

Giesselmann, Marco; Windzio, Michael

Article — Accepted Manuscript (Postprint)

Paneldaten in der Soziologie: Fixed Effects Paradigma und empirische Praxis in Panelregression und Ereignisanalyse

KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Giesselmann, Marco; Windzio, Michael (2014) : Paneldaten in der Soziologie: Fixed Effects Paradigma und empirische Praxis in Panelregression und Ereignisanalyse, KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie, ISSN 0023-2653, Springer Fachmedien, Wiesbaden, Vol. 66, Iss. 1, pp. 95-113,
<https://doi.org/10.1007/s11577-013-0248-z>

This Version is available at:

<https://hdl.handle.net/10419/109636>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Paneldaten in der Soziologie: Fixed Effects Paradigma und empirische Praxis in Panelregression und Ereignisanalyse

Marco Giesselmann und Michael Windzio

Original in: Kölner Zeitschrift für Soziologie und Sozialpsychologie (2014) 66, 95–113
DOI 10.1007/s11577-013-0248-z

Zusammenfassung

In der soziologischen Methodenforschung werden Paneldaten als Instrument diskutiert, welches eine fundiertere Überprüfung einer kausalen Hypothese als Querschnittsdaten ermöglicht. Zur Realisierung dieses Potentials ist allerdings die Isolierung intraindividuelle Zusammenhänge im Rahmen der Schätzprozedur notwendig. Bei multivariaten Regressionsanalysen leistet dies die *Fixed Effects Regression*, welche sich daher im methodologischen Diskurs als Standard zur Analyse von Paneldaten etabliert hat. In dieser Studie wird *erstens* dieser Standard aus verschiedenen Perspektiven begründet und zudem sein Äquivalent im ereignisanalytischen Kontext diskutiert. *Zweitens* stellen wir auf der Grundlage einer Vollerhebung von panelbasierten Studien in KZfSS und ZfS fest, dass in der soziologischen Praxis weder die Kausalitätsannäherung als Motiv zur Verwendung von Paneldaten dominiert, noch der *Fixed Effects*-Schätzer konsequent eingesetzt wird. Diese Inkonsistenz wird als Ausdruck eines *underusage* von Paneldaten in der empirischen Praxis gedeutet. Gleichzeitig erkennen wir hierin ein Indiz dafür, dass Motive zur Nutzung von Paneldaten in der Soziologie vielfältiger sind, als es der methodologische Diskurs nahelegt.

Schlüsselwörter: Paneldaten, Panelanalyse, Ereignisanalyse, Fixed Effects

Panel Data in Sociology: The Fixed Effects Paradigm and Empirical Practice in Panelregression and Event-History Analysis

Abstract

In sociology there is a growing body of literature discussing the ability of quantitative research methods to test causal inferences. This discourse introduces panel data as an important instrument to defend the interpretation of coefficients as effects: by focusing on what is going on *within* individuals, panel data allow for longitudinal empirical modelling, which reflects the inherent semantic of scientific hypotheses. As a consequence, unobserved heterogeneity is absorbed and alternative explanations to the presumed causal mechanisms are largely ruled out. However, a within-, or rather *Fixed Effects*-estimator is needed to realize this potential of panel data. Consequently, such within-estimators are well established as standard in the methodological discourse on panel data. We explain this standard with references to the mechanics of within-estimators in panel-regression and event-history analysis. Finally, inspection of contributions to the two major German language journals for sociology shows that within-estimators are rarely used in empirical practice. We conclude that the potential of panel data to control for heterogeneity is largely underused in empirical

practice. At the same time, our inspection of studies illustrates that the benefits of panel data are multidimensional and go beyond statistical virtues.

Key Words: Panel Data, Panel Analysis, Event-history Analysis, Fixed Effects

1 Einleitung

Personenbasierte Paneldaten sind in den Sozialwissenschaften zu einer zentralen empirischen Grundlage geworden. Dies verdeutlicht nicht nur die Forschungspraxis, sondern auch die zunehmende Bereitstellung von Paneldaten im Rahmen wissenschaftlicher Infrastrukturprojekte. Gerade in Deutschland ist die jüngere Förderpraxis von Erhebungsprojekten stark auf das Leitbild der Längsschnittstudie ausgerichtet: Neben dem Sozio-oekonomischen Panel (Wagner et al. 2007) wachsen derzeit mehrere Erhebungsprojekte, deren Themen im weitesten Sinne der Soziologie zugeordnet werden können, in ein komplexes und leistungsfähiges Längsschnittformat. Hierzu zählen insbesondere das Nationale Bildungspanel (Blossfeld et al. 2011), das Deutsche Beziehungs- und Familienpanel (Huinink et al. 2011) sowie das Panel *Arbeitsmarkt und soziale Sicherung* (Trappmann et al. 2010). Auch international vergleichend angelegte Infrastrukturprojekte stellen verstärkt auf Wiederholungsbefragungen ab. Dies betrifft beispielsweise die SHARE-Studie (Börsch-Supan und Jürges 2005) sowie die EU-SILC-Erhebung (Eurostat 2005). Da die Längsschnittpotentiale dieser Datensätze mit fortlaufender Erhebungsdauer immer bedeutsamer werden, lässt sich prophezeien, dass Paneldaten die soziologische Forschungskultur in den nächsten Jahren zunehmend prägen werden.

Dieser Prozess wird flankiert von einem Methodendiskurs, der die Frage nach der Berechtigung kausaler Interpretationen in der empirischen Soziologie in den Mittelpunkt rückt (Soebel 2000, Gangl 2010; Legewie 2012). Im Rahmen dieses Diskurses werden nun Paneldaten oft als Instrument diskutiert, welches die Möglichkeiten zur empirisch-statistischen Absicherung kausaler theoretischer Modellierungen substantiell verbessert.¹ Dieses Potential wird darauf zurückgeführt, dass Paneldaten (durch die Kontrolle *personenspezifischer Heterogenität*) den großzügigen Ausschluss von Alternativerklärungen zur angepeilten kausalen Interpretation ermöglichen (Brüderl 2010; Giesselmann und Windzio 2012; Andreß et al. 2013).

Ziel dieser Studie ist es nun, diesen Diskurs und das hierin entwickelte *Fixed Effects Analyseparadigma* für Paneldaten aufzubereiten und zusammenzufassen. Des Weiteren legen wir eine Übersicht vor, die aufzeigt, wie Paneldaten in der Soziologie eingesetzt werden. Dabei wird überprüft, inwieweit der oben zitierte methodische Diskurs durch den Status quo der empirischen Praxis gespiegelt wird. In der Auswertung unserer Übersicht sollen insbesondere Dissonanzen zwischen Methodendiskurs und Praxis dargelegt und gedeutet werden. Unsere Studie ist somit eng angelehnt an einen Literatursurvey, den Halaby (2004) vorgelegt hat. Wir beschränken uns dabei auf zwei zentrale Publikationsorgane der deutschsprachigen Soziologie: die *Kölner Zeitschrift für Soziologie und Sozialpsychologie* (KZfSS) sowie die *Zeitschrift für Soziologie* (ZfS). Im Rahmen dieses abgegrenzten Bezugsraums (und des Zeitfensters von 2000 bis 2009) führen wir eine Vollerhebung von

* Wir danken Henning Lohmann und Johannes Huinink für ihre kenntnisreichen Kommentare zum Manuskript.

¹ Weitere Methoden für kausale Analysen werden z.B. bei Winship und Morgan (2007) diskutiert.

personenbasierten Panelstudien durch – anders als Halaby, der Theorie/Praxis-Inkonsistenzen auf der Grundlage ausgewählter Beispiele verdeutlicht.

2 Methoden zur Analyse von Paneldaten

Innerhalb der soziologischen Methodenforschung hat sich eine Norm für die Analyse personenbezogener Paneldaten ausgebildet, nämlich die Verwendung eines *within*- bzw. *Fixed Effects*-Schätzers (Halaby 2004; Allison 2009). Diese Norm lässt sich gut begründen, und zwar durch den Verweis auf das zentrale statistische Potential von Paneldaten. Dieses liegt in der vollständigen Kontrolle zeitkonstanter Störvariablen. Das *Fixed Effects*-Verfahren (FE) realisiert dieses Potential, indem es ausschließlich auf intraindividuelle Zusammenhänge zur Konstruktion der Schätzer abstellt, also einen „within“-Vergleich im Rahmen des Regressionsverfahrens durchführt. So werden Niveauunterschiede zwischen personenspezifischen Messreihen vollständig ausgeblendet und damit schließlich kausale Interpretationen gegenüber dem Einfluss zeitkonstanter Merkmale abgesichert (Brüderl 2010; Giesselmann und Windzio 2012).²

Dieses Argument lässt sich in einen erweiterten Methodendiskurs eintakten, welcher an die Konsistenz zwischen der Mechanik eines Verfahrens und der Form der Fragestellung appelliert (King 1990). Sozialwissenschaftliche Hypothesen beziehen sich zumeist auf die Effektwirkung einer Variable und sind daher in einer längsschnittlichen Semantik angelegt: Sie treffen Aussagen darüber, wie sich die abhängige Variable verändert, wenn sich die unabhängige Variable von einem Ausgangswert entfernt. Diese Semantik wird in der technischen Umsetzung einer *within*-Analyse von Paneldaten – zumindest näherungsweise – aufgegriffen (Giesselmann und Windzio 2012). Zudem stellen nicht nur die Hypothesen, sondern auch die Auswertungen quantitativer Analysen zumeist auf kausale Zusammenhänge ab: Wie Sobel (2000) feststellt, haben viele Querschnittsmethoden (wie Standardregression und Pfadanalyse) die Kausalitätsattribution fest in ihrer Interpretationslogik verankert, ohne jedoch diese Proklamation auf statistischer Ebene einzulösen. Die Effektinterpretationen sind daher *extramathematisch*, spiegeln also nicht die Logik der zugrundeliegenden Verfahren wider. In diesem Sinne stärkt der FE-Schätzer die Bindung zwischen Fragestellung, Analysetechnik (bzw. *Mathematik*) und Interpretation, da er den Längsschnittcharakter von Hypothesen und Auswertungen in seiner Mechanik realisiert.

