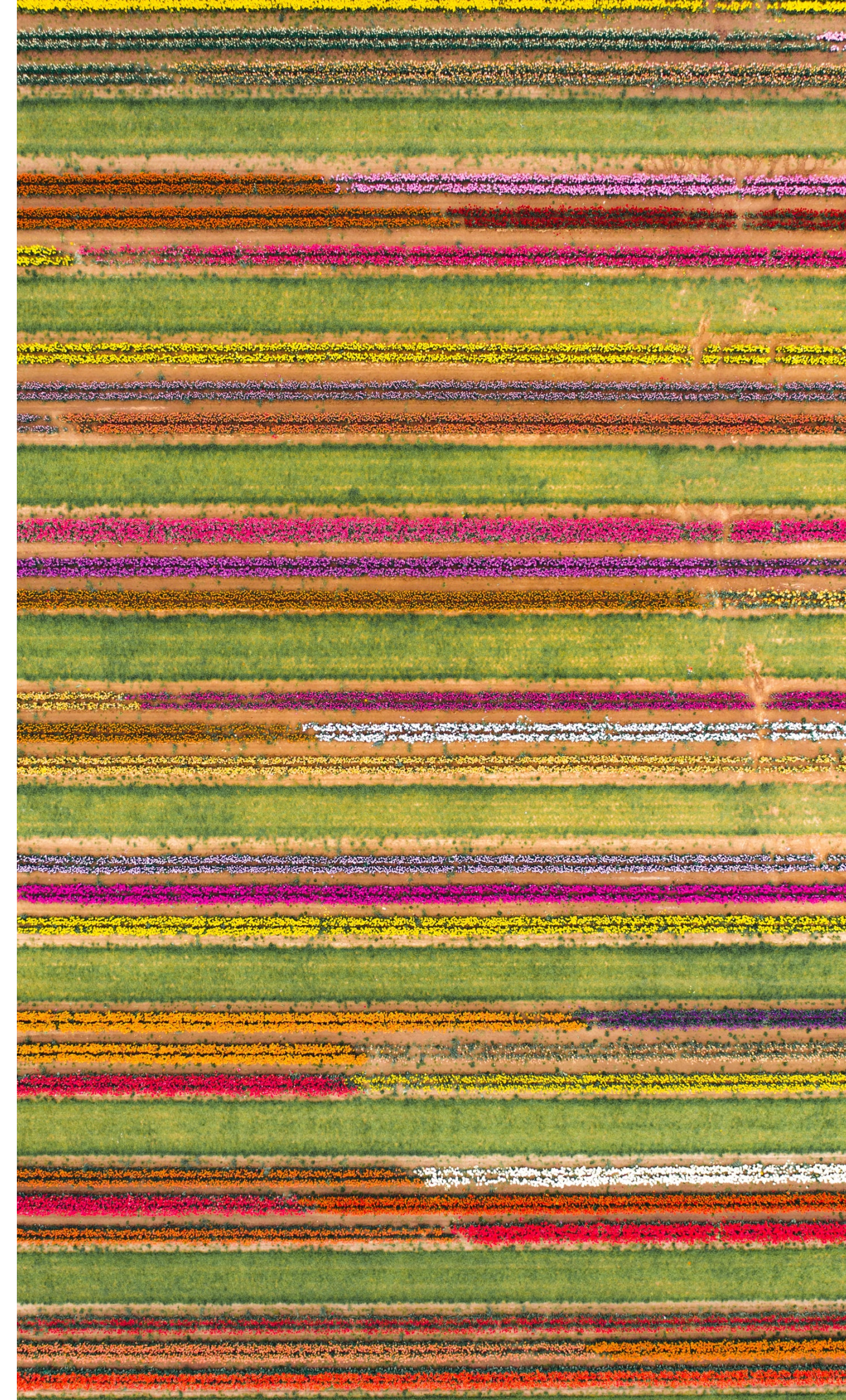


Marcel Raab |
StatistikTage Bamberg|Fürth |
14. Juli 2021 |

Sequenzanalyse

Ein vielseitiges Werkzeug zur Untersuchung familialer Prozesse



Gliederung

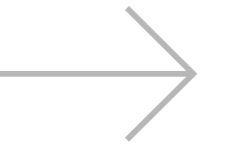
01 **Grundlagen**

02 **Anwendungsbeispiele**

03 **Neuere Entwicklungen**

Grundlagen

Was sind Sequenzen?



Was sind Sequenzen?

1

Geordnete Abfolgen kategorialer Zustände

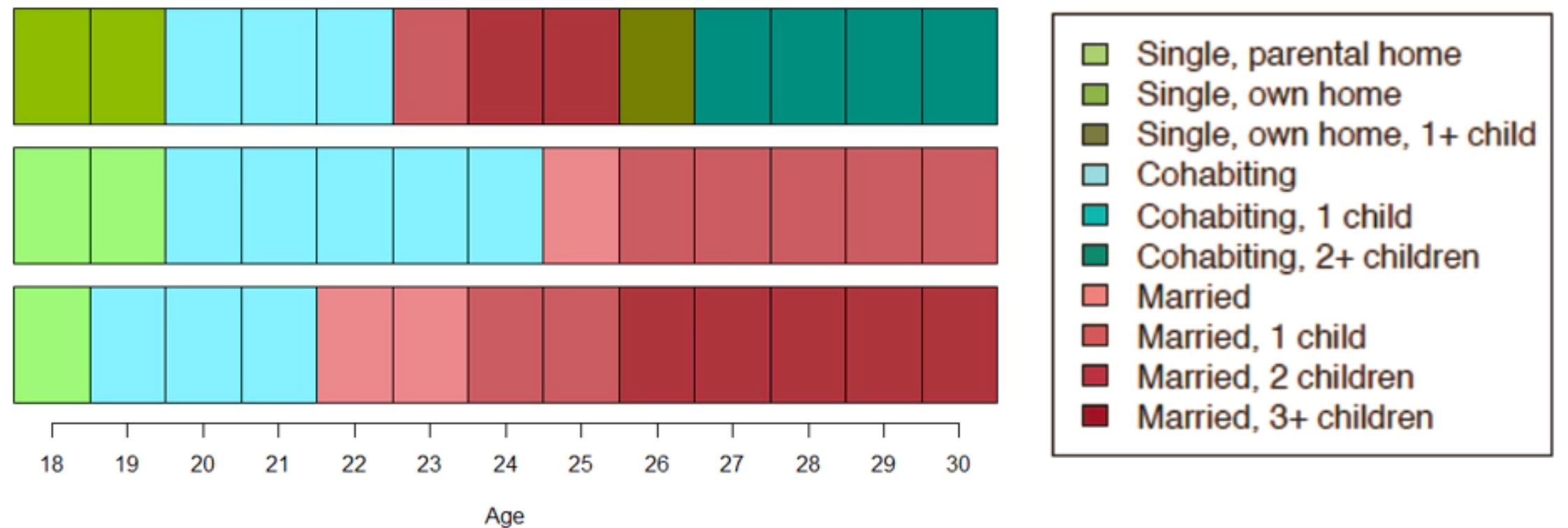
⌚ Typischer Weise zeitliche Abfolgen ...

2

... aber auch

- räumliche Ordnungen
- Präferenzhierarchien
- Kognitive Schemata und Skripte
- etc.

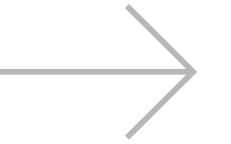
Illustration

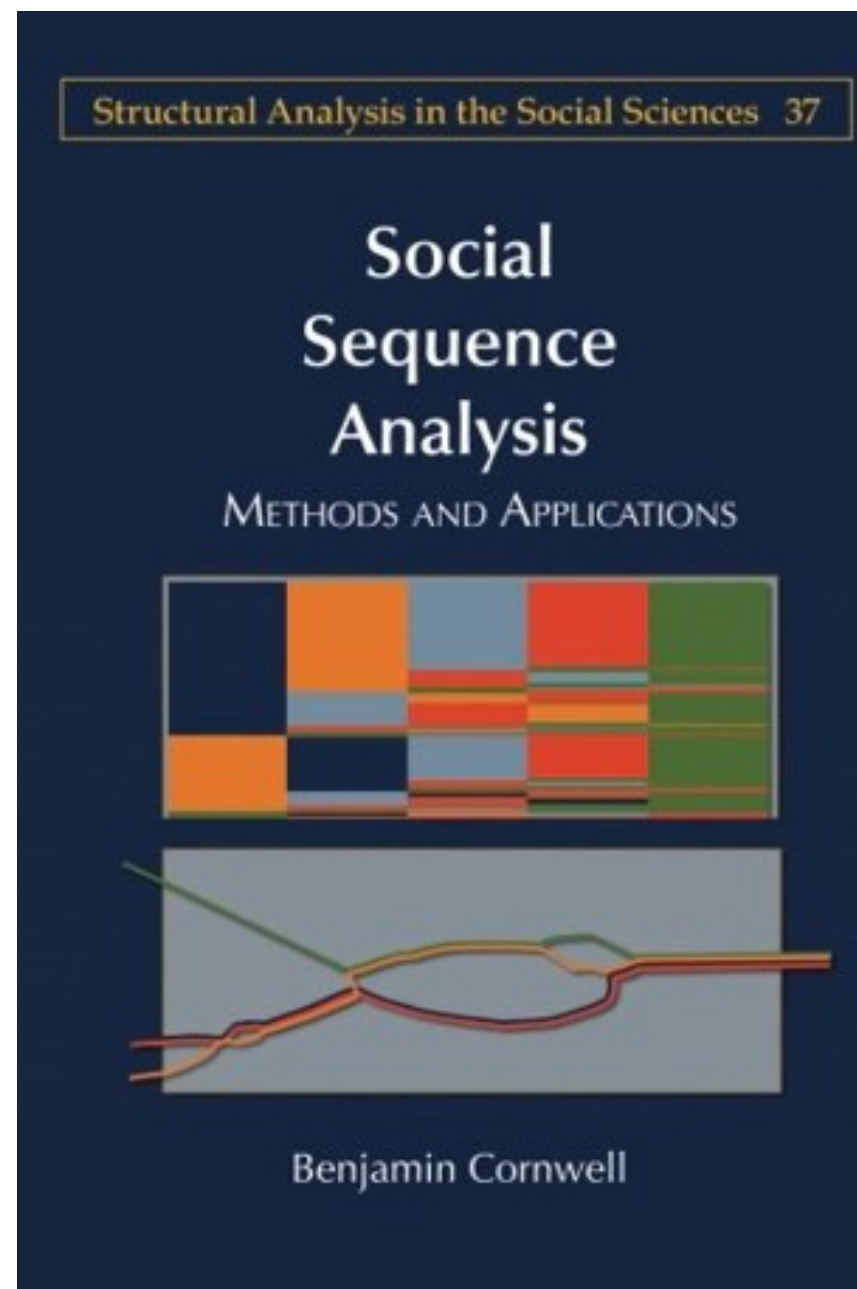


Drei **Familienbiografien** vom Alter 18 bis 30 mit jährlichen Informationen zu Partnerschaftsstatus, Wohnarrangement und Kinderanzahl

Grundlagen

Warum sind Sequenzen
interessant?



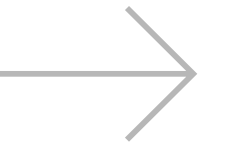


"Social actors of all kinds – individuals, families, organizations – are connected to each other and to larger society through ordered elements, states, and events. [...] examining the timing and order of these elements of social life is the only way to understand social structure."

