

**Methodendokumentation
der kriminologischen Schülerbefragung
in Duisburg 2002 bis 2019
– Dreizehn-Wellen-Panel –**

Anke Erdmann

**Schriftenreihe:
Jugendkriminalität in der modernen Stadt – Methoden
Nr. 27 / 2021**

ISSN 1610-2819

Informationen zur Schriftenreihe Jugendkriminalität in der modernen Stadt – Methoden

Herausgeber:

Prof. Dr. Klaus Boers

Institut für Kriminalwissenschaften

Abteilung Kriminologie

Bispinghof 24/25

48143 Münster

und

Prof. Dr. Jost Reinecke

Fakultät für Soziologie

Universität Bielefeld

Postfach 10 01 31

33 501 Bielefeld

Internet-Adressen

<http://www.jura.uni-muenster.de/kriminologie>

<http://www.uni-bielefeld.de/soz/krimstadt>

<http://www.krimstadt.de>

<http://www.crimoc.org>

ISSN 1610-2819

Übersicht der bisherigen Titel der Reihe:

1. Motzke, Katharina / Wittenberg, Jochen (1/2004): Methodendokumentation der kriminologischen Schülerbefragung in Münster 2000.
2. Wittenberg, Jochen (2/2004): Methodendokumentation der kriminologischen Schülerbefragung in Münster 2001.
3. Wittenberg, Jochen / Hilfert, Nicole (3/2004): Methodendokumentation der kriminologischen Schülerbefragung in Bocholt 2001.
4. Wittenberg, Jochen (4/2004): Methodendokumentation der kriminologischen Schülerbefragung in Münster 2002.
5. Motzke, Katharina / Brondies, Marc (5/2004): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002.
6. Brondies, Marc (6/2004): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2003.
7. Wittenberg, Jochen (7/2004): Methodendokumentation der kriminologischen Schülerbefragung in Münster 2003.
8. Brondies, Marc (8/2004): Methodendokumentation der Lehrerbefragung an Münsteraner und Duisburger Schulen 2003. Erhebung durchgeführter Präventionsmaßnahmen.
9. Pöge, Andreas (9/2005): Methodendokumentation der kriminologischen Schülerbefragung in Münster 2000–2003 (Vier-Wellen-Panel).
10. Hilfert, Nicole (10/2005): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2004.
11. Kunadt, Susann (11/2006): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2005.
12. Bentrup, Christina (12/2007): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2006.
13. Pöge, Andreas (13/2007): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002 bis 2005 (Vier-Wellen-Panel).
14. Pollich, Daniela (14/2007): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002 und 2003 (Zwei-Wellen-Panel).
15. Bentrup, Christina (15/2009): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2007.

-
16. Pollich, Daniela (16/2010): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002 bis 2007 (Sechs-Wellen-Panel).
 17. Bentrup, Christina (17/2010): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2008.
 18. Bentrup, Christina (18/2012): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2009.
 19. Schulte, Philipp (19/2014): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002 bis 2009 (Integriertes Hell- und Dunkelfeld-Panel).
 20. Bentrup, Christina / Verneuer, Lena (20/2014): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2011.
 21. Verneuer, Lena (21/2015): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2013.
 22. Verneuer, Lena (22/2017): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2015.
 23. Daniel, Andreas / Erdmann, Anke (23/2017): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002 bis 2013 (Zehn-Wellen-Panel).
 24. Bentrup, Christina / Schmid, Jule / Tzvetkova, Milena / Vreden, Carina (24/2018): Codebuch der kriminologischen Schülerbefragung in Münster (2000-2003).
 25. Kessler, Georg (25/2019): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2017.
 26. Kessler, Georg (26/2021): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2019.
 27. **Erdmann, Anke (27/2021): Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002 bis 2019 (Dreizehn-Wellen-Panel).**

Inhaltsverzeichnis

1	Einleitung	1
2	Studiendesign	2
3	Erstellung der Paneldaten	5
4	Realisierte Zuordnungen	15
4.1	Die verwendeten Datensätze und ihre Kurzbezeichnungen	15
4.2	Abschätzung und Bewertung der Zuordnungsgüte	18
5	Zusammensetzung der Paneldaten	20
5.1	Die Panelquerschnitte	20
5.1.1	Panelquerschnitt 2015 ($P_{Q_{t_{11}}}$)	21
5.1.2	Panelquerschnitt 2017 ($P_{Q_{t_{12}}}$)	21
5.1.3	Panelquerschnitt 2019 ($P_{Q_{t_{13}}}$)	22
5.2	Die Zwei-Wellen-Panel	23
5.2.1	Zwei-Wellen-Panel 2013 und 2015 ($P_{t_{10,11}}$)	23
5.2.2	Zwei-Wellen-Panel 2015 und 2017 ($P_{t_{11,12}}$)	24
5.2.3	Zwei-Wellen-Panel 2017 und 2019 ($P_{t_{12,13}}$)	25
5.3	Die Drei-Wellen-Panel	26
5.3.1	Drei-Wellen-Panel 2011, 2013 und 2015 ($P_{t_{9,10,11}}$)	26
5.3.2	Drei-Wellen-Panel 2013, 2015 und 2017 ($P_{t_{10,11,12}}$)	27
5.3.3	Drei-Wellen-Panel 2015, 2017 und 2019 ($P_{t_{11,12,13}}$)	28
5.4	Die Vier-Wellen-Panel	29
5.4.1	Vier-Wellen-Panel 2009, 2011, 2013 und 2015 ($P_{t_{8,9,10,11}}$)	29
5.4.2	Vier-Wellen-Panel 2011, 2013, 2015 und 2017 ($P_{t_{9,10,11,12}}$)	31
5.4.3	Vier-Wellen-Panel 2013, 2015, 2017 und 2019 ($P_{t_{10,11,12,13}}$)	32
5.5	Die Fünf-Wellen-Panel	33
5.5.1	Fünf-Wellen-Panel 2008, 2009, 2011, 2013 und 2015 ($P_{t_{7,8,9,10,11}}$)	33
5.5.2	Fünf-Wellen-Panel 2009, 2011, 2013, 2015 und 2017 ($P_{t_{8,9,10,11,12}}$)	35
5.5.3	Fünf-Wellen-Panel 2011, 2013, 2015, 2017 und 2019 ($P_{t_{9,10,11,12,13}}$)	37
5.6	Die Sechs-Wellen-Panel	39
5.6.1	Sechs-Wellen-Panel 2007, 2008, 2009, 2011, 2013 und 2015 ($P_{t_{6,7,8,9,10,11}}$)	39

5.6.2	Sechs-Wellen-Panel 2008, 2009, 2011, 2013, 2015 und 2017 ($P_{t_{7,8,9,10,11,12}}$)	41
5.6.3	Sechs-Wellen-Panel 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_{8,9,10,11,12,13}}$)	44
5.7	Die Sieben-Wellen-Panel	45
5.7.1	Sieben-Wellen-Panel 2006, 2007, 2008, 2009, 2011, 2013 und 2015 ($P_{t_{5,6,7,8,9,10,11}}$)	45
5.7.2	Sieben-Wellen-Panel 2007, 2008, 2009, 2011, 2013, 2015 und 2017 ($P_{t_{6,7,8,9,10,11,12}}$)	48
5.7.3	Sieben-Wellen-Panel 2008, 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_{7,8,9,10,11,12,13}}$)	50
5.8	Die Acht-Wellen-Panel	53
5.8.1	Acht-Wellen-Panel 2005, 2006, 2007, 2008, 2009, 2011, 2013 und 2015 ($P_{t_{4,5,6,7,8,9,10,11}}$)	53
5.8.2	Acht-Wellen-Panel 2006, 2007, 2008, 2009, 2011, 2013, 2015 und 2017 ($P_{t_{5,6,7,8,9,10,11,12}}$)	55
5.8.3	Acht-Wellen-Panel 2007, 2008, 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_{6,7,8,9,10,11,12,13}}$)	58
5.9	Die Neun-Wellen-Panel	60
5.9.1	Neun-Wellen-Panel 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013 und 2015 ($P_{t_{3,4,5,6,7,8,9,10,11}}$)	60
5.9.2	Neun-Wellen-Panel 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015 und 2017 ($P_{t_{4,5,6,7,8,9,10,11,12}}$)	64
5.9.3	Neun-Wellen-Panel 2006, 2007, 2008, 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_{5,6,7,8,9,10,11,12,13}}$)	66
5.10	Die Zehn-Wellen-Panel	69
5.10.1	Zehn-Wellen-Panel 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013 und 2015 ($P_{t_{2,3,4,5,6,7,8,9,10,11}}$)	69
5.10.2	Zehn-Wellen-Panel 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015 und 2017 ($P_{t_{3,4,5,6,7,8,9,10,11,12}}$)	73
5.10.3	Zehn-Wellen-Panel 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_{4,5,6,7,8,9,10,11,12,13}}$)	77
5.11	Die Elf-Wellen-Panel	79
5.11.1	Elf-Wellen-Panel 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013 und 2015 ($P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$)	79
5.11.2	Elf-Wellen-Panel 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015 und 2017 ($P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$)	84
5.11.3	Elf-Wellen-Panel 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$)	86

5.12 Die Zwölf-Wellen-Panel	89
5.12.1 Zwölf-Wellen-Panel 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015 und 2017 ($P_{t_{1,2,3,4,5,6,7,8,9,10,11,12}}$)	89
5.12.2 Zwölf-Wellen-Panel 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$)	95
5.13 Das Dreizehn-Wellen-Panel 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_{1,2,3,4,5,6,7,8,9,10,11,12,13}}$)	100
6 Zusammenfassung und Erklärung	105
Literaturverzeichnis	108
A Anhang	110
A.1 Codeblätter	110
A.2 Häufigkeiten der Codevariablen	113
A.3 Häufigkeiten der Gesamtcodes	121

Tabellenverzeichnis

3.1	Durchgeführte Kontrollschritte 2013/2015	8
3.2	Durchgeführte Kontrollschritte 2015/2017 – Eltern	8
3.3	Durchgeführte Kontrollschritte 2015/2017 – Nicht-Eltern	9
3.4	Durchgeführte Kontrollschritte 2017/2019	9
3.5	Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2013/2015	11
3.6	Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2015/2017 – Eltern	12
3.7	Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2015/2017 – Nicht-Eltern	13
3.8	Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2017/2019	14
4.2	Ausschöpfungen der durchgehenden Paneldatensätze	19
5.1	Zusammensetzung nach Geschlecht $P_{Q_{t_{11}}}$	21
5.2	Altersdurchschnitt $P_{Q_{t_{11}}}$	21
5.3	Zusammensetzung nach Geschlecht $P_{Q_{t_{12}}}$	22
5.4	Altersdurchschnitt $P_{Q_{t_{12}}}$	22
5.5	Zusammensetzung nach Geschlecht $P_{Q_{t_{13}}}$	22
5.6	Altersdurchschnitt $P_{Q_{t_{13}}}$	23
5.7	Zusammensetzung nach Geschlecht $P_{t_{10,11}}$	23
5.8	Altersdurchschnitt $P_{t_{10,11}}$	24
5.9	Zusammensetzung nach Geschlecht $P_{t_{11,12}}$	24
5.10	Altersdurchschnitt $P_{t_{11,12}}$	25
5.11	Zusammensetzung nach Geschlecht $P_{t_{12,13}}$	25
5.12	Altersdurchschnitt $P_{t_{12,13}}$	25
5.13	Zusammensetzung nach Geschlecht $P_{t_{9,10,11}}$	26
5.14	Altersdurchschnitt $P_{t_{9,10,11}}$	27
5.15	Zusammensetzung nach Geschlecht $P_{t_{10,11,12}}$	27
5.16	Altersdurchschnitt $P_{t_{10,11,12}}$	28
5.17	Zusammensetzung nach Geschlecht $P_{t_{11,12,13}}$	28
5.18	Altersdurchschnitt $P_{t_{11,12,13}}$	29
5.19	Zusammensetzung nach Geschlecht $P_{t_{8,9,10,11}}$	30
5.20	Altersdurchschnitt $P_{t_{8,9,10,11}}$	30
5.21	Zusammensetzung $P_{t_{8,9,10,11}}$ nach Schulform (Zeilenproz.)	31

5.22	Zusammensetzung nach Geschlecht $P_{t_9,10,11,12}$	31
5.23	Altersdurchschnitt $P_{t_9,10,11,12}$	32
5.24	Zusammensetzung nach Geschlecht $P_{t_{10,11,12,13}}$	32
5.25	Altersdurchschnitt $P_{t_{10,11,12,13}}$	33
5.26	Zusammensetzung nach Geschlecht $P_{t_{7,8,9,10,11}}$	34
5.27	Altersdurchschnitt $P_{t_{7,8,9,10,11}}$	35
5.28	Zusammensetzung $P_{t_{7,8,9,10,11}}$ nach Schulform (Zeilenproz.)	35
5.29	Zusammensetzung nach Geschlecht $P_{t_{8,9,10,11,12}}$	36
5.30	Altersdurchschnitt $P_{t_{8,9,10,11,12}}$	37
5.31	Zusammensetzung $P_{t_{8,9,10,11,12}}$ nach Schulform (Zeilenproz.)	37
5.32	Zusammensetzung nach Geschlecht $P_{t_{9,10,11,12,13}}$	38
5.33	Altersdurchschnitt $P_{t_{9,10,11,12,13}}$	38
5.34	Zusammensetzung nach Geschlecht $P_{t_{6,7,8,9,10,11}}$	39
5.35	Altersdurchschnitt $P_{t_{6,7,8,9,10,11}}$	40
5.36	Zusammensetzung $P_{t_{6,7,8,9,10,11}}$ nach Schulform (Zeilenproz.)	41
5.37	Zusammensetzung nach Geschlecht $P_{t_{7,8,9,10,11,12}}$	42
5.38	Altersdurchschnitt $P_{t_{7,8,9,10,11,12}}$	42
5.39	Zusammensetzung $P_{t_{7,8,9,10,11,12}}$ nach Schulform (Zeilenproz.)	43
5.40	Zusammensetzung nach Geschlecht $P_{t_{8,9,10,11,12,13}}$	44
5.41	Altersdurchschnitt $P_{t_{8,9,10,11,12,13}}$	45
5.42	Zusammensetzung $P_{t_{8,9,10,11,12,13}}$ nach Schulform (Zeilenproz.)	45
5.43	Zusammensetzung nach Geschlecht $P_{t_{5,6,7,8,9,10,11}}$	46
5.44	Altersdurchschnitt $P_{t_{5,6,7,8,9,10,11}}$	47
5.45	Zusammensetzung $P_{t_{5,6,7,8,9,10,11}}$ nach Schulform (Zeilenproz.)	47
5.46	Zusammensetzung nach Geschlecht $P_{t_{6,7,8,9,10,11,12}}$	48
5.47	Altersdurchschnitt $P_{t_{6,7,8,9,10,11,12}}$	49
5.48	Zusammensetzung $P_{t_{6,7,8,9,10,11,12}}$ nach Schulform (Zeilenproz.)	50
5.49	Zusammensetzung nach Geschlecht $P_{t_{7,8,9,10,11,12,13}}$	51
5.50	Altersdurchschnitt $P_{t_{7,8,9,10,11,12,13}}$	51
5.51	Zusammensetzung $P_{t_{7,8,9,10,11,12,13}}$ nach Schulform (Zeilenproz.)	52
5.52	Zusammensetzung nach Geschlecht $P_{t_{4,5,6,7,8,9,10,11}}$	53
5.53	Altersdurchschnitt $P_{t_{4,5,6,7,8,9,10,11}}$	54
5.54	Zusammensetzung $P_{t_{4,5,6,7,8,9,10,11}}$ nach Schulform (Zeilenproz.)	55
5.55	Zusammensetzung nach Geschlecht $P_{t_{5,6,7,8,9,10,11,12}}$	56
5.56	Altersdurchschnitt $P_{t_{5,6,7,8,9,10,11,12}}$	57
5.57	Zusammensetzung $P_{t_{5,6,7,8,9,10,11,12}}$ nach Schulform (Zeilenproz.)	57
5.58	Zusammensetzung nach Geschlecht $P_{t_{6,7,8,9,10,11,12,13}}$	58
5.59	Altersdurchschnitt $P_{t_{6,7,8,9,10,11,12,13}}$	59
5.60	Zusammensetzung $P_{t_{6,7,8,9,10,11,12,13}}$ nach Schulform (Zeilenproz.)	60
5.61	Zusammensetzung nach Geschlecht $P_{t_{3,4,5,6,7,8,9,10,11}}$	61

5.62	Altersdurchschnitt $P_{t_{3,4,5,6,7,8,9,10,11}}$	62
5.63	Zusammensetzung $P_{t_{3,4,5,6,7,8,9,10,11}}$ nach Schulform (Zeilenproz.)	63
5.64	Zusammensetzung nach Geschlecht $P_{t_{4,5,6,7,8,9,10,11,12}}$	64
5.65	Altersdurchschnitt $P_{t_{4,5,6,7,8,9,10,11,12}}$	65
5.66	Zusammensetzung $P_{t_{4,5,6,7,8,9,10,11,12}}$ nach Schulform (Zeilenproz.)	66
5.67	Zusammensetzung nach Geschlecht $P_{t_{5,6,7,8,9,10,11,12,13}}$	67
5.68	Altersdurchschnitt $P_{t_{5,6,7,8,9,10,11,12,13}}$	68
5.69	Zusammensetzung $P_{t_{5,6,7,8,9,10,11,12,13}}$ nach Schulform (Zeilenproz.)	68
5.70	Zusammensetzung nach Geschlecht $P_{t_{2,3,4,5,6,7,8,9,10,11}}$	70
5.71	Altersdurchschnitt $P_{t_{2,3,4,5,6,7,8,9,10,11}}$	71
5.72	Zusammensetzung $P_{t_{2,3,4,5,6,7,8,9,10,11}}$ nach Schulform (Zeilenproz.)	72
5.73	Zusammensetzung nach Geschlecht $P_{t_{3,4,5,6,7,8,9,10,11,12}}$	74
5.74	Altersdurchschnitt $P_{t_{3,4,5,6,7,8,9,10,11,12}}$	75
5.75	Zusammensetzung $P_{t_{3,4,5,6,7,8,9,10,11,12}}$ nach Schulform (Zeilenproz.)	76
5.76	Zusammensetzung nach Geschlecht $P_{t_{4,5,6,7,8,9,10,11,12,13}}$	77
5.77	Altersdurchschnitt $P_{t_{4,5,6,7,8,9,10,11,12,13}}$	78
5.78	Zusammensetzung $P_{t_{4,5,6,7,8,9,10,11,12,13}}$ nach Schulform (Zeilenproz.)	79
5.79	Zusammensetzung nach Geschlecht $P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$	81
5.80	Altersdurchschnitt $P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$	82
5.81	Zusammensetzung $P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$ nach Schulform (Zeilenproz.)	83
5.82	Zusammensetzung nach Geschlecht $P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	84
5.83	Altersdurchschnitt $P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	85
5.84	Zusammensetzung $P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$ nach Schulform (Zeilenproz.)	86
5.85	Zusammensetzung nach Geschlecht $P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	87
5.86	Altersdurchschnitt $P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	88
5.87	Zusammensetzung $P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$ nach Schulform (Zeilenproz.)	89
5.88	Zusammensetzung nach Geschlecht $P_{t_{1,2,3,4,5,6,7,8,9,10,11,12}}$	92
5.89	Altersdurchschnitt $P_{t_{1,2,3,4,5,6,7,8,9,10,11,12}}$	93
5.90	Zusammensetzung $P_{t_{1,2,3,4,5,6,7,8,9,10,11,12}}$ nach Schulform (Zeilenproz.)	94
5.91	Zusammensetzung nach Geschlecht $P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	97
5.92	Altersdurchschnitt $P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	98
5.93	Zusammensetzung $P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$ nach Schulform (Zeilenproz.)	99
5.94	Zusammensetzung nach Geschlecht $P_{t_{1,2,3,4,5,6,7,8,9,10,11,12,13}}$	102
5.95	Altersdurchschnitt $P_{t_{1,2,3,4,5,6,7,8,9,10,11,12,13}}$	103
5.96	Zusammensetzung $P_{t_{1,2,3,4,5,6,7,8,9,10,11,12,13}}$ nach Schulform (Zeilenproz.)	104
A.1	Häufigkeiten von Co001	113
A.2	Häufigkeiten von Co002	114
A.3	Häufigkeiten von Co003	115
A.4	Häufigkeiten von Co004	116

A.5	Häufigkeiten von Co005	117
A.6	Häufigkeiten von Co006	118
A.7	Häufigkeiten von Co011	119
A.8	Häufigkeiten von Co007 in Duisburg 2015	120
A.9	Häufigkeiten von Co007 in Duisburg 2017	120
A.10	Häufigkeiten von Co007 in Duisburg 2019	120
A.11	Häufigkeiten der Gesamtcodes in Duisburg 2015	121
A.12	Häufigkeiten der Gesamtcodes in Duisburg 2017	121
A.13	Häufigkeiten der Gesamtcodes in Duisburg 2019	121

Abbildungsverzeichnis

2.1	Erhebungsdesign der Studie „Kriminalität in der modernen Stadt“	4
A.1	Codeblatt Duisburg 2015	110
A.2	Codeblatt Duisburg 2017	111
A.3	Codeblatt Duisburg 2019	112

1 Einleitung

Die kriminologische und kriminalsoziologisch orientierte Studie *Kriminalität in der modernen Stadt*¹ hat zum Ziel, die Entwicklung und Entstehung delinquenter Handlungsmuster über die Zeit zu beschreiben und analytisch zu erklären. Ausgehend von einem theoretischen Analysemodell (vgl. Boers & Reinecke, 2019) werden die zu untersuchenden Phänomene sowohl aus der Perspektive der Makro- wie auch der Meso- und Mikroebene analysiert. Methodisch zeichnet sich die Studie durch die Verwendung eines kombinierten Panel-Kohorten-Designs aus, welches neben der Untersuchung individueller Entwicklungen über die Zeit zugleich die Betrachtung jahrgangsspezifischer Effekte ermöglicht. Kern der Erhebung war zunächst das Dunkelfeld der Delinquenz in Form von Selbstberichten. Zusätzlich zu diesen wurden in den eingesetzten Erhebungsinstrumenten unter anderem die Bereiche Erziehungsstile, Konfliktverhalten, Kriminalitätseinstellungen, Lebens-, Freizeit- und Konsumstile erfasst. Aufgrund des verwendeten Paneldesigns wurde der größte Teil der Fragen über alle Erhebungswellen unverändert beibehalten. Detailliertere Informationen zum Studiendesign sind den entsprechenden Methodenberichten und Kapitel 2 zu entnehmen. Da in diesem Methodenbericht die Dokumentation des *Panel*datensatzes der jüngeren Kohorte im Mittelpunkt steht, wird auf eine Beschreibung der Querschnittsdatsätze verzichtet. Auch hierfür können genauere Informationen den jeweiligen Methodenberichten zu den Querschnittsdatsätzen entnommen werden. Ebenso werden Ergebnisse und Analysen der Studie nicht dargestellt. Zusammenfassende Darstellungen der inhaltlichen Auswertungen finden sich unter anderem bei Boers et al. (2014) sowie bei Boers und Reinecke (2019).

1 Projektleitung: Prof. Dr. Klaus Boers, Institut für Kriminalwissenschaften, WWU Münster und Prof. Dr. Jost Reinecke, Fakultät für Soziologie, Universität Bielefeld, aktuelle DFG-Förderzeichen: Bo 1234/6-10 und Re 832/6-7

2 Studiendesign

Die Studie *Kriminalität in der modernen Stadt* wurde zunächst zwischen den Jahren 2000 und 2003 jährlich in der westfälischen Universitätsstadt Münster durchgeführt. Die Erhebung erfolgte hier ausschließlich im schulischen Kontext und umfasste die Schuljahre sieben bis zehn. Weitere Querschnittserhebungen unterschiedlicher Vergleichsjahrgänge erfolgten in Münster und Bocholt (siehe auch Abbildung 2.1).

Ab dem Jahr 2002 wurde in Duisburg eine weitere Erhebung begonnen. Ziel war eine Vollerhebung des gesamten siebten und neunten Jahrgangs. Während die Schülerinnen und Schüler der ehemals neunten Jahrgangsstufe lediglich ein weiteres Mal (2003) befragt wurden², erfolgte die Befragung der ehemaligen Siebtklässler/innen bis zum Jahr 2019. Der jährliche Erhebungsrhythmus wurde dabei bis zum Jahr 2009 beibehalten und darauffolgend auf einen zweijährigen Rhythmus umgestellt. Die weiteren Erhebungen erfolgten somit in den Jahren 2011, 2013, 2015, 2017 und 2019. Nachdem in den *ersten beiden Erhebungen* in den Jahren 2002 und 2003 noch eine Vollerhebung des jeweiligen Jahrgang angestrebt wurde (Motzke & Brondies, 2004; Brondies, 2004), konzentrierten sich die Erhebungen in den Folgejahren auf die Wiederbefragung der bisherigen Teilnehmerinnen und Teilnehmer. Die Erhebung im Klassenverband war in den Jahren 2002 bis 2005 dabei der einzige Erhebungsmodus. Mit dem zu erwartenden Abschluss der zehnten Klasse im Jahr 2005 (*vierte Erhebung*) wurde jedoch damit begonnen, eine Adressdatenbank der Befragten aufzubauen, um sie auch nach dem Beenden der sekundären Schullaufbahn noch erreichen zu können. Um einen möglichst großen Anteil jener Personen zu erreichen, die ab dem Jahr 2006 einen berufsschulischen Bildungsgang belegten, wurde die *fünfte Erhebung* im Jahr 2006 ebenfalls an Duisburger Berufskollegs durchgeführt. Da diese Befragung zu einer erheblichen Vergrößerung des Querschnittsdatensatzes führte, die neu hinzugekommenen Personen jedoch für die Konstruktion der Paneldatensätze keine Relevanz hatten, wurde in den folgenden Jahren auf eine weitere Erhebung in den Berufsschulen verzichtet. Ab dem Jahr 2006 wurde des Weiteren sowohl eine Befragung in Jugendzentren durchgeführt als auch mit einer postalischen Erhebung in Form einer Nachbefragung begonnen (Bentrup, 2007). Ab dem Jahr 2007 (*sechste Erhebungswelle*) wurde die Befragung zunehmend auf einen postalischen Modus umgestellt. Zusätzlich wurden jene Personen, die weder postalisch noch im Klassenverband befragt werden konnten, mittels einer Face-to-Face-Befragung kontaktiert (Bentrup, 2009). Die *siebte Erhebung* im Jahr 2008 wurde erstmals ausschließlich im postalischen Modus durchgeführt und durch eine anschließende

² Hieraus entstand ein Zwei-Wellen-Paneldatensatz, siehe Pollich (2007).

Face-to-Face-Befragung ergänzt. Diese kombinierte Erhebungsform wurde ab diesem Zeitpunkt für alle folgenden Befragungswellen fortgesetzt.

Mit dieser Umstellung des Erhebungsmodus stieg auch der zeitliche und administrative Aufwand der Befragung, sodass nach der Erhebung in 2009 eine Umstellung auf einen zweijährlichen Erhebungsrhythmus erfolgte. Im Jahr 2010 erfolgte daher erstmals eine Erhebungspause des zuvor jährlichen Befragungsturnus. Die *neunte* Erhebungswelle folgte somit im Jahr 2011 (Bentrop & Verneuer, 2014). In diesem Jahr wurde die Entlohnung im Rahmen der postalischen Erhebung von Bargeld auf Verrechnungsschecks umgestellt (die Entlohnung in der Face-to-Face-Befragung erfolgte weiterhin in bar). In diesem Zweijahresrhythmus wurden vier weitere Wellen erhoben: die *zehnte* Welle in 2013, die *elfte* Welle in 2015, die *zwölfte* Welle in 2017 und schließlich die *dreizehnte* und letzte Welle in 2019 (Verneuer, 2015; 2017; Kessler, 2019; 2021).

Somit liegen für die Duisburger Befragung insgesamt 13 Erhebungswellen vor, welche die Altersspanne vom 13. bis zum 30. Lebensjahr abdecken. Der vorliegende Bericht beschreibt die Entstehung und Repräsentativität des finalen Dreizehn-Wellen-Paneldatensatzes (2002 bis 2019). Dieser Bericht stellt damit die Fortführung der Berichte zum Vier-, Sechs- bzw. Zehn-Wellen-Paneldatensatz dar. Dementsprechend werden nur jene Datensätze behandelt, die nicht bereits in den vorangegangenen Berichten beschrieben wurden – also Datensätze, welche die Wellen 2015, 2017 und 2019 enthalten (d. h., die seit dem Bericht zum Zehn-Wellen-Panel neu erhobenen Wellen). Aufgrund der Tatsache, dass die Anzahl der sich ergebenden Datensätze mit jeder Welle exponentiell steigt, ist es aus Platzgründen nicht mehr möglich, jede mögliche Wellenkombination zu beschreiben. Dargestellt werden somit nur durchgehende Paneldatensätze mit aufeinanderfolgenden Wellen, also nur Paneldatensätze ohne Lücken.

3 Erstellung der Paneldaten

Mit der Durchführung einer Panelstudie sind mit der erforderlichen Wiederbefragung derselben Personen spezifische Probleme verbunden: Wie können Personen in den Befragungswellen der verschiedenen Jahre unter Berücksichtigung der Anonymität identifiziert werden und wie können die von ein und derselben Person ausgefüllten Fragebögen für die spätere Analyse zusammengebracht werden? Im Forschungsprojekt *Kriminalität in der modernen Stadt* entschied man sich für ein Verfahren, welches mit Hilfe eines persönlichen Codes die Fragebogenzuordnungen zwischen den Bögen der unterschiedlichen Erhebungsjahre leisten soll.

Um eine Zuordnung der Fragebögen aus den verschiedenen Erhebungswellen zu ermöglichen, wurden Codes verwendet, die über Codeblätter abgefragt wurden. Jede/r Befragte füllte im Zuge des Interviews ein Codeblatt aus, auf dem fünf, sechs (ab der Befragung in 2003) oder sieben (ab 2011) persönliche Fragen gestellt wurden, aus deren Antwort jeweils ein bestimmter Buchstabe oder eine Zahl notiert werden sollte. Die Fragen beziehen sich auf unveränderliche Merkmale einer Befragungsperson oder ihres Umfeldes (natürliche Haarfarbe, Name des Vaters etc.). Diese Buchstaben-Zahlen-Kombination bildet schließlich den gesamten Code. In jeder Erhebungswelle wurden die Codeblätter den Befragten zum Ausfüllen wieder vorgelegt, sodass bei einem stabilen Antwortverhalten die Codeblätter, die von ein und derselben Person in den verschiedenen Wellen ausgefüllt wurden, denselben Code aufweisen müssten.

Im Zuge der Befragungen sollte auf den Codeblättern³ Folgendes notiert werden:

- *Co001*: Der erste Buchstabe des Vornamens des Vaters
- *Co002*: Der erste Buchstabe des Vornamens der Mutter
- *Co003*: Der erste Buchstabe des eigenen Vornamens
- *Co004*: Die beiden Tagesziffern des eigenen Geburtstages
- *Co005*: Der letzte Buchstabe der eigenen Haarfarbe
- *Co006*: Der letzte Buchstabe der eigenen Augenfarbe
- *Co007*: Der letzte Buchstabe des eigenen Nachnamens⁴

Aufgrund der Erfahrungen mit dem Code und dem Zuordnungsverfahren aus den Münsteraner Befragungen (vgl. Pöge, 2005a, S. 7 ff.; Pöge, 2005b) wurden auf den

3 Abdrucke der Codeblätter der Erhebungsjahre 2015, 2017 und 2019 finden sich in den Abbildungen A.1, A.2 und A.3 im Anhang. Codeblätter früherer Erhebungen finden sich bei Pöge (2007, S. 88 ff.), Pollich (2010, S. 183 ff.) sowie bei Daniel und Erdmann (2017, S. 109 ff.)

4 Die Frage wurde erst im Jahr 2009 hinzugefügt. In den jeweiligen Datensätzen trägt sie den Variablennamen Co011. Co007 steht hier für die siebte *Codestelle*.

Duisburger Codeblättern zum einen einfachere Fragen mit einer gleichzeitig größeren Antwortvarianz gestellt, zum anderen wurde im Jahr 2003 der Code – ebenso wie in Münster – zunächst um eine Frage erweitert und das Layout des gesamten Blattes überarbeitet. Das Layout wurde dergestalt geändert, dass kein handschriftliches Ausfüllen mehr erfolgen musste, sondern alle Antwortvorgaben zum Ankreuzen aufgeführt wurden. Im Jahr 2009 wurde eine weitere Frage hinzugefügt (Co007). Aufgrund der Beendigung der Schullaufbahn der meisten Befragten spätestens im Jahr 2008, blieb von den bisher verwendeten zusätzlichen Fragen (zu einem Schulwechsel bzw. einer Nichtversetzung im vergangenen Jahr) lediglich die Frage nach der Befragungsteilnahme im Vorjahr⁵ erhalten (vgl. auch Pollich, 2010, S. 8).

- Co008: Befragungsteilnahme im Vorjahr (ja/nein)⁶

Zusätzlich zu den sieben Codefragen und der Zusatzfrage über die Teilnahme beim letzten Mal⁷ wurde für die Fragebogenzuordnungen das Geschlecht der Befragten sowie die (zuletzt) besuchte Schule⁸, die mit einer Kennnummer erhoben wurde, verwendet.

Bei der technischen Durchführung der Zuordnungen der Fragebögen aus den einzelnen Wellen traten prinzipiell die schon aus der Münsteraner Studie und den früheren Duisburger Panelwellen bekannten Probleme auf, sie konnten durch die oben beschriebenen Modifikationen und Verbesserungen jedoch deutlich abgeschwächt werden (vgl. Pöge, 2008, S. 60).⁹ Als generelle Problembereiche des Zuordnungsverfahrens waren schon in Münster die Probleme der *Identifizierung* und der *Reproduktion* der Codes aufgetreten. Während das Problem der Identifizierung durch Hinzunahme einer Codefrage weitgehend gelöst werden konnte (vgl. Pöge 2007, S. 6, 2008, S. 62 f. und siehe auch die geringen Häufigkeiten doppelt vorkommender Codes in den Tabellen A.11 bis A.13 im Anhang), bleibt das Problem der Reproduktion, also dass Befragte in den verschiedenen Wellen einen nicht übereinstimmenden Code angeben, prinzipiell bestehen. Jedoch konnte auch die Tragweite dieses Problems reduziert werden (vgl. Pöge, 2005a, S. 7 f.; Pöge, 2005b, S. 66).

In den ersten vier Wellen der jüngeren Duisburger Kohorte liegt der Anteil der Bögen mit exakt übereinstimmenden Codes an *allen gefundenen* passenden Zuordnungen recht stabil bei rund 60 % (berechnet nach Pöge, 2007, S. 7). In den Zuordnungsschritten 2005/2006 sowie 2006/2007 liegt dieser Wert bei 61,0 % bzw. bei 68,0 % (vgl. Pollich,

5 Um Unklarheiten durch die Verwendung des Begriffs „im Vorjahr“ beim zweijährigen Befragungsrhythmus zu beseitigen, wurde die Formulierung im Jahr 2019 auf „an der letzten Befragung 2017“ geändert.