Im Folgenden wird die konkrete Umsetzung des FE-Schätzers kurz beschrieben sowie sein hier herausgestellter Vorteil formal-statistisch begründet; zunächst im Kontext der Panelregression, danach für Ereignisanalysen.

2.1 Die *Fixed Effects Panelregression*

Bei der *Fixed Effects Regression* wird die kausale Analyselogik durch die Subtraktion personenspezifischer Mittelwerte umgesetzt. Mit einer einfachen Gleichung als Ausgangspunkt lässt sich das Verfahren formal wie folgt beschreiben:

$$(1) y_{it} = b_0 + b_1 * x_{it} + w_{it}, w_{it} = e_{it} + u_i$$

² Auch in FE-Analysen mit Paneldaten verbleiben allerdings einige Probleme kausaler Schlüsse, siehe z.B. Allison (2009: 93) und Andreß et al. (2013: 109).

b_1 kennzeichnet den Koeffizienten, der den Effekt von x auf y bestimmt.³ Durch den Index i werden Personen gekennzeichnet, t bezeichnet Zeit- bzw. Messpunkte. Der Fehlerterm w besteht aus zwei Komponenten: Der Personeneffekt u repräsentiert den Einfluss aller zeitkonstanten unbeobachteten Merkmale, e bezeichnet den idiosynkratischen, zeitabhängigen Fehler. Diese Dichotomie des Fehlerterms besteht grundsätzlich *immer*, auch im Querschnittskontext, und ist kein Spezifikum von Paneldaten. Lediglich Längsschnittdaten wohnt jedoch das Potential inne, die Größe der u_i abzubilden bzw. u systematisch zu eliminieren und damit den Einfluss aller zeitkonstanten unbeobachteten Merkmale zu kontrollieren. Letzteres leistet die FE-Subtraktionsprozedur, welche sich wie folgt als Erweiterung von (1) darstellen lässt (\bar{y}_i und \bar{x}_i bezeichnen dabei personenspezifische Mittelwerte):

$$(2) (y_{it} - \bar{y}_i) = b_1 * (x_{it} - \bar{x}_i) + w_{it}, \quad w_{it} = (e_{it} + u_i) - (\bar{e}_i + u_i),$$

Eine einfache Umformung der Gleichung (2) zeigt, dass der Fehlerterm nun dem idiosynkratischen Fehler gleicht und somit der Einfluss sämtlicher zeitkonstanter Störvariablen ausgegliedert ist:

$$(3) (e_{it} + u_i) - (\bar{e}_i + u_i) = e_{it} - \bar{e}_i + u_i - u_i = (e_{it} - 0) = e_{it}$$

Da der Einheiten effekt u zeitkonstant ist, entsprechen alle seine Ausprägungen dem jeweiligen einheitenspezifischen Mittelwert. Folglich fallen sämtliche u_i (sowie die Konstante) aus der Gleichung heraus. An dieser Eigenschaft lässt sich das Potential des within-Vergleichs (bzw. des FE-Schätzers) zur Kausalitätsannäherung nun statistisch-formal festmachen, denn mit der Elimination der u_i sind die Effekte aller zeitkonstanten Drittvariablen abgekoppelt: Die Bedingung $corr(u,x)=0$, die bei einfachen Regressionsanalysen zur Rechtfertigung der Unverzerrtheit der berechneten Effektschätzer notwendig ist, kann bei Durchführung einer *Fixed Effects*-Panelregression außer Acht gelassen werden.⁴

2.2 Der within-Schätzer der Ereignisanalyse

Das Äquivalent zur *Fixed Effects*-Panelregression ist der *within*-Schätzer der Ereignisanalyse.⁵ Die abhängige Variable in der Ereignisanalyse ist die *Übergangsrate*, die für jede der im Regressionsmodell untersuchten Subpopulationen die Anzahl der Ereignisse pro Zeiteinheit darstellt (Windzio 2013).

Ist die Zeit stetig gemessen, beschreibt die Übergangsrate die *augenblickliche* Neigung einer Untersuchungseinheit zu einem Übergang unter der Bedingung, dass bis zum aktuellen Zeitpunkt noch *kein* Ereignis eingetreten ist. Somit wird diese augenblickliche Neigung auf die jeweils aktuelle Risikopopulation konditioniert, was die folgende Gleichung zum Ausdruck bringt:

³ Alle hier dargestellten Zusammenhänge gelten auch für Modelle mit mehreren unabhängigen Variablen.

⁴ Zu Varianten der FE-Prozedur siehe Giesselmann und Windzio (2012).

⁵ Auch Ereignisanalysen basieren auf dem Regressionsprinzip, das hier allerdings explizit zur Modellierung von *Übergangsraten* (statt *zeitpunktspezifischen Zuständen*, wie bei der Panelregression) eingesetzt wird.

$$(4) \quad r(t) = \lim_{\Delta t \rightarrow 0} \frac{P_{i,m} \{t \leq T' < t + \Delta t | T \geq t\}}{\Delta t}$$

Man kann die Rate durch eine bedingte Wahrscheinlichkeit darstellen, in einem kleinen Zeitintervall ($\lim \Delta t \rightarrow 0$) ein Ereignis zu erleben, die durch die Länge des Intervalls dividiert wird. *Bedingt* ist die Wahrscheinlichkeit darauf, dass bis zum Beginn des kleinen Zeitintervalls noch kein Ereignis eingetreten ist ($| T \geq t$). Daraus folgt für eine Untersuchungseinheit, dass die durch die unabhängigen Variablen bestimmte Situation zu t auf die Neigung zu einem Ereignis zu $t + \Delta t$ wirkt, und damit ein *time-lag* zwischen Ursache und Wirkung angenommen wird. Im Rahmen ihres dynamischen Kausalansatzes heben Blossfeld et al. (1996: 34) darum hervor, dass Kausalität dadurch definiert sein sollte, dass „(...) eine zeitlich fixierbare Veränderung im Prozess Y_t^A zu einer zeitlich später fixierbaren Veränderung im Prozess $Y_{t'}^B$ führt“, was sich ausdrücken lässt durch:

$$(5) \quad \Delta Y_t^A \rightarrow \Delta Y_{t'}^B \quad t < t'$$

Aufgrund der Möglichkeiten der Modellierung von Effekten zeitveränderlicher Kovariaten betrachten die AutorInnen die Ereignisanalyse als ein angemessenes Verfahren zur Modellierung kausaler Beziehungen in nicht-experimentellen Untersuchungsdesigns. Die Einbeziehung zeitveränderlicher Kovariaten, die innerhalb einer Episode ihre Werte ändern, kann dabei über den datentechnischen Trick des *Episodensplittings* erreicht werden. Dabei wird eine Episode zum Zeitpunkt der Zustandsänderung in einem exogenen Prozess in zwei Teile zerlegt und die Zustandsänderung z.B. durch eine Dummyvariable gemessen. In der folgenden Tabelle 1 wird der Einfluss einer Heirat auf die Jobmobilität modelliert (Windzio 2013). Die Zeitangaben sind als so genannte *Jahrhundertmonate* dargestellt, also in Monaten, die seit Januar des Jahres 1900 vergangen sind. ID 23 begann eine Arbeitsstelle zum Zeitpunkt 1243 und beendete diese zum Zeitpunkt 1260. Sie hat aber währenddessen geheiratet, und zwar zum Zeitpunkt 1250. ID 67 hingegen hat während der aktuellen Arbeitsstelle nicht geheiratet. Im unteren Teil der Tabelle wurde das Episodensplitting vollzogen, die Episode der ID 23 wurde nun in zwei Subepisoden zerlegt, von denen die erste rechtszensiert ist, da zum Zeitpunkt des Splittings das zu erklärende Ereignis (*Ende des Jobs*) noch nicht eingetreten ist.