Benjamin Cornwell

Grundlagen

Methodenvergleich

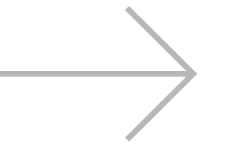


	Event history analysis	Sequence analysis
Theoretical concept	Transition/duration	Trajectory
Objective to identify . . .	Probability of single transitions/durations	Temporal patterns of sequential equivalence
Scientific tradition	Stochastic data modeling culture	Algorithmic modeling culture
Assumptions about data	Generated by stochastic process	None/“black box”

Aisenbrey & Fasang 2010: 424

Grundlagen

Typischer Ablauf einer
Sequenzanalyse



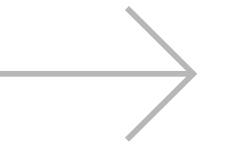
Klassische Sequenzanalyse

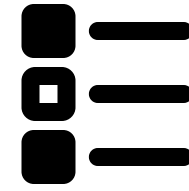




Grundlagen

Die Qual der Wahl





1

Alphabet

- Größe & Dimensionalität
- Häufigkeit der Zustände
- iterativer Prozess
- Rückbezug zur Forschungsfrage

Relativ robuste Ergebnisse bei verschiedenen
Alphabetspezifikationen

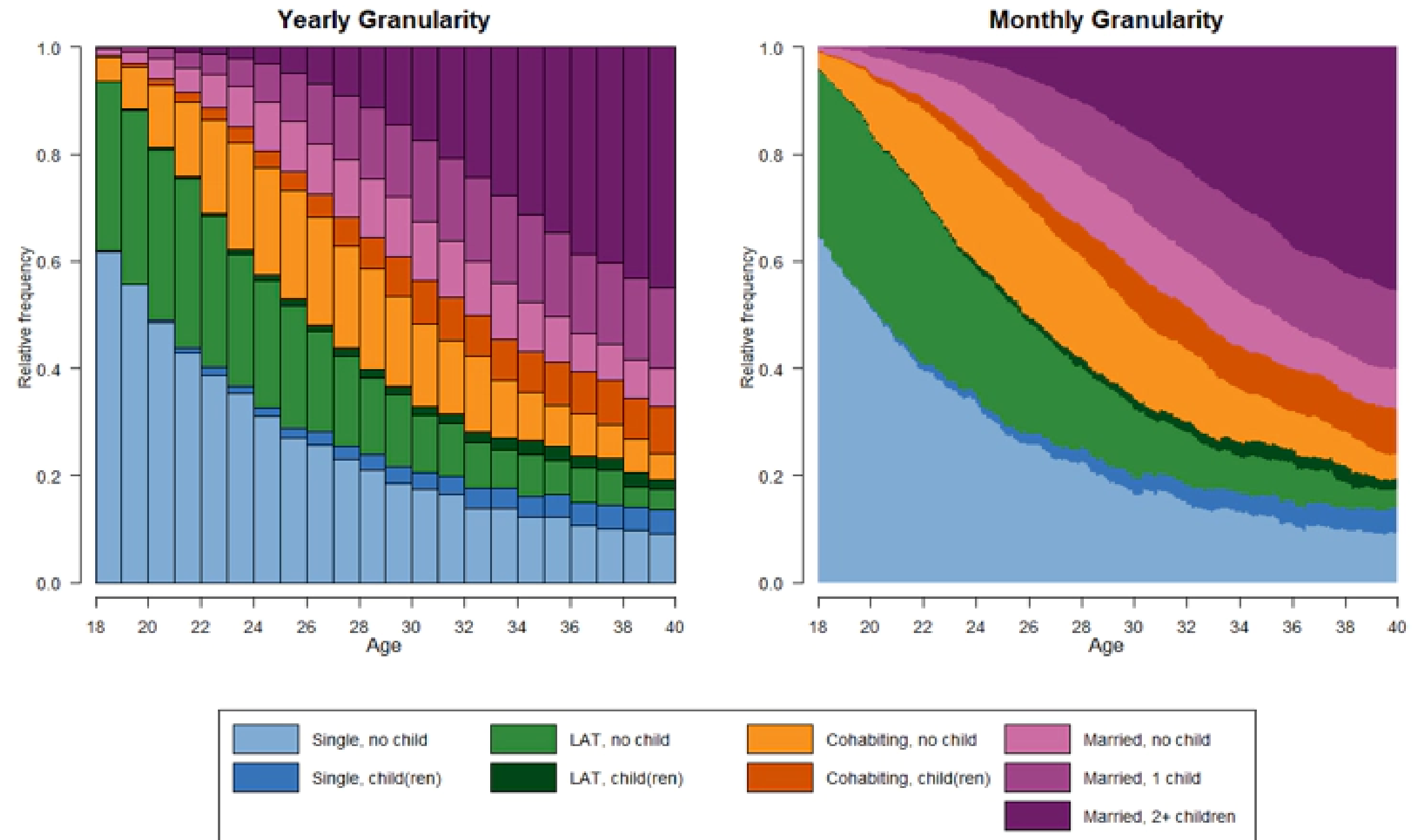
Forrest & Abbott 1990

2

Sequenzlänge & Granularität

- Prozess- vs. Kalenderzeit
- Bezugspunkt für Start und Ende
- Wichtigkeit kursorischer Episoden
- Sequenzen ungleicher Länge
- Umgang mit fehlenden Werten

Illustration

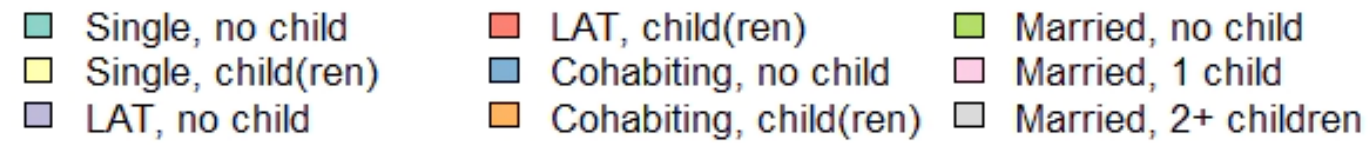
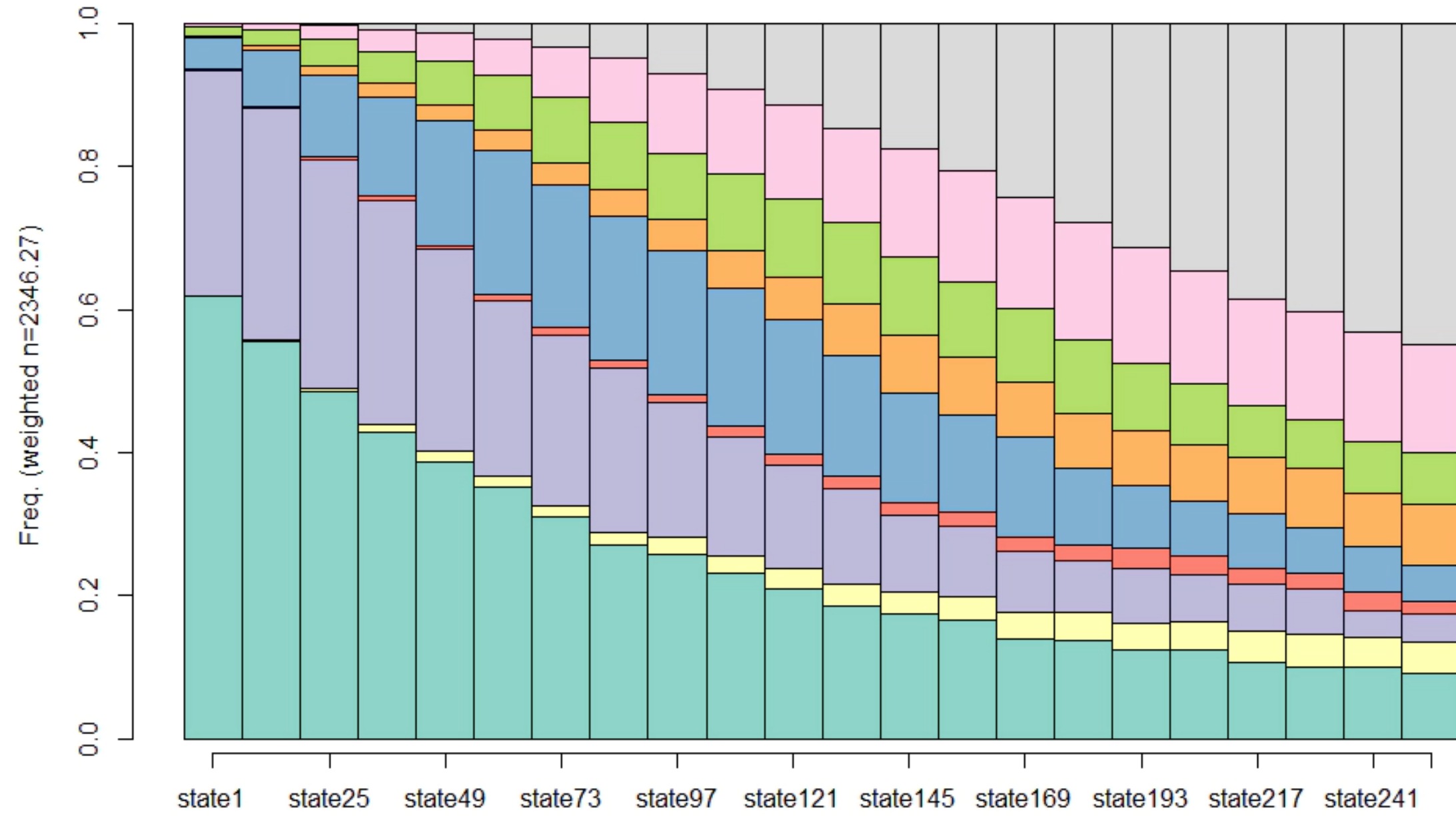


State Distribution Plot auf Basis von Daten des Beziehungs- und Familienpanels (pairfam) mit jährlicher und monatlicher Granularität

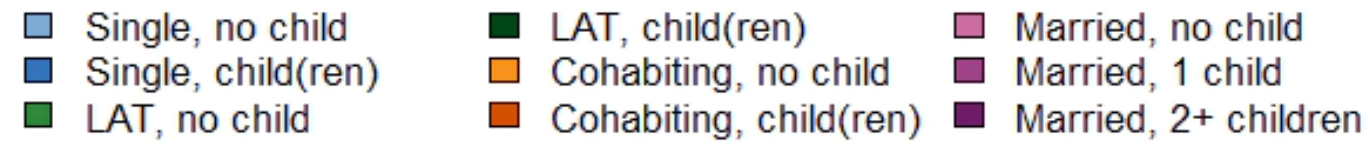
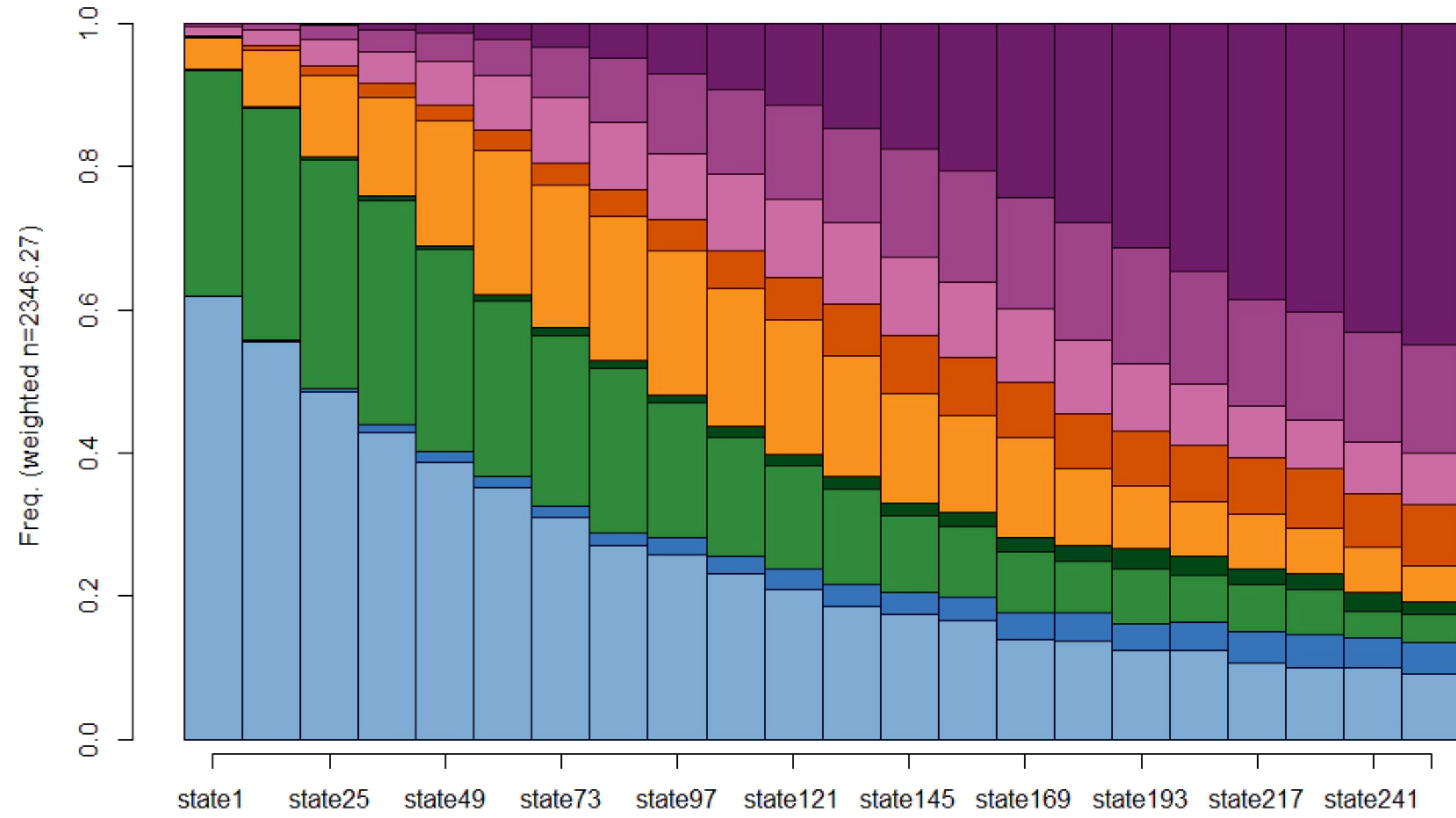
3

