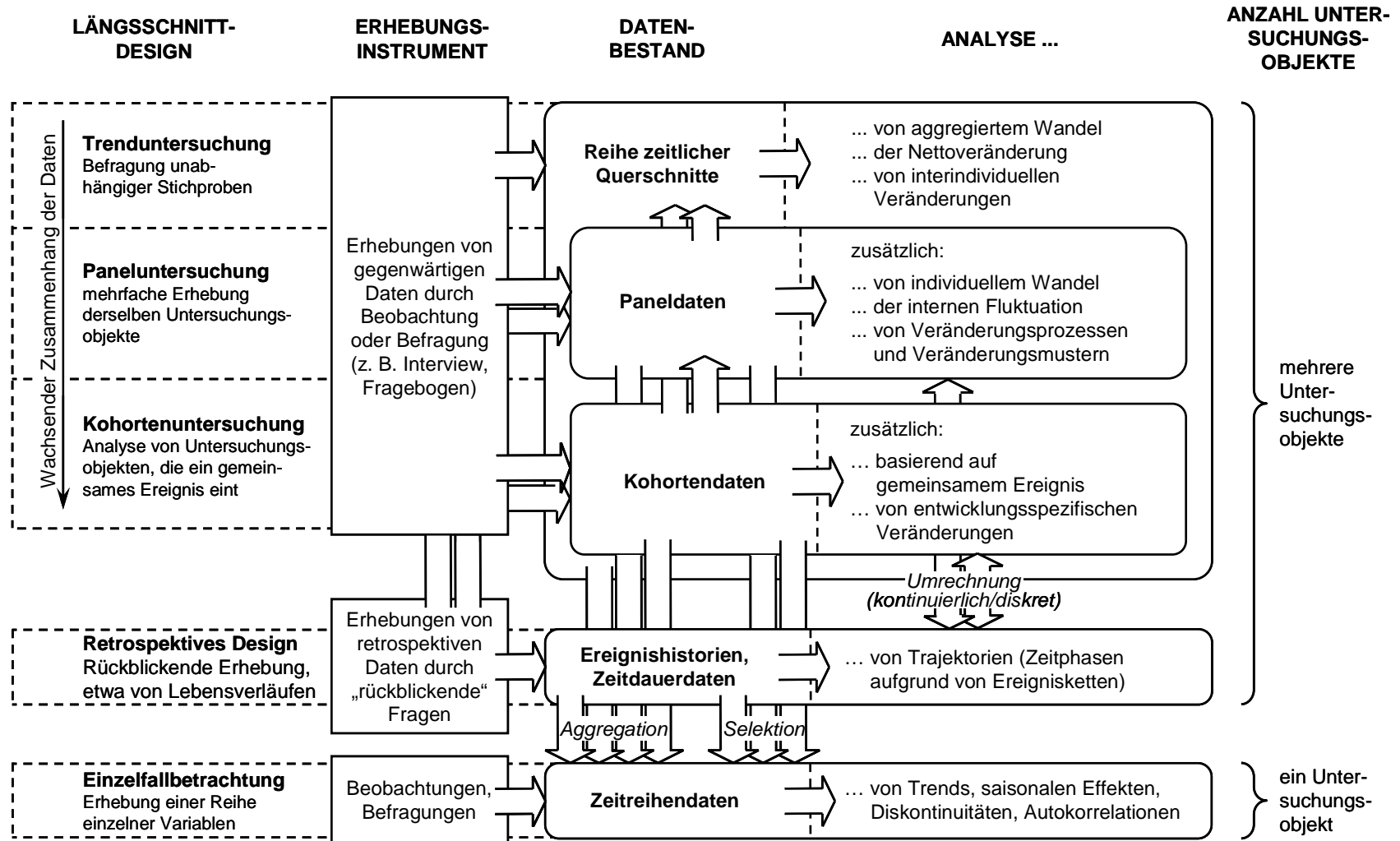
*Tab. 1: Vergleich der Schwerpunkte unterschiedlicher Längsschnittsdesigns*