<Tabelle 1 hier einfügen>

Für unsere Fragestellung ist es nun wichtig, dass bei der Episode der ID 67 *kein* Episodensplitting durchgeführt wurde. Die potenziell zeitabhängige Variable „Ehe“ ist bei dieser Person *zeitkonstant* und liefert somit auf der Seite der unabhängigen Variablen nur eine *Querschnittsinformation*. Aufgrund unbeobachteter Faktoren bleibt diese Person zumindest für die Dauer der Episode unverheiratet. Der Effekt der Variable „Ehe“ auf die Übergangsrate hat daher zwei Komponenten: erstens eine *within*-Komponente, die sich aus jenen Personen ergibt, die im Verlaufe der Episode eine Veränderung des Zustandes in x erleben (d.h. sie heiraten), und zweitens eine *between*-Komponente aufgrund jener Personen, die während einer Episode entweder dauerhaft verheiratet oder unverheiratet blieben. Anders als im Falle der ID 23 liefert das Episodensplitting für die Person ID 67 also keinen *within*-Schätzer, sondern einen *between*-Schätzer. Aufgrund der Tatsache, dass viele potenziell zeitveränderliche Kovariaten sich nur bei einer *Teilmenge* der Untersuchungseinheiten im Verlaufe einer Episode ändern, ist der Effekt dieser Kovariaten eine Mischung aus *within*-

und *between*-Effekt. In der Lebenslaufforschung wird das Verfahren des Episodensplittings häufig wie in der obigen Tabelle durchführt (Blossfeld et al. 2007: 141) (vgl. die Beispiele unten). Eine kausale Modellierung im Sinne des dynamischen Kausalansatzes, demzufolge zeitlich fixierbare Veränderungen im Prozess Y_t^A zu einer zeitlich später fixierbaren Veränderung im Prozess Y_t^B führen (Blossfeld et al. 1996: 34), wird durch die Berücksichtigung der *zeitkonstanten* Einflüsse der erklärenden Variablen allerdings unterlaufen. Nehmen wir an, in einer Studie sind nur sehr wenige Zustandswechsel in x (*Heirat*) feststellbar. Dann ergibt sich der Schätzer des Effektes der Heirat auf die Rate überwiegend durch die zeitkonstanten Effekte des (*un-*)*verheiratet Seins* – was jedoch nicht im Sinne des dynamischen Kausalansatzes ist, weil dieser zeitkonstante Zustand wiederum aus anderen unbeobachteten Faktoren resultieren kann!

Um die Potentiale der kausalen Modellierung auszuschöpfen sollte man die logischen Möglichkeiten der Stabilität der x -Variable (z.B. stabiler Zustand X_1 oder X_2) sowie die jeweiligen Veränderungen von x (z.B. von X_1 nach X_2 bzw. X_2 nach X_1) durch entsprechende Kovariaten im Modell berücksichtigen. Man könnte in Analogie zum *Fixed Effects*-Modell der Panelregression aber auch nur jene Episoden berücksichtigen, in die tatsächlich ein exogenes intervenierendes Ereignis eingelagert ist. Ein (*within*-)Schätzer ergäbe sich hier durch die Veränderung in x innerhalb der Episode. Die Isolierung solcher Spells führt dann zur Möglichkeit der systematischen Trennung von *within*- und *between*-Effekt im Rahmen der Ereignisanalyse.⁶

2.3 Querschnittsverfahren zur Analyse von Paneldaten

Verfahren zur Analyse von Paneldaten, die nicht ausschließlich auf intraindividuelle Kovariation zur Konstruktion von Koeffizienten abstellen, wie beispielsweise die *Random Effects* (RE) *Regression*, werden im methodischen Diskurs oft kritisch diskutiert (Halaby 2004; Brüderl 2010). Schließlich setzten diese Verfahren die (problembehaftete) Annahme voraus, dass die untersuchten Zusammenhänge unabhängig von unbeobachtbaren zeitkonstanten Drittvariablen sind.

Allerdings gibt es Analysesituationen, in denen auch bei personenspezifischen Paneldaten Verfahren angezeigt sind, die auf Unterschiede im Querschnitt abstellen. Dies ist dann der Fall, wenn eine zeitkonstante unabhängige Variable mit Paneldaten untersucht wird, also explizit Unterschiede *zwischen* Personen fokussiert werden. Die Legitimität dieser Konstellation von Daten- und Variablentyp – und damit letztlich auch von Standardregression oder RE im Panelkontext – beruht auf der Anerkennung alternativer Motive zur Verwendung von Paneldaten, *neben* der statistischen Annäherung an Kausalität.

So liegt ein weiterer zentraler analytischer Mehrwert von Panelerhebungen darin, dass sie die gemeinsame Modellierung von zeitpunkt- und lebensphasenübergreifenden Merkmalen erlauben. Durch das prospektive Erhebungsformat können dabei auch solche Merkmale untersucht werden, die in retrospektiven Erhebungen mit starken Messfehlern behaftet sind. Mit Paneldaten kann beispielsweise der Einfluss komplexer Sozialisationsbedingungen auf spätere Bildungs- oder Arbeitsmarktereignisse modelliert werden (Lohmann et al. 2009). Die erklärenden Faktoren (Lebensbedingungen zum Zeitpunkt t) fungieren im Rahmen solcher Designs als zeitkonstante Merkmale, die dementsprechend Auswertungsverfahren zur Modellierung von Querschnittsunterschieden bedingen. Zudem mag in bestimmten

⁶ Zum *within*-Schätzer für Fälle mit mehreren Episoden pro Subjekt siehe Allison (2009) sowie Brüderl (2010).

Situationen die Operationalisierung von unabhängiger oder abhängiger Variable auf Längsschnittinformation abstellen, obgleich in der Analyse explizit Unterschiede zwischen verschiedenen Merkmalsträgern untersucht werden (z. B. Giesselmann 2009). Das Motiv zur Verwendung von Paneldaten liegt in diesen Fällen also in der valideren Umsetzbarkeit der Konzepte. Diese Eigenschaft ist auch im Rahmen ereignisanalytischer Designs relevant, wenn der Einfluss eines zeitkonstanten unabhängigen Merkmals (zum Beispiel *Geschlecht* oder *Kohortenzugehörigkeit*) auf die Dauer bis zum Eintritt eines bestimmten Zustandswechsels untersucht wird, Zeiträume (bzw. *Verweildauern*) also exakt gemessen werden sollen.

Wenn dementsprechend im Rahmen der Analyse ein zeitkonstantes unabhängiges Merkmal untersucht wird, ist die Verwendung eines FE-Schätzers nicht sinnvoll: Bei *echten zeitkonstanten unabhängigen Merkmalen* führt dies zum Ausfall der untersuchten Variablen. Bei *quasi-zeitkonstanten Merkmalen* (die konzeptionell als zeitinvariant angelegt sind, aber marginale intraindividuelle Variation aufweisen) wird ausschließlich auf die marginale intraindividuelle Varianz abgestellt und ggf. ein irreführender Schätzer konstruiert (Giesselmann und Windzio 2012). Stattdessen ist ein Querschnittsverfahren angezeigt, also die Standardregression oder – falls mehrere (statistisch nicht unabhängige) Messungen pro Person in die Analyse eingehen – *Random Effects Regressionen* bzw. Verfahren mit Standardfehlerkorrektur.

Mit Blick auf die verschiedenen legitimen Motive zur Verwendung von Paneldaten (und die unterschiedlichen Frageformate und Variablentypen, die mit diesen Motiven zusammenhängen) können also Analyseverfahren im personenbezogenen Panelkontext nicht pauschal als *richtig* oder *falsch* eingeordnet werden. Allerdings lassen sich Muster von Motiv, Variablentyp und Analyseverfahren als *konsistent* oder *nicht konsistent* identifizieren. In einer Analyse mehrerer panelbasierter Aufsätze aus ZfS und KZfSS soll im Folgenden die Verbreitung konsistenter Muster überprüft werden. In Anlehnung an ein ähnliches Projekt von Halaby (2004) werden dabei Methode, Motiv und – zusätzlich – Variablentyp systematisch tabellarisch aufbereitet. Diese Aufbereitung soll folgende Einschätzungen ermöglichen: a.) *Welche Motive zur Verwendung von Paneldaten dominieren in der empirischen Praxis*, b.) *Inwieweit wird das im Methodendiskurs als zentral herausgearbeitete Potential von Paneldaten in der empirischen Praxis eingesetzt bzw. ausgeschöpft?* und c.) *Inwiefern werden Daten und Methoden im Panelkontext entsprechend der Forderung nach Konsistenz zwischen Fragestellung, Motiv, Variablentyp und Analyseverfahren eingesetzt?*

3 Panelstudien in der soziologischen Praxis

Als *Panelstudie* bezeichnen wir empirische Designs, bei denen Informationen aus mindestens zwei Wellen einer Panelerhebung verwendet werden. Dabei beschränken wir uns auf solche Studien, bei denen Personen die Einheiten der Analyse bilden. Unsere Vollerhebung so definierter Panelstudien umfasst 62 Artikel, die im Zeitraum zwischen 2000 und 2009 in der *Kölner Zeitschrift für Soziologie und Sozialpsychologie* (KZfSS) oder der *Zeitschrift für Soziologie* (ZfS) veröffentlicht wurden.⁷ 35 Studien stammen aus der KZfSS, 27 aus der ZfS. Studien mit Panelregressionen (34) haben ein leichtes Übergewicht gegenüber ereignisanalytischen Designs (28), wobei in 3 Studien beide Ansätze angewendet werden. Die Auswertung der Studien erfolgt im Folgenden getrennt nach Ereignisanalysen und Panelregressionen.

