

**Methodendokumentation  
der kriminologischen Schülerbefragung  
in Duisburg 2002 bis 2013**

– Zehn-Wellen-Panel –

Andreas Daniel

Anke Erdmann

**Schriftenreihe:  
Jugendkriminalität in der modernen Stadt – Methoden  
Nr. 23 / 2017**

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

48 143 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).**

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Studiendesign</b>	<b>2</b>
<b>3</b>	<b>Erstellung der Paneldaten</b>	<b>5</b>
<b>4</b>	<b>Realisierte Zuordnungen</b>	<b>15</b>
4.1	Die verwendeten Datensätze und ihre Kurzbezeichnungen . . . . .	15
4.2	Abschätzung und Bewertung der Zuordnungsgüte . . . . .	18
<b>5</b>	<b>Zusammensetzung der Paneldaten</b>	<b>20</b>
5.1	Die Panelquerschnitte . . . . .	20
5.1.1	Panelquerschnitt 2005 ( $P_{Q_{t_4}}$ ) . . . . .	21
5.1.2	Panelquerschnitt 2006 ( $P_{Q_{t_5}}$ ) . . . . .	22
5.1.3	Panelquerschnitt 2007 ( $P_{Q_{t_6}}$ ) . . . . .	23
5.1.4	Panelquerschnitt 2008 ( $P_{Q_{t_7}}$ ) . . . . .	24
5.1.5	Panelquerschnitt 2009 ( $P_{Q_{t_8}}$ ) . . . . .	26
5.1.6	Panelquerschnitt 2011 ( $P_{Q_{t_9}}$ ) . . . . .	27
5.1.7	Panelquerschnitt 2013 ( $P_{Q_{t_{10}}}$ ) . . . . .	27
5.2	Die Zwei-Wellen-Panel . . . . .	28
5.2.1	Zwei-Wellen-Panel 2007 und 2008 ( $P_{t_{6,7}}$ ) . . . . .	28
5.2.2	Zwei-Wellen-Panel 2008 und 2009 ( $P_{t_{7,8}}$ ) . . . . .	30
5.2.3	Zwei-Wellen-Panel 2009 und 2011 ( $P_{t_{8,9}}$ ) . . . . .	32
5.2.4	Zwei-Wellen-Panel 2011 und 2013 ( $P_{t_{9,10}}$ ) . . . . .	33
5.3	Die Drei-Wellen-Panel . . . . .	35
5.3.1	Drei-Wellen-Panel 2006, 2007 und 2008 ( $P_{t_{5,6,7}}$ ) . . . . .	35
5.3.2	Drei-Wellen-Panel 2007, 2008 und 2009 ( $P_{t_{6,7,8}}$ ) . . . . .	37
5.3.3	Drei-Wellen-Panel 2008, 2009 und 2011 ( $P_{t_{7,8,9}}$ ) . . . . .	39
5.3.4	Drei-Wellen-Panel 2009, 2011 und 2013 ( $P_{t_{8,9,10}}$ ) . . . . .	41
5.4	Die Vier-Wellen-Panel . . . . .	42
5.4.1	Vier-Wellen-Panel 2005, 2006, 2007 und 2008 ( $P_{t_{4,5,6,7}}$ ) . . . . .	42
5.4.2	Vier-Wellen-Panel 2006, 2007, 2008 und 2009 ( $P_{t_{5,6,7,8}}$ ) . . . . .	45
5.4.3	Vier-Wellen-Panel 2007, 2008, 2009 und 2011 ( $P_{t_{6,7,8,9}}$ ) . . . . .	47
5.4.4	Vier-Wellen-Panel 2008, 2009, 2011 und 2013 ( $P_{t_{7,8,9,10}}$ ) . . . . .	49

5.5	Die Fünf-Wellen-Panel . . . . .	51
5.5.1	Fünf-Wellen-Panel 2004, 2005, 2006, 2007 und 2008 ( $P_{t_{3,4,5,6,7}}$ ) . . . . .	51
5.5.2	Fünf-Wellen-Panel 2005, 2006, 2007, 2008 und 2009 ( $P_{t_{4,5,6,7,8}}$ ) . . . . .	53
5.5.3	Fünf-Wellen-Panel 2006, 2007, 2008, 2009 und 2011 ( $P_{t_{5,6,7,8,9}}$ ) . . . . .	55
5.5.4	Fünf-Wellen-Panel 2007, 2008, 2009, 2011 und 2013 ( $P_{t_{6,7,8,9,10}}$ ) . . . . .	57
5.6	Die Sechs-Wellen-Panel . . . . .	59
5.6.1	Sechs-Wellen-Panel 2003, 2004, 2005, 2006, 2007 und 2008 ( $P_{t_{2,3,4,5,6,7}}$ ) . . . . .	60
5.6.2	Sechs-Wellen-Panel 2004, 2005, 2006, 2007, 2008 und 2009 ( $P_{t_{3,4,5,6,7,8}}$ ) . . . . .	62
5.6.3	Sechs-Wellen-Panel 2005, 2006, 2007, 2008, 2009 und 2011 ( $P_{t_{4,5,6,7,8,9}}$ ) . . . . .	65
5.6.4	Sechs-Wellen-Panel 2006, 2007, 2008, 2009, 2011 und 2013 ( $P_{t_{5,6,7,8,9,10}}$ ) . . . . .	68
5.7	Die Sieben-Wellen-Panel . . . . .	70
5.7.1	Sieben-Wellen-Panel 2002, 2003, 2004, 2005, 2006, 2007 und 2008 ( $P_{t_{1,2,3,4,5,6,7}}$ ) . . . . .	70
5.7.2	Sieben-Wellen-Panel 2003, 2004, 2005, 2006, 2007, 2008 und 2009 ( $P_{t_{2,3,4,5,6,7,8}}$ ) . . . . .	73
5.7.3	Sieben-Wellen-Panel 2004, 2005, 2006, 2007, 2008, 2009 und 2011 ( $P_{t_{3,4,5,6,7,8,9}}$ ) . . . . .	76
5.7.4	Sieben-Wellen-Panel 2005, 2006, 2007, 2008, 2009, 2011 und 2013 ( $P_{t_{4,5,6,7,8,9,10}}$ ) . . . . .	79
5.8	Die Acht-Wellen-Panel . . . . .	82
5.8.1	Acht-Wellen-Panel 2002, 2003, 2004, 2005, 2006, 2007, 2008 und 2009 ( $P_{t_{1,2,3,4,5,6,7,8}}$ ) . . . . .	82
5.8.2	Acht-Wellen-Panel 2003, 2004, 2005, 2006, 2007, 2008, 2009 und 2011 ( $P_{t_{2,3,4,5,6,7,8,9}}$ ) . . . . .	86
5.8.3	Acht-Wellen-Panel 2004, 2005, 2006, 2007, 2008, 2009, 2011 und 2013 ( $P_{t_{3,4,5,6,7,8,9,10}}$ ) . . . . .	89
5.9	Die Neun-Wellen-Panel . . . . .	92
5.9.1	Neun-Wellen-Panel 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009 und 2011 ( $P_{t_{1,2,3,4,5,6,7,8,9}}$ ) . . . . .	92
5.9.2	Neun-Wellen-Panel 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011 und 2013 ( $P_{t_{2,3,4,5,6,7,8,9,10}}$ ) . . . . .	95
5.10	Das Zehn-Wellen-Panel 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011 und 2013 ( $P_{t_{1,2,3,4,5,6,7,8,9,10}}$ ) . . . . .	99

<b>6</b>	<b>Zusammenfassung und Erklärung</b>	<b>104</b>
----------	--------------------------------------	------------

<b>Literaturverzeichnis</b>	<b>107</b>
<b>A Anhang</b>	<b>109</b>
A.1 Codeblätter . . . . .	109
A.2 Häufigkeiten der Codevariablen . . . . .	113

## Tabellenverzeichnis

3.1	Durchgeführte Kontrollschritte 2007/2008 . . . . .	8
3.2	Durchgeführte Kontrollschritte 2008/2009 . . . . .	8
3.3	Durchgeführte Kontrollschritte 2009/2011 . . . . .	9
3.4	Durchgeführte Kontrollschritte 2011/2013 . . . . .	9
3.5	Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2007/2008	11
3.6	Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2008/2009	12
3.7	Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2009/2011	13
3.8	Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2011/2013	14
4.2	Ausschöpfungen der durchgehenden Paneldatensätze . . . . .	19
5.1	Zusammensetzung nach Geschlecht $P_{Q_{t_4}}$ . . . . .	21
5.2	Altersdurchschnitt $P_{Q_{t_4}}$ . . . . .	21
5.3	Zusammensetzung $P_{Q_{t_4}}$ nach Schulform (Zeilenproz.) . . . . .	22
5.4	Zusammensetzung nach Geschlecht $P_{Q_{t_5}}$ . . . . .	22
5.5	Altersdurchschnitt $P_{Q_{t_5}}$ . . . . .	22
5.6	Zusammensetzung $P_{Q_{t_5}}$ nach Schulform (Zeilenproz.) . . . . .	23
5.7	Zusammensetzung nach Geschlecht $P_{Q_{t_6}}$ . . . . .	23
5.8	Altersdurchschnitt $P_{Q_{t_6}}$ . . . . .	24
5.9	Zusammensetzung $P_{Q_{t_6}}$ nach Schulform (Zeilenproz.) . . . . .	24
5.10	Zusammensetzung nach Geschlecht $P_{Q_{t_7}}$ . . . . .	25
5.11	Altersdurchschnitt $P_{Q_{t_7}}$ . . . . .	25
5.12	Zusammensetzung $P_{Q_{t_7}}$ nach Schulform (Zeilenproz.) . . . . .	25
5.13	Zusammensetzung nach Geschlecht $P_{Q_{t_8}}$ . . . . .	26
5.14	Altersdurchschnitt $P_{Q_{t_8}}$ . . . . .	26
5.15	Zusammensetzung $P_{Q_{t_8}}$ nach Schulform (Zeilenproz.) . . . . .	26
5.16	Zusammensetzung nach Geschlecht $P_{Q_{t_9}}$ . . . . .	27
5.17	Altersdurchschnitt $P_{Q_{t_9}}$ . . . . .	27
5.18	Zusammensetzung nach Geschlecht $P_{Q_{t_{10}}}$ . . . . .	28
5.19	Altersdurchschnitt $P_{Q_{t_{10}}}$ . . . . .	28
5.20	Zusammensetzung nach Geschlecht $P_{t_{6,7}}$ . . . . .	29
5.21	Altersdurchschnitt $P_{t_{6,7}}$ . . . . .	29
5.22	Zusammensetzung $P_{t_{6,7}}$ nach Schulform (Zeilenproz.) . . . . .	30
5.23	Zusammensetzung nach Geschlecht $P_{t_{7,8}}$ . . . . .	30

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

5.64	Zusammensetzung nach Geschlecht $P_{t_{6,7,8,9,10}}$ . . . . .	58
5.65	Altersdurchschnitt $P_{t_{6,7,8,9,10}}$ . . . . .	59
5.66	Zusammensetzung $P_{t_{6,7,8,9,10}}$ nach Schulform (Zeilenproz.) . . . . .	59
5.67	Zusammensetzung nach Geschlecht $P_{t_{2,3,4,5,6,7}}$ . . . . .	60
5.68	Altersdurchschnitt $P_{t_{2,3,4,5,6,7}}$ . . . . .	61
5.69	Zusammensetzung $P_{t_{2,3,4,5,6,7}}$ nach Schulform (Zeilenproz.) . . . . .	62
5.70	Zusammensetzung nach Geschlecht $P_{t_{3,4,5,6,7,8}}$ . . . . .	63
5.71	Altersdurchschnitt $P_{t_{3,4,5,6,7,8}}$ . . . . .	64
5.72	Zusammensetzung $P_{t_{3,4,5,6,7,8}}$ nach Schulform (Zeilenproz.) . . . . .	65
5.73	Zusammensetzung nach Geschlecht $P_{t_{4,5,6,7,8,9}}$ . . . . .	66
5.74	Altersdurchschnitt $P_{t_{4,5,6,7,8,9}}$ . . . . .	67
5.75	Zusammensetzung $P_{t_{4,5,6,7,8,9}}$ nach Schulform (Zeilenproz.) . . . . .	67
5.76	Zusammensetzung nach Geschlecht $P_{t_{5,6,7,8,9,10}}$ . . . . .	68
5.77	Altersdurchschnitt $P_{t_{5,6,7,8,9,10}}$ . . . . .	69
5.78	Zusammensetzung $P_{t_{5,6,7,8,9,10}}$ nach Schulform (Zeilenproz.) . . . . .	70
5.79	Zusammensetzung nach Geschlecht $P_{t_{1,2,3,4,5,6,7}}$ . . . . .	71
5.80	Altersdurchschnitt $P_{t_{1,2,3,4,5,6,7}}$ . . . . .	72
5.81	Zusammensetzung $P_{t_{1,2,3,4,5,6,7}}$ nach Schulform (Zeilenproz.) . . . . .	73
5.82	Zusammensetzung nach Geschlecht $P_{t_{2,3,4,5,6,7,8}}$ . . . . .	74
5.83	Altersdurchschnitt $P_{t_{2,3,4,5,6,7,8}}$ . . . . .	75
5.84	Zusammensetzung $P_{t_{2,3,4,5,6,7,8}}$ nach Schulform (Zeilenproz.) . . . . .	76
5.85	Zusammensetzung nach Geschlecht $P_{t_{3,4,5,6,7,8,9}}$ . . . . .	77
5.86	Altersdurchschnitt $P_{t_{3,4,5,6,7,8,9}}$ . . . . .	78
5.87	Zusammensetzung $P_{t_{3,4,5,6,7,8,9}}$ nach Schulform (Zeilenproz.) . . . . .	79
5.88	Zusammensetzung nach Geschlecht $P_{t_{4,5,6,7,8,9,10}}$ . . . . .	80
5.89	Altersdurchschnitt $P_{t_{4,5,6,7,8,9,10}}$ . . . . .	81
5.90	Zusammensetzung $P_{t_{4,5,6,7,8,9,10}}$ nach Schulform (Zeilenproz.) . . . . .	82
5.91	Zusammensetzung nach Geschlecht $P_{t_{1,2,3,4,5,6,7,8}}$ . . . . .	83
5.92	Altersdurchschnitt $P_{t_{1,2,3,4,5,6,7,8}}$ . . . . .	84
5.93	Zusammensetzung $P_{t_{1,2,3,4,5,6,7,8}}$ nach Schulform (Zeilenproz.) . . . . .	85
5.94	Zusammensetzung nach Geschlecht $P_{t_{2,3,4,5,6,7,8,9}}$ . . . . .	86
5.95	Altersdurchschnitt $P_{t_{2,3,4,5,6,7,8,9}}$ . . . . .	87
5.96	Zusammensetzung $P_{t_{2,3,4,5,6,7,8,9}}$ nach Schulform (Zeilenproz.) . . . . .	88
5.97	Zusammensetzung nach Geschlecht $P_{t_{3,4,5,6,7,8,9,10}}$ . . . . .	89
5.98	Altersdurchschnitt $P_{t_{3,4,5,6,7,8,9,10}}$ . . . . .	90
5.99	Zusammensetzung $P_{t_{3,4,5,6,7,8,9,10}}$ nach Schulform (Zeilenproz.) . . . . .	91
5.100	Zusammensetzung nach Geschlecht $P_{t_{1,2,3,4,5,6,7,8,9}}$ . . . . .	93
5.101	Altersdurchschnitt $P_{t_{1,2,3,4,5,6,7,8,9}}$ . . . . .	94
5.102	Zusammensetzung $P_{t_{1,2,3,4,5,6,7,8,9}}$ nach Schulform (Zeilenproz.) . . . . .	95
5.103	Zusammensetzung nach Geschlecht $P_{t_{2,3,4,5,6,7,8,9,10}}$ . . . . .	96

5.104 Altersdurchschnitt $P_{t_{2,3,4,5,6,7,8,9,10}}$ . . . . .	97
5.105 Zusammensetzung $P_{t_{2,3,4,5,6,7,8,9,10}}$ nach Schulform (Zeilenproz.) . . . . .	99
5.106 Zusammensetzung nach Geschlecht $P_{t_{1,2,3,4,5,6,7,8,9,10}}$ . . . . .	100
5.107 Altersdurchschnitt $P_{t_{1,2,3,4,5,6,7,8,9,10}}$ . . . . .	101
5.108 Zusammensetzung $P_{t_{1,2,3,4,5,6,7,8,9,10}}$ nach Schulform (Zeilenproz.) . . . . .	103
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 2008 . . . . .	120
A.9 Häufigkeiten von Co007 in Duisburg 2009 . . . . .	120
A.10 Häufigkeiten von Co007 in Duisburg 2011 . . . . .	120
A.11 Häufigkeiten von Co007 in Duisburg 2013 . . . . .	120
A.12 Häufigkeiten von Co009 in Duisburg 2008 . . . . .	121
A.13 Häufigkeiten der Gesamtcodes in Duisburg 2008 . . . . .	121
A.14 Häufigkeiten der Gesamtcodes in Duisburg 2009 . . . . .	121
A.15 Häufigkeiten der Gesamtcodes in Duisburg 2011 . . . . .	121
A.16 Häufigkeiten der Gesamtcodes in Duisburg 2013 . . . . .	122

# Abbildungsverzeichnis

2.1	Erhebungsdesign der Studie „Kriminalität in der modernen Stadt“ . . . .	4
A.1	Codeblatt Duisburg 2008 . . . . .	109
A.2	Codeblatt Duisburg 2009 . . . . .	110
A.3	Codeblatt Duisburg 2011 . . . . .	111
A.4	Codeblatt Duisburg 2013 . . . . .	112

# 1 Einleitung

Die kriminologische und kriminalsoziologisch orientierte Studie *Kriminalität in der modernen Stadt*<sup>1</sup> 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, 2007) 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 *Paneldatensatzes* der jüngeren Kohorte im Mittelpunkt steht, wird auf eine Beschreibung der Querschnittdatensätze verzichtet. Auch Ergebnisse und Analysen der Studie werden nicht dargestellt<sup>2</sup>.

---

1 Projektleitung: Prof. Dr. Klaus Boers, Institut für Kriminalwissenschaften, WWU Münster und Prof. Dr. Jost Reinecke, Fakultät für Soziologie, Universität Bielefeld, DFG-Förderzeichen: Bo 1234/6-1, 6-2 und Re 832/4-1, 4-2.

2 Zusammenfassende Darstellungen der Analysen finden sich unter anderem bei Boers, Reinecke et al. (Boers et al., 2014) sowie Boers, Reinecke et al. (Boers, Reinecke, Mariotti & Seddig, 2010).

## 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 Schüler der ehemals neunten Jahrgangsstufe lediglich ein weiteres Mal (2003) befragt wurden<sup>3</sup>, erfolgte die Befragung der ehemaligen Siebtklässler bis zum Jahr 2015. Der jährliche Erhebungsrhythmus wurde dabei bis zum Jahr 2009 beibehalten und im Folgenden auf einen zweijährigen Rhythmus umgestellt. Die weiteren Erhebungen erfolgten somit in den Jahren 2011, 2013 und 2015. 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 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 Querschnittdatensatzes 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 wurde 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 Face-to-Face-Befragung ergänzt. Dieser Erhebungsform setzte sich

---

<sup>3</sup> Hieraus entstand ein Zwei-Wellen-Paneldatensatz, siehe Pollich (2007).

bis zur aktuellsten Erhebung im Jahr 2015 fort. Im Jahr 2010 erfolgte erstmals eine Erhebungspause. Die *neunte* Erhebungswelle folgte somit im Jahr 2011. 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). Die für diesen Bericht letzte relevante Erhebungswelle wurde im Jahr 2013 erhoben. Des Weiteren liegt aktuell bereits die elfte Erhebungswelle 2015 vor (vgl. Verneuer, 2017). Zwei weitere Erhebungswellen der Jahre 2017 und 2019 folgen, sodass für die Duisburger Befragung insgesamt 13 Erhebungswellen vorliegen werden, welche die Altersspanne vom 13. bis zum 30. Lebensjahr abdecken.

Der vorliegende Bericht beschreibt die Entstehung und Repräsentativität des Zehn-Wellen-Paneldatensatzes (2002 bis 2013). Dieser Bericht stellt damit die Fortführung der Berichte zu dem Vier- bzw. Sechs-Wellen-Paneldatensatz dar. Dementsprechend werden nur jene Datensätze behandelt, die nicht bereits in den vorangegangenen Berichten beschrieben wurden. Aufgrund der Tatsache, dass die Anzahl der sich ergebenden Datensätze mit jeder Welle exponentiell steigt, ist es aus Platzgründen nicht mehr möglich, diese vollumfänglich zu beschreiben. Dargestellt werden alle durchgehenden Paneldatensätze, sowie eine Auswahl weiterer exemplarischer Datensätze.



### 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. Jeder 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 des Probanden oder seines 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 Codeblättern<sup>4</sup> 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 Nachnamens<sup>5</sup>

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

4 Abdrucke der Codeblätter der Erhebungsjahre 2008, 2009, 2011 und 2013 finden sich in Abbildungen A.1, A.2, A.3 und A.4 (im Anhang). Codeblätter früherer Erhebungen finden sich bei Pöge (2007, S. 88 ff.) und Pollich (2010, S. 183 ff.)

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

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 lediglich die Frage nach der Befragungsteilnahme im Vorjahr erhalten (vgl. auch Pollich, 2010, S. 8).

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

Daher wurden zusätzlich zu den sechs bzw. sieben Codefragen und der Zusatzfrage<sup>7</sup> für die Fragebogenzuordnungen das Geschlecht und die Staatsangehörigkeit<sup>8</sup> der Befragten, sowie die (zuletzt) besuchte Schule<sup>9</sup>, 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).<sup>10</sup> 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.13 bis A.16 im Anhang), bleibt das Problem der Reproduktion, dass also 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, 2010, S. 8). Für den Zuordnungsschritt 2008/2009 kann die Rate der exakten Codereproduktion auf 71 % gesteigert werden, bevor sie in den Schritten 2009/2011 und 2011/2013

---

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.12 im Anhang.

8 1 *Deutsch*, 2 *Türkisch*, 3 *Italienisch*, 4 *ehem. Jugoslawien*, 5 *Polnisch*, 6 *Russisch*, 7 *Rumänien*, 8 *andere*, 9 *keine Angabe*.

9 25 Schulen aus Duisburg + Kategorie *eine andere Schule*.

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

auf etwa 64 % abfällt. Dieser recht starke Rückgang der Reproduktionsgüte kann vor allem durch die Umstellung auf einen Zweijahresrhythmus erklärt werden.

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, so dass sie für die nachfolgenden Zuordnungsschritte nicht mehr zur Verfügung standen. 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 das Geschlecht übereinstimmen. Zudem wurden die Befragten anhand einer Zusatzvariable (Codestelle *Co008*, Variable *Co007* im Datensatz) zugeordnet, durch die abgefragt wurde, ob sie im letzten Jahr an der Befragung teilgenommen haben. 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.

Die Verwendung der Schulinformation wurde, wie schon in Pollich (2010, S. 10) beschrieben, ab dem Verlassen der Sekundarstufe I problematisch, da die Schulwechsel zu einer Instabilität der Schulinformation führten. Der Wegfall der Schulinformationen wurde jedoch durch die Einführung der neuen Codevariable (siehe oben) ab dem Zuordnungsschritt 2009/2011 aufgefangen.

Durch die deutliche Erhöhung des Stichprobenumfangs im Querschnitt 2006 erhöhte sich ebenfalls die Zahl der notwendigen Kontrollen in allen Kontrollschritten, welche dieses Erhebungsjahr mit einschlossen. Dies machte es notwendig, die hierarchischen Restriktionen feingliedriger zu lockern( vgl. Pollich, 2010, S. 10). Ab dem Jahr 2007 verringerte sich der Kontrollaufwand wieder auf das zuvor gewohnte Maß. Aus diesem Grund konnte auch die Zahl der Kontrollschritte reduziert werden, ohne dass sich daraus ein überdurchschnittlicher Kontrollaufwand pro Kontrollschritt ergab. Eine Übersicht

**Tabelle 3.1:** Durchgeführte Kontrollschritte 2007/2008

	Schritt	Codevariablen	Zusatzvariablen
ohne Fehler	S1	ohne Fehler	ohne Fehler
	S2	ohne Fehler	keine Bedingung
ein Fehler	S3	ein Fehler	ohne Fehler
	S4	ein Fehler	keine Bedingung
zwei Fehler	S5	zwei Fehler	ohne Fehler
	S6	zwei Fehler	nur Geschlecht gleich
	S7	zwei Fehler	Geschlecht gleich; Teilnahme letztes Jahr: ja
	S8	zwei Fehler	Geschlecht gleich; Teilnahme letztes Jahr: k. A.
	S9	zwei Fehler	Geschlecht gleich; Teilnahme letztes Jahr: nein
drei Fehler	S10	drei Fehler	Teilnahme letztes Jahr: frei

**Tabelle 3.2:** Durchgeführte Kontrollschritte 2008/2009

	Schritt	Codevariablen	Zusatzvariablen
ohne Fehler	S1	ohne Fehler	ohne Fehler
	S2	ohne Fehler	Geschlecht gleich
ein Fehler	S3	ein Fehler	ohne Fehler
	S4	ein Fehler	keine Bedingung
zwei Fehler	S5	zwei Fehler	ohne Fehler
	S6	zwei Fehler	nur Geschlecht gleich
	S7	zwei Fehler	ohne Bedingung
drei Fehler	S8	drei Fehler	ohne Fehler
	S9	drei Fehler	nur Staatsangehörigkeit frei

über die durchgeführten Zuordnungsschritte und die Bedingungen, die der Code erfüllen musste, findet sich in den Tabellen 3.1 bis 3.4.

In 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. Es zeigt sich, dass der Anteil der passenden Bögen an den Gesamtkontrollen in den Zuordnungsschritten 2007/2008 (61,2 %) und 2008/2009 (68,1 %) gegenüber den Zuordnungsschritten 2005/2006 (43,2 %) und 2006/2007 (48,2 %) deutlich gesteigert werden konnte. Dies ist zum einen auf die geringere Anzahl an Gesamtkontrollen zurückzuführen, jedoch können auch absolut, im Vergleich zu den beiden vorangegangenen Kontrollschritten, gleich viele bzw. mehr Fragebögenpaare einander zugeordnet werden. Ohne die zusätzlichen Fälle des Jahres 2006 haben die Kontrollen somit das Effektivitätsniveau der Kontrollschritte 2002/2003 bis 2004/2005 erreicht. In den Kontrollschritten 2009/2011 und 2011/2013 erfolgt ein

**Tabelle 3.3:** Durchgeführte Kontrollschritte 2009/2011

	Schritt	Codevariablen	Zusatzvariablen
ohne Fehler	S1	ohne Fehler	ohne Fehler
	S2	ohne Fehler	Geschlecht gleich
	S3	ohne Fehler	keine Bedingung
ein Fehler	S4	ein Fehler	ohne Fehler
	S5	ein Fehler	keine Bedingung
zwei Fehler	S6	zwei Fehler	ohne Fehler
	S7	zwei Fehler	keine Bedingung
drei Fehler	S8	drei Fehler	ohne Fehler
	S9	drei Fehler	nur Teiln. letztes Jahr = frei

**Tabelle 3.4:** Durchgeführte Kontrollschritte 2011/2013

	Schritt	Codevariablen	Zusatzvariablen
ohne Fehler	S1	ohne Fehler	ohne Fehler
	S2	ohne Fehler	Geschlecht gleich
ein Fehler	S3	ein Fehler	ohne Fehler
	S4	ein Fehler	keine Bedingung
zwei Fehler	S5	zwei Fehler	ohne Fehler
	S6	zwei Fehler	keine Bedingung
drei Fehler	S7	drei Fehler	ohne Fehler
	S8	drei Fehler	nur Teiln. letztes Jahr = frei

deutlicher Abfall des Anteils passender Bögen an den Gesamtkontrollen (2009/2011 52,5 %; 2011/2013 55,0 %). Auch die Anzahl der absoluten Zuordnungen geht in beiden Jahren zurück. Dieser Rückgang war zu erwarten, da zwischen den Erhebungen zwei anstatt ein Jahr lagen und somit die Reproduktionsgüte der Codes nachlässt.

Wie in den Jahren zuvor wurde nach Abschluss jeder Erhebung 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.<sup>11</sup> Insgesamt belief sich der Aufwand dieser Vorkontrollen im Jahr 2009 auf 670, im Jahr 2011 auf 447 und im Jahr 2013 auf 371 durchgeführte Abgleiche.

Auch für die Kontrollen nach dem Jahr 2006 gilt, dass nicht direkt aufeinander folgenden Erhebungswellen ebenfalls kontrolliert wurden. Nicht im Panel zugeordnete Bögen eines Querschnittsdatensatzes wurden stets auch mit den nicht zugeordneten Bögen aus

<sup>11</sup> Dabei wurde zunächst jener Fragebogen mit dem früheren Teilnahmezeitpunkt behalten, ab dem Jahr 2011 wurde jener Fragebogen entfernt, der weniger brauchbare Informationen enthielt.

dem *übernächsten* 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.