6 In den jeweiligen Datensätzen trägt diese Codestelle den *Variablen*namen Co007.

7 Die Häufigkeitsverteilungen sämtlicher verwendeter Codes und Zusatzfragen finden sich in den Tabellen A.1 bis A.10 im Anhang.

8 25 Schulen aus Duisburg + Kategorie *eine andere Schule*.

9 Für eine weitergehende Darstellung des anonymisierten Zuordnungsverfahrens, sowie einen Vergleich mit unverschlüsselter Zuordnung, siehe Pöge (2011).

2010, S. 8). Für den Zuordnungsschritt 2008/2009 kann die Rate der exakten Code-reproduktion auf 71 % gesteigert werden, bevor sie in den Schritten 2009/2011 und 2011/2013 auf etwa 64 % abfällt. Dieser recht starke Rückgang der Reproduktionsgüte kann vor allem durch die Umstellung auf einen Zweijahresrhythmus erklärt werden. In den folgenden Erhebungen erhöht und stabilisiert sich die Reproduktionsgüte und liegt für die Zuordnungsschritte 2013/2015, 2015/2017 sowie 2017/2019 bei jeweils um 69 %.

Das zur Zuordnung der Bögen aus den unterschiedlichen Befragungsjahren angewendete *fehlertolerante Verfahren mit manuellem Handschriftenvergleich* bestand aus vier Hauptschritten: In einem ersten Schritt wurden maschinell alle *exakt übereinstimmenden Codes* aus zwei Erhebungswellen ermittelt. Die zusammengehörigen Fragebögen wurden daraufhin einer manuellen Handschriftenkontrolle unterzogen. Diejenigen Bogenpaare, die offensichtlich von derselben Person ausgefüllt wurden, wurden aus den Datensätzen genommen, sodass sie für die nachfolgenden Zuordnungsschritte nicht mehr zur Verfügung standen, da sie bereits als zusammengehörig identifiziert werden konnten. Nicht passende Bögen verblieben in den Datensätzen, um möglicherweise in einem der nächsten Zuordnungsschritte als passend identifiziert zu werden. Im zweiten Schritt wurde nach Codeübereinstimmungen unter Zulassung *eines Fehlers* und im dritten Schritt unter Zulassung von *zwei Fehlern* im Code gesucht, die zugehörigen Bogennummern herausgeschrieben und manuell kontrolliert. In einem letzten, vierten Schritt wurden die Kontrollen schließlich unter Zulassen von *drei Fehlern* im Code vorgenommen. Bei jedem Hauptkontrollschritt mussten zunächst neben den Codes (mit der jeweils zugelassenen Fehleranzahl) auch ausgewählte Zusatzmerkmale übereinstimmen (Geschlecht, Teilnahme beim letzten Mal, zuletzt besuchte Schule). Anschließend wurden in jedem Kontrollschritt – unter Beibehaltung der vorgegebenen Fehlerzahl im persönlichen Code – die zu erfüllenden Zusatzbedingungen sukzessive gelockert und teilweise ganz weggelassen. Das geschilderte Verfahren ist damit codegeleitet und hierarchisch: Das erste Kriterium für eine Zuordnung ist der Code, denn es werden nur offensichtliche Fehlzuordnungen ausgesondert. Überdies wird durch die schrittweise Durchführung gewährleistet, dass Zuordnungsmöglichkeiten mit weniger Fehlern im Code bevorzugt werden.

Eine Übersicht über die durchgeführten Zuordnungsschritte und die Bedingungen, die der Code erfüllen musste, findet sich in den Tabellen 3.1 bis 3.4. Dabei ist für die Zuordnung zwischen den Jahren 2015 und 2017 eine Besonderheit zu beachten: Aufgrund damals aktueller inhaltlicher Analysen war es nötig, die im Datensatz enthaltenen *Elternteile* zuerst ans Panel anzubinden, damit für diese Teilpopulation der Paneldatensatz früher zur Verfügung stand. Aus diesem Grund erfolgte die Zuordnung in 2015/2017 in zwei Schritten, nämlich erst *Eltern* und im Anschluss *Nicht-Eltern*. Die Zuordnungsschritte unterscheiden sich zwischen beiden Teildatensätzen etwas, da aufgrund der unterschied-

Tabelle 3.1: Durchgeführte Kontrollschritte 2013/2015

Schritt	Fehler im Code	Zusatzvariablen		
		Geschlecht	Teilnahme letztes Mal	Schule
S1	ohne Fehler	passt	ja	passt
S2	ohne Fehler	offen	offen	offen
S3	ein Fehler	passt	ja	passt
S4	ein Fehler	offen	offen	offen
S5	zwei Fehler	passt	ja	passt
S6	zwei Fehler	offen	offen	offen
S7	drei Fehler	passt	ja	passt
S8	drei Fehler	passt	ja	offen
S9	drei Fehler	passt	offen	offen

Tabelle 3.2: Durchgeführte Kontrollschritte 2015/2017 – Eltern

Schritt	Fehler im Code	Zusatzvariablen		
		Geschlecht	Teilnahme letztes Mal	Schule
S1	ohne Fehler	passt	ja	passt
S2	ohne Fehler	offen	offen	offen
S3	ein Fehler	passt	ja	passt
S4	ein Fehler	offen	offen	offen
S5	zwei Fehler	passt	ja	passt
S6	zwei Fehler	offen	offen	offen
S7	drei Fehler	passt	ja	offen
S8	drei Fehler	passt	offen	offen

lich großen Fallzahl die sukzessive Lockerung der betrachteten Zusatzmerkmale anders gestaltet werden musste.

In den Tabellen 3.5 bis 3.8 sind jeweils die Anzahlen der in den einzelnen Schritten durchgeführten Handschriftenkontrollen sowie die jeweiligen Anteile passender und nicht passender Bogenpaare aufgeführt. Für die Zuordnung 2013/2015 liegt der Anteil passender Bögen an den Gesamtkontrollen bei 69 %. Für die Zuordnung zwischen 2015 und 2017 (*Eltern* und *Nicht-Eltern* zusammengenommen) liegt der Anteil hingegen bei lediglich 55 %. Bei der Zuordnung 2017/2019 liegt der Anteil mit 75 % sehr hoch. Auch wenn diese Zahlen sehr unterschiedlich ausfallen, ist jedoch zu beachten, dass die absolute Anzahl an passenden Fragebogenpaaren bei jeder Erhebung tatsächlich etwa gleich groß ist (2013/2015: 2 335; 2015/2017: 2 339; 2017/2019: 2 346). Die absolute Anzahl insgesamt durchgeführter Kontrollen variiert jedoch teilweise stark. Insbesondere

Tabelle 3.3: Durchgeführte Kontrollschritte 2015/2017 – Nicht-Eltern

Schritt	Fehler im Code	Zusatzvariablen		
		Geschlecht	Teilnahme letztes Mal	Schule
S1	ohne Fehler	passt	ja	passt
S2	ohne Fehler	offen	offen	offen
S3	ein Fehler	passt	ja	passt
S4	ein Fehler	offen	offen	offen
S5	zwei Fehler	passt	ja	passt
S6	zwei Fehler	offen	offen	offen
S7	drei Fehler	passt	ja	passt
S8	drei Fehler	passt	ja	offen
S9	drei Fehler	passt	offen	offen
S10	drei Fehler	offen	offen	offen

Tabelle 3.4: Durchgeführte Kontrollschritte 2017/2019

Schritt	Fehler im Code	Zusatzvariablen		
		Geschlecht	Teilnahme letztes Mal	Schule
S1	ohne Fehler	passt	ja	passt
S2	ohne Fehler	offen	offen	offen
S3	ein Fehler	passt	ja	passt
S4	ein Fehler	offen	offen	offen
S5	zwei Fehler	passt	ja	passt
S6	zwei Fehler	offen	offen	offen
S7	drei Fehler	passt	ja	passt
S8	drei Fehler	passt	offen	offen

bei der Zuordnung zwischen 2015 und 2017, bei welcher eine Unterteilung in *Eltern* und *Nicht-Eltern* stattfand, wurden insgesamt deutlich mehr Kontrollen durchgeführt, sodass der relative Anteil gefundener Matches entsprechend kleiner ausfällt.

Auch für alle Kontrollen nach dem Jahr 2006 gilt, dass nicht direkt aufeinander folgende Erhebungswellen ebenfalls kontrolliert wurden. Nicht im Panel zugeordnete Bögen eines Querschnittsdatensatzes wurden stets auch mit den nicht zugeordneten Bögen aus dem *vorletzten* Jahr verglichen, um Personen, die nur in einem Befragungsjahr nicht teilgenommen haben, nicht endgültig aus dem Panel zu verlieren. Die Zuordnung dieser Paarungen erfolgte analog zu der oben beschriebenen Prozedur; aus Kapazitätsgründen

erfolgt keine gesonderte Darstellung.

Wie in den Jahren zuvor wurde nach Abschluss jeder Erhebung und vor Beginn der Abgleiche für das Panel zunächst überprüft, ob doppelte Teilnahmen innerhalb eines Jahres vorlagen. Zu diesem Zweck wurden mittels des beschriebenen hierarchischen Verfahrens vermeintliche Doppelungen identifiziert und bei Verifizierung der Doppelung ausgeschlossen.¹⁰ Insgesamt belief sich der Aufwand dieser Vorkontrollen im Jahr 2015 auf 1 404 durchgeführte Abgleiche, im Jahr 2017 auf 1 276 durchgeführte Abgleiche sowie im Jahr 2019 auf 1 146 durchgeführte Abgleiche.

Zusätzlich wurde ab dem Erhebungsjahr 2015 ein weiterer Schritt zur Verbesserung der Datenqualität eingeführt, welcher als „Fragebogen-Inventur“ bezeichnet werden kann. Bereits in den Vorjahren zeigte sich, dass stets ein gewisser Anteil an *fehlenden* Fragebögen vorliegt (Pöge, 2007; Pollich, 2010; Daniel & Erdmann, 2017) – also an Bögen, die während der Kontrollen nicht aufgefunden werden können. Ab der Welle des Jahres 2015 wurde daher vor Beginn der Kontrollen geprüft, welche Fragebögen tatsächlich vorliegen. Dazu wurden alle vorliegenden Fragebögen durchgegangen und in einer Liste abgehakt, welche Fragebogennummern von 0001 bis 5500 physisch vorhanden sind. Diese Liste an vorhandenen Bögen wurde mit den im Datensatz vorhandenen Fragebogennummern abgeglichen. Dabei zeigte sich, dass es Papierfragebögen gab, deren Nummern nicht im Datensatz auftauchten, bzw. umgekehrt, dass im Datensatz Fragebogennummern vorhanden waren, die nicht in Papierform vorlagen. Dabei war allerdings offensichtlich, dass es sich oftmals jeweils um „Zahlendreher“ voneinander handelte, d. h. es traten vereinzelt Fehler bei der Eingabe der Fragebogennummer auf. Die Fragebogennummern konnten somit schließlich korrigiert werden, sodass das Auftauchen vereinzelter, fehlender Fragebögen nahezu vollständig eliminiert werden konnte.

¹⁰ Dabei wurde in den früheren Wellen zunächst jener Fragebogen mit dem früheren Teilnahmezeitpunkt behalten. Ab dem Jahr 2011 wurde jener Fragebogen entfernt, der weniger brauchbare Informationen enthielt. Im Falle, dass beide Bögen von der Qualität her gleichwertig waren, wurde jeweils der neuere behalten.

Tabelle 3.5: Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2013/2015

(a) absolut durchgeführte Kontrollen pro Zuordnungsschritt

		Kontrollen gesamt	passt	passt nicht	fehlt ^a
S1	ohne Fehler	1 114	1 053	58	3
S2	ohne Fehler	594	573	17	4
	gesamt	1 708	1 626	75	7
S3	ein Fehler	315	307	7	1
S4	ein Fehler	238	212	22	4
	gesamt	553	519	29	5
S5	zwei Fehler	77	71	5	1
S6	zwei Fehler	177	59	113	5
	gesamt	254	130	118	6
S7	drei Fehler	86	27	58	1
S8	drei Fehler	538	19	516	3
S9	drei Fehler	242	14	228	0
	gesamt	866	60	802	4
Summe		3 381	2 335	1 024	

^a Die Bögen in dieser Spalte werden nicht zu einer Gesamtzahl aufsummiert, da als fehlend deklarierte Bögen im nächsten Kontrollschritt stets erneut einbezogen wurden, um Ausfälle möglichst zu minimieren.

(b) prozentuale Ergebnisse (Zeilenprozentuierung)

		Kontrollen gesamt	passt	passt nicht	fehlt ^a
S1	ohne Fehler	100,0	94,5	5,2	0,3
S2	ohne Fehler	100,0	96,5	2,9	0,7
	gesamt	100,0	95,2	4,4	0,4
S3	ein Fehler	100,0	97,5	2,2	0,3
S4	ein Fehler	100,0	89,1	9,2	1,7
	gesamt	100,0	93,9	5,2	0,9
S5	zwei Fehler	100,0	92,2	6,5	1,3
S6	zwei Fehler	100,0	33,3	63,8	2,8
	gesamt	100,0	51,2	46,5	2,4
S7	drei Fehler	100,0	31,4	67,4	1,2
S8	drei Fehler	100,0	3,5	95,9	0,6
S9	drei Fehler	100,0	5,8	94,2	0,0
	gesamt	100,0	6,9	92,6	0,5
Summe		100,0	69,1	30,3	

Tabelle 3.6: Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2015/2017 – Eltern

(a) absolut durchgeführte Kontrollen pro Zuordnungsschritt

		Kontrollen gesamt	passt	passt nicht	fehlt ^a
S1	ohne Fehler	264	238	26	0
S2	ohne Fehler	143	143	0	0
	gesamt	407	381	26	0
S3	ein Fehler	87	85	2	0
S4	ein Fehler	84	78	6	0
	gesamt	171	163	8	0
S5 ^b	zwei Fehler	—			
S6	zwei Fehler	227	60	167	0
	gesamt	227	60	167	0
S7	drei Fehler	630	13	617	0
S8	drei Fehler	354	2	352	0
	gesamt	984	15	969	0
Summe		1 789	619	1 170	

^a Die Bögen in dieser Spalte werden nicht zu einer Gesamtzahl aufsummiert, da als fehlend deklarierte Bögen im nächsten Kontrollschritt stets erneut einbezogen wurden, um Ausfälle möglichst zu minimieren.

^b Der Schritt S5 wurde übersprungen, da die Anzahl an Abgleichen bei der vorgesehenen Merkmalskombination sehr gering war. Die Abgleiche sind im Schritt S6 integriert.

(b) prozentuale Ergebnisse (Zeilenprozentuierung)

		Kontrollen gesamt	passt	passt nicht	fehlt ^a
S1	ohne Fehler	100,0	90,2	9,8	0,0
S2	ohne Fehler	100,0	100,0	0,0	0,0
	gesamt	100,0	93,6	6,4	0,0
S3	ein Fehler	100,0	97,7	2,3	0,0
S4	ein Fehler	100,0	92,9	7,1	0,0
	gesamt	100,0	95,3	4,7	0,0
S5 ^b	zwei Fehler	—			
S6	zwei Fehler	100,0	26,4	73,6	0,0
	gesamt	100,0	26,4	73,6	0,0
S7	drei Fehler	100,0	2,1	97,9	0,0
S8	drei Fehler	100,0	0,6	99,4	0,0
	gesamt	100,0	1,5	98,5	0,0
Summe		100,0	34,6	65,4	

Tabelle 3.7: Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2015/2017 – Nicht-Eltern

(a) absolut durchgeführte Kontrollen pro Zuordnungsschritt

		Kontrollen gesamt	passt	passt nicht	fehlt ^a
S1	ohne Fehler	846	844	2	0
S2	ohne Fehler	438	435	3	0
	gesamt	1 284	1 279	5	0
S3	ein Fehler	231	227	3	1
S4	ein Fehler	170	160	10	0
	gesamt	401	387	13	1
S5	zwei Fehler	47	44	3	0
S6	zwei Fehler	97	39	58	0
	gesamt	144	83	61	0
S7 ^b	drei Fehler	—			
S8	drei Fehler	272	25	247	0
S9	drei Fehler	155	5	150	0
S10	drei Fehler	279	1	278	0
	gesamt	706	31	675	0
Summe		2 535	1 780	754	

^a Die Bögen in dieser Spalte werden nicht zu einer Gesamtzahl aufsummiert, da als fehlend deklarierte Bögen im nächsten Kontrollschritt stets erneut einbezogen wurden, um Ausfälle möglichst zu minimieren.

^b Der Schritt S7 wurde übersprungen, da die Anzahl an Abgleichen bei der vorgesehenen Merkmalskombination sehr gering war. Die Abgleiche sind im Schritt S8 integriert.

(b) prozentuale Ergebnisse (Zeilenprozentuierung)

		Kontrollen gesamt	passt	passt nicht	fehlt ^a
S1	ohne Fehler	100,0	99,8	0,2	0,0
S2	ohne Fehler	100,0	99,3	0,7	0,0
	gesamt	100,0	99,6	0,4	0,0
S3	ein Fehler	100,0	98,3	1,3	0,4
S4	ein Fehler	100,0	94,1	5,9	0,0
	gesamt	100,0	96,5	3,2	0,3
S5	zwei Fehler	100,0	93,6	6,4	0,0
S6	zwei Fehler	100,0	40,2	59,8	0,0
	gesamt	100,0	57,6	42,4	0,0
S7 ^b	drei Fehler	—			
S8	drei Fehler	100,0	9,2	90,8	0,0
S9	drei Fehler	100,0	3,2	96,8	0,0
S10	drei Fehler	100,0	0,4	99,6	0,0
	gesamt	100,0	4,4	95,6	0,0
Summe		100,0	70,2	29,7	

Tabelle 3.8: Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2017/2019

(a) absolut durchgeführte Kontrollen pro Zuordnungsschritt

		Kontrollen gesamt	passt	passt nicht	fehlt ^a
S1	ohne Fehler	1 199	1 150	49	0
S2	ohne Fehler	495	485	10	0
	gesamt	1 694	1 635	59	0
S3	ein Fehler	363	352	10	1
S4	ein Fehler	205	186	17	2
	gesamt	568	538	27	3
S5 ^b	zwei Fehler	—			
S6	zwei Fehler	241	127	111	3
	gesamt	241	127	111	3
S7	drei Fehler	45	19	26	0
S8	drei Fehler	595	27	565	3
	gesamt	640	46	591	3
Summe		3 143	2 346	788	

^a Die Bögen in dieser Spalte werden nicht zu einer Gesamtzahl aufsummiert, da als fehlend deklarierte Bögen im nächsten Kontrollschritt stets erneut einbezogen wurden, um Ausfälle möglichst zu minimieren.

^b Der Schritt S5 wurde übersprungen, da die Anzahl an Abgleichen bei der vorgesehenen Merkmalskombination sehr gering war. Die Abgleiche sind im Schritt S6 integriert.

(b) prozentuale Ergebnisse (Zeilenprozentuierung)

		Kontrollen gesamt	passt	passt nicht	fehlt ^a
S1	ohne Fehler	100,0	95,9	4,1	0,0
S2	ohne Fehler	100,0	98,0	2,0	0,0
	gesamt	100,0	96,5	3,5	0,0
S3	ein Fehler	100,0	97,0	2,8	0,3
S4	ein Fehler	100,0	90,7	8,3	1,0
	gesamt	100,0	94,7	4,8	0,5
S5 ^b	zwei Fehler	—			
S6	zwei Fehler	100,0	52,7	46,1	1,2
	gesamt	100,0	52,7	46,1	1,2
S7	drei Fehler	100,0	42,2	57,8	0,0
S8	drei Fehler	100,0	4,5	95,0	0,5
	gesamt	100,0	7,2	92,3	0,5
Summe		100,0	74,6	25,1	

4 Realisierte Zuordnungen

Mit Hilfe des geschilderten Zuordnungsverfahrens konnten für die neu entstandenen Paneldatensätze die in Tabelle 4.1 aufgeführten Fallzahlen erreicht werden. Für längsschnittliche Analysen sind neben den Zwei-Wellen-Panels die *lückenlosen* Paneldatensätze mit drei und mehr Wellen am bedeutsamsten. Daher wird sich die Darstellung auf diese Datensätze konzentrieren. Das Resultat der vollständigen Erhebung in Duisburg ist ein Dreizehn-Wellen-Paneldatensatz der Jahre 2002 bis 2019 ($t_{1,2,3,4,5,6,7,8,9,10,11,12,13}$) mit 736 durchgehend¹¹ und 4 076 insgesamt (mit Lücken) enthaltenen Fällen.¹²

4.1 Die verwendeten Datensätze und ihre Kurzbezeichnungen

Im Folgenden sollen der Einfachheit halber Kurzbezeichnungen für die verwendeten Datensätze benutzt werden.

Die Bezeichnung » S_{t_x} « steht für die Daten der amtlichen Schulstatistik desjenigen Schuljahres, innerhalb dessen die jeweilige Befragung stattfand (Stichtag der Schulstatistik ist jeweils der 15. 10. des entsprechenden Schuljahres). Bei Angabe von Fallzahlen aus der Schulstatistik sind nur diejenigen Schulen berücksichtigt, die im jeweiligen Jahr an den Befragungen teilgenommen haben. Jedoch findet die Schulstatistik lediglich zur Abschätzung der Repräsentativität der ersten drei Erhebungswellen Verwendung, da ab t_4 auch Jugendliche befragt wurden, die ein Schuljahr wiederholen mussten, und somit die Zugehörigkeit aller Befragten zu *einer* Jahrgangsstufe nicht mehr gewährleistet ist (vgl. hierzu genauer Pöge, 2007, S. 11). Diese Maßnahme, die das Ziel hatte, möglichst alle bis dahin im Panel enthaltenen Personen wieder zu befragen, führte dazu, dass die Schulstatistik in diesem Erhebungsjahr nicht mehr als maßgebliche Grundgesamtheit zu betrachten ist.

Da seit der Erhebungswelle t_5 zudem sukzessive immer mehr Schülerinnen und Schüler die allgemeinbildenden Schulen verlassen haben und für die Berufsschulen keine hier verwendbaren amtlichen Statistiken verfügbar sind, wird das Heranziehen der Schulstatistik ab dem Befragungszeitpunkt t_5 obsolet (vgl. Bentrup, 2007, S. 5). Ab dem Zeitpunkt t_4 werden daher ausschließlich die Querschnittsdaten zur Beurteilung der Repräsentativität der einzelnen Paneldatensätze herangezogen. Zwar ist diese Grundlage

11 Berücksichtigt sind hier nur die gültigen Fälle, die auch den Plausibilitätskontrollen der einzelnen Querschnittsdatsätze standhielten.

12 Hierbei ist zu beachten, dass sich die Größe des Datensatzes mit Lücken im Vergleich zu Pollich (2010, S. 15) verändert hat. Dies ist auf eine nachträgliche Korrektur der Paneldaten zurückzuführen.

– insbesondere zu t_5 , wo beabsichtigte Abweichungen vorliegen – weniger verlässlich als die amtliche Statistik, jedoch lässt sich eine Grundgesamtheit wie in den ersten Wellen ohnehin nur noch schwerlich definieren. Es handelt sich bei den Querschnittsdaten um die einzig verfügbare Informationsquelle, die durchaus eine Abschätzung ermöglicht, ob Geschlechter oder Schulformen in den Paneldatensätzen systematisch verzerrt repräsentiert sind.

Ein $\gg Q_t \ll$ mit einem Zeitpunkt als Index steht für einen Querschnittsdatensatz aus einer spezifischen Erhebungswelle. Angegeben werden ausschließlich diejenigen Fälle, die den in jedem Erhebungsjahr durchgeführten Plausibilitätskontrollen standhielten. $Q_{t_{11}}$ beinhaltet zum Beispiel alle 2 754 gültigen Fälle des Querschnittsdatensatzes aus dem Jahr 2015.

Ein $\gg P_{Q_t} \ll$ mit nur einem Zeitpunkt als Index bezeichnet einen so genannten Panelquerschnitt. Diese Panelquerschnitte beinhalten *alle* vorhandenen Fälle des gesamten Dreizehn-Wellen-Paneldatensatzes (mit Lücken) aus *einem* Erhebungszeitpunkt. So enthält zum Beispiel $P_{Q_{t_{11}}}$ alle 1 812 gültigen Fälle die aus t_{11} im Gesamtpaneldatensatz enthalten sind. Dies sind im Unterschied zum Querschnitt $Q_{t_{11}}$ alle Fälle, die zum Zeitpunkt t_{11} befragt wurden und im Zuge der Erstellung des Dreizehn-Wellen-Panels *mindestens einem* Fall eines beliebigen anderen Zeitpunktes zugeordnet werden konnten.

Die eigentlichen Paneldatensätze werden durch ein $\gg P_{t_x, y, \dots} \ll$ mit einem Index, der die jeweils einbezogenen Zeitpunkte umfasst, bezeichnet. So steht beispielsweise $P_{t_{10, 11}}$ für den Zwei-Wellen-Paneldatensatz mit allen 1 699 *gültigen* Fällen, die zwischen t_{10} und t_{11} einander zugeordnet werden konnten. Sind mehr als zwei Zeitpunkte im Index vermerkt, so enthält der betreffende Datensatz genau so viele Erhebungswellen wie Zeitpunkte aufgeführt sind. Zu den angegebenen Zeitpunkten sind im betreffenden Datensatz dann alle Fälle dieser Erhebungswellen *durchgängig* vorhanden.

In Tabelle 4.1 sind *alle* Teildatensätze und deren Fallzahlen aufgeführt, die in diesem Bericht dargestellt werden. Aus Kapazitätsgründen können in diesem Methodenbericht nur jene Datensätze beschrieben werden, die *keine* Lücken aufweisen. Weitere Darstellungen von Datensätzen finden sich bei Pöge (2007), Pollich (2010) sowie bei Daniel und Erdmann (2017).

Tabelle 4.1: Bezeichnungen und Fallzahlen aller betrachteten Panel- und Referenzdatensätze 2002 bis 2019

Abkürzung	Fallzahl	Beschreibung
S_{t_1}	3 910	Amtliche Schulstatistik, Schuljahr 2001/2002 (Stichtag: 15. 10.) ^a
S_{t_2}	3 763	Amtliche Schulstatistik, Schuljahr 2002/2003 (Stichtag: 15. 10.) ^a
S_{t_3}	3 925	Amtliche Schulstatistik, Schuljahr 2003/2004 (Stichtag: 15. 10.) ^a
Q_{t_1}	3 411	Querschnittsdaten des ersten Befragungsjahres 2002
Q_{t_2}	3 392	Querschnittsdaten des zweiten Befragungsjahres 2003
Q_{t_3}	3 339	Querschnittsdaten des dritten Befragungsjahres 2004
Q_{t_4}	3 405	Querschnittsdaten des vierten Befragungsjahres 2005
Q_{t_5}	4 548	Querschnittsdaten des fünften Befragungsjahres 2006
Q_{t_6}	3 336	Querschnittsdaten des sechsten Befragungsjahres 2007
Q_{t_7}	3 086	Querschnittsdaten des siebten Befragungsjahres 2008
Q_{t_8}	3 090	Querschnittsdaten des achten Befragungsjahres 2009
Q_{t_9}	3 050	Querschnittsdaten des neunten Befragungsjahres 2011
$Q_{t_{10}}$	2 849	Querschnittsdaten des zehnten Befragungsjahres 2013
$Q_{t_{11}}$	2 754	Querschnittsdaten des elften Befragungsjahres 2015
$Q_{t_{12}}$	2 778	Querschnittsdaten des zwölften Befragungsjahres 2017
$Q_{t_{13}}$	2 698	Querschnittsdaten des dreizehnten Befragungsjahres 2019
$P_{Q_{t_{11}}}$	1 812	Panelquerschnittsdaten des Jahres 2015
$P_{Q_{t_{12}}}$	1 760	Panelquerschnittsdaten des Jahres 2017
$P_{Q_{t_{13}}}$	1 671	Panelquerschnittsdaten des Jahres 2019
$P_{t_{10,11}}$	1 699	Paneldaten der Jahre 2013 und 2015
$P_{t_{11,12}}$	1 686	Paneldaten der Jahre 2015 und 2017
$P_{t_{12,13}}$	1 604	Paneldaten der Jahre 2017 und 2019
$P_{t_{9,10,11}}$	1 625	Paneldaten der Jahre 2011, 2013, 2015
$P_{t_{10,11,12}}$	1 596	Paneldaten der Jahre 2013, 2015, 2017
$P_{t_{11,12,13}}$	1 549	Paneldaten der Jahre 2015, 2017, 2019
$P_{t_{8,9,10,11}}$	1 573	Paneldaten der Jahre 2009, 2011, 2013, 2015
$P_{t_{9,10,11,12}}$	1 527	Paneldaten der Jahre 2011, 2013, 2015, 2017
$P_{t_{10,11,12,13}}$	1 474	Paneldaten der Jahre 2013, 2015, 2017, 2019
$P_{t_{7,8,9,10,11}}$	1 520	Paneldaten der Jahre 2008, 2009, 2011, 2013, 2015
$P_{t_{8,9,10,11,12}}$	1 480	Paneldaten der Jahre 2009, 2011, 2013, 2015, 2017
$P_{t_{9,10,11,12,13}}$	1 414	Paneldaten der Jahre 2011, 2013, 2015, 2017, 2019

Tabelle 4.1: Bezeichnungen und Fallzahlen aller betrachteten Panel- und Referenzdatensätze 2002 bis 2019

$P_{6,7,8,9,10,11}$	1 479	Paneldaten der Jahre 2007, 2008, 2009, 2011, 2013, 2015
$P_{7,8,9,10,11,12}$	1 430	Paneldaten der Jahre 2008, 2009, 2011, 2013, 2015, 2017
$P_{8,9,10,11,12,13}$	1 372	Paneldaten der Jahre 2009, 2011, 2013, 2015, 2017, 2019
$P_{5,6,7,8,9,10,11}$	1 419	Paneldaten der Jahre 2006, 2007, 2008, 2009, 2011, 2013, 2015
$P_{6,7,8,9,10,11,12}$	1 394	Paneldaten der Jahre 2007, 2008, 2009, 2011, 2013, 2015, 2017
$P_{7,8,9,10,11,12,13}$	1 327	Paneldaten der Jahre 2008, 2009, 2011, 2013, 2015, 2017, 2019
$P_{4,5,6,7,8,9,10,11}$	1 285	Paneldaten der Jahre 05, 06, 07, 08, 09, 11, 13, 15
$P_{5,6,7,8,9,10,11,12}$	1 339	Paneldaten der Jahre 06, 07, 08, 09, 11, 13, 15, 17
$P_{6,7,8,9,10,11,12,13}$	1 294	Paneldaten der Jahre 07, 08, 09, 11, 13, 15, 17, 19
$P_{3,4,5,\dots,9,10,11}$	1 163	Paneldaten der Jahre 04, 05, 06, 07, 08, 09, 11, 13, 15
$P_{4,5,6,\dots,10,11,12}$	1 214	Paneldaten der Jahre 05, 06, 07, 08, 09, 11, 13, 15, 17
$P_{5,6,7,\dots,11,12,13}$	1 246	Paneldaten der Jahre 06, 07, 08, 09, 11, 13, 15, 17, 19
$P_{2,3,4,\dots,9,10,11}$	1 000	Paneldaten der Jahre 03, 04, 05, 06, 07, 08, 09, 11, 13, 15
$P_{3,4,5,\dots,10,11,12}$	1 100	Paneldaten der Jahre 04, 05, 06, 07, 08, 09, 11, 13, 15, 17
$P_{4,5,6,\dots,11,12,13}$	1 127	Paneldaten der Jahre 05, 06, 07, 08, 09, 11, 13, 15, 17, 19
$P_{1,2,3,\dots,9,10,11}$	832	Paneldaten der Jahre 02, 03, 04, 05, 06, 07, 08, 09, 11, 13, 15
$P_{2,3,4,\dots,10,11,12}$	950	Paneldaten der Jahre 03, 04, 05, 06, 07, 08, 09, 11, 13, 15, 17
$P_{3,4,5,\dots,11,12,13}$	1 022	Paneldaten der Jahre 04, 05, 06, 07, 08, 09, 11, 13, 15, 17, 19
$P_{1,2,3,\dots,10,11,12}$	790	Paneldaten der Jahre 02, 03, 04, 05, 06, 07, 08, 09, 11, 13, 15, 17
$P_{2,3,4,\dots,11,12,13}$	884	Paneldaten der Jahre 03, 04, 05, 06, 07, 08, 09, 11, 13, 15, 17, 19
$P_{1,2,3,\dots,11,12,13}$	736	Paneldaten der Jahre 02, 03, 04, 05, 06, 07, 08, 09, 11, 13, 15, 17, 19

^a nur teilnehmende Schulen

4.2 Abschätzung und Bewertung der Zuordnungsgüte

Wie schon bei Pollich (2010) ist eine Abschätzung der Zuordnungsgüte auf Basis der Schulstatistik ab dem Jahr 2006 nicht mehr möglich. Es ist jedoch durchaus möglich zu betrachten, welche prozentualen Anteile Befragter aus einem durchgehenden Paneldatensatz im jeweils nächsten, d. h. um eine Welle erweiterten, durchgehenden Paneldatensatz noch enthalten sind (vgl. Pollich, 2010, S. 18 f.).

Bei Betrachtung der Tabelle 4.2 wird deutlich, dass die Stabilität von Erhebungsjahr zu Erhebungsjahr zunimmt. Erst zwischen dem Fünf- und Sechs-Wellen-Panel ist ein leichter Einbruch zu verzeichnen. Jedoch schon im Vergleich mit dem folgenden Sieben-Wellen-Panel erhöht sich die Stabilität erneut deutlich auf 91%. Dieses Niveau kann bei den weiteren Paneldatensätzen gehalten werden und liegt teilweise sogar bei 93 bzw. 95%. Es wird also deutlich, dass Teilnehmer/innen, die bereits mehrere Wellen durchgehend teilgenommen haben, auch im Folgejahr mit großer Wahrscheinlichkeit

Tabelle 4.2: Ausschöpfungen der durchgehenden Paneldatensätze

(a) Ausschöpfungen Q_{t_1} bis $P_{t_{1,2,3,4,5,6}}$						
Datensatz	Q_{t_1}	$P_{t_{1,2}}$	$P_{t_{1,2,3}}$	$P_{t_{1,2,3,4}}$	$P_{t_{1,2,3,4,5}}$	$P_{t_{1,2,3,4,5,6}}$
Fallzahl	3411	2472	2012	1769	1552	1307
Anteil der Fallzahl des durchgehenden Datensatzes aus $t - 1$ in %		72	81	88	88	84

(b) Ausschöpfungen $P_{t_{1,2,3,4,5,6,7}}$ bis $P_{t_{1,2,3,4,5,6,7,8,9,10}}$				
Datensatz	$P_{t_{1,2,3,4,5,6,7}}$	$P_{t_{1,2,3,4,5,6,7,8}}$	$P_{t_{1,2,3,4,5,6,7,8,9}}$	$P_{t_{1,2,3,4,5,6,7,8,9,10}}$
Fallzahl	1189	1103	1002	911
Anteil der Fallzahl des durchgehenden Datensatzes aus $t - 1$ in %	91	93	91	91

(c) Ausschöpfungen $P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$ bis $P_{t_{1,2,3,4,5,6,7,8,9,10,11,12,13}}$			
Datensatz	$P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$	$P_{t_{1,2,3,4,5,6,7,8,9,10,11,12}}$	$P_{t_{1,2,3,4,5,6,7,8,9,10,11,12,13}}$
Fallzahl	832	790	736
Anteil der Fallzahl des durchgehenden Datensatzes aus $t - 1$ in %	91	95	93

erneut teilnehmen. Der durchgehende Paneldatensatz erweist sich somit als äußerst robust gegenüber Ausfällen.