⁷ Alle Informationen zu den Studien haben wir nach bestem Wissen und Gewissen aufbereitet. Sollten AutorInnen ihr Design nicht angemessen repräsentiert vorfinden, bitten wir um Entschuldigung und Mitteilung. Eine Referenzliste zu allen Artikeln befindet sich im Anhang des Artikels.

3.1 Praxis der Panelregression

Um zu ermitteln, wie Panelregressionen in das Design der Studie eingebettet sind, unterscheiden wir zwischen vier Typen von unabhängigen Variablen: *zeitkonstanten*, *quasi-zeitkonstanten*, *zeitveränderlichen* sowie, als Sonderfall der zeitveränderlichen Variablen, *Trendvariablen*. Als quasi-zeitkonstante Variablen gelten solche Merkmale, die zwar konzeptionell als zeitinvariante Merkmale angelegt sind, aber trotzdem marginale Variation auf individueller Ebene aufweisen. Neben dem Variablentyp unterscheiden wir zwei verschiedene Typen von Schätzern: Schätzer, die ausschließlich intraindividuelle Kovariation berücksichtigen (FE) sowie Schätzer, welche anteilig Querschnittsvariation aufnehmen (RE und Standardregression).

[Tabelle 2 hier einfügen]

Des Weiteren differenzieren wir die Studien nach vier Motiven zur Verwendung von Paneldaten. Dabei berücksichtigen wir ausschließlich Motive, die von den AutorInnen angeführt werden. Sämtliche Erwähnungen, die auf das Potential von Paneldaten hindeuten, Kausalität durch Kontrolle zeitkonstanter Drittvariablen anzunähern (bzw. *personenspezifische Heterogenität* zu kontrollieren), fassen wir in der Kategorie *Annäherung Kausalität* zusammen. Tabelle 2 zeigt, dass dieses Motiv lediglich bei 9 der 34 Studien im Regressionsdesign genannt wird. Eine weitere Studie stellt in den Hypothesen auf zeitveränderliche unabhängige Variablen ab, ohne dabei aber explizit das Motiv der Kausalitätsannäherung zu bemühen. Somit bieten sich also 10 Fragestellungen für eine empirische Längsschnittmodellierung an. Bei 7 dieser 10 Studien wird tatsächlich der *within*-Effekt über eine *Fixed Effects*-Modellierung freigelegt und damit das zentrale Potential von Paneldaten zur Annäherung von Kausalität realisiert. Folglich weisen 3 Studien ein inkonsistentes Muster auf: Im Widerspruch zum methodologischen Standard wird hier auch Querschnittsvariation zur Konstruktion der Koeffizienten genutzt.

Zwei dieser Studien (Pollmann-Schult 2008, Isengard 2005) legitimieren die Verwendung des RE-Schätzers durch Verweis auf die Ähnlichkeit zu FE-Resultaten. Ein systematischer Test auf Koeffizientenunterschiede zwischen FE und RE ist allerdings in beiden Studien nicht dokumentiert. Zudem wird bei Pollmann-Schult RE ausdrücklich als Verfahren eingeführt, mit welchem "*für Selektionsprozesse kontrolliert werden kann und somit eine Kausalanalyse erst ermöglicht wird*" (Pollmann-Schult 2008: 505). Damit ist, im Lichte des methodologischen Diskurses und der statistischen Eigenschaften von RE, das Potential dieses Verfahrens überschätzt.

Eine ähnliche Legitimierung von RE bemüht Strauß (2009: 656). RE wird hier als Methode beschrieben, welche "*durch die Aufspaltung des Fehlerterms in einen personenspezifischen und einen idiosynkratischen Teil die zeitkonstante unbeobachtete Heterogenität kontrolliert*". Obgleich RE tatsächlich die Dichotomie des Fehlerterms modelliert, wird der Personeneffekt dabei nicht systematisch spezifiziert, was wiederum Grundlage der – somit fälschlich behaupteten – Kontrolle zeitkonstanter unbeobachteter Heterogenität wäre: Die Varianz der Personeneffekte wird bei RE zwar berücksichtigt, die Einflussfaktoren der Heterogenitäten dadurch jedoch nicht statistisch kontrolliert. Gleichzeitig verwirft die Autorin FE-Modelle mit der Begründung der fehlenden Analysierbarkeit zeitkonstanter unabhängiger Variablen. Nach unserer Auffassung ist diese Bewertung der Variable *ehrenamtliches Engagement* zumindest fragwürdig und sollte weiter ausgeführt werden – zumal in verschiedenen Teilen der Studie eine Längsschnittterminologie bei der Interpretation der Koeffizienten angewendet wird. Schließlich ersetzt ein Vergleich zwischen Standardregression und RE einen systematischen Test auf Koeffizientenunterschiede zwischen FE und RE. Damit werden also zwei Analysen,

die beide auf der Annahme *keiner* zeitkonstanter Störvariablen beruhen, gegeneinander geprüft.

Bei Isengard (2005) wird zwar nicht ausdrücklich auf die kausalitätsabsichernden Potentiale von Paneldaten verwiesen, aber trotzdem ein analytischer Rahmen gewählt und beschrieben, welcher an die Verwendung eines *within*-Schätzers appelliert: Einige der als zentral eingeführten unabhängigen Variablen, z.B. *Einkommen*, sind klar zeitabhängig und bieten sich somit für eine *Fixed Effects*-Modellierung an. Im Lichte des ausgestellten Anspruchs der Studie, Alternativerklärungen auszuschließen und "*unverzerrte Schätzungen zu erhalten*" (Isengard 2005: 264) ist die Auswahl des Analyseverfahrens also diskussionswürdig.

Der überwiegende Anteil an Studien, nämlich 24, nutzt Paneldaten nicht im Zusammenhang mit zeitveränderlichen unabhängigen Variablen. Die meisten dieser Studien (16) verwenden Paneldaten, um *zeitversetzt erhobene Variablen* gemeinsam zu modellieren. Wesentlich häufiger als es der methodologische Diskurs um die Potentiale von Paneldaten nahelegt werden Paneldaten in den untersuchten Studien also *nicht* eingesetzt, um Kausalitätsinterpretationen zeitveränderlicher Variablen zu verteidigen, sondern um ein anderes spezifisches Potential von Paneldaten zu nutzen: Den zeitpunktübergreifenden Fokus der Daten auf individueller Ebene, der den simultanen Zugriff auf verschiedene Punkte im Lebenslauf ermöglicht. Sämtliche Studien, welche entsprechende Motive ausflaggen, weisen ein konsistentes Muster von Motiv, Variablentyp und Methode auf.

Bei 6 der 24 Studien mit zeitkonstanten UV wird die Verwendung von Paneldaten durch die Möglichkeit zur Durchführung eines Trenddesigns motiviert, bei zwei weiteren dienen Paneldaten lediglich einer Vergrößerung der Stichprobe. Somit verwenden ein Viertel der untersuchten Studien Paneldaten, ohne dabei auf genuine Eigenschaften des Panelformates zuzugreifen.

3.2 Praxis der Ereignisanalyse

In den Zeitschriften ZfS und KZfSS zählen wir für die Jahre 2000-2009 28 Aufsätze, in denen Ereignisanalysen geschätzt wurden. Tabelle 3 zeigt, dass in vier dieser Studien tatsächlich über die Bildung der erklärenden Variablen *between*- und *within*-Effekte konsequent getrennt werden. Somit wird in der überwiegenden Zahl der publizierten Analysen der dynamische Kausalansatz aufgrund der Vermengung von *between*- und *within*-Effekt letztlich nicht umgesetzt – auch, wenn das Episodensplitting angewandt wird. Zeitkonstante und zeitveränderliche Beiträge von x werden also oftmals nicht getrennt und Kausalität damit lediglich an Sequentialität festgemacht. Ein zentrales Potential von Paneldaten zur Absicherung der Kausalitätsattribution – nämlich die Kontrolle zeitkonstanter Störvariablen – wird nicht genutzt.