Farben

Default Colors



Custom Colors



4

Sequenzvergleich

"The assignment of transformation costs haunts all optimal matching analyses"

Stovel et al. 1996: 394

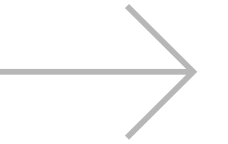
6

Clusteranalyse

... und anschließende Regression

Anwendungsfelder

Familienforschung

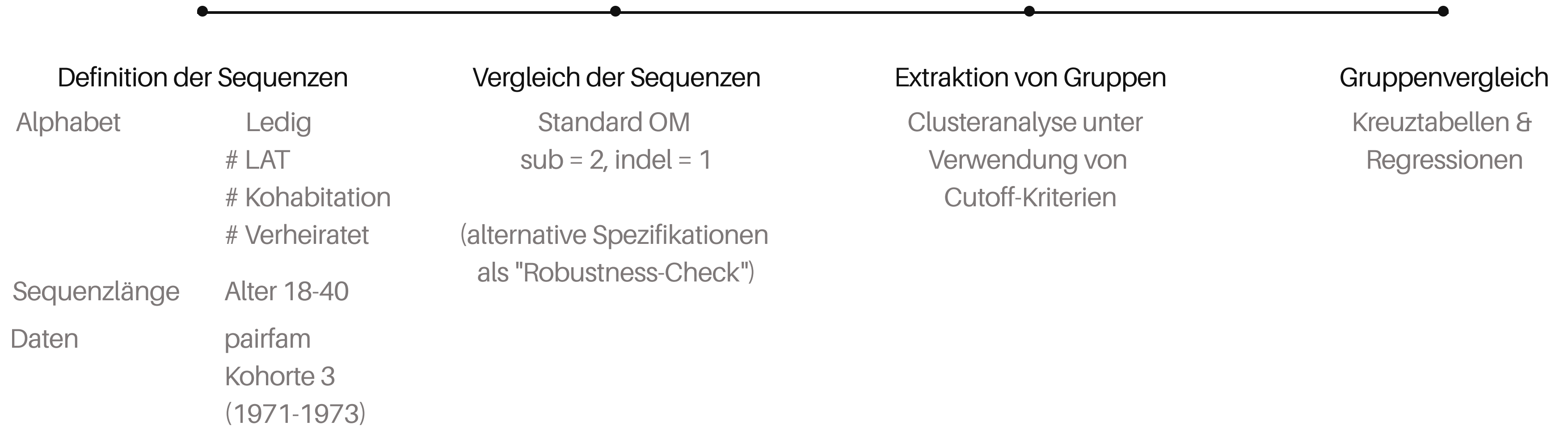


Partnerschaftsbiografien von Kinderlosen

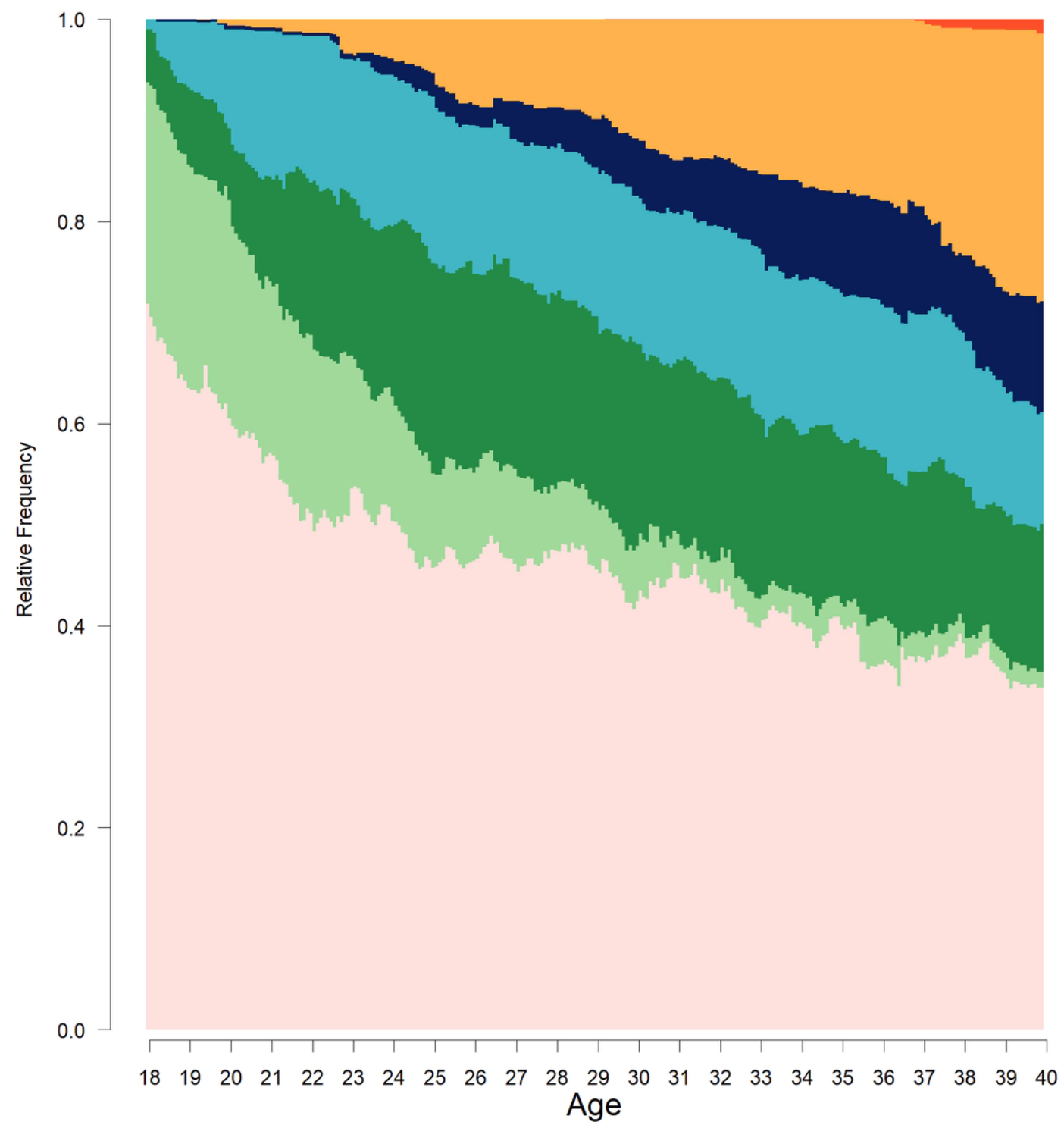
Raab & Struffolino

Raab, M., Struffolino, E. The Heterogeneity of Partnership Trajectories to Childlessness in Germany.
Eur J Population 36, 53–70 (2020). <https://doi.org/10.1007/s10680-019-09519-y>

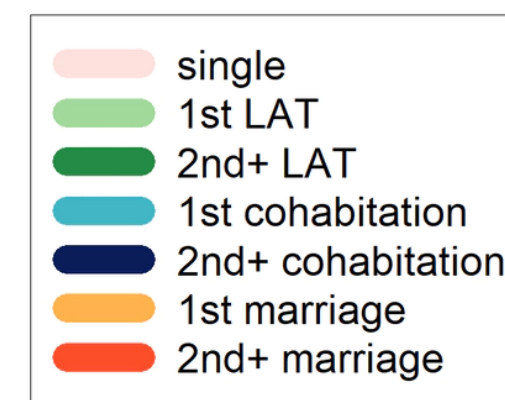
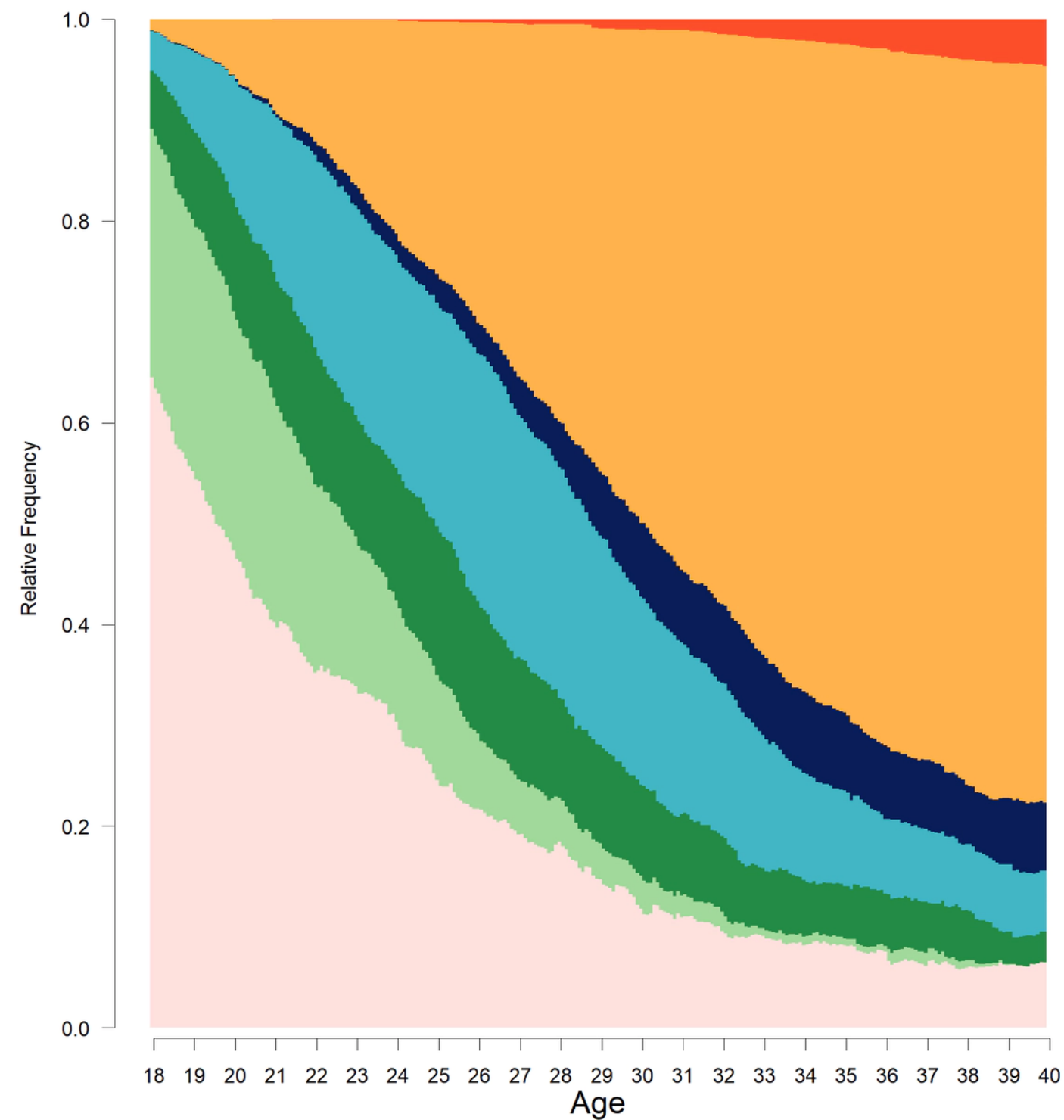
Klassische Sequenzanalyse