**Tabelle 3.5:** Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2007/2008

**(a)** absolut durchgeführte Kontrollen pro Zuordnungsschritt

		Kontrollen gesamt	passt	passt nicht	fehlt <sup>b</sup>
ohne Fehler	S1	1403	1343	30	30
	S2	584	506	36	42
	gesamt	1 987	1 849	66	72
ein Fehler	S3	415	370	16	29
	S4	333	251	68	14
	S5	38	7	22	9
	gesamt	786	628	106	52
zwei Fehler	S6	138	104	23	11
	S7	451	65	367	19
	S8	739	24	684	31
	gesamt	1 328	193	1 074	61
drei Fehler	S9	194	24	154	16
	S10	112	4	88	20
	gesamt	306	28	242	36
<b>Summe</b>		<b>4 407</b>	<b>2 698</b>	<b>1 488</b>	

<sup>b</sup> 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 <sup>b</sup>
ohne Fehler	S1	100,0	95,7	2,1	2,1
	S2	100,0	86,6	6,2	7,2
	gesamt	100,0	93,1	3,3	3,6
ein Fehler	S3	100,0	89,2	3,9	7,0
	S4	100,0	75,4	20,4	4,2
	S5	100,0	18,4	57,9	23,7
	gesamt	100,0	79,9	13,5	6,6
zwei Fehler	S6	100,0	75,4	16,7	8,0
	S7	100,0	14,4	81,4	4,2
	S8	100,0	3,2	92,6	4,2
	gesamt	100,0	18,9	80,9	4,6
drei Fehler	s9	100,0	12,4	79,4	8,2
	S10	100,0	3,6	78,6	17,9
	gesamt	100,0	9,2	79,1	11,8
<b>Summe</b>		<b>100,0</b>	<b>61,2</b>	<b>33,8</b>	

**Tabelle 3.6:** Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2008/2009

**(a)** absolut durchgeführte Kontrollen pro Zuordnungsschritt

		Kontrollen gesamt	passt	passt nicht	fehlt <sup>b</sup>
ohne Fehler	S1	444	441	2	1
	S2	1505	1463	40	2
	gesamt	1 949	1 904	42	3
ein Fehler	S3	138	131	6	1
	S4	619	472	145	2
	gesamt	757	603	151	3
zwei Fehler	S5	76	39	36	1
	S6	651	152	487	12
	S7	293	2	278	13
	gesamt	1 020	193	801	26
drei Fehler	S8	180	6	174	0
	S9	69	0	69	0
	gesamt	249	6	243	0
<b>Summe</b>		<b>3 975</b>	<b>2 706</b>	<b>1 237</b>	

<sup>b</sup> 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 <sup>b</sup>
ohne Fehler	S1	100,0	99,3	0,4	0,2
	S2	100,0	97,2	2,7	0,1
	gesamt	100,0	97,7	2,2	0,2
ein Fehler	S3	100,0	94,9	4,4	0,7
	S4	100,0	76,3	23,4	0,3
	gesamt	100,0	79,7	19,9	0,4
zwei Fehler	S5	100,0	51,3	47,4	1,3
	S6	100,0	23,3	74,8	1,8
	S7	100,0	0,7	94,9	4,4
	gesamt	100,0	18,9	78,5	2,5
drei Fehler	S8	100,0	3,3	96,7	0,0
	S9	100,0	0	100	0
	gesamt	100,0	2,4	97,6	0
<b>Summe</b>		<b>100,0</b>	<b>68,1</b>	<b>31,1</b>	

**Tabelle 3.7:** Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2009/2011

**(a)** absolut durchgeführte Kontrollen pro Zuordnungsschritt

		Kontrollen gesamt	passt	passt nicht	fehlt <sup>b</sup>
ohne Fehler	S1	1 448	1351	67	30
	S2	351	263	61	27
	S3	40	29	5	6
	gesamt	1 839	1 643	133	63
ein Fehler	S4	597	522	64	11
	S5	152	129	19	4
	gesamt	749	651	83	15
zwei Fehler	S6	229	147	76	6
	S7	166	35	127	4
	gesamt	395	182	203	10
drei Fehler	S8	684	26	644	14
	S9	1125	13	1103	9
	gesamt	1 809	39	1 747	23
Summe		4 792	2 515	2 166	

<sup>b</sup> die Bögen in dieser Spalte werden nicht zu einer Gesamtzahl aufsummiert, da als fehlend deklarierte Bögen im nächsten Kontrollschritt stets wieder einbezogen wurden, um Ausfälle möglichst zu minimieren

**(b)** prozentuale Ergebnisse (Zeilenprozentuierung)

		Kontrollen gesamt	passt	passt nicht	fehlt <sup>b</sup>
ohne Fehler	S1	100	93,3	4,6	2,1
	S2	100	74,9	17,4	7,7
	S3	100	72,5	12,5	15,0
	gesamt	100	89,3	7,2	3,4
ein Fehler	S4	100	87,4	10,7	1,8
	S5	100	84,9	12,5	2,6
	gesamt	100	86,9	11,1	2,0
zwei Fehler	S6	100	64,2	33,2	2,6
	S7	100	21,1	76,5	2,4
	gesamt	100	46,1	51,4	2,5
drei Fehler	S8	100	3,8	94,2	2,0
	S9	100	1,2	98,0	0,8
	gesamt	100	2,2	96,6	1,3
Summe		100	52,5	45,2	

**Tabelle 3.8:** Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2011/2013

**(a)** absolut durchgeführte Kontrollen pro Zuordnungsschritt

		Kontrollen gesamt	passt	passt nicht	fehlt <sup>b</sup>
ohne Fehler	S1	1 388	1356	26	6
	S2	230	216	9	5
	gesamt	1 618	1572	35	11
ein Fehler	S3	545	512	29	4
	S4	135	107	24	4
	gesamt	680	619	53	8
zwei Fehler	S5	238	147	87	4
	S6	137	38	90	9
	gesamt	375	185	177	13
drei Fehler	S7	471	31	434	6
	S8	1260	14	1241	5
	gesamt	1 731	45	1 675	11
<b>Summe</b>		<b>4 404</b>	<b>2 421</b>	<b>1 940</b>	

<sup>b</sup> die Bögen in dieser Spalte werden nicht zu einer Gesamtzahl aufsummiert, da als fehlend deklarierte Bögen im nächsten Kontrollschritt stets wieder einbezogen wurden, um Ausfälle möglichst zu minimieren

**(b)** prozentuale Ergebnisse (Zeilenprozentuierung)

		Kontrollen gesamt	passt	passt nicht	fehlt <sup>b</sup>
ohne Fehler	S1	100	97,7	1,9	0,4
	S2	100	93,9	3,9	2,2
	gesamt	100	97,2	2,2	0,7
ein Fehler	S4	100	93,9	5,3,7	0,7
	S5	100	79,3	17,8	3,0
	gesamt	100	91,0	7,8	1,2
zwei Fehler	S6	100	61,8	36,6	1,7
	S7	100	27,7	65,7	6,6
	gesamt	100	49,3	47,2	3,5
drei Fehler	S8	100	6,6	92,1	1,3
	S9	100	1,1	98,5	0,4
	gesamt	100	2,6	96,8	0,6
<b>Summe</b>		<b>100</b>	<b>55,0</b>	<b>44,1</b>	

## 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 Drei-, Vier-, Fünf-, Sechs-, Sieben-, Acht-, Neun- und das Zehn-Wellen-Panel am bedeutsamsten, daher wird sich die Darstellung auf diese Datensätze konzentrieren. Das vorläufige Resultat der jährlichen Erhebungen in Duisburg ist ein Zehn-Wellen-Paneldatensatz der Jahre 2002 bis 2013 ( $t_{1,2,3,4,5,6,7,8,9,10}$ ) mit 911 durchgehend<sup>12</sup> und 4 076 insgesamt (mit Lücken) enthaltenen Fällen.<sup>13</sup>

### 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 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).

---

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

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

Ab dem Zeitpunkt  $t_4$  werden daher ausschließlich die Querschnittsdaten zur Beurteilung der Repräsentativität der einzelnen Paneldatensätze herangezogen. Zwar ist diese Grundlage – 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 ermöglichen, ob Geschlechter oder Schulformen in den Paneldatensätzen systematisch verzerrt repräsentiert sind.

Ein » $Q_{t_x}$ « 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.  $Q_{t_7}$  beinhaltet zum Beispiel alle 3 086 gültigen Fälle des Querschnittsdatensatzes aus dem Jahr 2008.

Ein » $P_{Q_{t_x}}$ « mit nur einem Zeitpunkt als Index bezeichnet einen so genannten Panelquerschnitt. Diese Panelquerschnitte beinhalten *alle* vorhandenen Fälle des gesamten Zehn-Wellen-Paneldatensatzes (mit Lücken) aus *einem* Erhebungszeitpunkt. So enthält zum Beispiel  $P_{Q_{t_7}}$  alle 2 412 gültigen Fälle die aus  $t_7$  im Gesamtpaneldatensatz enthalten sind. Dies sind im Unterschied zum Querschnitt  $Q_{t_7}$  alle Fälle, die zum Zeitpunkt  $t_7$  befragt wurden und im Zuge der Erstellung des Zehn-Wellen-Panels *mindestens einem* Fall eines beliebigen anderen Zeitpunktes zugeordnet werden konnten.

Die eigentlichen Paneldatensätze werden durch ein » $P_{t_x, y, \dots}$ « mit einem Index, der die jeweils einbezogenen Zeitpunkte umfasst, bezeichnet. So steht beispielsweise  $P_{t_6, 7}$  für den Zwei-Wellen-Paneldatensatz mit allen 2 299 *gültigen* Fällen, die zwischen  $t_6$  und  $t_7$  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. Insgesamt sind 380 Datensätze im Verlauf der Panelanalysen entstanden, diese können jedoch aus Kapazitätsgründen nicht umfassend dargestellt werden. Daher werden in diesem Methodenbericht nur jene Datensätze dargestellt, die *keine* Lücken aufweisen. Weitere Darstellungen von Datensätzen finden sich bei Pöge (2007) und Pollich (2010).

#### 4 Realisierte Zuordnungen

**Tabelle 4.1:** *Bezeichnungen und Fallzahlen aller betrachteten Panel- und Referenzdatensätze 2002 bis 2013*

Abkürzung	Fallzahl	Beschreibung
$S_{t_1}$	3 910	Amtliche Schulstatistik, Schuljahr 2001/2002 (Stichtag: 15. 10.) <sup>a</sup>
$S_{t_2}$	3 763	Amtliche Schulstatistik, Schuljahr 2002/2003 (Stichtag: 15. 10.) <sup>a</sup>
$S_{t_3}$	3 925	Amtliche Schulstatistik, Schuljahr 2003/2004 (Stichtag: 15. 10.) <sup>a</sup>
$Q_{t_1}$	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
$P_{Q_{t_4}}$	3 195	Panelquerschnittsdaten des Jahres 2005
$P_{Q_{t_5}}$	3 032	Panelquerschnittsdaten des Jahres 2006
$P_{Q_{t_6}}$	2 587	Panelquerschnittsdaten des Jahres 2007
$P_{Q_{t_7}}$	2 412	Panelquerschnittsdaten des Jahres 2008
$P_{Q_{t_8}}$	2 304	Panelquerschnittsdaten des Jahres 2009
$P_{Q_{t_9}}$	2 100	Panelquerschnittsdaten des Jahres 2011
$P_{Q_{t_{10}}}$	1 912	Panelquerschnittsdaten des Jahres 2013
$P_{t_{6,7}}$	2 299	Paneldaten der Jahre 2007 und 2008
$P_{t_{7,8}}$	2 209	Paneldaten der Jahre 2008 und 2009
$P_{t_{8,9}}$	1 997	Paneldaten der Jahre 2009 und 2011
$P_{t_{9,10}}$	1 813	Paneldaten der Jahre 2011 und 2013
$P_{t_{5,6,7}}$	2 163	Paneldaten der Jahre 2006, 2007, 2008
$P_{t_{6,7,8}}$	2 129	Paneldaten der Jahre 2007, 2008, 2009
$P_{t_{7,8,9}}$	1 920	Paneldaten der Jahre 2008, 2009, 2011
$P_{t_{8,9,10}}$	1 746	Paneldaten der Jahre 2009, 2011, 2013
$P_{t_{4,5,6,7}}$	1 951	Paneldaten der Jahre 2005, 2006, 2007, 2008
$P_{t_{5,6,7,8}}$	2 017	Paneldaten der Jahre 2006, 2007, 2008, 2009
$P_{t_{6,7,8,9}}$	1 854	Paneldaten der Jahre 2007, 2008, 2009, 2011
$P_{t_{7,8,9,10}}$	1 684	Paneldaten der Jahre 2008, 2009, 2011, 2013
$P_{t_{3,4,5,6,7}}$	1 741	Paneldaten der Jahre 2004, 2005, 2006, 2007, 2008
$P_{t_{4,5,6,7,8}}$	1 812	Paneldaten der Jahre 2005, 2006, 2007, 2008, 2009
$P_{t_{5,6,7,8,9}}$	1 768	Paneldaten der Jahre 2006, 2007, 2008, 2009, 2011
$P_{t_{6,7,8,9,10}}$	1 636	Paneldaten der Jahre 2007, 2008, 2009, 2011, 2013

**Tabelle 4.1:** Bezeichnungen und Fallzahlen aller betrachteten Panel- und Referenzdatensätze 2002 bis 2013

$P_{I_{2,3,4,5,6,7}}$	1 465	Paneldaten der Jahre 2003, 2004, 2005, 2006, 2007, 2008
$P_{I_{3,4,5,6,7,8}}$	1 616	Paneldaten der Jahre 2004, 2005, 2006, 2007, 2008, 2009
$P_{I_{4,5,6,7,8,9}}$	1 586	Paneldaten der Jahre 2005, 2006, 2007, 2008, 2009, 2011
$P_{I_{5,6,7,8,9,10}}$	1 567	Paneldaten der Jahre 2006, 2007, 2008, 2009, 2011, 2013
$P_{I_{1,2,3,4,5,6,7}}$	1 189	Paneldaten der Jahre 2002, 2003, 2004, 2005, 2006, 2007, 2008
$P_{I_{2,3,4,5,6,7,8}}$	1 365	Paneldaten der Jahre 2003, 2004, 2005, 2006, 2007, 2008, 2009
$P_{I_{3,4,5,6,7,8,9}}$	1 424	Paneldaten der Jahre 2004, 2005, 2006, 2007, 2008, 2009, 2011
$P_{I_{4,5,6,7,8,9,10}}$	1 415	Paneldaten der Jahre 2005, 2006, 2007, 2008, 2009, 2011, 2013
$P_{I_{1,2,3,4,5,6,7,8}}$	1 103	Paneldaten der Jahre 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009
$P_{I_{2,3,4,5,6,7,8,9}}$	1 222	Paneldaten der Jahre 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011
$P_{I_{3,4,5,6,7,8,9,10}}$	1 274	Paneldaten der Jahre 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013
$P_{I_{1,2,3,4,5,6,7,8,9}}$	1 002	Paneldaten der Jahre 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011
$P_{I_{2,3,4,5,6,7,8,9,10}}$	1 098	Paneldaten der Jahre 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013
$P_{I_{1,2,3,4,5,6,7,8,9,10}}$	911	Paneldaten der Jahre 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011, 2013

<sup>a</sup> 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%. Abgesehen von einem kleineren positiven Ausreißer (Acht-Wellen-Panel), bleibt diese Stabilität auch in den folgenden Jahren erhalten. Der durchgehende Paneldatensatz erweist sich somit als äußerst robust gegenüber Ausfällen. Anders als die in den Tabellen 3.7 und 3.8 in Kapitel 3 dargestellten totalen Zuordnungen vermuten lassen, zeigt auch der Wechsel auf einen zweijährigen Befragungsrhythmus keine gravierenden negativen Auswirkungen auf die Ausschöpfungsquote. Zwar ist ein Abfall von 93 auf 91 Prozent zu beobachten, dies stellt jedoch lediglich einen Rückgang auf das Ausgangsniveau dar. Insgesamt kann folglich festgehalten werden, dass die Teilnehmer dem Panel auch in der dritten Lebensdekade weitestgehend erhalten bleiben.

**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

## 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.<sup>14</sup> 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 Erhebungsjahren  $t_9$  und  $t_{10}$  so gut wie keine Personen mehr eine Schule besuchen, weshalb in den Tabellen auf die Aufschlüsselung nach besuchter Schulform für diese beiden 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 (ebenfalls gültigen) Bogen zumindest eines weiteren Zeitpunktes zugeordnet werden konnten. Sie enthalten dementsprechend nur diejenigen Fälle aus den Querschnitten nicht mehr, zu denen kein passender Bogen irgendeines weiteren Erhebungsjahres gefunden werden konnte. Bei den Panelquerschnitten handelt es sich um stark konstruierte Daten, die nur unter bestimmten Analysegesichtspunkten von Bedeutung sind.<sup>15</sup>

---

14 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.

15 Wenn also genau die Population der überhaupt zuordenbaren Fälle von Interesse ist.

### 5.1.1 Panelquerschnitt 2005 ( $P_{Q_{t_4}}$ )

Der Panelquerschnitt 2005 wurde bereits bei Pollich (2010) beschrieben, jedoch wurde die Fallzahl aufgrund einer vorgenommenen Panelkorrektur nach unten korrigiert. Aus diesem Grund wird der Datensatz hier nochmals dargestellt. Der Panelquerschnitt 2005 umfasst 3 195 Fälle und enthält damit 93,8 Prozent der Fälle des ursprünglichen Querschnittsdatsatzes.

Hinsichtlich des Merkmals Geschlecht liegt eine leichte Verschiebung im Panelquerschnitt vor, wie aus Tabelle 5.1 hervorgeht. Weibliche Befragte sind mit einer Differenz von knapp zwei Prozent im Panelquerschnitt stärker vertreten als im Querschnitt  $Q_{t_4}$ .

**Tabelle 5.1:** Zusammensetzung nach Geschlecht  $P_{Q_{t_4}}$

	männlich		weiblich		gesamt (100 %)
	Hfk	%	Hfk	%	
$Q_{t_4}$	1 717	50,4	1 688	49,6	3 405
$P_{Q_{t_4}}$	1 557	48,7	1 638	51,3	3 195

Bezüglich des durchschnittlichen Alters, wie es in Tabelle 5.2 abgebildet ist, lassen sich zwischen Querschnitt und Panelquerschnitt keine nennenswerten Differenzen feststellen: Im Querschnitt  $Q_{t_4}$  liegt das durchschnittliche Alter um 0,1 Jahre höher als im Panelquerschnitts des Zeitpunktes  $t_4$ .

**Tabelle 5.2:** Altersdurchschnitt  $P_{Q_{t_4}}$

	Alter	Std.-abw.	gesamt
$Q_{t_4}$	16,1	0,8	3 313
$P_{Q_{t_4}}$	16,0	0,8	3 138

Auch hinsichtlich der Zusammensetzung nach Schulform weisen die beiden Datensätze keine besonders auffälligen Differenzen zueinander auf (vgl. Tabelle 5.3). Im Panelquerschnitt  $P_{Q_{t_4}}$  sind 0,5 Prozent mehr Schülerinnen und Schüler an Gymnasien vertreten als im Querschnitt, umgekehrt ebenso 0,5 Prozent weniger Gesamtschülerinnen und -schüler.

**Tabelle 5.3:** Zusammensetzung  $P_{Q_{t_4}}$  nach Schulform (Zeilenprozentuierung)

	Schulform <sup>a</sup> (%)								gesamt (100 %)
	HS	RS	GS	GY	BK	SO	andere	keine	
$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
$P_{Q_{t_4}}$	21,8	21,8	32,2	22,5	0,8	0,3	0,5	0,1	3181

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

### 5.1.2 Panelquerschnitt 2006 ( $P_{Q_{t_5}}$ )

Der Panelquerschnitt 2006 wurde bereits bei Pollich (2010) beschrieben, jedoch wurde die Fallzahl aufgrund einer vorgenommenen Panelkorrektur nach unten korrigiert. Aus diesem Grund wird der Datensatz hier nochmals dargestellt. Der Panelquerschnitt 2006 umfasst 3 032 Fälle und enthält damit 66,7 Prozent der Fälle des ursprünglichen Querschnittsdatensatzes.

**Tabelle 5.4:** Zusammensetzung nach Geschlecht  $P_{Q_{t_5}}$

	männlich		weiblich		gesamt (100 %)
	Hfk	%	Hfk	%	
$Q_{t_5}$	2265	49,8	2282	50,2	4547
$P_{Q_{t_5}}$	1387	45,8	1645	54,2	3032

Hinsichtlich der Zusammensetzung des Geschlechts (Tabelle 5.4) lässt sich im Vergleich mit dem Querschnitt  $Q_{t_5}$  eine Verzerrung feststellen: Im Panelquerschnitt  $P_{Q_{t_5}}$  sind weibliche Befragte mit einer Differenz von vier Prozentpunkten überrepräsentiert. Männliche Befragte sind entsprechend im selben Maße unterrepräsentiert.

**Tabelle 5.5:** Altersdurchschnitt  $P_{Q_{t_5}}$

	Alter	Std.-abw.	gesamt
$Q_{t_5}$	17,4	1,3	4468
$P_{Q_{t_5}}$	17,1	0,9	2297

Bezüglich des durchschnittlichen Alters lässt sich eine Abweichung zwischen Querschnitt und Panelquerschnitt feststellen (vgl. Tabelle 5.5). Mit 17,4 Jahren liegt das Durchschnittsalter im Querschnitt  $Q_{t_5}$  etwas höher, gleichzeitig ist eine stärkere Heterogenität der Altersverteilung festzustellen. Anzumerken ist dabei, dass zum Zeitpunkt  $t_5$  eine temporäre Ausweitung der Stichprobe stattgefunden hat, was auch an der deutlich höheren Fallzahl zu erkennen ist.

**Tabelle 5.6:** Zusammensetzung  $P_{Q_{t_5}}$  nach Schulform (Zeilenprozentuierung)

	Schulform <sup>a</sup> (%)								gesamt (100 %)
	HS	RS	GS	GY	BK	SO	andere	keine	
$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
$P_{Q_{t_5}}$	0,5	0,4	20,4	28,3	50,2	0,2	—	—	2786

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

Auch bei der Zusammensetzung nach Schulform zeigen sich Differenzen zwischen den beiden Datensätzen, wie sich Tabelle 5.6 entnehmen lässt. Gymnasialschülerinnen und -schüler sind mit einer Differenz von etwa acht Prozentpunkten im Panelquerschnitt stärker vertreten als im Querschnitt. Selbes gilt für den Schultyp der Gesamtschule mit einer Differenz von gut vier Prozent. Hingegen sind Personen an Berufskollegs mit einer Abweichung von gut zwölf Prozent im Panelquerschnitt weniger stark vertreten als im Querschnitt  $Q_{t_5}$ . Die Differenzen sind an dieser Stelle aufgrund der temporären Ausweitung der Stichprobe durch Hinzunahme von Berufsschulen jedoch streng genommen wenig aussagekräftig und daher mit Vorsicht zu interpretieren.

### 5.1.3 Panelquerschnitt 2007 ( $P_{Q_{t_6}}$ )

Der Panelquerschnitt 2007 wurde bereits bei Pollich (2010) beschrieben, jedoch wurde die Fallzahl aufgrund einer vorgenommenen Panelkorrektur nach unten korrigiert. Aus diesem Grund wird der Datensatz hier nochmals dargestellt. Der Panelquerschnitt 2007 umfasst 2 587 Fälle und enthält damit 77,5 Prozent der Fälle des ursprünglichen Querschnittsdatsatzes.

**Tabelle 5.7:** Zusammensetzung nach Geschlecht  $P_{Q_{t_6}}$

	männlich		weiblich		gesamt (100 %)
	Hfk	%	Hfk	%	
$Q_{t_6}$	1 567	47,0	1 768	53,0	3 335
$P_{Q_{t_6}}$	1 118	43,2	1 469	56,8	2 587

Aus Tabelle 5.7 geht hervor, dass im Panelquerschnitt  $P_{Q_{t_6}}$  eine recht deutliche Verschiebung hinsichtlich des Geschlechts vorliegt: Vergleicht man die Geschlechterverteilung im Panelquerschnitt mit dem Querschnittsdatsatz des entsprechenden Erhebungsjahres ( $Q_{t_6}$ ), so wird deutlich, dass der Anteil der weiblichen Befragten im Panelquerschnitt um 3,8 Prozentpunkte höher liegt.

Das durchschnittliche Alter der Befragten in den Panelquerschnitten kann jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden:

**Tabelle 5.8:** Altersdurchschnitt  $P_{Q_{t_6}}$

	Alter	Std.-abw.	gesamt
$Q_{t_6}$	18,1	1,1	3320
$P_{Q_{t_6}}$	18,0	0,8	2576

Ein Vergleich zwischen den Querschnittsdaten  $Q_{t_6}$  und dem Panelquerschnitt zeigt, dass sich das Durchschnittsalter der Befragten in beiden Datensätzen nicht unterscheidet. Jedoch zeigt sich eine leicht höhere Standardabweichung in den Querschnittsdaten (siehe Tabelle 5.8).

**Tabelle 5.9:** Zusammensetzung  $P_{Q_{t_6}}$  nach Schulform (Zeilenprozentuierung)

	Schulform <sup>a</sup> (%)							gesamt (100 %)	
	HS	RS	GS	GY	BK	SO	andere keine		
$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
$P_{Q_{t_6}}$	—	—	17,7	27,8	31,3	—	10,8	12,5	2518

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

Die Betrachtung der Schulform in Tabelle 5.9 zeigt, dass Realschule, Hauptschule und Sonderschule in den betrachteten Datensätzen keine Rolle mehr spielen. Demgegenüber bleibt die Bedeutung der Schulen der Sekundarstufe II (Gymnasium, Gesamtschule), sowie der berufsbildenden Schulen erhalten. Der Panelquerschnitt  $P_{Q_{t_6}}$  zeigt sich im Vergleich zum Querschnitt  $Q_{t_6}$  insgesamt leicht nach besuchter Schulform verzerrt: Der Anteil der Schülerinnen und -schüler an Gesamt- und Berufsschulen, wie auch der Anteil jener Personen, die eine andere Schulform als die genannten besuchen unterscheidet sich kaum zwischen den Datensätzen. Der Anteil der Gymnasiastinnen und Gymnasiasten liegt im Panelquerschnitt mit 1,6 Prozentpunkten über denen des Querschnittsdatensatzes. Demgegenüber sind Personen die keine Schule mehr besuchen, mit 1,6 Prozentpunkten im Panelquerschnitt leicht unterrepräsentiert.

#### 5.1.4 Panelquerschnitt 2008 ( $P_{Q_{t_7}}$ )

Der Panelquerschnitt 2008 umfasst 2412 Fälle und enthält damit 78,2 Prozent der Fälle des ursprünglichen Querschnittsdatensatzes.

Aus Tabelle 5.10 geht hervor, dass im Panelquerschnitt  $P_{Q_{t_7}}$  eine recht deutliche Verschiebung hinsichtlich des Geschlechts vorliegt: Vergleicht man die Geschlechterverteilung im Panelquerschnitt mit dem Querschnittsdatensatz des entsprechenden Erhebungsjahres ( $Q_{t_7}$ ), so wird deutlich, dass der Anteil der weiblichen Befragten im Panelquerschnitt um 3,7 Prozentpunkte höher liegt.

**Tabelle 5.10:** Zusammensetzung nach Geschlecht  $P_{Q_{t_7}}$

	männlich		weiblich		gesamt
	Hfk	%	Hfk	%	(100 %)
$Q_{t_7}$	1 452	47,1	1 634	52,9	3 086
$P_{Q_{t_7}}$	1 048	43,4	1 364	56,6	2 412

**Tabelle 5.11:** Altersdurchschnitt  $P_{Q_{t_7}}$

	Alter	Std.-abw.	gesamt
$Q_{t_7}$	19,1	0,8	3 081
$P_{Q_{t_7}}$	19,0	0,8	2 410

Das durchschnittliche Alter der Befragten in den Panelquerschnitten kann jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Ein Vergleich zwischen den Querschnittsdaten  $Q_{t_7}$  und dem Panelquerschnitt zeigt, dass sich das Durchschnittsalter der Befragten in beiden Datensätzen nicht unterscheidet (siehe Tabelle 5.11).

**Tabelle 5.12:** Zusammensetzung  $P_{Q_{t_7}}$  nach Schulform (Zeilenprozentuierung)

	Schulform <sup>a</sup> (%)								gesamt
	HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2 975
$P_{Q_{t_7}}$	—	—	14,9	26,2	27,3	—	13,7	18,0	2 345

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

Die Betrachtung der Schulform in Tabelle 5.12 zeigt, dass Realschule, Hauptschule und Sonderschule in den betrachteten Datensätzen keine Rolle mehr spielen. Demgegenüber bleibt die Bedeutung der Schulen der Sekundarstufe II (Gymnasium, Gesamtschule), sowie der berufsbildenden Schulen erhalten. Der Panelquerschnitt  $P_{Q_{t_7}}$  zeigt sich im Vergleich zum Querschnitt  $Q_{t_7}$  insgesamt weitestgehend systematisch nach besuchter Schulform verzerrt: Der Anteil der Gesamtschülerinnen und -schüler liegt etwa einen und jener der Gymnasiasten etwa dreieinhalb Prozentpunkte über denen des Querschnittsdatensatzes. Demgegenüber sind Schülerinnen und Schüler der Berufsschulen 0,8, jene Personen die angaben keine Schule zu besuchen rund drei und jene die eine andere Schulform besuchten um einen Prozentpunkt im Panelquerschnitt unterrepräsentiert.

### 5.1.5 Panelquerschnitt 2009 ( $P_{Q_{t_8}}$ )

Der Panelquerschnitt 2009 umfasst 2 304 Fälle und enthält damit 74,6 Prozent der Fälle des ursprünglichen Querschnittsdatensatzes.

**Tabelle 5.13:** Zusammensetzung nach Geschlecht  $P_{Q_{t_8}}$

	männlich		weiblich		gesamt (100 %)
	Hfk	%	Hfk	%	
$Q_{t_8}$	1 446	46,8	1 644	53,2	3 090
$P_{Q_{t_8}}$	973	42,2	1 331	57,8	2 304

Aus Tabelle 5.13 geht hervor, dass im Panelquerschnitt  $P_{Q_{t_8}}$  eine recht deutliche Verschiebung hinsichtlich des Geschlechts vorliegt: Vergleicht man die Geschlechterverteilung im Panelquerschnitt mit dem Querschnittsdatensatz des entsprechenden Erhebungsjahres ( $Q_{t_8}$ ), so wird deutlich, dass der Anteil der weiblichen Befragten im Panelquerschnitt um 4,6 Prozentpunkte höher liegt.

**Tabelle 5.14:** Altersdurchschnitt  $P_{Q_{t_8}}$

	Alter	Std.-abw.	gesamt
$Q_{t_8}$	20,1	0,8	3 073
$P_{Q_{t_8}}$	20,0	0,8	2 297

Ein Vergleich des durchschnittlichen Alters der Befragten in den Querschnittsdaten  $Q_{t_8}$  und dem Panelquerschnitt zeigt, dass sich das Durchschnittsalter der Befragten in beiden Datensätzen nahezu nicht unterscheidet (siehe Tabelle 5.14).

**Tabelle 5.15:** Zusammensetzung  $P_{Q_{t_8}}$  nach Schulform (Zeilenprozentuierung)

	Schulform <sup>a</sup> (%)								gesamt (100 %)
	HS	RS	GS	GY	BK	SO	andere	keine	
$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2 934
$P_{Q_{t_8}}$	—	—	4,4	3,9	22,3	—	18,5	51,0	2 206

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

Die Betrachtung der Schulform in Tabelle 5.15 zeigt, dass die Hälfte aller Befragten keine Schule mehr besuchen und weitere 19 Prozent einen anderen Schultyp als die genannten. Diese Zahlen weisen kaum Abweichungen zwischen dem Panelquerschnitt  $P_{Q_{t_8}}$  und dem Querschnitt  $Q_{t_8}$  auf. Auch die Unterschiede hinsichtlich der Anteile der Schülerinnen und Schüler der Gesamt- und Berufsschulen erweisen sich als gering.