## Abbildungen



*Abb. 1: Forschungsprozess bei Längsschnittstudien*

Abb. 2: Zusammenhang und Systematisierung von Längsschnittdesigns



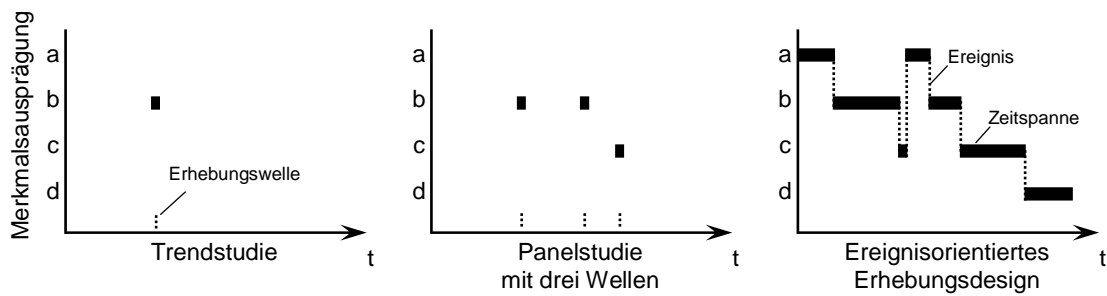
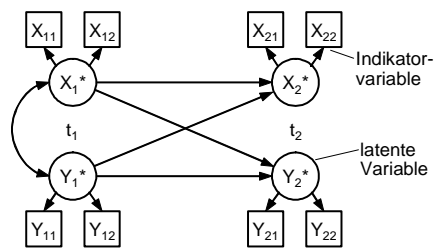


Abb. 3: Kontinuierlich und diskret aufgezeichnete Daten

	Modelle für diskrete Variablen	Modelle für diskrete und kontinuierliche Variablen	Modelle für kontinuierliche Variablen
Modelle mit manifesten Variablen	<ul style="list-style-type: none"> <li>• Markov-Modelle</li> <li>• Log-lineare Modelle</li> <li>• Sequenzanalyse</li> </ul>	<ul style="list-style-type: none"> <li>• Varianzanalyse (ANOVA)</li> <li>• Diskriminanzanalyse</li> <li>• Logit- und Probit-Modelle</li> <li>• Mehrebenenmodelle</li> <li>• Ereignisdatenanalyse</li> </ul>	<ul style="list-style-type: none"> <li>• Regressionsmodelle</li> <li>• Pfadmodelle</li> <li>• Zeitreihenanalyse</li> </ul>
Modelle mit latenten Variablen	<ul style="list-style-type: none"> <li>• Latente Klassenmodelle</li> <li>• Latentes Markov-Modell</li> <li>• Log-lineare Modelle mit latenten Variablen</li> <li>• Strukturgleichungsmodelle für ordinale Variablen</li> </ul>		<ul style="list-style-type: none"> <li>• Strukturgleichungsmodelle</li> <li>• Latente Wachstums-kurvenmodelle</li> </ul>

Abb. 4: Übersicht der Methoden zur Analyse von Längsschnittdaten



**a. Exemplarisches Strukturgleichungsmodell für zwei Erhebungswellen**

		$t_1$	$t_2$	$t_3$
$b_1$	$a_1$			
	$a_2$			
$b_2$	$a_1$			
	$a_2$			

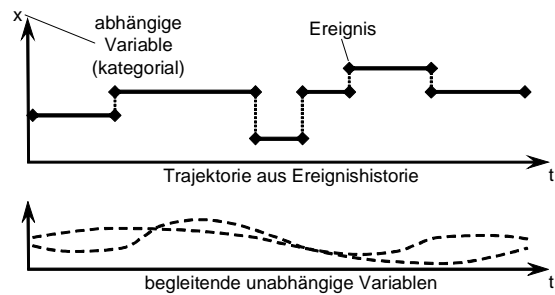
Messzeitpunkte  
Zellhäufigkeit, abhängig von Spalten- und Zeileneffekten  
 $m_{abl}$