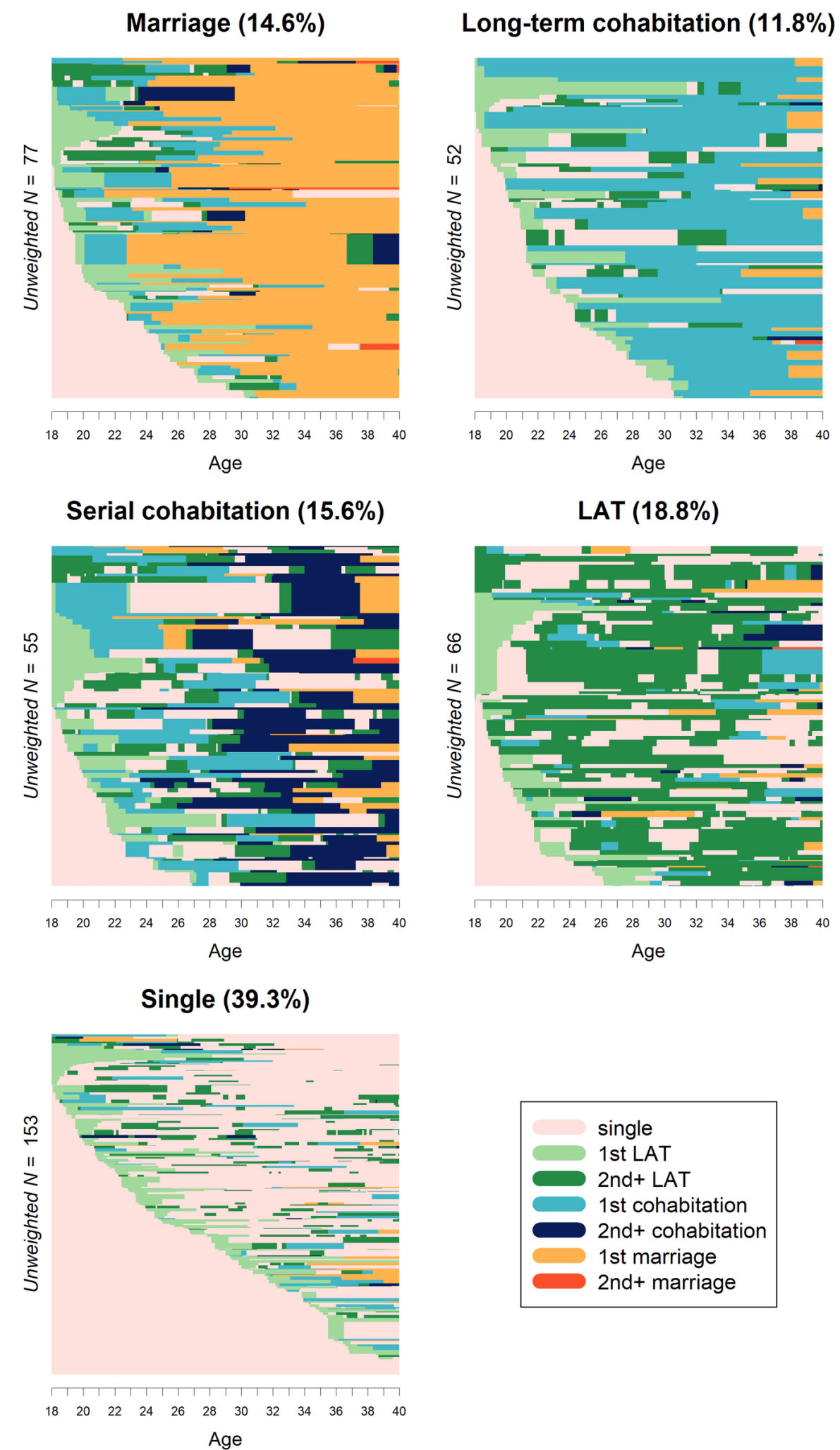
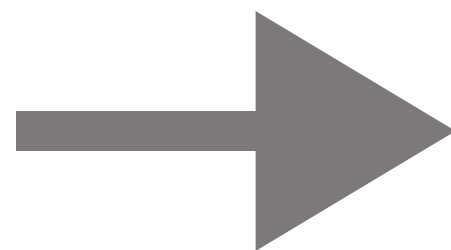
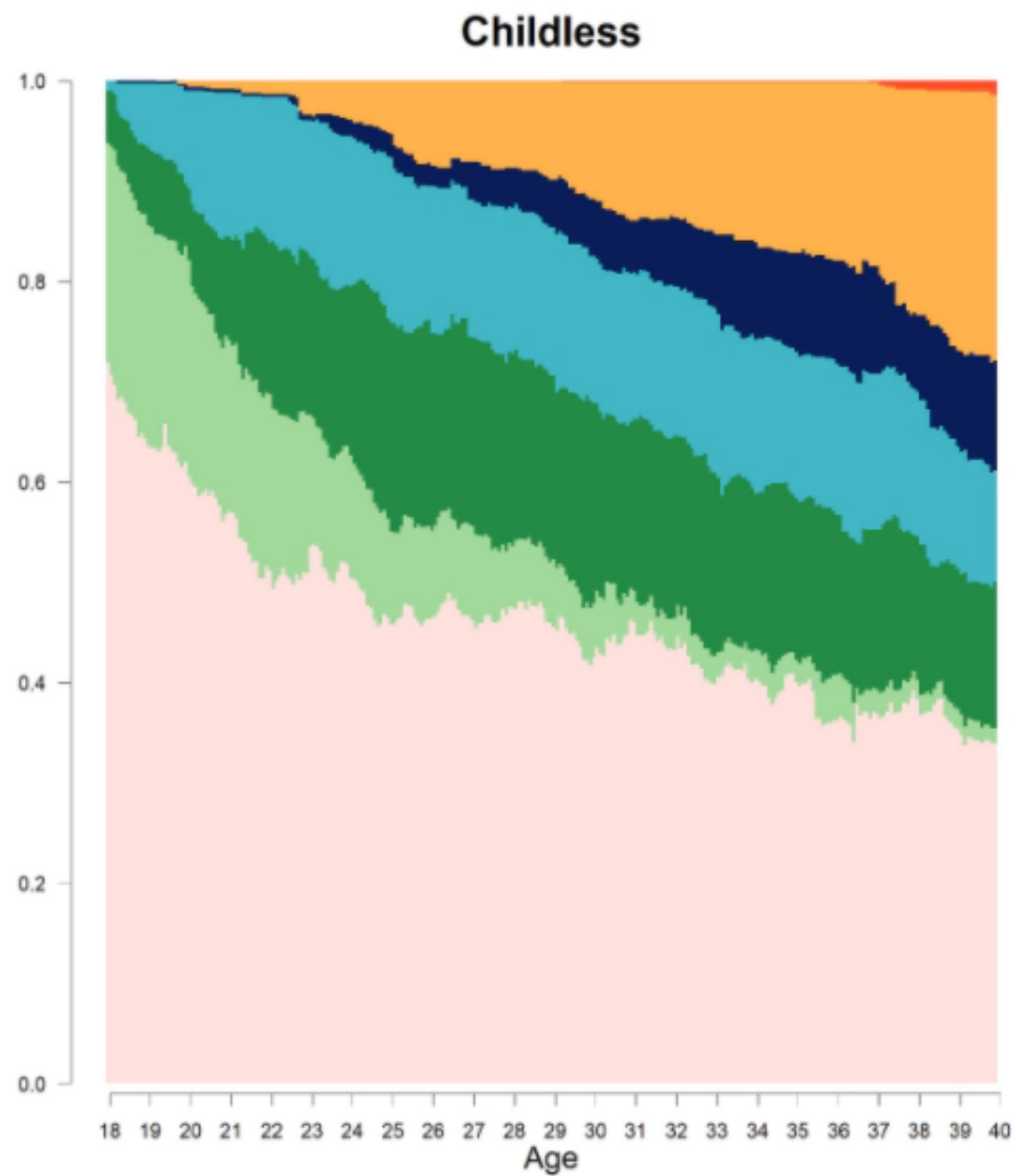


Childless



Parents





Turbulence of partnership trajectories

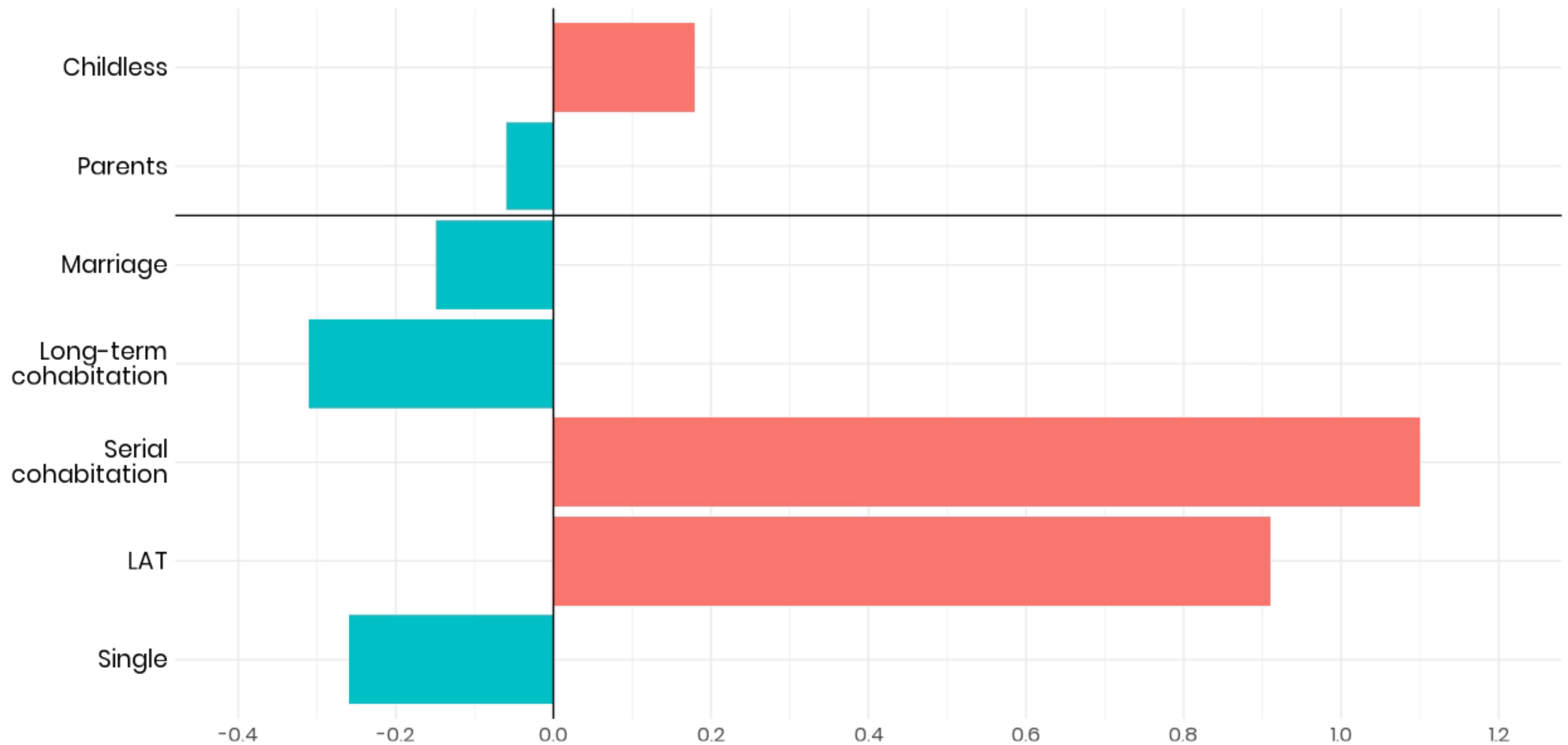


Table 2 Adjusted predictions and marginal effects at group-specific means

	Marriage			Long-term cohabitation			Serial cohabitation			LAT			Single		
	Female	Male	Δ gender	Female	Male	Δ gender	Female	Male	Δ gender	Female	Male	Δ gender	Female	Male	Δ gender
Overall	.218	.082	.136***	.122	.119	.002	.163	.157	.006	.205	.165	.041	.292	.477	−.185***
Education															
<i>No Abitur</i>	.326	.100	.226***	.117	.132	−.015	.176	.153	.023	.133	.230	−.096†	.247	.385	−.138*
<i>Abitur</i>	.144	.067	.077†	.117	.105	.013	.143	.154	−.012	.279	.118	.162**	.317	.556	−.240***
Δ education	−.183**	−.034	.149*	.000	−.027	−.028	−.033	.002	.035	.146*	−.112**	−.258***	.069	.171**	.102