Lediglich die Schülerinnen und Schüler des Gymnasiums sind mit einem Prozentpunkt im Panelquerschnitt leicht überrepräsentiert. Insgesamt muss jedoch festgehalten werden, dass die Auswertung der Schulform spätestens für die Datensätze der Jahre 2011 und 2013 keinen inhaltlichen Wert mehr besitzt, da die weitaus meisten Personen zu diesen Zeitpunkten die Schule bereits verlassen haben. Aus diesem Grund wird in den folgenden Darstellungen auf die Betrachtung der Schulform verzichtet.

### 5.1.6 Panelquerschnitt 2011 ( $P_{Q_{t_9}}$ )

Der Panelquerschnitt 2011 umfasst 2 100 Fälle und enthält damit 68,9 Prozent der Fälle des ursprünglichen Querschnittsdatsatzes.

**Tabelle 5.16:** *Zusammensetzung nach Geschlecht  $P_{Q_{t_9}}$*

	männlich		weiblich		gesamt
	Hfk	%	Hfk	%	(100 %)
$Q_{t_9}$	1 423	46,8	1 620	53,2	3 043
$P_{Q_{t_9}}$	860	41,0	1 240	59,0	2 100

Aus Tabelle 5.16 geht hervor, dass im Panelquerschnitt  $P_{Q_{t_9}}$  eine recht deutliche Verschiebung hinsichtlich des Geschlechts vorliegt: Vergleicht man die Geschlechterverteilung im Panelquerschnitt mit dem Querschnittsdatsatz des entsprechenden Erhebungsjahres ( $Q_{t_9}$ ), so wird deutlich, dass der Anteil der weiblichen Befragten im Panelquerschnitt um etwa 5,8 Prozentpunkte höher liegt.

**Tabelle 5.17:** *Altersdurchschnitt  $P_{Q_{t_9}}$*

	Alter	Std.-abw.	gesamt
$Q_{t_9}$	22,1	0,9	3 021
$P_{Q_{t_9}}$	22,0	0,8	2 089

Ein Vergleich des durchschnittlichen Alters der Befragten in den Querschnittsdaten  $Q_{t_9}$  und dem Panelquerschnitt zeigt, dass sich das Durchschnittsalter der Befragten in beiden Datensätzen nahezu nicht unterscheidet (siehe Tabelle 5.17).

### 5.1.7 Panelquerschnitt 2013 ( $P_{Q_{t_{10}}}$ )

Der Panelquerschnitt 2013 umfasst 1 912 Fälle und enthält damit 67,1 Prozent der Fälle des ursprünglichen Querschnittsdatsatzes.

**Tabelle 5.18:** Zusammensetzung nach Geschlecht  $P_{Q_{t_{10}}}$

	männlich		weiblich		gesamt
	Hfk	%	Hfk	%	(100 %)
$Q_{t_{10}}$	1 302	45,7	1 546	54,3	2 848
$P_{Q_{t_{10}}}$	1 748	39,1	1 164	60,9	1 912

Aus Tabelle 5.18 geht hervor, dass im Panelquerschnitt  $P_{Q_{t_{10}}}$  eine recht deutliche Verschiebung hinsichtlich des Geschlechts vorliegt: Vergleicht man die Geschlechterverteilung im Panelquerschnitt mit dem Querschnittsdatensatz des entsprechenden Erhebungsjahres ( $Q_{t_{10}}$ ), so wird deutlich, dass der Anteil der weiblichen Befragten im Panelquerschnitt um etwa 6,6 Prozentpunkte höher liegt.

**Tabelle 5.19:** Altersdurchschnitt  $P_{Q_{t_{10}}}$

	Alter	Std.-abw.	gesamt
$Q_{t_{10}}$	24,1	1,2	2 839
$P_{Q_{t_{10}}}$	24,0	0,9	1 909

Ein Vergleich des durchschnittlichen Alters der Befragten in den Querschnittsdaten  $Q_{t_{10}}$  und dem Panelquerschnitt zeigt, dass sich das Durchschnittsalter der Befragten in beiden Datensätzen nahezu nicht unterscheidet, jedoch ist eine leicht höhere Standardabweichung als in den vorangegangenen Zeitpunkten zu beobachten (siehe Tabelle 5.19).

## 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). Dabei sind auch jene Datensätze enthalten, die, aufgrund der Änderung des Erhebungsrhythmus, in einem Zweijahresabstand erhoben wurden. Weitere Zwei-Wellen-Panel werden bei (Pöge, 2007) und (Pollich, 2010) beschrieben.

### 5.2.1 Zwei-Wellen-Panel 2007 und 2008 ( $P_{t_{6,7}}$ )

Das Zwei-Wellen-Panel der Jahre 2007 und 2008 enthält 2 299 Fälle. Aus Tabelle 5.20 geht hervor, dass in den Erhebungszeitpunkten  $t_6$  und  $t_7$  eine Verschiebung hinsichtlich des Geschlechts vorliegt:

**Tabelle 5.20:** Zusammensetzung nach Geschlecht  $P_{t_6,7}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_6$	$Q_{t_6}$	1 567	47,0	1 768	53,0	3 335
	$P_{t_6,7}$	975	42,4	1 324	57,6	2 299
$t_7$	$Q_{t_7}$	1 452	47,1	1 634	52,9	3 086
	$P_{t_6,7}$	976	42,5	1 323	57,5	2 299

Der Anteil der weiblichen Befragten im Zwei-Wellen-Paneldatensatz liegt im Vergleich zu den Querschnittsdaten der Erhebungszeitpunkte  $t_6$  und  $t_7$  etwa viereinhalb Prozentpunkte höher. Hinsichtlich des Durchschnittsalters zeigt der Vergleich zwischen dem Zwei-Wellen-Paneldatensatz  $P_{t_6,7}$  und den Querschnittsdatensätzen, dass das Durchschnittsalter der Befragten im Paneldatensatz nur geringfügig von dem in den Querschnittsdaten abweicht (Tabelle 5.21). Im Zeitpunkt  $t_6$  ist die Altershomogenität im Paneldatensatz größer, als im vergleichbaren Querschnittsdatensatz.

**Tabelle 5.21:** Altersdurchschnitt  $P_{t_6,7}$

		Alter	Std.-abw.	gesamt
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_6,7}$	18,0	0,8	2 291
$t_7$	$Q_{t_7}$	19,1	0,8	3 081
	$P_{t_6,7}$	19,0	0,8	2 297

Tabelle 5.22 zeigt, dass beim Zwei-Wellen-Paneldatensatz  $P_{t_6,7}$  im Vergleich zu den Querschnittsdaten Abweichungen bei der Verteilung nach besuchter Schulform vorliegen: Im Erhebungszeitpunkt  $t_6$  liegt der Anteil der Schülerinnen und Schüler an Gesamtschulen einen halben Prozentpunkt unterhalb des Querschnitts. Im folgenden Zeitpunkt  $t_7$  erhöht sich dieser Unterschied auf knapp eineinhalb Prozentpunkte. Die Schülerinnen und Schüler der Gymnasien sind im Unterschied dazu zu beiden Zeitpunkten im Paneldatensatz deutlich überrepräsentiert. So liegt der Anteil im Erhebungszeitpunkt  $t_6$  um gut zwei, im folgenden Zeitpunkt  $t_7$  um knapp vier Prozentpunkte über dem der vergleichbaren Querschnittsdatensätze. Während die Schülerinnen und Schüler der Berufskollegs im Erhebungszeitpunkt  $t_6$  – im Vergleich zu dem entsprechenden Querschnittsdatensatz – mit knapp einem halben Prozentpunkt noch leicht überrepräsentiert sind, wandelt sich dies im folgenden Zeitpunkt  $t_7$  in eine Unterrepräsentanz von einem Prozentpunkt. Der Anteil jener Personen, welche angaben eine andere, als die genannten Schulen zu besuchen, unterscheidet sich im Erhebungszeitpunkt  $t_6$  kaum zwischen dem Zwei-Wellen-Paneldatensatz  $P_{t_6,7}$  und dem entsprechenden Querschnittsdatensatz.

Im folgenden Erhebungszeitpunkt  $t_7$  ist dieser Personenkreis jedoch mit gut einem Prozentpunkt im Querschnittsdatensatz überrepräsentiert. Der Anteil jener Personen, die angaben, keine Schule mehr zu besuchen, ist in dem Querschnittsdatensatz des Erhebungszeitpunkts  $t_6$  mit knapp zweieinhalb und jenem des Erhebungszeitpunkts  $t_7$  mit gut drei Prozentpunkten deutlich überrepräsentiert. Schülerinnen und Schüler der Haupt-, Sonder- und Realschulen sind in den betreffenden Zeitpunkten der Quer- und Paneldatensätze nicht mehr enthalten.

**Tabelle 5.22:** Zusammensetzung  $P_{t_6,7}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							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	3 219
	$P_{t_6,7}$	—	—	17,5	28,4	31,4	—	10,9	11,7	2 240
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2 975
	$P_{t_6,7}$	—	—	15,2	26,7	27,1	—	13,5	17,4	2 237

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

### 5.2.2 Zwei-Wellen-Panel 2008 und 2009 ( $P_{t_7,8}$ )

Das Zwei-Wellen-Panel der Jahre 2008 und 2009 enthält 2 209 Fälle. Aus Tabelle 5.23 geht hervor, dass in den Erhebungszeitpunkten  $t_7$  und  $t_8$  eine Verschiebung hinsichtlich des Geschlechts vorliegt:

**Tabelle 5.23:** Zusammensetzung nach Geschlecht  $P_{t_7,8}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_7$	$Q_{t_7}$	1 452	47,1	1 634	52,9	3 086
	$P_{t_7,8}$	928	42,0	1 281	58,0	2 209
$t_8$	$Q_{t_8}$	1 446	46,8	1 644	53,2	3 090
	$P_{t_7,8}$	928	42,0	1 281	58,0	2 209

Der Anteil der weiblichen Befragten im Zwei-Wellen-Paneldatensatz  $P_{t_7,8}$  liegt, im Vergleich zu den Querschnittsdaten der Erhebungszeitpunkte  $t_7$  und  $t_8$ , etwa fünf Prozentpunkte höher. Hinsichtlich des Durchschnittsalters zeigt der Vergleich zwischen dem Zwei-Wellen-Paneldatensatz und den Querschnittsdatensätzen, dass das Durchschnittsalter der Befragten im Paneldatensatz nur geringfügig von dem in den Querschnittsdaten

abweicht (Tabelle 5.24). Der Zwei-Wellen-Paneldatensatz  $P_{t_{7,8}}$  und die entsprechenden Querschnitte weisen zu allen Erhebungszeitpunkten zudem eine gleichförmige Altershomogenität auf.

**Tabelle 5.24:** Altersdurchschnitt  $P_{t_{7,8}}$

		Alter	Std.-abw.	gesamt
$t_7$	$Q_{t_7}$	19,1	0,8	3081
	$P_{t_{7,8}}$	19,0	0,8	2208
$t_8$	$Q_{t_8}$	20,1	0,8	3073
	$P_{t_{7,8}}$	20,0	0,8	2202

**Tabelle 5.25:** Zusammensetzung  $P_{t_{7,8}}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								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
	$P_{t_{7,8}}$	—	—	15,4	27,2	26,8	—	13,4	17,3	2155
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{7,8}}$	—	—	4,4	3,8	22,6	—	18,7	50,5	2115

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

Tabelle 5.25 zeigt, dass der Zwei-Wellen-Paneldatensatz  $P_{t_{7,8}}$  nach besuchter Schulform leicht verzerrt ist: Im Erhebungszeitpunkt  $t_7$  liegt der Anteil der Schülerinnen und Schüler an Gesamtschulen etwa eineinhalb Prozentpunkte über jenem des Querschnitts. Im folgenden Zeitpunkt  $t_8$  nivelliert sich dieser Unterschied nahezu auf knapp einen halben Prozentpunkt. Der Anteil der Schülerinnen und Schüler ist zu diesem Zeitpunkt jedoch deutlich gesunken. Die Schülerinnen und Schüler der Gymnasien sind zu beiden Zeitpunkten im Paneldatensatz überrepräsentiert. So liegt der Anteil im Erhebungszeitpunkt  $t_7$  um etwa viereinhalb, im folgenden Zeitpunkt  $t_8$  um knapp einen halben Prozentpunkt über dem der vergleichbaren Querschnittsdatensätze. Im Erhebungszeitpunkt  $t_8$  besuchen jedoch nur noch wenige Personen das Gymnasium. Die Schülerinnen und Schüler der Berufskollegs sind im Erhebungszeitpunkt  $t_7$  – im Vergleich zu dem entsprechenden Querschnittsdatensatz – mit etwa einem Prozentpunkt unterrepräsentiert. Im folgenden Zeitpunkt  $t_8$  besteht zwischen beiden Datensätzen kaum noch ein Unterschied. Der Anteil jener Personen, welche angaben, eine andere als die genannten Schulen zu besuchen, ist im Zwei-Wellen-Paneldatensatz  $P_{t_{7,8}}$  zum Erhebungszeitpunkt  $t_7$  etwa einen Prozentpunkt höher, als im entsprechenden Querschnittsdatensatz. Im folgenden

Erhebungszeitpunkt  $t_8$  bestehen kaum noch Unterschiede zwischen den Datensätzen. Der Anteil jener Personen, die angaben, keine Schule mehr zu besuchen, ist im Zwei-Wellen-Paneldatensatz  $P_{t_7,8}$  zum Erhebungszeitpunkt  $t_7$  – im Vergleich mit dem entsprechenden Querschnittsdatsatz – mit gut drei Prozentpunkten deutlich überrepräsentiert. Im Erhebungszeitpunkt  $t_8$  ist kaum noch ein Unterschied vorhanden. Schülerinnen und Schüler der Haupt-, Sonder- und Realschulen sind in den betreffenden Zeitpunkten der Quer- und Paneldatensätze nicht mehr enthalten.

### 5.2.3 Zwei-Wellen-Panel 2009 und 2011 ( $P_{t_8,9}$ )

Das Zwei-Wellen-Panel der Jahre 2009 und 2011 enthält 1 997 Fälle. Es ist der erste Paneldatensatz ohne Erhebungslücke, der einen Zweijahresabstand aufweist. Aus Tabelle 5.26 geht hervor, dass auch in den Erhebungszeitpunkten  $t_8$  und  $t_9$  eine Verschiebung hinsichtlich des Geschlechts vorliegt:

**Tabelle 5.26:** Zusammensetzung nach Geschlecht  $P_{t_8,9}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_8$	$Q_{t_8}$	1 446	46,8	1 644	53,2	3 090
	$P_{t_8,9}$	792	39,7	1 205	60,3	1 997
$t_9$	$Q_{t_9}$	1 423	46,8	1 620	53,2	3 043
	$P_{t_8,9}$	796	39,9	1 201	60,1	1 997

Der Anteil der weiblichen Befragten im Zwei-Wellen-Paneldatensatz  $P_{t_8,9}$  liegt, im Vergleich zu den Querschnittsdaten der Erhebungszeitpunkte  $t_8$  und  $t_9$ , etwa sieben Prozentpunkte höher. Hinsichtlich des Durchschnittsalters zeigt der Vergleich zwischen dem Zwei-Wellen-Paneldatensatz  $P_{t_8,9}$  und den Querschnittsdatsätzen, dass das Durchschnittsalter der Befragten im Paneldatensatz nur geringfügig von dem in den Querschnittsdaten abweicht (Tabelle 5.27). In beiden betrachteten Erhebungszeitpunkten ist die Altershomogenität im Zwei-Wellen-Paneldatensatz  $P_{t_8,9}$  größer, als in den vergleichbaren Querschnittsdatsätzen.

**Tabelle 5.27:** Altersdurchschnitt  $P_{t_8,9}$

		Alter	Std.-abw.	gesamt
$t_8$	$Q_{t_8}$	20,1	0,8	3 073
	$P_{t_8,9}$	20,0	0,7	1 991
$t_9$	$Q_{t_9}$	22,1	0,9	3 021
	$P_{t_8,9}$	22,0	0,8	1 988

**Tabelle 5.28:** Zusammensetzung  $P_{8,9}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							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
	$P_{8,9}$	—	—	4,0	4,1	22,4	—	18,8	50,7	1923

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

Die Informationen bezüglich der Schulform liegen lediglich für den Erhebungszeitpunkt  $t_8$  vor. Diesbezüglich geht aus Tabelle 5.28 hervor, dass der Zwei-Wellen-Paneldatensatz  $P_{8,9}$  zum Zeitpunkt  $t_8$  nur noch leicht nach besuchter Schulform verzerrt ist: So bestehen für Gesamtschülerinnen und -schüler keine nennenswerten Unterschiede zwischen den Datensätzen. Schülerinnen und Schüler, die noch das Gymnasium besuchen, sind mit etwa einem Prozentpunkt im Zwei-Wellen-Paneldatensatz überrepräsentiert. Berufsschülerinnen und -schüler sind hingegen – im Vergleich zum entsprechenden Querschnittsdatsatz – mit einem halben Prozentpunkt im Zwei-Wellen-Paneldatensatz  $P_{8,9}$  leicht unterrepräsentiert. Der Anteil jener Personen, welche angaben eine andere, als die genannten Schulen zu besuchen, ist im Zwei-Wellen-Paneldatensatz zum Erhebungszeitpunkt  $t_8$  etwa einen halben Prozentpunkt höher, als im entsprechenden Querschnittsdatsatz. Der Anteil jener Personen, die angaben keine Schule mehr zu besuchen, ist im Zwei-Wellen-Paneldatensatz  $P_{8,9}$  zum Erhebungszeitpunkt  $t_8$  – im Vergleich mit dem entsprechenden Querschnittsdatsatz – mit gut einem halben Prozentpunkt unterrepräsentiert. Schülerinnen und Schüler der Haupt-, Sonder- und Realschulen sind in den betreffenden Zeitpunkten der Quer- und Paneldatensätze nicht mehr enthalten.

#### 5.2.4 Zwei-Wellen-Panel 2011 und 2013 ( $P_{t_9,t_{10}}$ )

Das Zwei-Wellen-Panel der Jahre 2011 und 2013 enthält 1 813 Fälle. Aus Tabelle 5.29 geht hervor, dass in den Erhebungszeitpunkten  $t_9$  und  $t_{10}$  eine Verschiebung hinsichtlich des Geschlechts vorliegt:

Der Anteil der weiblichen Befragten im Zwei-Wellen-Paneldatensatz  $P_{t_9,t_{10}}$  liegt, im Vergleich zu den Querschnittsdaten der Erhebungszeitpunkte  $t_9$  und  $t_{10}$ , etwa sieben-einhalb Prozentpunkte höher. Hinsichtlich des Durchschnittsalters zeigt der Vergleich zwischen dem Zwei-Wellen-Paneldatensatz und den Querschnittsdatsätzen, dass das Durchschnittsalter der Befragten im Zwei-Wellen-Paneldatensatz nur geringfügig von dem in den Querschnittsdaten abweicht (Tabelle 5.30). In beiden betrachteten Erhebungszeitpunkten ist die Altershomogenität im Zwei-Wellen-Paneldatensatz größer, als in den vergleichbaren Querschnittsdatsätzen.

**Tabelle 5.29:** Zusammensetzung nach Geschlecht  $P_{t_9, t_{10}}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_9$	$Q_{t_9}$	1423	46,8	1620	53,2	3043
	$P_{t_9, t_{10}}$	697	38,4	1116	61,6	1813
$t_{10}$	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_9, t_{10}}$	692	38,2	1121	61,8	1813

**Tabelle 5.30:** Altersdurchschnitt  $P_{t_9, t_{10}}$

		Alter	Std.-abw.	gesamt
$t_9$	$Q_{t_9}$	22,1	0,9	3021
	$P_{t_9, t_{10}}$	22,0	0,8	1805
$t_{10}$	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_9, t_{10}}$	24,0	0,9	1810

Für die Zeitpunkte  $t_9$  und  $t_{10}$  erübrigt sich eine Auswertung hinsichtlich der Schulform, da nur noch ein Bruchteil der Befragten zu diesen Zeitpunkten eine schulische Bildungseinrichtung besucht.

### 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). Dabei sind auch jene Datensätze enthalten, die aufgrund der Änderung des Erhebungsrhythmus in einem Zweijahresabstand erhoben wurden.

#### 5.3.1 Drei-Wellen-Panel 2006, 2007 und 2008 ( $P_{t_5,6,7}$ )

Das Drei-Wellen-Panel der Jahre 2006 bis 2008 enthält 2 163 Fälle. Aus Tabelle 5.31 geht hervor, dass in den Erhebungszeitpunkten  $t_5$  bis  $t_7$  eine Verschiebung hinsichtlich des Geschlechts vorliegt:

**Tabelle 5.31:** Zusammensetzung nach Geschlecht  $P_{t_5,6,7}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_5,6,7}$	901	41,7	1262	58,3	2163
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{t_5,6,7}$	902	41,7	1261	58,3	2163
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_5,6,7}$	902	41,7	1261	58,3	2163

Der Anteil der weiblichen Befragten im Drei-Wellen-Paneldatensatz  $P_{t_5,6,7}$  liegt, im Vergleich zu den Querschnittsdaten des Erhebungszeitpunkts  $t_5$ , etwa acht Prozentpunkte höher. Dieser Unterschied ist mit der deutlichen Erhöhung der Fallzahl im Erhebungszeitpunkt  $t_5$  zu erklären, der in dieser Hinsicht eine Ausnahme darstellt (siehe Abschnitt 2). In den Erhebungszeitpunkten  $t_6$  und  $t_8$  sind die weiblichen Befragten mit etwa fünf Prozentpunkten im Drei-Wellen-Paneldatensatz überrepräsentiert. Hinsichtlich des Durchschnittsalters zeigt der Vergleich zwischen dem Drei-Wellen-Paneldatensatz  $P_{t_5,6,7}$  und den Querschnittsdatensätzen, dass das Durchschnittsalter der Befragten im Paneldatensatz zu den Zeitpunkten  $t_6$  und  $t_7$  nur geringfügig von dem in den entsprechenden Querschnittsdaten abweicht. Im Erhebungszeitpunkt  $t_5$  zeigt sich ein etwas höheres Durchschnittsalter der Befragten im entsprechenden Querschnittsdatensatz (Tabelle 5.35). In den Erhebungszeitpunkten  $t_5$  und  $t_6$  ist die Altershomogenität im Drei-Wellen-Paneldatensatz  $P_{t_5,6,7}$  größer, als in den vergleichbaren Querschnittsdatensätzen. Im letzten Erhebungszeitpunkt besteht hingegen kein Unterschied.

**Tabelle 5.32:** Altersdurchschnitt  $P_{t_5,6,7}$ 

		Alter	Std.-abw.	gesamt
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_5,6,7}$	17,0	0,8	2146
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{t_5,6,7}$	18,0	0,8	2155
$t_7$	$Q_{t_7}$	19,1	0,8	3081
	$P_{t_5,6,7}$	19,0	0,8	2161

Tabelle 5.33 zeigt, dass der Drei-Wellen-Paneldatensatz  $P_{t_5,6,7}$  nach besuchter Schulform verzerrt ist: Im Erhebungszeitpunkt  $t_5$  liegt der Anteil der Schülerinnen und Schüler an Gesamtschulen sechs Prozentpunkte über jenem des Querschnitts. Im folgenden Zeitpunkt  $t_6$  nivelliert sich dieser Unterschied nahezu. Im letzten betrachteten Erhebungszeitpunkt  $t_7$  sind die Gesamtschülerinnen und -schüler mit knapp zwei Prozentpunkten im Drei-Wellen-Paneldatensatz  $P_{t_5,6,7}$  überrepräsentiert. Die Schülerinnen und Schüler der Gymnasien sind zum Erhebungszeitpunkt  $t_5$  mit gut 13 Prozentpunkten Differenz zum Querschnitt im Paneldatensatz deutlich überrepräsentiert. Zwar wird dieser Anteil in den folgenden beiden Zeitpunkten geringer, jedoch bleibt die Überrepräsentanz mit etwa drei und fünf Prozentpunkten bestehen. Die Schülerinnen und Schüler der Berufskollegs sind im Erhebungszeitpunkt  $t_5$  – im Vergleich zu dem entsprechenden Querschnittsdatensatz – mit knapp 20 Prozentpunkten unterrepräsentiert. Auch dies ist auf die bereits erwähnte temporäre Vergrößerung der Stichprobe mit Berufsschülern zurückzuführen. Im folgenden Zeitpunkt  $t_6$  besteht zwischen beiden Datensätzen kaum noch ein Unterschied. Im Zeitpunkt  $t_7$  zeigt sich mit etwa eineinhalb Prozentpunkten eine leichte Überrepräsentanz im Querschnittsdatensatz. In den Zeitpunkten  $t_6$  und  $t_7$  sind jene Personen, die angaben, eine andere oder keine Schule mehr zu besuchen, im Vergleich mit den entsprechenden Querschnittsdatensätzen mit knapp zwei und vier Prozentpunkten leicht unterrepräsentiert. Schülerinnen und Schüler der Haupt-, Sonder- und Realschulen sind lediglich im Erhebungszeitpunkt  $t_5$  vorhanden, doch sind die Anteile so gering, dass ein Vergleich wenig sinnvoll erscheint.

**Tabelle 5.33:** Zusammensetzung  $P_{t_5,6,7}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_5,6,7}$	0,5	0,6	22,2	33,5	43,1	0,1	—	—	1996
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_5,6,7}$	—	—	18,1	29,5	31,2	—	10,9	10,3	2113
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_5,6,7}$	—	—	15,6	27,8	26,6	—	13,2	16,9	2109

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

### 5.3.2 Drei-Wellen-Panel 2007, 2008 und 2009 ( $P_{t_6,7,8}$ )

Das Drei-Wellen-Panel der Jahre 2007 bis 2009 enthält 2 129 Fälle. Aus Tabelle 5.34 geht hervor, dass in den Erhebungszeitpunkten  $t_6$  bis  $t_8$  eine Verschiebung hinsichtlich des Geschlechts vorliegt:

**Tabelle 5.34:** Zusammensetzung nach Geschlecht  $P_{t_6,7,8}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{t_6,7,8}$	876	41,2	1253	58,8	2129
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_6,7,8}$	876	41,2	1253	58,8	2129
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{t_6,7,8}$	876	41,2	1253	58,8	2129

Der Anteil der weiblichen Befragten im Drei-Wellen-Paneldatensatz  $P_{t_6,7,8}$  liegt, im Vergleich zu den Querschnittsdaten der Erhebungszeitpunkte  $t_6$  und  $t_7$ , etwa sechs Prozentpunkte höher. Im Erhebungszeitpunkt  $t_8$  sind die weiblichen Befragten mit etwa fünfeinhalb Prozentpunkten im Drei-Wellen-Paneldatensatz überrepräsentiert. Hinsichtlich des Durchschnittsalters zeigt der Vergleich zwischen dem Drei-Wellen-Paneldatensatz  $P_{t_6,7,8}$  und den Querschnittsdatensätzen, dass das Durchschnittsalter der Befragten im Paneldatensatz zu den Zeitpunkten  $t_6$  und  $t_8$  nur geringfügig von dem in den entsprechenden Querschnittsdaten abweicht. Im Erhebungszeitpunkt  $t_7$  zeigt sich ein etwas höheres Durchschnittsalter der Befragten im entsprechenden Querschnittsdatensatz (Tabel-

le 5.35). In den Erhebungszeitpunkten  $t_7$  und  $t_8$  besteht kein Unterschied hinsichtlich der Altershomogenität zwischen den Datensätzen. Im letzten Erhebungszeitpunkt besteht hingegen kein Unterschied. Im Erhebungszeitpunkt  $t_6$  zeigt sich eine etwas größere Altershomogenität im Drei-Wellen-Paneldatensatz  $P_{t_6,7,8}$  als in dem vergleichbaren Querschnittsdatsatz.

**Tabelle 5.35:** Altersdurchschnitt  $P_{t_6,7,8}$

		Alter	Std.-abw.	gesamt
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_6,7,8}$	18,0	0,8	2 123
$t_7$	$Q_{t_7}$	19,1	0,8	3 081
	$P_{t_6,7,8}$	19,0	0,8	2 128
$t_8$	$Q_{t_8}$	20,1	0,8	3 073
	$P_{t_6,7,8}$	20,0	0,8	2 122

**Tabelle 5.36:** Zusammensetzung  $P_{t_6,7,8}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							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	3 219
	$P_{t_6,7,8}$	—	—	17,7	29,3	31,3	—	10,4	11,41	2 077
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2 975
	$P_{t_6,7,8}$	—	—	15,7	27,5	26,9	—	13,2	16,8	2 077
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2 934
	$P_{t_6,7,8}$	—	—	4,4	3,7	22,6	—	18,7	50,7	2 042

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

Tabelle 5.36 zeigt, dass der Drei-Wellen-Paneldatensatz  $P_{t_6,7,8}$  bezüglich des Merkmals Schulform Abweichungen zu den vergleichbaren Querschnittsdatsätzen aufweist: Im Erhebungszeitpunkt  $t_6$  liegt der Anteil der Schülerinnen und Schüler an Gesamtschulen leicht unter jenem des Querschnitts. Im folgenden Zeitpunkt  $t_7$  dreht sich dies um, sodass nun die Schülerinnen und Schüler der Gesamtschule im Paneldatensatz mit etwa drei Prozentpunkten überrepräsentiert sind. Im letzten betrachteten Erhebungszeitpunkt  $t_8$  liegt der Anteil der Gesamtschülerinnen und -schüler im Drei-Wellen-Paneldatensatz  $P_{t_6,7,8}$  nur noch leicht über jenem in dem entsprechenden Querschnittsdatsatz. Die Schülerinnen und Schüler der Gymnasien sind zum Erhebungszeitpunkt  $t_6$  mit etwa drei Prozentpunkten im Paneldatensatz überrepräsentiert. Im folgenden Zeitpunkt sind knapp fünf Prozent

mehr Schülerinnen und Schüler des Gymnasiums im Drei-Wellen-Paneldatensatz  $P_{t_{6,7,8}}$  vertreten. Im letzten betrachteten Zeitpunkt besuchen kaum noch Personen das Gymnasium, es verbleibt jedoch eine leichte Überrepräsentanz im Paneldatensatz. Zwischen Schülerinnen und Schüler der Berufskollegs sind im Erhebungszeitpunkt  $t_6$  – im Vergleich zu dem entsprechenden Querschnittsdatsatz – kaum Unterschiede zum Paneldatensatz festzustellen. Im folgenden Zeitpunkt  $t_7$  besteht zwischen beiden Datensätzen eine leichte Verschiebung von etwa einem Prozentpunkt in Richtung des Paneldatensatzes  $P_{t_{6,7,8}}$ . Zum Zeitpunkt  $t_8$  besteht kaum noch ein Unterschied zwischen den beiden Datensätzen. Im Erhebungszeitpunkt  $t_6$  bestehen zwischen dem Drei-Wellen-Paneldatensatz  $P_{t_{6,7,8}}$  und dem entsprechenden Querschnittsdatsatz kaum Differenzen hinsichtlich jener Personen, welche angaben, eine andere als die genannten Schulen zu besuchen. Im Erhebungszeitpunkt  $t_7$  besteht mit etwa zwei Prozentpunkten eine leichte Unterrepräsentanz im Paneldatensatz. Im letzten betrachteten Zeitpunkt ( $t_8$ ) bestehen wiederum kaum Unterschiede zwischen den Datensätzen. Der Anteil jener Personen, die angaben keine Schule mehr zu besuchen, ist im Drei-Wellen-Paneldatensatz  $P_{t_{6,7,8}}$  zum Erhebungszeitpunkt  $t_5$  – im Vergleich zu dem entsprechenden Querschnittsdatsatz – um etwa drei Prozentpunkte niedriger. Im Zeitpunkt  $t_7$  erhöht sich diese Differenz auf knapp vier Prozentpunkte. Im letzten betrachteten Zeitpunkt  $t_8$  liegt Anteil im Paneldatensatz noch einen halben Prozentpunkt niedriger, als im vergleichbaren Querschnittsdatsatz. Schülerinnen und Schüler der Haupt-, Sonder- und Realschulen sind in den betreffenden Zeitpunkten der betrachteten Quer- und Paneldatensätze nicht mehr enthalten.