[Tabelle 3 hier einfügen]

Auffällig ist, dass in einigen Fällen explizit auf den dynamischen Kausalansatz verwiesen wurde, mit dem das Problem der kausalen Modellierung gelöst werden könne. Dabei wird jedoch allein auf die zeitliche Anordnung von Ursache und Wirkung verwiesen bzw. in Anlehnung an Blossfeld et al. (1996) auf die zeitliche Lagerung der Veränderungen in x und y (Brose 2008: 37). Nicht diskutiert wird das Problem der Trennung zwischen *between*- und *within*-Effekten. In vier Studien wurden *between*- und *within*-Effekte für manche Variablen getrennt, für manche hingegen nicht. Explizit wird das Problem von Schröder und Brüderl (2008) angesprochen und angemessen in der empirischen Umsetzung behandelt. Dies geschieht dadurch, dass neben den konstanten Zuständen auch die jeweiligen Veränderungen im Status der erklärenden Variablen explizit modelliert wurden.

4 Diskussion

Trotz eines Methodendiskurses, der das Potential von Paneldaten zur Absicherung von Effektinterpretationen betont und entsprechende Analysemethoden ausführlich verhandelt hat, finden sich in unserer Überblicksstudie mehrere Analysen, die trotz klarem Kausalitätsmotiv auf problematische Querschnittsschätzer abstellen. Hierbei fällt auf, dass die betreffenden RE-Panelregressionen allesamt über die Modellierbarkeit zeitkonstanter Variablen gerechtfertigt werden (Isengard 2005: 265; Pollmann-Schult 2006: 504; Strauß 2009: 656). Dieses Argument ist aus zwei Gründen problematisch.

Erstens werden in den entsprechenden Analysen zeitveränderliche Faktoren modelliert und auch in Titel sowie Hypothesen als zentral ausgestellt. Zeitkonstante Variablen sind dabei teils ausdrücklich als Kontrollvariablen eingeführt. Gerade die perfekte Kontrolle zeitkonstanter Merkmale wird aber im Rahmen einer *within*-Regression geleistet – der Ausfall zeitkonstanter Variablen stellt somit keinen Verlust dar, sondern indiziert vielmehr das zentrale, auch explizit motivierte Potential von Paneldaten zur Kausalitätsannäherung. Zudem lassen sich im FE Modell *Interaktionen* zeitkonstanter und zeitveränderlicher Variablen problemlos spezifizieren (und so die *Mechanismen* aufklären, die zu Gruppenunterschieden führen können). *Zweitens* ermöglichen sog. *hybride Regressionen* die gemeinsame Spezifikation von *within*- und *between*-Effekt. Solche Modelle erlauben also die Modellierung zeitkonstanter Variablen bei gleichzeitiger Spezifikation des *within*-Schätzers zeitveränderlicher Variablen (Allison 2009; Giesselmann und Windzio 2012; Andreß et al. 2013; Schunck 2013). Besteht demnach im spezifischen Analysefall nicht nur ein Interesse an den Koeffizienten zeitveränderlicher Variablen, sondern ein zusätzliches substantielles Interesse an den Effekten oder Mediationseigenschaften zeitkonstanter Merkmale, so wäre die Verwendung der Hybrid-Regression angezeigt. Offensichtlich aber ist diese Modelloption in der Praxis noch nicht hinreichend etabliert – keine der hier zitierten Studien wendet dieses Verfahren an.

Grundsätzlich können wir jedoch feststellen, dass das Potential von Paneldaten zur Absicherung kausaler Interpretationen im Rahmen von Panelregressionen – sofern entsprechende Variablen untersucht und Motive ausgeflaggt werden – zumeist (konkret: in 70% der betreffenden Studien) realisiert wird. Im Rahmen ereignisanalytischer Modellierungen ist dies allerdings nicht der Fall. Hier werden selbst bei zeitveränderlichen unabhängigen Variablen mehrheitlich analytische Designs gewählt, die zwar die Sequentialität einer Kausalbeziehung abbilden, jedoch nicht unbeobachtete personenspezifische Heterogenität kontrollieren. Besonders kennzeichnend für diesen Unterschied zwischen den beiden Analysetypen *Panelregression* und *Ereignisanalyse* sind diejenigen Studien, welche beide analytischen Zugänge nutzen (mit denen sich quasi die *Fixed Effects*-Logik innerhalb unseres Literatursurveys umsetzen lässt): In den Artikeln von Pollmann-Schult und Diewald (2007), Grunow et al. (2007) sowie Kalter (2006) wird jeweils explizit das Motiv der Kausalitätsannäherung genannt und im Rahmen der Panelregression auch stringent umgesetzt. Die flankierenden Ereignisanalysen jedoch legen in allen drei Studien nicht den *within*-Vergleich frei, nutzen also nicht das volle Potential der Daten zur Absicherung kausaler Interpretationen aus. Offenbar hat sich im Rahmen des ereignisanalytischen Frameworks also bisher kein methodischer Standard in der Praxis etabliert, der auf beide Aspekte des Potentials von Längsschnittdaten zur Kausalitätsabsicherung (Sequentialität *und* Heterogenitätskontrolle) abstellt. Auch

diagnostizieren wir vorsichtig eine fehlende Anerkennung des Unterschieds zwischen *within*- und *between*-Schätzern im Kontext von Ereignisanalysen.⁸

Zudem fällt mit Blick auf die analytischen Muster sämtlicher ausgewerteter Studien auf, dass Methodendiskurs und empirische Praxis jeweils unterschiedliche Potentiale von Paneldaten in den Vordergrund rücken. Das in vielen methodischen Publikationen als entscheidend herausgearbeitete Potential von Paneldaten, nämlich die *Kontrolle von Heterogenität* ist *nicht* das zentrale Motiv zur Verwendung von Paneldaten in der empirischen soziologischen Praxis. Diese Diskrepanz zwischen methodischen Standards und Empirie kann auf verschiedene Weisen interpretiert werden. Sie mag als Hinweis darauf gedeutet werden, dass das zentrale Analysepotential von Paneldaten in der Soziologie nicht ausgeschöpft wird (und entsprechend auf einen *underusage* von Paneldaten verweisen). In dieser Perspektive trifft der Aufwand, der grade in Deutschland betrieben wird, um Daten im Längsschnittformat bereitzustellen, nicht auf eine entsprechende Analysebereitschaft seitens der Forschungsgemeinschaft. Diese Sichtweise mündet in einem Appell an die Disziplin, stärker die genuinen Stärken von Paneldaten zur Verteidigung kausaler theoretischer Modelle zu nutzen. Auch sendet unser Literatursurvey die Botschaft an die Praxis, bereits publizierte Studienergebnisse intensiv mit dem leistungsfähigen Instrumentarium, welches Paneldaten anbieten, zu überprüfen und herauszufordern.

Aus einer anderen Sichtweise auf das Phänomen ergibt sich jedoch eine Aufforderung an die soziologische Methodenforschung. So mag es grade für die Soziologie spezifisch sein, dass sie Einflüsse zeitkonstanter Konstrukte (*Herkunft, Bildung, Sozialisation*) zu bestimmen sucht. Wenn aber solche Heterogenitäten *systematisch modelliert* werden, entfalten sich die Potentiale von Paneldaten zwangsläufig entlang anderer Eigenschaften als der Kontrolle (bzw. Elimination) von Heterogenität. Vielmehr motiviert in solchen Fällen die Möglichkeit der gemeinsamen Modellierung zeitpunktübergreifend erhobener Merkmale die Verwendung von Paneldaten. So werden in den von uns betrachteten Studien vielfach Arbeitsmarkt- und Bildungsergebnisse als Konsequenz früherer Sozialisationsbedingungen modelliert (z.B. Kratzmann und Schneider 2009). Da solche Bedingungen als konstante personenspezifische Merkmale behandelt werden, sind die Hypothesen dieser Studien nicht in einer Längsschnitterminologie angelegt, sondern stellen explizit auf Unterschiede zwischen Personengruppen ab. Angesichts der augenscheinlichen Praxisrelevanz solcher *lebensverlaufsbezogener Querschnittshypothesen* wäre es eine verengte Perspektive, würde man die Potentiale von Paneldaten ausschließlich bei zeitveränderlichen Merkmalen verorten und kategorisch die Freilegung von Längsschnittinformation durch FE einfordern. Somit ergibt sich aus dieser Perspektive ein Appell an die Methodenforschung, stärker die unterschiedlichen Motive der Verwendung von Paneldaten anzuerkennen und in konstruktiver Weise die Probleme der Modellierung auch solcher unabhängiger Merkmale zu beleuchten, die konzeptionell als zeitkonstant angelegt sind.

⁸ Zur Modellierung von Effekten auf *Formparameter* siehe Windzio 2013.