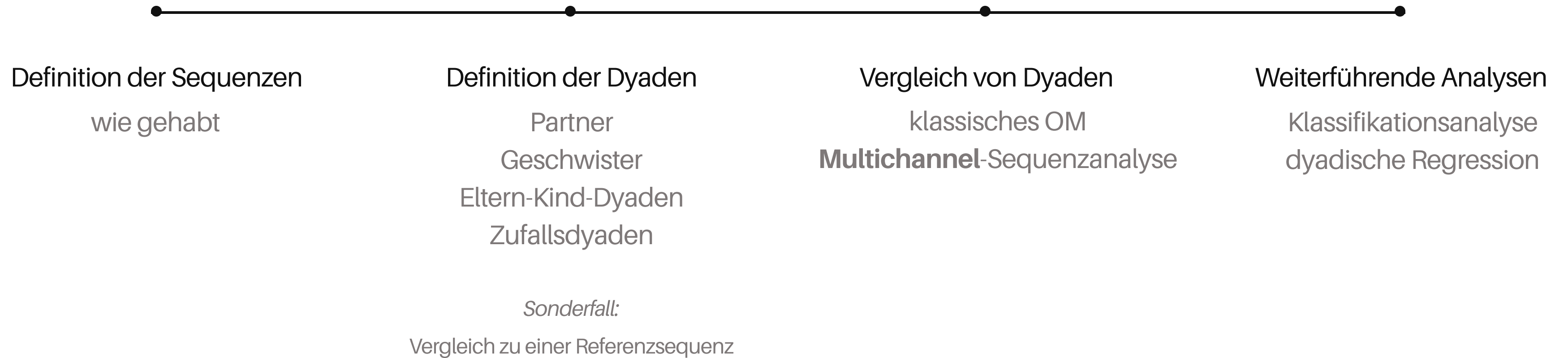
Data: pairfam, release 9.0; results are weighted by attrition-adjusted post-stratification weights

*** $p \leq 0.001$; ** $p \leq 0.01$; * $p \leq 0.05$; † $p \leq 0.10$

Alternative Anwendungsfelder

Dyadische Analysen

Dyadische Sequenzanalyse



Geschwisterähnlichkeiten

Karhula, Erola,
Raab & Fasang

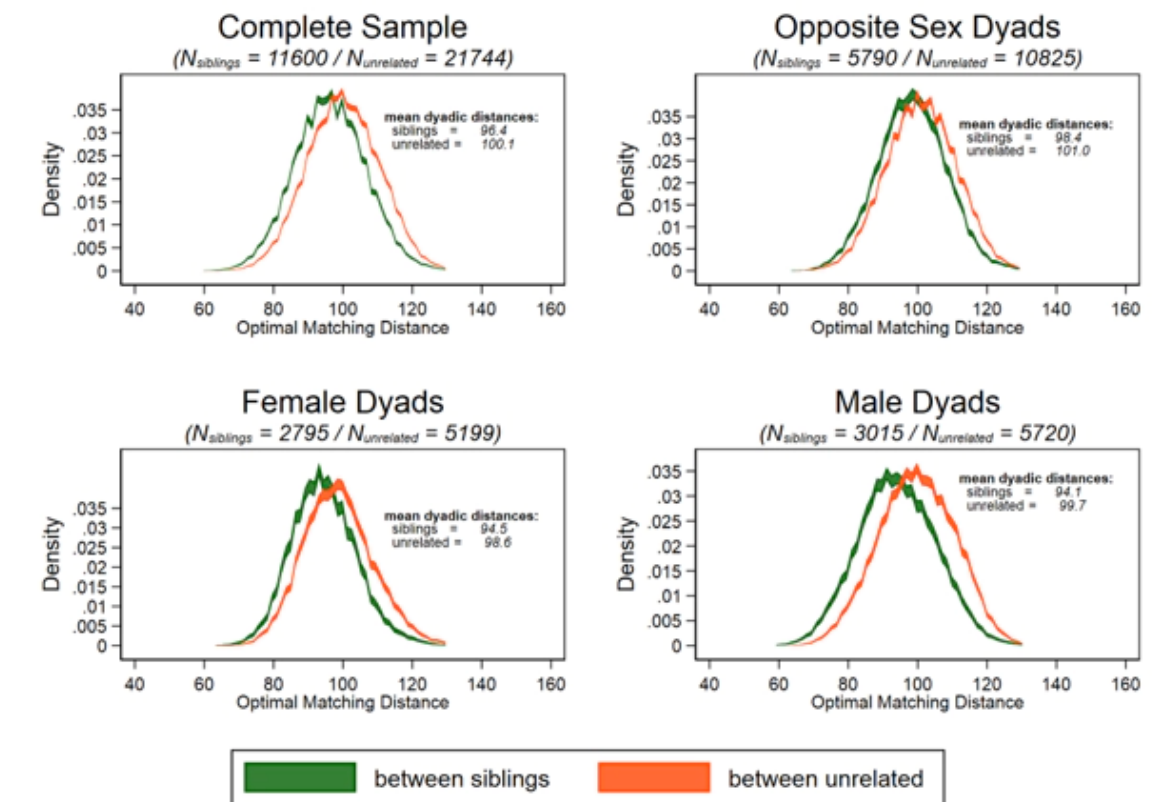
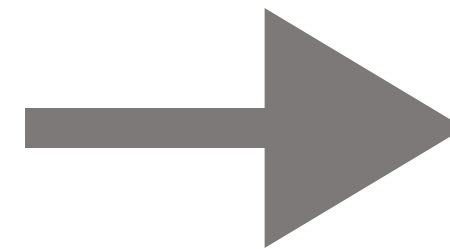
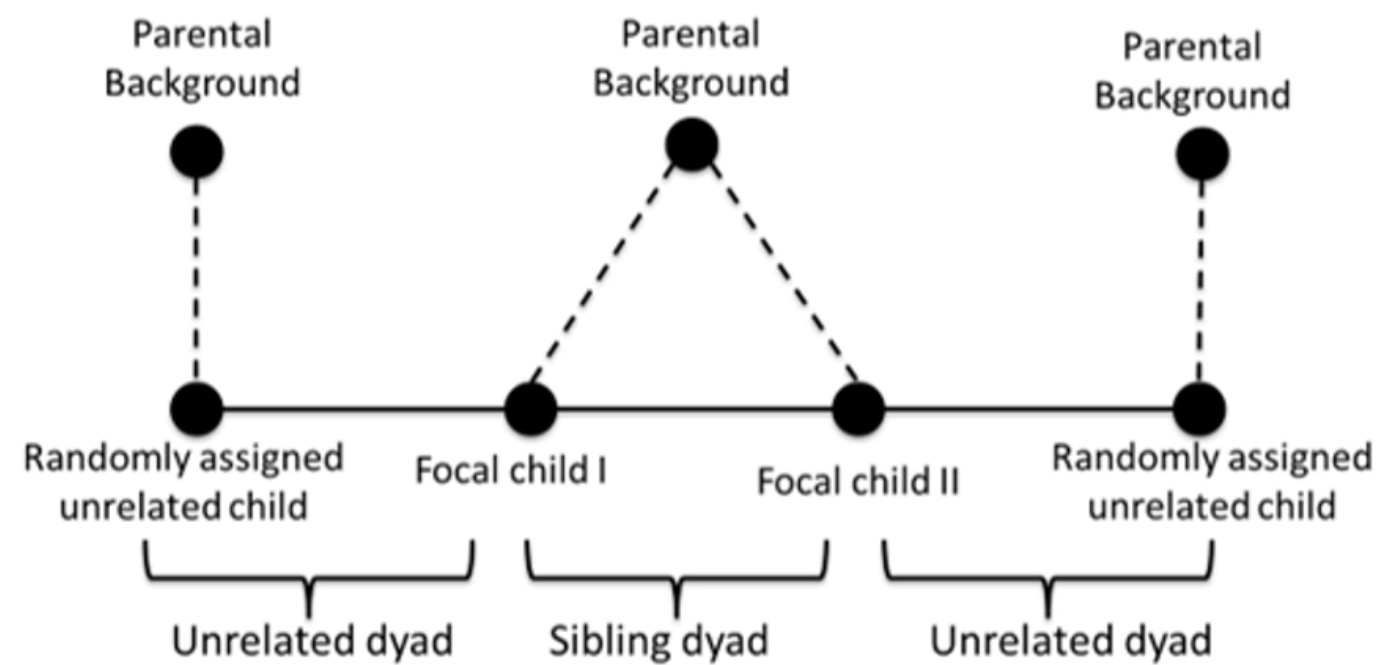




Karhula, A., Erola, J., Raab, M., & Fasang, A. (2019). Destination as a process: Sibling similarity in early socioeconomic trajectories. *Advances in Life Course Research*, 40, 85–98. <https://doi.org/10.1016/j.alcr.2019.04.015>

Forschungsfrage 1

Sind sich Geschwister ähnlicher?

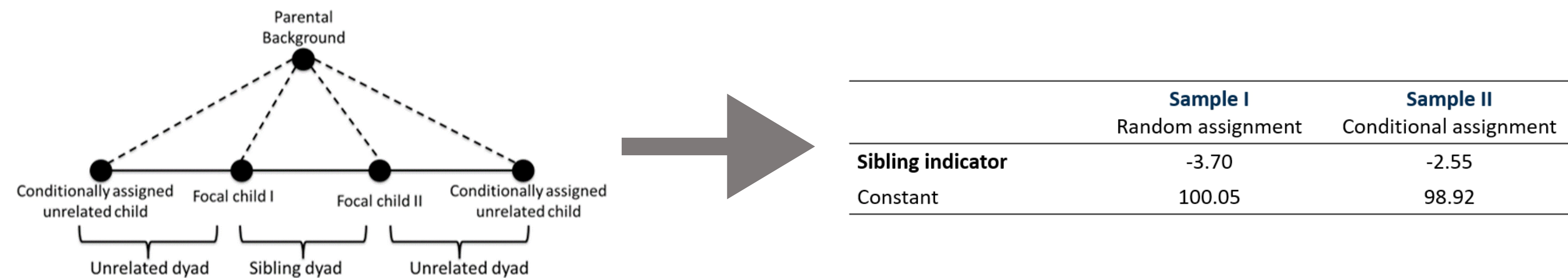


Ja.

Vor allem Geschwister selben Geschlechts

Forschungsfrage 2

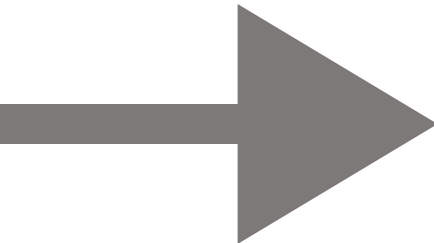
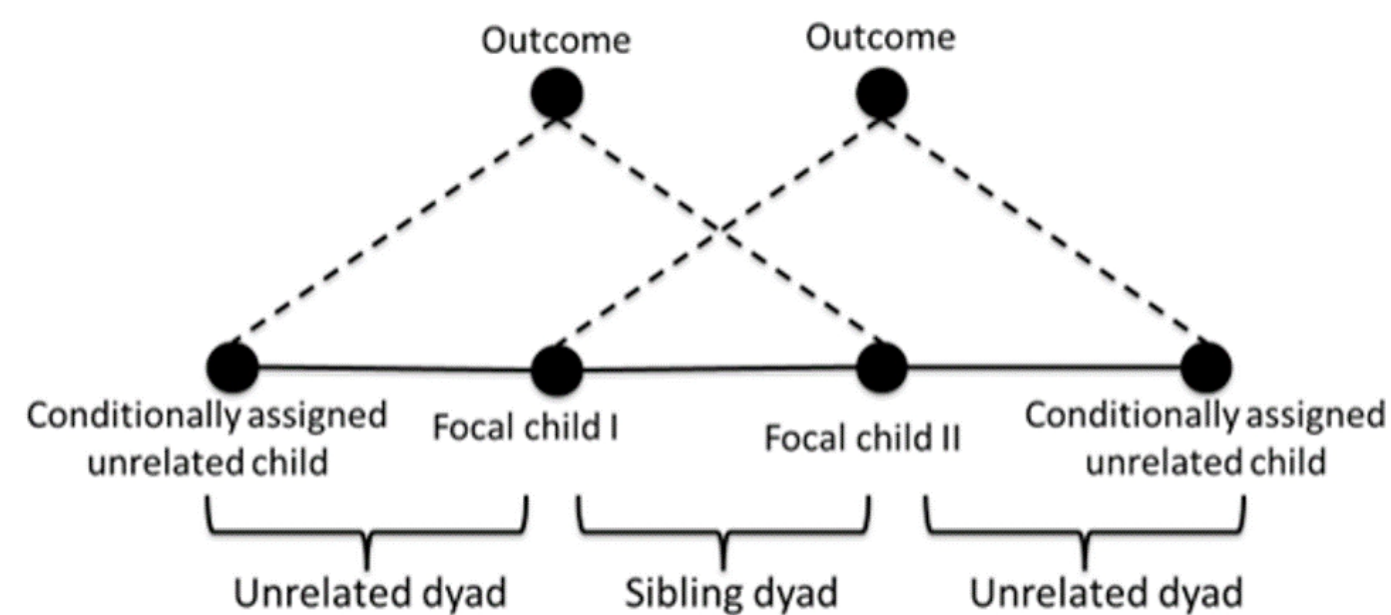
Lässt sich die Ähnlichkeit durch geteilten elterlichen Hintergrund erklären?