### 5.3.3 Drei-Wellen-Panel 2008, 2009 und 2011 ( $P_{t_{7,8,9}}$ )

Das Drei-Wellen-Panel der Jahre 2008 bis 2011 enthält 1 920 Fälle. In Tabelle 5.37 wird ersichtlich, dass in den Erhebungszeitpunkten  $t_7$  bis  $t_9$  eine Verschiebung bezüglich des Geschlechts vorliegt.

**Tabelle 5.37:** Zusammensetzung nach Geschlecht  $P_{t_{7,8,9}}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_7$	$Q_{t_7}$	1 452	47,1	1 634	52,9	3 086
	$P_{t_{7,8,9}}$	759	39,5	1 161	60,5	1 920
$t_8$	$Q_{t_8}$	1 446	46,8	1 644	53,2	3 090
	$P_{t_{7,8,9}}$	759	39,5	1 161	60,5	1 920
$t_9$	$Q_{t_9}$	1 423	46,8	1 620	53,2	3 043
	$P_{t_{7,8,9}}$	761	39,6	1 159	60,4	1 920

Weibliche Befragte sind im Drei-Wellen-Paneldatensatz  $P_{t_7,8,9}$  leicht überrepräsentiert. Der Prozentanteil weiblicher Befragter liegt im Vergleich zu den Querschnitten der jeweiligen Erhebungszeitpunkte  $t_7$ ,  $t_8$  und  $t_9$  jeweils etwa sieben Prozentpunkte höher. Entsprechend fällt umgekehrt der Anteil männlicher Befragter im Panel geringer aus als in den jeweiligen Querschnitten.

Bezüglich des Durchschnittsalters zeigen sich nur sehr geringe Diskrepanzen zwischen dem Drei-Wellen-Panel  $P_{t_7,8,9}$  und den Querschnittsdatsätzen der jeweiligen Zeitpunkte. Die Standardabweichung zeigt an, dass die verglichenen Datensätze und Zeitpunkte keine wesentlichen Unterschiede bezüglich der Altershomogenität aufweisen.

**Tabelle 5.38:** Altersdurchschnitt  $P_{t_7,8,9}$

		Alter	Std.-abw.	gesamt
$t_7$	$Q_{t_7}$	19,1	0,8	3081
	$P_{t_7,8,9}$	19,0	0,8	1920
$t_8$	$Q_{t_8}$	20,1	0,8	3073
	$P_{t_7,8,9}$	20,0	0,7	1914
$t_9$	$Q_{t_9}$	22,1	0,9	3021
	$P_{t_7,8,9}$	22,0	0,8	1913

Mit Blick auf die Zusammensetzung nach Schulform, welche in Tabelle 5.39 dargestellt ist, lässt sich bezüglich dieses Merkmals eine leichte Verzerrung des Drei-Wellen-Panels  $P_{t_7,8,9}$  erkennen. Zum Zeitpunkt  $t_7$  liegt der Anteil an Gesamtschülerinnen und -schülern im Paneldatensatz rund zwei Prozentpunkte höher als im Querschnitt. Im darauffolgenden Zeitpunkt unterscheidet sich dieser Anteil jedoch kaum mehr.

Eine deutlichere Überrepräsentierung zeigt sich jedoch beim Gymnasium: Für das Jahr 2008 ( $t_7$ ) liegt der Anteil dieser Schulform im Drei-Wellen-Panel gut sechs Prozent höher als im Querschnitt. Zum Zeitpunkt  $t_8$  liegt ebenso eine derart gerichtete Differenz vor, sie fällt jedoch mit 0,6% deutlich geringer aus.

Bei den Schülerinnen und Schülern an Berufskollegs zeigt sich für das Panel zum Zeitpunkt  $t_7$  im Vergleich zum Querschnitt dieses Jahres eine leichte Unterrepräsentanz, die jedoch zum darauffolgenden Zeitpunkt  $t_8$  kaum mehr besteht.

Darüber hinaus ist erkennbar, dass der Anteil an Personen, die keine Schule mehr besuchen zwischen  $t_7$  und  $t_8$  deutlich zunimmt (von ca. 21% auf rund 51% im Querschnitt und von rund 17% auf etwa 50% im Drei-Wellen-Panel). Zum Zeitpunkt  $t_9$  besuchen kaum mehr Befragte eine schulische Bildungseinrichtung, sodass zu diesem Zeitpunkt ein Vergleich wenig sinnvoll ist und demnach in der vergleichenden Darstellung darauf verzichtet wird.

**Tabelle 5.39:** Zusammensetzung  $P_{t_7,8,9}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							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
	$P_{t_7,8,9}$	—	—	15,4	29,1	26,3	—	12,7	16,5	1877
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_7,8,9}$	—	—	4,0	4,0	22,6	—	19,1	50,3	1848

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

### 5.3.4 Drei-Wellen-Panel 2009, 2011 und 2013 ( $P_{t_8,9,10}$ )

Die Fallzahl des Drei-Wellen-Panels  $P_{t_8,9,10}$  beträgt 1 746. Die in Tabelle 5.40 dargestellten Prozentzahlen lassen den Schluss zu, dass verglichen mit den Querschnittsdaten im Paneldatensatz eine Verzerrung bezüglich des Merkmals Geschlecht vorliegt. Zu allen drei betrachteten Zeitpunkten  $t_8$ ,  $t_9$  und  $t_{10}$  liegt der Anteil weiblicher Befragter im Drei-Wellen-Panel acht bis neun Prozentpunkte höher als im Querschnittsdatensatz desselben Jahres.

**Tabelle 5.40:** Zusammensetzung nach Geschlecht  $P_{t_8,9,10}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{t_8,9,10}$	656	37,6	1090	62,4	1746
$t_9$	$Q_{t_9}$	1423	46,8	1620	53,2	3043
	$P_{t_8,9,10}$	662	37,9	1084	62,1	1746
$t_{10}$	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_8,9,10}$	655	37,5	1091	62,5	1746

Mit Blick auf das Durchschnittsalter in Tabelle 5.41 zeigen sich nur sehr geringfügige Abweichungen zwischen Drei-Wellen-Panel und Querschnittsdaten und das Durchschnittsalter entspricht den zu erwartenden Werten. Auffällig ist, dass im Querschnittsdatensatz zum Zeitpunkt  $t_{10}$  die Standardabweichung etwas höher liegt als in den Vergleichsdaten. Dies weist auf eine größere Altersheterogenität im Querschnittsdatensatz des Erhebungsjahres 2013 hin.

Beim Vergleich nach Schulform zeigen sich geringfügige Abweichungen zwischen dem Drei-Wellen-Panel und dem Querschnittsdatensatz zum Zeitpunkt  $t_8$ . Der Anteil an Gesamtschülerinnen und -schülern im Panel  $P_{t_8,9,10}$  ist marginal geringer als im

**Tabelle 5.41:** Altersdurchschnitt  $P_{8,9,10}$

		Alter	Std.-abw.	gesamt
$t_8$	$Q_{t_8}$	20,1	0,8	3073
	$P_{8,9,10}$	20,0	0,7	1744
$t_9$	$Q_{t_9}$	22,1	0,9	3021
	$P_{8,9,10}$	22,0	0,8	1740
$t_{10}$	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{8,9,10}$	24,0	0,8	1743

Querschnitt. Schülerinnen und Schülern, die ein Gymnasium besuchen, sind hingegen im Drei-Wellen-Panel beim Vergleich mit dem Querschnitt leicht überrepräsentiert - der Anteil dieser Schulform liegt im Panel rund ein Prozent höher.

Da ab dem neunten Erhebungszeit – das durchschnittliche Alter der Befragten liegt hier bei 22 Jahren – kaum Befragte mehr eine Schule besuchen, erübrigt sich an dieser Stelle ein Vergleich der Schulformen für diese Zeitpunkte.

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

		Schulform <sup>a</sup> (%)							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
	$P_{8,9,10}$	—	—	3,9	4,2	22,2	—	18,8	50,8	1684

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

## 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 nur jene Datensätze beschrieben, die Teilnahmen in vier direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke). Ebenso sind Datensätze enthalten, die aufgrund der Änderung des Erhebungsrhythmus in einem Abstand von zwei Jahren erhoben wurden.

### 5.4.1 Vier-Wellen-Panel 2005, 2006, 2007 und 2008 ( $P_{t_{4,5,6,7}}$ )

Das Vier-Wellen-Panel  $P_{t_{4,5,6,7}}$  der Jahre 2005 bis 2008 enthält insgesamt 1 951 Fälle. Betrachtet man die Prozentuierung des Merkmals Geschlecht (Tabelle 5.43) so wird

deutlich, dass bezüglich dieses Merkmals eine Verschiebung vorliegt. Zu den Zeitpunkten  $t_4$  und  $t_5$  liegt der Anteil weiblicher Befragter im Vier-Wellen-Panel jeweils rund neun Prozentpunkte höher als im Querschnitt desselben Jahres. In den späteren Zeitpunkten  $t_6$  und  $t_7$  liegt eine derart gerichtete Diskrepanz ebenso vor, sie fällt mit knapp sechs Prozent jedoch geringer aus. Entsprechend verhält es sich bei den männlichen Befragten genau umgekehrt: Im Paneldatensatz sind männliche Personen im Vergleich zum Querschnitt im selben Maße unterrepräsentiert.

**Tabelle 5.43:** Zusammensetzung nach Geschlecht  $P_{t_4,5,6,7}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_4$	$Q_{t_4}$	1 717	50,4	1 688	49,6	3 405
	$P_{t_4,5,6,7}$	811	41,6	1 140	58,4	1 951
$t_5$	$Q_{t_5}$	2 265	49,8	2 282	50,2	4 547
	$P_{t_4,5,6,7}$	807	41,4	1 144	58,6	1 951
$t_6$	$Q_{t_6}$	1 567	47,0	1 768	53,0	3 335
	$P_{t_4,5,6,7}$	806	41,3	1 145	58,7	1 951
$t_7$	$Q_{t_7}$	1 452	47,1	1 634	52,9	3 086
	$P_{t_4,5,6,7}$	810	41,5	1 141	58,5	1 951

Das in Tabelle 5.44 dargestellte Alter zeigt zu allen vier Zeitpunkten sowohl im Vier-Wellen-Panel als auch im Querschnitt die zu erwartenden Durchschnittswerte. Differenzen zwischen dem Panel  $P_{t_4,5,6,7}$  und den Querschnitten fallen nur sehr gering aus. Lediglich zum Zeitpunkt  $t_5$  liegt das durchschnittliche Alter im Querschnitt 0,4 Jahre höher als im Vier-Wellen-Panel. Dabei weisen die Querschnittsdaten zu diesem Zeitpunkt  $t_5$  – ebenso wie in etwas geringerem Maße auch zum Zeitpunkt  $t_6$  – eine etwas größere Streuung auf. In den Querschnittsdaten der Jahre 2005 und 2006 ist das Alter also etwas weniger homogen verteilt als in den Vergleichsdaten.

Auch bei der Zusammensetzung nach Schulform, welche in Tabelle 5.45 dargestellt ist, lassen sich Verschiebungen erkennen. Zum Zeitpunkt  $t_4$  ist der Anteil an Personen, die eine Hauptschule besuchen, im Panel  $P_{t_4,5,6,7}$  etwa fünf Prozentpunkte geringer als im Querschnittsdatensatz dieses Zeitpunktes. Im Vergleich dazu sind in  $t_4$  Schülerinnen und Schüler eines Gymnasiums im Panel mit einer Differenz von 6,5 Prozent etwas stärker vertreten als im Querschnitt. Bei Real- und Gesamtschulen zeigen sich ebenso Differenzen, die jedoch mit je rund einem Prozent nur sehr gering ausfallen.

Zum Zeitpunkt  $t_5$  – also mit durchschnittlich 17 Jahren – nimmt der generelle Anteil an Hauptschülerinnen und -schülern zwischen  $t_4$  und  $t_5$  erwartungsgemäß stark ab, da jene Personen zu diesem Zeitpunkt größtenteils die Schule beendet haben. Entsprechend nimmt der Anteil an Personen, die ein Berufskolleg besuchen, stark zu. Des Weiteren

**Tabelle 5.44:** Altersdurchschnitt  $P_{t_{4,5,6,7}}$

		Alter	Std.-abw.	gesamt
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_{4,5,6,7}}$	16,0	0,7	1 931
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_{4,5,6,7}}$	17,0	0,8	1 941
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_{4,5,6,7}}$	18,0	0,8	1 945
$t_7$	$Q_{t_7}$	19,1	0,8	3 081
	$P_{t_{4,5,6,7}}$	19,0	0,8	1 949

zeigen sich auch zum Zeitpunkt  $t_5$  Diskrepanzen beim Vergleich von Querschnitts- und Paneldatensatz: Befragte, die eine Gesamtschule besuchen, sind im Panel  $P_{t_{4,5,6,7}}$  im Vergleich zum Querschnitt  $Q_{t_5}$  knapp fünf Prozent, Gymnasiastinnen und Gymnasiasten sogar fast 14 Prozentpunkte überrepräsentiert. Umgekehrt liegt der Anteil der Schulform Berufskolleg im Vier-Wellen-Panel 19 Prozent niedriger als im Querschnitt des Jahres 2006.

In den späteren Zeitpunkten  $t_6$  und  $t_7$  bleibt die Differenz, dass Schülerinnen und Schüler von Gymnasien im Panel stärker vertreten sind als im Querschnitt, bestehen, sie fällt jedoch mit etwa drei bzw. fünf Prozent geringer aus. Beim Berufskolleg zeigen sich in  $t_6$  und  $t_7$  zwischen Querschnitt und Panel nur noch sehr geringe Differenzen. Allerdings fällt auf, dass Personen, die keine Schule (mehr) besuchen, in diesen beiden Zeitpunkten im Panel weniger stark vertreten sind als im Querschnitt. Dies ist möglicherweise darauf zurückzuführen, dass in diese Kategorie vor allem jene Personen fallen, die zuvor an Haupt- und Realschulen waren und die Schule nun beendet haben – insbesondere Personen an Hauptschulen sind generell, wie den Beschreibungen der anderen Paneldatensätze zu entnehmen ist, im Panel meist unterrepräsentiert.

**Tabelle 5.45:** Zusammensetzung  $P_{t_4,5,6,7}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								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
	$P_{t_4,5,6,7}$	16,3	22,7	31,3	28,5	0,7	0,3	0,3	0,1	1 946
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4 253
	$P_{t_4,5,6,7}$	0,6	0,6	21,1	34,0	43,7	0,1	—	—	1 791
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3 219
	$P_{t_4,5,6,7}$	—	—	17,5	29,4	31,8	—	10,8	10,6	1 903
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2 975
	$P_{t_4,5,6,7}$	—	—	15,1	27,6	27,2	—	13,7	16,5	1 907

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

#### 5.4.2 Vier-Wellen-Panel 2006, 2007, 2008 und 2009 ( $P_{t_5,6,7,8}$ )

Das Vier-Wellen-Panel der Zeitpunkte  $t_5$ ,  $t_6$ ,  $t_7$  und  $t_8$  beinhaltet insgesamt 2 017 Fälle. Nachstehende Tabelle 5.46 zeigt die Zusammensetzung nach Geschlecht für die Vier-Wellen-Panel- und Querschnittsdatensätze zu den vier betrachteten Zeitpunkten. Es wird deutlich, dass insbesondere im Panel, in kleinerem Ausmaß aber auch im Querschnitt, eine Verschiebung bezüglich des Geschlechtes vorliegt. Im Panel  $P_{t_5,6,7,8}$  liegt der Anteil weiblicher Befragter innerhalb aller Zeitpunkte mindestens sechs Prozentpunkte höher als in den Querschnitten desselben Jahres. Zum Zeitpunkt  $t_5$  beträgt diese Differenz sogar rund neun Prozent.

An dieser Stelle sei allerdings darauf hingewiesen – was auch in den vorherigen Tabellen zur Zusammensetzung nach Geschlecht bereits ersichtlich war – dass auch im Querschnitt wohlmöglich eine leichte Überrepräsentanz weiblicher Personen vorliegt. In den offiziellen Schulstatistiken früherer Zeitpunkte<sup>16</sup> lag eine annähernde Gleichverteilung von männlichen und weiblichen Befragten vor. Mit dem Prozess des Ausscheidens aus der Schule und der Umstellung auf eine postalische Erhebung verschiebt sich dieses ursprünglich gleiche Verhältnis auch im Querschnitt in Richtung einer leichten Überrepräsentanz weiblicher Personen. Ab dem gleichen Zeitpunkt liegen wie bereits erwähnt allerdings keine Schulstatistiken zum Vergleich mehr vor.

Betrachtet man in Tabelle 5.46 den Zeitpunkt  $t_5$ , so lässt sich in diesem Jahr im Querschnittsdatensatz noch eine annähernd gleiche Zusammensetzung nach Geschlecht erkennen. Im Vier-Wellen-Panel hingegen liegt der Anteil an Frauen deutlich höher

<sup>16</sup> Siehe hierfür die Zahlen der Schulstatistiken  $S_{t_1}$ ,  $S_{t_2}$  und  $S_{t_3}$  im weiteren Verlauf dieses Berichts.

und bleibt auch in den Folgezeitpunkten auf diesem Niveau. In den Querschnitten der weiteren Zeitpunkte  $t_6$ ,  $t_7$  und  $t_8$  steigt der Anteil an Personen weiblichen Geschlechts auf durchschnittlich 53 Prozent. Dass zum Zeitpunkt  $t_5$  die Differenz zwischen Querschnitt und Panel neun Prozent beträgt und zu den weiteren drei Zeitpunkten nur noch sechs, ist also darauf zurückzuführen, dass weibliche Personen im Querschnitt ebenso zunehmend stärker vertreten sind – die Verschiebung, die im Vier-Wellen-Panel vorliegt, bleibt hingegen über die Zeitpunkte hinweg unverändert.

**Tabelle 5.46:** Zusammensetzung nach Geschlecht  $P_{t_5,6,7,8}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_5$	$Q_{t_5}$	2 265	49,8	2 282	50,2	4 547
	$P_{t_5,6,7,8}$	820	40,7	1 197	59,4	2 017
$t_6$	$Q_{t_6}$	1 567	47,0	1 768	53,0	3 335
	$P_{t_5,6,7,8}$	821	40,7	1 196	59,3	2 017
$t_7$	$Q_{t_7}$	1 452	47,1	1 634	52,9	3 086
	$P_{t_5,6,7,8}$	820	40,7	1 197	59,4	2 017
$t_8$	$Q_{t_8}$	1 446	46,8	1 644	53,2	3 090
	$P_{t_5,6,7,8}$	819	40,6	1 198	59,4	2 017

Das durchschnittliche Alter der Befragungspersonen weist sowohl im Querschnitt als auch im Vier-Wellen-Panel die zu erwartenden Werte von 17 bis 20 Jahren auf. Es liegen nur geringe Abweichungen zwischen Querschnittsdatensätzen und dem Vier-Wellen-Panel vor. Im Jahr 2006 (Zeitpunkt  $t_5$ ) liegt das Durchschnittsalter unter Verwendung des Querschnittsdatensatzes knapp ein halbes Jahr über dem Durchschnitt des Panels  $P_{t_5,6,7,8}$  und es zeigt sich außerdem – ebenso wie in etwas geringerem Maße auch zum Zeitpunkt  $t_6$  – eine etwas höhere Standardabweichung des Mittelwerts. Die Verteilung des Alters im Querschnitt ist zu diesen Zeitpunkten also etwas weniger homogen verteilt als in den Vergleichsdatensätzen bzw. Vergleichszeitpunkten.

Die Tabelle 5.48 zeigt die Zusammensetzung nach Schulform. Zum Zeitpunkt  $t_5$  besuchen nur noch sehr wenige Personen eine Haupt-, Real- oder Sonderschule. In den späteren Zeitpunkten spielen diese Schulformen keine Rolle mehr, da sich die Befragten mit durchschnittlich 17 Jahren vorrangig auf Gesamtschulen, Gymnasien oder Berufskollegs befinden oder die schulische Ausbildung gänzlich beendet haben.

Insbesondere zum Zeitpunkt  $t_5$  zeigt sich eine starke Diskrepanz zwischen Querschnittsdaten und Vier-Wellen-Panel. Schülerinnen und Schüler auf Gesamtschulen sind mit gut sechs Prozent im Panel überrepräsentiert, Personen auf Gymnasien sogar mit 14 Prozent. Umgekehrt liegt im Panel der Anteil an Personen, die ein Berufskolleg besuchen, 19 Prozentpunkte unter dem entsprechenden Wert des Querschnitts. Derar-

**Tabelle 5.47:** Altersdurchschnitt  $P_{t_{5,6,7,8}}$

		Alter	Std.-abw.	gesamt
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_{5,6,7,8}}$	17,0	0,8	2001
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{t_{5,6,7,8}}$	18,0	0,8	2011
$t_7$	$Q_{t_7}$	19,1	0,8	3081
	$P_{t_{5,6,7,8}}$	19,0	0,8	2016
$t_8$	$Q_{t_8}$	20,1	0,8	3073
	$P_{t_{5,6,7,8}}$	20,0	0,8	2011

tige Abweichungen liegen in den späteren Zeitpunkten ebenso vor, sind aber deutlich schwächer ausgeprägt. Zum Zeitpunkt  $t_8$  sind die Differenzen zwischen Querschnitts- und Paneldatensatz nur noch marginal.

**Tabelle 5.48:** Zusammensetzung  $P_{t_{5,6,7,8}}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							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
	$P_{t_{5,6,7,8}}$	0,5	0,5	22,4	34,2	42,3	0,1	—	—	1869
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{5,6,7,8}}$	—	—	18,1	30,4	30,9	—	9,8	10,9	1972
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{5,6,7,8}}$	—	—	15,9	28,5	26,3	—	13,0	16,4	1972
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{5,6,7,8}}$	—	—	4,2	3,7	22,2	—	19,1	50,8	1938

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

### 5.4.3 Vier-Wellen-Panel 2007, 2008, 2009 und 2011 ( $P_{t_{6,7,8,9}}$ )

Die Fallzahl des Vier-Wellen-Panels  $P_{t_{6,7,8,9}}$  beträgt 1 854. Auch hier liegt eine Verschiebung der Zusammensetzung nach Geschlecht vor (vgl. Tabelle 5.49), wie sie auch bereits in den vorherigen Abschnitten beschrieben wurde. Der Anteil weiblicher Befragter liegt im Querschnitt zu allen Zeitpunkten bei etwa 53 Prozent. Im Vier-Wellen-Panel sind Frauen mit jeweils rund 61 Prozent deutlich stärker vertreten.

**Tabelle 5.49:** Zusammensetzung nach Geschlecht  $P_{t_{6,7,8,9}}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_6$	$Q_{t_6}$	1 567	47,0	1 768	53,0	3 335
	$P_{t_{6,7,8,9}}$	719	38,8	1 135	61,2	1 854
$t_7$	$Q_{t_7}$	1 452	47,1	1 634	52,9	3 086
	$P_{t_{6,7,8,9}}$	720	38,8	1 134	61,2	1 854
$t_8$	$Q_{t_8}$	1 446	46,8	1 644	53,2	3 090
	$P_{t_{6,7,8,9}}$	720	38,8	1 134	61,2	1 854
$t_9$	$Q_{t_9}$	1 423	46,8	1 620	53,2	3 043
	$P_{t_{6,7,8,9}}$	722	38,9	1 132	61,1	1 854

Betrachtet man das Durchschnittsalter in Tabelle 5.50 so zeigen sich nur sehr geringe Abweichungen zwischen Querschnitts- und Paneldaten. Auch die Streuung ist zu allen betrachteten Zeitpunkten sowohl im Querschnitt als auch im Vier-Wellen-Panel gleich. Lediglich zum Zeitpunkt  $t_6$  ist im Querschnitt eine etwas höhere Standardabweichung zu erkennen, was auf eine etwas stärker ausgeprägte Heterogenität im Alter der Befragten hinweist.

**Tabelle 5.50:** Altersdurchschnitt  $P_{t_{6,7,8,9}}$

		Alter	Std.-abw.	gesamt
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_{6,7,8,9}}$	18,0	0,8	1 849
$t_7$	$Q_{t_7}$	19,1	0,8	3 081
	$P_{t_{6,7,8,9}}$	19,0	0,8	1 854
$t_8$	$Q_{t_8}$	20,1	0,8	3 073
	$P_{t_{6,7,8,9}}$	20,0	0,8	1 848
$t_9$	$Q_{t_9}$	22,1	0,9	3 021
	$P_{t_{6,7,8,9}}$	22,0	0,8	1 848

Tabelle 5.51 weist die Zusammensetzung der Daten nach Schulform aus. Die Daten sind hierbei nur bis zum Zeitpunkt  $t_8$  (durchschnittliches Alter 20 Jahre) dargestellt, da ab diesem Alter nahezu alle Befragten die schulischen Bildungseinrichtungen verlassen haben, sodass die Schulform ab diesem Zeitpunkt keine Rolle mehr spielt. Auch bereits in  $t_6$ ,  $t_7$  und  $t_8$  befinden sich keine Personen mehr auf Haupt-, Real- und Sonderschulen. bei den Zeitpunkten  $t_6$  und  $t_7$  zeigen sich Abweichungen derart, dass Gymnasiastinnen und Gymnasiasten im Panel stärker vertreten sind als im Querschnitt. Umgekehrt sind im

Panel Personen an Berufskollegs im Vergleich zum Querschnitt etwas unterrepräsentiert. Selbiges gilt für Personen, die keine Schule mehr besuchen. Zum Zeitpunkt  $t_8$  fallen diese Abweichungen zwischen Querschnittsdatensatz und Vier-Wellen-Panel nur noch gering aus.

**Tabelle 5.51:** Zusammensetzung  $P_{t_6,7,8,9}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							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
	$P_{t_6,7,8,9}$	—	—	17,7	31,1	30,4	—	10,4	10,4	1815
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_6,7,8,9}$	—	—	15,6	29,5	26,2	—	12,7	16,1	1813
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_6,7,8,9}$	—	—	4,0	3,9	22,6	—	19,0	50,5	1789

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

#### 5.4.4 Vier-Wellen-Panel 2008, 2009, 2011 und 2013 ( $P_{t_7,8,9,10}$ )

Das Vier-Wellen-Panel der Jahre 2008 bis 2013 umfasst 1 684 Fälle. Die Überrepräsentierung weiblicher Befragungspersonen im Panel nimmt im Vergleich mit den vorherigen Vier-Wellen-Panels weiter zu und liegt zu allen Zeitpunkten bei gut 62 Prozent. Die Differenz zum Querschnitt beträgt zu den einzelnen Zeitpunkten somit jeweils acht bis neun Prozentpunkte.

**Tabelle 5.52:** Zusammensetzung nach Geschlecht  $P_{t_7,8,9,10}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_7,8,9,10}$	634	37,7	1050	62,4	1684
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{t_7,8,9,10}$	634	37,7	1050	62,4	1684
$t_9$	$Q_{t_9}$	1423	46,8	1620	53,2	3043
	$P_{t_7,8,9,10}$	638	37,9	1046	62,1	1684
$t_{10}$	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_7,8,9,10}$	633	37,6	1051	62,4	1684

Bezüglich des Durchschnittsalters (s. Tabelle 5.53) zeigen sich nur sehr geringe Abweichungen zwischen den Querschnittsdaten und den Daten des Vier-Wellen-Panels  $P_{t_{7,8,9,10}}$ . Im letzten Erhebungszeitpunkt  $t_{10}$  weist das Alter im Querschnitt eine leicht höhere Streuung auf, woraus sich schlussfolgern lässt, dass das Alter im Querschnitt des Jahres 2013 eine etwas geringere Homogenität aufweist als in den Vergleichsdaten.