Auch der Wechsel auf einen zweijährigen Befragungsrhythmus (von $P_{t_{1,2,3,4,5,6,7,8}}$ auf $P_{t_{1,2,3,4,5,6,7,8,9}}$) zeigt keine gravierenden negativen Auswirkungen auf die Ausschöpfungsquote des Paneldatensatzes. Zwar ist ein leichter Rückgang von 93 auf 91% zu beobachten, dies entspricht jedoch dem Ausgangsniveau des vorherigen Panels $P_{t_{1,2,3,4,5,6,7}}$. Dieser hohe Wert kann auch mit den weiter hinzukommenden Wellen gehalten werden. Bei den letzten beiden Erhebungen steigt die Ausschöpfungsquote sogar auf 95 bzw. 93% an. Über die gesamte Länge des Panels zeigt sich also eine Erhöhung der Ausschöpfungsquote sowie eine Stabilisierung auf einem hohen Niveau. Dies bedeutet, dass Personen, die bereits viele Zeitpunkte an der Befragung teilgenommen haben, zum größten Teil auch an den weiteren Erhebungswellen teilnehmen. Insgesamt kann folglich festgehalten werden, dass die Teilnehmer/innen dem Panel auch in der dritten Lebensdekade weitestgehend erhalten bleiben.

5 Zusammensetzung der Paneldaten

Im folgenden Kapitel werden die einzelnen Paneldatensätze, welche bereits in Abschnitt 4.1 aufgelistet wurden, hinsichtlich ihrer Zusammensetzung nach ausgewählten Merkmalen beschrieben. Als Referenz werden den Paneldatensätzen jeweils die Daten aus den Querschnitten gegenübergestellt.¹³ Zusätzlich werden Daten der Schulstatistik herangezogen – dies ist jedoch nur für die ersten drei Erhebungswellen möglich (vgl. Kapitel 4.1).

Bei den vergleichenden Darstellungen werden die zentralen Variablen Geschlecht, Alter und Schulform betrachtet. Dabei ist anzumerken, dass die Bedeutung der Schulform für die Befragten mit steigendem Alter immer mehr abnimmt und in den späteren Erhebungszeitpunkten zur Begutachtung der Repräsentativität der Paneldaten zunehmend unzuverlässiger wird. Zudem wird die Zusammensetzung nach Schulform nur bis zur Welle t_8 untersucht, da in den darauf folgenden Erhebungsjahren t_9 bis t_{13} so gut wie keine Personen mehr eine Schule besuchen, weshalb in den Tabellen auf die Aufschlüsselung nach besuchter Schulform für diese Zeitpunkte verzichtet wird.

Bei der Darstellung der Zusammensetzung jener Teilpanels, die mehrere Wellen umfassen, sind bei den Merkmalen Geschlecht und Alter leichte Inkonsistenzen möglich. Der Grund liegt darin, dass bei *jeder* Frage fehlerhafte bzw. inkonsistente Antworten auftreten. Dies gilt auch für scheinbar eindeutig zu beantwortende Fragen wie denen nach dem Alter oder Geschlecht.

5.1 Die Panelquerschnitte

Wie bereits erwähnt, enthalten die Panelquerschnitte der einzelnen Jahre alle Fälle, deren zugehörige (und nach Plausibilitätskontrollen gültige) Fragebögen einem Bogen mindestens eines weiteren Zeitpunktes zugeordnet werden konnten. Aus den Querschnitten Q_{t_x} tauchen dementsprechend jene Fälle *nicht* im Panelquerschnitt auf, denen kein passender Fragebogen irgendeines weiteren Erhebungsjahres zugeordnet werden konnte. Bei den Panelquerschnitten handelt es sich demnach um stark konstruierte Daten, die nur unter sehr spezifischen Analysegesichtspunkten von Bedeutung sind.¹⁴ Im Folgenden werden die Panelquerschnitte der Erhebungswellen 2015, 2017 und 2019 (Zeitpunkte t_{11} , t_{12} und

¹³ Aufgrund nachträglicher Korrekturen liegen für die Querschnitte Q_{t_1} , Q_{t_2} und Q_{t_5} geringe Abweichungen zwischen diesem Bericht und den Panel-Methodenberichten von Pöge (2007) und Pollich (2010) vor.

¹⁴ Beispielsweise dann, wenn die Population der überhaupt zuordenbaren Fälle von Interesse ist.

t_{13}) im Vergleich zum Querschnittsdatensatz des jeweiligen Erhebungsjahres beschrieben. Dabei werden diese zwei Arten von Datensätzen hinsichtlich ihrer Verteilung nach Geschlecht und des durchschnittlichen Alters gegenübergestellt. Auf die Auswertung der Schulform wird dabei verzichtet, da schon seit etwa der neunten Erhebungswelle in 2011 die weitaus meisten Personen die Schule bereits verlassen hatten. Somit erübrigt sich eine Betrachtung der Schulform ab diesem Zeitpunkt.

5.1.1 Panelquerschnitt 2015 ($P_{Q_{t_{11}}}$)

Der Panelquerschnitt 2015 umfasst 1 812 Fälle und enthält damit 65,4 Prozent der Fälle des ursprünglichen Querschnittsdatensatzes.

Hinsichtlich des Merkmals Geschlecht liegt eine auffällige Verschiebung im Panelquerschnitt vor, wie aus Tabelle 5.1 hervorgeht. Weibliche Befragte sind mit einer Differenz von rund sieben Prozent im Panelquerschnitt stärker vertreten als im Querschnitt $Q_{t_{11}}$. Äquivalent dazu sind männliche Befragte im Panelquerschnitt verglichen mit dem Querschnitt um rund sieben Prozentpunkte unterrepräsentiert.

Tabelle 5.1: Zusammensetzung nach Geschlecht $P_{Q_{t_{11}}}$

	männlich		weiblich		gesamt (100 %)
	<i>n</i>	%	<i>n</i>	%	
$Q_{t_{11}}$	1 252	45,5	1 502	54,5	2 754
$P_{Q_{t_{11}}}$	700	38,6	1 112	61,4	1 812

Bezüglich des durchschnittlichen Alters, wie es in Tabelle 5.2 abgebildet ist, lassen sich zwischen Querschnitt und Panelquerschnitt keinerlei Differenzen feststellen. Lediglich die Standardabweichung fällt im Panelquerschnitt $P_{Q_{t_{11}}}$ etwas geringer aus als im Querschnitt $Q_{t_{11}}$, was für eine etwas homogenere Altersverteilung spricht.

Tabelle 5.2: Altersdurchschnitt $P_{Q_{t_{11}}}$

	Alter	Std.-abw.	gesamt
$Q_{t_{11}}$	26,0	1,0	2 745
$P_{Q_{t_{11}}}$	26,0	0,8	1 806

5.1.2 Panelquerschnitt 2017 ($P_{Q_{t_{12}}}$)

Der Panelquerschnitt 2017 umfasst 1 760 Fälle und enthält damit 62,9 Prozent der Fälle des ursprünglichen Querschnittsdatensatzes.

Tabelle 5.3: Zusammensetzung nach Geschlecht $P_{Q_{t_{12}}}$

	männlich		weiblich		gesamt
	<i>n</i>	%	<i>n</i>	%	(100 %)
$Q_{t_{12}}$	1 269	45,7	1 509	54,3	2 778
$P_{Q_{t_{12}}}$	672	38,2	1 088	61,8	1 760

Hinsichtlich der Zusammensetzung des Geschlechts (Tabelle 5.3) lässt sich im Vergleich mit dem Querschnitt $Q_{t_{12}}$ eine Verzerrung feststellen: Im Panelquerschnitt $P_{Q_{t_{12}}}$ sind weibliche Befragte mit einer Differenz von 7,5 Prozentpunkten überrepräsentiert. Männliche Befragte sind entsprechend im selben Maße unterrepräsentiert. Bezüglich des durchschnittlichen Alters lässt sich eine minimale Abweichung zwischen Querschnitt und Panelquerschnitt feststellen (vgl. Tabelle 5.4). Mit 28,1 Jahren liegt das Durchschnittsalter im Querschnitt $Q_{t_{12}}$ um 0,1 Jahre höher als im Panelquerschnitt, gleichzeitig ist anhand der Standardabweichung eine stärkere Heterogenität der Altersverteilung festzustellen.

Tabelle 5.4: Altersdurchschnitt $P_{Q_{t_{12}}}$

	Alter	Std.-abw.	gesamt
$Q_{t_{12}}$	28,1	1,2	2 753
$P_{Q_{t_{12}}}$	28,0	0,8	1 749

5.1.3 Panelquerschnitt 2019 ($P_{Q_{t_{13}}}$)

Der Panelquerschnitt 2007 umfasst 1 671 Fälle und enthält damit 61,7 Prozent der Fälle des ursprünglichen Querschnittsdatensatzes.

Tabelle 5.5: Zusammensetzung nach Geschlecht $P_{Q_{t_{13}}}$

	männlich		weiblich		gesamt
	<i>n</i>	%	<i>n</i>	%	(100 %)
$Q_{t_{13}}$	1 238	45,9	1 460	54,1	2 698
$P_{Q_{t_{13}}}$	629	37,6	1 042	62,4	1 671

Aus Tabelle 5.5 geht hervor, dass im Panelquerschnitt $P_{Q_{t_{13}}}$ eine Verschiebung hinsichtlich des Geschlechts vorliegt: Vergleicht man die Geschlechterverteilung im Panelquerschnitt mit dem Querschnittsdatensatz des entsprechenden Erhebungsjahres ($Q_{t_{13}}$), so wird deutlich, dass der Anteil der weiblichen Befragten im Panelquerschnitt um

gut acht Prozentpunkte höher liegt. Der Anteil männlicher Befragter fällt entsprechend geringer aus.

Tabelle 5.6: Altersdurchschnitt $P_{Q_{t_{13}}}$

	Alter	Std.-abw.	gesamt
$Q_{t_{13}}$	30,1	0,9	2677
$P_{Q_{t_{13}}}$	30,0	0,8	1668

Ein Vergleich zwischen den Querschnittsdaten $Q_{t_{13}}$ und dem Panelquerschnitt zeigt, dass sich das Durchschnittsalter der Befragten in beiden Datensätzen nur unwesentlich unterscheidet. Jedoch zeigt sich eine leicht höhere Standardabweichung in den Querschnittsdaten (vgl. Tabelle 5.6).

5.2 Die Zwei-Wellen-Panel

Alle (nach den Plausibilitätskontrollen) gültigen Fälle, die mindestens einem weiteren Fall im Datensatz zugeordnet werden konnten, bilden die Zwei-Wellen-Panel. In diesem Bericht werden all jene Datensätze betrachtet, die Teilnahmen in zwei direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke) und nicht bereits in vorherigen Methodenberichten (Pöge, 2007; Pollich, 2010; Daniel & Erdmann, 2017) beschrieben wurden.

5.2.1 Zwei-Wellen-Panel 2013 und 2015 ($P_{t_{10,11}}$)

Das Zwei-Wellen-Panel der Jahre 2013 und 2015 enthält 1 699 Fälle. Aus Tabelle 5.7 geht hervor, dass in den Erhebungszeitpunkten t_{10} und t_{11} eine Verschiebung hinsichtlich des Geschlechts vorliegt:

Tabelle 5.7: Zusammensetzung nach Geschlecht $P_{t_{10,11}}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_{10}	$Q_{t_{10}}$	1 302	45,7	1 546	54,3	2 848
	$P_{t_{10,11}}$	643	37,8	1 056	62,2	1 699
t_{11}	$Q_{t_{11}}$	1 252	45,5	1 502	54,5	2 754
	$P_{t_{10,11}}$	643	37,8	1 056	62,2	1 699

Der Anteil der weiblichen Befragten im Zwei-Wellen-Paneldatensatz liegt im Vergleich zu den Querschnittsdaten der Erhebungszeitpunkte t_{10} und t_{11} rund acht Prozentpunkte höher. Hinsichtlich der Altersverteilung (Tabelle 5.8) zeigt der Vergleich

zwischen dem Zwei-Wellen-Paneldatensatz $P_{t_{10,11}}$ und den Querschnittsdatsätzen keine Unterschiede im durchschnittlichen Alter der Befragten. Jedoch ist die Altershomogenität im Paneldatensatz $P_{t_{10,11}}$ etwas höher als in den Querschnittsdaten der jeweiligen Zeitpunkte.

Tabelle 5.8: Altersdurchschnitt $P_{t_{10,11}}$

		Alter	Std.-abw.	gesamt
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_{10,11}}$	24,0	0,9	1697
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_{10,11}}$	26,0	0,8	1694

5.2.2 Zwei-Wellen-Panel 2015 und 2017 ($P_{t_{11,12}}$)

Das Zwei-Wellen-Panel der Jahre 2015 und 2017 enthält 1 686 Fälle. Auch hier wird beim Vergleich mit den Querschnittsdatsätzen erneut eine Verzerrung hinsichtlich der Verteilung nach Geschlecht ersichtlich (Tabelle 5.9).

Tabelle 5.9: Zusammensetzung nach Geschlecht $P_{t_{11,12}}$

		männlich		weiblich		gesamt
		n	%	n	%	(100 %)
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_{11,12}}$	645	38,3	1041	61,7	1686
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_{11,12}}$	644	38,2	1042	61,8	1686

Der Anteil der weiblichen Befragten im Zwei-Wellen-Paneldatensatz liegt im Vergleich zu den Querschnittsdaten der Erhebungszeitpunkte t_{11} und t_{12} etwa gut sieben Prozentpunkte höher. Hinsichtlich des Durchschnittsalters (Tabelle 5.10) sind nur geringfügige Unterschiede zu erkennen: Zum Zeitpunkt t_{11} unterscheidet sich das durchschnittliche Alter zwischen den Datensätzen nicht, zum Zeitpunkt t_{12} liegt nur eine marginale Differenz von 0,1 Jahren vor. Jedoch ist die Altershomogenität im Paneldatensatz größer, wie sich an den geringeren Standardabweichungen im Vergleich zum Querschnitt ablesen lässt.

Tabelle 5.10: Altersdurchschnitt $P_{t_{11},t_{12}}$

		Alter	Std.-abw.	gesamt
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_{11},t_{12}}$	26,0	0,8	1680
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_{11},t_{12}}$	28,0	0,8	1676

5.2.3 Zwei-Wellen-Panel 2017 und 2019 ($P_{t_{12},t_{13}}$)

Das Zwei-Wellen-Panel der Jahre 2017 und 2019 enthält 1 604 Fälle. In Tabelle 5.11 ist zu erkennen, dass in den Erhebungszeitpunkten t_{12} und t_{13} auch hier eine Verschiebung hinsichtlich des Geschlechts vorliegt.

Tabelle 5.11: Zusammensetzung nach Geschlecht $P_{t_{12},t_{13}}$

		männlich		weiblich		gesamt
		n	%	n	%	(100 %)
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_{12},t_{13}}$	601	37,5	1003	62,5	1604
t_{13}	$Q_{t_{13}}$	1238	45,9	1460	54,1	2698
	$P_{t_{12},t_{13}}$	601	37,5	1003	62,5	1604

Der Anteil der weiblichen Befragten im Paneldatensatz $P_{t_{12},t_{13}}$ liegt in beiden Erhebungszeitpunkten gut acht Prozent höher als im vergleichbaren Querschnittdatensatz. Männliche Befragte sind im selben Maße unterrepräsentiert. Hinsichtlich des Alters (Tabelle 5.12) ist zu erkennen, dass das Durchschnittsalter in beiden Zeitpunkten im Paneldatensatz um 0,1 Jahre unter dem des jeweiligen Querschnittdatensatzes liegt. Die Querschnitte sind dabei von einer etwas größeren Streuung in der Altersverteilung gekennzeichnet, wobei sich Panel- und Querschnittdatensatz insbesondere zum Zeitpunkt t_{12} stärker unterscheiden.

Tabelle 5.12: Altersdurchschnitt $P_{t_{12},t_{13}}$

		Alter	Std.-abw.	gesamt
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_{12},t_{13}}$	28,0	0,8	1594
t_{13}	$Q_{t_{13}}$	30,1	0,9	2677
	$P_{t_{12},t_{13}}$	30,0	0,8	1602

5.3 Die Drei-Wellen-Panel

Alle (nach den Plausibilitätskontrollen) gültigen Fälle, die mindestens zwei weiteren Fällen im Datensatz zugeordnet werden konnten, bilden die Drei-Wellen-Panel. In diesem Bericht werden all jene Datensätze betrachtet, die Teilnahmen in drei direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke) und nicht bereits in vorherigen Methodenberichten (Pöge, 2007; Pollich, 2010; Daniel & Erdmann, 2017) beschrieben wurden.

5.3.1 Drei-Wellen-Panel 2011, 2013 und 2015 ($P_{t_9,10,11}$)

Das Drei-Wellen-Panel der Jahre 2011 bis 2015 enthält 1 625 Fälle. Aus Tabelle 5.13 geht hervor, dass in den Erhebungszeitpunkten t_9 bis t_{11} eine Verzerrung hinsichtlich des Geschlechts vorliegt.

Der Anteil der weiblichen Befragten fällt im Drei-Wellen-Paneldatensatz $P_{t_9,10,11}$ im Vergleich zu den Querschnittsdaten der Erhebungszeitpunkte t_9 bis t_{10} zwischen 8,3 und 9,4 Prozentpunkte höher aus. Der Anteil männlicher Befragter ist entsprechend im selben Maße im Paneldatensatz niedriger.

Tabelle 5.13: Zusammensetzung nach Geschlecht $P_{t_9,10,11}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_9	Q_{t_9}	1 423	46,8	1 620	53,2	3 043
	$P_{t_9,10,11}$	608	37,4	1 017	62,6	1 625
t_{10}	$Q_{t_{10}}$	1 302	45,7	1 546	54,3	2 848
	$P_{t_9,10,11}$	604	37,2	1 021	62,8	1 625
t_{11}	$Q_{t_{11}}$	1 252	45,5	1 502	54,5	2 754
	$P_{t_9,10,11}$	604	37,2	1 021	62,8	1 625

Hinsichtlich des Durchschnittsalters (Tabelle 5.14) ist zu erkennen, dass – wenn überhaupt – nur geringfügige Differenzen zwischen Paneldatensatz und den entsprechenden Querschnitten vorliegen. Der Paneldatensatz ist dabei jedoch von einer etwas größeren Altershomogenität geprägt, was aus der geringeren Standardabweichung ersichtlich wird.

Tabelle 5.14: Altersdurchschnitt $P_{t_9,10,11}$

		Alter	Std.-abw.	gesamt
t_9	Q_{t_9}	22,1	0,9	3 021
	$P_{t_9,10,11}$	22,0	0,8	1 620
t_{10}	$Q_{t_{10}}$	24,1	1,2	2 839
	$P_{t_9,10,11}$	24,0	0,9	1 623
t_{11}	$Q_{t_{11}}$	26,0	1,0	2 745
	$P_{t_9,10,11}$	26,0	0,8	1 620

5.3.2 Drei-Wellen-Panel 2013, 2015 und 2017 ($P_{t_{10,11,12}}$)

Das Drei-Wellen-Panel der Jahre 2013, 2015 und 2017 enthält 1 596 Fälle. In Tabelle 5.15 lässt sich auch an dieser Stelle im Paneldatensatz eine etwas andere Geschlechterverteilung erkennen als in den Querschnittsdatensätzen der Erhebungszeitpunkte t_{10} , t_{11} und t_{12} . Mit einer Differenz von gerundet acht Prozentpunkten sind weibliche Befragte zu jedem Zeitpunkt im Paneldatensatz im Vergleich zum Querschnitt überrepräsentiert bzw. männliche Befragte in selbem Maße unterrepräsentiert.

Tabelle 5.15: Zusammensetzung nach Geschlecht $P_{t_{10,11,12}}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_{10}	$Q_{t_{10}}$	1 302	45,7	1 546	54,3	2 848
	$P_{t_{10,11,12}}$	600	37,6	996	62,4	1 596
t_{11}	$Q_{t_{11}}$	1 252	45,5	1 502	54,5	2 754
	$P_{t_{10,11,12}}$	601	37,7	995	62,3	1 596
t_{12}	$Q_{t_{12}}$	1 269	45,7	1 509	54,3	2 778
	$P_{t_{10,11,12}}$	601	37,7	995	62,3	1 596

Hinsichtlich des Alters (Tabelle 5.16) lässt sich festhalten, dass erneut nur sehr geringe Differenzen in Höhe von 0,1 Jahren zwischen Querschnitt- und Paneldatensatz auftreten; zum Zeitpunkt t_{11} liegt sogar keine Differenz vor. Auch bei diesem Vergleich zeigt sich, dass die Streuung in den Querschnittsdaten um 0,2 bis 0,3 Einheiten höher ausfällt als im Paneldatensatz, was eine etwas höhere Altersheterogenität in den Querschnittsdaten bedeutet.

Tabelle 5.16: Altersdurchschnitt $P_{t_{10,11,12}}$

		Alter	Std.-abw.	gesamt
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_{10,11,12}}$	24,0	0,9	1595
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_{10,11,12}}$	26,0	0,8	1591
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_{10,11,12}}$	28,0	0,8	1587

5.3.3 Drei-Wellen-Panel 2015, 2017 und 2019 ($P_{t_{11,12,13}}$)

Das Drei-Wellen-Panel der Jahre 2015 bis 2019 enthält 1 549 Fälle. In Tabelle 5.17 ist zu erkennen, dass der Anteil weiblicher Befragungspersonen im Paneldatensatz zu jedem Zeitpunkt gut acht Prozentpunkte über den Anteilen in den Querschnittsdatsätzen liegt. Analog dazu fällt der Anteil männlicher Befragter im Panel geringer aus.

Tabelle 5.17: Zusammensetzung nach Geschlecht $P_{t_{11,12,13}}$

		männlich		weiblich		gesamt
		n	%	n	%	(100 %)
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_{11,12,13}}$	580	37,4	969	62,6	1549
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_{11,12,13}}$	579	37,4	970	62,6	1549
t_{13}	$Q_{t_{13}}$	1238	45,9	1460	54,1	2698
	$P_{t_{11,12,13}}$	579	37,4	970	62,6	1549

Mit Blick auf Tabelle 5.18 zeigen sich keine bzw. nur sehr geringfügige Abweichungen zwischen Paneldatensatz und den Querschnittsdatsätzen bezüglich des durchschnittlichen Alters der Befragten. Insgesamt fällt die Streuung der Altersverteilung im Paneldatensatz jedoch etwas geringer aus, sodass eine etwas größere Altershomogenität angenommen werden kann.

Tabelle 5.18: Altersdurchschnitt $P_{t_{11},12,13}$

		Alter	Std.-abw.	gesamt
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_{11},12,13}$	26,0	0,8	1543
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_{11},12,13}$	28,0	0,8	1539
t_{13}	$Q_{t_{13}}$	30,1	0,9	2677
	$P_{t_{11},12,13}$	30,0	0,8	1547

5.4 Die Vier-Wellen-Panel

Alle (nach den Plausibilitätskontrollen) gültigen Fälle, die mindestens drei weiteren Fällen im Datensatz zugeordnet werden konnten, bilden die Vier-Wellen-Panel. In diesem Bericht werden all jene Datensätze betrachtet, die Teilnahmen in vier direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke) und nicht bereits in vorherigen Methodenberichten (Pöge, 2007; Pollich, 2010; Daniel & Erdmann, 2017) beschrieben wurden.

5.4.1 Vier-Wellen-Panel 2009, 2011, 2013 und 2015 ($P_{t_{8,9,10,11}}$)

Das Vier-Wellen-Panel $P_{t_{8,9,10,11}}$ der Jahre 2009 bis 2015 enthält insgesamt 1 573 Fälle. Betrachtet man die Prozentuierung des Merkmals Geschlecht (Tabelle 5.19) so wird deutlich, dass im Paneldatensatz weibliche Befragte im Vergleich zu den Querschnittsdaten über- und männliche Befragte unterrepräsentiert sind. Die Diskrepanz zwischen Panel- und Querschnittsdaten liegt in den jeweiligen Erhebungszeitpunkten bei rund neun bis zehn Prozentpunkten.

Das in Tabelle 5.20 dargestellte Alter zeigt zu allen vier Zeitpunkten sowohl im Vier-Wellen-Panel als auch in den Querschnitten die zu erwartenden Durchschnittswerte. Differenzen zwischen dem Panel $P_{t_{8,9,10,11}}$ und den Querschnitten fallen nur sehr gering aus. Dabei weisen die Querschnittsdaten zu allen Zeitpunkten eine etwas größere Streuung auf. In den Querschnittsdaten ist das Alter also etwas weniger homogen verteilt als in den Paneldaten.

Die Zusammensetzung nach Schulform (Tabelle 5.21) kann nur für den Zeitpunkt t_8 betrachtet werden, da die Befragungspersonen zu diesem Zeitpunkt durchschnittlich 20 Jahre alt sind und entsprechend bereits die Hälfte aller Personen keine Schule mehr besucht. Die Betrachtung der Zusammensetzung nach Schulform erübrigt sich daher ab dem Zeitpunkt t_9 . Zwischen Panel und Querschnitt bestehen zum Zeitpunkt t_8 nur geringfügige Differenzen. Während Gesamtschüler/innen im Panel leicht unterre-

Tabelle 5.19: Zusammensetzung nach Geschlecht $P_{t_8,9,10,11}$

		männlich		weiblich		gesamt (100 %)
		<i>n</i>	%	<i>n</i>	%	
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_8,9,10,11}$	576	36,6	997	63,4	1573
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_8,9,10,11}$	581	36,9	992	63,1	1573
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_8,9,10,11}$	575	36,6	998	63,4	1573
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_8,9,10,11}$	576	36,6	997	63,4	1573

Tabelle 5.20: Altersdurchschnitt $P_{t_8,9,10,11}$

		Alter	Std.-abw.	gesamt
		t_8	Q_{t_8}	20,1
	$P_{t_8,9,10,11}$	20,0	0,7	1571
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_8,9,10,11}$	22,0	0,8	1568
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_8,9,10,11}$	23,9	0,8	1571
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_8,9,10,11}$	26,0	0,8	1568

präsentiert sind, befinden sich etwas mehr Gymnasiast/innen im Längsschnittdatensatz. Die Unterschiede sind allerdings marginal und betragen nur 0,8 Prozentpunkte.

Tabelle 5.21: Zusammensetzung $P_{8,9,10,11}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)							gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{8,9,10,11}$	—	—	3,3	4,2	22,7	—	19,1	50,7	1520

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.4.2 Vier-Wellen-Panel 2011, 2013, 2015 und 2017 ($P_{t_9,10,11,12}$)

Das Vier-Wellen-Panel $P_{t_9,10,11,12}$ der Jahre 2011 bis 2017 enthält 1 527 Fälle. Hinsichtlich der Verteilung nach Geschlecht lassen sich Diskrepanzen zwischen Querschnitts- und Paneldaten feststellen (Tabelle 5.22). Je nach Erhebungszeitpunkt beträgt die Abweichung zwischen Querschnitt und Panel 8,5 bis 9,6 Prozentpunkte dahingehend, dass weibliche Personen im Paneldatensatz stärker vertreten sind als im Querschnittsdatsatz. Umgekehrt gilt entsprechend, dass männliche Personen im selben Maße unterrepräsentiert sind.

Tabelle 5.22: Zusammensetzung nach Geschlecht $P_{t_9,10,11,12}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_9,10,11,12}$	568	37,2	959	62,8	1527
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_9,10,11,12}$	564	36,9	963	63,1	1527
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_9,10,11,12}$	565	37,0	962	63,0	1527
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_9,10,11,12}$	565	37,0	962	63,0	1527

Das in Tabelle 5.23 dargestellte Alter zeigt zu allen vier Zeitpunkten die zu erwartenden Durchschnittswerte. Unterschiede zwischen Paneldatensatz und den Querschnitten sind dabei marginal und betragen lediglich maximal 0,2 Jahre. Im Paneldatensatz lässt sich im Vergleich zum jeweiligen Querschnitt in allen Zeitpunkten eine leicht geringere Standardabweichung feststellen, was für eine größere Altershomogenität im Paneldatensatz spricht.

Tabelle 5.23: Altersdurchschnitt $P_{t_9,10,11,12}$

		Alter	Std.-abw.	gesamt
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_9,10,11,12}$	22,0	0,8	1522
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_9,10,11,12}$	23,9	0,9	1526
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_9,10,11,12}$	26,0	0,8	1522
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_9,10,11,12}$	28,0	0,8	1518

5.4.3 Vier-Wellen-Panel 2013, 2015, 2017 und 2019 ($P_{t_{10,11,12,13}}$)

Das Vier-Wellen-Panel $P_{t_{10,11,12,13}}$ der Jahre 2013 bis 2019 enthält insgesamt 1 474 Fälle. Bei der Verteilung nach Geschlecht (Tabelle 5.24) ist zu erkennen, dass weibliche Befragte im Paneldatensatz stärker vertreten sind als in den Querschnittsdatensätzen. Männliche Befragungspersonen weisen im Panel entsprechend einen geringeren Anteil auf. Diese Diskrepanz zwischen Querschnittsdatensatz und dem Panel $P_{t_{10,11,12,13}}$ beträgt je nach Zeitpunkt zwischen 8,6 und 9,0 Prozentpunkte.

Tabelle 5.24: Zusammensetzung nach Geschlecht $P_{t_{10,11,12,13}}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_{10,11,12,13}}$	543	36,8	931	63,2	1474
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_{10,11,12,13}}$	544	36,9	930	63,1	1474
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_{10,11,12,13}}$	544	36,9	930	63,1	1474
t_{13}	$Q_{t_{13}}$	1238	45,9	1460	54,1	2698
	$P_{t_{10,11,12,13}}$	544	36,9	930	63,1	1474

Hinsichtlich des Alters in Tabelle 5.25 ist zu erkennen, dass nur sehr geringfügige Unterschiede zwischen den Datensatztypen vorliegen. Die Differenz zwischen Querschnitt und Paneldatensatz beträgt dabei 0,1 bis 0,2 Jahre. Zum Zeitpunkt t_{11} unterscheiden sich beide Datensätze hinsichtlich des Durchschnittsalters nicht. Allerdings lässt sich feststellen, dass die Standardabweichung in den Querschnittsdatensätzen stets etwas

höher ausfällt als im Paneldatensatz. Die Altersverteilung in den Querschnitten ist also etwas heterogener als im Paneldatensatz der Wellen t_{10} bis t_{13} .

Tabelle 5.25: Altersdurchschnitt $P_{t_{10},t_{11},t_{12},t_{13}}$

		Alter	Std.-abw.	gesamt
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_{10},t_{11},t_{12},t_{13}}$	23,9	0,9	1473
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_{10},t_{11},t_{12},t_{13}}$	26,0	0,8	1469
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_{10},t_{11},t_{12},t_{13}}$	28,0	0,8	1465
t_{13}	$Q_{t_{13}}$	30,1	0,9	2677
	$P_{t_{10},t_{11},t_{12},t_{13}}$	30,0	0,8	1472

5.5 Die Fünf-Wellen-Panel

Alle (nach den Plausibilitätskontrollen) gültigen Fälle, die mindestens vier weiteren Fällen im Datensatz zugeordnet werden konnten, bilden die Fünf-Wellen-Panel. In diesem Bericht werden all jene Datensätze betrachtet, die Teilnahmen in fünf direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke) und nicht bereits in vorherigen Methodenberichten (Pollich, 2010; Daniel & Erdmann, 2017) beschrieben wurden.

5.5.1 Fünf-Wellen-Panel 2008, 2009, 2011, 2013 und 2015 ($P_{t_{7,8,9,10,11}}$)

Das Fünf-Wellen-Panel der Erhebungsjahre 2008 bis 2015 umfasst 1 520 Fälle.

Die in Tabelle 5.26 gezeigte Zusammensetzung dieses Panels nach Geschlecht lässt erkennen, dass im Vergleich zu den Referenzdaten des Querschnitts eine Verschiebung hinsichtlich der Geschlechterverteilung im Panel $P_{t_{7,8,9,10,11}}$ vorliegt. Während der Paneldatensatz mehr weibliche Befragte enthält als die Querschnittsdaten, sind männliche Befragte im Panel entsprechend weniger vertreten. Diese Abweichung zwischen den Datensatztypen liegt je nach Zeitpunkt bei rund neun bis zehn Prozentpunkten.

Tabelle 5.27 zeigt das durchschnittliche Alter im Panel $P_{t_{7,8,9,10,11}}$ sowie vergleichend die Werte der Querschnittserhebungen. In vier von fünf Zeitpunkten liegt das Durchschnittsalter im Querschnitt leicht über dem im Fünf-Wellen-Panel. Diese Abweichung beläuft sich dabei allerdings nur auf maximal 0,2 Jahre. Zum Zeitpunkt t_{11} liegt keine Abweichung im Durchschnittsalter vor. Zudem weist das Panel in den meisten Zeitpunkten eine leicht größere Altershomogenität auf, wie an der etwas geringeren Standardabweichung zu erkennen ist.

Tabelle 5.26: Zusammensetzung nach Geschlecht $P_{t_7,8,9,10,11}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t_7,8,9,10,11}$	559	36,8	961	63,2	1520
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_7,8,9,10,11}$	559	36,8	961	63,2	1520
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_7,8,9,10,11}$	562	37,0	958	63,0	1520
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_7,8,9,10,11}$	558	36,7	962	63,3	1520
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_7,8,9,10,11}$	560	36,8	960	63,2	1520

Eine Betrachtung der Zusammensetzung nach Schulform erübrigt sich ab dem Erhebungszeitpunkt t_9 , da ab diesem Zeitpunkt bereits der Großteil der Befragten die Schule verlassen hat und die Auswertung keinen inhaltlichen Wert mehr besitzt. In Tabelle 5.28 sind daher nur die Angaben in den Zeitpunkten t_7 und t_8 dargestellt. Dabei zeigen sich zwischen Paneldatensatz und den Querschnittsdatsätzen nur geringfügige Abweichungen. Im Zeitpunkt t_7 ist zu erkennen, dass der Paneldatensatz $P_{t_7,8,9,10,11}$ gut acht Prozent mehr Gymnasiast/innen enthält als der Querschnittsdatsatz Q_{t_7} . Zum Zeitpunkt t_8 fallen die Abweichungen geringer aus, allerdings besucht hier nur noch etwa die Hälfte aller Befragten überhaupt eine Schule.