Ja, teilweise.
ca. 30% des Geschwistereffekts

Forschungsfrage 3

Wäre es nicht ausreichend nur die Ähnlichkeit zum Ende des Prozesses anzuschauen um das volle Ausmaß des Geschwistereffekts zu erfassen?



	Random assignment	Conditional assignment on ... (at age 35)		
		... education	... earnings	... both
Sibling indicator	-3.70	-2.28	-2.79	-1.79
Constant	100.05	98.65	99.16	98.15
Decrease in sibling effect (%) compared to random assignment		38.4	24.6	51.6

Nein.

Der Zielzustand ist zwar wichtig, aber der Herkunftseffekt beeinflusst auch maßgeblich den Weg dorthin

Zeitverwendung in Familien

Raab & Schulz

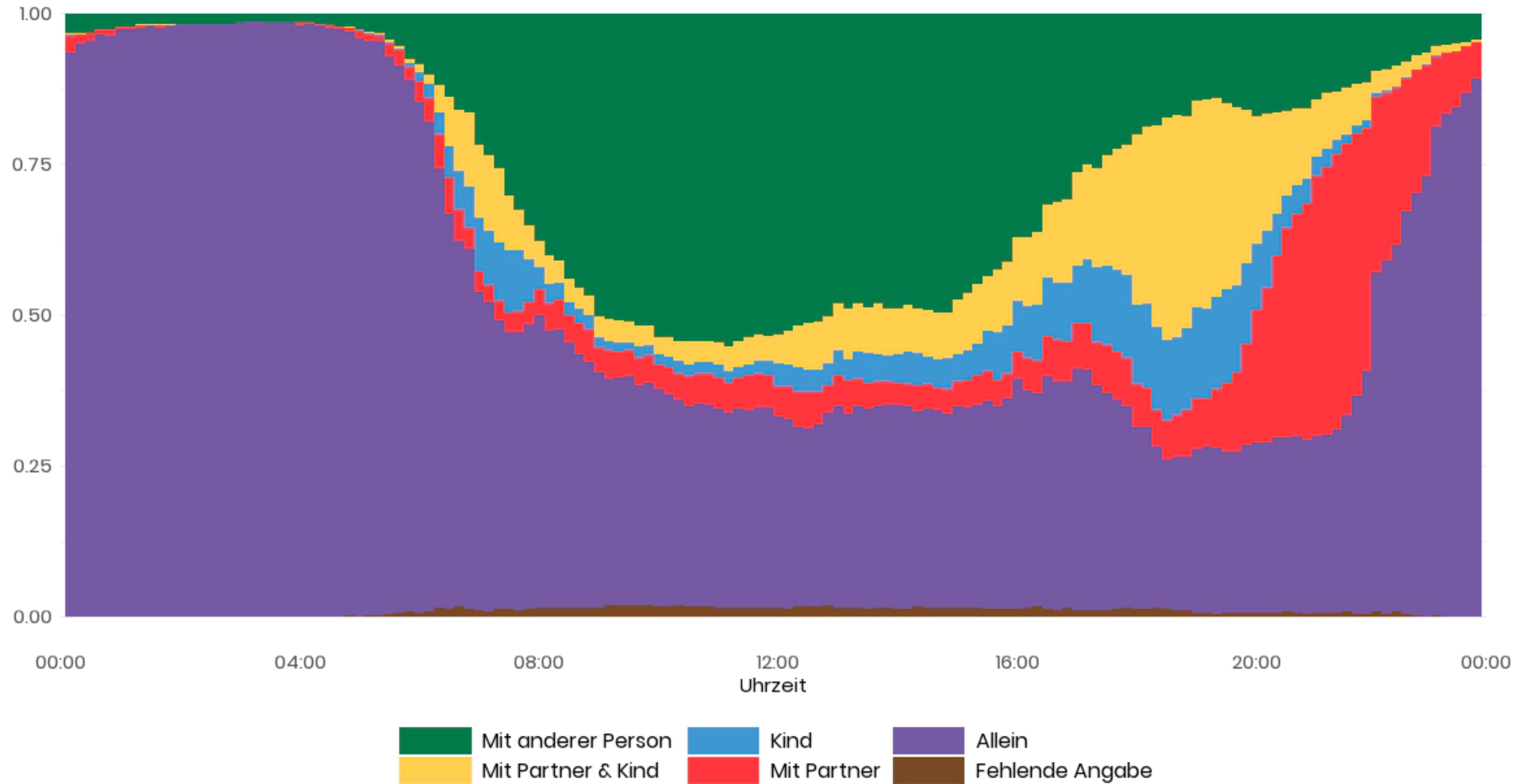
Erwerbstätigkeit & Elternschaft

Die Abbildungen basieren auf vorläufigen Analysen von 954 Tagebucheinträgen für Werktage von 494 Paaren, deren jüngstes Kind 3 bis 10 Jahre alt ist.



Mit wem verbringen Eltern ihre Zeit?

Die Abbildungen basiert auf vorläufigen Analysen von 954 Tagebucheinträgen für Werktage von 494 Paaren, deren jüngstes Kind 3 bis 10 Jahre alt ist.

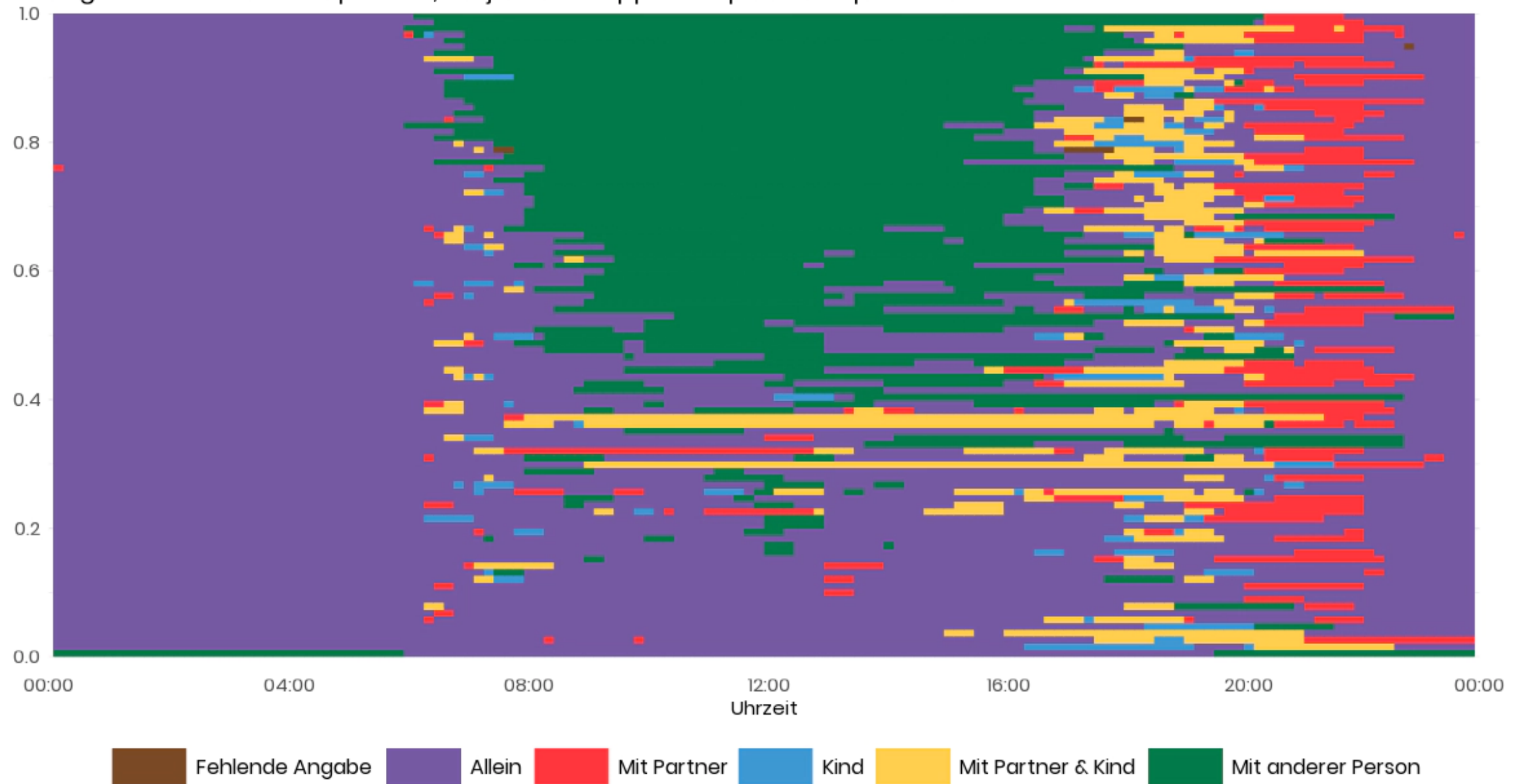


Daten: Zeitverwendungserhebung 2012/2013

Berechnungen: Florian Schulz

Mit wem verbringen Eltern ihre Zeit?

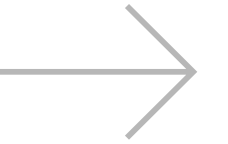
Die Abbildungen basiert auf vorläufigen Analysen von 954 Tagebucheinträgen für Werkzeuge von 494 Paaren, deren jüngstes Kind 3 bis 10 Jahre alt ist. Die 100 horizontalen Linien sind sogenannte Medoidsequenzen, die jeweils knapp 10 Sequenzen repräsentieren



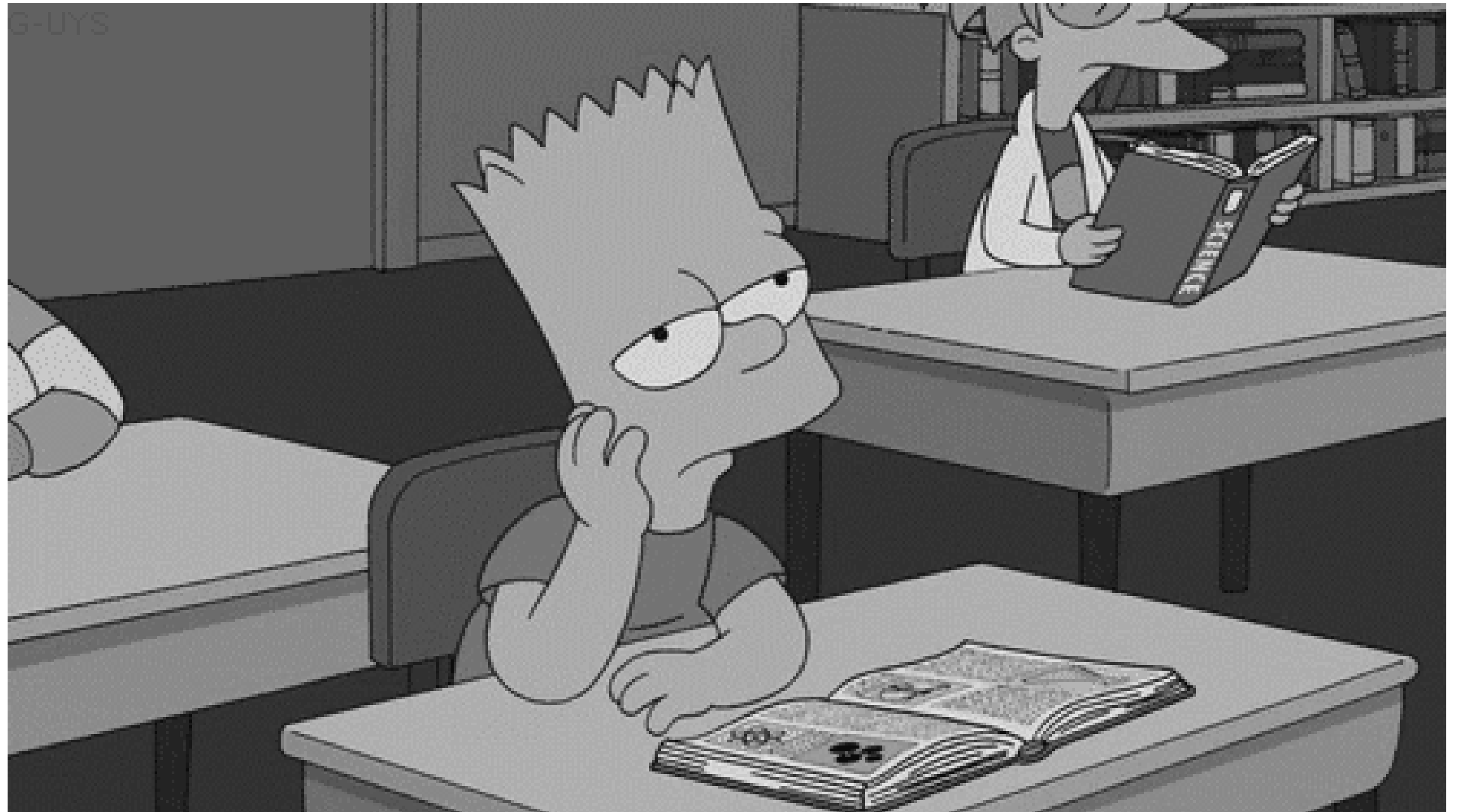
Daten: Zeitverwendungserhebung 2012/2013