Literatur

- Allison, Paul D. 2009. *Fixed effects regression models*. Los Angeles: Sage.^
- Andreß, Hans-Jürgen, Katrin Golsch, und Alexander Schmidt, 2013, *Applied panel data analysis for economic and social surveys*. London: Springer.
- Blossfeld, Hans-Peter, Karin Golsch, und Götz Rohwer. 2007. *Event history analysis with Stata*. Mahwah: Erlbaum.
- Blossfeld, Hans-Peter, Erik Klijzing, Katharina Pohl, und Götz Rohwer. 1996. Die Modellierung paralleler interdependenter Prozesse in der Bevölkerungswissenschaft: Konzepte und Methoden am Beispiel der Heiratsneigung nichtehelicher Lebensgemeinschaften bei Geburt des ersten Kindes. *Zeitschrift für Bevölkerungswissenschaft* 22: 29–56.
- Blossfeld, Hans-Peter, Hans-Günther Roßbach, und Jutta von Maurice, Hrsg. 2011. *Education as a lifelong process. The German National Educational Panel Study (NEPS)*. Wiesbaden: VS.
- Börsch-Supan, Axel, und Hendrik Jürges, Hrsg. 2005. *The survey of health, aging, and retirement in Europe*. Mannheim: MEA.
- Brose, Nicole. 2008. Entscheidung unter Unsicherheit - Familiengründung und -erweiterung im Erwerbsverlauf. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 60: 30-52.
- Brüderl, Josef. 2010. Kausalanalyse mit Paneldaten. In *Handbuch der sozialwissenschaftlichen Datenanalyse*, Hrsg. Christof Wolf und Henning Best, 963-994. Wiesbaden: VS.
- Eurostat. 2005. *EU-SILC Database Description*. European Commission: Luxembourg.
- Gangl, Markus. 2010. Causal Inference in Sociological Research. *Annual Review of Sociology* 36: 21-47.
- Giesselmann, Marco. 2009. Arbeitsmarktpolitischer Wandel in Deutschland seit 1991 und das Working Poor - Problem: Einsteiger als Verlierer des Reformprozesses? *Zeitschrift für Soziologie* 38: 215-238.
- Giesselmann, Marco, und Michael Windzio. 2012. *Regressionsmodelle zur Analyse von Paneldaten*. Wiesbaden: Springer VS.
- Grunow, Daniela, Florian Schulz und Hans-Peter Blossfeld. 2007. Was erklärt die Traditionalisierungsprozesse häuslicher Arbeitsteilung im Eheverlauf: soziale Normen oder ökonomische Ressourcen? *Zeitschrift für Soziologie* 36: 162-181.
- Halaby, Charles N. 2004. Panel Models In Sociological Research: Theory into Practice. *Annual Review of Sociology* 30: 507-544.
- Huinink, Johannes, Josef Brüderl, Bernhard Nauck, Sabine Walper, Laura Castiglioni und Michael Feldhaus. 2011. Panel Analysis of Intimate Relationships and Family Dynamics (pairfam): Conceptual framework and design. *Zeitschrift für Familienforschung* 23: 77-101.
- Isengard, Bettina. 2005. Freizeitverhalten als Ausdruck sozialer Ungleichheiten oder Ergebnis individualisierter Lebensführung? Zur Bedeutung von Einkommen und Bildung im Zeitverlauf. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 57: 254-277.

- Kalter, Frank. 2006. Auf der Suche nach einer Erklärung für die spezifischen Arbeitsmarktnachteile Jugendlicher türkischer Herkunft. *Zeitschrift für Soziologie* 35: 144-160.
- King, Gary. 1990. On Political Methodology. *Political Analysis* 2: 1-29.
- Kratzman, Jens, und Thorsten Schneider. 2009. Soziale Ungleichheit beim Schulstart. Empirische Untersuchungen zur Bedeutung der sozialen Herkunft und des Kindergartenbesuchs auf den Zeitpunkt der Einschulung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 61: 211-234.
- Legewie, Joscha. 2012. Die Schätzung von kausalen Effekten: Überlegungen zu Methoden der Kausalanalyse anhand von Kontexteffekten in der Schule. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 64: 123-153.
- Lohmann, Henning, C. K. Spieß, Olaf Groh-Samberg, und Jürgen Schupp. 2009. Analysepotenziale des Sozio-oekonomischen Panels (SOEP) für die empirische Bildungsforschung. *Zeitschrift für Erziehungswissenschaft* 12: 252-280.
- Pollmann-Schult, Matthias. 2006. Ausmaß und Struktur von arbeitnehmerinduzierter Abstiegsmobilität. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 58: 573-591.
- Pollmann-Schult, Matthias. 2008. Familiengründung und gewünschter Erwerbsumfang von Männern - Eine Längsschnittanalyse für die alten Bundesländer. *Zeitschrift für Soziologie* 37: 498-515.
- Pollmann-Schult, Matthias, und Martin Diewald. 2007. Auswirkungen der Familiengründung auf den Berufsverlauf von Männern. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 59: 440-458.
- Schröder, Jette, und Josef Brüderl. 2008. Der Effekt der Erwerbstätigkeit von Frauen auf die Fertilität: Kausalität oder Selbstselektion? *Zeitschrift für Soziologie* 37: 117-136.
- Schunck, Reinhard. 2013. Within- and Between-Estimates in Random Effects Models. Advantages and Drawbacks of Correlated Random Effects and Hybrid Models, *The Stata Journal* 13: 65-76.
- Sobel, Michael E. 2000. Causal Inference in the Social Sciences. *Journal of the American Statistical Association* 95: 647-651.
- Strauß, Susanne. 2009. Ehrenamt in Deutschland und Großbritannien – Sprungbrett zurück auf den Arbeitsmarkt? *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 61: 647-670.
- Trappmann, Mark, Stefanie Gundert, Claudia Wenzig, und Daniel Gebhard. 2010. PASS: a household panel survey for research on unemployment and poverty. *Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften* 130: 609-622.
- Wagner, Gert G., Joachim R. Frick, und Jürgen Schupp. 2007. The German socio-economic panel study (SOEP). Scope, evolution and enhancements. *Schmollers Jahrbuch: Zeitschrift für Wirtschafts- und Sozialwissenschaften* 127: 139-169.
- Windzio, Michael. 2013. *Regressionsmodelle für Zustände und Ereignisse*. Wiesbaden: Springer VS.
- Winship, Christopher, und Stephen Lawrence Morgan. 2007. *Counterfactuals and causal inference. Methods and principles for social research*. Cambridge: Univ. Press.

Tabelle 1: Datentransformation beim Episodensplitting

vor dem Episodensplitting:					
ID	StartT	EndeT	Dauer	Des	HeiratT
23	1243	1260	17	1	1250
67	960	1178	218	0	.

nach dem Episodensplitting:								
ID	StartT	EndeT	Stspl	Endspl	Dauer	Des	HeiratT	Ehe
23	1243	1260	1243	<u>1250</u>	7	0	1250	0
23	1243	1260	<u>1250</u>	1260	10	1	1250	1
67	960	1178	.	.	218	0	.	0

Tabelle 2: Panelregressionen in KZfSS und ZfS 2000-2009: Variablen und Methoden

Autoren	Variablentyp (UV)	Genanntes Motiv Paneldaten	Within- Effekt isoliert?
Best 2009 (ZfS)	veränderlich	Annäherung Kausalität	Ja
Esser 2009 (ZfS)	veränderlich	Annäherung Kausalität	Ja
Grunow et al. 2007 (ZfS)	veränderlich	Annäherung Kausalität	Ja
Isengard 2005 (KZfSS)	veränderlich	Trendanalyse	Nein
Kalter 2006 (ZfS)	veränderlich	Annäherung Kausalität	Ja
Pollmann-Schult 2008 (ZfS)	veränderlich	Annäherung Kausalität	Nein
Pollmann-Schult und Diewald 2007 (KZfSS)	veränderlich	Annäherung Kausalität	Ja
Strauß 2009 (KZfSS)	veränderlich	Annäherung Kausalität	Nein
Wolter und Schiener 2009 (KZfSS)	veränderlich	Annäherung Kausalität	Ja
Ziefle 2004 (KZfSS)	veränderlich	Annäherung Kausalität	Ja
Andreß und Seeck 2007 (KZfSS)	trend	Trendanalyse	Nein
Birkelbach 2003 (ZfS)	trend	Trendanalyse	Nein
Erlinghagen 2005 (ZfS)	trend	Trendanalyse	Nein
Giesecke und Verwiebe 2008 (ZfS)	trend	Trendanalyse	Nein
Klein und Fischer-Kerli 2000 (ZfS)	trend	Verknüpfung zeitv. Information	Nein
Lengfeld und Hirschle 2009 (ZfS)	trend	Trendanalyse	Nein
Wobbe und Otte 2000 (ZfS)	trend	Trendanalyse	Nein
Giesselmann 2009 (ZfS)	quasi-konstant	Verknüpfung zeitv. Information	Nein
Glowsky 2007 (ZfS)	quasi-konstant	Vergrößerung Stichprobe	Nein
Klein und Rosar 2007 (KZfSS)	quasi-konstant	Verknüpfung zeitv. Information	Nein
Schoen 2003 (KZfSS)	quasi-konstant	Verknüpfung zeitv. Information	Nein
Windzio und Grotheer 2002 (ZfS)	quasi-konstant	Verknüpfung zeitv. Information	Nein
Becker 2000 (KZfSS)	konstant	Verknüpfung zeitv. Information	Nein
Becker 2007 (KZfSS)	konstant	Verknüpfung zeitv. Information	Nein
Fuchs und Sixt 2007 (KZfSS)	konstant	Verknüpfung zeitv. Information	Nein
Groh-Samberg 2004 (KZfSS)	konstant	Verknüpfung zeitv. Information	Nein
Klein 2005 (ZfS)	konstant	Verknüpfung zeitv. Information	Nein
Klein und Pötschke 2004 (KZfSS)	konstant	Verknüpfung zeitv. Information	Nein
Korupp et al. 2002 (KZfSS)	konstant	Verknüpfung zeitv. Information	Nein
Kratzmann und Schneider 2009 (KZfSS)	konstant	Verknüpfung zeitv. Information	Nein
Lauterbach und Sacher 2001 (KZfSS)	konstant	Verknüpfung zeitv. Information	Nein
Leuze und Strauss 2009 (ZfS)	konstant	Verknüpfung zeitv. Information	Nein
Schneider 2004 (ZfS)	konstant	Verknüpfung zeitv. Information	Nein
Uhlig et al. 2009 (ZfS)	konstant	Vergrößerung Stichprobe	Nein