Tabelle 5.27: Altersdurchschnitt $P_{t_{7,8,9,10,11}}$

		Alter	Std.-abw.	gesamt
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t_{7,8,9,10,11}}$	18,9	0,8	1520
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t_{7,8,9,10,11}}$	20,0	0,7	1518
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_{7,8,9,10,11}}$	22,0	0,8	1516
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_{7,8,9,10,11}}$	23,9	0,8	1518
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_{7,8,9,10,11}}$	26,0	0,8	1515

Tabelle 5.28: Zusammensetzung $P_{t_{7,8,9,10,11}}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)							gesamt (100 %)	
		HS	RS	GS	GY	BK	SO	andere		keine
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{7,8,9,10,11}}$	—	—	15,3	31,2	26,1	—	12,4	15,0	1488
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{7,8,9,10,11}}$	—	—	3,3	4,2	23,0	—	19,3	50,3	1468

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.5.2 Fünf-Wellen-Panel 2009, 2011, 2013, 2015 und 2017 ($P_{t_{8,9,10,11,12}}$)

Das Fünf-Wellen-Panel der Erhebungsjahre 2009 bis 2017 beinhaltet insgesamt 1 480 Fälle. Bei der Zusammensetzung dieses Panels nach Geschlecht wird im Vergleich zu den Querschnittsdaten eine Verzerrung deutlich (Tabelle 5.29). Während weibliche Personen im Paneldatensatz im Vergleich zum Querschnitt überrepräsentiert sind, sind männliche Personen unterrepräsentiert. Diese Abweichung liegt je nach Zeitpunkt bei rund neun bis zehn Prozentpunkten.

Hinsichtlich des Durchschnittsalters (Tabelle 5.30) zeigen sich nur marginale Differenzen zwischen Panel und Querschnitt dahingehend, dass das durchschnittliche Alter im Querschnitt leicht höher ist. Der Unterschied liegt jedoch bei maximal 0,2 Jahren. Zudem ist zu erkennen, dass in den Querschnittsdatensätzen oftmals eine leicht höhere Standardabweichung vorliegt, was für eine etwas stärkere Altersheterogenität in den Querschnittsdatensätzen spricht.

Tabelle 5.29: Zusammensetzung nach Geschlecht $P_{t_8,9,10,11,12}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_8,9,10,11,12}$	538	36,4	942	63,6	1480
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_8,9,10,11,12}$	543	36,7	937	63,3	1480
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_8,9,10,11,12}$	537	36,3	943	63,7	1480
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_8,9,10,11,12}$	539	36,4	941	63,6	1480
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_8,9,10,11,12}$	539	36,4	941	63,6	1480

Die Zusammensetzung nach Schulform kann für das Panel $P_{t_8,9,10,11,12}$ lediglich für den Zeitpunkt t_8 betrachtet werden, da ab dem Zeitpunkt t_9 der Großteil der Befragten die Schule bereits verlassen hat, sodass sich eine Betrachtung nach Schulform ab dieser Erhebungswelle erübrigt. Es lassen sich zwischen dem Fünf-Wellen-Panel und dem Querschnitt Q_{t_8} nur leichte Abweichungen feststellen. Im Panel sind etwas mehr Gymnasiast/innen vertreten, im Querschnitt hingegen etwas mehr Gesamtschüler/innen. Die Abweichung liegt jedoch jeweils bei unter einem Prozent.

Tabelle 5.30: Altersdurchschnitt $P_{t_8,9,10,11,12}$

		Alter	Std.-abw.	gesamt
t_8	Q_{t_8}	20,1	0,8	3 073
	$P_{t_8,9,10,11,12}$	20,0	0,8	1 478
t_9	Q_{t_9}	22,1	0,9	3 021
	$P_{t_8,9,10,11,12}$	22,0	0,8	1 475
t_{10}	$Q_{t_{10}}$	24,1	1,2	2 839
	$P_{t_8,9,10,11,12}$	23,9	0,8	1 479
t_{11}	$Q_{t_{11}}$	26,0	1,0	2 745
	$P_{t_8,9,10,11,12}$	26,0	0,8	1 475
t_{12}	$Q_{t_{12}}$	28,1	1,2	2 753
	$P_{t_8,9,10,11,12}$	28,0	0,8	1 472

Tabelle 5.31: Zusammensetzung $P_{t_8,9,10,11,12}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)							gesamt (100 %)	
		HS	RS	GS	GY	BK	SO	andere		keine
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2 934
	$P_{t_8,9,10,11,12}$	—	—	3,2	4,2	22,9	—	19,3	50,4	1 432

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.5.3 Fünf-Wellen-Panel 2011, 2013, 2015, 2017 und 2019 ($P_{t_9,10,11,12,13}$)

Das Fünf-Wellen-Panel der Erhebungsjahre 2011 bis 2019 enthält 1 414 Fälle. Tabelle 5.32 zeigt die Zusammensetzung nach Geschlecht im Panel $P_{t_9,10,11,12,13}$ im Vergleich zu den Querschnittsdatensätzen der Zeitpunkte t_9 bis t_{13} . Dabei wird eine Verschiebung deutlich dahingehend, dass der Paneldatensatz mehr weibliche Personen enthält als die Querschnittsdaten bzw. umgekehrt weniger männliche Personen. Die Differenz zwischen Panel und Querschnitten liegt dabei je nach Zeitpunkt zwischen gut neun bis zehn Prozentpunkten.

Bei Betrachtung des Durchschnittsalters (Tabelle 5.33) wird deutlich, dass sich dieses zwischen Panel- und Querschnittsdaten nur unwesentlich unterscheidet. Sofern überhaupt Abweichungen vorliegen, so betragen diese maximal 0,2 Jahre. Zudem lässt sich im Paneldatensatz $P_{t_9,10,11,12,13}$ eine leicht geringere Standardabweichung im Vergleich zu den Querschnittsdatensätzen feststellen. Das Alter ist im Paneldatensatz also etwas homogener verteilt.

Tabelle 5.32: Zusammensetzung nach Geschlecht $P_{t_9,10,11,12,13}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_9,10,11,12,13}$	513	36,3	901	63,7	1414
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_9,10,11,12,13}$	510	36,1	904	63,9	1414
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_9,10,11,12,13}$	511	36,1	903	63,9	1414
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_9,10,11,12,13}$	511	36,1	903	63,9	1414
t_{13}	$Q_{t_{13}}$	1238	45,9	1460	54,1	2698
	$P_{t_9,10,11,12,13}$	511	36,1	903	63,9	1414

Tabelle 5.33: Altersdurchschnitt $P_{t_9,10,11,12,13}$

		Alter	Std.-abw.	gesamt
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_9,10,11,12,13}$	22,0	0,8	1410
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_9,10,11,12,13}$	23,9	0,9	1413
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_9,10,11,12,13}$	26,0	0,8	1409
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_9,10,11,12,13}$	28,0	0,8	1405
t_{13}	$Q_{t_{13}}$	30,1	0,9	2677
	$P_{t_9,10,11,12,13}$	30,0	0,8	1412

5.6 Die Sechs-Wellen-Panel

Alle (nach den Plausibilitätskontrollen) gültigen Fälle, die mindestens fünf weiteren Fällen im Datensatz zugeordnet werden konnten, bilden die Sechs-Wellen-Panel. In diesem Bericht werden all jene Datensätze betrachtet, die Teilnahmen in sechs direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke) und nicht bereits in vorherigen Methodenberichten (Pollich, 2010; Daniel & Erdmann, 2017) beschrieben wurden.

5.6.1 Sechs-Wellen-Panel 2007, 2008, 2009, 2011, 2013 und 2015

$(P_{t_{6,7,8,9,10,11}})$

Das Sechs-Wellen-Panel der Erhebungsjahre 2007 bis 2015 umfasst 1 479 Fälle. Hinsichtlich der Zusammensetzung nach Geschlecht, wie sie in Tabelle 5.34 dargestellt ist, lässt sich eine Verzerrung des Paneldatensatzes im Vergleich zu den Querschnittsdaten feststellen. Während weibliche Personen im Paneldatensatz überrepräsentiert sind, sind männliche Personen in gleichem Maße unterrepräsentiert. Die Abweichung zwischen Panel und Querschnitt beträgt dabei je nach Zeitpunkt rund neun bis elf Prozentpunkte.

Tabelle 5.34: Zusammensetzung nach Geschlecht $P_{t_{6,7,8,9,10,11}}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_6	Q_{t_6}	1 567	47,0	1 768	53,0	3 335
	$P_{t_{6,7,8,9,10,11}}$	534	36,1	945	63,9	1 479
t_7	Q_{t_7}	1 452	47,1	1 634	52,9	3 086
	$P_{t_{6,7,8,9,10,11}}$	535	36,2	944	63,8	1 479
t_8	Q_{t_8}	1 446	46,8	1 644	53,2	3 090
	$P_{t_{6,7,8,9,10,11}}$	535	36,2	944	63,8	1 479
t_9	Q_{t_9}	1 423	46,8	1 620	53,2	3 043
	$P_{t_{6,7,8,9,10,11}}$	537	36,3	942	63,7	1 479
t_{10}	$Q_{t_{10}}$	1 302	45,7	1 546	54,3	2 848
	$P_{t_{6,7,8,9,10,11}}$	534	36,1	945	63,9	1 479
t_{11}	$Q_{t_{11}}$	1 252	45,5	1 502	54,5	2 754
	$P_{t_{6,7,8,9,10,11}}$	536	36,2	943	63,8	1 479

Hinsichtlich des Durchschnittsalters (Tabelle 5.35) sind nur geringfügige Abweichungen zwischen dem Paneldatensatz und den jeweiligen Querschnitten zu erkennen. Das durchschnittliche Alter liegt dabei in den Querschnitten oftmals etwas höher – die Abweichung beträgt allerdings maximal 0,2 Jahre. Zusätzlich ist das Alter in den Quer-

schnittsdatensätze etwas heterogener verteilt als im Paneldatensatz $P_{t_{6,7,8,9,10,11}}$, was sich an der etwas höheren Standardabweichung erkennen lässt.

Tabelle 5.35: Altersdurchschnitt $P_{t_{6,7,8,9,10,11}}$

		Alter	Std.-abw.	gesamt
t_6	Q_{t_6}	18,1	1,1	3320
	$P_{t_{6,7,8,9,10,11}}$	18,0	0,8	1476
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t_{6,7,8,9,10,11}}$	18,9	0,8	1479
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t_{6,7,8,9,10,11}}$	20,0	0,7	1477
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_{6,7,8,9,10,11}}$	22,0	0,8	1476
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_{6,7,8,9,10,11}}$	23,9	0,8	1477
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_{6,7,8,9,10,11}}$	26,0	0,8	1475

Die Zusammensetzung nach Schulform (Tabelle 5.36) für die Zeitpunkte t_6 , t_7 und t_8 zeigt, dass eine Verschiebung des Paneldatensatzes vorliegt. Zum Zeitpunkt t_6 sind im Sechs-Wellen-Panel sieben Prozent mehr Schüler/innen an Gymnasien vertreten als im Querschnitt dieses Erhebungsjahres. Zum Zeitpunkt t_7 beträgt diese Differenz 8,6 Prozent. Umgekehrt enthält der Paneldatensatz zu diesen beiden Zeitpunkten bis zu rund sechs Prozent weniger Schülerinnen und Schüler, die zu diesem Zeitpunkt keine Schule mehr besuchen. Zum Zeitpunkt t_8 sind nur noch geringe Abweichungen hinsichtlich der Schulform zu erkennen, was in erster Linie daran liegt, dass auch nur noch etwa die Hälfte aller Befragungspersonen überhaupt noch eine Schule besucht.

Tabelle 5.36: Zusammensetzung $P_{6,7,8,9,10,11}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)								gesamt (100 %)
		HS	RS	GS	GY	BK	SO	andere	keine	
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3 219
	$P_{6,7,8,9,10,11}$	—	—	17,4	33,2	29,2	—	10,2	9,9	1 454
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2 975
	$P_{6,7,8,9,10,11}$	—	—	15,5	31,4	26,0	—	12,2	14,8	1 447
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2 934
	$P_{6,7,8,9,10,11}$	—	—	3,3	4,1	23,0	—	19,0	50,6	1 432

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.6.2 Sechs-Wellen-Panel 2008, 2009, 2011, 2013, 2015 und 2017

($P_{7,8,9,10,11,12}$)

Das Sechs-Wellen-Panel der Erhebungsjahre 2008 bis 2017 umfasst 1 430 Fälle. Hinsichtlich der Verteilung nach Geschlecht (Tabelle 5.37) ist zu erkennen, dass die Zusammensetzung im Panel von der in den Querschnitten abweicht. Während weibliche Personen im Panel überrepräsentiert sind, sind männliche Personen entsprechend weniger vertreten. Diese Diskrepanz beträgt dabei je nach Zeitpunkt rund neun bis elf Prozentpunkte.

Tabelle 5.38 zeigt das durchschnittliche Alter im Sechs-Wellen-Panel $P_{7,8,9,10,11,12}$ sowie vergleichend die Werte der Querschnittserhebungen. Sofern überhaupt Abweichungen bestehen, so liegt das durchschnittliche Alter im Querschnitt meist geringfügig höher als im Paneldatensatz. Diese Abweichung fällt mit maximal 0,2 Jahren allerdings sehr gering aus. Zudem ist die Altersverteilung im Panel an vier von sechs Zeitpunkten von einer etwas geringeren Standardabweichung geprägt, was für eine größere Altershomogenität im Paneldatensatz spricht.

Die Zusammensetzung nach Schulform (Tabelle 5.39) zeigt, dass zum Zeitpunkt t_7 Abweichungen zwischen Panel- und Querschnittsdatsatz bestehen. Während im Paneldatensatz etwa neun Prozent mehr Gymnasiast/innen vertreten sind als im Querschnitt Q_{t_7} , so befinden sich umgekehrt gut fünf Prozent weniger Personen im Panel, die keine Schule mehr besuchen. Zum Zeitpunkt t_8 bestehen nur noch geringfügige Differenzen zwischen beiden Datensatztypen – in dieser Erhebungswelle besucht bereits die Hälfte aller Befragungspersonen keine Schule mehr.

Tabelle 5.37: Zusammensetzung nach Geschlecht $P_{t_7,8,9,10,11,12}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t_7,8,9,10,11,12}$	521	36,4	909	63,6	1430
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_7,8,9,10,11,12}$	521	36,4	909	63,6	1430
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_7,8,9,10,11,12}$	524	36,6	906	63,4	1430
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_7,8,9,10,11,12}$	520	36,4	910	63,6	1430
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_7,8,9,10,11,12}$	523	36,6	907	63,4	1430
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_7,8,9,10,11,12}$	522	36,5	908	63,5	1430

Tabelle 5.38: Altersdurchschnitt $P_{t_7,8,9,10,11,12}$

		Alter	Std.-abw.	gesamt
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t_7,8,9,10,11,12}$	18,9	0,8	1430
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t_7,8,9,10,11,12}$	20,0	0,8	1428
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_7,8,9,10,11,12}$	22,0	0,8	1426
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_7,8,9,10,11,12}$	23,9	0,8	1429
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_7,8,9,10,11,12}$	26,0	0,8	1425
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_7,8,9,10,11,12}$	28,0	0,8	1422

Tabelle 5.39: Zusammensetzung $P_{t7,8,9,10,11,12}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)							gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t7,8,9,10,11,12}$	—	—	14,8	31,4	25,8	—	12,6	15,3	1401
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t7,8,9,10,11,12}$	—	—	3,2	4,1	23,2	—	19,5	50,0	1382

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.6.3 Sechs-Wellen-Panel 2009, 2011, 2013, 2015, 2017 und 2019

 $(P_{t_{8,9,10,11,12,13}})$

Das Sechs-Wellen-Panel der Erhebungsjahre 2009 bis 2019 umfasst 1 372 Fälle. Unter Betrachtung der in Tabelle 5.40 gezeigten Zusammensetzung dieses Panels nach Geschlecht wird im Vergleich zu den Referenzdaten des Querschnitts deutlich, dass eine Verschiebung des Panels $P_{t_{8,9,10,11,12,13}}$ vorliegt. Der Anteil weiblicher Befragter liegt im Paneldatensatz etwa zehn bis elf Prozentpunkte über dem Anteil im Querschnitt. Der Anteil männlicher Befragter ist im Paneldatensatz entsprechend vergleichsweise geringer.

Tabelle 5.40: Zusammensetzung nach Geschlecht $P_{t_{8,9,10,11,12,13}}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_8	Q_{t_8}	1 446	46,8	1 644	53,2	3 090
	$P_{t_{8,9,10,11,12,13}}$	487	35,5	885	64,5	1 372
t_9	Q_{t_9}	1 423	46,8	1 620	53,2	3 043
	$P_{t_{8,9,10,11,12,13}}$	491	35,8	881	64,2	1 372
t_{10}	$Q_{t_{10}}$	1 302	45,7	1 546	54,3	2 848
	$P_{t_{8,9,10,11,12,13}}$	486	35,4	886	64,6	1 372
t_{11}	$Q_{t_{11}}$	1 252	45,5	1 502	54,5	2 754
	$P_{t_{8,9,10,11,12,13}}$	488	35,6	884	64,4	1 372
t_{12}	$Q_{t_{12}}$	1 269	45,7	1 509	54,3	2 778
	$P_{t_{8,9,10,11,12,13}}$	488	35,6	884	64,4	1 372
t_{13}	$Q_{t_{13}}$	1 238	45,9	1 460	54,1	2 698
	$P_{t_{8,9,10,11,12,13}}$	488	35,6	884	64,4	1 372

Beim Durchschnittsalter (Tabelle 5.41) zeigen sich nur wenige Differenzen zwischen Paneldaten und den Querschnitten. In den meisten Zeitpunkten liegt das durchschnittliche Alter im Querschnitt etwas höher, wobei diese Abweichung lediglich 0,1 bis 0,2 Jahre beträgt. Im Paneldatensatz zeigt sich jedoch oftmals eine etwas geringere Streuung des Merkmals Alter, was für eine etwas größere Altershomogenität im Längsschnittdatensatz spricht.

Die Zusammensetzung nach Schulform (Tabelle 5.42) kann für das Panel $P_{t_{8,9,10,11,12,13}}$ lediglich für den Zeitpunkt t_8 betrachtet werden, da im weiteren Verlauf kaum noch Personen eine Schule besuchen und sich die Betrachtung dieses Merkmals ab dem Erhebungszeitpunkt t_9 erübrigt. Zum Zeitpunkt t_8 bestehen nur marginale Abweichungen zwischen dem Sechs-Wellen-Panel und dem Querschnitt Q_{t_8} . Bereits die Hälfte aller Befragungspersonen besucht zu diesem Zeitpunkt keine Schule mehr.

Tabelle 5.41: Altersdurchschnitt $P_{8,9,10,11,12,13}$

		Alter	Std.-abw.	gesamt
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{8,9,10,11,12,13}$	19,9	0,8	1370
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{8,9,10,11,12,13}$	22,0	0,8	1368
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{8,9,10,11,12,13}$	23,9	0,8	1371
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{8,9,10,11,12,13}$	26,0	0,8	1367
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{8,9,10,11,12,13}$	28,0	0,8	1364
t_{13}	$Q_{t_{13}}$	30,1	0,9	2677
	$P_{8,9,10,11,12,13}$	30,0	0,8	1370

Tabelle 5.42: Zusammensetzung $P_{8,9,10,11,12,13}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)							gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{8,9,10,11,12,13}$	—	—	3,4	4,0	23,0	—	19,7	50,0	1327

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.7 Die Sieben-Wellen-Panel

Alle (nach den Plausibilitätskontrollen) gültigen Fälle, die mindestens sechs weiteren Fällen im Datensatz zugeordnet werden konnten, bilden die Sieben-Wellen-Panel. In diesem Bericht werden all jene Datensätze betrachtet, die Teilnahmen in sieben direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke) und nicht bereits im vorherigen Methodenbericht (Daniel & Erdmann, 2017) beschrieben wurden.

5.7.1 Sieben-Wellen-Panel 2006, 2007, 2008, 2009, 2011, 2013 und 2015

($P_{5,6,7,8,9,10,11}$)

Das Sieben-Wellen-Panel der Erhebungsjahre 2006 bis 2015 umfasst 1 419 Fälle. Unter Betrachtung der in Tabelle 5.43 gezeigten Zusammensetzung dieses Panels nach Geschlecht wird im Vergleich zu den Referenzdaten des Querschnitts deutlich, dass eine Verschiebung des Panels $P_{5,6,7,8,9,10,11}$ vorliegt. Im Paneldatensatz liegt der Anteil

weiblicher Befragter stets über den Referenzwerten der Querschnittsdatensätze. Diese Diskrepanz reicht dabei von rund neun bis zu 14 Prozentpunkte. Männliche Befragte sind entsprechend im selben Maße im Paneldatensatz unterrepräsentiert.

Tabelle 5.43: Zusammensetzung nach Geschlecht $P_{t_{5,6,7,8,9,10,11}}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_5	Q_{t_5}	2265	49,8	2282	50,2	4547
	$P_{t_{5,6,7,8,9,10,11}}$	510	35,9	909	64,1	1419
t_6	Q_{t_6}	1567	47,0	1768	53,0	3335
	$P_{t_{5,6,7,8,9,10,11}}$	511	36,0	908	64,0	1419
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t_{5,6,7,8,9,10,11}}$	511	36,0	908	64,0	1419
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_{5,6,7,8,9,10,11}}$	510	35,9	909	64,1	1419
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_{5,6,7,8,9,10,13}}$	513	36,2	906	63,8	1419
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_{5,6,7,8,9,10,11}}$	510	35,9	909	64,1	1419
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_{5,6,7,8,9,10,11}}$	512	36,1	907	63,9	1419

Hinsichtlich des Durchschnittsalters (Tabelle 5.44) zeigen sich leichte Abweichungen zwischen Panel und Querschnitt dahingehend, dass das durchschnittliche Alter in den Querschnittsdatensätzen meist etwas höher ausfällt. Diese Diskrepanz ist mit 0,1 bis 0,2 Jahren jedoch sehr gering. Lediglich zum Zeitpunkt t_5 fällt sie mit einer Differenz von 0,4 Jahren etwas höher aus.

Hinsichtlich der Zusammensetzung der Datensätze nach Schulform (Tabelle 5.45) liegt ebenfalls eine Verzerrung des Paneldatensatzes im Vergleich mit den Querschnittsdatensätzen vor. Beispielsweise befinden sich zum Zeitpunkt t_5 über 18 Prozent mehr Schüler/innen an Gymnasien im Paneldatensatz $P_{t_{5,6,7,8,9,10,11}}$ als im Querschnitt Q_{t_5} . Auch Gesamtschüler/innen sind im Panel etwas stärker vertreten. Umgekehrt weist der Paneldatensatz einen geringeren Anteil an Schüler/innen an Berufskollegs auf – die Diskrepanz beträgt hier knapp 24 Prozentpunkte. Zu den Zeitpunkten t_6 und t_7 gestalten sich die Abweichungen ähnlich, wobei sich die Unterschiede mit zunehmendem Erhebungszeitpunkt nivellieren, da immer weniger Personen überhaupt noch eine Schule besuchen.

Tabelle 5.44: Altersdurchschnitt $P_{t5,6,7,8,9,10,11}$

		Alter	Std.-abw.	gesamt
t_5	Q_{t_5}	17,4	1,3	4468
	$P_{t5,6,7,8,9,10,11}$	17,0	0,8	1408
t_6	Q_{t_6}	18,1	1,1	3320
	$P_{t5,6,7,8,9,10,11}$	17,9	0,8	1416
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t5,6,7,8,9,10,11}$	18,9	0,8	1419
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t5,6,7,8,9,10,11}$	19,9	0,7	1418
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t5,6,7,8,9,10,11}$	21,9	0,8	1417
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t5,6,7,8,9,10,11}$	23,9	0,8	1417
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t5,6,7,8,9,10,11}$	26,0	0,8	1415

Tabelle 5.45: Zusammensetzung $P_{t5,6,7,8,9,10,11}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)								gesamt (100 %)
		HS	RS	GS	GY	BK	SO	andere	keine	
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t5,6,7,8,9,10,11}$	0,2	0,5	21,8	38,4	39,1	0,1	—	—	1313
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t5,6,7,8,9,10,11}$	—	—	17,7	34,2	28,8	—	9,7	9,6	1396
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t5,6,7,8,9,10,11}$	—	—	15,8	32,3	25,4	—	12,2	14,4	1390
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t5,6,7,8,9,10,11}$	—	—	3,3	4,2	22,5	—	19,4	50,6	1374

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.7.2 Sieben-Wellen-Panel 2007, 2008, 2009, 2011, 2013, 2015 und 2017

$(P_{t_{6,7,8,9,10,11,12}})$

Das Sieben-Wellen-Panel der Erhebungsjahre 2007 bis 2017 umfasst 1 394 Fälle. Mit Blick auf die Verteilung nach Geschlecht in Tabelle 5.46 wird deutlich, dass das Panel hinsichtlich der Zusammensetzung verzerrt ist. Im Paneldatensatz befinden sich im Vergleich zu den Querschnittsdatsätzen je nach Zeitpunkt etwa neun bis elf Prozent mehr weibliche Befragte. Umgekehrt fällt der Anteil männlicher Befragter vergleichsweise geringer aus.

Tabelle 5.46: *Zusammensetzung nach Geschlecht $P_{t_{6,7,8,9,10,11,12}}$*

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_6	Q_{t_6}	1 567	47,0	1 768	53,0	3 335
	$P_{t_{6,7,8,9,10,11,12}}$	500	35,9	894	64,1	1 394
t_7	Q_{t_7}	1 452	47,1	1 634	52,9	3 086
	$P_{t_{6,7,8,9,10,11,12}}$	501	35,9	893	64,1	1 394
t_8	Q_{t_8}	1 446	46,8	1 644	53,2	3 090
	$P_{t_{6,7,8,9,10,11,12}}$	501	35,9	893	64,1	1 394
t_9	Q_{t_9}	1 423	46,8	1 620	53,2	3 043
	$P_{t_{6,7,8,9,10,11,12}}$	503	36,1	891	63,9	1 394
t_{10}	$Q_{t_{10}}$	1 302	45,7	1 546	54,3	2 848
	$P_{t_{6,7,8,9,10,11,12}}$	500	35,9	894	64,1	1 394
t_{11}	$Q_{t_{11}}$	1 252	45,5	1 502	54,5	2 754
	$P_{t_{6,7,8,9,10,11,12}}$	503	36,1	891	63,9	1 394
t_{12}	$Q_{t_{12}}$	1 269	45,7	1 509	54,3	2 778
	$P_{t_{6,7,8,9,10,11,12}}$	502	36,0	892	64,0	1 394

Beim durchschnittlichen Alter der Befragten lassen sich nur sehr geringe Unterschiede zwischen den Datensatztypen feststellen (Tabelle 5.47). Das Durchschnittsalter liegt im Querschnitt meist etwas höher als im Paneldatensatz, jedoch fällt diese Diskrepanz mit maximal 0,2 Jahren sehr gering aus. Zusätzlich weist die Altersverteilung im Panel eine geringere Streuung auf, sodass von einer größeren Altershomogenität im Paneldatensatz ausgegangen werden kann.

Die Zusammensetzung der Datensätze nach Schulform (Tabelle 5.48) zeigt, dass zu den Zeitpunkten t_6 und t_7 Gymnasiast/innen im Paneldatensatz überrepräsentiert sind, wohingegen Personen an Berufskollegs im Vergleich zum Querschnitt leicht unterrepräsentiert sind. Diese Unterschiede gestalten sich jedoch mit zunehmender Erhebungs-

Tabelle 5.47: Altersdurchschnitt $P_{6,7,8,9,10,11,12}$

		Alter	Std.-abw.	gesamt
t_6	Q_{t_6}	18,1	1,1	3 320
	$P_{6,7,8,9,10,11,12}$	18,0	0,8	1 391
t_7	Q_{t_7}	19,1	0,8	3 081
	$P_{6,7,8,9,10,11,12}$	18,9	0,8	1 394
t_8	Q_{t_8}	20,1	0,8	3 073
	$P_{6,7,8,9,10,11,12}$	19,9	0,8	1 392
t_9	Q_{t_9}	22,1	0,9	3 021
	$P_{6,7,8,9,10,11,12}$	22,0	0,8	1 391
t_{10}	$Q_{t_{10}}$	24,1	1,2	2 839
	$P_{6,7,8,9,10,11,12}$	23,9	0,8	1 393
t_{11}	$Q_{t_{11}}$	26,0	1,0	2 745
	$P_{6,7,8,9,10,11,12}$	26,0	0,8	1 390
t_{12}	$Q_{t_{12}}$	28,1	1,2	2 753
	$P_{6,7,8,9,10,11,12}$	28,0	0,8	1 386

dauer niedriger, sodass zum Zeitpunkt t_8 kaum noch nennenswerte Differenzen zwischen Panel- und Querschnittsdatensatz vorliegen.

Tabelle 5.48: Zusammensetzung $P_{t_6,7,8,9,10,11,12}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)							gesamt (100 %)	
		HS	RS	GS	GY	BK	SO	andere		keine
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3 219
	$P_{t_6,7,8,9,10,11,12}$	—	—	17,1	33,5	29,1	—	10,7	9,7	1 369
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2 975
	$P_{t_6,7,8,9,10,11,12}$	—	—	15,2	31,6	25,7	—	12,4	15,2	1 365
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2 934
	$P_{t_6,7,8,9,10,11,12}$	—	—	3,2	4,1	23,2	—	19,2	50,4	1 350

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.7.3 Sieben-Wellen-Panel 2008, 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_7,8,9,10,11,12,13}$)

Das Sieben-Wellen-Panel der Erhebungsjahre 2008 bis 2019 umfasst 1 327 Fälle. Bei Betrachtung der Geschlechterverteilung im Datensatz (Tabelle 5.49) wird eine Verzerrung des Panels $P_{t_7,8,9,10,11,12,13}$ deutlich. Der Längsschnittdatensatz enthält wesentlich mehr weibliche Personen als der Querschnittdatensatz. Die Differenz beträgt je nach Zeitpunkt zwischen etwa zehn bis gut elf Prozentpunkte. Umgekehrt sind männliche Personen im Paneldatensatz weniger vertreten als im Querschnitt.

Hinsichtlich des durchschnittlichen Alters der Befragten (Tabelle 5.50) zeigen sich nur leichte Abweichungen zwischen den Querschnittdatensätzen und dem Sieben-Wellen-Panel $P_{t_7,8,9,10,11,12,13}$. Das Durchschnittsalter ist im Querschnitt meist etwas höher – die Differenz ist mit 0,1 bis 0,2 Jahren jedoch marginal. An einigen Stellen zeigt sich außerdem, dass die Streuung in den Querschnittdatensätzen etwas höher ist, was auf eine stärkere Altersheterogenität der Befragten in den Querschnittdatensätzen hinweist.

Die Zusammensetzung nach Schulform (Tabelle 5.51) kann nur für die Zeitpunkte t_7 und t_8 betrachtet werden, da ab dem Zeitpunkt t_9 kaum noch Personen eine Schule besuchen und sich die Untersuchung der Zusammensetzung nach Schultyp daher ab dieser Erhebung erübrigt. Zum Erhebungszeitpunkt t_7 liegen Abweichungen derart vor, dass Schülerinnen und Schüler an Gymnasien etwas stärker im Paneldatensatz vertreten sind als im Querschnitt Q_{t_7} . Umgekehrt enthält das Panel etwas weniger Personen, die keine Schule mehr besuchen. Zum Zeitpunkt t_8 besucht bereits rund die Hälfte aller Befragten keine Schule mehr, sodass sich auch kaum bedeutsame Unterschiede zwischen den beiden Datensatztypen feststellen lassen.

Tabelle 5.49: Zusammensetzung nach Geschlecht $P_{t7,8,9,10,11,12,13}$

		männlich		weiblich		gesamt (100 %)
		<i>n</i>	%	<i>n</i>	%	
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t7,8,9,10,11,12,13}$	472	35,6	855	64,4	1327
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t7,8,9,10,11,12,13}$	472	35,6	855	64,4	1327
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t7,8,9,10,11,12,13}$	474	35,7	853	64,3	1327
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t7,8,9,10,11,12,13}$	471	35,5	856	64,5	1327
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t7,8,9,10,11,12,13}$	474	35,7	853	64,3	1327
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t7,8,9,10,11,12,13}$	473	35,6	854	64,4	1327
t_{13}	$Q_{t_{13}}$	1238	45,9	1460	54,1	2698
	$P_{t7,8,9,10,11,12,13}$	473	35,6	854	64,4	1327

Tabelle 5.50: Altersdurchschnitt $P_{t7,8,9,10,11,12,13}$

		Alter	Std.-abw.	gesamt
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t7,8,9,10,11,12,13}$	18,9	0,8	1327
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t7,8,9,10,11,12,13}$	19,9	0,8	1325
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t7,8,9,10,11,12,13}$	22,0	0,8	1323
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t7,8,9,10,11,12,13}$	23,9	0,8	1326
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t7,8,9,10,11,12,13}$	26,0	0,8	1322
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t7,8,9,10,11,12,13}$	28,0	0,8	1319
t_{13}	$Q_{t_{13}}$	30,1	0,9	2677
	$P_{t7,8,9,10,11,12,13}$	30,0	0,8	1325

Tabelle 5.51: Zusammensetzung $P_{7,8,9,10,11,12,13}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)							gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{7,8,9,10,11,12,13}$	—	—	14,2	31,4	26,4	—	12,7	15,2	1301
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{7,8,9,10,11,12,13}$	—	—	3,4	4,0	23,2	—	20,0	49,4	1282

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.8 Die Acht-Wellen-Panel

Alle (nach den Plausibilitätskontrollen) gültigen Fälle, die mindestens sieben weiteren Fällen im Datensatz zugeordnet werden konnten, bilden die Acht-Wellen-Panel. In diesem Bericht werden all jene Datensätze betrachtet, die Teilnahmen in acht direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke) und nicht bereits im vorherigen Methodenbericht (Daniel & Erdmann, 2017) beschrieben wurden.