Berechnungen: Florian Schulz

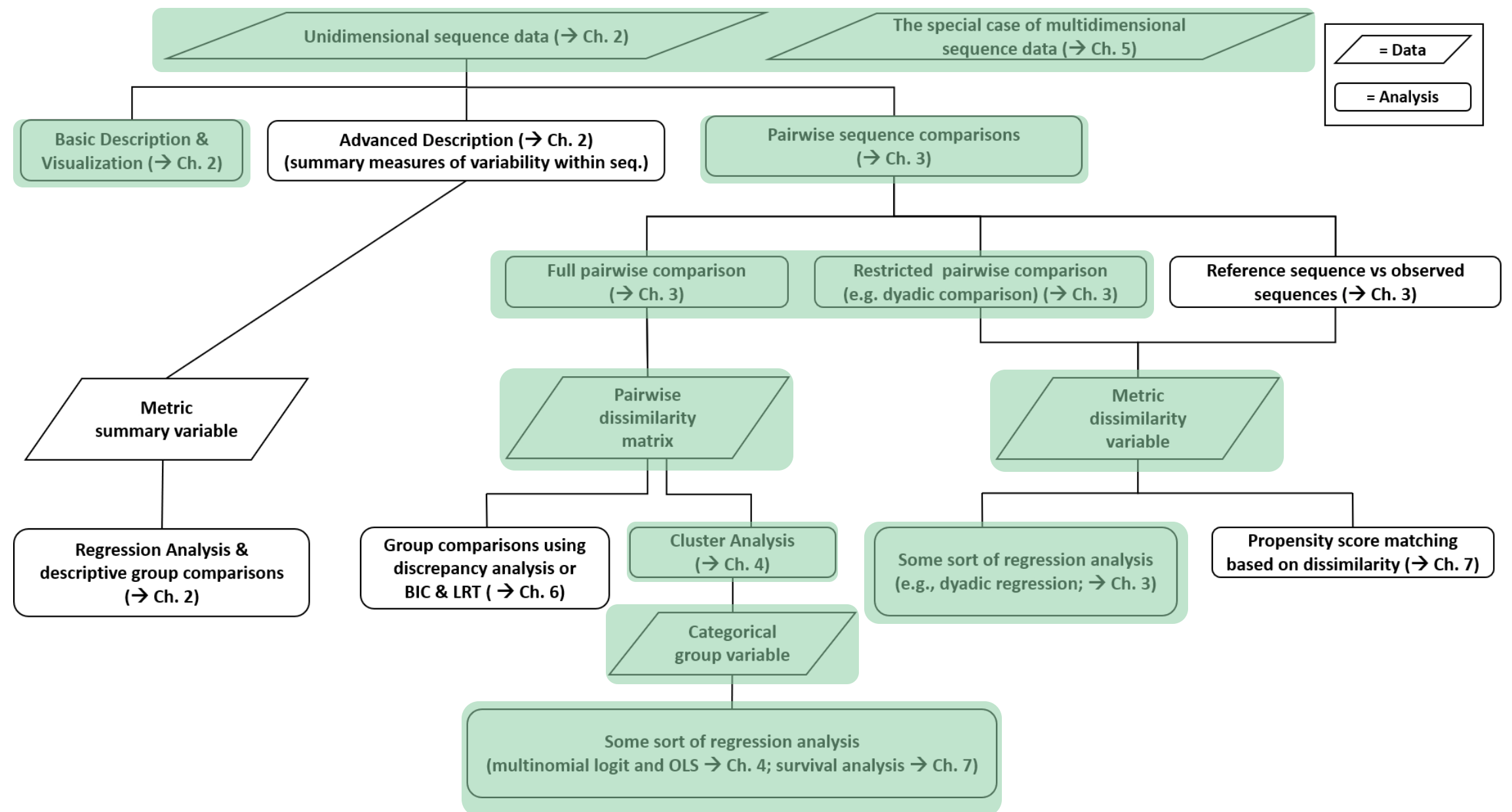
Kritik & Alternative Ansätze



War das
schon alles?



War das schon alles?



1

Vorbehalte gegenüber Distanzberechnung

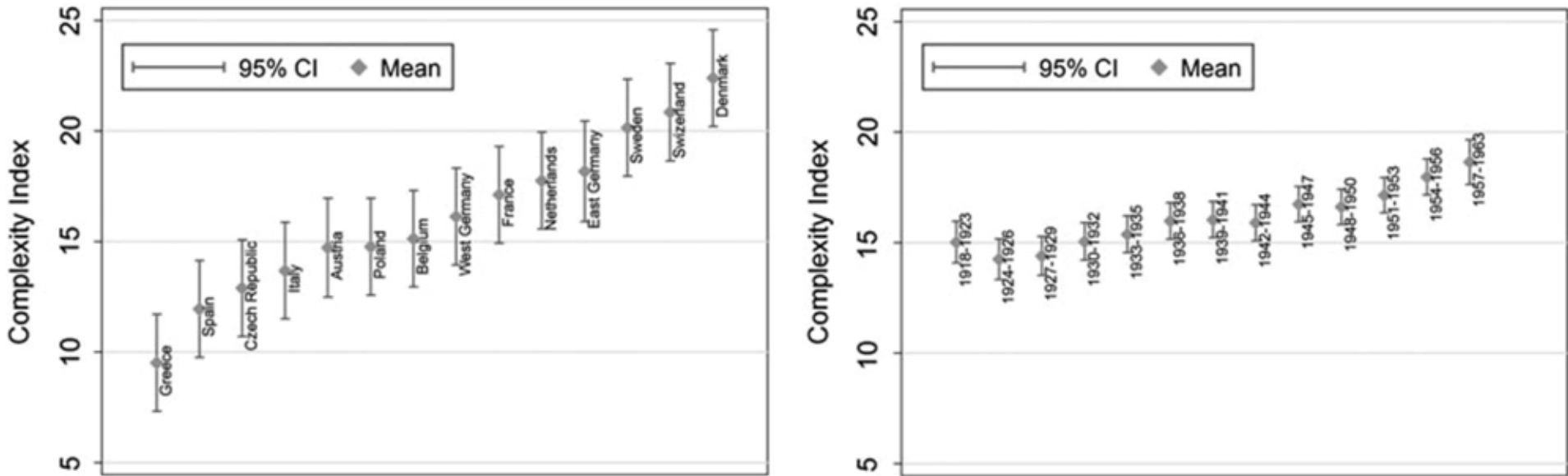
Verwendung von Indizes

Komplexität & Qualität der Sequenzen

Komplexität



Figure 5. Mean employment complexity by country and cohort calculated from model 1 in table 3