Tabelle 3: Ereignisanalysen in KZfSS und ZfS 2000-2009: Variablen und Methoden

Autoren	Variablentyp (UV)	Within-Effekt isoliert?
Bender et al. 2000 (KZfSS)	veränderlich	nein
Brose 2008 (KZfSS)	veränderlich	nein
Erlinghagen 2000 (KZfSS)	veränderlich	ja
Gebel und Gieseke 2009 (ZfS)	veränderlich	ja
Gieseke und Groß 2002 (KZfSS)	veränderlich	ja
Grunow et al. 2007 (ZfS)	veränderlich	nein
Hank 2003 (KZfSS)	veränderlich	nein
Hank et al. 2004 (ZfS)	veränderlich	nein
Klein und Becker 2008 (ZfS)	veränderlich	nein
Klein und Unger 2001 (KZfSS)	veränderlich	nein
Pollmann-Schult 2006 (KZfSS)	veränderlich	nein
Pollmann-Schult und Diewald 2007 (KZfSS)	veränderlich	nein
Schneider et al. 2001 (ZfS)	veränderlich	nein
Schröder und Brüderl 2008 (ZfS)	veränderlich	ja
Schulz und Blossfeld 2006 (KZfSS)	veränderlich	nein
Struck et al. 2007 (KZfSS)	veränderlich	nein
Windzio 2001 (ZfS)	veränderlich	nein
Windzio 2004 (KZfSS)	veränderlich	nein
Wingens et al. 2000 (KZfSS)	veränderlich	nein
Asendorpf 2008 (KZfSS)	quasi-konstant	nein
Lois 2008 (KZfSS)	quasi-konstant	nein
Brandt 2006 (KZfSS)	konstant	nein
Haak und Rasner 2009 (KZfSS)	konstant	nein
Kalter 2006 (ZfS)	konstant	nein
Klein et al. 2001 (ZfS)	konstant	nein
Konietzka 2001 (KZfSS)	konstant	nein
Nollmann 2009 (KZfSS)	konstant	nein
Uhlendorff 2004 (KZfSS)	konstant	nein

Anhang

Verwendete Studien im Literatursurvey

- Andreß, Hans-Jürgen, und Till Seeck. 2007. Ist das Normalarbeitsverhältnis noch armutsvermeidend? Erwerbstätigkeit in Zeiten deregulierter Arbeitsmärkte und des Umbaus sozialer Sicherungssysteme. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 59: 459-492.
- Asendorpf, Jens B. 2008. Living Apart Together: Alters- und Kohortenabhängigkeit einer heterogenen Lebensform. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 60: 749-764.
- Becker, Rolf. 2000. Klassenlage und Bildungsentscheidungen. Eine empirische Anwendung der Wert-Erwartungstheorie. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 52: 450-474.
- Becker, Rolf. 2007. Wie nachhaltig sind die Bildungsaufstiege wirklich? Eine Reanalyse der Studie von Fuchs und Sixt (2007) über die soziale Vererbung von Bildungserfolgen in der Generationenabfolge. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 59: 512-523.
- Bender, Stefan, Dirk Konietzka, und Peter Sopp. 2000. Diskontinuität im Erwerbsverlauf und betrieblicher Kontext. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 52: 475-499.
- Best, Henning. 2009. Kommt erst das Fressen und dann die Moral? Eine feldexperimentelle Überprüfung der Low-Cost-Hypothese und des Modells der Frame-Selektion. *Zeitschrift für Soziologie* 38: 131-151.
- Birkelbach, Klaus. 2003. Ärzteverbände im Urteil ihrer Mitglieder. Eine empirische Untersuchung der Zufriedenheit von Ärztinnen und Ärzten mit ihren Verbänden in den Jahren 1992 und 1998/99. *Zeitschrift für Soziologie* 32: 156-177.
- Brandt, Martina. 2006. Soziale Kontakte als Weg aus der Erwerbslosigkeit. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 58: 468-488.
- Brose, Nicole. 2008. Entscheidung unter Unsicherheit - Familiengründung und -erweiterung im Erwerbsverlauf. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 60: 30-52.
- Erlinghagen, Marcel. 2000. Arbeitslosigkeit und ehrenamtliche Tätigkeit im Zeitverlauf. Eine Längsschnittanalyse der westdeutschen Stichprobe des Sozio-oekonomischen Panels (SOEP) für die Jahre 1992 und 1996. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 52: 291-310.
- Erlinghagen, Marcel. 2005. Entlassungen und Beschäftigungssicherheit im Zeitverlauf. Zur Entwicklung unfreiwilliger Arbeitsmarktmobilität in Deutschland. *Zeitschrift für Soziologie* 34: 147-168.
- Esser, Hartmut. 2009. Pluralisierung oder Assimilation? Effekte der multiplen Inklusion auf die Integration von Migranten. *Zeitschrift für Soziologie* 38: 358-378.
- Fuchs, Marek, und Michaela Sixt. 2007. Zur Nachhaltigkeit von Bildungsaufstiegen. Soziale Vererbung von Bildungserfolgen über mehrere Generationen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 59: 1-29.
- Gebel, Michael, und Johannes Giesecke. 2009. Ökonomische Unsicherheit und Fertilität. Die Wirkung von Beschäftigungsunsicherheit und Arbeitslosigkeit auf die Familiengründung in Ost- und Westdeutschland. *Zeitschrift für Soziologie* 38: 399-417.
- Giesecke, Johannes, und Martin Groß. 2002. Befristete Beschäftigung: Chance oder Risiko? *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 54, 85-108.
- Giesecke, Johannes, und Roland Verwiebe. 2008. Die Zunahme der Lohnungleichheit in der Bundesrepublik. Aktuelle Befunde für den Zeitraum von 1998 bis 2005. *Zeitschrift für Soziologie* 37: 403-422.
- Giesselmann, Marco. 2009. Arbeitsmarktpolitischer Wandel in Deutschland seit 1991 und das Working Poor - Problem: Einsteiger als Verlierer des Reformprozesses? *Zeitschrift für Soziologie* 38: 215-238.
- Glowsky, David. 2007. Staatsbürgerschaft als Ressource bei der Heirat ausländischer Frauen. Eine Analyse mit Daten des Sozio-oekonomischen Panel. *Zeitschrift für Soziologie* 36: 282-301.
- Groh-Samberg, Olaf. 2004. Armut und Klassenstruktur. Zur Kritik der „Entgrenzungsthese“ aus einer multidimensionalen Perspektive. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 56: 653-682.
- Grunow, Daniela, Florian Schulz und Hans-Peter Blossfeld. 2007. Was erklärt die Traditionalisierungsprozesse häuslicher Arbeitsteilung im Eheverlauf: soziale Normen oder ökonomische Ressourcen? *Zeitschrift für Soziologie* 36: 162-181.