5.8.1 Acht-Wellen-Panel 2005, 2006, 2007, 2008, 2009, 2011, 2013 und 2015 ($P_{t_{4,5,6,7,8,9,10,11}}$)

Das Acht-Wellen-Panel der Erhebungsjahre 2005 bis 2015 umfasst 1 285 Fälle. Mit Blick auf die Verteilung nach Geschlecht wird eine Verzerrung dieses Paneldatensatzes im Vergleich zu den Referenzdaten des Querschnitts deutlich. Im Paneldatensatz $P_{t_{4,5,6,7,8,9,10,11}}$ sind deutlich mehr weibliche Personen vertreten als in den Querschnittsdatsätzen. Diese Differenz beträgt je nach Zeitpunkt zwischen etwa neun und 14 Prozentpunkten. Männliche Personen sind entsprechend unterrepräsentiert. Am stärksten besteht diese Diskrepanz zum Zeitpunkt t_4 .

Tabelle 5.52: Zusammensetzung nach Geschlecht $P_{t_{4,5,6,7,8,9,10,11}}$

		männlich		weiblich		gesamt (100 %)
		<i>n</i>	%	<i>n</i>	%	
t_4	Q_{t_4}	1 717	50,4	1 688	49,6	3 405
	$P_{t_{4,5,6,7,8,9,10,11}}$	462	36,0	823	64,0	1 285
t_5	Q_{t_5}	2 265	49,8	2 282	50,2	4 547
	$P_{t_{4,5,6,7,8,9,10,11}}$	462	36,0	823	64,0	1 285
t_6	Q_{t_6}	1 567	47,0	1 768	53,0	3 335
	$P_{t_{4,5,6,7,8,9,10,11}}$	461	35,9	824	64,1	1 285
t_7	Q_{t_7}	1 452	47,1	1 634	52,9	3 086
	$P_{t_{4,5,6,7,8,9,10,11}}$	463	36,0	822	64,0	1 285
t_8	Q_{t_8}	1 446	46,8	1 644	53,2	3 090
	$P_{t_{4,5,6,7,8,9,10,11}}$	462	36,0	823	64,0	1 285
t_9	Q_{t_9}	1 423	46,8	1 620	53,2	3 043
	$P_{t_{4,5,6,7,8,9,10,11}}$	465	36,2	820	63,8	1 285
t_{10}	$Q_{t_{10}}$	1 302	45,7	1 546	54,3	2 848
	$P_{t_{4,5,6,7,8,9,10,11}}$	460	35,8	825	64,2	1 285
t_{11}	$Q_{t_{11}}$	1 252	45,5	1 502	54,5	2 754
	$P_{t_{4,5,6,7,8,9,10,11}}$	464	36,1	821	63,9	1 285

Das Durchschnittsalter der Befragten (Tabelle 5.53) liegt in den Querschnittsdatensätzen etwas höher als im Acht-Wellen-Panel $P_{t_{4,5,6,7,8,9,10,11}}$. Die Differenzen fallen jedoch mit 0,1 bis 0,2 Jahren marginal aus. Lediglich zum Zeitpunkt t_5 ist die Differenz mit 0,5 Jahren etwas höher. Zudem weisen die Querschnittsdatensätze durchweg eine etwas höhere Standardabweichung auf. Das Alter ist im Paneldatensatz also etwas homogener verteilt.

Tabelle 5.53: Altersdurchschnitt $P_{t_{4,5,6,7,8,9,10,11}}$

		Alter	Std.-abw.	gesamt
t_4	Q_{t_4}	16,1	0,8	3 313
	$P_{t_{4,5,6,7,8,9,10,11}}$	15,9	0,7	1 274
t_5	Q_{t_5}	17,4	1,3	4 468
	$P_{t_{4,5,6,7,8,9,10,11}}$	16,9	0,8	1 279
t_6	Q_{t_6}	18,1	1,1	3 320
	$P_{t_{4,5,6,7,8,9,10,11}}$	17,9	0,7	1 283
t_7	Q_{t_7}	19,1	0,8	3 081
	$P_{t_{4,5,6,7,8,9,10,11}}$	18,9	0,7	1 285
t_8	Q_{t_8}	20,1	0,8	3 073
	$P_{t_{4,5,6,7,8,9,10,11}}$	19,9	0,7	1 284
t_9	Q_{t_9}	22,1	0,9	3 021
	$P_{t_{4,5,6,7,8,9,10,11}}$	21,9	0,7	1 284
t_{10}	$Q_{t_{10}}$	24,1	1,2	2 839
	$P_{t_{4,5,6,7,8,9,10,11}}$	23,9	0,8	1 284
t_{11}	$Q_{t_{11}}$	26,0	1,0	2 745
	$P_{t_{4,5,6,7,8,9,10,11}}$	25,9	0,7	1 283

Auch mit Blick auf die Zusammensetzung nach der Schulform lassen sich Abweichungen zwischen dem Paneldatensatz und der Verteilung in den Querschnittsdatensätzen feststellen (Tabelle 5.54). Zu allen Erhebungszeitpunkten ist der Anteil an Gymnasiast/innen im Paneldatensatz im Vergleich zum jeweiligen Querschnitt erhöht. Andererseits fällt der Anteil an Personen, die keine Schule mehr besuchen im Datensatz $P_{t_{4,5,6,7,8,9,10,11}}$ kleiner aus, dabei insbesondere zu den Zeitpunkten t_6 und t_7 . Zum letzten Zeitpunkt t_8 bestehen nur noch geringe Abweichungen. Allerdings ist zu erwähnen, dass der Großteil der Befragten die Schule zu diesem Erhebungszeitpunkt bereits verlassen hat, sodass die Betrachtung der Verteilung nach Schulform nur noch wenig bedeutsam ist, weshalb sie für die späteren Zeitpunkte auch nicht mehr betrachtet wird.

Tabelle 5.54: Zusammensetzung $P_{t_{4,5,6,7,8,9,10,11}}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)								gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3 386
	$P_{t_{4,5,6,7,8,9,10,11}}$	12,2	24,0	30,3	32,8	0,2	0,2	0,3	—	1 282
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4 253
	$P_{t_{4,5,6,7,8,9,10,11}}$	0,2	0,4	21,1	38,9	39,3	0,1	—	—	1 185
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3 219
	$P_{t_{4,5,6,7,8,9,10,11}}$	—	—	17,5	34,0	29,1	—	10,3	9,1	1 263
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2 975
	$P_{t_{4,5,6,7,8,9,10,11}}$	—	—	15,4	32,1	26,0	—	12,6	13,9	1 263
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2 934
	$P_{t_{4,5,6,7,8,9,10,11}}$	—	—	3,2	4,2	23,0	—	19,0	50,6	1 246

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.8.2 Acht-Wellen-Panel 2006, 2007, 2008, 2009, 2011, 2013, 2015 und 2017 ($P_{t_{5,6,7,8,9,10,11,12}}$)

Das Acht-Wellen-Panel der Erhebungsjahre 2006 bis 2017 umfasst 1 339 Fälle. Wie Tabelle 5.55 zu entnehmen ist, liegen dabei hinsichtlich der Verteilung nach Geschlecht Unterschiede zwischen diesem Paneldatensatz und den Querschnitten der jeweiligen Erhebungswellen vor. Diese Differenzen gestalten sich derart, dass weibliche Personen im Paneldatensatz stärker vertreten sind als im Querschnitt. Entsprechend sind männliche Personen im Längsschnittdatensatz unterrepräsentiert. Je nach Zeitpunkt beträgt die Diskrepanz zwischen Panel- und Querschnittsdaten zwischen rund zehn bis 14 Prozentpunkte.

Tabelle 5.56 zeigt das durchschnittliche Alter der Befragten der Zeitpunkte t_5 bis t_8 jeweils für den Paneldatensatz sowie die Querschnitte der jeweiligen Erhebungswellen. Dabei wird ersichtlich, dass sich das Durchschnittsalter zwischen den Datensatztypen nur unwesentlich unterscheidet. In den Querschnitten liegt das Alter in den meisten Fällen etwas höher, diese Differenz ist mit 0,1 bis 0,2 Jahren jedoch sehr gering. Lediglich zum Zeitpunkt t_5 liegt mit einer Differenz von 0,5 Jahren ein etwas höherer Unterschied vor.

Auch hinsichtlich der Zusammensetzung nach Schulform, wie sie in Tabelle 5.57 dargestellt ist, lassen sich Unterschiede zwischen dem Acht-Wellen-Panel und den jeweiligen Querschnitten identifizieren. Es zeigt sich, dass Gymnasiast/innen sowie Gesamtschüler/innen zum Zeitpunkt t_5 im Paneldatensatz stärker vertreten sind als im

Tabelle 5.55: Zusammensetzung nach Geschlecht $P_{t_5,6,7,8,9,10,11,12}$

		männlich		weiblich		gesamt (100 %)
		<i>n</i>	%	<i>n</i>	%	
t_5	Q_{t_5}	2 265	49,8	2 282	50,2	4 547
	$P_{t_5,6,7,8,9,10,11,12}$	477	35,6	862	64,4	1 339
t_6	Q_{t_6}	1 567	47,0	1 768	53,0	3 335
	$P_{t_5,6,7,8,9,10,11,12}$	478	35,7	861	64,3	1 339
t_7	Q_{t_7}	1 452	47,1	1 634	52,9	3 086
	$P_{t_5,6,7,8,9,10,11,12}$	478	35,7	861	64,3	1 339
t_8	Q_{t_8}	1 446	46,8	1 644	53,2	3 090
	$P_{t_5,6,7,8,9,10,11,12}$	477	35,6	862	64,4	1 339
t_9	Q_{t_9}	1 423	46,8	1 620	53,2	3 043
	$P_{t_5,6,7,8,9,10,11,12}$	480	35,8	859	64,2	1 339
t_{10}	$Q_{t_{10}}$	1 302	45,7	1 546	54,3	2 848
	$P_{t_5,6,7,8,9,10,11,12}$	477	35,6	862	64,4	1 339
t_{11}	$Q_{t_{11}}$	1 252	45,5	1 502	54,5	2 754
	$P_{t_5,6,7,8,9,10,11,12}$	480	35,8	859	64,2	1 339
t_{12}	$Q_{t_{12}}$	1 269	45,7	1 509	54,3	2 778
	$P_{t_5,6,7,8,9,10,11,12}$	479	35,8	860	64,2	1 339

Querschnitt Q_{t_5} . Personen an Berufskollegs hingegen sind im Panel unterrepräsentiert. Diese Unterschiede nivellieren sich mit im Laufe der späteren Erhebungszeitpunkte. Zum Zeitpunkt t_8 bestehen kaum noch relevante Differenzen zwischen den zwei Datensätzen und die Hälfte aller Befragten hat die Schule zu diesem Zeitpunkt bereits verlassen, sodass die Verteilung nach Schulform für die weiteren Zeitpunkte nicht mehr betrachtet wird.

Tabelle 5.56: Altersdurchschnitt $P_{t_{5,6,7,8,9,10,11,12}}$

		Alter	Std.-abw.	gesamt
t_5	Q_{t_5}	17,4	1,3	4468
	$P_{t_{5,6,7,8,9,10,11,12}}$	17,0	0,9	1329
t_6	Q_{t_6}	18,1	1,1	3320
	$P_{t_{5,6,7,8,9,10,11,12}}$	17,9	0,8	1336
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t_{5,6,7,8,9,10,11,12}}$	18,9	0,8	1339
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t_{5,6,7,8,9,10,11,12}}$	19,9	0,7	1338
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_{5,6,7,8,9,10,11,12}}$	21,9	0,8	1337
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_{5,6,7,8,9,10,11,12}}$	23,9	0,8	1338
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_{5,6,7,8,9,10,11,12}}$	26,0	0,8	1335
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_{5,6,7,8,9,10,11,12}}$	28,0	0,8	1332

Tabelle 5.57: Zusammensetzung $P_{t_{5,6,7,8,9,10,11,12}}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)							gesamt (100 %)	
		HS	RS	GS	GY	BK	SO	andere		keine
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{5,6,7,8,9,10,11,12}}$	0,2	0,5	21,5	38,8	39,0	0,1	—	—	1239
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{5,6,7,8,9,10,11,12}}$	—	—	17,4	34,5	28,6	—	10,1	9,3	1316
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{5,6,7,8,9,10,11,12}}$	—	—	15,4	32,5	25,1	—	12,3	14,7	1313
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{5,6,7,8,9,10,11,12}}$	—	—	3,2	4,2	22,7	—	19,6	50,2	1297

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.8.3 Acht-Wellen-Panel 2007, 2008, 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_{6,7,8,9,10,11,12,13}}$)

Das Acht-Wellen-Panel der Erhebungsjahre 2007 bis 2019 umfasst 1 294 Fälle. Auch hier lässt sich erneut beim Vergleich von Panel- und Querschnittsdaten eine Verschiebung der Zusammensetzung nach Geschlecht feststellen (Tabelle 5.58). Weibliche Personen sind im Paneldatensatz deutlich häufiger vertreten als im Querschnitt, männliche Befragte im Panel hingegen seltener. Je nach betrachtetem Zeitpunkt liegt diese Differenz bei etwa elf bis zwölf Prozentpunkten.

Tabelle 5.58: *Zusammensetzung nach Geschlecht $P_{t_{6,7,8,9,10,11,12,13}}$*

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_6	Q_{t_6}	1 567	47,0	1 768	53,0	3 335
	$P_{t_{6,7,8,9,10,11,12,13}}$	454	35,1	840	64,9	1 294
t_7	Q_{t_7}	1 452	47,1	1 634	52,9	3 086
	$P_{t_{6,7,8,9,10,11,12,13}}$	455	35,2	839	64,8	1 294
t_8	Q_{t_8}	1 446	46,8	1 644	53,2	3 090
	$P_{t_{6,7,8,9,10,11,12,13}}$	455	35,2	839	64,8	1 294
t_9	Q_{t_9}	1 423	46,8	1 620	53,2	3 043
	$P_{t_{6,7,8,9,10,11,12,13}}$	456	35,2	838	64,8	1 294
t_{10}	$Q_{t_{10}}$	1 302	45,7	1 546	54,3	2 848
	$P_{t_{6,7,8,9,10,11,12,13}}$	454	35,1	840	64,9	1 294
t_{11}	$Q_{t_{11}}$	1 252	45,5	1 502	54,5	2 754
	$P_{t_{6,7,8,9,10,11,12,13}}$	457	35,3	837	64,7	1 294
t_{12}	$Q_{t_{12}}$	1 269	45,7	1 509	54,3	2 778
	$P_{t_{6,7,8,9,10,11,12,13}}$	456	35,2	838	64,8	1 294
t_{13}	$Q_{t_{13}}$	1 238	45,9	1 460	54,1	2 698
	$P_{t_{6,7,8,9,10,11,12,13}}$	456	35,2	838	64,8	1 294

Tabelle 5.59 zeigt das durchschnittliche Alter im Panel $P_{t_{6,7,8,9,10,11,12,13}}$ sowie vergleichend die Werte der Querschnitterhebungen. Bei nahezu allen Zeitpunkten liegt das Durchschnittsalter im Querschnitt leicht über dem in Acht-Wellen-Panel. Diese Abweichung beläuft sich dabei auf maximal 0,2 Jahre. Des Weiteren ist zu beobachten, dass für das Panel durchweg eine größere Altershomogenität festzustellen ist als in den Querschnitten, wie sich an der geringeren Standardabweichung ablesen lässt.

Auch hinsichtlich der Zusammensetzung nach Schulform sind Verschiebungen zwischen Panel und Querschnitt zu erkennen (Tabelle 5.60). Beispielsweise ist der Anteil an Schüler/innen, die ein Gymnasium besuchen, im Paneldatensatz höher als im Quer-

Tabelle 5.59: Altersdurchschnitt $P_{t_6,7,8,9,10,11,12,13}$

		Alter	Std.-abw.	gesamt
t_6	Q_{t_6}	18,1	1,1	3320
	$P_{t_6,7,8,9,10,11,12,13}$	17,9	0,8	1291
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t_6,7,8,9,10,11,12,13}$	18,9	0,8	1294
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t_6,7,8,9,10,11,12,13}$	19,9	0,8	1292
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_6,7,8,9,10,11,12,13}$	21,9	0,8	1291
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_6,7,8,9,10,11,12,13}$	23,9	0,8	1293
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_6,7,8,9,10,11,12,13}$	26,0	0,8	1290
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_6,7,8,9,10,11,12,13}$	27,9	0,8	1286
t_{13}	$Q_{t_{13}}$	30,1	0,9	2677
	$P_{t_6,7,8,9,10,11,12,13}$	30,0	0,8	1292

schnitt. Umgekehrt ist der Anteil an Schüler/innen an Berufskollegs sowie der Anteil an Personen, die keine Schule mehr besuchen, zu den Zeitpunkten t_6 und t_7 im Paneldatensatz geringer als im Querschnitt. In der achten Erhebungswelle bestehen kaum noch nennenswerte Differenzen in der Verteilung nach Schulform zwischen den Datensätzen, da der Großteil der Befragten die Schule zu diesem Zeitpunkt bereits verlassen hat. Eine Betrachtung der Verteilung nach Schulform erübrigt sich daher für die weiteren Erhebungszeitpunkte.

Tabelle 5.60: Zusammensetzung $P_{t_6,7,8,9,10,11,12,13}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)							gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_6,7,8,9,10,11,12,13}$	—	—	16,5	33,6	29,3	—	10,7	9,8	1269
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_6,7,8,9,10,11,12,13}$	—	—	14,6	31,5	26,3	—	12,5	15,1	1268
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_6,7,8,9,10,11,12,13}$	—	—	3,4	3,9	23,3	—	19,7	49,7	1253

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.9 Die Neun-Wellen-Panel

Alle (nach den Plausibilitätskontrollen) gültigen Fälle, die mindestens acht weiteren Fällen im Datensatz zugeordnet werden konnten, bilden die Neun-Wellen-Panel. In diesem Bericht werden all jene Datensätze betrachtet, die Teilnahmen in neun direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke) und nicht bereits im vorherigen Methodenbericht (Daniel & Erdmann, 2017) beschrieben wurden.

5.9.1 Neun-Wellen-Panel 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013 und 2015 ($P_{t_3,4,5,6,7,8,9,10,11}$)

Das Neun-Wellen-Panel der Erhebungsjahre 2004 bis 2015 umfasst 1 163 Fälle. Unter Betrachtung der in Tabelle 5.61 gezeigten Zusammensetzung dieses Panels nach Geschlecht wird im Vergleich zu den Referenzdaten des Querschnitts deutlich, dass eine Verschiebung dieses Panels vorliegt. Die Schulstatistiken des Zeitpunkts t_3 verdeutlicht, dass ein weitgehend ausgeglichenes Verhältnis im Anteil von weiblichen und männlichen Befragten vorliegt. Ab t_6 ist für die Querschnitte eine leichte Verschiebung dieser Relation erkennbar – der Frauenanteil liegt ab diesem Zeitpunkt stets bei rund 53 bis 54 Prozent. Dieser Anteil ist im Paneldatensatz $P_{t_3,4,5,6,7,8,9,10,11}$ nochmals erhöht und liegt je nach Zeitpunkt zwischen rund zehn bis 15 Prozentpunkte über dem Anteil in den Querschnitten. Entsprechend fällt der Anteil männlicher Befragter im Panel geringer aus.

Hinsichtlich des Durchschnittsalters sind nur geringfügige Abweichungen zwischen Querschnittsdatsätzen und dem Neun-Wellen-Panel erkennbar (Tabelle 5.62). Im Schnitt sind die Befragten im Querschnitt etwas älter, wobei diese Abweichung mit 0,1 bis 0,2 Jahren in den meisten Fällen sehr gering ist. Lediglich zum Zeitpunkt t_5 liegt mit

Tabelle 5.61: Zusammensetzung nach Geschlecht $P_{t_3,4,5,6,7,8,9,10,11}$

		männlich		weiblich		gesamt (100 %)
		<i>n</i>	%	<i>n</i>	%	
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_3,4,5,6,7,8,9,10,11}$	406	34,9	757	65,1	1163
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_3,4,5,6,7,8,9,10,11}$	411	35,3	752	64,7	1163
t_5	Q_{t_5}	2265	49,8	2282	50,2	4547
	$P_{t_3,4,5,6,7,8,9,10,11}$	411	35,3	752	64,7	1163
t_6	Q_{t_6}	1567	47,0	1768	53,0	3335
	$P_{t_3,4,5,6,7,8,9,10,11}$	410	35,3	753	64,7	1163
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t_3,4,5,6,7,8,9,10,11}$	412	35,4	751	64,6	1163
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_3,4,5,6,7,8,9,10,11}$	411	35,3	752	64,7	1163
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_3,4,5,6,7,8,9,10,11}$	412	35,4	751	64,6	1163
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_3,4,5,6,7,8,9,10,11}$	409	35,2	754	64,8	1163
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_3,4,5,6,7,8,9,10,11}$	412	35,4	751	64,6	1163

einer Differenz von 0,5 Jahren ein etwas höherer Unterschied vor. Zudem zeigt sich in den Querschnittsdatensätzen durchgehend eine etwas höhere Standardabweichung der Altersverteilung, was für eine größere Altershomogenität im Paneldatensatz spricht.

Auch hinsichtlich der Verteilung nach Schulform (Tabelle 5.63) sind Unterschiede zwischen den Querschnittsdatensätzen und dem Neun-Wellen-Panel erkennbar. So zeigt sich, dass der Anteil an Gymnasiast/innen im Paneldatensatz stets deutlich höher ist als in den Querschnitten der jeweiligen Erhebungswellen. Umgekehrt ist der Anteil an Schüler/innen an Hauptschulen im Panel niedriger. Ein Vergleich mit der Schulstatistik des Zeitpunktes t_3 zeigt, dass auch bereits der Querschnittsdatensatz Q_{t_3} nach dieser Systematik leicht von der offiziellen Verteilung abweicht. Mit zunehmender Erhebungsdauer reduzieren sich die Unterschiede zwischen den beiden Datensatztypen, was vor allem darauf zurückzuführen ist, dass ab t_8 bereits die Hälfte aller Personen keine Schule mehr besucht, sodass sich die Untersuchung der Zusammensetzung nach Schulform für die weiteren Zeitpunkte erübrigt.

Tabelle 5.62: Altersdurchschnitt $P_{t_{3,4,5,6,7,8,9,10,11}}$

		Alter	Std.-abw.	gesamt
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	14,9	0,7	1156
t_4	Q_{t_4}	16,1	0,8	3313
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	15,9	0,7	1155
t_5	Q_{t_5}	17,4	1,3	4468
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	16,9	0,8	1158
t_6	Q_{t_6}	18,1	1,1	3320
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	17,9	0,7	1162
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	18,9	0,7	1163
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	19,9	0,7	1163
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	21,9	0,7	1162
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	23,9	0,8	1162
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	25,9	0,7	1162

Tabelle 5.63: Zusammensetzung $P_{t_{3,4,5,6,7,8,9,10,11}}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)							gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_3	S_{t_3}	24,9	21,5	32,3	20,0	—	1,3	—	—	3 925
	Q_{t_3}	22,6	22,9	31,9	22,2	—	0,5	—	—	3 339
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	12,6	24,7	29,7	33,0	—	—	—	—	1 163
t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3 386
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	11,9	24,7	29,7	33,1	0,3	—	0,3	—	1 160
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4 253
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	0,2	0,5	21,5	38,9	38,9	—	—	—	1 071
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3 219
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	—	—	17,5	34,1	28,6	—	10,4	9,4	1 142
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2 975
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	—	—	15,7	32,4	25,4	—	12,5	14,0	1 142
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2 934
	$P_{t_{3,4,5,6,7,8,9,10,11}}$	—	—	3,3	4,2	22,7	—	19,1	50,8	1 128

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.9.2 Neun-Wellen-Panel 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015 und 2017 ($P_{t_{4,5,6,7,8,9,10,11,12}}$)

Das Neun-Wellen-Panel der Erhebungsjahre 2005 bis 2017 umfasst 1 214 Fälle. Die Zusammensetzung nach Geschlecht (Tabelle 5.64) zeigt, dass sich dieses Panel hinsichtlich der Verteilung nach Geschlecht von den Querschnittsdatensätzen unterscheidet. Im Paneldatensatz liegt der Anteil weiblicher Befragter je nach Zeitpunkt zwischen etwa zehn bis 15 Prozentpunkte über dem Anteil im jeweiligen Querschnitt. Entsprechend fällt der Anteil männlicher Befragter im Paneldatensatz im selben Maße geringer aus.

Tabelle 5.64: Zusammensetzung nach Geschlecht $P_{t_{4,5,6,7,8,9,10,11,12}}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_4	Q_{t_4}	1 717	50,4	1 688	49,6	3 405
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	431	35,5	783	64,5	1 214
t_5	Q_{t_5}	2 265	49,8	2 282	50,2	4 547
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	431	35,5	783	64,5	1 214
t_6	Q_{t_6}	1 567	47,0	1 768	53,0	3 335
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	430	35,4	784	64,6	1 214
t_7	Q_{t_7}	1 452	47,1	1 634	52,9	3 086
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	432	35,6	782	64,4	1 214
t_8	Q_{t_8}	1 446	46,8	1 644	53,2	3 090
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	431	35,5	783	64,5	1 214
t_9	Q_{t_9}	1 423	46,8	1 620	53,2	3 043
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	434	35,7	780	64,3	1 214
t_{10}	$Q_{t_{10}}$	1 302	45,7	1 546	54,3	2 848
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	429	35,3	785	64,7	1 214
t_{11}	$Q_{t_{11}}$	1 252	45,5	1 502	54,5	2 754
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	434	35,7	780	64,3	1 214
t_{12}	$Q_{t_{12}}$	1 269	45,7	1 509	54,3	2 778
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	433	35,7	781	64,3	1 214

Bezüglich des Durchschnittsalters, welches in Tabelle 5.65 dargestellt ist, lassen sich nur wenige Diskrepanzen zwischen den Datensätzen feststellen. In den Querschnitten liegt das durchschnittliche Alter stets etwas über dem Wert im Neun-Wellen-Panel, diese Differenz ist mit 0,1 bis 0,2 Jahren jedoch als gering zu bewerten. Lediglich zum Zeitpunkt t_5 fällt die Differenz mit 0,5 Jahren etwas höher aus. Zudem weisen die Querschnittsdatensätze eine etwas höhere Standardabweichung auf als der Paneldatensatz, woraus sich eine größere Altersheterogenität der Querschnittsdaten schließen lässt.

Tabelle 5.65: Altersdurchschnitt $P_{t_{4,5,6,7,8,9,10,11,12}}$

		Alter	Std.-abw.	gesamt
t_4	Q_{t_4}	16,1	0,8	3313
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	15,9	0,7	1203
t_5	Q_{t_5}	17,4	1,3	4468
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	16,9	0,8	1208
t_6	Q_{t_6}	18,1	1,1	3320
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	17,9	0,7	1212
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	18,9	0,7	1214
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	19,9	0,7	1213
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	21,9	0,7	1213
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	23,9	0,8	1214
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	25,9	0,7	1212
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	27,9	0,8	1209

Mit Blick auf die Verteilung nach Schulform in Tabelle 5.66 ist zu sehen, dass das Neun-Wellen-Panel von der Verteilung in den Querschnitten abweicht. Insbesondere zu den Zeitpunkten t_4 bis t_7 wird ersichtlich, dass der Anteil an Schüler/innen an Gymnasien im Panel deutlich erhöht ist. Umgekehrt sind Schüler/innen an Hauptschulen im Panel weniger vertreten, wie beispielsweise zum Zeitpunkt t_4 zu erkennen ist. Mit zunehmendem Alter nivellieren sich die Unterschiede der Verteilung nach Schulform, da immer weniger Personen überhaupt eine Schule besuchen. Daher erübrigt sich die Betrachtung der Verteilung nach Schulform ab dem Zeitpunkt t_9 .

Tabelle 5.66: Zusammensetzung $P_{t_{4,5,6,7,8,9,10,11,12}}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)								gesamt (100 %)
		HS	RS	GS	GY	BK	SO	andere	keine	
t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3 386
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	12,4	24,1	29,7	33,0	0,2	0,2	0,3	—	1 211
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4 253
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	0,2	0,4	20,8	39,4	39,1	0,1	—	—	1 120
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3 219
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	—	—	17,1	34,5	28,9	—	10,7	8,8	1 192
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2 975
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	—	—	15,0	32,5	25,7	—	12,6	14,2	1 195
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2 934
	$P_{t_{4,5,6,7,8,9,10,11,12}}$	—	—	3,2	4,2	23,3	—	19,1	50,1	1 178

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.9.3 Neun-Wellen-Panel 2006, 2007, 2008, 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_{5,6,7,8,9,10,11,12,13}}$)

Das Neun-Wellen-Panel der Erhebungsjahre 2006 bis 2019 umfasst 1 246 Fälle. Betrachtet man die Zusammensetzung dieses Panels nach Geschlecht (Tabelle 5.67), wird im Vergleich zu den Referenzdaten des Querschnitts eine Verschiebung des Panels $P_{t_{5,6,7,8,9,10,11,12,13}}$ deutlich. Der Anteil weiblicher Befragter fällt im Panel stets höher aus als im Querschnittsdatensatz – je nach Zeitpunkt beträgt diese Differenz zwischen rund elf und 15 Prozentpunkte. Umgekehrt sind männliche Personen im Längsschnittdatensatz im selben Maße unterrepräsentiert.

Tabelle 5.68 zeigt das durchschnittliche Alter der Befragten in den Querschnittsdatensätzen und dem Neun-Wellen-Panel. Insgesamt liegen nur geringfügige Abweichungen vor. Das Durchschnittsalter liegt dabei im Querschnitt in den meisten Fällen etwas höher. Diese Differenz ist mit 0,1 bis 0,2 Jahren jedoch als gering zu bewerten. Lediglich zum Zeitpunkt t_5 ist eine etwas höhere Abweichung in Höhe von 0,4 Jahren festzustellen. Zudem weisen die Querschnittsdaten meist eine etwas höhere Standardabweichung auf als der Paneldatensatz. Demnach liegt im Panel eine größere Altershomogenität vor.

Mit Blick auf die Verteilung nach Schulform (Tabelle 5.69) werden erneut Abweichungen zwischen dem Paneldatensatz und den Querschnitten ersichtlich. Insbesondere zu den Zeitpunkten t_4 , t_5 und t_6 fällt auf, dass der Anteil an Gymnasiast/innen im Panel deutlich höher ausfällt als in den Querschnitten der jeweiligen Zeitpunkte. Gleichzeitig

Tabelle 5.67: Zusammensetzung nach Geschlecht $P_{t_{5,6,7,8,9,10,11,12,13}}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_5	Q_{t_5}	2265	49,8	2282	50,2	4547
	$P_{t_{5,6,7,8,9,10,11,12,13}}$	432	34,7	814	65,3	1246
t_6	Q_{t_6}	1567	47,0	1768	53,0	3335
	$P_{t_{5,6,7,8,9,10,11,12,13}}$	434	34,8	812	65,2	1246
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t_{5,6,7,8,9,10,11,12,13}}$	434	34,8	812	65,2	1246
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_{5,6,7,8,9,10,11,12,13}}$	433	34,8	813	65,2	1246
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_{5,6,7,8,9,10,11,12,13}}$	435	34,9	811	65,1	1246
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_{5,6,7,8,9,10,11,12,13}}$	433	34,8	813	65,2	1246
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_{5,6,7,8,9,10,11,12,13}}$	436	35,0	810	65,0	1246
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_{5,6,7,8,9,10,11,12,13}}$	435	34,9	811	65,1	1246
t_{13}	$Q_{t_{13}}$	1238	45,9	1460	54,1	2698
	$P_{t_{5,6,7,8,9,10,11,12,13}}$	435	34,9	811	65,1	1246

ist der Anteil an Schüler/innen an Berufskollegs im Panel etwas geringer sowie auch der Anteil an der Personen, die keine Schule mehr besuchen. Zum Zeitpunkt t_8 haben sich die Unterschiede zwischen den Datensatztypen weitgehend nivelliert, was vor allem darauf zurück zu führen ist, dass zu diesem Erhebungszeitpunkt nur noch die Hälfte aller Personen überhaupt eine Schule besucht. Eine Betrachtung der Zusammensetzung nach Schulform erübrigt sich daher für die weiteren Zeitpunkte.

Tabelle 5.68: Altersdurchschnitt $P_{t_5,6,7,8,9,10,11,12,13}$

		Alter	Std.-abw.	gesamt
t_5	Q_{t_5}	17,4	1,3	4468
	$P_{t_5,6,7,8,9,10,11,12,13}$	17,0	0,9	1237
t_6	Q_{t_6}	18,1	1,1	3320
	$P_{t_5,6,7,8,9,10,11,12,13}$	17,9	0,8	1243
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t_5,6,7,8,9,10,11,12,13}$	18,9	0,8	1246
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t_5,6,7,8,9,10,11,12,13}$	19,9	0,8	1245
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_5,6,7,8,9,10,11,12,13}$	21,9	0,8	1244
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_5,6,7,8,9,10,11,12,13}$	23,9	0,8	1245
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_5,6,7,8,9,10,11,12,13}$	26,0	0,8	1242
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_5,6,7,8,9,10,11,12,13}$	27,9	0,8	1239
t_{13}	$Q_{t_{13}}$	30,1	0,9	2677
	$P_{t_5,6,7,8,9,10,11,12,13}$	30,0	0,8	1244

Tabelle 5.69: Zusammensetzung $P_{t_5,6,7,8,9,10,11,12,13}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)							andere	keine	gesamt (100 %)
		HS	RS	GS	GY	BK	SO				
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253	
	$P_{t_5,6,7,8,9,10,11,12,13}$	0,2	0,4	20,6	38,8	39,9	0,1	—	—	1155	
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
	$P_{t_5,6,7,8,9,10,11,12,13}$	—	—	16,8	34,6	28,8	—	10,2	9,6	1223	
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975	
	$P_{t_5,6,7,8,9,10,11,12,13}$	—	—	14,7	32,4	25,8	—	12,3	14,8	1223	
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934	
	$P_{t_5,6,7,8,9,10,11,12,13}$	—	—	3,4	4,1	22,7	—	20,0	49,8	1207	

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.10 Die Zehn-Wellen-Panel

Alle (nach den Plausibilitätskontrollen) gültigen Fälle, die mindestens neun weiteren Fällen im Datensatz zugeordnet werden konnten, bilden die Zehn-Wellen-Panel. In diesem Bericht werden all jene Datensätze betrachtet, die Teilnahmen in zehn direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke) und nicht bereits im vorherigen Methodenbericht (Daniel & Erdmann, 2017) beschrieben wurden.

5.10.1 Zehn-Wellen-Panel 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013 und 2015 ($P_{t_{2,3,4,5,6,7,8,9,10,11}}$)

Das Zehn-Wellen-Panel der Erhebungsjahre 2003 bis 2015 umfasst 1 000 Fälle. Unter Betrachtung der in Tabelle 5.70 gezeigten Zusammensetzung dieses Panels nach Geschlecht wird im Vergleich zu den Referenzdaten des Querschnitts deutlich, dass eine Verschiebung dieses Panels vorliegt. Die Schulstatistiken der Zeitpunkte t_2 und t_3 verdeutlichen, dass ein weitgehend ausgeglichenes Verhältnis im Anteil von weiblichen und männlichen Befragten vorliegt. Ab t_6 ist für die Querschnitte eine leichte Verschiebung dieser Relation erkennbar – der Frauenanteil liegt ab diesem Zeitpunkt stets bei rund 53 bis 54 Prozent. Dieser Anteil ist im Paneldatensatz $P_{t_{2,3,4,5,6,7,8,9,10,11}}$ nochmals erhöht und liegt je nach Zeitpunkt zwischen elf und 16 Prozentpunkte über dem Anteil in den Querschnitten. Entsprechend fällt der Anteil männlicher Befragter im Panel geringer aus.