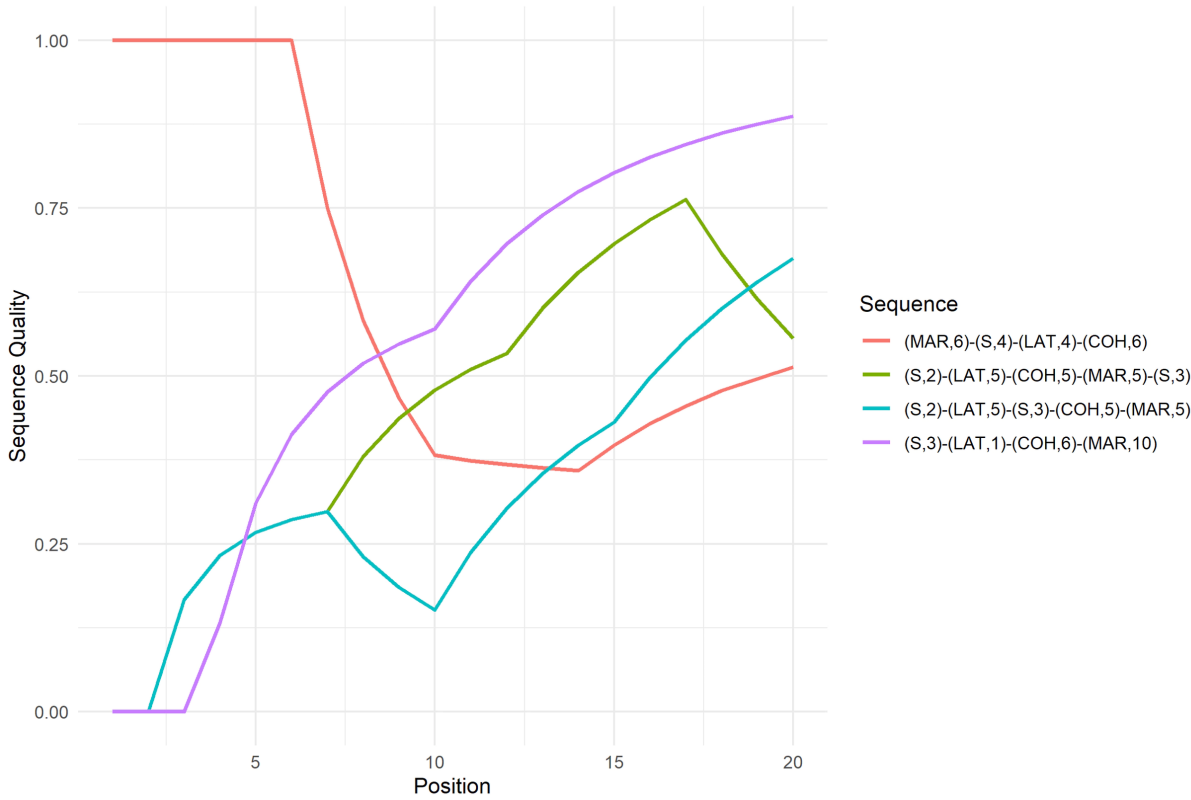
Van Winkle & Fasang 2017

Qualität

erfordert viele Annahmen



<https://github.com/maraab23/seqquality>



2

Vorbehalte gegenüber Clusteranalyse

Gruppenvergleiche ohne Clusteranalyse

Diskrepanzanalyse

Bayesianisches Informationskriterium & Likelihood-Quotienten-Test

Diskrepanz- analyse I

Table 6-1 Group comparisons of discrepancies and complexities

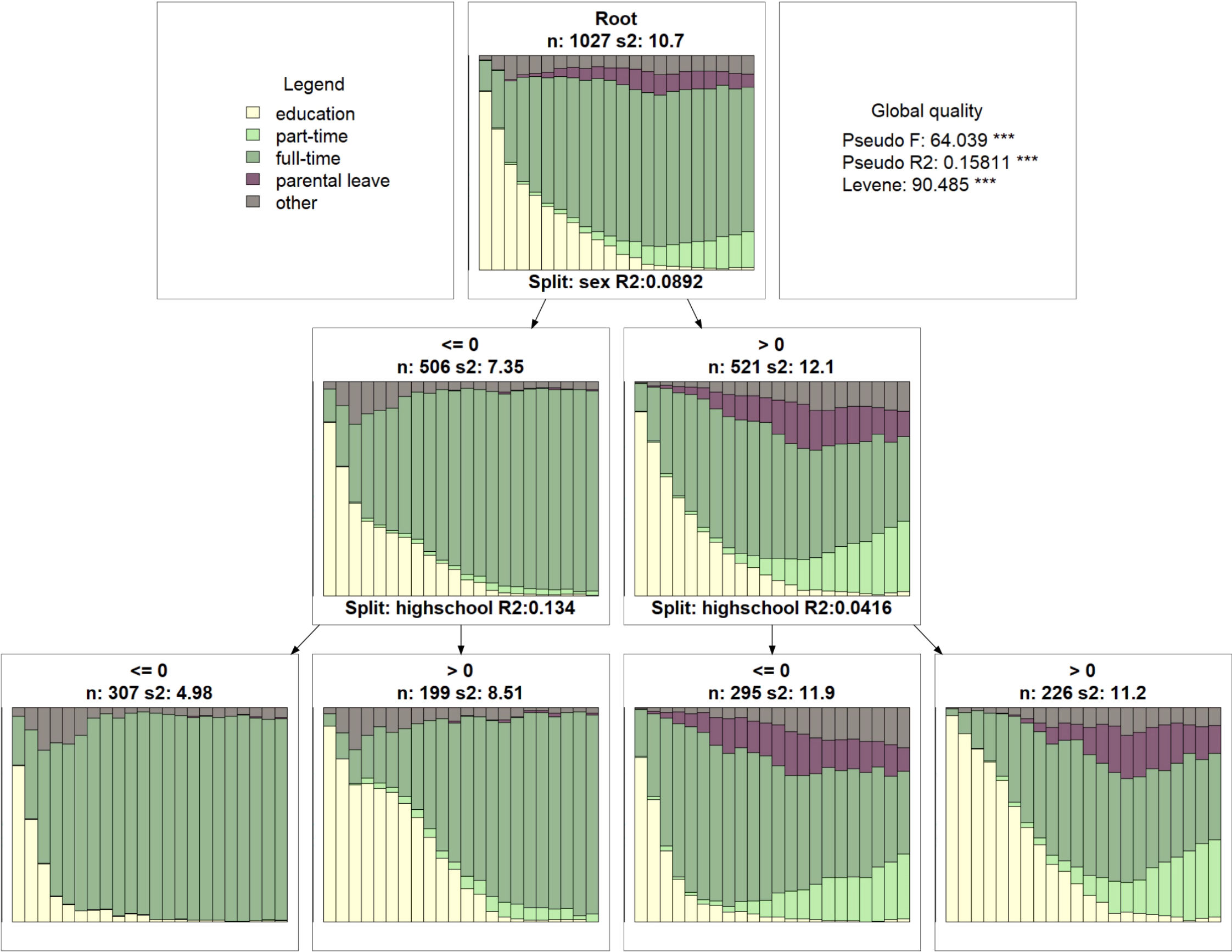
		Region			Gender			High school certificate		
		West	East	p-value ¹	Men	Women	p-value ¹	No	Yes	p-value ¹
Family sequences	N	1208	658		827	1039		1115	751	
	Discrepancy	13.80	14.79	0.001	13.67	14.41	0.001	14.84	13.15	0.001
	Complexity	0.33	0.32	0.013	0.33	0.32	0.075	0.31	0.35	0.000
Activity sequences	N	629	398		506	521		602	425	
	Discrepancy	10.75	10.46	0.507	7.35	12.11	0.001	9.54	10.65	0.011
	Complexity	0.22	0.23	0.200	0.17	0.27	0.000	0.20	0.25	0.000

¹Discrepancy: Levene-Test; Complexity: t-Test

Table 6-2 Associations of covariates with trajectories

	Family sequences (bivariate)		Activity sequences (bivariate)		Activity sequences (multivariate)	
	R ²	p-value	R ²	p-value	ΔR ²	p-value
Region	0.01	0.001	0.01	0.001	0.00	0.001
Gender	0.01	0.001	0.09	0.001	0.09	0.001
High-school certificate	0.01	0.001	0.07	0.001	0.06	0.001
Total					R ² _{total} 0.16	0.001

Diskrepanz- analyse II



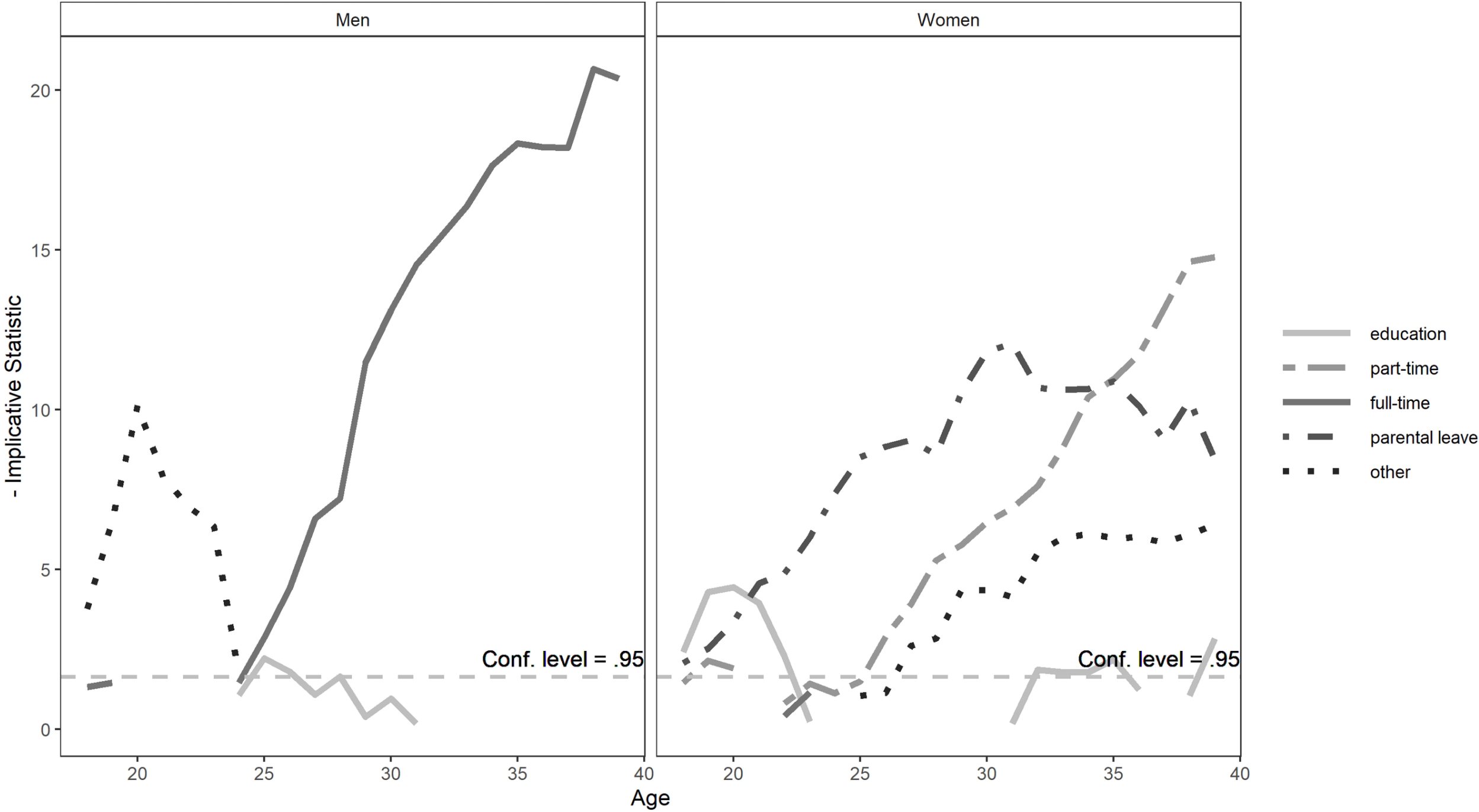
Diskrepanz- analyse III

BIC & LRT

Table 6-3 Group comparisons using different methods

	Δ BIC	LR-Test p-value	Pseudo- R ²	F-Test (perm.) p-value	N
West vs. East					
Overall	0.01	0.028	0.01	0.001	1027
Men	3.69	0.013	0.00	0.019	506
Women	0.68	0.018	0.02	0.001	521
Men vs. Women					
Overall	24.85	0.000	0.09	0.001	1027
West	26.72	0.000	0.11	0.001	629
East	14.41	0.000	0.07	0.001	398
No high school certificate	50.67	0.000	0.12	0.001	602
West	62.98	0.000	0.17	0.001	338
East	19.91	0.000	0.09	0.001	264
High school certificate	17.40	0.000	0.06	0.001	425
West	17.80	0.000	0.08	0.001	291
East	10.09	0.000	0.04	0.001	134

Implicative Statistic



3

**Fehlende Möglichkeiten zur
Berücksichtigung zeitveränderlicher Variablen und
Identifikation kausaler Effekte**

Zwei Ansätze

Verknüpfung von SA & EHA

Statt Analyse von diskreten Übergängen
Untersuchung von Übergängen in Sub-Episoden

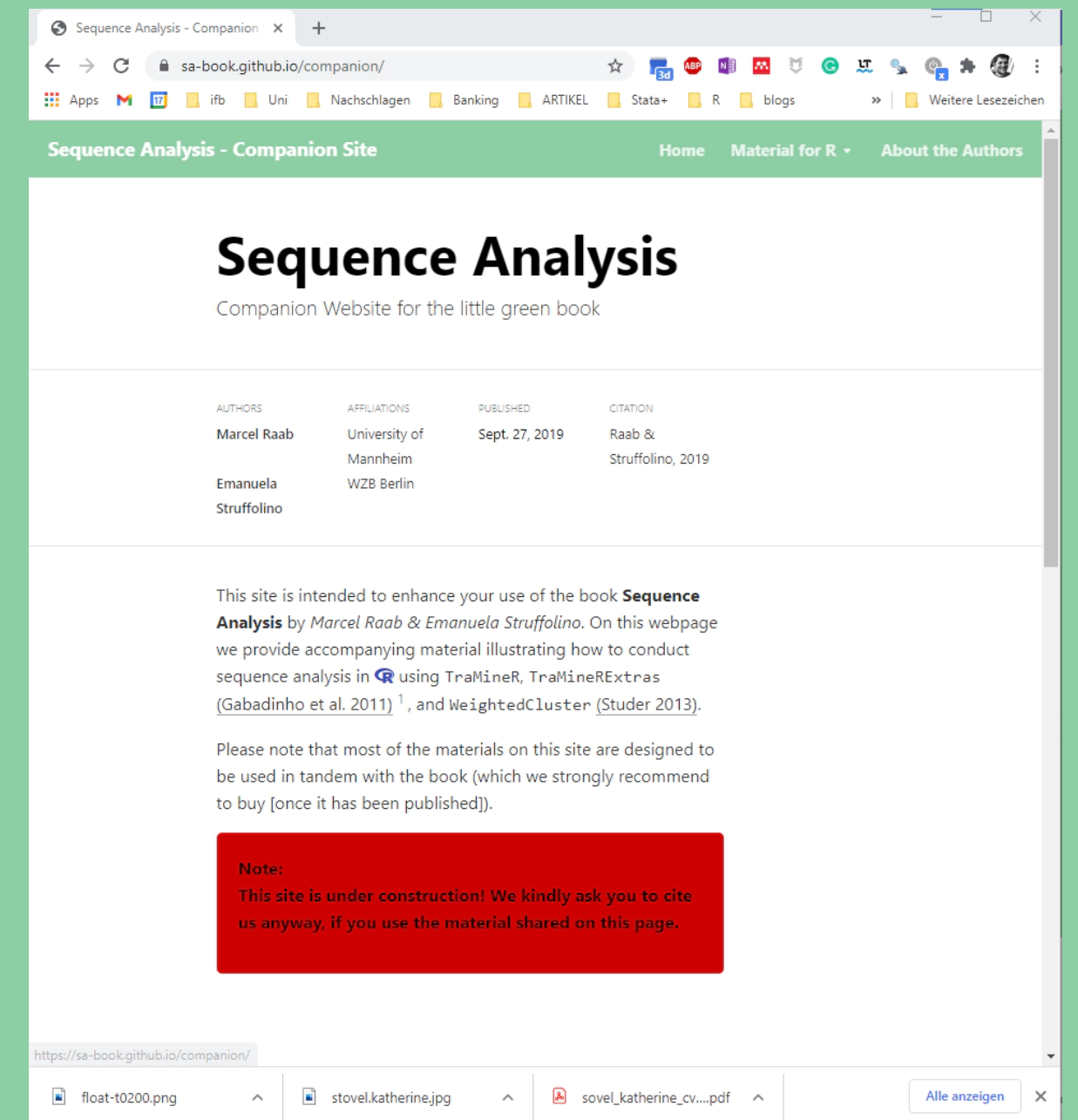
Kausales Matching mit OM

Statistische Zwillinge werden über Optimal Matching anstelle
eines Propensity Scores ermittelt

beide Verfahren können auch kombiniert werden

Vielen Dank

marcel.raab@ifb-uni-bamberg.de
marcelraab.de



<https://sa-book.github.io/companion/>