**Tabelle 5.53:** Altersdurchschnitt  $P_{t_{7,8,9,10}}$

		Alter	Std.-abw.	gesamt
$t_7$	$Q_{t_7}$	19,1	0,8	3081
	$P_{t_{7,8,9,10}}$	18,9	0,8	1684
$t_8$	$Q_{t_8}$	20,1	0,8	3073
	$P_{t_{7,8,9,10}}$	20,0	0,7	1682
$t_9$	$Q_{t_9}$	22,1	0,9	3021
	$P_{t_{7,8,9,10}}$	22,0	0,8	1680
$t_{10}$	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_{7,8,9,10}}$	23,9	0,8	1681

Ein Vergleich der Zusammensetzung nach Schulform ist für das Vier-Wellen-Panel  $P_{t_{7,8,9,10}}$  nur noch für die Zeitpunkte  $t_7$  und  $t_8$  (19. und 20. Lebensjahr) sinnvoll. Anhand der in Tabelle 5.54 abgebildeten Daten lässt sich eine Verzerrung des Paneldatensatzes erkennen: Zum Zeitpunkt  $t_7$  sind Schülerinnen und Schüler an Gesamtschulen sowie insbesondere an Gymnasien im Paneldatensatz häufiger vertreten als im Querschnitt dieses Erhebungszeitpunktes. Schülerinnen und Schüler eines Berufskollegs sowie jene Personen, die eine andere oder keine Schule besuchen, sind verglichen mit dem Querschnitt im Panel unterrepräsentiert. Im darauffolgenden Erhebungsjahr  $t_8$  sind die Diskrepanzen zwischen den beiden Arten von Datensätzen allerdings nur noch schwach ausgeprägt.

**Tabelle 5.54:** Zusammensetzung  $P_{t_{7,8,9,10}}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							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
	$P_{t_{1,2,3,4,5,6,7,8,9,10}}$	—	—	16,7	30,4	25,7	—	12,5	15,7	1648
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{1,2,3,4,5,6,7,8,9,10}}$	—	—	3,9	4,1	22,6	—	19,1	50,3	1623

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

## 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 nur jene Datensätze beschrieben, die Teilnahmen in fünf direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke). Dabei sind auch die Datensätze enthalten, die aufgrund der Änderung des Erhebungsrhythmus in einem Abstand von zwei Jahren erhoben wurden.

### 5.5.1 Fünf-Wellen-Panel 2004, 2005, 2006, 2007 und 2008 ( $P_{t_{3,4,5,6,7}}$ )

Das Fünf-Wellen-Panel  $P_{t_{3,4,5,6,7}}$  beinhaltet 1 741 Fälle. Mit Blick auf Tabelle 5.55 lässt sich feststellen, dass in den Datensätzen eine Verzerrung bezüglich des Merkmals Geschlechts vorliegt. Die Schulstatistik des Jahres 2004 (Erhebungszeitpunkt  $t_3$ ) weist aus, dass sich rund 51 Prozent männliche und 49 Prozent weibliche Befragte an den erhobenen Schulen befinden. Im Querschnittsdatsatz  $Q_{t_3}$  dieses Jahres liegt jedoch eine genau umgekehrte Prozentuierung vor. Dieses Verhältnis verschiebt sich in den Querschnittsdatsätzen der Folgejahre ( $Q_{t_4}$  bis  $Q_{t_7}$ ) zunehmend, sodass letztlich zum Erhebungszeitpunkt  $t_7$  die weiblichen Befragten mit rund 53 Prozent leicht überwiegen. Diese Verzerrung zeigt sich im Fünf-Wellen-Panel  $P_{t_{3,4,5,6,7}}$  nochmals stärker. Dort liegt der Anteil an weiblichen Befragten zu allen Erhebungszeitpunkten zwischen 59 und knapp 60 Prozent – Frauen sind demnach im Panel überrepräsentiert.

**Tabelle 5.55:** Zusammensetzung nach Geschlecht  $P_{t_{3,4,5,6,7}}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_{3,4,5,6,7}}$	698	40,1	1043	59,9	1741
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_{3,4,5,6,7}}$	707	40,6	1034	59,4	1741
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_{3,4,5,6,7}}$	705	40,5	1036	59,5	1741
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{t_{3,4,5,6,7}}$	704	40,4	1037	59,6	1741
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_{3,4,5,6,7}}$	708	40,7	1033	59,3	1741

Beim durchschnittlichen Alter (Tabelle 5.56) zeigen sich nur unbedeutende Abweichungen von jeweils 0,1 Jahren zwischen den Datensätzen der Querschnitte und des

Fünf-Wellen-Panels. Lediglich zum Zeitpunkt  $t_5$  tritt eine etwas größere Differenz von 0,4 Jahren auf. Zu diesem Zeitpunkt sowie auch zum darauffolgenden Zeitpunkt  $t_6$  liegt eine etwas größere Standardabweichung als in den Vergleichsdaten vor, was ein Hinweis auf eine etwas größere Heterogenität der Altersverteilung zu diesen beiden Zeitpunkten ist.

**Tabelle 5.56:** Altersdurchschnitt  $P_{t_3,4,5,6,7}$

		Alter	Std.-abw.	gesamt
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_3,4,5,6,7}$	15,0	0,7	1 723
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_3,4,5,6,7}$	16,0	0,7	1 727
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_3,4,5,6,7}$	17,0	0,8	1 733
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_3,4,5,6,7}$	17,9	0,7	1 737
$t_7$	$Q_{t_7}$	19,1	0,8	3 081
	$P_{t_3,4,5,6,7}$	18,9	0,7	1 741

Zieht man die Zusammensetzung der Panel- und Querschnittsdaten nach der Schulform heran (s. Tabelle 5.57), so ist auch hier eine Verzerrung des Fünf-Wellen-Panels  $P_{t_3,4,5,6,7}$  erkennbar. Vergleicht man zum Zeitpunkt  $t_3$  die Verteilung der Schulform im Paneldatensatz mit jener der Schulstatistik, so lässt sich feststellen, dass mit knapp 8 Prozentpunkten deutlich weniger Hauptschülerinnen und -schüler im Panel vertreten sind als die Schulstatistik ausweist. Umgekehrt verhält es sich mit Schülerinnen und Schülern an Gymnasien, die mit rund neun Prozent im Panel überrepräsentiert sind. Selbe Verschiebung liegt auch beim Vergleich vom Querschnitt  $Q_{t_3}$  mit der Schulstatistik  $S_{t_3}$  vor, jedoch sind die Differenzen hier mit nur rund zwei Prozent deutlich geringer.

Beim Vergleich von Querschnitt und Fünf-Wellen-Panel verschärft sich diese Verzerrung nach Schulform im Erhebungszeitpunkt  $t_5$ . Hier sind Personen an Gymnasien um rund 14 Prozentpunkte im Panel stärker vertreten als im Querschnitt. Bei den Gesamtschulen beträgt diese Überrepräsentierung immerhin noch rund fünf Prozent. Hingegen befinden sich im Panel 19 Prozent weniger Berufsschülerinnen und -schüler als im Querschnitt. Diese Abweichungen sind jedoch aufgrund der erwähnten Erweiterung der Stichprobe zum Zeitpunkt  $t_5$  mit Vorsicht zu interpretieren.

Diese Tendenz bleibt im Großen und Ganzen auch in den weiteren Erhebungszeitpunkten erhalten, jedoch sind die Differenzen geringer. Im Erhebungsjahr 2008 (Zeitpunkt  $t_7$ ) liegen beim Vergleich mit dem Querschnitt  $Q_{t_7}$  im Panel nur noch rund fünf Prozent mehr Gymnasiastinnen und Gymnasiasten vor; die Unterrepräsentierung der Berufsschü-

lerinnen und -schüler beläuft sich nur noch auf rund ein Prozent. Zusätzlich befinden sich im Panel  $P_{t_{3,4,5,6,7}}$  etwa vier Prozent weniger Personen, die keine Schule (mehr) besuchen, als es im Querschnitt der Fall ist.

**Tabelle 5.57:** Zusammensetzung  $P_{t_{3,4,5,6,7}}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								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
	$P_{t_{3,4,5,6,7}}$	17,1	23,2	31,0	28,7	—	0,1	—	—	1741
$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_{3,4,5,6,7}}$	16,0	23,2	31,1	28,6	0,8	0,1	0,4	0,1	1736
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{3,4,5,6,7}}$	0,6	0,6	21,1	33,8	43,8	—	—	—	1599
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{3,4,5,6,7}}$	—	—	17,4	29,4	31,6	—	11,2	10,4	1698
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{3,4,5,6,7}}$	—	—	15,2	28,0	27,0	—	13,6	16,3	1702

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

### 5.5.2 Fünf-Wellen-Panel 2005, 2006, 2007, 2008 und 2009 ( $P_{t_{4,5,6,7,8}}$ )

Das Fünf-Wellen-Panel  $P_{t_{4,5,6,7,8}}$ , das die Erhebungsjahre von 2005 bis 2009 umfasst, beinhaltet insgesamt 1 812 Personen. Aus Tabelle 5.58 wird ersichtlich, dass sowohl in Querschnitt als auch im Fünf-Wellen-Panel Verzerrungen im Hinblick auf die Zusammensetzung nach Geschlecht vorliegen. Zu den Erhebungszeitpunkten  $t_4$  und  $t_5$  ist die Verteilung nach männlich und weiblich in den Querschnittsdatensätzen  $Q_{t_4}$  und  $Q_{t_5}$  relativ ausgeglichen. In den folgenden Wellen  $t_6$ ,  $t_7$  und  $t_8$  überwiegt jedoch leicht der Anteil weiblicher Personen, der jeweils bei ca. 53 Prozent liegt.

Im Fünf-Wellen-Panel ist diese Verzerrung nochmals stärker ausgeprägt. Dort liegt der Anteil weiblicher Befragungspersonen in allen Erhebungswellen bei knapp 60 Prozent und somit deutlich über den Werten der Querschnitte.

In Tabelle 5.59 ist das durchschnittliche Alter der Befragten in den Querschnittsdatensätzen und im Panel dargestellt. Es zeigen sich in den einzelnen Erhebungszeitpunkten nur geringfügige Abweichungen zwischen Querschnitt und Panel. Einzig im Querschnitt  $Q_{t_5}$  liegt das durchschnittliche Alter um 0,4 Jahre höher als es im Fünf-Wellen-Panel der

**Tabelle 5.58:** Zusammensetzung nach Geschlecht  $P_{t_4,5,6,7,8}$

		männlich		weiblich		gesamt (100 %)
		Hfk	%	Hfk	%	
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_4,5,6,7,8}$	732	40,4	1080	59,6	1812
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_4,5,6,7,8}$	730	40,3	1082	59,7	1812
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{t_4,5,6,7,8}$	729	40,2	1083	59,8	1812
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_4,5,6,7,8}$	732	40,4	1080	59,6	1812
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{t_4,5,6,7,8}$	731	40,3	1081	59,7	1812

Fall ist. Zugleich ist zu diesem Zeitpunkt – ebenso wie auch beim Querschnitt  $Q_{t_6}$  – das Alter etwas weniger homogen verteilt als im Paneldatensatz.

**Tabelle 5.59:** Altersdurchschnitt  $P_{t_4,5,6,7,8}$

		Alter	Std.-abw.	gesamt
		$t_4$	$Q_{t_4}$	16,1
	$P_{t_4,5,6,7,8}$	16,0	0,7	1 793
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_4,5,6,7,8}$	17,0	0,8	1 803
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_4,5,6,7,8}$	18,0	0,8	1 808
$t_7$	$Q_{t_7}$	19,1	0,8	3081
	$P_{t_4,5,6,7,8}$	18,9	0,8	1 811
$t_8$	$Q_{t_8}$	20,1	0,8	3073
	$P_{t_4,5,6,7,8}$	20,0	0,7	1 806

Die Zusammensetzung nach Schulform für das Fünf-Wellen-Panel  $P_{t_4,5,6,7,8}$  weist das übliche Verschiebungsmuster auf, wie es auch bereits in den vorherigen Teilpaneldatensätzen erkennbar war: Zum Erhebungszeitpunkt  $t_4$  sind im Panel etwas weniger Schülerinnen und Schüler an Hauptschulen vertreten als im Querschnitt. Dafür sind umgekehrt verglichen mit dem Querschnitt Personen an Gymnasien im Panel überrepräsentiert. Zusätzlich sind Personen an Gesamtschulen beim Vergleich der Datensätze im Panel stärker vorhanden und umgekehrt Besucherinnen und Besucher von Berufs-

kollegs deutlich unterrepräsentiert – diese Tendenz wird insbesondere am Zeitpunkt  $t_5$  ersichtlich.

In den späteren Zeitpunkten verliert die Schulform an Bedeutung, da die Befragten mit zunehmenden Alter ihre schulischen Bildungseinrichtungen verlassen. Entsprechend liegt im letzten dargestellten Zeitpunkt  $t_8$  (Durchschnittsalter 20 Jahre) ein hoher Anteil an Personen vor, die keine Schule mehr besuchen und es treten nur noch unbedeutende Diskrepanzen zwischen Querschnitts- und Paneldatensatz bezüglich der Schulform auf.

**Tabelle 5.60:** Zusammensetzung  $P_{t_4,5,6,7,8}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								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
	$P_{t_4,5,6,7,8}$	15,3	22,8	31,4	29,1	0,7	0,2	0,3	0,1	1 808
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4 253
	$P_{t_4,5,6,7,8}$	0,6	0,5	21,2	34,7	42,9	0,1	—	—	1 671
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3 219
	$P_{t_4,5,6,7,8}$	—	—	17,5	30,2	31,5	—	10,2	10,5	1 769
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2 975
	$P_{t_4,5,6,7,8}$	—	—	15,4	28,4	26,9	—	13,5	15,9	1 776
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2 934
	$P_{t_4,5,6,7,8}$	—	—	3,9	3,7	22,6	—	19,2	50,6	1 743

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

### 5.5.3 Fünf-Wellen-Panel 2006, 2007, 2008, 2009 und 2011 ( $P_{t_5,6,7,8,9}$ )

Das fünf Wellen umfassende Panel  $P_{t_5,6,7,8,9}$  beinhaltet 1 768 Fälle. Vergleicht man die Querschnitts- und Teilpaneldatensätze bezüglich ihrer Zusammensetzung nach Geschlecht (Tabelle 5.61), so wird eine Verzerrung des Fünf-Wellen-Panels erkennbar. Mit jeweils rund 63 Prozent liegt der Anteil weiblicher Befragter rund zehn Prozentpunkte höher als in den Querschnitten der jeweiligen Erhebungsjahre.

Betrachtet man den Altersdurchschnitt, welcher in Tabelle 5.62 abgebildet ist, so zeigen sich keine wesentlichen Differenzen zwischen Quer- und Längsschnitt. Die Abweichung zwischen Querschnitt und dem Fünf-Wellen-Panel beträgt maximal 0,2 Jahre und das Ausmaß an Streuung ist zu allen Zeitpunkten annähernd gleich. Lediglich zum Erhebungszeitpunkt  $t_5$  liegt das durchschnittliche Alter mit einem Wert von 17,4 um 0,4 Jahre als im Fünf-Wellen-Panel zu diesem Zeitpunkt. Zugleich ist die Altersverteilung

**Tabelle 5.61:** Zusammensetzung nach Geschlecht  $P_{5,6,7,8,9}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{5,6,7,8,9}$	679	38,4	1089	61,6	1768
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3336
	$P_{5,6,7,8,9}$	679	38,4	1089	61,6	1768
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{5,6,7,8,9}$	679	38,4	1089	61,6	1768
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{5,6,7,8,9}$	678	38,4	1090	61,6	1768
$t_9$	$Q_{t_9}$	1423	46,8	1620	53,2	3043
	$P_{5,6,7,8,9}$	680	38,5	1088	61,5	1768

zu diesem Zeitpunkt, und auch zum Zeitpunkt  $t_6$  etwas heterogener als in den übrigen Zeitpunkten, was sich in einer leicht höheren Standardabweichung widerspiegelt.

**Tabelle 5.62:** Altersdurchschnitt  $P_{5,6,7,8,9}$

		Alter	Std.-abw.	gesamt
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{5,6,7,8,9}$	17,0	0,8	1755
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{5,6,7,8,9}$	18,0	0,8	1763
$t_7$	$Q_{t_7}$	19,1	0,8	3081
	$P_{5,6,7,8,9}$	18,9	0,8	1768
$t_8$	$Q_{t_8}$	20,1	0,8	3073
	$P_{5,6,7,8,9}$	20,0	0,7	1763
$t_9$	$Q_{t_9}$	22,1	0,9	3021
	$P_{5,6,7,8,9}$	22,0	0,8	1763

Die Zusammensetzung nach Schulform in Tabelle 5.63 lässt auf eine Verzerrung des Paneldatensatzes nach diesem Merkmal schließen. Die Tendenz geht dahin, dass Schülerinnen und Schüler an Gesamtschulen und Gymnasien im Fünf-Wellen-Panel vergleichen mit dem Querschnitt überrepräsentiert sind. Im Gegenzug sind Personen, die ein Berufskolleg oder keine Schule besuchen, im Panel weniger stark vertreten als im Querschnitt. Diese Diskrepanz ist zum Zeitpunkt  $t_5$  am stärksten ausgeprägt, wobei zu diesem Zeitpunkt die Differenzen aufgrund einer veränderten Stichprobe vorsichtig

zu betrachten sind. Derartige Differenzen zwischen Querschnitt und Fünf-Wellen-Panel gehen jedoch in den späteren Erhebungszeitpunkten zurück: Im letzten Zeitpunkt  $t_8$  bestehen nur noch wenig bedeutsame Differenzen zwischen dem Querschnitt und dem Paneldatensatz.

**Tabelle 5.63:** Zusammensetzung  $P_{t_5,6,7,8,9}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_5,6,7,8,9}$	0,2	0,4	21,9	37,7	39,8	0,1	—	—	1454
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_5,6,7,8,9}$	—	—	18,1	33,4	29,4	—	9,8	9,3	1541
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_5,6,7,8,9}$	—	—	16,1	31,7	25,1	—	12,3	14,9	1534
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_5,6,7,8,9}$	—	—	3,9	4,2	22,1	—	19,4	50,5	1516

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

#### 5.5.4 Fünf-Wellen-Panel 2007, 2008, 2009, 2011 und 2013 ( $P_{t_6,7,8,9,10}$ )

Das Fünf-Wellen-Panel  $P_{t_6,7,8,9,10}$  der Erhebungsjahre 2007 bis 2013 enthält 1 636 Fälle. Aus Tabelle 5.64, welche die Zusammensetzung nach Geschlecht für Panel- und Querschnittsdatensätze der Erhebungszeitpunkte darstellt, wird eine Verzerrung dieses Fünf-Wellen-Panels deutlich. Der Anteil weiblicher Personen beläuft sich in den Zeitpunkten von  $t_6$  bis  $t_9$  auf jeweils rund 53 Prozent, zum Erhebungszeitpunkt  $t_{10}$  liegt dieser Anteil bei 54 Prozent. Im Fünf-Wellen-Panel beträgt der Anteil an Frauen in allen Wellen etwa 63 Prozent und liegt damit bis zu zehn Prozentpunkte höher als im Querschnitt – Frauen sind im Fünf-Wellen-Panel überrepräsentiert.

Ein Vergleich des durchschnittlichen Alters (s. Tabelle 5.65) zeigt nur marginale Abweichungen zwischen Querschnittsdatensätzen und dem Fünf-Wellen-Panel: In den Querschnitt liegt das Durchschnittsalter in den einzelnen Erhebungszeitpunkten um 0,1 bis 0,2 Jahre höher als im Fünf-Wellen-Panel. Die Standardabweichung weist auf eine gleichbleibend homogene Streuung des Alters hin. Lediglich in den Zeitpunkten  $t_6$  und  $t_{10}$  ist das Alter in den Querschnittsdatensätzen etwas weniger homogen verteilt als in den Vergleichsdaten.

Mit zunehmendem Alter verlassen die Befragten nach und nach die Schule; aus diesem Grund ist in Tabelle 5.66 die Zusammensetzung der Datensätze nach Schulform nur bis

**Tabelle 5.64:** Zusammensetzung nach Geschlecht  $P_{6,7,8,9,10}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_6$	$Q_{t_6}$	1 567	47,0	1 768	53,0	3 335
	$P_{6,7,8,9,10}$	604	36,9	1 032	63,1	1 636
$t_7$	$Q_{t_7}$	1 452	47,1	1 634	52,9	3 086
	$P_{6,7,8,9,10}$	606	37,0	1 030	63,0	1 636
$t_8$	$Q_{t_8}$	1 446	46,8	1 644	53,2	3 090
	$P_{6,7,8,9,10}$	606	37,0	1 030	63,0	1 636
$t_9$	$Q_{t_9}$	1 423	46,8	1 620	53,2	3 043
	$P_{6,7,8,9,10}$	610	37,3	1 026	62,7	1 636
$t_{10}$	$Q_{t_{10}}$	1 302	45,7	1 546	54,3	2 848
	$P_{6,7,8,9,10}$	606	37,0	1 030	63,0	1 636

zum Zeitpunkt  $t_8$  – die Befragten sind hier durchschnittlich 20 Jahre alt – ausgewertet. Es ist zu erkennen, dass sich keine Befragungspersonen mehr an Haupt-, Real- oder Sonderschulen befinden. Im Bezug auf die übrigen Schulformen wird ersichtlich, dass vor allem Schülerinnen und Schüler an Gymnasien im Fünf-Wellen-Panel stärker repräsentiert sind als im Querschnitt – zum Zeitpunkt  $t_7$  beträgt diese Differenz beispielsweise acht Prozentpunkte. Entgegengesetzt sind zum selben Zeitpunkt Besucherinnen und Besucher von Berufskollegs um 2,5 Prozent weniger im Panel vertreten als im Querschnitt; bei Personen, welche keine Schule besuchen, beträgt diese Abweichung etwas über fünf Prozent.

Im Erhebungsjahr 2009 (Zeitpunkt  $t_8$ ) weisen Querschnittsdatensatz und Fünf-Wellen-Panel nur noch sehr geringe Diskrepanzen zueinander auf.

**Tabelle 5.65:** Altersdurchschnitt  $P_{6,7,8,9,10}$

		Alter	Std.-abw.	gesamt
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{t_6,7,8,9,10}$	18,0	0,8	1632
$t_7$	$Q_{t_7}$	19,1	0,8	3081
	$P_{t_6,7,8,9,10}$	18,9	0,8	1636
$t_8$	$Q_{t_8}$	20,1	0,8	3073
	$P_{t_6,7,8,9,10}$	20,0	0,8	1634
$t_9$	$Q_{t_9}$	22,1	0,9	3021
	$P_{t_6,7,8,9,10}$	22,0	0,8	1633
$t_{10}$	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_6,7,8,9,10}$	23,9	0,8	1633

**Tabelle 5.66:** Zusammensetzung  $P_{6,7,8,9,10}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							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	3219
	$P_{t_6,7,8,9,10}$	—	—	17,8	32,4	29,8	—	10,3	9,8	1606
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_6,7,8,9,10}$	—	—	15,9	30,8	25,6	—	12,4	15,4	1600
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_6,7,8,9,10}$	—	—	3,9	4,1	22,4	—	18,9	50,7	1582

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

## 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 nur jene Datensätze beschrieben, die Teilnahmen in sechs direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke). Ebenso sind Datensätze enthalten, die aufgrund der Änderung des Erhebungsrythmus in einem Abstand von zwei Jahren erhoben wurden.

5.6.1 Sechs-Wellen-Panel 2003, 2004, 2005, 2006, 2007 und 2008

$(P_{t_{2,3,4,5,6,7}})$

Das sechs Wellen umfassende Panel von 2003 bis 2008 enthält 1 465 Fälle. Anhand der in Tabelle 5.67 dargestellten Zusammensetzung nach Geschlecht lässt sich für dieses Panel  $P_{t_{2,3,4,5,6,7}}$  eine Verzerrung feststellen. Zu den Zeitpunkten  $t_2$  und  $t_3$  weist die Schulstatistik ( $S_{t_2}$  und  $S_{t_3}$ ) einen Anteil von rund 49 Prozent weiblicher Befragter aus. Bis zum Jahre 2006 (Erhebungszeitpunkt  $t_5$ ) weicht dieser Anteil im Querschnitt nur leicht hiervon ab : Von  $Q_{t_2}$  bis  $Q_{t_5}$  liegt der Anteil an Frauen zwischen 49 und 51 Prozent. Mit zunehmender Befragungsdauer erhöht sich dieser Anteil jedoch und liegt zu den Zeitpunkten  $t_6$  und  $t_7$  bei 53 Prozent. Eine deutlich ausgeprägtere Verzerrung zeigt sich im Paneldatensatz; hier liegt der Anteil weiblicher Befragter durchgehend zu allen dargestellten Zeitpunkten bei rund 61 Prozent. Im Panel  $P_{t_{2,3,4,5,6,7}}$  liegt also eine Verschiebung vor derart, dass Frauen über- und Männer unterrepräsentiert sind.

**Tabelle 5.67:** Zusammensetzung nach Geschlecht  $P_{t_{2,3,4,5,6,7}}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{t_{2,3,4,5,6,7}}$	571	39,0	894	61,0	1465
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_{2,3,4,5,6,7}}$	565	38,6	900	61,4	1465
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_{2,3,4,5,6,7}}$	572	39,0	893	61,0	1465
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_{2,3,4,5,6,7}}$	569	38,8	896	61,2	1465
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{t_{2,3,4,5,6,7}}$	569	38,8	896	61,2	1465
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_{2,3,4,5,6,7}}$	572	39,0	893	61,0	1465

Bei Betrachtung des durchschnittlichen Alters (Tabelle 5.68) finden sich nur geringe Abweichungen zwischen Panel- und Querschnittsdatsätzen von 2003 bis 2008. In allen hier betrachteten Wellen liegt das gemittelte Alter im Querschnitt 0,1 bis 0,2 Jahre über dem des Panels  $P_{t_{2,3,4,5,6,7}}$  zu den jeweiligen Erhebungszeitpunkten. lediglich zum Zeitpunkt  $t_5$  liegt das Durchschnittsalter im Querschnitt  $Q_{t_5}$  ein halbes Jahr über dem Vergleichswert für das Sechs-Wellen-Panel. An dieser Stelle – ebenso wie in etwas

geringerem Maße auch bei  $t_6$  – lässt sich außerdem eine etwas größere Heterogenität der Altersverteilung feststellen.

**Tabelle 5.68:** Altersdurchschnitt  $P_{2,3,4,5,6,7}$

		Alter	Std.-abw.	gesamt
$t_2$	$Q_{t_2}$	14,0	0,8	3 270
	$P_{2,3,4,5,6,7}$	13,9	0,7	1 448
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{2,3,4,5,6,7}$	14,9	0,7	1 453
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{2,3,4,5,6,7}$	15,9	0,7	1 453
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{2,3,4,5,6,7}$	16,9	0,8	1 457
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{2,3,4,5,6,7}$	17,9	0,7	1 463
$t_7$	$Q_{t_7}$	19,1	0,8	3 081
	$P_{2,3,4,5,6,7}$	18,9	0,7	1 465

Die Tabelle 5.69 präsentiert die Zusammensetzung des Sechs-Wellen-Panels  $P_{2,3,4,5,6,7}$  und der Querschnittsdatensätze von  $t_2$  bis  $t_7$  nach der Schulform. Es ist zu erkennen, dass eine Verzerrung der erhobenen Datensätze nach diesem Merkmal vorliegt. Zu den Zeitpunkten  $t_2$  und  $t_3$ , bei denen ein Vergleich mit der Schulstatistik möglich ist, zeigt sich, dass bereits im Querschnitt der Anteil an Gymnasiastinnen und Gymnasiasten leicht erhöht und umgekehrt der Anteil an Schülerinnen und Schülern an Hauptschulen verglichen mit der offiziellen Statistik etwas niedriger ist. Diese Differenz ist beim Vergleich von Schulstatistik und Sechs-Wellen-Panel noch deutlicher ausgeprägt: Die derart gerichtete Abweichung beträgt hier rund neun (Gymnasium) bis elf (Hauptschule) Prozentpunkte.

Ab dem Zeitpunkt  $t_4$  sind nur noch Vergleiche zwischen Querschnitt und Panel möglich, da eine Gegenüberstellung mit der Schulstatistik ab diesem Zeitpunkt aufgrund schrittweiser Beendigung der Schullaufbahn hinfällig ist. Doch auch bei diesem Vergleich lässt sich dennoch eine verschärfte Verzerrung des Panels  $P_{2,3,4,5,6,7}$  festhalten, welche insbesondere zum Zeitpunkt  $t_5$  deutlich wird: Im Sechs-Wellen-Panel liegt der Anteil an Schülerinnen und Schülern von Gymnasien knapp 13 Prozentpunkte über dem Wert des Querschnitts  $Q_{t_5}$ . Gegensätzlich dazu sind zum selben Zeitpunkt im Panel rund 19 Prozent weniger Personen vertreten, die ein Berufskolleg besuchen, als es im Querschnitt der Fall ist.

Diese Differenzen reduzieren sich in den weiteren Wellen und mit abnehmender Bedeutung der Schule etwas, dennoch bleibt auch im letzten betrachteten Erhebungs-

zeitpunkt  $t_7$  eine leichte Abweichung bestehen: Der Anteil der Schulform Gymnasium liegt im Panel rund fünf Prozent über dem Wert des Querschnitts, der Anteil an Personen, die keine Schule (mehr) besuchen, umgekehrt rund fünf Prozent darunter.

**Tabelle 5.69:** Zusammensetzung  $P_{t_{2,3,4,5,6,7}}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							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
	$P_{t_{2,3,4,5,6,7}}$	16,4	23,6	32,3	27,7	—	—	—	—	1465
$t_3$	$S_{t_3}$	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
	$P_{t_{2,3,4,5,6,7}}$	16,5	23,9	32,5	27,1	—	—	—	—	1465
$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_{2,3,4,5,6,7}}$	15,9	23,7	32,3	27,4	0,5	—	0,3	—	1464
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{2,3,4,5,6,7}}$	0,7	0,6	22,2	32,8	43,8	—	—	—	1361
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{2,3,4,5,6,7}}$	—	—	18,5	28,8	31,9	—	11,2	9,5	1426
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{2,3,4,5,6,7}}$	—	—	15,9	27,5	27,1	—	13,8	15,7	1430

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

### 5.6.2 Sechs-Wellen-Panel 2004, 2005, 2006, 2007, 2008 und 2009 ( $P_{t_{3,4,5,6,7,8}}$ )

Das Sechs-Wellen-Panel  $P_{t_{3,4,5,6,7,8}}$  beinhaltet 1 616 Fälle. Bei Betrachtung der Zusammensetzung nach Geschlecht (Tabelle 5.70) wird eine Verzerrung des Panels hinsichtlich dieses Merkmals deutlich. Die Schulstatistik  $S_{t_3}$  weist für das Jahr 2004 einen Anteil an Schülerinnen von 49 Prozent aus. Im Querschnitt dieses Zeitpunkts sind weibliche Personen jedoch mit 51 Prozent etwas stärker vertreten. Dieser Wert nimmt im Laufe der Erhebung zu, sodass der Anteil weiblicher Personen im Querschnitt  $Q_{t_8}$  letztlich gut 53 Prozent beträgt. Diese Verschiebung hinsichtlich einer Überrepräsentanz weiblicher Personen tritt im Sechs-Wellen-Panel nochmals stärker auf: In den Panelquerschnitten aller Zeitpunkte von 2004 bis 2009 liegt der Anteil an Frauen in der Stichprobe bei jeweils rund 61 Prozent.