Hinsichtlich des Durchschnittsalters sind nur geringfügige Abweichungen zwischen Querschnittsdatsätzen und dem hier betrachteten Zehn-Wellen-Panel erkennbar (Tabelle 5.71). Im Schnitt sind die Befragten in den Querschnittsdatsätzen etwas älter, wobei diese Abweichung mit 0,1 bis 0,3 Jahren in den meisten Fällen sehr gering ist. Lediglich zum Zeitpunkt t_5 liegt mit einer Differenz von 0,5 Jahren ein etwas höherer Unterschied vor. Zudem zeigt sich in den Querschnittsdatsätzen durchgehend eine etwas höhere Standardabweichung der Altersverteilung, was für eine größere Altershomogenität im Paneldatensatz spricht.

Auch hinsichtlich der Verteilung nach Schulform (Tabelle 5.72) sind Unterschiede zwischen den Querschnittsdatsätzen und dem betrachteten Zehn-Wellen-Panel erkennbar. So zeigt sich, dass der Anteil an Gymnasiast/innen im Paneldatensatz stets deutlich höher ist als in den Querschnitten der jeweiligen Erhebungswellen. Umgekehrt ist der Anteil an Schüler/innen an Hauptschulen im Panel niedriger. Ein Vergleich mit der Schulstatistik der Zeitpunkte t_2 und t_3 zeigt, dass auch bereits die Querschnittsdatsätze nach dieser Systematik leicht von der offiziellen Verteilung abweichen. Mit zunehmender Erhebungsdauer reduzieren sich die Unterschiede zwischen den beiden Datensatztypen, was vor allem darauf zurückzuführen ist, dass ab t_8 bereits die Hälfte aller Personen

Tabelle 5.70: Zusammensetzung nach Geschlecht $P_{t_2,3,4,5,6,7,8,9,10,11}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_2,3,4,5,6,7,8,9,10,11}$	345	34,5	655	65,5	1000
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_2,3,4,5,6,7,8,9,10,11}$	341	34,1	659	65,9	1000
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_2,3,4,5,6,7,8,9,10,11}$	344	34,4	656	65,6	1000
t_5	Q_{t_5}	2265	49,8	2282	50,2	4547
	$P_{t_2,3,4,5,6,7,8,9,10,11}$	343	34,3	657	65,7	1000
t_6	Q_{t_6}	1567	47,0	1768	53,0	3335
	$P_{t_2,3,4,5,6,7,8,9,10,11}$	343	34,3	657	65,7	1000
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t_2,3,4,5,6,7,8,9,10,11}$	345	34,5	655	65,5	1000
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_2,3,4,5,6,7,8,9,10,11}$	344	34,4	656	65,6	1000
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_2,3,4,5,6,7,8,9,10,11}$	345	34,5	655	65,5	1000
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_2,3,4,5,6,7,8,9,10,11}$	342	34,2	658	65,8	1000
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_2,3,4,5,6,7,8,9,10,11}$	345	34,5	655	65,5	1000

keine Schule mehr besucht, sodass sich die Untersuchung der Zusammensetzung nach Schulform für die weiteren Zeitpunkte erübrigt.

Tabelle 5.71: Altersdurchschnitt $P_{2,3,4,5,6,7,8,9,10,11}$

		Alter	Std.-abw.	gesamt
t_2	Q_{t_2}	14,0	0,8	3270
	$P_{2,3,4,5,6,7,8,9,10,11}$	13,8	0,7	988
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{2,3,4,5,6,7,8,9,10,11}$	14,9	0,7	996
t_4	Q_{t_4}	16,1	0,8	3313
	$P_{2,3,4,5,6,7,8,9,10,11}$	15,8	0,7	993
t_5	Q_{t_5}	17,4	1,3	4468
	$P_{2,3,4,5,6,7,8,9,10,11}$	16,9	0,8	995
t_6	Q_{t_6}	18,1	1,1	3320
	$P_{2,3,4,5,6,7,8,9,10,11}$	17,8	0,7	999
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{2,3,4,5,6,7,8,9,10,11}$	18,8	0,7	1000
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{2,3,4,5,6,7,8,9,10,11}$	19,9	0,7	1000
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{2,3,4,5,6,7,8,9,10,11}$	21,9	0,7	999
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{2,3,4,5,6,7,8,9,10,11}$	23,8	0,8	999
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{2,3,4,5,6,7,8,9,10,11}$	25,9	0,7	999

Tabelle 5.72: Zusammensetzung $P_{t_{2,3,4,5,6,7,8,9,10,11}}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)							gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_2	S_{t_2}	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
	Q_{t_2}	24,6	23,6	31,4	20,1	—	0,4	—	—	3392
	$P_{t_{2,3,4, \dots, 9,10,11}}$	12,6	25,0	30,7	31,7	—	—	—	—	1000
t_3	S_{t_3}	24,9	21,5	32,3	20,0	—	1,3	—	—	3925
	Q_{t_3}	22,6	22,9	31,9	22,2	—	0,5	—	—	3339
	$P_{t_{2,3,4, \dots, 9,10,11}}$	12,7	25,6	30,8	30,9	—	—	—	—	1000
t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_{2,3,4, \dots, 9,10,11}}$	12,2	25,5	30,8	31,0	0,1	—	0,3	—	999
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{2,3,4, \dots, 9,10,11}}$	0,2	0,5	23,1	36,9	39,2	—	—	—	929
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{2,3,4, \dots, 9,10,11}}$	—	—	19,2	32,4	29,1	—	10,8	8,4	982
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{2,3,4, \dots, 9,10,11}}$	—	—	17,1	31,1	25,7	—	12,6	13,5	981
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{2,3,4, \dots, 9,10,11}}$	—	—	3,6	4,5	22,8	—	18,9	50,1	971

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.10.2 Zehn-Wellen-Panel 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015 und 2017 ($P_{t_{3,4,5,6,7,8,9,10,11,12}}$)

Das Zehn-Wellen-Panel der Erhebungsjahre 2004 bis 2017 umfasst 1 100 Fälle. Unter Betrachtung der in Tabelle 5.73 gezeigten Zusammensetzung dieses Panels nach Geschlecht wird im Vergleich zu den Referenzdaten des Querschnitts deutlich, dass eine Verschiebung dieses Panels vorliegt. Die Schulstatistik des Zeitpunktes t_3 sowie die Querschnittsdaten der Zeitpunkte t_3 bis t_5 verdeutlichen, dass ein weitgehend ausgeglichenes Verhältnis im Anteil von weiblichen und männlichen Befragten vorliegt. Ab t_6 ist für die Querschnitte eine leichte Verschiebung dieser Relation erkennbar – der Frauenanteil liegt ab diesem Zeitpunkt stets bei rund 53 bis 54 Prozent. Dieser Anteil ist im Paneldatensatz $P_{t_{3,4,5,6,7,8,9,10,11,12}}$ nochmals erhöht und liegt je nach Zeitpunkt zwischen rund zehn und 15 Prozentpunkten über dem Anteil in den Querschnitten. Entsprechend fällt der Anteil männlicher Befragter im Panel geringer aus.

Tabelle 5.74 zeigt das durchschnittliche Alter im hier betrachteten Zehn-Wellen-Panel sowie vergleichend die Werte der jeweiligen Querschnittserhebungen. Zu allen zehn Zeitpunkten liegt das Durchschnittsalter im Querschnitt leicht über dem im Panel. Diese Abweichung beläuft sich dabei auf 0,1 bis 0,3 Jahre, einzig zum Zeitpunkt t_5 auf 0,5 Jahre. Des Weiteren ist zu beobachten, dass für das Panel durchweg eine größere Altershomogenität festzustellen ist als in den Querschnitten.

Auch hinsichtlich der Verteilung nach Schulform in Tabelle 5.75 sind Unterschiede zwischen den Querschnittsdatensätzen und dem hier betrachteten Panel erkennbar. Unter Hinzunahme der Daten der Schulstatistik zum Zeitpunkt t_3 zeigt sich, dass im Paneldatensatz nur etwa halb so viele Hauptschüler/innen vertreten sind wie in den Daten S_{t_3} . Umgekehrt enthält das Panel verglichen mit dieser Schulstatistik über 13 Prozent mehr Schüler/innen an Gymnasien. Auch beim Vergleich von Schulstatistik und Querschnitt fällt auf, dass auch die Querschnittsdatensätze nach dieser Systematik bereits leicht verzerrt sind. Mit zunehmender Erhebungsdauer reduzieren sich die Unterschiede zwischen den beiden Datensatztypen, was vor allem darauf zurückzuführen ist, dass ab t_8 bereits die Hälfte aller Personen keine Schule mehr besucht, sodass sich die Untersuchung der Zusammensetzung nach Schulform für die weiteren Zeitpunkte erübrigt.

Tabelle 5.73: Zusammensetzung nach Geschlecht $P_{t_{3,4,5,6,7,8,9,10,11,12}}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	381	34,6	719	65,4	1100
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	386	35,1	714	64,9	1100
t_5	Q_{t_5}	2265	49,8	2282	50,2	4547
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	386	35,1	714	64,9	1100
t_6	Q_{t_6}	1567	47,0	1768	53,0	3335
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	385	35,0	715	65,0	1100
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	387	35,2	713	64,8	1100
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	386	35,1	714	64,9	1100
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	387	35,2	713	64,8	1100
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	384	34,9	716	65,1	1100
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	388	35,3	712	64,7	1100
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	388	35,3	712	64,7	1100

Tabelle 5.74: Altersdurchschnitt $P_{t_{3,4,5,6,7,8,9,10,11,12}}$

		Alter	Std.-abw.	gesamt
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	14,9	0,7	1093
t_4	Q_{t_4}	16,1	0,8	3313
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	15,9	0,7	1092
t_5	Q_{t_5}	17,4	1,3	4468
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	16,9	0,8	1095
t_6	Q_{t_6}	18,1	1,1	3320
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	17,9	0,7	1099
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	18,9	0,7	1100
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	19,9	0,7	1100
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	21,9	0,7	1099
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	23,8	0,8	1100
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	25,9	0,7	1099
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_{3,4,5,6,7,8,9,10,11,12}}$	27,9	0,8	1095

Tabelle 5.75: Zusammensetzung $P_{3,4,5,6,7,8,9,10,11,12}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)							gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_3	S_{t_3}	24,9	21,5	32,3	20,0	—	1,3	—	—	3925
	Q_{t_3}	22,6	22,9	31,9	22,2	—	0,5	—	—	3339
	$P_{t_{3,4,5, \dots, 10,11,12}}$	12,6	24,8	29,2	33,4	—	—	—	—	1100
t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_{3,4,5, \dots, 10,11,12}}$	11,9	24,8	29,2	33,4	0,3	—	0,4	—	1097
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{3,4,5, \dots, 10,11,12}}$	0,2	0,5	21,1	39,6	38,6	—	—	—	1013
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{3,4,5, \dots, 10,11,12}}$	—	—	17,1	34,8	28,4	—	10,8	9,1	1079
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{3,4,5, \dots, 10,11,12}}$	—	—	15,2	32,9	25,0	—	12,6	14,3	1082
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{3,4,5, \dots, 10,11,12}}$	—	—	3,3	4,2	23,0	—	19,2	50,3	1068

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.10.3 Zehn-Wellen-Panel 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_{4,5,6,7,8,9,10,11,12,13}}$)

Das Zehn-Wellen-Panel der Erhebungsjahre 2004 bis 2017 umfasst 1 127 Fälle.

Unter Betrachtung der in Tabelle 5.76 gezeigten Zusammensetzung dieses Panels nach Geschlecht wird im Vergleich zu den Referenzdaten des Querschnitts deutlich, dass eine Verschiebung dieses Panels vorliegt. Je nach Zeitpunkt liegt der Anteil weiblicher Befragter zwischen elf und knapp 16 Prozentpunkte über dem Anteil im Querschnitt. Männliche Befragte sind entsprechend im selben Maße im Panel unterrepräsentiert.

Tabelle 5.76: Zusammensetzung nach Geschlecht $P_{t_{4,5,6,7,8,9,10,11,12,13}}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	390	34,6	737	65,4	1127
t_5	Q_{t_5}	2265	49,8	2282	50,2	4547
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	390	34,6	737	65,4	1127
t_6	Q_{t_6}	1567	47,0	1768	53,0	3335
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	390	34,6	737	65,4	1127
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	392	34,8	735	65,2	1127
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	391	34,7	736	65,3	1127
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	393	34,9	734	65,1	1127
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	389	34,5	738	65,5	1127
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	394	35,0	733	65,0	1127
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	393	34,9	734	65,1	1127
t_{13}	$Q_{t_{13}}$	1238	45,9	1460	54,1	2698
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	393	34,9	734	65,1	1127

Tabelle 5.77 zeigt das durchschnittliche Alter im Panel $P_{t_{4,5,6,7,8,9,10,11,12,13}}$ sowie vergleichend die Werte der jeweiligen Querschnittserhebungen. Zu allen zehn Zeitpunkten liegt das Durchschnittsalter im Querschnitt leicht über dem im Panel. Diese Abweichung beläuft sich dabei meist auf 0,1 bis 0,2 Jahre, einzig zum Zeitpunkt t_5 auf 0,5 Jahre. Des Weiteren ist zu beobachten, dass für das Panel durchweg eine größere Al-

tershomogenität festzustellen ist als in den Querschnitten, was anhand der geringeren Standardabweichung zu erkennen ist.

Tabelle 5.77: Altersdurchschnitt $P_{t_{4,5,6,7,8,9,10,11,12,13}}$

		Alter	Std.-abw.	gesamt
t_4	Q_{t_4}	16,1	0,8	3 313
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	15,9	0,7	1 116
t_5	Q_{t_5}	17,4	1,3	4 468
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	16,9	0,8	1 121
t_6	Q_{t_6}	18,1	1,1	3 320
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	17,9	0,7	1 125
t_7	Q_{t_7}	19,1	0,8	3 081
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	18,9	0,7	1 127
t_8	Q_{t_8}	20,1	0,8	3 073
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	19,9	0,7	1 126
t_9	Q_{t_9}	22,1	0,9	3 021
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	21,9	0,7	1 126
t_{10}	$Q_{t_{10}}$	24,1	1,2	2 839
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	23,9	0,8	1 127
t_{11}	$Q_{t_{11}}$	26,0	1,0	2 745
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	25,9	0,7	1 125
t_{12}	$Q_{t_{12}}$	28,1	1,2	2 753
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	27,9	0,8	1 122
t_{13}	$Q_{t_{13}}$	30,1	0,9	2 677
	$P_{t_{4,5,6,7,8,9,10,11,12,13}}$	29,9	0,7	1 126

Hinsichtlich der Verteilung nach Schulform (Tabelle 5.78) sind Unterschiede zwischen den Querschnittsdatensätzen und dem hier betrachteten Panel erkennbar. Beim Vergleich von Paneldatensatz und Querschnitt zu den Zeitpunkten t_4 und t_5 fällt auf, dass der Längsschnittdatensatz mehr Gymnasiast/innen enthält als die Querschnitte, umgekehrt aber weniger Schüler/innen an Hauptschulen. Im weiteren Verlauf ist zu den Zeitpunkten t_6 und t_7 zu erkennen, dass das Panel $P_{t_{4,5,6,7,8,9,10,11,12,13}}$ weniger Personen enthält, die keine Schule mehr besuchen, als die jeweiligen Querschnitte dieser Erhebungsjahre. Hierbei handelt es sich vermutlich um die Personen, die zuvor an Haupt- und Realschulen waren und in den vorherigen Zeitpunkten bereits im Panel unterrepräsentiert waren. Mit zunehmender Erhebungsdauer reduzieren sich die Unterschiede zwischen den beiden Datensatztypen, was vor allem darauf zurückzuführen ist, dass ab t_8 bereits die Hälfte aller Personen keine Schule mehr besucht, sodass sich die Untersuchung der Zusammensetzung nach Schulform für die weiteren Zeitpunkte erübrigt.

Tabelle 5.78: Zusammensetzung $P_{t_{4,5,6,7,8,9,10,11,12,13}}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)								gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3 386
	$P_{t_{4,5,6, \dots, 11,12,13}}$	12,6	24,2	29,4	33,1	0,2	0,2	0,4	—	1 124
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4 253
	$P_{t_{4,5,6, \dots, 11,12,13}}$	0,2	0,4	20,1	39,3	40,0	0,1	—	—	1 041
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3 219
	$P_{t_{4,5,6, \dots, 11,12,13}}$	—	—	16,7	34,5	29,0	—	10,9	9,0	1 105
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2 975
	$P_{t_{4,5,6, \dots, 11,12,13}}$	—	—	14,5	32,2	26,3	—	12,8	14,2	1 111
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2 934
	$P_{t_{4,5,6, \dots, 11,12,13}}$	—	—	3,4	4,0	23,3	—	19,5	49,8	1 093

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.11 Die Elf-Wellen-Panel

Alle (nach den Plausibilitätskontrollen) gültigen Fälle, die mindestens zehn weiteren Fällen im Datensatz zugeordnet werden konnten, bilden die Elf-Wellen-Panel. In diesem Bericht werden nur jene Datensätze beschrieben, die Teilnahmen in elf direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke).

5.11.1 Elf-Wellen-Panel 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013 und 2015 ($P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$)

Das Elf-Wellen-Panel der Erhebungsjahre 2002 bis 2015 umfasst 832 Fälle. Unter Betrachtung der in Tabelle 5.70 gezeigten Zusammensetzung dieses Panels nach Geschlecht wird im Vergleich zu den Referenzdaten des Querschnitts deutlich, dass eine Verschiebung dieses Panels vorliegt. Die Schulstatistiken der Zeitpunkte t_1 bis t_3 sowie die Querschnitte der Wellen t_1 bis t_5 verdeutlichen, dass ein weitgehend ausgeglichenes Verhältnis im Anteil von weiblichen und männlichen Befragten vorliegt. Ab t_6 ist für die Querschnitte eine leichte Verschiebung dieser Relation erkennbar – der Frauenanteil liegt ab diesem Zeitpunkt stets bei rund 53 bis 54 Prozent. Dieser Anteil ist im Paneldatensatz $P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$ nochmals erhöht und liegt je nach Zeitpunkt zwischen rund zwölf und gut 17 Prozentpunkten über dem Anteil in den Querschnitten. Entsprechend fällt der Anteil männlicher Befragter im Panel geringer aus.

Mit Blick auf das Durchschnittsalter sind nur geringfügige Abweichungen zwischen Querschnittsdatensätzen und dem hier betrachteten Elf-Wellen-Panel erkennbar (Tabelle 5.80). Im Durchschnitt sind die Befragten in den Querschnittsdatensätzen etwas älter, wobei diese Abweichung mit 0,2 bis 0,3 Jahren in den meisten Fällen sehr gering ist. Lediglich zum Zeitpunkt t_5 liegt mit einer Differenz von 0,5 Jahren ein etwas höherer Unterschied vor. Zudem zeigt sich, dass in den Querschnittsdatensätzen durchgehend eine etwas höhere Standardabweichung der Altersverteilung vorliegt, was für eine größere Altershomogenität im Paneldatensatz spricht.

Die in Tabelle 5.81 dargestellte Verteilung nach Schulform zeigt Unterschiede zwischen den Querschnittsdatensätzen und dem hier betrachteten Panel $P_{t_1,2,3,4,5,6,7,8,9,10,11}$. Beim Vergleich mit der Schulstatistik der Erhebungswellen t_1 bis t_3 fällt auf, dass im Paneldatensatz deutlich weniger Schüler/innen an Hauptschulen vertreten sind, umgekehrt aber wesentlich mehr Schüler/innen an Gymnasien. Es zeigt sich zudem, dass auch bereits die Querschnittsdatensätze leicht nach dieser Systematik verzerrt sind. Mit zunehmender Erhebungsdauer reduzieren sich die Unterschiede zwischen den beiden Datensatztypen, was vor allem darauf zurückzuführen ist, dass ab t_8 bereits die Hälfte aller Personen keine Schule mehr besucht, sodass sich die Untersuchung der Zusammensetzung nach Schulform für die weiteren Zeitpunkte erübrigt.

Tabelle 5.79: Zusammensetzung nach Geschlecht $P_{t_1,2,3,4,5,6,7,8,9,10,11}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_1	S_{t_1}	2003	51,2	1907	48,8	3910
	Q_{t_1}	1728	50,7	1679	49,3	3407
	$P_{t_1,2,3,4,5,6,7,8,9,10,11}$	276	33,2	556	66,8	832
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_1,2,3,4,5,6,7,8,9,10,11}$	281	33,8	551	66,2	832
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_1,2,3,4,5,6,7,8,9,10,11}$	277	33,3	555	66,7	832
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_1,2,3,4,5,6,7,8,9,10,11}$	278	33,4	554	66,6	832
t_5	Q_{t_5}	2265	49,8	2282	50,2	4547
	$P_{t_1,2,3,4,5,6,7,8,9,10,11}$	280	33,7	552	66,3	832
t_6	Q_{t_6}	1567	47,0	1768	53,0	3335
	$P_{t_1,2,3,4,5,6,7,8,9,10,11}$	279	33,5	553	66,5	832
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t_1,2,3,4,5,6,7,8,9,10,11}$	281	33,8	551	66,2	832
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_1,2,3,4,5,6,7,8,9,10,11}$	280	33,7	552	66,3	832
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_1,2,3,4,5,6,7,8,9,10,11}$	280	33,7	552	66,3	832
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_1,2,3,4,5,6,7,8,9,10,11}$	278	33,4	554	66,6	832
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_1,2,3,4,5,6,7,8,9,10,11}$	280	33,7	552	66,3	832

Tabelle 5.80: Altersdurchschnitt $P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$

		Alter	Std.-abw.	gesamt
t_1	Q_{t_1}	13,0	0,8	3346
	$P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$	12,8	0,6	829
t_2	Q_{t_2}	14,0	0,8	3270
	$P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$	13,8	0,6	824
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$	14,9	0,6	828
t_4	Q_{t_4}	16,1	0,8	3313
	$P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$	15,8	0,6	827
t_5	Q_{t_5}	17,4	1,3	4468
	$P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$	16,9	0,6	829
t_6	Q_{t_6}	18,1	1,1	3320
	$P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$	17,8	0,7	831
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$	18,8	0,7	832
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$	19,8	0,7	832
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$	21,8	0,7	832
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$	23,8	0,8	831
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_{1,2,3,4,5,6,7,8,9,10,11}}$	25,8	0,7	831

Tabelle 5.81: Zusammensetzung $P_{t_1,2,3,4,5,6,7,8,9,10,11}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)								gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_1	S_{t_1}	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
	Q_{t_1}	21,9	23,6	31,2	22,8	—	0,4	—	—	3411
	$P_{t_1,2,3, \dots, 9,10,11}$	11,5	26,7	30,2	31,6	—	—	—	—	832
t_2	S_{t_2}	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
	Q_{t_2}	24,6	23,6	31,4	20,1	—	0,4	—	—	3392
	$P_{t_1,2,3, \dots, 9,10,11}$	11,9	26,8	30,2	31,1	—	—	—	—	832
t_3	S_{t_3}	24,9	21,5	32,3	20,0	—	1,3	—	—	3925
	Q_{t_3}	22,6	22,9	31,9	22,2	—	0,5	—	—	3339
	$P_{t_1,2,3, \dots, 9,10,11}$	12,1	27,0	30,2	30,6	—	—	—	—	832
t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_1,2,3, \dots, 9,10,11}$	11,7	27,1	29,8	30,9	0,1	—	0,4	—	831
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_1,2,3, \dots, 9,10,11}$	0,3	0,5	23,3	37,4	38,6	—	—	—	773
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_1,2,3, \dots, 9,10,11}$	—	—	19,0	32,4	29,0	—	11,4	8,2	817
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_1,2,3, \dots, 9,10,11}$	—	—	16,8	30,8	25,4	—	13,3	13,7	819
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_1,2,3, \dots, 9,10,11}$	—	—	3,8	4,5	21,6	—	19,2	50,9	806

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.11.2 Elf-Wellen-Panel 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015 und 2017 ($P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$)

Das Elf-Wellen-Panel der Erhebungsjahre 2003 bis 2017 umfasst 950 Fälle. Betrachtet man die Zusammensetzung dieses Panels im Vergleich zu den Querschnittsdatensätzen, so wird eine Verzerrung des Paneldatensatzes deutlich. Der Anteil weiblicher Befragter liegt im Panel $P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$ je nach Zeitpunkt zwischen etwa elf und 16 Prozentpunkten über dem Anteil im jeweiligen Querschnitt. Männliche Befragte sind entsprechend im selben Maße unterrepräsentiert.

Tabelle 5.82: Zusammensetzung nach Geschlecht $P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$

		männlich		weiblich		gesamt
		<i>n</i>	%	<i>n</i>	%	(100 %)
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	325	34,2	625	65,8	950
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	321	33,8	629	66,2	950
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	324	34,1	626	65,9	950
t_5	Q_{t_5}	2265	49,8	2282	50,2	4547
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	323	34,0	627	66,0	950
t_6	Q_{t_6}	1567	47,0	1768	53,0	3335
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	323	34,0	627	66,0	950
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	325	34,2	625	65,8	950
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	324	34,1	626	65,9	950
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	325	34,2	625	65,8	950
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	322	33,9	628	66,1	950
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	326	34,3	624	65,7	950
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	325	34,2	625	65,8	950

Beim Durchschnittsalter sind nur geringfügige Abweichungen zwischen den Querschnittsdatensätzen und dem hier betrachteten Elf-Wellen-Panel erkennbar (Tabelle 5.83). Zu allen Zeitpunkten liegt das durchschnittliche Alter im Querschnitt etwas höher als im Längsschnittdatensatz, die Abweichung ist mit 0,1 bis 0,3 Jahren jedoch als gering einzustufen. Lediglich zum Zeitpunkt t_5 liegt mit einer Abweichung von 0,5 Jahren eine etwas höhere Differenz vor. Zudem zeigt sich, dass in den Querschnittsdatensätzen durchgehend eine etwas höhere Standardabweichung der Altersverteilung vorliegt, was für eine größere Heterogenität der Altersverteilung in den Querschnittsdaten spricht.

Tabelle 5.83: Altersdurchschnitt $P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$

		Alter	Std.-abw.	gesamt
t_2	Q_{t_2}	14,0	0,8	3270
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	13,8	0,7	938
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	14,9	0,7	946
t_4	Q_{t_4}	16,1	0,8	3313
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	15,9	0,7	943
t_5	Q_{t_5}	17,4	1,3	4468
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	16,9	0,8	945
t_6	Q_{t_6}	18,1	1,1	3320
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	17,8	0,7	949
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	18,8	0,7	950
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	19,8	0,7	950
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	21,9	0,7	949
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	23,8	0,8	950
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	25,9	0,7	949
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$	27,9	0,7	946

Auch bei der Verteilung nach Schulform (Tabelle 5.84) treten Unterschiede zwischen dem Paneldatensatz $P_{t_{2,3,4,5,6,7,8,9,10,11,12}}$ und den jeweiligen Querschnittsdatensätzen auf. Beim Vergleich mit der Schulstatistik der Erhebungszeitpunkte t_2 und t_3 wird deutlich, dass deutlich mehr Befragte an Gymnasien im Panel vertreten sind, umgekehrt aber wesentlich weniger Schüler/innen an Hauptschulen. Eine derartige Verzerrung tritt auch

bereits beim Vergleich der Querschnittsdatensätze mit der Schulstatistik auf, jedoch in geringerem Ausmaß. Mit zunehmender Erhebungsdauer reduzieren sich die Unterschiede zwischen den beiden Datensatztypen, was vor allem darauf zurückzuführen ist, dass ab t_8 bereits die Hälfte aller Personen keine Schule mehr besucht, sodass sich die Untersuchung der Zusammensetzung nach Schulform für die weiteren Zeitpunkte erübrigt.

Tabelle 5.84: Zusammensetzung $P_{2,3,4,5,6,7,8,9,10,11,12}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)								gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_2	S_{t_2}	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
	Q_{t_2}	24,6	23,6	31,4	20,1	—	0,4	—	—	3392
	$P_{2,3,4, \dots, 10,11,12}$	12,8	24,8	30,2	32,1	—	—	—	—	950
t_3	S_{t_3}	24,9	21,5	32,3	20,0	—	1,3	—	—	3925
	Q_{t_3}	22,6	22,9	31,9	22,2	—	0,5	—	—	3339
	$P_{2,3,4, \dots, 10,11,12}$	12,9	25,5	30,3	31,3	—	—	—	—	950
t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{2,3,4, \dots, 10,11,12}$	12,4	25,4	30,3	31,4	0,1	—	0,3	—	949
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{2,3,4, \dots, 10,11,12}$	0,2	0,6	22,9	37,6	38,7	—	—	—	881
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{2,3,4, \dots, 10,11,12}$	—	—	18,9	33,0	28,6	—	11,2	8,3	932
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{2,3,4, \dots, 10,11,12}$	—	—	16,7	31,6	25,4	—	12,5	13,7	933
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{2,3,4, \dots, 10,11,12}$	—	—	3,6	4,5	23,1	—	18,8	50,0	924

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.11.3 Elf-Wellen-Panel 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$)

Das Elf-Wellen-Panel der Erhebungsjahre 2004 bis 2019 umfasst 1 022 Fälle. Tabelle 5.85 zeigt die Zusammensetzung dieses Panels nach Geschlecht im Vergleich zu den Referenzdaten der Querschnitte und der Schulstatistik. Dabei wird deutlich, dass der Anteil weiblicher Befragter im Paneldatensatz zu jedem Zeitpunkt deutlich über dem

Anteil im jeweiligen Querschnitt liegt. Je nach Welle liegt diese Differenz zwischen elf und 16 Prozentpunkten. Weibliche Befragte sind im Panel also deutlich überrepräsentiert und männliche Befragte entsprechend im selben Maße unterrepräsentiert.

Tabelle 5.85: Zusammensetzung nach Geschlecht $P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$

		männlich		weiblich		gesamt (100 %)
		<i>n</i>	%	<i>n</i>	%	
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	345	33,8	677	66,2	1022
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	350	34,2	672	65,7	1022
t_5	Q_{t_5}	2265	49,8	2282	50,2	4547
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	350	34,2	672	65,7	1022
t_6	Q_{t_6}	1567	47,0	1768	53,0	3335
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	350	34,2	672	65,7	1022
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	352	34,4	670	65,6	1022
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	351	34,3	671	65,7	1022
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	352	34,4	670	65,6	1022
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	349	34,1	673	65,9	1022
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	353	34,5	669	65,5	1022
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	353	34,5	669	65,5	1022
t_{13}	$Q_{t_{13}}$	1238	45,9	1460	54,1	2698
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	352	34,4	670	65,6	1022

Das Durchschnittsalter in jeder Erhebungswelle ist für das Panel $P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$ und die jeweiligen Querschnittserhebungen in Tabelle 5.86 abgebildet. Dabei zeigen sich nur marginale Abweichungen zwischen Querschnitts- und Paneldaten. In den Querschnitten liegt das durchschnittliche Alter stets um 0,1 bis 0,3 Jahre höher, einzig zum Zeitpunkt t_5 sind es 0,5 Jahre Abweichung. Zudem weist der Paneldatensatz eine durchweg geringere Standardabweichung der Altersverteilung auf, was für eine größere Altershomogenität im Paneldatensatz spricht.

Tabelle 5.86: Altersdurchschnitt $P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$

		Alter	Std.-abw.	gesamt
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	14,9	0,7	1015
t_4	Q_{t_4}	16,1	0,8	3313
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	15,9	0,7	1014
t_5	Q_{t_5}	17,4	1,3	4468
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	16,9	0,8	1017
t_6	Q_{t_6}	18,1	1,1	3320
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	17,9	0,7	1021
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	18,9	0,7	1022
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	19,9	0,7	1022
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	21,9	0,7	1021
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	23,8	0,8	1022
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	25,9	0,7	1021
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	27,9	0,8	1017
t_{13}	$Q_{t_{13}}$	30,1	0,9	2677
	$P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$	29,9	0,7	1021

In Tabelle 5.87 ist die Zusammensetzung des Panels $P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$ nach Schulform im Vergleich zu den Querschnitterhebungen der jeweiligen Zeitpunkte sowie der Schulstatistik S_{t_3} dargestellt. Dabei zeigen sich Abweichungen zwischen den Datensatztypen. Ein Vergleich mit der Schulstatistik zum Zeitpunkt t_3 zeigt, dass sich im Paneldatensatz zu diesem Zeitpunkt deutlich mehr Gymnasiast/innen befinden als in der Schulstatistik ausgewiesen ist. Umgekehrt enthält der Längsschnittdatensatz bedeutend weniger Befragte an Hauptschulen. Eine derartige Verzerrung nach Schulform lässt sich auch für den Querschnitt Q_{t_3} identifizieren, wobei die Abweichungen zur Schulstatistik hierbei geringer ausfallen. Mit zunehmender Erhebungsdauer reduzieren sich die Unterschiede zwischen den beiden Datensatztypen, was vor allem darauf zurückzuführen ist, dass ab t_8 bereits die Hälfte aller Personen keine Schule mehr besucht, sodass sich die Untersuchung der Zusammensetzung nach Schulform für die weiteren Zeitpunkte erübrigt.

Tabelle 5.87: Zusammensetzung $P_{t_{3,4,5,6,7,8,9,10,11,12,13}}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)								gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_3	S_{t_3}	24,9	21,5	32,3	20,0	—	1,3	—	—	3925
	Q_{t_3}	22,6	22,9	31,9	22,2	—	0,5	—	—	3339
	$P_{t_{3,4,5, \dots, 11,12,13}}$	12,9	25,0	29,1	33,1	—	—	—	—	1022
t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_{3,4,5, \dots, 11,12,13}}$	12,2	25,0	28,9	33,3	0,2	—	0,4	—	1019
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{3,4,5, \dots, 11,12,13}}$	0,2	0,4	20,6	39,3	39,5	—	—	—	942
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{3,4,5, \dots, 11,12,13}}$	—	—	16,9	34,6	28,3	—	10,9	9,4	1001
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{3,4,5, \dots, 11,12,13}}$	—	—	14,8	32,5	25,6	—	12,8	14,3	1007
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{3,4,5, \dots, 11,12,13}}$	—	—	3,4	4,0	22,9	—	19,7	50,0	992

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.12 Die Zwölf-Wellen-Panel

Alle (nach den Plausibilitätskontrollen) gültigen Fälle, die mindestens elf weiteren Fällen im Datensatz zugeordnet werden konnten, bilden die Zwölf-Wellen-Panel. In diesem Bericht werden nur jene Datensätze beschrieben, die Teilnahmen in zwölf direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke).