**b. Log-lineares Modell zur Untersuchung von Längsschnittdaten**

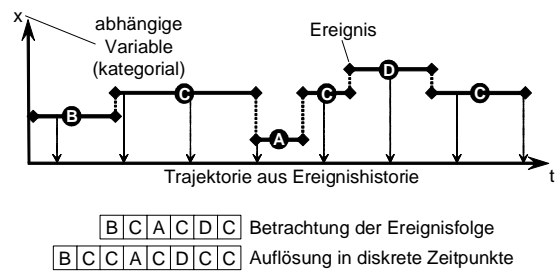
		$t_1$	$t_2$	$t_3$
$t_1$	$a_1$			
	$a_2$			
	$a_3$			

Zustandsveränderung  
Zustandsstabilität  
 $m_{ij}$

**c. Markov-Modell zur Analyse von Zustandsübergängen**



**d. Ereignisdatenanalyse zur Untersuchung von Längsschnittdaten**



**e. Sequenzanalyse in der Längsschnittanalyse**

Abb. 5: Veranschaulichung einiger Methoden zur Analyse von Längsschnittdaten

## Literatur

- Abbott, A.*, A Primer on Sequence Models, in: *G. P. Huber, A. H. Van de Ven* (Hrsg.), *Longitudinal Field Research Methods: Studying Processes of Organizational Change*, Thousand Oaks 1995, S. 204–227.
- Abbott, A., A. Hrycak*, Measuring Resemblance in Sequence Data: An Optimal Matching Analysis of Musicians' Careers, in: *American Journal of Sociology*, Vol. 96 (1990), Nr. 1, S. 144–185.
- Agresti, A.*, *Analysis of Ordinal Categorical Data*, New York 1984.
- Allison, P. D.*, *Event History Analysis: Regression for Longitudinal Event Data*, London 1990.
- Baltes, P. B., J. R. Nesselroade*, History and Rationale of Longitudinal Research, in: *J. R. Nesselroade* (Hrsg.), *Longitudinal Research in the Study of Behavior and Development*, New York 1979, S. 1–39.
- Blossfeld, H.-P., A. Hamerle, K. U. Mayer*, *Ereignisanalyse: statistische Theorie und Anwendung in den Wirtschafts- und Sozialwissenschaften*, Frankfurt 1986.
- Bollen, K. A.*, *Structural Equations with Latent Variables*, New York 1989.
- Campbell, R. T.*, Integrating Conceptualization, Design, and Analysis in Panel Studies, in: *K. W. Schaie, R. T. Campbell, W. Meredith, S. C. Rawlings* (Hrsg.), *Methodological Issues in Aging Research*, New York 1988, S. 43–70.
- Coleman, J. S.*, *Longitudinal Data Analysis*, New York 1981.
- Dierkes, M.*, Die Analyse von Zeitreihen und Longitudinalstudien, in: *J. von Koolwijk, M. Wieken-Mayser* (Hrsg.), *Erhebungsmethoden: Datenanalyse*, 7. Bd., München/Wien 1977, S. 111–169.
- Engel, U., W. Meyer*, Structural Analysis in the Study of Social Change, in: *U. Engel, J. Reinecke* (Hrsg.), *Analysis of Change: Advanced Techniques in Panel Data Analysis*, Berlin 1996, S. 221–252.
- Engel, U., J. Reinecke*, *Panelanalyse*, Berlin 1994.
- Finkel, S. E.*, *Causal Analysis with Panel Data*, Thousand Oaks 1995.
- Firebaugh, G.*, *Analyzing Repeated Surveys*, Thousand Oaks 1997.