**Tabelle 5.70:** Zusammensetzung nach Geschlecht  $P_{t_3,4,5,6,7,8}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_3,4,5,6,7,8}$	629	38,9	987	61,1	1616
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_3,4,5,6,7,8}$	636	39,4	980	60,6	1616
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_3,4,5,6,7,8}$	635	39,3	981	60,7	1616
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{t_3,4,5,6,7,8}$	634	39,2	982	60,8	1616
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_3,4,5,6,7,8}$	637	39,4	979	60,6	1616
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{t_3,4,5,6,7,8}$	636	39,4	980	60,6	1616

Eine Analyse des durchschnittlichen Alters (Tabelle 5.71) in den verschiedenen Datensätzen zeigt auch im Sechs-Wellen-Panel  $P_{t_3,4,5,6,7,8}$  weiterhin keine problematischen Differenzen. Der Altersdurchschnitt entspricht zu allen Zeitpunkten den zu erwartenden Werten und die Abweichung zwischen Panel- und allgemeinem Querschnitt beträgt 0,1 bis 0,2 Jahre. Wie auch bereits in den Beschreibungen der vorherigen Paneldatensätze liegt diese Abweichung einzig zum Zeitpunkt  $t_5$  mit einer Differenz von 0,4 etwas höher. Mit einer Standardabweichung von 1,3 bzw. 1,1 ist außerdem zu den Erhebungszeitpunkten  $t_5$  und  $t_6$  das Alter im Querschnitt etwas weniger homogen verteilt als in den Vergleichsdatensätzen.

Die Zusammensetzung der Datensätze nach Schulform, wie sie in Tabelle 5.72 abgebildet ist, zeigt auch beim Vergleich von Sechs-Wellen-Panel  $P_{t_3,4,5,6,7,8}$  mit den Querschnittsdatensätzen der Jahre 2004 bis 2009 eine Verzerrung in der Form, wie sie auch bereits in den vorherigen Kapiteln und Abschnitten beschrieben wurde.

Beim Vergleich mit der Schulstatistik  $S_{t_3}$  ist zu erkennen, dass sowohl im Quer- als auch im Längsschnitt eine Verschiebung vorliegt: Zum Zeitpunkt  $t_3$  sind Schülerinnen und Schüler an Gymnasien im Querschnitt  $Q_{t_3}$  mit etwas über zwei Prozent stärker vertreten als die Schulstatistik ausweist, im Sechs-Wellen-Panel mit rund zehn Prozent mehr. Entgegengesetzt befinden sich gegenüber  $S_{t_3}$  im Querschnitt  $Q_{t_3}$  gut zwei Prozent weniger Hauptschülerinnen und -schüler; verglichen mit dem Sechs-Wellen-Panel beträgt diese Differenz knapp neun Prozentpunkte.

**Tabelle 5.71:** Altersdurchschnitt  $P_{t_{3,4,5,6,7,8}}$

		Alter	Std.-abw.	gesamt
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_{3,4,5,6,7,8}}$	15,0	0,7	1 600
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_{3,4,5,6,7,8}}$	15,9	0,7	1 603
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_{3,4,5,6,7,8}}$	17,0	0,8	1 609
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_{3,4,5,6,7,8}}$	17,9	0,7	1 613
$t_7$	$Q_{t_7}$	19,1	0,8	3 081
	$P_{t_{3,4,5,6,7,8}}$	18,9	0,7	1 616
$t_8$	$Q_{t_8}$	20,1	0,8	3 073
	$P_{t_{3,4,5,6,7,8}}$	20,0	0,7	1 611

Diese Diskrepanzen verringern sich mit den weiteren Erhebungswellen, da zeitgleich die Bedeutung der Schule stetig abnimmt – ab dem Erhebungszeitpunkt  $t_6$  befinden sich keine Personen mehr an Haupt-, Real- oder Sonderschulen, stattdessen steigt zunehmend der Anteil an Personen, die keine Schule (mehr) besuchen. Die Abweichung zwischen Sechs-Wellen-Panel und Querschnitt bezüglich der Schulform Gymnasium beträgt zum Zeitpunkt  $t_6$  nur noch rund vier Prozent, derart, dass diese Schulform überrepräsentiert ist. Hingegen befinden sich im Panel  $P_{t_{3,4,5,6,7,8}}$  beim Vergleich mit  $Q_{t_6}$  vier Prozent weniger Personen, die angeben, keine Schule zu besuchen. Mit fortschreitender Erhebungsdauer reduzieren sich derartige Abweichungen zunehmend, sodass sich in der letzten hier abgebildeten Erhebungswelle  $t_8$  nur noch geringfügige Differenzen in der Verteilung der Schulformen zwischen Querschnitt und Panel finden lassen.

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

		Schulform <sup>a</sup> (%)								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
	$P_{t_{3,4,5,6,7,8}}$	16,2	23,2	31,1	29,5	—	0,1	—	—	1616
$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_{3,4,5,6,7,8}}$	15,0	23,3	31,1	29,3	0,8	—	0,4	0,1	1612
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{3,4,5,6,7,8}}$	0,6	0,6	21,3	34,6	43,0	—	—	—	1490
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{3,4,5,6,7,8}}$	—	—	17,4	30,3	31,4	—	10,6	10,3	1576
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{3,4,5,6,7,8}}$	—	—	15,5	28,7	26,7	—	13,3	15,8	1583
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{3,4,5,6,7,8}}$	—	—	4,1	3,6	22,2	—	19,4	50,7	1555

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

### 5.6.3 Sechs-Wellen-Panel 2005, 2006, 2007, 2008, 2009 und 2011 ( $P_{t_{4,5,6,7,8,9}}$ )

Das Sechs-Wellen-Panel der Erhebungsjahre 2005 bis 2009 beinhaltet 1 586 Personen. Hinsichtlich des Merkmals Geschlechts liegt eine Verzerrung dieses Sechs-Wellen-Panels vor (Tabelle 5.73). In den Erhebungswellen 2005 und 2006 ( $t_4$  und  $t_5$ ) ist das Verhältnis zwischen männlichen und weiblichen Personen ungefähr ausgewogen. In den späteren Erhebungszeitpunkten ab  $t_6$  verschiebt sich diese Relation jedoch, sodass der Anteil an Frauen jeweils rund 53 Prozent beträgt. Diese überproportionale Häufigkeit an weiblichen Befragten ist im Sechs-Wellen-Panel  $P_{t_{4,5,6,7,8,9}}$  nochmals stärker ausgeprägt: Frauen sind hier mit jeweils rund 62 Prozent zu allen betrachteten Zeitpunkten im Längsschnittdatensatz stärker vertreten als Männer.

Der in Tabelle 5.74 dargestellte Altersdurchschnitt in den einzelnen Datensätzen lässt keine problematischen Abweichungen erkennen. Die Abweichung im Durchschnittsalter zwischen dem Sechs-Wellen-Panel und den Querschnitten beträgt zu allen Zeitpunkten 0,1 bis 0,2 Jahre dahingehend, dass das Alter im Querschnitt um diesen Wert höher liegt. Nur zum Zeitpunkt  $t_5$  fällt diese Differenz mit 0,4 geringfügig größer aus. Zudem liegt zu diesem Zeitpunkt im Querschnitt auch eine etwas höhere Standardabweichung als in

**Tabelle 5.73:** Zusammensetzung nach Geschlecht  $P_{t_{4,5,6,7,8,9}}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_4$	$Q_{t_4}$	1 717	50,4	1 688	49,6	3 405
	$P_{t_{4,5,6,7,8,9}}$	603	38,0	983	62,0	1 586
$t_5$	$Q_{t_5}$	2 265	49,8	2 282	50,2	4 547
	$P_{t_{4,5,6,7,8,9}}$	603	38,0	983	62,0	1 586
$t_6$	$Q_{t_6}$	1 567	47,0	1 768	53,0	3 335
	$P_{t_{4,5,6,7,8,9}}$	601	37,9	985	62,1	1 586
$t_7$	$Q_{t_7}$	1 452	47,1	1 634	52,9	3 086
	$P_{t_{4,5,6,7,8,9}}$	604	38,1	982	61,9	1 586
$t_8$	$Q_{t_8}$	1 446	46,8	1 644	53,2	3 090
	$P_{t_{4,5,6,7,8,9}}$	603	38,0	983	62,0	1 586
$t_9$	$Q_{t_9}$	1 423	46,8	1 620	53,2	3 043
	$P_{t_{4,5,6,7,8,9}}$	605	38,2	981	61,9	1 586

den Vergleichsdaten vor, was darauf hinweist, dass in diesem Datensatz das Alter eine etwas geringere Homogenität aufweist. Selbe Tendenz gilt auch für den Querschnitt des Erhebungszeitpunktes  $t_6$ .

Auch bei der Verteilung nach der Schulform lassen sich Abweichungen des Panels  $P_{t_{4,5,6,7,8,9}}$  zu den Querschnitten der jeweiligen Jahre feststellen (Tabelle 5.75). Im ersten hier betrachteten Zeitpunkt  $t_4$  liegt mit 14,3 Prozent der Anteil an Hauptschülerinnen und -schülern gut sieben Prozent unter dem Wert des Querschnitts. Dementgegen liegt der Anteil an Personen an Gymnasien knapp neun Prozent höher. Im Folgejahr nehmen die Diskrepanzen weiter zu: Im Sechs-Wellen-Panel ist die Schulform Gymnasium mit über 16 Prozent Differenz stärker repräsentiert als im Querschnitt, Berufskollegs hingegen mit rund 22 Prozent weniger. Die Differenzen sind zum Zeitpunkt  $t_5$  jedoch aufgrund besagter Stichprobenerweiterung um Berufskollegs streng genommen nicht aussagekräftig. Mit den weiteren Erhebungswellen und abnehmender Relevanz der Schule reduzieren sich auch diese Abweichungen zwischen Quer- und Längsschnitt. Zum Zeitpunkt  $t_8$  zeigen sich nur noch marginale Diskrepanzen.

**Tabelle 5.74:** Altersdurchschnitt  $P_{t_{4,5,6,7,8,9}}$

		Alter	Std.-abw.	gesamt
$t_4$	$Q_{t_4}$	16,1	0,8	3313
	$P_{t_{4,5,6,7,8,9}}$	15,9	0,7	1572
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_{4,5,6,7,8,9}}$	17,0	0,8	1579
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{t_{4,5,6,7,8,9}}$	17,9	0,7	1583
$t_7$	$Q_{t_7}$	19,1	0,8	3081
	$P_{t_{4,5,6,7,8,9}}$	18,9	0,7	1586
$t_8$	$Q_{t_8}$	20,1	0,8	3073
	$P_{t_{4,5,6,7,8,9}}$	19,9	0,7	1581
$t_9$	$Q_{t_9}$	22,1	0,9	3021
	$P_{t_{4,5,6,7,8,9}}$	22,0	0,7	1582

**Tabelle 5.75:** Zusammensetzung  $P_{t_{4,5,6,7,8,9}}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							keine	gesamt (100 %)
		HS	RS	GS	GY	BK	SO	andere		
$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_{4,5,6,7,8,9}}$	14,3	23,0	31,0	30,8	0,3	0,3	0,3	0,1	1583
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{4,5,6,7,8,9}}$	0,6	0,5	21,0	36,7	41,0	0,1	—	—	1472
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{4,5,6,7,8,9}}$	—	—	17,4	32,1	30,5	—	10,4	9,5	1554
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{4,5,6,7,8,9}}$	—	—	15,4	30,4	26,2	—	13,1	14,9	1557
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{4,5,6,7,8,9}}$	—	—	3,6	4,0	22,5	—	19,5	50,4	1534

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.6.4 Sechs-Wellen-Panel 2006, 2007, 2008, 2009, 2011 und 2013

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

Das sechs Wellen umfassende Panel  $P_{t_{5,6,7,8,9,10}}$  beinhaltet 1 567 Fälle. Wie in tabelle 5.76 ersichtlich wird, liegt eine Verzerrung dieses Panels bezüglich des Merkmals Geschlecht vor. Zum Zeitpunkt  $t_5$  liegt im Querschnitt ein nahezu gleicher Anteil an Frauen und Männern vor. In den weiteren Erhebungswellen verschiebt sich dieses Verhältnis, sodass der Anteil weiblicher Befragter in den Querschnittsdatensätzen rund 53 bis 54 beträgt. Im Sechs-Wellen-Panel  $P_{t_{5,6,7,8,9,10}}$  liegt dieser Anteil zu allen Zeitpunkten bei über 63 Prozent und somit folglich jeweils etwa zehn Prozent höher als in den Querschnitten der jeweiligen Jahre.

**Tabelle 5.76:** Zusammensetzung nach Geschlecht  $P_{t_{5,6,7,8,9,10}}$

		männlich		weiblich		gesamt (100 %)
		Hfk	%	Hfk	%	
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_{5,6,7,8,9,10}}$	573	36,6	994	63,4	1567
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{t_{5,6,7,8,9,10}}$	573	36,6	994	63,4	1567
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_{5,6,7,8,9,10}}$	574	36,6	993	63,4	1567
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{t_{5,6,7,8,9,10}}$	573	36,6	994	63,4	1567
$t_9$	$Q_{t_9}$	1423	46,8	1620	53,2	3043
	$P_{t_{5,6,7,8,9,10}}$	577	36,8	990	63,2	1567
$t_{10}$	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_{5,6,7,8,9,10}}$	574	36,6	993	63,4	1567

Tabelle 5.77 stellt das durchschnittliche Alter der Querschnittsdatensätze und des Sechs-Wellen-Panels gegenüber. Es lässt sich dabei festhalten, dass das Durchschnittsalter der Befragten – wie auch bereits in den Abschnitten zuvor – im Panel  $P_{t_{5,6,7,8,9,10}}$  mit einer Abweichung von 0,1 bis 0,2 Jahren geringfügig unter dem jeweiligen Vergleichsmittelwert für die Querschnitte liegt. Zusätzlich ist erkennbar, dass das Befragtenalter im Sechs-Wellen-Panel etwas homogener verteilt ist als in den Querschnitten. Dies wird insbesondere in den Zeitpunkten  $t_5$ ,  $t_6$  und  $t_{10}$  deutlich: Hier weisen die Querschnittsdatensätze gegenüber dem Panel eine höhere Standardabweichung der Altersverteilung auf.

Die Zusammensetzung nach Schulform ist für das Teilpanel  $P_{t_{5,6,7,8,9,10}}$  in Tabelle 5.78 nur bis einschließlich 2009 (Zeitpunkt  $t_8$ ) abgebildet. Eine Auswertung der Schulform

**Tabelle 5.77:** Altersdurchschnitt  $P_{t_5,6,7,8,9,10}$

		Alter	Std.-abw.	gesamt
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_5,6,7,8,9,10}$	17,0	0,8	1556
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{t_5,6,7,8,9,10}$	17,9	0,8	1563
$t_7$	$Q_{t_7}$	19,1	0,8	3081
	$P_{t_5,6,7,8,9,10}$	18,9	0,8	1567
$t_8$	$Q_{t_8}$	20,1	0,8	3073
	$P_{t_5,6,7,8,9,10}$	20,0	0,7	1566
$t_9$	$Q_{t_9}$	22,1	0,9	3021
	$P_{t_5,6,7,8,9,10}$	22,0	0,8	1565
$t_{10}$	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_5,6,7,8,9,10}$	23,9	0,8	1564

erübrigt sich ab diesem Zeitpunkt, da die meisten Befragten nach dem 20. Lebensjahr ihre schulische Einrichtung verlassen haben. Die in der Tabelle aufgeschlüsselte Zusammensetzung umfasst entsprechend den Zeitraum von durchschnittlich 17 bis 20 Jahren. Zu diesen dargestellten Zeitpunkten besuchen nur noch sehr wenige bzw. keine Befragungspersonen mehr eine Haupt-, Real- oder Sonderschule. Die relevanten Schulformen in dieser Altersspanne sind bis  $t_7$  vorrangig Gesamtschulen, Gymnasien und Berufskollegs; in  $t_8$  besucht bereits die Hälfte aller Befragungspersonen keine schulische Einrichtung mehr.

Die Kontrastierung von Querschnitt und Sechs-Wellen-Panel zeigt eine systematische Verzerrung nach der besuchten Schulform. Zum Erhebungszeitpunkt  $t_5$  beträgt der Anteil an Gymnasien im Querschnitt  $Q_{t_5}$  rund 20 Prozent. Im Sechs-Wellen-Panel liegt dieser Anteil mit knapp 38 Prozent deutlich darüber. Ein umgekehrtes Verhältnis besteht bei der Schulform des Berufskollegs: Während der Anteil zu  $t_5$  im Querschnitt bei rund 63 Prozent liegt, sind im Panel  $P_{t_5,6,7,8,9,10}$  nur etwa 40 Prozent Berufsschülerinnen und -schüler vertreten.

Diese Diskrepanz nimmt mit den weiteren Erhebungswellen ab. So beträgt die Abweichung bei der Schulform Gymnasium in  $t_7$  nur noch rund neun Prozent, bei den Berufskollegs nur noch drei Prozentpunkte. Gleichzeitig wird jedoch die Ausprägung, keine Schule mehr zu besuchen, zunehmend relevanter und auch hier liegt eine Verzerrung vor: Im Sechs-Wellen-Panel liegt beispielsweise in  $t_7$  der Anteil dieser Kategorie knapp sechs Prozent unter dem Anteil im Querschnitt. Diese Über- bzw. Unterrepräsentanz der jeweiligen Schulformen mindert sich im weiteren Verlauf der Befragung.

Zum Zeitpunkt  $t_8$  weisen Quer- und Längsschnittdaten nur noch geringe Abweichungen zueinander auf.

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

		Schulform <sup>a</sup> (%)							gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{5,6,7,8,9,10}}$	0,2	0,4	21,9	37,7	39,8	0,1	—	—	1454
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{5,6,7,8,9,10}}$	—	—	18,1	33,4	29,4	—	9,8	9,3	1541
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{5,6,7,8,9,10}}$	—	—	16,1	31,7	25,1	—	12,3	14,9	1534
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{5,6,7,8,9,10}}$	—	—	3,9	4,2	22,1	—	19,4	50,5	1516

<sup>a</sup>HS=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 nur jene Datensätze beschrieben, die Teilnahmen in sieben direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke). Ebenso sind Datensätze enthalten, die aufgrund der Änderung des Erhebungsrhythmus in einem Abstand von zwei Jahren erhoben wurden.

### 5.7.1 Sieben-Wellen-Panel 2002, 2003, 2004, 2005, 2006, 2007 und 2008 ( $P_{t_{1,2,3,4,5,6,7}}$ )

Das Sieben-Wellen-Panel  $P_{t_{1,2,3,4,5,6,7}}$  enthält 1 189 Fälle. Die in Tabelle 5.79 dargestellten Zahlen zur Zusammensetzung nach Geschlecht lassen eine systematische Verzerrung dieses Teilpanels erkennen.

In den Erhebungsjahren 2002, 2003 und 2004 ( $t_1$ ,  $t_2$  und  $t_3$ ) ist ein Vergleich mit der Schulstatistik möglich. Hierbei ist zu erkennen, dass sich der Anteil an Schülerinnen zu allen drei Zeitpunkten auf 49 Prozent beläuft, der Anteil an Schülern entsprechend auf rund 51 Prozent. In den Querschnitten  $Q_{t_1}$  und  $Q_{t_2}$  weicht diese Verteilung nur in sehr geringem Maße von den Kennziffern der Schulstatistik ab. In den weiteren

Erhebungszeitpunkten lässt sich jedoch für die Querschnittsdaten feststellen, dass sich der Anteil weiblicher Befragter in der Stichprobe leicht erhöht und in  $t_6$  und  $t_7$  letztlich bei rund 53 Prozent liegt. Eine stärkere Verschiebung wird bei der Gegenüberstellung des Sieben-Wellen-Panels mit den Kennziffern aus Schulstatistik und Querschnitten sichtbar: Der Anteil weiblicher Befragungspersonen liegt für alle Erhebungszeitpunkte durchgängig bei etwa 62 Prozent und weist damit eine bedeutende Differenz zu den Vergleichsdaten auf.

**Tabelle 5.79:** Zusammensetzung nach Geschlecht  $P_{t_{1,2,3,4,5,6,7}}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_1$	$S_{t_1}$	2003	51,2	1907	48,8	3910
	$Q_{t_1}$	1728	50,7	1679	49,3	3407
	$P_{t_{1,2,3,4,5,6,7}}$	448	37,7	741	62,3	1189
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{t_{1,2,3,4,5,6,7}}$	454	38,2	735	61,8	1189
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_{1,2,3,4,5,6,7}}$	448	37,7	741	62,3	1189
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_{1,2,3,4,5,6,7}}$	452	38,0	737	62,0	1189
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_{1,2,3,4,5,6,7}}$	452	38,0	737	62,0	1189
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{t_{1,2,3,4,5,6,7}}$	452	38,0	737	62,0	1189
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_{1,2,3,4,5,6,7}}$	454	38,2	735	61,8	1189

Vergleicht man das Durchschnittsalter der Befragten zu den einzelnen Zeitpunkten zwischen Quer- und Längsschnittdaten (vgl. Tabelle 5.80), so wird ersichtlich, dass das durchschnittliche Alter im Sieben-Wellen-Panel geringfügig unter dem der Querschnitte liegt. Zu den einzelnen Zeitpunkten sind Abweichungen von 0,1 bis 0,3 Jahren erkennbar. In der Erhebungswelle  $t_5$  fällt diese Differenz etwas höher aus – hier liegt das Durchschnittsalter im Querschnitt  $Q_{t_5}$  ein halbes Jahr höher als im Sieben-Wellen-Panel. Außerdem lässt sich festhalten, dass im Panel  $P_{t_{1,2,3,4,5,6,7}}$  eine etwas homogenere Altersverteilung vorliegt als in den Querschnitten, was an den jeweils kleineren Standardabweichungen zu erkennen ist. Dieser Unterschied wird insbesondere zu den Zeitpunkten

$t_5$  und  $t_6$  deutlich; hier weist die Altersverteilung die vergleichsweise größte Streuung auf.

**Tabelle 5.80:** Altersdurchschnitt  $P_{t_1,2,3,4,5,6,7}$

		Alter	Std.-abw.	gesamt
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{t_1,2,3,4,5,6,7}$	12,9	0,7	1 183
$t_2$	$Q_{t_2}$	14,0	0,8	3 270
	$P_{t_1,2,3,4,5,6,7}$	13,8	0,7	1 177
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_1,2,3,4,5,6,7}$	14,9	0,7	1 182
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_1,2,3,4,5,6,7}$	15,9	0,7	1 181
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_1,2,3,4,5,6,7}$	16,9	0,7	1 184
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_1,2,3,4,5,6,7}$	17,9	0,7	1 187
$t_7$	$Q_{t_7}$	19,1	0,8	3 081
	$P_{t_1,2,3,4,5,6,7}$	18,8	0,7	1 189

Eine vergleichende Betrachtung der Zusammensetzung nach Schulform lässt den Schluss auf eine systematische Verzerrung des Sieben-Wellen-Panels zu (vgl. Tabelle 5.81).

In Referenz auf die Schulstatistiken der Jahre 2002 bis 2004 ( $t_1$  bis  $t_3$ ) wird ersichtlich, dass Schülerinnen und Schüler an Hauptschulen im Sieben-Wellen-Panel mit einer Differenz von rund neun bis zwölf Prozent unterrepräsentiert sind, während Personen an Gymnasien mit einer Abweichung von etwa sieben bis neun Prozentpunkten stärker vertreten sind. Auch bei den anderen Schulformen liegen Abweichungen zwischen Querschnitt und Panel vor, jedoch in geringerem Ausmaß.

Ausgeprägtere Differenzen lassen sich zum Zeitpunkt  $t_5$  feststellen: Der Prozentanteil an Schülerinnen und Schülern der Gesamtschulen liegt verglichen mit dem Querschnitt  $Q_{t_5}$  im Sieben-Wellen-Panel rund sechs Punkte, bei der Schulform Gymnasium 13 Punkte höher. Umgekehrt ist zu diesem Zeitpunkt der Anteil an Berufsschülerinnen und -schülern im Sieben-Wellen-Panel über 19 Prozent geringer als im Querschnitt des Erhebungszeitpunktes  $t_5$ .

Diese Verschiedenheiten in der Zusammensetzung nach Schulform sind in den weiteren Erhebungszeitpunkten weniger stark ausgeprägt. Beispielsweise sind Gymnasiastinnen und Gymnasiasten im Paneldatensatz in der Welle  $t_7$  mit rund fünf Prozent Differenz

häufiger vertreten als im Querschnitt; Personen, die keine Schule besuchen, mit etwa fünf Prozent weniger.

**Tabelle 5.81:** Zusammensetzung  $P_{t_{1,2,3,4,5,6,7}}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							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
	$P_{t_{1,2,3,4,5,6,7}}$	15,0	24,8	32,0	28,2	—	—	—	—	1189
$t_2$	$S_{t_2}$	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
	$P_{t_{1,2,3,4,5,6,7}}$	15,6	25,0	32,0	27,4	—	—	—	—	1189
$t_3$	$S_{t_3}$	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
	$P_{t_{1,2,3,4,5,6,7}}$	15,6	25,3	32,0	27,0	—	—	—	—	1189
$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_{1,2,3,4,5,6,7}}$	15,1	25,3	31,6	27,4	0,4	—	0,3	—	1188
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{1,2,3,4,5,6,7}}$	0,6	0,5	22,4	33,2	43,2	—	—	—	1106
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{1,2,3,4,5,6,7}}$	—	—	18,5	29,0	31,6	—	11,7	9,2	1157
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{1,2,3,4,5,6,7}}$	—	—	15,9	27,4	26,5	—	14,1	16,1	1163

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

### 5.7.2 Sieben-Wellen-Panel 2003, 2004, 2005, 2006, 2007, 2008 und 2009 ( $P_{t_{2,3,4,5,6,7,8}}$ )

Das Sieben-Wellen-Panel der Jahre 2003 bis 2009 beinhaltet 1 365 Fälle. Dieser Teilpaneldatensatz  $P_{t_{2,3,4,5,6,7,8}}$  weist eine Verzerrung bezüglich der Verteilung nach Geschlecht auf, wie sich in Tabelle 5.82 ablesen lässt.

Unter Hinzuziehung der Schulstatistik der Zeitpunkte  $t_2$  und  $t_3$  wird ersichtlich, dass das prozentuale Verhältnis zwischen männlichen und weiblichen Befragten weitgehend ausgewogen ist: Der Anteil an Männern beträgt 51 Prozent, der Anteil an Frauen 49. In den Querschnitten zeigt sich, dass ab der Erhebungswelle  $t_6$  der Anteil an weiblichen

Person fortan bei etwa 53 Prozent liegt. Im Sieben-Wellen-Panel liegt dieser Anteil zu allen Zeitpunkten deutlich höher. Mit jeweils ca. 62 Prozent liegt hier der Frauenanteil wesentlich über den Vergleichswerten von Schulstatistik und Querschnitterhebung.

**Tabelle 5.82:** Zusammensetzung nach Geschlecht  $P_{t_{2,3,4,5,6,7,8}}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{t_{1,2,3,4,5,6,7,8}}$	517	37,9	848	62,1	1365
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_{2,3,4,5,6,7,8}}$	511	37,4	854	62,6	1365
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_{2,3,4,5,6,7,8}}$	516	37,8	849	62,2	1365
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_{2,3,4,5,6,7,8}}$	514	37,7	851	62,3	1365
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{t_{2,3,4,5,6,7,8}}$	514	37,7	851	62,3	1365
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_{2,3,4,5,6,7,8}}$	517	37,9	848	62,1	1365
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{t_{2,3,4,5,6,7,8}}$	516	37,8	849	62,2	1365

Ein Blick auf das Durchschnittsalter der einzelnen Zeitpunkte in Tabelle 5.83 macht deutlich, dass das gemittelte Alter im Querschnitt zu allen Zeitpunkten über dem des Sieben-Wellen-Panels liegt. Die Differenz fällt mit jeweils 0,1 bis 0,2 Jahren jedoch eher gering aus. Lediglich zum Zeitpunkt  $t_5$  ist das Durchschnittsalter im Querschnitt etwas höher und weist eine Differenz von 0,5 Jahren zum Sieben-Wellen-Panel auf. Im selben Zeitpunkt sowie auch bei der darauffolgenden Welle  $t_6$  sind für die Querschnitte  $Q_{t_5}$  und  $Q_{t_6}$  die Standardabweichungen vergleichsweise höher, woraus eine etwas geringere Altershomogenität zu schlussfolgern ist. Generell wird zudem ersichtlich, dass sich die Befragten bezüglich ihres Alters im Paneldatensatz  $P_{t_{2,3,4,5,6,7,8}}$  ähnlicher sind als in den Querschnitten.

Die Zusammensetzung des Sieben-Wellen-Panels  $P_{t_{2,3,4,5,6,7,8}}$  nach der Schulform kann Tabelle 5.84 entnommen werden. Ein Vergleich mit der Schulstatistik zu den Zeitpunkten  $t_2$  und  $t_3$  zeigt auf, dass sich vor allem bei den Schulformen Hauptschule und Gymnasium Verzerrungen des Längsschnittdatensatzes bemerkbar machen. Im Sieben-Wellen-Panel sind beispielsweise in  $t_2$  zwölf Prozent weniger Schülerinnen und Schüler an Hauptschu-

**Tabelle 5.83:** Altersdurchschnitt  $P_{t_2,3,4,5,6,7,8}$

		Alter	Std.-abw.	gesamt
$t_2$	$Q_{t_2}$	14,0	0,8	3270
	$P_{t_2,3,4,5,6,7,8}$	13,9	0,7	1350
$t_3$	$Q_{t_3}$	15,1	0,8	3270
	$P_{t_2,3,4,5,6,7,8}$	14,9	0,7	1353
$t_4$	$Q_{t_4}$	16,1	0,8	3313
	$P_{t_2,3,4,5,6,7,8}$	15,9	0,7	1353
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_2,3,4,5,6,7,8}$	16,9	0,8	1358
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{t_2,3,4,5,6,7,8}$	17,9	0,7	1363
$t_7$	$Q_{t_7}$	19,1	0,8	3081
	$P_{t_2,3,4,5,6,7,8}$	18,9	0,7	1365
$t_8$	$Q_{t_8}$	20,1	0,8	3073
	$P_{t_2,3,4,5,6,7,8}$	19,9	0,7	1360

len vertreten als die Schulstatistik ausweist. Umgekehrt liegen im Datensatz  $P_{t_2,3,4,5,6,7,8}$  ca. zehn Prozent mehr Gymnasiastinnen und Gymnasiasten vor als in der offiziellen Statistik. Dementsprechend lässt sich festhalten, dass eine Verschiebung vorliegt in der Hinsicht, dass bis zum Zeitpunkt  $t_4$  höhere Schulformen im Panel über- und niedrigere Schulformen unterrepräsentiert sind.