5.12.1 Zwölf-Wellen-Panel 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015 und 2017 ($P_{t_{1,2,3,4,5,6,7,8,9,10,11,12}}$)

Das Zwölf-Wellen-Panel der Erhebungsjahre 2002 bis 2017 umfasst 790 Fälle. Mit Blick auf die Verteilung nach Geschlecht (Tabelle 5.88) ist eine Verzerrung dieses Paneldatensatzes im Vergleich zu den Referenzdaten der Schulstatistik und der Querschnittsdatsätze zu erkennen. Die Schulstatistiken der Zeitpunkte t_1 bis t_3 sowie die Querschnitte von t_1 bis t_5 zeigen, dass ein weitgehend ausgeglichenes Verhältnis beim Anteil von weiblichen und männlichen Befragten vorliegt. Ab t_6 ist für die Querschnitte eine leichte Verschiebung dieser Relation erkennbar – der Frauenanteil liegt ab diesem Zeitpunkt stets bei rund 53 bis 54 Prozent. Noch deutlicher spiegelt sich dieses Un-

gleichgewicht im Panel $P_{t_{1,2,3,4,5,6,7,8,9,10,11,12}}$ wider: Zu allen Zeitpunkten liegt der Anteil an Personen weiblichen Geschlechts bei rund 67 Prozent und damit je nach Zeitpunkt zwischen zwölf und rund 18 Prozentpunkten über dem Anteil im Querschnitt. Das hier betrachtete Zwölf-Wellen-Panel ist also deutlich in Richtung einer Überrepräsentanz weiblicher Befragter verzerrt: Es enthält zu gut zwei Dritteln weibliche Befragte und nur ein Drittel männliche Befragte, sodass die Geschlechterverteilung stark von der Relation der ursprünglichen Grundgesamtheit abweicht.

Tabelle 5.89 zeigt das Durchschnittsalter der Querschnittsdatensätze und des Panels $P_{t_{1,2,3,4,5,6,7,8,9,10,11,12}}$ für die betrachteten Zeitpunkte. In jeder Erhebungswelle entspricht das durchschnittliche Alter der Befragten den zu erwartenden Werten. Die Querschnitte und das Panel unterscheiden sich dabei nur geringfügig. Der Mittelwert des Alters liegt im Querschnitt durchweg etwas höher, diese Differenz ist mit 0,2 bis 0,3 Jahren jedoch als gering zu bewerten. Lediglich zum Zeitpunkt t_5 ist die Differenz mit 0,6 Jahren leicht erhöht. In diesem Jahr fand zusätzlich eine Erhebung an Berufsschulen statt, was diese etwas höhere Differenz erklärt. Zudem gestaltet sich die Altersverteilung im betrachteten Paneldatensatz etwas homogener als in den Querschnittsdatensätzen wie an der stets geringeren Standardabweichung zu erkennen ist. Auch hier fällt die Differenz zum Zeitpunkt t_5 am deutlichsten aus.

Bei der Zusammensetzung nach Schulform (Tabelle 5.90) wird auch für dieses Merkmal eine systematische Verzerrung des Paneldatensatzes erkennbar. Beim Vergleich mit der Schulstatistik zum Erhebungszeitpunkt t_1 wird ersichtlich, dass mit einer Differenz von gut zwölf Prozentpunkten Schülerinnen und Schüler an Hauptschulen im Panel $P_{t_{1,2,3,4,5,6,7,8,9,10,11,12}}$ unterrepräsentiert sind. Umgekehrt befinden sich in diesem Panel 4,5 Prozent mehr Realschülerinnen und -schüler sowie knapp elf Prozent mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik. Eine Verschiebung dieser Art lässt sich ebenso für die weiteren Zeitpunkte und auch beim Vergleich vom Panel $P_{t_{1,2,3,4,5,6,7,8,9,10,11,12}}$ mit dem Querschnitt feststellen. Dazu sei zu erwähnen, dass auch bereits im Querschnitt leichte Verzerrungen auftreten, wenn man vergleichend die Schulstatistik betrachtet.

Ab dem Zeitpunkt t_5 besuchen nur noch sehr wenige Personen eine Haupt- oder Realschule, sodass eine Betrachtung der Differenzen an dieser Stelle entfällt. Dafür wird von vielen Befragten angegeben, eine Berufsschule zu besuchen. Die Abweichung zwischen Quer- und Längsschnitt ist bei diesem Schultypus in t_5 besonders stark ausgeprägt. Dies ist darauf zurückzuführen, dass in t_5 eine temporäre Ausweitung der Querschnittsstichprobe durch Hinzunahme von Berufsschulen vorgenommen wurde. Ein Vergleich der Datensätze bezüglich ihrer Zusammensetzung nach Schulform ist für diesen Zeitpunkt daher weniger aussagekräftig, obgleich sich auch hier dieselben Tendenzen der Überrepräsentanz von Gymnasien und auch Gesamtschulen abzeichnen. Zu den Zeitpunkten t_6 und t_7 ist zu erkennen, dass der Anteil an Personen, die keine Schule mehr besuchen, im Panel geringer ausfällt als im Querschnitt. Hierbei

handelt es sich vermutlich unter anderem um jene Personen, die zuvor an Haupt- und Realschulen waren, welche zu den früheren Zeitpunkten ebenfalls im Paneldatensatz unterrepräsentiert waren.

Mit andauernder Erhebung und geringerer Relevanz der Schulform nivellieren sich auch zunehmend die gefundenen Differenzen. Zum Zeitpunkt t_8 lassen sich nur noch marginale Abweichungen zwischen dem Querschnitt Q_{t_8} und dem betrachteten Zwölf-Wellen-Panel feststellen. Bereits die Hälfte aller Befragten besucht in t_8 keine Schule mehr. Auf eine Betrachtung der Zusammensetzung nach Schulform wird daher ab dem Zeitpunkt t_9 verzichtet, da ab dieser Erhebungswelle so gut wie keine Befragten mehr eine Schule besuchen.

Tabelle 5.88: Zusammensetzung nach Geschlecht $P_{t_1,2,3,4,5,6,7,8,9,10,11,12}$

		männlich		weiblich		gesamt (100 %)
		<i>n</i>	%	<i>n</i>	%	
t_1	S_{t_1}	2003	51,2	1907	48,8	3910
	Q_{t_1}	1728	50,7	1679	49,3	3407
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12}$	260	32,9	530	67,1	790
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12}$	265	33,5	525	66,5	790
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12}$	261	33,0	529	67,0	790
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12}$	262	33,2	528	66,8	790
t_5	Q_{t_5}	2265	49,8	2282	50,2	4547
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12}$	264	33,4	526	66,6	790
t_6	Q_{t_6}	1567	47,0	1768	53,0	3335
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12}$	263	33,3	527	66,7	790
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12}$	265	33,5	525	66,5	790
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12}$	264	33,4	526	66,6	790
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12}$	264	33,4	526	66,6	790
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12}$	262	33,2	528	66,8	790
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12}$	265	33,5	525	66,5	790
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12}$	264	33,4	526	66,6	790

Tabelle 5.89: Altersdurchschnitt $P_{1,2,3,4,5,6,7,8,9,10,11,12}$

		Alter	Std.-abw.	gesamt
t_1	Q_{t_1}	13,0	0,8	3 346
	$P_{1,2,3,4,5,6,7,8,9,10,11,12}$	12,8	0,6	787
t_2	Q_{t_2}	14,0	0,8	3 270
	$P_{1,2,3,4,5,6,7,8,9,10,11,12}$	13,8	0,6	782
t_3	Q_{t_3}	15,1	0,8	3 270
	$P_{1,2,3,4,5,6,7,8,9,10,11,12}$	14,8	0,6	786
t_4	Q_{t_4}	16,1	0,8	3 313
	$P_{1,2,3,4,5,6,7,8,9,10,11,12}$	15,8	0,6	785
t_5	Q_{t_5}	17,4	1,3	4 468
	$P_{1,2,3,4,5,6,7,8,9,10,11,12}$	16,8	0,6	787
t_6	Q_{t_6}	18,1	1,1	3 320
	$P_{1,2,3,4,5,6,7,8,9,10,11,12}$	17,8	0,7	789
t_7	Q_{t_7}	19,1	0,8	3 081
	$P_{1,2,3,4,5,6,7,8,9,10,11,12}$	18,8	0,7	790
t_8	Q_{t_8}	20,1	0,8	3 073
	$P_{1,2,3,4,5,6,7,8,9,10,11,12}$	19,8	0,7	790
t_9	Q_{t_9}	22,1	0,9	3 021
	$P_{1,2,3,4,5,6,7,8,9,10,11,12}$	21,8	0,7	790
t_{10}	$Q_{t_{10}}$	24,1	1,2	2 839
	$P_{1,2,3,4,5,6,7,8,9,10,11,12}$	23,8	0,8	790
t_{11}	$Q_{t_{11}}$	26,0	1,0	2 745
	$P_{1,2,3,4,5,6,7,8,9,10,11,12}$	25,8	0,7	789
t_{12}	$Q_{t_{12}}$	28,1	1,2	2 753
	$P_{1,2,3,4,5,6,7,8,9,10,11,12}$	27,8	0,7	786

Tabelle 5.90: Zusammensetzung $P_{t_{1,2,3,4,5,6,7,8,9,10,11,12}}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)								gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_1	S_{t_1}	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
	Q_{t_1}	21,9	23,6	31,2	22,8	—	0,4	—	—	3411
	$P_{t_{1,2,3, \dots, 10,11,12}}$	11,6	26,5	29,9	32,0	—	—	—	—	790
t_2	S_{t_2}	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
	Q_{t_2}	24,6	23,6	31,4	20,1	—	0,4	—	—	3392
	$P_{t_{1,2,3, \dots, 10,11,12}}$	12,0	26,6	29,9	31,5	—	—	—	—	790
t_3	S_{t_3}	24,9	21,5	32,3	20,0	—	1,3	—	—	3925
	Q_{t_3}	22,6	22,9	31,9	22,2	—	0,5	—	—	3339
	$P_{t_{1,2,3, \dots, 10,11,12}}$	12,3	26,8	29,9	31,0	—	—	—	—	790
t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_{1,2,3, \dots, 10,11,12}}$	11,8	26,9	29,5	31,3	0,1	—	0,4	—	789
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{1,2,3, \dots, 10,11,12}}$	0,3	0,5	23,1	38,1	38,1	—	—	—	733
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{1,2,3, \dots, 10,11,12}}$	—	—	18,6	33,2	28,6	—	11,7	7,9	775
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{1,2,3, \dots, 10,11,12}}$	—	—	16,4	31,4	25,1	—	13,1	14,0	778
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{1,2,3, \dots, 10,11,12}}$	—	—	3,9	4,6	21,9	—	18,9	50,7	767

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.12.2 Zwölf-Wellen-Panel 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$)

Das Zwölf-Wellen-Panel der Erhebungsjahre 2003 bis 2019 umfasst 884 Fälle. Hinsichtlich der Verteilung nach Geschlecht in Tabelle 5.91 ist eine Verzerrung dieses Paneldatensatzes im Vergleich zu den Referenzdaten der Schulstatistik und der Querschnittsdatsätze erkennbar. Die Schulstatistiken der Zeitpunkte t_1 bis t_3 sowie die Querschnitte von t_1 bis t_5 zeigen ein weitgehend ausgeglichenes Verhältnis beim Anteil von weiblichen und männlichen Befragten. Ab t_6 ist für die Querschnitte eine leichte Verschiebung dieser Relation erkennbar – der Frauenanteil liegt ab diesem Zeitpunkt stets bei rund 53 bis 54 Prozent. Noch deutlicher spiegelt sich dieses Ungleichgewicht im Panel $P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$ wider: Zu allen Zeitpunkten liegt der Anteil an Personen weiblichen Geschlechts bei rund 67 Prozent und damit je nach Zeitpunkt zwischen zwölf und gut 17 Prozentpunkten über dem Anteil im Querschnitt. Das hier betrachtete Zwölf-Wellen-Panel ist also deutlich in Richtung einer Überrepräsentanz weiblicher Befragter verzerrt: Es enthält zu gut zwei Dritteln weibliche Befragte und nur ein Drittel männliche Befragte, sodass die Geschlechterverteilung stark von der Relation der ursprünglichen Grundgesamtheit abweicht.

Tabelle 5.92 zeigt das Durchschnittsalter der Querschnittsdatsätze und des Panels $P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$ für die betrachteten Zeitpunkte. In jeder Erhebungswelle entspricht das durchschnittliche Alter der Befragten den zu erwartenden Werten. Die Querschnitte und das Panel unterscheiden sich dabei nur geringfügig. Der Mittelwert des Alters liegt im Querschnitt durchweg etwas höher, diese Differenz ist mit 0,1 bis 0,3 Jahren jedoch als gering zu bewerten. Lediglich zum Zeitpunkt t_5 ist die Differenz mit 0,5 Jahren leicht erhöht. In diesem Jahr fand zusätzlich eine Erhebung an Berufsschulen statt, was diese etwas höhere Differenz erklärt. Zudem gestaltet sich die Altersverteilung im betrachteten Paneldatensatz etwas homogener als in den Querschnittsdatsätzen wie an der stets geringeren Standardabweichung zu erkennen ist.

Bei der Zusammensetzung nach Schulform (Tabelle 5.93) wird auch für dieses Merkmal eine systematische Verzerrung des Paneldatensatzes erkennbar. Beim Vergleich mit der Schulstatistik zum Erhebungszeitpunkt t_2 wird ersichtlich, dass mit einer Differenz von gut 14 Prozentpunkten Schülerinnen und Schüler an Hauptschulen im Panel $P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$ unterrepräsentiert sind. Umgekehrt befinden sich in diesem Panel gut zwei Prozent mehr Realschülerinnen und -schüler sowie gut 13 Prozent mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik. Eine Verschiebung dieser Art lässt sich ebenso für die weiteren Zeitpunkte und auch beim Vergleich vom Panel $P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$ mit dem Querschnitt feststellen. Dazu sei zu erwähnen, dass auch bereits im Querschnitt leichte Verzerrungen auftreten, wenn man vergleichend die Schulstatistik betrachtet.

Ab dem Zeitpunkt t_5 besuchen nur noch sehr wenige Personen eine Haupt- oder Realschule, sodass eine Betrachtung der Differenzen an dieser Stelle entfällt. Dafür wird von vielen Befragten angegeben, eine Berufsschule zu besuchen. Die Abweichung zwischen Quer- und Längsschnitt ist bei diesem Schultypus in t_5 besonders stark ausgeprägt. Dies ist darauf zurückzuführen, dass in t_5 eine temporäre Ausweitung der Querschnittsstichprobe durch Hinzunahme von Berufsschulen vorgenommen wurde. Ein Vergleich der Datensätze bezüglich ihrer Zusammensetzung nach Schulform ist für diesen Zeitpunkt daher weniger aussagekräftig, obgleich sich auch hier dieselben Tendenzen der Überrepräsentanz von Gymnasien und auch Gesamtschulen abzeichnen. Zu den Zeitpunkten t_6 und t_7 ist zu erkennen, dass der Anteil an Personen, die keine Schule mehr besuchen, im Panel geringer ausfällt als im Querschnitt. Hierbei handelt es sich vermutlich unter anderem um jene Personen, die zuvor an Haupt- und Realschulen waren, welche zu den früheren Zeitpunkten ebenfalls im Paneldatensatz unterrepräsentiert waren.

Mit andauernder Erhebung und geringerer Relevanz der Schulform nivellieren sich auch zunehmend die gefundenen Differenzen. Zum Zeitpunkt t_8 lassen sich nur noch marginale Abweichungen zwischen dem Querschnitt Q_{t_8} und dem hier betrachteten Zwölf-Wellen-Panel feststellen. Bereits die Hälfte aller Befragten besucht in t_8 keine Schule mehr. Auf eine Betrachtung der Zusammensetzung nach Schulform wird daher ab dem Zeitpunkt t_9 verzichtet, da ab dieser Erhebungswelle so gut wie keine Befragten mehr eine Schule besuchen.

Tabelle 5.91: Zusammensetzung nach Geschlecht $P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$

		männlich		weiblich		gesamt (100 %)
		<i>n</i>	%	<i>n</i>	%	
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	293	33,1	591	66,9	884
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	289	32,7	595	67,3	884
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	292	33,0	592	67,0	884
t_5	Q_{t_5}	2265	49,8	2282	50,2	4547
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	291	32,9	593	67,1	884
t_6	Q_{t_6}	1567	47,0	1768	53,0	3335
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	292	33,0	592	67,0	884
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	294	33,3	590	66,7	884
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	293	33,1	591	66,9	884
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	294	33,3	590	66,7	884
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	291	32,9	593	67,1	884
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	295	33,4	589	66,6	884
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	294	33,3	590	66,7	884
t_{13}	$Q_{t_{13}}$	1238	45,9	1460	54,1	2698
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	295	33,4	589	66,6	884

Tabelle 5.92: Altersdurchschnitt $P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$

		Alter	Std.-abw.	gesamt
t_2	Q_{t_2}	14,0	0,8	3 270
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	13,8	0,7	872
t_3	Q_{t_3}	15,1	0,8	3 270
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	14,9	0,7	880
t_4	Q_{t_4}	16,1	0,8	3 313
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	15,9	0,7	877
t_5	Q_{t_5}	17,4	1,3	4 468
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	16,9	0,8	879
t_6	Q_{t_6}	18,1	1,1	3 320
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	17,8	0,7	883
t_7	Q_{t_7}	19,1	0,8	3 081
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	18,8	0,7	884
t_8	Q_{t_8}	20,1	0,8	3 073
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	19,9	0,7	884
t_9	Q_{t_9}	22,1	0,9	3 021
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	21,9	0,7	883
t_{10}	$Q_{t_{10}}$	24,1	1,2	2 839
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	23,8	0,8	884
t_{11}	$Q_{t_{11}}$	26,0	1,0	2 745
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	25,9	0,7	883
t_{12}	$Q_{t_{12}}$	28,1	1,2	2 753
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	27,9	0,7	880
t_{13}	$Q_{t_{13}}$	30,1	0,9	2 677
	$P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$	29,9	0,7	883

Tabelle 5.93: Zusammensetzung $P_{t_{2,3,4,5,6,7,8,9,10,11,12,13}}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)							gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_2	S_{t_2}	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
	Q_{t_2}	24,6	23,6	31,4	20,1	—	0,4	—	—	3392
	$P_{t_{2,3,4, \dots, 11,12,13}}$	13,3	24,9	29,9	31,9	—	—	—	—	884
t_3	S_{t_3}	24,9	21,5	32,3	20,0	—	1,3	—	—	3925
	Q_{t_3}	22,6	22,9	31,9	22,2	—	0,5	—	—	3339
	$P_{t_{2,3,4, \dots, 11,12,13}}$	13,5	25,6	30,0	31,0	—	—	—	—	884
t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_{2,3,4, \dots, 11,12,13}}$	12,9	25,5	30,1	31,1	—	—	0,3	—	883
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{2,3,4, \dots, 11,12,13}}$	0,2	0,5	22,4	37,3	39,6	—	—	—	821
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{2,3,4, \dots, 11,12,13}}$	—	—	18,8	32,9	28,5	—	11,3	8,4	866
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{2,3,4, \dots, 11,12,13}}$	—	—	16,4	31,3	25,9	—	12,9	13,6	870
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{2,3,4, \dots, 11,12,13}}$	—	—	3,7	4,4	22,8	—	19,2	49,9	860

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.13 Das Dreizehn-Wellen-Panel 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013, 2015, 2017 und 2019 ($P_{t_{1,2,3,4,5,6,7,8,9,10,11,12,13}}$)

Alle (nach den Plausibilitätskontrollen) gültigen Fälle, die zwölf weiteren Fällen im Datensatz zugeordnet werden konnten – also entsprechend in allen Erhebungswellen der Duisburger Studie vorhanden sind – bilden das vollständige Dreizehn-Wellen-Panel. In diesem Panel sind also ausschließlich Personen enthalten, die an jedem der Befragungszeitpunkte von 2002 bis 2019 teilgenommen haben und anhand ihres persönlichen Codes zugeordnet werden konnten. Dieses Dreizehn-Wellen-Panel ist der umfassendste Datensatz des gesamten Erhebungszeitraums der jüngeren Duisburger Kohorte von 2002 bis 2019 und bildet entsprechend die Altersspanne von 13 bis 30 Jahren ab. Dieses vollständige Panel umfasst 736 Fälle.

Mit Blick auf die Verteilung nach Geschlecht, die in Tabelle 5.94 dargestellt ist, lässt sich eine Verzerrung dieses Paneldatensatzes im Vergleich zu den Referenzdaten der Schulstatistik und der Querschnittsdatsätze feststellen. Die Schulstatistiken der Zeitpunkte t_1 bis t_3 sowie die Querschnitte von t_1 bis t_5 zeigen, dass ein weitgehend ausgeglichenes Verhältnis beim Anteil von weiblichen und männlichen Befragten vorliegt. Ab t_6 ist für die Querschnitte eine leichte Verschiebung dieser Relation erkennbar – der Frauenanteil liegt ab diesem Zeitpunkt stets bei rund 53 bis 54 Prozent. Noch deutlicher spiegelt sich dieses Ungleichgewicht im Dreizehn-Wellen-Panel wider: Zu allen Zeitpunkten liegt der Anteil an Personen weiblichen Geschlechts bei rund 67 bis 68 Prozent. Der Paneldatensatz ist also in Richtung einer Überrepräsentanz weiblicher Befragter verzerrt: Er enthält zu gut zwei Dritteln weibliche Befragte und nur ein Drittel männliche Befragte, sodass die Geschlechterverteilung stark von der Relation der ursprünglichen Grundgesamtheit abweicht.

Tabelle 5.95 zeigt das Durchschnittsalter der Querschnittsdatsätze und des Dreizehn-Wellen-Panels für alle Erhebungswellen. Zu jedem Zeitpunkt entspricht das durchschnittliche Alter der Befragten den zu erwartenden Werten. Die Querschnitte und das Panel unterscheiden sich dabei nur geringfügig. Der Mittelwert des Alters liegt im Querschnitt durchweg etwas höher, diese Differenz ist mit 0,2 bis 0,3 Jahren jedoch als gering zu bewerten. Lediglich zum Zeitpunkt t_5 ist die Differenz mit 0,6 Jahren leicht erhöht. In diesem Jahr fand zusätzlich eine Erhebung an Berufsschulen statt, was diese leicht erhöhte Differenz erklärt. Zudem gestaltet sich die Altersverteilung im Paneldatensatz etwas homogener als in den Querschnittsdatsätzen wie an der stets geringeren Standardabweichung zu erkennen ist. Auch hier fällt die Differenz zum Zeitpunkt t_5 am deutlichsten aus.

Tabelle 5.95

Bei der Zusammensetzung nach Schulform (Tabelle 5.96) wird auch für dieses Merkmal eine systematische Verzerrung des Paneldatensatzes erkennbar. Beim Vergleich mit der Schulstatistik zum Erhebungszeitpunkt t_1 wird ersichtlich, dass mit einer Differenz von knapp zwölf Prozentpunkten Schülerinnen und Schüler an Hauptschulen im Dreizehn-Wellen-Panel unterrepräsentiert sind. Umgekehrt befinden sich in diesem Panel 4,5 Prozent mehr Realschülerinnen und -schüler sowie gut zehn Prozent mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik. Eine Verschiebung dieser Art lässt sich ebenso für die weiteren Zeitpunkte und auch beim Vergleich vom Dreizehn-Wellen-Panel mit dem Querschnitt feststellen. Dazu sei zu erwähnen, dass auch bereits im Querschnitt Verzerrungen auftreten, wenn man vergleichend die Schulstatistik betrachtet.

Ab dem Zeitpunkt t_5 besuchen nur noch sehr wenige Personen eine Haupt- oder Realschule, sodass eine Betrachtung der Differenzen an dieser Stelle entfällt. Dafür wird von vielen Befragten angegeben, eine Berufsschule zu besuchen. Die Abweichung zwischen Quer- und Längsschnitt ist bei diesem Schultypus in t_5 besonders stark ausgeprägt. Dies ist darauf zurückzuführen, dass in t_5 eine temporäre Ausweitung der Querschnittsstichprobe durch Hinzunahme von Berufsschulen vorgenommen wurde. Ein Vergleich der Datensätze bezüglich ihrer Zusammensetzung nach Schulform ist für diesen Zeitpunkt daher weniger aussagekräftig, obgleich sich auch hier dieselben Tendenzen der Überrepräsentanz von Gymnasien und auch Gesamtschulen abzeichnen. Zu den Zeitpunkten t_6 und t_7 ist zu erkennen, dass der Anteil an Personen, die keine Schule mehr besuchen, im Panel geringer ausfällt als im Querschnitt. Hierbei handelt es sich vermutlich unter anderem um jene Personen, die zuvor an Haupt- und Realschulen waren, welche zu den früheren Zeitpunkten ebenfalls im Paneldatensatz unterrepräsentiert waren.

Mit andauernder Erhebung und geringerer Relevanz der Schulform nivellieren sich auch zunehmend die gefundenen Differenzen. Zum Zeitpunkt t_8 lassen sich nur noch marginale Abweichungen zwischen dem Querschnitt Q_{t_8} und dem Dreizehn-Wellen-Panel feststellen. Auf eine Betrachtung der Zeitpunkte t_9 und t_{13} wird auch an dieser Stelle wie gehabt verzichtet, da in diesen Erhebungsjahren – das durchschnittliche Alter der Befragten liegt hier bei 22 bis 30 Jahren – so gut wie keine Befragten mehr eine schulische Einrichtung besuchen.

Tabelle 5.94: Zusammensetzung nach Geschlecht $P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$

		männlich		weiblich		gesamt (100 %)
		<i>n</i>	%	<i>n</i>	%	
t_1	S_{t_1}	2003	51,2	1907	48,8	3910
	Q_{t_1}	1728	50,7	1679	49,3	3407
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	233	31,7	503	68,3	736
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	238	32,3	498	67,7	736
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	234	31,8	502	68,2	736
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	235	31,9	501	68,1	736
t_5	Q_{t_5}	2265	49,8	2282	50,2	4547
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	237	32,2	499	67,8	736
t_6	Q_{t_6}	1567	47,0	1768	53,0	3335
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	237	32,2	499	67,8	736
t_7	Q_{t_7}	1452	47,1	1634	52,9	3086
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	239	32,5	497	67,5	736
t_8	Q_{t_8}	1446	46,8	1644	53,2	3090
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	238	32,3	498	67,7	736
t_9	Q_{t_9}	1423	46,8	1620	53,2	3043
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	238	32,3	498	67,7	736
t_{10}	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	236	32,1	500	67,9	736
t_{11}	$Q_{t_{11}}$	1252	45,5	1502	54,5	2754
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	239	32,5	497	67,5	736
t_{12}	$Q_{t_{12}}$	1269	45,7	1509	54,3	2778
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	238	32,3	498	67,7	736
t_{13}	$Q_{t_{13}}$	1238	45,9	1460	54,1	2698
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	240	32,6	496	67,4	736

Tabelle 5.95: Altersdurchschnitt $P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$

		Alter	Std.-abw.	gesamt
t_1	Q_{t_1}	13,0	0,8	3346
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	12,8	0,6	733
t_2	Q_{t_2}	14,0	0,8	3270
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	13,8	0,6	728
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	14,9	0,7	732
t_4	Q_{t_4}	16,1	0,8	3313
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	15,8	0,6	731
t_5	Q_{t_5}	17,4	1,3	4468
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	16,8	0,6	733
t_6	Q_{t_6}	18,1	1,1	3320
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	17,8	0,7	735
t_7	Q_{t_7}	19,1	0,8	3081
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	18,8	0,7	736
t_8	Q_{t_8}	20,1	0,8	3073
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	19,8	0,7	736
t_9	Q_{t_9}	22,1	0,9	3021
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	21,8	0,7	736
t_{10}	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	23,8	0,8	736
t_{11}	$Q_{t_{11}}$	26,0	1,0	2745
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	25,8	0,7	735
t_{12}	$Q_{t_{12}}$	28,1	1,2	2753
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	27,8	0,7	732
t_{13}	$Q_{t_{13}}$	30,1	0,9	2677
	$P_{t_1,2,3,4,5,6,7,8,9,10,11,12,13}$	29,8	0,7	735

Tabelle 5.96: Zusammensetzung $P_{t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8}$ nach Schulform (Zeilenprozentuierung)

		Schulform ^a (%)								gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
t_1	S_{t_1}	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
	Q_{t_1}	21,9	23,6	31,2	22,8	—	0,4	—	—	3411
	$P_{t_1,2,3, \dots, 11,12,13}$	12,0	26,5	29,9	31,7	—	—	—	—	736
t_2	S_{t_2}	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
	Q_{t_2}	24,6	23,6	31,4	20,1	—	0,4	—	—	3392
	$P_{t_1,2,3, \dots, 11,12,13}$	12,4	26,6	29,9	31,1	—	—	—	—	736
t_3	S_{t_3}	24,9	21,5	32,3	20,0	—	1,3	—	—	3925
	Q_{t_3}	22,6	22,9	31,9	22,2	—	0,5	—	—	3339
	$P_{t_1,2,3, \dots, 11,12,13}$	12,6	26,9	29,9	30,6	—	—	—	—	736
t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_1,2,3, \dots, 11,12,13}$	12,1	26,9	29,7	30,9	—	—	0,4	—	735
t_5	Q_{t_5}	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_1,2,3, \dots, 11,12,13}$	0,3	0,4	22,7	37,5	39,1	—	—	—	683
t_6	Q_{t_6}	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_1,2,3, \dots, 11,12,13}$	—	—	18,7	32,7	28,7	—	11,9	7,9	721
t_7	Q_{t_7}	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_1,2,3, \dots, 11,12,13}$	—	—	16,4	30,7	25,6	—	13,5	13,8	726
t_8	Q_{t_8}	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_1,2,3, \dots, 11,12,13}$	—	—	4,1	4,5	21,8	—	19,3	50,3	714

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

6 Zusammenfassung und Erklärung

Bei zusammenfassender Betrachtung dieses Methodenberichts kann abschließend festgehalten werden, dass die Paneldatensätze Verzerrungen aufweisen.

Dies betrifft vor allem die Zusammensetzung der Datensätze nach dem Geschlecht. Die Schulstatistik, welche in den ersten drei Zeitpunkten des Erhebungszeitraumes als Referenz verwendet werden kann, weist aus, dass weibliche und männliche Personen in einem annähernd ausgewogenen Verhältnis zueinander stehen. Der Anteil an weiblichen Personen beträgt in dieser Schulstatistik meist rund 49 Prozent, der Anteil männlicher Befragter entsprechend 51 Prozent. In den ersten Erhebungsjahren spiegelt sich diese Relation auch in den Querschnittsdaten wider. Ab dem Erhebungsjahr 2007 beginnt sich dieses Verhältnis in den Querschnittsdaten jedoch zunehmend leicht in Richtung eines größeren Frauenanteils zu verschieben, welcher letztlich 54 Prozent erreicht. Dies bedeutet, dass auch bereits die Querschnittsdaten von – wenn auch geringen – Verzerrungen nach dem Merkmal Geschlecht betroffen sind.

Diese Überrepräsentanz weiblicher Befragter wird in den Paneldatensätzen nochmals besonders deutlich: Dort liegt der Frauenanteil stets wesentlich über den Referenzwerten aus Querschnitt und Schulstatistik. Diese Verzerrung ist außerdem umso größer, je umfangreicher das Panel wird. So erreicht der Anteil an weiblichen Personen im umfassendsten Paneldatensatz – dem Dreizehn-Wellen-Panel – einen Wert von rund 68 Prozent. Somit besteht dieser vollständige, alle Wellen umfassende Paneldatensatz zu gut zwei Dritteln aus weiblichen Personen und entsprechend nur zu einem Drittel aus männlichen Befragten.

Auch bei Betrachtung der besuchten Schulform sind systematische Verschiebungen erkennbar. Vergleiche mit der Schulstatistik zeigen, dass Schülerinnen und Schüler an Hauptschulen in den Paneldatensätzen durchweg unterrepräsentiert sind, während Gymnasiastinnen und Gymnasiasten stets überrepräsentiert sind. Für die Jahre, in denen keine Daten der Schulstatistik mehr vorliegen, wurden die Paneldatensätze den jeweiligen Querschnitten dieser Erhebungsjahre gegenübergestellt und auch bei diesem Vergleich wird die genannte Verschiebung sichtbar. Hierbei ist anzumerken, dass allerdings auch bereits die Querschnitte beim Vergleich mit der Schulstatistik derartige Verzerrungen aufweisen, allerdings in geringerem Maße. Demnach ist davon auszugehen, dass die Verzerrungen des Panels, die am Vergleich zwischen Querschnitt und Panel abgeleitet werden, eigentlich nochmals stärker sind, da auch bereits der Querschnitt verzerrte Werte enthält. Für die anderen Schulformen wie zum Beispiel Real- und Gesamtschule finden sich auch Diskrepanzen zwischen Querschnitts- und Paneldatensätzen, allerdings sind

die Differenzen hierbei meist geringer und weisen eine weniger prägnante Systematik auf als die oben genannten Schulformen.

In den späteren Erhebungszeitpunkten wechselt die Relevanz der verschiedenen Schulformen. So sind ab dem Jahr 2006 (Zeitpunkt t_5 und durchschnittliches Alter 17 Jahre) Haupt-, Real- und Sonderschulen kaum mehr in der Stichprobe vertreten, dafür befindet sich ein großer Anteil an Personen nun auf einem Berufskolleg. Insgesamt gilt für diese Schulform die Tendenz, dass auch hier eine leichte Unterrepräsentanz im Panel vorliegt. Allerdings ist zu beachten, dass Abweichungen insbesondere zum Zeitpunkt t_5 mit Vorsicht zu interpretieren sind. Wie bereits in Kapitel 2 beschrieben, wurde zu diesem Zeitpunkt die Stichprobe temporär ausgeweitet, was auch an der hohen Fallzahl dieses Erhebungsjahres sichtbar wird. Auch die immer wiederkehrende Feststellung, dass zu t_5 das Durchschnittsalter im Querschnitt leicht erhöht und zudem heterogener verteilt ist, kann hierauf zurückgeführt werden.

Diese Erweiterung ist für den großen Anteil an Berufsschülerinnen und -schülern im Erhebungsjahr 2006 verantwortlich, woraus auch eine große Differenz zwischen Anteil im Querschnitt und Anteil im Panel resultiert. Streng genommen sind für den Zeitpunkt t_5 Vergleiche zwischen Querschnittsdaten und den Paneldatensätzen nicht zulässig und auch wenig aussagekräftig. Ob das Panel repräsentativ ist, lässt sich also in dieser Welle nicht zuverlässig an einem Vergleich mit dem Querschnitt erörtern.