- Haak, Carroll, und Anika Rasner. 2009. Search (f)or Work: Der Übergang vom Studium in den Beruf. Geisteswissenschaftler im interdisziplinären Vergleich. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 61: 235-258.
- Hank, Karsten. 2003. Eine Mehrebenenanalyse regionaler Einflüsse auf die Familiengründung westdeutscher Frauen in den Jahren 1984 bis 1999. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 55: 79-98.
- Hank, Karsten, Michaela Kreyenfeld und Katharina Spieß. 2004. Kinderbetreuung und Fertilität in Deutschland. *Zeitschrift für Soziologie* 33: 228-244.
- Isengard, Bettina. 2005. Freizeitverhalten als Ausdruck sozialer Ungleichheiten oder Ergebnis individualisierter Lebensführung? Zur Bedeutung von Einkommen und Bildung im Zeitverlauf. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 57: 254-277.
- Kalter, Frank. 2006. Auf der Suche nach einer Erklärung für die spezifischen Arbeitsmarktnachteile Jugendlicher türkischer Herkunft. Zugleich eine Replik auf den Beitrag von Holger Seibert und Heike Solga: "Gleiche Chancen dank einer abgeschlossenen Ausbildung?" (*ZfS* 5/2005). *Zeitschrift für Soziologie* 35: 144-160.
- Klein, Markus. 2005. Der Einfluss der beiden TV-Duelle im Vorfeld der Bundestagswahl 2002 auf die Wahlbeteiligung und die Wahlentscheidung. Eine log-lineare Pfadanalyse auf der Grundlage von Paneldaten. *Zeitschrift für Soziologie* 34: 207-222.
- Klein, Markus, und Manuela Pötschke. 2004. Die intra-individuelle Stabilität gesellschaftlicher Wertorientierungen. Eine Mehrebenenanalyse auf der Grundlage des Sozio-oekonomischen Panels (SOEP). *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 56: 432-456.
- Klein, Markus, und Ulrich Rosar. 2007. Wirkungen des TV-Duells im Vorfeld der Bundestagswahl 2005 auf die Wahlentscheidung. Eine empirische Analyse unter besonderer Berücksichtigung von Medieneinflüssen auf die Siegerwahrnehmung und subjektiven Erwartungshaltung an die Debattenperformance der Kandidaten. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 59: 81-104.
- Klein, Thomas, und Simone Becker. 2008. Gibt es wirklich eine Reduzierung sportlicher Aktivität im Lebenslauf? Eine Analyse alters- und kohortenbezogener Unterschiede der Sportaktivität. *Zeitschrift für Soziologie* 37: 226-245.
- Klein, Thomas, und David Fischer-Kerli. 2000. Die Zuverlässigkeit retrospektiv erhobener Lebensverlaufsdaten. Analysen zur Partnerschaftsbiografie des Familiensurvey. *Zeitschrift für Soziologie* 29: 294-312.
- Klein, Thomas, Sven Schneider und Hannelore Löwel. 2001. Bildung und Mortalität. Die Bedeutung gesundheitsrelevanter Aspekte des Lebensstils. *Zeitschrift für Soziologie* 30: 384-400.
- Klein, Thomas, und Rainer Unger. 2001. Einkommen, Gesundheit und Mortalität in Deutschland, Großbritannien und den USA. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 53: 96-111.
- Konietzka, Dirk. 2001. Hat sich das duale System in den neuen Ländern erfolgreich etabliert? Ausbildung, Arbeitslosigkeit und Berufseinstieg in Ost- und Westdeutschland in der Mitte der 1990er Jahre. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 53: 50-75.
- Korupp, Sylvia, Harry B. Ganzeboom und Karin Sanders. 2002. Wie die Mutter, so die Tochter? Dynamiken und Trends im Statusvererbungsprozess zwischen Müttern und Töchtern in Westdeutschland und den Niederlanden. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 54: 1-26.
- Kratzman, Jens, und Thorsten Schneider. 2009. Soziale Ungleichheit beim Schulstart. Empirische Untersuchungen zur Bedeutung der sozialen Herkunft und des Kindergartenbesuchs auf den Zeitpunkt der Einschulung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 61: 211-234.
- Lauterbach, Wolfgang, und Matthias Sacher. 2001. Erwerbseinstieg und erste Erwerbsjahre. Ein Vergleich von vier westdeutschen Geburtskohorten. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 53: 258-282.
- Lengfeld, Holger, und Jochen Hirschle. 2009. Die Angst der Mittelschicht vor dem sozialen Abstieg. Eine Längsschnittanalyse 1984-2007. *Zeitschrift für Soziologie* 38: 379-398.
- Leuze, Kathrin, und Susanne Strauß. 2009. Lohnungleichheiten zwischen Akademikerinnen und Akademikern: Der Einfluss von fachlicher Spezialisierung, frauendominierten Fächern und beruflicher Segregation. *Zeitschrift für Soziologie* 38: 262-281.

- Lois, Daniel. 2008. Arbeitsteilung, Berufsorientierung und Partnerschaftsstabilität - Ehen und nichteheliche Lebensgemeinschaften im Vergleich. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 60: 53-77.
- Nollmann, Gerd. 2009. Working Poor. Eine vergleichende Längsschnittstudie für Deutschland und die USA. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 61: 33-55.
- Pollmann-Schult, Matthias. 2006. Ausmaß und Struktur von arbeitnehmerinduzierter Abstiegsmobilität. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 58: 573-591.
- Pollmann-Schult, Matthias. 2008. Familiengründung und gewünschter Erwerbsumfang von Männern - Eine Längsschnittanalyse für die alten Bundesländer. *Zeitschrift für Soziologie* 37: 498-515.
- Pollmann-Schult, Matthias, und Martin Diewald. 2007. Auswirkungen der Familiengründung auf den Berufsverlauf von Männern. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 59: 440-458.
- Schneider, Thorsten. 2004. Der Einfluss des Einkommens der Eltern auf die Schulwahl. *Zeitschrift für Soziologie* 33: 471-492.
- Schneider, Thorsten, Sonja Drobnic und Hans-Peter Blossfeld. 2001. Pflegebedürftige Personen im Haushalt und das Erwerbsverhalten verheirateter Frauen. *Zeitschrift für Soziologie* 30: 362-383.
- Schoen, Harald. 2003. Gleich und gleich gesellt sich gern - auch im Wahlkampf? Zur politischen Prägekraft der Sozialstruktur in Abhängigkeit von der Nähe zum Wahltag. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 55: 299-320.
- Schröder, Jette, und Josef Brüderl. 2008. Der Effekt der Erwerbstätigkeit von Frauen auf die Fertilität: Kausalität oder Selbstselektion? *Zeitschrift für Soziologie* 37: 117-136.
- Schulz, Florian, und Hans-Peter Blossfeld. 2006. Wie verändert sich die häusliche Arbeitsteilung im Eheverlauf? Eine Längsschnittstudie der ersten 14 Ehejahre in Westdeutschland. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 58: 23-49.
- Strauß, Susanne. 2009. Ehrenamt in Deutschland und Großbritannien – Sprungbrett zurück auf den Arbeitsmarkt? *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 61: 647-670.
- Struck, Olaf, Michael Grotheer, Tim Schröder und Christoph Köhler. 2007. Instabile Beschäftigung. Ergebnisse zu einer alten Kontroverse. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 59: 271-293.
- Uhlendorff, Arne. 2004. Der Einfluss von Persönlichkeitseigenschaften und sozialen Ressourcen auf die Arbeitslosigkeitsdauer. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 56, 279-303.
- Uhlig, Johannes, Heike Solga und Jürgen Schupp. 2009. Bildungsungleichheiten und blockierte Lernpotenziale: Welche Bedeutung hat die Persönlichkeitsstruktur für diesen Zusammenhang? *Zeitschrift für Soziologie* 38: 418-440.
- Windzio, Michael. 2001. Organisationsökologie und Arbeitsmarktmobilität im sozialen Wandel. Eine empirische Analyse am Beispiel Ostdeutschland. *Zeitschrift für Soziologie* 30: 116-134.
- Windzio, Michael. 2004. Kann der regionale Kontext zur "Arbeitslosenfalle" werden? Der Einfluss der Arbeitslosigkeit auf die Mobilität zwischen regionalen Arbeitsmärkten in Westdeutschland. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 56: 257-278.
- Windzio, Michael, und Michael Grotheer. 2002. Bleiben die Erfolgreichen übrig? Die Kombination von Sequenzmusteranalysen und log-linearen Pfadmodellen bei der Analyse des Zusammenhangs von Berufserfolg und Panelmortalität. *Zeitschrift für Soziologie* 31: 514-528.
- Wingens, Matthias, Reinhold Sackmann und Michael Grotheer. 2000. Berufliche Qualifizierung für Arbeitslose. Zur Effektivität AFG-finanzierter Weiterbildung im Transformationsprozeß. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 52: 60-80.
- Wobbe, Theresa, und Roland Otte. 2000. Politische Institutionen im gesellschaftlichen Wandel. Einbürgerung zwischen Erwartungen von Migranten und staatlicher Vorgabe. *Zeitschrift für Soziologie* 29: 444-462.
- Wolter, Felix, und Jürgen Schiener. 2009. Governanceregimes im Wandel. Eine Analyse des außeruniversitären, staatlich finanzierten Forschungssektors in Deutschland. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 61: 90-117.
- Ziefle, Andrea. 2004. Die individuellen Kosten des Erziehungsurlaubs: Eine empirische Analyse der kurz- und längerfristigen Folgen für den Karriereverlauf von Frauen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 56: 213-231.