Derart gerichtete Verzerrungen lassen sich auch bei den weiteren Zeitpunkten identifizieren, da ab  $t_5$  jedoch vorrangig Haupt-, Real- und Sonderschulen aufgrund des Alters der Befragten keine Rolle mehr spielen, zeigt sich die Diskrepanz zwischen Quer- und Längsschnitt an anderen Stellen.

Zum Erhebungszeitpunkt  $t_5$  liegt der Anteil an Personen, die angeben ein Gymnasium zu besuchen, etwa 13 Prozentpunkte über dem Wert für den Querschnitt. Ähnliches gilt für die Gesamtschulen, wobei hier die Differenz mit rund sechs Prozent geringer ausfällt. Dem entgegengesetzt ist der Anteil an Schülerinnen und Schülern an Berufskollegs im Sieben-Wellen-Panel deutlich geringer als im Querschnitt: Im Panel beträgt dieser Anteil etwa 43 Prozent während der Querschnitt  $Q_{t_5}$  zum selben Zeitpunkt knapp 63 Prozent beinhaltet. Die Differenzen zu diesem Zeitpunkt sind aufgrund der erwähnten Ausweitung der Stichprobe jedoch nur begrenzt aussagekräftig.

In den weiteren Zeitpunkten  $t_6$ ,  $t_7$  und  $t_8$  verlassen zunehmend mehr Personen die schulischen Bildungseinrichtungen – in  $t_8$  liegt der Anteil an Personen, die keine Schulform angeben, bereits bei 50 bzw. 51 Prozent in Längsschnitt- und Querschnitt. Die Differenzen zwischen den Querschnitten und dem Sieben-Wellen-Panel bezüglich der

Verteilung nach Schulform sind in diesen späteren Zeitpunkten nur noch geringfügig vorhanden.

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

		Schulform <sup>a</sup> (%)							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
	$P_{t_{2,3,4,5,6,7,8}}$	15,5	23,9	32,2	28,4	—	—	—	—	1365
$t_3$	$S_{t_3}$	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
	$P_{t_{2,3,4,5,6,7,8}}$	15,6	24,0	32,5	27,9	—	—	—	—	1365
$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_{2,3,4,5,6,7,8}}$	14,8	23,9	32,3	28,2	0,5	—	0,3	—	1364
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{2,3,4,5,6,7,8}}$	0,7	0,6	22,3	33,5	42,9	—	—	—	1274
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{2,3,4,5,6,7,8}}$	—	—	18,7	29,8	31,5	—	10,6	9,5	1329
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{2,3,4,5,6,7,8}}$	—	—	16,5	28,2	26,6	—	13,5	15,2	1335
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{2,3,4,5,6,7,8}}$	—	—	4,1	3,9	22,1	—	19,5	50,4	1317

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

### 5.7.3 Sieben-Wellen-Panel 2004, 2005, 2006, 2007, 2008, 2009 und 2011 ( $P_{3,4,5,6,7,8,9}$ )

Das Sieben-Wellen-Panel  $P_{3,4,5,6,7,8,9}$  umfasst Daten der Jahre 2004 bis 2011 und beinhaltet insgesamt 1 424 Beobachtungen. Mit Bezug auf die Daten aus Querschnitt und Schulstatistik wird für dieses Sieben-Wellen-Panel eine Verzerrung bezüglich der Verteilung nach Geschlecht erkennbar (vgl. Tabelle 5.85).

Die Schulstatistik zum Zeitpunkt  $t_3$  weist ein annähernd ausgewogenes Verhältnis von Männern und Frauen mit 51 zu 49 Prozent aus. In den Querschnittsdatensätzen verhält es sich innerhalb der Erhebungszeitpunkte  $t_4$  und  $t_5$  ähnlich ausgeglichen. Mit fortschreitender Erhebung erhöht sich der Anteil weiblicher Befragter im Querschnitt

jedoch und liegt ab  $t_6$  konstant bei rund 53 Prozent. Eine derartige Überrepräsentanz weiblicher Personen, die im Querschnitt nur leicht vorliegt, ist im Sieben-Wellen-Panel bedeutsam stärker ausgeprägt. Die Quote an Personen weiblichen Geschlechts liegt zu allen Zeitpunkten bei rund 63 Prozent und somit je nach Welle zwischen zehn und 13 Prozentpunkten über den Vergleichswerten der betrachteten Querschnitte.

**Tabelle 5.85:** Zusammensetzung nach Geschlecht  $P_{t_3,4,5,6,7,8,9}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_3,4,5,6,7,8,9}$	527	37,0	897	63,0	1424
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_3,4,5,6,7,8,9}$	531	37,3	893	62,7	1424
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_3,4,5,6,7,8,9}$	531	37,3	893	62,7	1424
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{t_3,4,5,6,7,8,9}$	529	37,2	895	62,9	1424
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_3,4,5,6,7,8,9}$	523	37,4	892	62,6	1424
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{t_3,4,5,6,7,8,9}$	531	37,3	893	62,7	1424
$t_9$	$Q_{t_9}$	1423	46,8	1620	53,2	3043
	$P_{t_3,4,5,6,7,8,9}$	531	37,3	893	62,7	1424

Die Berechnung des durchschnittlichen Alters in den unterschiedlichen Datensätzen (vgl. Tabelle 5.86) macht deutlich, dass das gemittelte Alter im Querschnitt durchgehend leicht über dem Durchschnittsalter im Sieben-Wellen-Panel liegt. Die Differenz ist mit 0,1 bis 0,2 jedoch als eher gering zu bewerten. Einzig zum Zeitpunkt  $t_5$  fällt diese Differenz mit 0,4 Jahren etwas höher aus. Die Standardabweichung weist darauf hin, dass die Altersverteilung in den Querschnitten etwas heterogener ist als im Panel  $P_{t_3,4,5,6,7,8,9}$ . Besonders in den Querschnitten  $Q_{t_5}$  und  $Q_{t_6}$  ist die Variation in den Angaben des Alters der Befragten gegenüber dem Panel erhöht.

Die Gegenüberstellung des Sieben-Wellen-Panels mit der Schulstatistik bzw. den Querschnittserhebungen (vgl. Tabelle 5.87) zeigt auch für das Teilpanel  $P_{t_3,4,5,6,7,8,9}$  eine systematische Verzerrung nach der Schulform, wie sie auch bereits in den vorherigen Abschnitten beschrieben wurde.

Unter Hinzuziehung der Schulstatistik zum Zeitpunkt  $t_3$  lässt sich insbesondere festhalten, dass der Anteil an Personen, die eine Hauptschule besuchen, im betrachteten

**Tabelle 5.86:** Altersdurchschnitt  $P_{t_{3,4,5,6,7,8,9}}$

		Alter	Std.-abw.	gesamt
$t_3$	$Q_{t_3}$	15,1	0,8	3270
	$P_{t_{3,4,5,6,7,8,9}}$	15,0	0,7	1415
$t_4$	$Q_{t_4}$	16,1	0,8	3313
	$P_{t_{3,4,5,6,7,8,9}}$	15,9	0,7	1414
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_{3,4,5,6,7,8,9}}$	17,0	0,8	1419
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{t_{3,4,5,6,7,8,9}}$	17,9	0,7	1422
$t_7$	$Q_{t_7}$	19,1	0,8	3081
	$P_{t_{3,4,5,6,7,8,9}}$	18,9	0,7	1424
$t_8$	$Q_{t_8}$	20,1	0,8	3073
	$P_{t_{3,4,5,6,7,8,9}}$	19,9	0,7	1420
$t_9$	$Q_{t_9}$	22,1	0,9	3021
	$P_{t_{3,4,5,6,7,8,9}}$	21,9	0,7	1420

Sieben-Wellen-Panel zehn Prozentpunkte geringer liegt als im Vergleichsdatensatz  $S_{t_3}$ . Der Anteil der Schulform Gymnasium liegt hingegen elf Prozentpunkte höher.

Ab dem Zeitpunkt  $t_4$  sind nur noch Vergleiche mit den Querschnitten der jeweiligen Wellen möglich, doch auch hier bleiben derart gerichtete Abweichungen bestehen: Im Sieben-Wellen-Panel liegt für den Zeitpunkt  $t_4$  der Anteil an Hauptschülerinnen und -schülern ca. acht Prozent niedriger als im Querschnitt  $Q_{t_4}$ , umgekehrt der Anteil an Gymnasiastinnen und Gymnasiasten im Panel neun Prozent höher. Ab  $t_5$  fallen Haupt-, Real- und Sonderschulen als besuchte Schulform weg.

Stattdessen liegt eine Unterrepräsentanz, wie sie bisher für die Hauptschulen zu bemerken war, ab diesem Zeitpunkt für die Berufskollegs vor. In Erhebungswelle  $t_5$  sind nur rund 41 Prozent Schülerinnen und Schüler an Berufskollegs im Panel vertreten, wohingegen dieser Anteil für den Querschnitt  $Q_{t_5}$  nahezu 63 Prozent beträgt, was eine merkliche Abweichung von über 21 Prozentpunkten bedeutet. Diese Differenzen sollte jedoch mit Vorsicht betrachtet werden, da zum Zeitpunkt  $t_5$  eine temporäre Ausweitung der Stichprobe vorgenommen wurde, was auch an der erhöhten Fallzahl des Querschnitts  $Q_{t_5}$  ersichtlich ist. Eine umgekehrte Differenz gilt für Gymnasien und Gesamtschulen: Erstere weisen im Vergleich von Querschnitt und Sieben-Wellen-Panel in  $t_5$  eine Differenz von rund 16 Prozent auf, zweitere fünf Prozent.

Zu zunehmenden Alter der Befragten und der damit einhergehenden abnehmenden Relevanz der Schule nähern sich auch die Anteilswerte aus Querschnitt und Sieben-Wellen-Panel an: Zum Erhebungszeitpunkt  $t_8$  lassen sich nur noch geringfügige Abwei-

chungen zwischen den Betrachteten Datensätzen ausfindig machen. Eine Analyse der Schulform für den Zeitpunkt  $t_9$  erübrigt sich, da schulische Einrichtungen in diesem Alter (durchschnittlich 22 Jahre) von den Befragten kaum noch besucht werden.

**Tabelle 5.87:** Zusammensetzung  $P_{3,4,5,6,7,8,9}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								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
	$P_{t_{3,4,5,6,7,8,9}}$	14,8	23,6	30,6	31,0	—	—	—	—	1424
$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_{3,4,5,6,7,8,9}}$	13,9	23,7	30,7	31,0	0,4	—	0,4	0,1	1421
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{3,4,5,6,7,8,9}}$	0,6	0,6	21,2	36,5	41,1	—	—	—	1320
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{3,4,5,6,7,8,9}}$	—	—	17,4	32,1	30,0	—	10,8	9,8	1394
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{3,4,5,6,7,8,9}}$	—	—	15,4	30,7	25,9	—	13,1	15,0	1396
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{3,4,5,6,7,8,9}}$	—	—	3,7	3,9	22,3	—	19,6	50,4	1376

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

#### 5.7.4 Sieben-Wellen-Panel 2005, 2006, 2007, 2008, 2009, 2011 und 2013 ( $P_{t_{4,5,6,7,8,9,10}}$ )

Das sieben Wellen umfassende Panel  $P_{t_{4,5,6,7,8,9,10}}$  enthält 1 415 Fälle. Beim Vergleich dieser Längsschnittdaten mit denen des Querschnitts wird eine Verzerrung des Sieben-Wellen-Panels nach Geschlecht deutlich (Tabelle 5.88).

Zu den Zeitpunkten  $t_4$  und  $t_5$  liegt ein nahezu gleiches Verhältnis an männlichen und weiblichen Befragungspersonen vor. Ab der Welle  $t_6$  lässt sich für die Querschnitte ein leicht erhöhter Anteil an Frauen von 53 bis 54 Prozent feststellen. Eine deutlichere Überrepräsentierung von Frauen lässt sich jedoch vor allem im Sieben-Wellen-Panel bemerken: Der Anteil weiblicher Personen liegt für das Panel  $P_{t_{4,5,6,7,8,9,10}}$  zu allen Zeitpunkten bei rund 63 bis 64 Prozent und somit je nach Zeitpunkt bis zu 14 Prozent über den Vergleichswerten des Querschnitts.

**Tabelle 5.88:** Zusammensetzung nach Geschlecht  $P_{t_4,5,6,7,8,9,10}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_4,5,6,7,8,9,10}$	515	36,4	900	63,6	1415
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_4,5,6,7,8,9,10}$	515	36,4	900	63,6	1415
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{t_4,5,6,7,8,9,10}$	513	36,3	902	63,8	1415
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_4,5,6,7,8,9,10}$	516	36,5	899	63,5	1415
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{t_4,5,6,7,8,9,10}$	515	36,4	900	63,6	1415
$t_9$	$Q_{t_9}$	1423	46,8	1620	53,2	3043
	$P_{t_4,5,6,7,8,9,10}$	519	36,7	896	63,3	1415
$t_{10}$	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_4,5,6,7,8,9,10}$	514	36,3	901	63,7	1415

Bei der Gegenüberstellung von Querschnitten und Sieben-Wellen-Panel hinsichtlich des durchschnittlichen Alters (Tabelle 5.89) lässt sich zusammenfassen, dass der Mittelwert des Befragtenalters zu allen Zeitpunkten im Querschnitt geringfügig höher liegt als im betrachteten Panel. Die Differenzen fallen mit 0,1 bis 0,2 jedoch eher marginal aus. Lediglich zum Zeitpunkt  $t_5$  lässt sich eine etwas größere Abweichung von 0,4 Jahren feststellen. Es ist außerdem zu erkennen, dass das Sieben-Wellen-Panel im Vergleich zu den Querschnitten durchgängig eine etwas stärkere Altershomogenität aufweist; die Standardabweichungen der Altersverteilungen sind in den Querschnitten zu allen Erhebungszeitpunkten höher. Insbesondere innerhalb der Zeitpunkte  $t_5$  und  $t_6$  liegt für die Querschnitte  $Q_{t_5}$ ,  $Q_{t_6}$  sowie  $Q_{t_{10}}$  eine stärkere Heterogenität des Befragtenalters vor.

Die Tabelle 5.90 zeigt die Zusammensetzung des Sieben-Wellen-Panels  $P_{t_4,5,6,7,8,9,10}$  nach der angegebenen Schulform und vergleicht diese mit den Querschnittsdaten der jeweiligen Jahre. Dabei sind die Zeitpunkte  $t_9$  und  $t_{10}$  nicht dargestellt, da sich in diesem Alter (durchschnittlich 22 Jahre in  $t_9$  und 24 Jahre in  $t_{10}$ ) so gut wie keine Befragten mehr an Schulen befinden.

Unter Betrachtung der Anteilsdifferenzen zwischen Längs- und Querschnitt fällt eine Verzerrung des Sieben-Wellen-Panels hinsichtlich der Verteilung nach Schulform auf. So befinden sich beispielsweise zum Zeitpunkt  $t_4$  im Querschnitt  $Q_{t_4}$  knapp 22 Prozent der Personen an einer Hauptschule. Im Panel beträgt dieser Anteil zum selben Zeitpunkt hingegen nur knapp 13 Prozent. Umgekehrt verhält es sich mit der Schulform

**Tabelle 5.89:** Altersdurchschnitt  $P_{t_4,5,6,7,8,9,10}$

		Alter	Std.-abw.	gesamt
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_4,5,6,7,8,9,10}$	15,9	0,7	1 404
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_4,5,6,7,8,9,10}$	17,0	0,8	1 409
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_4,5,6,7,8,9,10}$	17,9	0,7	1 413
$t_7$	$Q_{t_7}$	19,1	0,8	3 081
	$P_{t_4,5,6,7,8,9,10}$	18,9	0,7	1 415
$t_8$	$Q_{t_8}$	20,1	0,8	3 073
	$P_{t_4,5,6,7,8,9,10}$	19,9	0,7	1 414
$t_9$	$Q_{t_9}$	22,1	0,9	3 021
	$P_{t_4,5,6,7,8,9,10}$	21,9	0,7	1 414
$t_{10}$	$Q_{t_{10}}$	24,1	1,2	2 839
	$P_{t_4,5,6,7,8,9,10}$	23,9	0,8	1 413

Gymnasium: Der Anteil im Querschnitt  $Q_{t_4}$  beträgt 22 Prozent, während das Panel einen Anteil von rund 32 Prozent verzeichnet.

Im darauffolgenden Zeitpunkt  $t_5$  zeigt sich eine derartige Verschiebung bei der ab diesem Alter (17 Jahre) nun relevanten Schulform des Berufskollegs. Im Sieben-Wellen-Panel sind rund 22 Prozent weniger Schülerinnen und Schüler dieser Schulform vertreten als im Querschnitt  $Q_{t_5}$ . Weiterhin erhöht ist der Anteil jener Personen, die ein Gymnasium besuchen: Dieser Anteil liegt im Panel  $P_{t_4,5,6,7,8,9,10}$  um fast 18 Prozentpunkte höher als im Querschnitt.

Derartige Differenzen einer Überrepräsentanz von insbesondere Gymnasien und umgekehrt einer Unterrepräsentanz anderer Schulformen liegen auch in den weiteren Zeitpunkten vor, sind dort jedoch zunehmend schwächer ausgeprägt. Zum Erhebungszeitpunkt  $t_8$  lassen sich nur noch sehr leichte Abweichungen zwischen Quer- und Längsschnittdaten festhalten.

**Tabelle 5.90:** Zusammensetzung  $P_{4,5,6,7,8,9,10}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								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
	$P_{t_{4,5,6,7,8,9,10}}$	12,8	23,4	31,2	31,9	0,3	0,2	0,3	—	1 412
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4 253
	$P_{t_{4,5,6,7,8,9,10}}$	0,2	0,4	21,0	38,1	40,3	0,1	—	—	1 308
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3 219
	$P_{t_{4,5,6,7,8,9,10}}$	—	—	17,7	33,1	29,9	—	10,4	8,9	1 391
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2 975
	$P_{t_{4,5,6,7,8,9,10}}$	—	—	15,7	31,5	25,8	—	12,7	14,3	1 389
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2 934
	$P_{t_{4,5,6,7,8,9,10}}$	—	—	3,7	4,2	22,6	—	19,2	50,4	1 372

<sup>a</sup>HS=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 nur jene Datensätze beschrieben, die Teilnahmen in acht direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke). Dabei sind auch die Datensätze enthalten, die aufgrund der Änderung des Erhebungsrhythmus in einem Abstand von zwei Jahren erhoben wurden.

### 5.8.1 Acht-Wellen-Panel 2002, 2003, 2004, 2005, 2006, 2007, 2008 und 2009 ( $P_{t_{1,2,3,4,5,6,7,8}}$ )

Das Acht-Wellen-Panel der Jahre 2002 bis 2009 umfasst 1 103 Fälle. Die in Tabelle 5.91 dargestellte Zusammensetzung der Stichprobe nach dem Geschlecht legt eine Verzerrung dieses Teilpanels  $P_{t_{1,2,3,4,5,6,7,8}}$  nahe.

Zu den ersten drei Zeitpunkten von 2002 bis 2004 liegen Daten der Schulstatistik vor, die einen Vergleich mit der tatsächlichen Geschlechterverteilung ermöglichen. Aus ihnen geht hervor, dass mit 49 Prozent Frauen- und 51 Prozent Männeranteil das Verhältnis zwischen weiblichen und männlichen befragten mehr oder minder ausgeglichen ist. Dieses Verhältnis spiegelt sich auch in den Querschnittsdatsätzen  $Q_{t_1}$  bis  $Q_{t_5}$  wider; in  $Q_{t_3}$  überwiegen allerdings weibliche Personen mit 51 Prozent. Ab dem Zeitpunkt  $t_6$  sind Frauen im Querschnitt mit jeweils 53 Prozent etwas stärker vertreten als Männer.

Dieser überproportionale Anteil weiblicher Befragter ist insbesondere im Acht-Wellen-Panel stark ausgeprägt. Zu allen Zeitpunkten liegt der Frauenanteil bei rund 63 bis 64 Prozent und somit bis zu 15 Prozentpunkte über den Vergleichswerten aus Querschnitt und Schulstatistik.

**Tabelle 5.91:** Zusammensetzung nach Geschlecht  $P_{1,2,3,4,5,6,7,8}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_1$	$S_{t_1}$	2003	51,2	1907	48,8	3910
	$Q_{t_1}$	1728	50,7	1679	49,3	3407
	$P_{1,2,3,4,5,6,7,8}$	398	36,1	705	63,9	1103
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{1,2,3,4,5,6,7,8}$	405	36,7	698	63,3	1103
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{1,2,3,4,5,6,7,8}$	399	36,2	704	63,8	1103
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{1,2,3,4,5,6,7,8}$	402	36,5	701	63,6	1103
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{1,2,3,4,5,6,7,8}$	403	36,5	700	63,5	1103
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{1,2,3,4,5,6,7,8}$	403	36,5	700	63,5	1103
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{1,2,3,4,5,6,7,8}$	405	36,7	698	63,3	1103
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{1,2,3,4,5,6,7,8}$	404	36,6	699	63,4	1103

In Tabelle 5.92 ist das durchschnittliche Alter in den Querschnitten von 2002 bis 2009 sowie dem Acht-Wellen-Panel zu diesen Zeitpunkten gegenübergestellt. Es fällt auf, dass in allen Erhebungswellen das Durchschnittsalter im Querschnitt etwas über dem des Panels liegt. Dazu sei zu erwähnen, dass die Differenzen mit Werten von meist 0,2 bis 0,3 etwas höher ausfallen als es in den bisherigen Beschreibungen der Paneldatensätze der Fall war – die Differenz betrug dort bislang nur 0,1 bis 0,2. Die Ausnahme bildet wie gehabt der Zeitpunkt  $t_5$ , der ein etwas höheres Durchschnittsalter im Querschnitt aufweist und somit eine ausgeprägtere Differenz von 0,5 Jahren zum Acht-Wellen-Panel. Auch ist an dieser Stelle, ebenso wie im darauffolgenden Zeitpunkt  $t_6$ , die Homogenität des Befragtenalters im Querschnitt etwas geringer als in den Vergleichszeitpunkten.

Dies lässt sich mit der bereits erwähnten temporären Erweiterung der Stichprobe zum Zeitpunkt  $t_5$  begründen.

**Tabelle 5.92:** Altersdurchschnitt  $P_{t_1,2,3,4,5,6,7,8}$

		Alter	Std.-abw.	gesamt
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{t_1,2,3,4,5,6,7,8}$	12,8	0,6	1 097
$t_2$	$Q_{t_2}$	14,0	0,8	3 270
	$P_{t_1,2,3,4,5,6,7,8}$	13,8	0,7	1 093
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_1,2,3,4,5,6,7,8}$	14,9	0,7	1 096
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_1,2,3,4,5,6,7,8}$	15,8	0,7	1 095
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_1,2,3,4,5,6,7,8}$	16,9	0,7	1 099
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_1,2,3,4,5,6,7,8}$	17,8	0,7	1 101
$t_7$	$Q_{t_7}$	19,1	0,8	3 081
	$P_{t_1,2,3,4,5,6,7,8}$	18,8	0,7	1 103
$t_8$	$Q_{t_8}$	20,1	0,8	3 073
	$P_{t_1,2,3,4,5,6,7,8}$	19,9	0,7	1 100

Auch die in Tabelle 5.93 abgebildete Zusammensetzung nach der Schulform macht deutlich, dass in dem Acht-Wellen-Panel  $P_{t_1,2,3,4,5,6,7,8}$  eine Verzerrung hinsichtlich dieses Merkmals vorliegt. Der Vergleich von Schulstatistik und Acht-Wellen-Panel zu den Zeitpunkten  $t_1$ ,  $t_2$  und  $t_3$  macht deutlich, dass Schülerinnen und Schüler an Hauptschulen mit einer Differenz von nahezu zehn Prozent (in  $t_1$ ) im Panel unterrepräsentiert sind. In den Wellen  $t_2$  und  $t_3$  beläuft sich diese Diskrepanz auf 13 bzw. zehn Prozent. Umgekehrt sind in diesem Teilpanel in  $t_1$  drei Prozent mehr Realschülerinnen und -schüler vertreten als die Schulstatistik ausweist sowie knapp acht Prozent mehr Personen an Gymnasien. Ähnliche Verschiedenheiten sind für diese beiden Schulformen auch für  $t_2$  und  $t_3$  auffindbar.

Zum Zeitpunkt  $t_5$  befinden sich kaum noch Personen an Haupt- und Realschulen, sodass eine Betrachtung der Differenzen für diese beiden Schulformen obsolet wird. Allerdings bestehen weiterhin Abweichungen zwischen Querschnitt und Acht-Wellen-Panel für die Schulform Gymnasium: Im Längsschnitt geben 14 Prozent mehr Personen an, diese Schulform zu besuchen. Für die Gesamtschulen sind es gut sechs Prozent mehr.

Umkehrtes Verhältnis gilt für die Personen an Berufskollegs. Dort beträgt der Anteil im Acht-Wellen-Panel 42 Prozent, im Querschnitt  $Q_{t_5}$  hingegen fast 63 Prozent. Dieser

deutliche Überschuss im Querschnitt ist auf die temporäre Ausweitung der Stichprobe durch Hinzunahme von Berufsschulen zurückzuführen.

Im weiteren Verlauf der Erhebung geben immer mehr Befragte an, keine Schule mehr zu besuchen, sodass die bisherige Verzerrung nach der Schulform zunehmend unwesentlicher wird. Im letzten betrachteten Zeitpunkt  $t_8$  sind die auftretenden Differenzen zwischen den Datensätzen kaum mehr relevant.

**Tabelle 5.93:** Zusammensetzung  $P_{t_{1,2,3,4,5,6,7,8}}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								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
	$P_{t_{1,2,3,4,5,6,7,8}}$	14,1	25,0	32,0	28,9	—	—	—	—	1103
$t_2$	$S_{t_2}$	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
	$P_{t_{1,2,3,4,5,6,7,8}}$	14,5	25,2	32,0	28,3	—	—	—	—	1103
$t_3$	$S_{t_3}$	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
	$P_{t_{1,2,3,4,5,6,7,8}}$	14,6	25,4	32,0	28,0	—	—	—	—	1103
$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_{1,2,3,4,5,6,7,8}}$	13,9	25,4	31,6	28,4	0,5	—	0,3	—	1102
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{1,2,3,4,5,6,7,8}}$	0,7	0,6	22,4	34,2	42,1	—	—	—	1031
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{1,2,3,4,5,6,7,8}}$	—	—	18,5	30,3	31,0	—	11,2	9,0	1074
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{1,2,3,4,5,6,7,8}}$	—	—	16,4	28,5	25,7	—	13,9	15,5	1081
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{1,2,3,4,5,6,7,8}}$	—	—	4,2	4,0	21,2	—	19,9	50,7	1064

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

5.8.2 Acht-Wellen-Panel 2003, 2004, 2005, 2006, 2007, 2008, 2009 und 2011 ( $P_{t_{2,3,4,5,6,7,8,9}}$ )

Das acht Erhebungswellen umfassende Panel  $P_{t_{2,3,4,5,6,7,8,9}}$  beinhaltet 1 222 Fälle. Die in Tabelle 5.94 dargestellte Aufschlüsselung nach dem Merkmal Geschlecht lässt auf eine Verzerrung des Acht-Wellen-Panels schließen.

Die Schulstatistik der Zeitpunkte  $t_2$  und  $t_3$  zeigt mit 49 zu 51 Prozent ein annähernd ausgewogenes Verhältnis zwischen weiblichen und männlichen Befragten auf. Auch in den Querschnitten ist dieses Verhältnis in dieser Form bis  $t_5$  vorzufinden. Ab dem Zeitpunkt  $t_6$  bildet sich jedoch ein leichter Überschuss weiblicher Personen heraus, der fortan bei rund 53 Prozent liegt.

Im betrachteten Acht-Wellen-Panel weicht die Verteilung nach dem Geschlecht deutlich ab. Hier beträgt der Anteil an Personen weiblichen Geschlechts zu allen aufgelisteten Zeitpunkten rund 64 Prozent. Dieser Prozentsatz liegt damit bis zu 15 Prozentpunkte höher als in den Vergleichsdaten auf Querschnittserhebung und Schulstatistik.

**Tabelle 5.94:** Zusammensetzung nach Geschlecht  $P_{t_{2,3,4,5,6,7,8,9}}$

		männlich		weiblich		gesamt (100 %)
		Hfk	%	Hfk	%	
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3 763
	$Q_{t_2}$	1703	50,2	1687	49,8	3 390
	$P_{t_{2,3,4,5,6,7,8,9}}$	445	36,4	777	63,6	1 222
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3 925
	$Q_{t_3}$	1635	49,0	1704	51,0	3 339
	$P_{t_{2,3,4,5,6,7,8,9}}$	440	36,0	782	64,0	1 222
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3 405
	$P_{t_{2,3,4,5,6,7,8,9}}$	443	36,3	779	63,8	1 222
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4 547
	$P_{t_{2,3,4,5,6,7,8,9}}$	442	36,2	780	63,8	1 222
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3 335
	$P_{t_{2,3,4,5,6,7,8,9}}$	441	36,1	781	63,9	1 222
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3 086
	$P_{t_{2,3,4,5,6,7,8,9}}$	444	36,3	778	63,7	1 222
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3 090
	$P_{t_{2,3,4,5,6,7,8,9}}$	443	36,3	779	63,8	1 222
$t_9$	$Q_{t_9}$	1423	46,8	1620	53,2	3 043
	$P_{t_{2,3,4,5,6,7,8,9}}$	441	36,1	781	63,9	1 222

Das gemittelte Alter der Befragungspersonen ist für die Querschnitte und das Acht-Wellen-Panel in Tabelle 5.95 kontrastiert. Es ist zu erkennen, dass das durchschnittliche

Alter in den Querschnittsdatensätzen jeweils zwischen 0,1 bis 0,2 Jahren über dem Wert für das Acht-Wellen-Panel  $P_{t_{2,3,4,5,6,7,8,9}}$  zu den jeweiligen Zeitpunkten liegt. Einzig zum Zeitpunkt  $t_5$  ist die Differenz mit 0,5 Jahren Unterschied etwas ausgeprägter. Außerdem weist das Acht-Wellen-Panel durchgehend eine etwas größere Altershomogenität auf, da die Standardabweichung in den Querschnitten durchgehend größer ist als zu den jeweiligen Zeitpunkten unter Verwendung des Panels. Insbesondere die Querschnitte  $Q_{t_5}$  und  $Q_{t_6}$  weisen eine auffällig erhöhte Heterogenität der Altersverteilung auf.