- Frederiksen, C. H., J. A. Rotondo*, Time-Series Models and the Study of Longitudinal Change, in: *J. R. Nesselroade* (Hrsg.), *Longitudinal Research in the Study of Behavior and Development*, New York 1979, S. 1–39.
- Freitag, K.*, *Zeitreihenanalyse: Methoden und Verfahren*, Lohmar 2003.
- Galtung, J.*, *Theory and Methods of Social Research*, London 1973.
- Glogg, C. C., J. W. Shockey*, Multivariate Analysis of Discrete Data, in: *J. R. Nesselroade, B. C. Raymond* (Hrsg.), *Handbook of Multivariate Experimental Psychology*, 2. Aufl., New York 1988, S. 337–365.
- Goldstein, H.*, *The Design and Analysis of Longitudinal Studies: Their Role in the Measurement of Change*, London 1979.
- Hagenaars, J. A.*, *Loglinear Models with Latent Variables*, Newbury Park 1993.
- Hakim, C.*, *Research Design: Successful Designs for Social and Economic Research*, 2. Aufl., London 2000.
- Hartung, J., B. Elpelt, K.-H. Klösener*, *Statistik: Lehr- und Handbuch der angewandten Statistik*, 11. Aufl., München 1998.
- Helmreich, R.*, *Strategien zur Auswertung von Längsschnittdaten: ein Beitrag zur Messung von Veränderung in der empirischen Sozialforschung*, Stuttgart 1977.
- Kline, R. B.*, *Principles and Practice of Structural Equation Modeling*, New York 1998.
- Kratochwill, T. R.*, Foundations of Time-series Research, in: *T. R. Kratochwill* (Hrsg.), *Single Subject Research: Strategies for Evaluating Change*, New York 1978, S. 1–100.
- Lazarsfeld, P. F., M. Rosenberg, W. Thielens*, Die Panel-Befragung, in: *R. König* (Hrsg.), *Das Interview: Formen, Technik, Auswertung*, 7. Aufl., Köln 1972, S. 253–268.
- Markus, G. B.*, *Analyzing Panel Data*, Beverly Hills 1979.
- Menard, S.*, *Longitudinal Research*, Newbury Park 1991.
- Petermann, F.*, Einzelfallanalyse – Definitionen, Ziele und Entwicklungslinien, in: *F. Petermann* (Hrsg.), *Einzelfallanalyse*, 2. Aufl., München 1989, S. 1–17.
- Petersen, T.*, Analyzing Event Histories, in: *A. von Eye* (Hrsg.), *Statistical Methods in Longitudinal Research: Time Series and Categorical Longitudinal Data*, 2. Bd., Boston 1990, S. 259–288.

- Plewis, I.*, *Analysing Change: Measurement and Explanation Using Longitudinal Data*, Chichester 1985.
- Rogosa, D.*, Causal Models in Longitudinal Research: Rationale, Formulation, and Interpretation, in: *J. R. Nesselroade* (Hrsg.), *Longitudinal Research in the Study of Behavior and Development*, New York 1979, S. 263–302.
- Ruspini, E.*, *Introduction to Longitudinal Research*, London 2002.
- Ryder, N. B.*, Cohort Analysis, in: *D. L. Sills* (Hrsg.), *International Encyclopedia of the Social Sciences*, 2. Bd., New York 1968, S. 546–550.
- Sankoff, D., J. B. Kruskal*, *Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison*, Reading 1983.
- Stovel, K., M. Bolan*, Residential Trajectories: Using Optimal Alignment to Reveal the Structure of Residential Mobility, in: *Sociological Methods and Research*, Vol. 32 (2004), Nr. 4, S. 559–598.
- Taris, T. W.*, *A Primer in Longitudinal Data Analysis*, London 2000.
- Urban, D.*, *Neue Methoden der Längsschnittanalyse*, Münster 2004.
- Winker, P.*, *Empirische Wirtschaftsforschung*, Berlin/Heidelberg 1997.
- Yamaguchi, K.*, *Event History Analysis*, Newbury Park 1991.