Unter Betrachtung der weiteren Erhebungszeitpunkte ab t_6 ist erkennbar, dass sich die Verzerrungen bezüglich der Zusammensetzung nach Schulform zunehmend mindern. So weichen die Kennziffern aus Quer- und Längsschnitt zum letzten betrachteten Zeitpunkt t_8 nur noch marginal voneinander ab. Dies ist damit zu begründen, dass ab t_8 – die Befragten sind jetzt durchschnittlich 20 Jahre alt – kaum noch Personen eine schulische Einrichtung besuchen, sodass das Merkmal, welches bisher die systematische Verzerrung hervorgerufen hat, schlicht nicht mehr vorhanden ist.

Eine mögliche und naheliegende Erklärung für die Verzerrungen bezüglich des Geschlechts, wie sie sich im Querschnitt zeigen, ist die Umstellung des Befragungsmodus. Es fällt auf, dass die Verschiebungen im Querschnitt ab dem Zeitpunkt t_6 beginnen – in diesem Jahr wurde der Erhebungsmodus weitgehend auf eine postalische Paper-and-Pencil-Befragung umgestellt. In den vorherigen Jahren fand die Befragung im Klassenverband statt, wodurch ein größerer Einfluss auf die Teilnahmebereitschaft vorhanden war. Mit der postalischen Erhebung geht eine weniger kontrollierbare Interviewsituation einher, sodass die Stichprobe durch unterschiedliche Teilnahmebereitschaften selektiert wird. Frauen scheinen demnach eher bereit zu sein, an der Befragung teilzunehmen, was sich in der Verschiebung nach Geschlecht widerspiegelt. Durch wiederholte Befragungen verstärkt sich dieser Effekt weiter, sodass diese Verzerrung schließlich im Längsschnitt besonders deutlich zum Ausdruck kommt. Dabei gilt, dass solche Effekte zusätzlich stärker sind, je umfassender das Panel wird:

Beispielsweise liegt für das von Pöge (2007) beschriebene totale Vier-Wellen-Panel $P_{t_{1,2,3,4}}$ ein Frauenanteil von rund 56 Prozent vor. Bei Pollich (2010) beträgt dieser Anteil für das Sechs-Wellen-Panel $P_{t_{1,2,3,4,5,6}}$ 61 Prozent und im Zehn-Wellen-Panel $P_{t_{1,2,3,4,5,6,7,8,9,10}}$ (Daniel & Erdmann, 2017) bereits 66 Prozent. Im vorliegenden Methodenbericht zum Dreizehn-Wellen-Panel $P_{t_{1,2,3,4,5,6,7,8,9,10,11,12,13}}$ erhöht sich dieser Anteil schließlich weiter und erreicht rund 68 Prozent.

Solche Selektionseffekte durch unterschiedliche Teilnahmebereitschaft oder die Möglichkeit dazu können auch für die Verzerrungen nach Schulform verantwortlich sein, die auftreten, obwohl im Klassenverband befragt wird. Mögliche Gründe für Nicht-Teilnahme sind dabei Krankheit, Schwänzen oder Verweigerung (vgl. Pöge, 2007). Es kann angenommen werden, dass diese Ausfallgründe nach Schulform, teilweise aber auch nach Geschlecht variieren.

Eine weitere große Bedeutung für das Panel und eventuelle Verzerrungen hat das Zuordnungsverfahren anhand des persönlichen Codes. Wie bei Pöge (2007) ausgeführt, haben manche Personengruppen Probleme mit der korrekten Reproduktion ihres persönlichen Codes. Hierbei bestehen Zusammenhänge mit Geschlecht und Bildungsniveau. Es zeigt sich, dass weibliche Personen weniger Schwierigkeiten haben, ihr Codeblatt fehlerfrei auszufüllen als männliche Personen. Auch die Schulbildung hat einen zu erwartenden Effekt: Je höher die Schulbildung, desto weniger Fehler werden beim Ausfüllen des Codeblattes gemacht. Trotz der Einfachheit der Codefragen bestehen also Hürden bezüglich Teilnahmefähigkeit bzw. generell Teilnahmewille. Dies führt dazu, dass die Zuordnung im Panel für bestimmte Personengruppen besser und zuverlässiger funktioniert als für andere. Derartige Unterschiede bezüglich Teilnahmebereitschaft und Code-Reproduktion begründen die beschriebenen Verzerrungen, die insbesondere im Paneldatensatz auftreten und sich dort zunehmend verstärken je umfangreicher der Paneldatensatz wird.

Literaturverzeichnis

- Bentrup, C. (2007). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2006* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 12). Münster und Bielefeld.
- Bentrup, C. (2009). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2007* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 15). Münster und Bielefeld.
- Bentrup, C. & Verneuer, L. (2014). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2011* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 20). Münster und Bielefeld.
- Boers, K. & Reinecke, J. (Hrsg.). (2019). *Delinquenz im Altersverlauf: Erkenntnisse der Langzeitstudie Kriminalität in der modernen Stadt*. Münster und New York: Waxmann.
- Boers, K., Reinecke, J., Bentrup, C., Daniel, A., Kanz, K.-M., Schulte, P., . . . Walburg, C. (2014). Vom Jugend- zum frühen Erwachsenenalter. Delinquenzverläufe und Erklärungszusammenhänge in der Verlaufsstudie "Kriminalität in der modernen Stadt". *Monatsschrift für Kriminologie und Strafrechtsreform*, 97 (3), 183–202.
- Brondies, M. (2004). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2003* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 6). Münster und Trier.
- Daniel, A. & Erdmann, A. (2017). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002 bis 2013 - Zehn-Wellen-Panel* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 23). Münster und Bielefeld.
- Kessler, G. (2019). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2017* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 25). Münster und Bielefeld.
- Kessler, G. (2021). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2019* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 26). Münster und Bielefeld.
- Motzke, K. & Brondies, M. (2004). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 5). Münster und Trier.
- Pöge, A. (2005a). *Methodendokumentation der kriminologischen Schülerbefragung in*

- Münster 2000-2003 - Vier-Wellen-Panel* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 9). Münster und Trier.
- Pöge, A. (2005b). Persönliche Codes bei Längsschnittstudien: Ein Erfahrungsbericht. *ZA-Information* (56), 50–69.
- Pöge, A. (2007). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002 bis 2005 - Vier-Wellen-Panel* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 13). Münster und Bielefeld.
- Pöge, A. (2008). Persönliche Codes "reloaded". *Methoden – Daten – Analysen. Zeitschrift für Empirische Sozialforschung*, 2 (1), 59–70.
- Pöge, A. (2011). Persönliche Codes bei Längsschnittuntersuchungen III. *Methoden – Daten – Analysen. Zeitschrift für Empirische Sozialforschung*, 5 (1), 109–134.
- Pollich, D. (2007). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002 und 2003 - Zwei-Wellen-Panel* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 14). Münster und Bielefeld.
- Pollich, D. (2010). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002 bis 2007 - Sechs-Wellen-Panel* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 16). Münster und Bielefeld.
- Verneuer, L. (2015). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2013* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 21). Münster und Bielefeld.
- Verneuer, L. (2017). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2015* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 22). Münster und Bielefeld.

A Anhang

A.1 Codeblätter

Liebe Teilnehmerin, lieber Teilnehmer,

da wir Ihren Fragebogen dem des letzten Jahres ohne Ihren Namen zuordnen wollen, ist es wichtig, dass Sie sich an Ihren persönlichen Code vom letzten Jahr erinnern. Denn nur so können Ihre Fragebögen einander zugeordnet werden, ohne dass jemand herausfinden kann, wer diese Fragebögen ausgefüllt hat. Wichtig ist also, dass Sie denselben Code noch wissen. Aus diesem Grund haben wir die nachfolgenden Fragen formuliert, die Ihnen helfen sollen, sich an Ihre persönliche Kombination zu erinnern.

*Bitte kreuzen Sie bei jeder der sieben Fragen immer nur ein Feld an!
Wenn Sie eine der Fragen überhaupt nicht beantworten können, kreuzen Sie bitte kein Feld an!*

Hier nun die sieben Fragen zur Erstellung Ihres persönlichen Codes:

1	<p>Bitte kreuzen Sie den ersten Buchstaben des Vornamens Ihres Vaters (oder einer Person, die für Sie einem Vater am nächsten kommt) an. (z. B. <input type="checkbox"/>Anton, <input type="checkbox"/>Bernd, <input type="checkbox"/>Hans-Peter usw.).</p> <table border="1"> <tr> <td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td><td>h</td><td>i</td><td>j</td><td>k</td><td>l</td><td>m</td><td>n</td><td>o</td> </tr> <tr> <td>p</td><td>q</td><td>r</td><td>s</td><td>t</td><td>u</td><td>v</td><td>w</td><td>x</td><td>y</td><td>z</td><td>ä</td><td>ö</td><td>ü</td><td>ß</td> </tr> </table>	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß	
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o																		
p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß																		
2	<p>Bitte kreuzen Sie den ersten Buchstaben des Vornamens Ihrer Mutter (oder einer Person, die für Sie einer Mutter am nächsten kommt) an. (z. B. <input type="checkbox"/>Anna, <input type="checkbox"/>Beate, <input type="checkbox"/>Jutta, <input type="checkbox"/>Maria, usw.).</p> <table border="1"> <tr> <td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td><td>h</td><td>i</td><td>j</td><td>k</td><td>l</td><td>m</td><td>n</td><td>o</td> </tr> <tr> <td>p</td><td>q</td><td>r</td><td>s</td><td>t</td><td>u</td><td>v</td><td>w</td><td>x</td><td>y</td><td>z</td><td>ä</td><td>ö</td><td>ü</td><td>ß</td> </tr> </table>	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß	
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o																		
p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß																		
3	<p>Bitte kreuzen Sie den ersten Buchstaben Ihres Vornamens an (z. B. <input type="checkbox"/>Michael, <input type="checkbox"/>Thomas, <input type="checkbox"/>Ute usw.).</p> <table border="1"> <tr> <td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td><td>h</td><td>i</td><td>j</td><td>k</td><td>l</td><td>m</td><td>n</td><td>o</td> </tr> <tr> <td>p</td><td>q</td><td>r</td><td>s</td><td>t</td><td>u</td><td>v</td><td>w</td><td>x</td><td>y</td><td>z</td><td>ä</td><td>ö</td><td>ü</td><td>ß</td> </tr> </table>	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß	
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o																		
p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß																		
4	<p>Bitte kreuzen Sie den Tag Ihres Geburtsdatums an (z. B. Geburtstag am 7. Januar = <input type="checkbox"/>7, am 12. Mai = <input type="checkbox"/>12, am 31. Oktober = <input type="checkbox"/>31).</p> <table border="1"> <tr> <td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td> </tr> <tr> <td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td><td>31</td> </tr> </table>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15																		
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31																	
5	<p>Bitte kreuzen Sie den letzten Buchstaben Ihrer natürlichen Haarfarbe an (z. B. braun, Glatze, schwarz, usw.).</p> <table border="1"> <tr> <td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td><td>h</td><td>i</td><td>j</td><td>k</td><td>l</td><td>m</td><td>n</td><td>o</td> </tr> <tr> <td>p</td><td>q</td><td>r</td><td>s</td><td>t</td><td>u</td><td>v</td><td>w</td><td>x</td><td>y</td><td>z</td><td>ä</td><td>ö</td><td>ü</td><td>ß</td> </tr> </table>	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß	
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o																		
p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß																		
6	<p>Bitte kreuzen Sie den letzten Buchstaben Ihrer Augenfarbe an (z. B. braun, grün, grau, usw.).</p> <table border="1"> <tr> <td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td><td>h</td><td>i</td><td>j</td><td>k</td><td>l</td><td>m</td><td>n</td><td>o</td> </tr> <tr> <td>p</td><td>q</td><td>r</td><td>s</td><td>t</td><td>u</td><td>v</td><td>w</td><td>x</td><td>y</td><td>z</td><td>ä</td><td>ö</td><td>ü</td><td>ß</td> </tr> </table>	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß	
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o																		
p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß																		
7	<p>Bitte kreuzen Sie den letzten Buchstaben Ihres Nachnamens an. (Sollten Sie Ihren Namen gewechselt haben, nehmen Sie Ihren Geburtsnamen!)</p> <table border="1"> <tr> <td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td><td>h</td><td>i</td><td>j</td><td>k</td><td>l</td><td>m</td><td>n</td><td>o</td> </tr> <tr> <td>p</td><td>q</td><td>r</td><td>s</td><td>t</td><td>u</td><td>v</td><td>w</td><td>x</td><td>y</td><td>z</td><td>ä</td><td>ö</td><td>ü</td><td>ß</td> </tr> </table>	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß	
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o																		
p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß																		

Seite 2

Abbildung A.1: Codeblatt Duisburg 2015

Liebe Teilnehmerin, lieber Teilnehmer,

da wir Ihren Fragebogen dem des letzten Jahres ohne Ihren Namen zuordnen wollen, ist es wichtig, dass Sie sich an Ihren persönlichen Code vom letzten Jahr erinnern. Denn nur so können Ihre Fragebögen einander zugeordnet werden, ohne dass jemand herausfinden kann, wer diese Fragebögen ausgefüllt hat. Wichtig ist also, dass Sie denselben Code noch wissen. Aus diesem Grund haben wir die nachfolgenden Fragen formuliert, die Ihnen helfen sollen, sich an Ihre persönliche Kombination zu erinnern.

*Bitte kreuzen Sie bei jeder der sieben Fragen immer nur ein Feld an!
Wenn Sie eine der Fragen überhaupt nicht beantworten können, kreuzen Sie bitte kein Feld an!*

Hier nun die sieben Fragen zur Erstellung Ihres persönlichen Codes:

1	Bitte kreuzen Sie den ersten Buchstaben des Vornamens Ihres Vaters (oder einer Person, die für Sie einem Vater am nächsten kommt) an. (z. B. <input type="checkbox"/> Anton, <input type="checkbox"/> Bernd, <input type="checkbox"/> Hans-Peter usw.). <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß
2	Bitte kreuzen Sie den ersten Buchstaben des Vornamens Ihrer Mutter (oder einer Person, die für Sie einer Mutter am nächsten kommt) an. (z. B. <input type="checkbox"/> Anna, <input type="checkbox"/> Cate, <input type="checkbox"/> Jutta, <input type="checkbox"/> Maria, usw.). <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß
3	Bitte kreuzen Sie den ersten Buchstaben Ihres Vornamens an (z. B. <input type="checkbox"/> Michael, <input type="checkbox"/> Thomas, <input type="checkbox"/> Ute usw.). <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß
4	Bitte kreuzen Sie den Tag Ihres Geburtsdatums an (z. B. Geburtstag am 7. Januar = <input type="checkbox"/> 7, am 12. Mai = <input type="checkbox"/> 12, am 31. Oktober = <input type="checkbox"/> 31). <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/> 8 <input type="checkbox"/> 9 <input type="checkbox"/> 10 <input type="checkbox"/> 11 <input type="checkbox"/> 12 <input type="checkbox"/> 13 <input type="checkbox"/> 14 <input type="checkbox"/> 15 <input type="checkbox"/> 16 <input type="checkbox"/> 17 <input type="checkbox"/> 18 <input type="checkbox"/> 19 <input type="checkbox"/> 20 <input type="checkbox"/> 21 <input type="checkbox"/> 22 <input type="checkbox"/> 23 <input type="checkbox"/> 24 <input type="checkbox"/> 25 <input type="checkbox"/> 26 <input type="checkbox"/> 27 <input type="checkbox"/> 28 <input type="checkbox"/> 29 <input type="checkbox"/> 30 <input type="checkbox"/> 31
5	Bitte kreuzen Sie den letzten Buchstaben Ihrer natürlichen Haarfarbe an (z. B. braun, <input type="checkbox"/> Glatz, <input type="checkbox"/> schwarz, usw.). <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß
6	Bitte kreuzen Sie den letzten Buchstaben Ihrer Augenfarbe an (z. B. braun, grün, <input type="checkbox"/> grau, usw.). <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß
7	Bitte kreuzen Sie den letzten Buchstaben Ihres Nachnamens an. (Sollten Sie Ihren Namen gewechselt haben, nehmen Sie Ihren Geburtsnamen!) <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß

Seite 2

Abbildung A.2: Codeblatt Duisburg 2017

Liebe Teilnehmerin, lieber Teilnehmer,

da wir Ihren Fragebogen dem der letzten Erhebung ohne Ihren Namen zuordnen wollen, ist es wichtig, dass Sie sich an Ihren persönlichen Code vom vorletzten Jahr erinnern. Denn nur so können Ihre Fragebögen einander zugeordnet werden, ohne dass jemand herausfinden kann, wer diese Fragebögen ausgefüllt hat. Wichtig ist also, dass Sie denselben Code noch wissen. Aus diesem Grund haben wir die nachfolgenden Fragen formuliert, die Ihnen helfen sollen, sich an Ihre persönliche Kombination zu erinnern.

*Bitte kreuzen Sie bei jeder der sieben Fragen immer nur ein Feld an!
Wenn Sie eine der Fragen überhaupt nicht beantworten können, kreuzen Sie bitte kein Feld an!*

Hier nun die sieben Fragen zur Erstellung Ihres persönlichen Codes:

1	<p>Bitte kreuzen Sie den ersten Buchstaben des Vornamens Ihres Vaters (oder einer Person, die für Sie einem Vater am nächsten kommt) an. (z. B. <input type="checkbox"/>Anton, <input type="checkbox"/>Bernd, <input type="checkbox"/>Hans-Peter usw.).</p> <table border="1"> <tr><td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td><td>h</td><td>i</td><td>j</td><td>k</td><td>l</td><td>m</td><td>n</td><td>o</td></tr> <tr><td>p</td><td>q</td><td>r</td><td>s</td><td>t</td><td>u</td><td>v</td><td>w</td><td>x</td><td>y</td><td>z</td><td>ä</td><td>ö</td><td>ü</td><td>ß</td></tr> </table>	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß	
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o																		
p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß																		
2	<p>Bitte kreuzen Sie den ersten Buchstaben des Vornamens Ihrer Mutter (oder einer Person, die für Sie einer Mutter am nächsten kommt) an. (z. B. <input type="checkbox"/>Anna, <input type="checkbox"/>Beate, <input type="checkbox"/>Jutta, <input type="checkbox"/>Maria, usw.).</p> <table border="1"> <tr><td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td><td>h</td><td>i</td><td>j</td><td>k</td><td>l</td><td>m</td><td>n</td><td>o</td></tr> <tr><td>p</td><td>q</td><td>r</td><td>s</td><td>t</td><td>u</td><td>v</td><td>w</td><td>x</td><td>y</td><td>z</td><td>ä</td><td>ö</td><td>ü</td><td>ß</td></tr> </table>	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß	
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o																		
p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß																		
3	<p>Bitte kreuzen Sie den ersten Buchstaben Ihres Vornamens an (z. B. <input type="checkbox"/>Michael, <input type="checkbox"/>Thomas, <input type="checkbox"/>Ute, usw.).</p> <table border="1"> <tr><td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td><td>h</td><td>i</td><td>j</td><td>k</td><td>l</td><td>m</td><td>n</td><td>o</td></tr> <tr><td>p</td><td>q</td><td>r</td><td>s</td><td>t</td><td>u</td><td>v</td><td>w</td><td>x</td><td>y</td><td>z</td><td>ä</td><td>ö</td><td>ü</td><td>ß</td></tr> </table>	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß	
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o																		
p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß																		
4	<p>Bitte kreuzen Sie den Tag Ihres Geburtsdatums an (z. B. Geburtstag am 7. Januar = <input type="checkbox"/>7, am 12. Mai = <input type="checkbox"/>12, am 31. Oktober = <input type="checkbox"/>31).</p> <table border="1"> <tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td></tr> <tr><td>16</td><td>17</td><td>18</td><td>19</td><td>20</td><td>21</td><td>22</td><td>23</td><td>24</td><td>25</td><td>26</td><td>27</td><td>28</td><td>29</td><td>30</td><td>31</td></tr> </table>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15																		
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31																	
5	<p>Bitte kreuzen Sie den letzten Buchstaben Ihrer natürlichen Haarfarbe an (z. B. braun <input type="checkbox"/>, Glatz <input type="checkbox"/>, schwarz <input type="checkbox"/>, usw.).</p> <table border="1"> <tr><td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td><td>h</td><td>i</td><td>j</td><td>k</td><td>l</td><td>m</td><td>n</td><td>o</td></tr> <tr><td>p</td><td>q</td><td>r</td><td>s</td><td>t</td><td>u</td><td>v</td><td>w</td><td>x</td><td>y</td><td>z</td><td>ä</td><td>ö</td><td>ü</td><td>ß</td></tr> </table>	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß	
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o																		
p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß																		
6	<p>Bitte kreuzen Sie den letzten Buchstaben Ihrer Augenfarbe an (z. B. braun <input type="checkbox"/>, grün <input type="checkbox"/>, grau <input type="checkbox"/>, usw.).</p> <table border="1"> <tr><td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td><td>h</td><td>i</td><td>j</td><td>k</td><td>l</td><td>m</td><td>n</td><td>o</td></tr> <tr><td>p</td><td>q</td><td>r</td><td>s</td><td>t</td><td>u</td><td>v</td><td>w</td><td>x</td><td>y</td><td>z</td><td>ä</td><td>ö</td><td>ü</td><td>ß</td></tr> </table>	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß	
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o																		
p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß																		
7	<p>Bitte kreuzen Sie den letzten Buchstaben Ihres Nachnamens an. (Sollten Sie Ihren Namen gewechselt haben, nehmen Sie Ihren Geburtsnamen!)</p> <table border="1"> <tr><td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td>f</td><td>g</td><td>h</td><td>i</td><td>j</td><td>k</td><td>l</td><td>m</td><td>n</td><td>o</td></tr> <tr><td>p</td><td>q</td><td>r</td><td>s</td><td>t</td><td>u</td><td>v</td><td>w</td><td>x</td><td>y</td><td>z</td><td>ä</td><td>ö</td><td>ü</td><td>ß</td></tr> </table>	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß	
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o																		
p	q	r	s	t	u	v	w	x	y	z	ä	ö	ü	ß																		

Seite 2

Abbildung A.3: Codeblatt Duisburg 2019

A.2 Häufigkeiten der Codevariablen

Tabelle A.1: Häufigkeiten von Co001 (erster Buchstabe Vorname Vater)

	2015		2017		2019	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
a	221	8,0	220	7,9	225	8,3
ä	—	—	—	—	—	—
b	81	2,9	75	2,7	78	2,9
c	57	2,1	56	2,0	63	2,3
d	119	4,3	130	4,6	123	4,5
e	76	2,7	84	3,0	82	3,0
f	114	4,1	121	4,3	107	3,9
g	115	4,1	117	4,2	120	4,4
h	323	11,7	317	11,3	301	11,1
i	44	1,6	51	1,8	48	1,8
j	171	6,2	164	5,9	153	5,6
k	132	4,8	135	4,8	141	5,2
l	39	1,4	39	1,4	31	1,1
m	280	10,1	273	9,8	276	10,2
n	72	2,6	67	2,4	69	2,5
o	34	1,2	37	1,3	30	1,1
ö	5	0,2	5	0,2	3	0,1
p	105	3,8	102	3,6	96	3,5
q	1	0,0	—	—	1	0,0
r	212	7,6	213	7,6	202	7,5
s	105	3,8	112	4,0	97	3,6
ß	—	—	—	—	—	—
t	83	3,0	92	3,3	81	3,0
u	105	3,8	110	3,9	99	3,7
ü	2	0,1	4	0,1	3	0,1
v	54	1,9	62	2,2	58	2,1
w	141	5,1	136	4,9	137	5,1
x	1	0,0	—	—	4	0,1
y	27	1,0	21	0,8	24	0,9
z	14	0,5	15	0,5	13	0,5
fehlend	39	1,4	39	1,4	44	1,6
gesamt	2772	100	2797	100,0	2709	100,0

Tabelle A.2: Häufigkeiten von Co002 (erster Buchstabe Vorname Mutter)

	2015		2017		2019	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
a	240	8,7	245	8,8	246	9,1
ä	1	0,0	1	0,0	1	0,0
b	248	8,9	233	8,3	230	8,5
c	156	5,6	168	6,0	152	5,6
d	82	3,0	78	2,8	77	2,8
e	166	6,0	164	5,9	162	6,0
f	62	2,2	62	2,2	63	2,3
g	150	5,4	161	5,8	153	5,6
h	198	7,1	195	7,0	187	6,9
i	112	4,0	113	4,0	109	4,0
j	38	1,4	40	1,4	33	1,2
k	100	3,6	101	3,6	101	3,7
l	40	1,4	45	1,6	43	1,6
m	315	11,4	312	11,2	297	11,0
n	114	4,1	108	3,9	101	3,7
o	12	0,4	13	0,5	11	0,4
ö	1	0,0	1	0,0	1	0,0
p	90	3,2	90	3,2	91	3,4
q	—	—	—	—	—	—
r	119	4,3	120	4,3	119	4,4
s	275	9,9	293	10,5	286	10,6
ß	—	—	—	—	—	—
t	32	1,2	29	1,0	30	1,1
u	97	3,5	98	3,5	91	3,4
ü	1	0,0	1	0,0	1	0,0
v	30	1,1	30	1,1	33	1,2
w	13	0,5	12	0,4	12	0,4
x	2	0,1	3	0,1	2	0,1
y	15	0,5	13	0,5	13	0,5
z	26	0,9	31	1,1	28	1,0
fehlend	37	1,3	37	1,3	36	1,3
gesamt	2772	100,0	2797	100,0	2709	100,0

Tabelle A.3: Häufigkeiten von Co003 (erster Buchstabe eigener Vorname)

	2015		2017		2019	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
a	237	8,5	247	8,8	231	8,5
ä	—	—	—	—	—	—
b	68	2,5	70	2,5	66	2,4
c	147	5,3	151	5,4	150	5,5
d	209	7,5	209	7,5	204	7,5
e	80	2,9	81	2,9	86	3,2
f	79	2,8	81	2,9	73	2,7
g	26	0,9	30	1,1	31	1,1
h	59	2,1	70	2,5	63	2,3
i	39	1,4	35	1,3	40	1,5
j	249	9,0	230	8,2	236	8,7
k	145	5,2	154	5,5	140	5,2
l	102	3,7	97	3,5	94	3,5
m	327	11,8	347	12,4	330	12,2
n	133	4,8	128	4,6	119	4,4
o	22	0,8	19	0,7	19	0,7
ö	10	0,4	11	0,4	9	0,3
p	90	3,2	86	3,1	85	3,1
q	—	—	—	—	—	—
r	87	3,1	82	2,9	76	2,8
s	370	13,3	374	13,4	362	13,4
ß	—	—	—	—	1	0,0
t	121	4,4	119	4,3	117	4,3
u	6	0,2	6	0,2	11	0,4
ü	—	—	1	0,0	2	0,1
v	59	2,1	63	2,3	60	2,2
w	6	0,2	6	0,2	7	0,3
x	—	—	1	0,0	1	0,0
y	44	1,6	45	1,6	42	1,6
z	18	0,6	14	0,5	11	0,4
fehlend	39	1,4	40	1,4	43	1,6
gesamt	2772	100,0	2797	100,0	2709	100,0

Tabelle A.4: Häufigkeiten von Co004 (Geburtstag)

	2015		2017		2019	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
1	88	3,2	95	3,4	89	3,3
2	66	2,4	69	2,5	72	2,7
3	84	3,0	82	2,9	82	3,0
4	93	3,4	110	3,9	94	3,5
5	97	3,5	98	3,5	96	3,5
6	82	3,0	78	2,8	80	3,0
7	90	3,2	87	3,1	84	3,1
8	105	3,8	98	3,5	95	3,5
9	107	3,9	107	3,8	101	3,7
10	102	3,7	96	3,4	100	3,7
11	90	3,2	95	3,4	89	3,3
12	80	2,9	83	3,0	90	3,3
13	73	2,6	82	2,9	77	2,8
14	93	3,4	87	3,1	87	3,2
15	98	3,5	101	3,6	92	3,4
16	101	3,6	104	3,7	98	3,6
17	92	3,3	92	3,3	88	3,2
18	108	3,9	110	3,9	97	3,6
19	98	3,5	94	3,4	91	3,4
20	87	3,1	87	3,1	83	3,1
21	83	3,0	77	2,8	86	3,2
22	88	3,2	88	3,1	80	3,0
23	79	2,8	83	3,0	83	3,1
24	90	3,2	91	3,3	88	3,2
25	90	3,2	97	3,5	89	3,3
26	101	3,6	92	3,3	96	3,5
27	74	2,7	77	2,8	70	2,6
28	77	2,8	71	2,5	72	2,7
29	80	2,9	79	2,8	79	2,9
30	68	2,5	80	2,9	80	3,0
31	54	1,9	53	1,9	48	1,8
fehlend	54	1,9	54	1,9	53	2,0
gesamt	2772	100,0	2797	100,0	2709	100,0

Tabelle A.5: Häufigkeiten von Co005 (letzter Buchstabe Haarfarbe)

	2015		2017		2019	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
a	—	—	—	—	—	—
ä	—	—	—	—	—	—
b	88	3,2	95	3,4	87	3,2
c	—	—	—	—	—	—
d	954	34,4	969	34,6	906	33,4
e	13	0,5	14	0,5	16	0,6
f	—	—	—	—	—	—
g	1	0,0	1	0,0	3	0,1
h	—	—	1	0,0	—	—
i	—	—	—	—	—	—
j	—	—	—	—	—	—
k	—	—	—	—	—	—
l	1	0,0	—	—	—	—
m	2	0,1	2	0,1	—	—
n	1379	49,7	1372	49,1	1372	50,6
o	—	—	—	—	—	—
ö	—	—	—	—	—	—
p	—	—	—	—	—	—
q	—	—	—	—	—	—
r	2	0,1	3	0,1	2	0,1
s	22	0,8	17	0,6	26	1,0
ß	—	—	—	—	—	—
t	28	1,0	27	1,0	24	0,9
u	3	0,1	2	0,1	2	0,1
ü	—	—	—	—	—	—
v	—	—	1	0,0	1	0,0
w	—	—	—	—	1	0,0
x	—	—	—	—	1	0,0
y	1	0,0	1	0,0	—	—
z	235	8,5	250	8,9	218	8,0
fehlend	43	1,6	42	1,5	50	1,8
gesamt	2772	100,0	2797	100,0	2709	100,0

Tabelle A.6: Häufigkeiten von Co006 (letzter Buchstabe Augenfarbe)

	2015		2017		2019	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
a	1	0,0	—	—	—	—
ä	—	—	—	—	—	—
b	35	1,3	37	1,3	34	1,3
c	—	—	—	—	—	—
d	1	0,0	1	0,0	5	0,2
e	—	—	—	—	1	0,0
f	—	—	—	—	—	—
g	13	0,5	17	0,6	8	0,3
h	1	0,0	1	0,0	1	0,0
i	1	0,0	—	—	—	—
j	—	—	—	—	—	—
k	—	—	—	—	—	—
l	—	—	—	—	—	—
m	2	0,1	1	0,0	1	0,0
n	1809	65,3	1832	65,5	1762	65,0
o	—	—	—	—	—	—
ö	1	0,0	—	—	1	0,0
p	—	—	—	—	—	—
q	—	—	—	—	1	0,0
r	—	—	1	0,0	1	0,0
s	2	0,1	—	—	2	0,1
ß	—	—	—	—	—	—
t	—	—	—	—	—	—
u	847	30,6	851	30,4	835	30,8
ü	—	—	—	—	—	—
v	—	—	1	0,0	1	0,0
w	—	—	—	—	—	—
x	—	—	—	—	1	0,0
y	—	—	—	—	—	—
z	14	0,5	12	0,4	11	0,4
fehlend	45	1,6	43	1,5	44	1,6
gesamt	2772	100,0	2797	100,0	2709	100,0

Tabelle A.7: Häufigkeiten von Co011 (letzter Buchstabe eigener Nachname)

	2015		2017		2019	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
a	128	4,6	132	4,7	135	5,0
ä	—	—	—	—	1	0,0
b	13	0,5	14	0,5	15	0,6
c	56	2,0	60	2,1	59	2,2
d	54	1,9	57	2,0	60	2,2
e	157	5,7	159	5,7	163	6,0
f	27	1,0	29	1,0	28	1,0
g	83	3,0	79	2,8	77	2,8
h	129	4,7	127	4,5	119	4,4
i	164	5,9	153	5,5	152	5,6
j	8	0,3	13	0,5	10	0,4
k	207	7,5	195	7,0	184	6,8
l	134	4,8	139	5,0	140	5,2
m	49	1,8	44	1,6	43	1,6
n	441	15,9	431	15,4	416	15,4
o	38	1,4	39	1,4	35	1,3
ö	—	—	—	—	—	—
p	21	0,8	23	0,8	21	0,8
q	1	0,0	1	0,0	1	0,0
r	396	14,3	411	14,7	400	14,8
s	208	7,5	217	7,8	201	7,4
ß	14	0,5	17	0,6	17	0,6
t	144	5,2	149	5,3	143	5,3
u	35	1,3	46	1,6	39	1,4
ü	3	0,1	5	0,2	4	0,1
v	4	0,1	5	0,2	4	0,1
w	18	0,6	27	1,0	19	0,7
x	6	0,2	5	0,2	7	0,3
y	40	1,4	38	1,4	40	1,5
z	150	5,4	143	5,1	139	5,1
fehlend	44	1,6	39	1,4	37	1,4
gesamt	2772	100,0	2797	100,0	2709	100,0

Tabelle A.8: Häufigkeiten von Co007 in Duisburg 2015 (Befragungsteilnahme letzte Erhebung ja/nein)

	Häufigkeit	%
ja	2412	87,0
nein	285	10,3
k. A.	75	2,7
unplausibel	—	—
Gesamt	2772	100,0

Tabelle A.9: Häufigkeiten von Co007 in Duisburg 2017 (Befragungsteilnahme letzte Erhebung ja/nein)

	Häufigkeit	%
ja	2349	84,0
nein	363	13,0
k. A.	85	3,0
unplausibel	—	—
Gesamt	2797	100,0

Tabelle A.10: Häufigkeiten von Co007 in Duisburg 2019 (Befragungsteilnahme letzte Erhebung ja/nein)

	Häufigkeit	%
ja	2503	92,4
nein	144	5,3
k. A.	62	2,3
unplausibel	—	—
Gesamt	2709	100,0

A.3 Häufigkeiten der Gesamtcodes

Tabelle A.11: *Häufigkeiten der Gesamtcodes in Duisburg 2015*

	Häufigkeit	%
1	2733	98,6
2	4	0,1
fehlend	35	1,3
Gesamt	2772	100,0

Tabelle A.12: *Häufigkeiten der Gesamtcodes in Duisburg 2017*

	Häufigkeit	%
1	2759	98,6
2	4	0,1
fehlend	34	1,2
Gesamt	2797	100,0

Tabelle A.13: *Häufigkeiten der Gesamtcodes in Duisburg 2019*

	Häufigkeit	%
1	2668	98,5
2	8	0,3
fehlend	33	1,2
Gesamt	2709	100,0