**Tabelle 5.95:** Altersdurchschnitt  $P_{t_{2,3,4,5,6,7,8,9}}$

		Alter	Std.-abw.	gesamt
$t_2$	$Q_{t_2}$	14,0	0,8	3270
	$P_{t_{2,3,4,5,6,7,8,9}}$	13,9	0,7	1208
$t_3$	$Q_{t_3}$	15,1	0,8	3270
	$P_{t_{2,3,4,5,6,7,8,9}}$	14,9	0,7	1217
$t_4$	$Q_{t_4}$	16,1	0,8	3313
	$P_{t_{2,3,4,5,6,7,8,9}}$	15,9	0,7	1213
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_{2,3,4,5,6,7,8,9}}$	16,9	0,8	1217
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{t_{2,3,4,5,6,7,8,9}}$	17,9	0,7	1221
$t_7$	$Q_{t_7}$	19,1	0,8	3081
	$P_{t_{2,3,4,5,6,7,8,9}}$	18,9	0,7	1222
$t_8$	$Q_{t_8}$	20,1	0,8	3073
	$P_{t_{2,3,4,5,6,7,8,9}}$	19,9	0,7	1218
$t_9$	$Q_{t_9}$	22,1	0,9	3021
	$P_{t_{2,3,4,5,6,7,8,9}}$	21,9	0,7	1218

Die Gegenüberstellung der Paneldaten aus  $P_{t_{2,3,4,5,6,7,8,9}}$  mit Werten, wie sie in Schulstatistik sowie Querschnittserhebung vorzufinden sind, weist auf eine Verzerrung des Panels bezüglich der Verteilung nach Schulform hin (vgl. Tabelle 5.96). Der Vergleich mit den Schulstatistiken  $S_{t_2}$  und  $S_{t_3}$  macht deutlich, dass im Acht-Wellen-Panel einerseits deutlich mehr Schülerinnen und Schüler von Gymnasien vertreten und andererseits weniger Personen an Hauptschulen. Auch für die Schultypen der Real- und Gesamtschule finden sich Abweichungen, die jedoch ein weniger starkes Ausmaß aufweisen. Ab dem Zeitpunkt  $t_4$  ist nur noch ein Abgleich des Acht-Wellen-Panels mit dem Querschnitt möglich, da sich die Betrachtung von Schulstatistiken ab diesem Zeitpunkt erübrigt, doch es zeigen sich auch bei diesem Vergleich Differenzen der eben genannten Tendenz.

Ab  $t_5$  spielen Schultypen wie Haupt- oder Realschule aufgrund des erreichten Alters der Befragten keine Rolle mehr. Stattdessen werden Berufskollegs relevant. Die Differenz

von  $Q_{t_5}$  und  $P_{t_{2,3,4,5,6,7,8,9}}$  ist in  $t_5$  für diesen Ausbildungstyp besonders hoch, da im Querschnitt die Stichprobe vorübergehend um Berufsschulen erweitert wurde, damit möglichst viele Befragte der Ausgangsstichprobe weiterverfolgt werden können (vgl. Kunadt, 2006). Die Abweichung zwischen Anteil im Querschnitt und im Acht-Wellen-Panel ist daher einmalig besonders ausgeprägt, da Berufskollegs im Querschnitt  $Q_{t_5}$  stark überrepräsentiert sind. Zudem sind umgekehrt in  $t_5$  Personen an Gesamtschulen sowie insbesondere an Gymnasien im Acht-Wellen-Panel stärker vertreten als im Querschnitt.

In den späteren Erhebungszeitpunkten gleichen sich mit abnehmender Relevanz der Schule die Anteilswerte von Quer- und Längsschnitt zunehmend an. Im Zeitpunkt  $t_8$  lassen sich nur noch geringfügige Diskrepanzen zwischen den verglichenen Datensätzen ausmachen.

**Tabelle 5.96:** Zusammensetzung  $P_{t_{2,3,4,5,6,7,8,9}}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							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	—	—	3 763
	$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,4	—	—	3 392
	$P_{t_{2,3,4,5,6,7,8,9}}$	14,7	24,3	31,5	29,5	—	—	—	—	1 222
$t_3$	$S_{t_3}$	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
	$P_{t_{2,3,4,5,6,7,8,9}}$	14,7	24,6	31,8	29,0	—	—	—	—	1 222
$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3 386
	$P_{t_{2,3,4,5,6,7,8,9}}$	14,1	24,5	31,7	29,2	0,3	—	0,3	—	1 221
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4 253
	$P_{t_{2,3,4,5,6,7,8,9}}$	0,7	0,6	22,6	35,0	41,1	—	—	—	1 142
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3 219
	$P_{t_{2,3,4,5,6,7,8,9}}$	—	—	18,9	30,9	30,0	—	11,0	9,1	1 196
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2 975
	$P_{t_{2,3,4,5,6,7,8,9}}$	—	—	16,7	29,8	25,8	—	13,3	14,4	1 196
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2 934
	$P_{t_{2,3,4,5,6,7,8,9}}$	—	—	4,0	4,3	22,3	—	19,5	49,9	1 182

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

### 5.8.3 Acht-Wellen-Panel 2004, 2005, 2006, 2007, 2008, 2009, 2011 und 2013 ( $P_{t_{3,4,5,6,7,8,9,10}}$ )

Das Acht-Wellen-Panel der Erhebungsjahre 2004 bis 2013 umfasst 1 274 Fälle. Anhand der in Tabelle 5.97 dargestellten Auswertung nach Geschlecht wird eine Verzerrung dieses Panels  $P_{t_{3,4,5,6,7,8,9,10}}$  deutlich.

In den Querschnitten der Zeitpunkte  $t_3$  bis  $t_5$  sowie in der Schulstatistik des Erhebungszeitpunktes  $t_3$  ist zu sehen, dass die Relation zwischen männlichen und weiblichen Befragungspersonen weitgehend ausgeglichen ist. In den späteren Querschnitten ab dem Zeitpunkt  $t_6$  überwiegt leicht der Anteil der Frauen, welcher fortan bei rund 53 Prozent liegt. Im Acht-Wellen-Panel macht sich diese Überrepräsentierung weiblicher Personen deutlich stärker bemerkbar: Zu allen Zeitpunkten liegt der Frauenanteil bei 64 bis 65 Prozent, sodass im Vergleich mit den Referenzdaten eine merkliche Verzerrung hinsichtlich eines Überschusses weiblicher Befragter vorliegt.

**Tabelle 5.97:** Zusammensetzung nach Geschlecht  $P_{t_{3,4,5,6,7,8,9,10}}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_{3,4,5,6,7,8,9,10}}$	450	35,3	824	64,7	1274
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_{3,4,5,6,7,8,9,10}}$	454	35,6	820	64,4	1274
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_{3,4,5,6,7,8,9,10}}$	454	35,6	820	64,4	1274
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{t_{3,4,5,6,7,8,9,10}}$	452	35,5	822	64,5	1274
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_{3,4,5,6,7,8,9,10}}$	455	35,7	819	64,3	1274
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{t_{3,4,5,6,7,8,9,10}}$	454	35,6	820	64,4	1274
$t_9$	$Q_{t_9}$	1423	46,8	1620	53,2	3043
	$P_{t_{3,4,5,6,7,8,9,10}}$	456	35,8	818	64,2	1274
$t_{10}$	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_{3,4,5,6,7,8,9,10}}$	453	35,6	821	64,4	1274

Aus den in Tabelle 5.98 gezeigten Werten geht hervor, dass das durchschnittliche Befragtenalter im Querschnitt stets leicht über den Mittelwerten für das Acht-Wellen-Panel liegt. Die Abweichung ist dabei eher gering und beträgt überall 0,2 Jahre. Einzig zu

Zeitpunkt  $t_5$  ist diese Differenz etwas größer und beläuft sich auf 0,5 Jahre. Anzumerken ist dabei, dass in diesem Erhebungsjahr die Ausgangsstichprobe aus methodischen Gründen temporär ausgeweitet wurde (vgl. hierzu Kunadt, 2006), wodurch das leicht erhöhte Durchschnittsalter zu begründen wäre. Zusätzlich geht an dieser Stelle – sowie in kleinerem Maße auch noch im Folgejahr – eine vergleichsweise größere Altersheterogenität einher. Auch im aktuell letzten Erhebungszeitpunkt  $t_{10}$  ist das Alter heterogener verteilt als in den Vergleichsdaten. Das Acht-Wellen-Panel  $P_{t_{3,4,5,6,7,8,9,10}}$  ist hingegen durch eine durchgehend größere Altershomogenität gekennzeichnet.

**Tabelle 5.98:** Altersdurchschnitt  $P_{t_{3,4,5,6,7,8,9,10}}$

		Alter	Std.-abw.	gesamt
$t_3$	$Q_{t_3}$	15,1	0,8	3270
	$P_{t_{3,4,5,6,7,8,9,10}}$	14,9	0,7	1266
$t_4$	$Q_{t_4}$	16,1	0,8	3313
	$P_{t_{3,4,5,6,7,8,9,10}}$	15,9	0,7	1266
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_{3,4,5,6,7,8,9,10}}$	16,9	0,8	1269
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{t_{3,4,5,6,7,8,9,10}}$	17,9	0,7	1273
$t_7$	$Q_{t_7}$	19,1	0,8	3081
	$P_{t_{3,4,5,6,7,8,9,10}}$	18,9	0,7	1274
$t_8$	$Q_{t_8}$	20,1	0,8	3073
	$P_{t_{3,4,5,6,7,8,9,10}}$	19,9	0,7	1274
$t_9$	$Q_{t_9}$	22,1	0,9	3021
	$P_{t_{3,4,5,6,7,8,9,10}}$	21,9	0,7	1273
$t_{10}$	$Q_{t_{10}}$	24,1	1,2	2839
	$P_{t_{3,4,5,6,7,8,9,10}}$	23,9	0,8	1272

Der Zusammensetzung nach Schulform in Tabelle 5.99 ist zu entnehmen, dass eine systematische Verzerrung des Acht-Wellen-Panels  $P_{t_{3,4,5,6,7,8,9,10}}$  nach diesem Merkmal vorliegt. Verglichen mit der Schulstatistik  $S_{t_3}$  sind im Panel zum Zeitpunkt  $t_3$  Befragte an Hauptschulen mit einer Differenz von rund zwölf Prozent unterrepräsentiert. Dagegen sind Befragte, die angeben ein Gymnasium zu besuchen, mit einer ebenso großen Differenz überrepräsentiert. Kleinere Abweichungen im Umfang von zwei bis drei Prozentpunkten sind auch bei den Schulformen Real- und Gesamtschule zu ersichtlich.

Im darauffolgenden Zeitpunkt  $t_4$  ist kein Vergleich des Acht-Wellen-Panels mit der Schulstatistik mehr möglich, sondern nur noch mit der Querschnittserhebung  $Q_{t_4}$ , doch auch hier finden sich derart gerichtete Abweichungen.

Ein leicht anderes Bild zeichnet sich zum Zeitpunkt  $t_5$ : Haupt-, Real- und Sonderschulen werden nur noch von sehr wenigen Befragungspersonen besucht, sodass sich ein Vergleich der Datensätze für diese Schultypen ab dieser Erhebungswelle erübrigt. Dennoch liegt weiterhin ein größerer Anteil an Gymnasiastinnen und Gymnasiasten im Panel vor als im Querschnitt. Ebenso fällt auf, dass der Schultyp des Berufskollegs im Querschnitt mit einer Proportion von knapp 63 Prozent sehr stark vertreten ist. Im Acht-Wellen-Panel hingegen fällt dieser Anteil mit 40 Prozent wesentlich geringer aus. Grund hierfür ist – wie auch an der Fallzahl ersichtlich – eine temporäre Erweiterung der Ausgangsstichprobe durch Hinzunahme von Berufsschulen.

Im weiteren Verlauf nivellieren sich derartige Abweichungen in der Zusammensetzung nach der Schulform für Quer- und Längsschnittdatensatz. Zum Zeitpunkt  $t_8$  lassen sich kaum mehr Differenzen ausmachen. Des Weiteren erübrigt sich auch die Betrachtung der Schulform für die Zeitpunkte  $t_9$  und  $t_{10}$ , da in diesem Alter (Durchschnittsalter 22 bzw. 24 Jahre) so gut wie keine Befragten mehr eine schulische Ausbildungsstätte besuchen.

**Tabelle 5.99:** Zusammensetzung  $P_{t_{3,4,5,6,7,8,9,10}}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							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
	$P_{t_{3,4,5,6,7,8,9,10}}$	13,2	24,2	30,5	32,2	—	—	—	—	1 274
$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3 386
	$P_{t_{3,4,5,6,7,8,9,10}}$	12,4	24,1	30,7	32,2	0,3	—	0,3	—	1 271
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4 253
	$P_{t_{3,4,5,6,7,8,9,10}}$	0,3	0,4	21,2	38,1	40,1	—	—	—	1 177
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3 219
	$P_{t_{3,4,5,6,7,8,9,10}}$	—	—	17,6	33,2	29,5	—	10,5	9,2	1 252
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2 975
	$P_{t_{3,4,5,6,7,8,9,10}}$	—	—	15,8	31,8	25,2	—	12,8	14,4	1 249
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2 934
	$P_{t_{3,4,5,6,7,8,9,10}}$	—	—	3,9	4,2	22,4	—	19,3	50,2	1 235

<sup>a</sup>HS=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 nur jene Datensätze beschrieben, die Teilnahmen in neun direkt aufeinander folgenden Zeitpunkten aufweisen (ohne Lücke). Dabei sind auch die Datensätze enthalten, die aufgrund der Änderung des Erhebungsrhythmus in einem Abstand von zwei Jahren erhoben wurden.

### 5.9.1 Neun-Wellen-Panel 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009 und 2011 ( $P_{t_{1,2,3,4,5,6,7,8,9}}$ )

Das Neun-Wellen-Panel  $P_{t_{1,2,3,4,5,6,7,8,9}}$  umfasst 1 002 Fälle. Mit Blick auf die Tabelle 5.100 lässt sich beobachten, dass dieser Paneldatensatz bezüglich der Verteilung nach Geschlecht Verzerrungen aufweist.

Aus der Schulstatistik der Erhebungszeitpunkte  $t_1$  bis  $t_3$  geht eine annähernde Gleichverteilung von rund 50 zu 50 bzw. 49 zu 51 Prozent Frauen- und Männeranteil hervor. In den Querschnitten  $Q_{t_1}$  bis  $Q_{t_5}$  liegt ebenso ein derartiges Gleichgewicht hervor. Mit zunehmender Erhebungsdauer erhöht sich ab  $t_6$  jedoch im Querschnitt der Anteil an weiblichen Personen und liegt für die weiteren Zeitpunkte konstant bei rund 53 Prozent.

Eine bedeutsamere Verschiebung wird im Neun-Wellen-Panel ersichtlich: Der Frauenanteil in der Stichprobe erreicht unter Verwendung dieser Paneldaten zu allen Zeitpunkten 64 bzw. 65 Prozent. Weibliche Befragungsteilnehmer sind im Neun-Wellen-Panel der Jahre 2002 bis 2011 folglich stark überrepräsentiert.

Die Betrachtung des Durchschnittsalters (Tabelle 5.101) enthüllt, dass in den Querschnitten stets ein leicht höheres durchschnittliches Befragtenalter vorliegt. Die Differenzen zu den Mittelwerten des Neun-Wellen-Panel liegen zwischen 0,2 und 0,3 Jahren. Wie auch bereits in den vorherigen Kapiteln und Abschnitten berichtet, liegt nur zum Zeitpunkt  $t_5$  eine etwas größere Abweichung von einem halben Jahr vor. Gleichzeitig tritt hier eine etwas größere Heterogenität in der Stichprobe auf, die sich auch im nächsten Zeitpunkt  $t_6$  noch fortsetzt. Zu begründen ist dies mit der bereits diskutierten Ausweitung der ursprünglichen Stichprobe, was sich vorrangig auf die Zeitpunkte  $t_5$  und  $t_6$  auswirkt. Insgesamt ist beim Vergleich mit den Querschnitten im Neun-Wellen-Panel zu allen Zeitpunkten eine etwas größere Homogenität der Altersverteilung zu erkennen.

Die Tabelle 5.102 gibt Aufschluss über die Zusammensetzung nach Schulform für das Neun-Wellen-Panel  $P_{t_{1,2,3,4,5,6,7,8,9}}$ . Dabei werden dieselben Abweichungen erkenntlich wie auch bereits bei den einzelnen Paneldatensätzen der Kapitel zuvor.

Vergleicht man die Daten dieses Neun-Wellen-Panels mit den offiziellen Kennziffern der Schulstatistiken  $S_{t_1}$ ,  $S_{t_2}$  und  $S_{t_3}$ , so kann insbesondere eine Verschiebung hinsichtlich der Schultypen Hauptschule und Gymnasium festgehalten werden.

**Tabelle 5.100:** Zusammensetzung nach Geschlecht  $P_{1,2,3,4,5,6,7,8,9}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_1$	$S_{t_1}$	2003	51,2	1907	48,8	3910
	$Q_{t_1}$	1728	50,7	1679	49,3	3407
	$P_{1,2,3,4,5,6,7,8,9}$	352	35,1	650	64,9	1002
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{1,2,3,4,5,6,7,8,9}$	357	35,6	645	64,4	1002
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{1,2,3,4,5,6,7,8,9}$	352	35,1	650	64,9	1002
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{1,2,3,4,5,6,7,8,9}$	353	35,2	649	64,8	1002
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{1,2,3,4,5,6,7,8,9}$	355	35,4	647	64,6	1002
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{1,2,3,4,5,6,7,8,9}$	354	35,3	648	64,7	1002
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{1,2,3,4,5,6,7,8,9}$	356	35,5	646	64,5	1002
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{1,2,3,4,5,6,7,8,9}$	355	35,4	647	64,6	1002
$t_9$	$Q_{t_9}$	1423	46,8	1620	53,2	3043
	$P_{1,2,3,4,5,6,7,8,9}$	353	35,2	649	64,8	1002

Beispielsweise sind zum Zeitpunkt  $t_1$  im Neun-Wellen-Panel gut zehn Prozent weniger Hauptschülerinnen und -schüler vertreten als die Schulstatistik dieses Zeitpunktes ausweist. Andersherum enthält das Panel knapp neun Prozent mehr Personen an Gymnasien. Auch für die weiteren Erhebungszeitpunkte bis einschließlich  $t_4$  sind derartige Diskrepanzen zwischen Quer- und Längsschnittdatensatz festzustellen.

Eine Besonderheit weist der Vergleich von Neun-Wellen-Panel mit dem Querschnitt zum Zeitpunkt  $t_5$  auf: Durch die bereits erwähnte Vergrößerung der Stichprobe durch Hinzunehmen von Berufsschulen liegt zum Zeitpunkt  $t_5$  eine deutlich umfassendere Querschnittsstichprobe vor, in der sich entsprechend viele Schülerinnen und -schüler von Berufskollegs befinden. Der Anteil dieser Schulform ist demnach in  $Q_{t_5}$  und dem Panel  $P_{1,2,3,4,5,6,7,8,9}$  zum Zeitpunkt  $t_5$  stark verschieden. Überdies finden sich weiterhin die auch zuvor bereits auffälligen Differenzen dahingehend, dass Gymnasiastinnen und

**Tabelle 5.101:** Altersdurchschnitt  $P_{t_1,2,3,4,5,6,7,8,9}$

		Alter	Std.-abw.	gesamt
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{t_1,2,3,4,5,6,7,8,9}$	12,8	0,6	998
$t_2$	$Q_{t_2}$	14,0	0,8	3 270
	$P_{t_1,2,3,4,5,6,7,8,9}$	13,8	0,6	992
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_1,2,3,4,5,6,7,8,9}$	14,9	0,7	998
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_1,2,3,4,5,6,7,8,9}$	15,8	0,6	996
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_1,2,3,4,5,6,7,8,9}$	16,9	0,6	999
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_1,2,3,4,5,6,7,8,9}$	17,8	0,7	1 001
$t_7$	$Q_{t_7}$	19,1	0,8	3 081
	$P_{t_1,2,3,4,5,6,7,8,9}$	18,8	0,7	1 002
$t_8$	$Q_{t_8}$	20,1	0,8	3 073
	$P_{t_1,2,3,4,5,6,7,8,9}$	19,9	0,7	999
$t_9$	$Q_{t_9}$	22,1	0,9	3 021
	$P_{t_1,2,3,4,5,6,7,8,9}$	21,9	0,7	999

Gymnasiasten im Neun-Wellen-Panel stärker vertreten sind als im Querschnitt. Selbiges gilt für den Schultyp der Gesamtschule.

Mit fortschreitender Erhebung nimmt der Anteil an Personen, die keine Schule (mehr) besuchen, stetig zu und liegt zum Zeitpunkt  $t_8$  bereits bei etwa der Hälfte der Stichprobe. Für den weiteren Zeitpunkt  $t_9$  ist eine Betrachtung der Schulform daher wenig sinnvoll, weswegen in der Tabelle 5.102 darauf verzichtet wird. Gleichzeitig lässt sich auch festhalten, dass mit abnehmender Bedeutung der Schulform sich auch die Verzerrungen des Paneldatensatzes im Vergleich zu den Querschnitten zunehmend mindern: Zum Zeitpunkt  $t_8$  sind keine nennenswerten Differenzen mehr zu erkennen.

**Tabelle 5.102:** Zusammensetzung  $P_{1,2,3,4,5,6,7,8,9}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								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
	$P_{1,2,3,4,5,6,7,8,9}$	13,5	25,3	31,3	29,9	—	—	—	—	1002
$t_2$	$S_{t_2}$	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
	$P_{1,2,3,4,5,6,7,8,9}$	13,9	25,6	31,3	29,2	—	—	—	—	1002
$t_3$	$S_{t_3}$	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
	$P_{1,2,3,4,5,6,7,8,9}$	14,0	25,6	31,3	28,9	—	—	—	—	1002
$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{1,2,3,4,5,6,7,8,9}$	13,5	25,8	30,9	29,4	0,2	—	0,3	—	1001
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{1,2,3,4,5,6,7,8,9}$	0,8	0,5	22,5	35,5	40,7	—	—	—	936
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{1,2,3,4,5,6,7,8,9}$	—	—	18,7	31,0	30,0	—	11,6	8,8	981
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{1,2,3,4,5,6,7,8,9}$	—	—	16,5	29,6	25,4	—	13,8	14,7	983
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{1,2,3,4,5,6,7,8,9}$	—	—	4,0	4,3	21,4	—	19,8	50,4	968

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

### 5.9.2 Neun-Wellen-Panel 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2011 und 2013 ( $P_{2,3,4,5,6,7,8,9,10}$ )

Das Neun-Wellen-Panel des Erhebungszeitraums von 2003 bis 2013 beinhaltet 1 098 Fälle.

Unter Begutachtung der Zusammensetzung der Daten nach Geschlecht [Tabelle 5.103] wird eine Verzerrung dieses Paneldatensatzes deutlich. Sowohl die zum Vergleich herangezogene Schulstatistik  $S_{t_2}$  und  $S_{t_3}$  als auch die Querschnitte der Zeitpunkte von  $t_2$  bis einschließlich  $t_5$  weisen ein annähernd gleiches Verhältnis von männlichen und weiblichen Befragten aus. Die Größen beider Gruppen weichen um maximal zwei Prozent voneinander ab.

Ab dem Zeitpunkt  $t_5$  des Erhebungsjahres 2006 weicht diese Zusammensetzung in den Querschnitten zunehmend von der ursprünglichen Verteilung der Schulstatistik ab: Der Frauenanteil in der Stichprobe liegt fortan bei 53 bis 54 Prozent.

Noch deutlicher wird diese Überrepräsentanz weiblicher Probanden im Neun-Wellen-Panel  $P_{t_{2,3,4,5,6,7,8,9,10}}$ : Der Anteil an weiblichen Befragten liegt in diesem Datensatz zu allen Zeitpunkten stets bei rund 65 Prozent. Es sind also nahezu zwei Drittel aller Befragten in der Stichprobe dieses Teilpanels Frauen statt wie ursprünglich in der Schulstatistik die Hälfte.

**Tabelle 5.103:** Zusammensetzung nach Geschlecht  $P_{t_{2,3,4,5,6,7,8,9,10}}$

		männlich		weiblich		gesamt (100 %)
		Hfk	%	Hfk	%	
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,1	3390
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	385	35,1	713	64,9	1098
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	380	34,6	718	65,4	1098
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_{1,2,3,4,5,6,7,8,9,10}}$	383	34,9	715	65,1	1098
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	382	34,8	716	65,2	1098
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	381	34,7	717	65,3	1098
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	384	35,0	714	65,0	1098
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	383	34,9	715	65,1	1098
$t_9$	$Q_{t_9}$	1423	46,8	1620	53,2	3043
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	383	34,9	715	65,1	1098
$t_{10}$	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	382	34,8	716	65,2	1098

In Tabelle 5.104 ist das durchschnittliche Alter der Befragten für die Zeitpunkte des Neun-Wellen-Panels  $P_{t_{2,3,4,5,6,7,8,9,10}}$  sowie zum Vergleich das Durchschnittsalter in den Querschnitten abgebildet. Es zeigt sich, dass das gemittelte Alter im Querschnitt stets etwas über dem des Panels liegt. Die Abweichung zwischen den beiden Arten von Datensätzen beträgt dabei jeweils 0,2 bis 0,3 Jahre. Einzige Ausnahme bildet der

Zeitpunkt  $t_5$ , bei welchem die Differenz mit einem Wert von 0,5 Jahren etwas höher ausfällt.

Bezüglich der Streuung der Altersverteilung ist zu erkennen, dass sich die Befragten im Neun-Wellen-Panel bezüglich ihres Alters ähnlicher sind; die Standardabweichungen im Längsschnittdatensatz liegen durchgängig unter den Vergleichswerten der Querschnitte. Besonders stechen die Kennziffern der Querschnitte  $Q_{t_5}$ ,  $Q_{t_6}$  und  $Q_{t_{10}}$  hervor. In diesen drei Querschnittsdatsätzen weist die Stichprobe die vergleichsweise geringste Altershomogenität auf.

**Tabelle 5.104:** Altersdurchschnitt  $P_{t_{2,3,4,5,6,7,8,9,10}}$

		Alter	Std.-abw.	gesamt
$t_2$	$Q_{t_2}$	14,0	0,8	3 270
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	13,8	0,7	1 085
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	14,9	0,7	1 094
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	15,9	0,7	1 091
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	16,9	0,8	1 093
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	17,9	0,7	1 097
$t_7$	$Q_{t_7}$	19,1	0,8	3 081
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	18,8	0,7	1 098
$t_8$	$Q_{t_8}$	20,1	0,8	3 073
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	19,9	0,7	1 098
$t_9$	$Q_{t_9}$	22,1	0,9	3 021
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	21,9	0,7	1 097
$t_{10}$	$Q_{t_{10}}$	24,1	1,2	2 839
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	23,8	0,8	1 096

Vergleicht man die Zusammensetzung nach Schulform im Neun-Wellen-Panel  $P_{t_{2,3,4,5,6,7,8,9,10}}$  mit der Verteilung in der Schulstatistik oder den Querschnittserhebungen (vgl. Tabelle 5.105), so wird auch für dieses Merkmal eine systematische Verzerrung des Längsschnittdatensatzes erkennbar.

Zu den Zeitpunkten  $t_2$  und  $t_3$  liegen Zahlen der Schulstatistik als Referenz vor. Dabei wird ersichtlich, dass Hauptschülerinnen- und schüler im Neun-Wellen-Panel weniger stark vertreten sind als in der offiziellen Statistik. Umgekehrt befinden sich im Panel prozentual mehr Schülerinnen und Schüler von Gymnasien. Auch Personen an Real-

und Gesamtschulen sind zum Zeitpunkt  $t_2$  im Panel überrepräsentiert, die Abweichung ist hier im Vergleich zu Hauptschulen und Gymnasien jedoch eher gering.

Eine derart gerichtete Verschiebung lässt sich aber nicht nur für das Panel beobachten, sondern tritt auch bereits im Querschnitt  $Q_{t_2}$  auf, wenn man diesen der Schulstatistik  $S_{t_2}$  gegenüberstellt. Tendenzen dieser Art finden sich auch in den weiteren Erhebungszeitpunkten.

Zum Zeitpunkt  $t_5$  spielen Haupt- und Realschulen keine Rolle mehr, dafür werden Berufskollegs relevant. Der Anteil dieser Schulform ist aufgrund der bereits beschriebenen Besonderheit im Erhebungsmodus zum Zeitpunkt  $t_5$  besonders hoch, sodass ein Vergleich zwischen Querschnitt und Neun-Wellen-Panel an dieser Stelle wenig aussagekräftig ist. Nichtsdestotrotz sei zu erwähnen, dass sich auch zu diesem Zeitpunkt eine Überrepräsentanz von Gesamtschulen und Gymnasien im Panel beim Vergleich mit dem Querschnitt feststellen lässt.

In den weiteren Zeitpunkten nimmt die Bedeutung der Schule für die Befragten immer mehr ab – in den Erhebungswellen  $t_9$  und  $t_{10}$  besuchen so gut wie keine Personen mehr eine Schule, sodass in Tabelle 5.105 auf die Darstellung dieser Zeitpunkte verzichtet wird. Zudem werden gleichzeitig die bisher gefundenen Differenzen in der Schulformverteilung zwischen Querschnitt und Neun-Wellen-Panels zunehmend unbedeutender: Zum Zeitpunkt  $t_8$  weichen beide Datensätze nur noch marginal voneinander ab.

**Tabelle 5.105:** Zusammensetzung  $P_{t_{2,3,4,5,6,7,8,9,10}}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								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
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	13,2	24,7	31,4	30,7	—	—	—	—	1098
$t_3$	$S_{t_3}$	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
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	13,2	25,1	31,5	30,2	—	—	—	—	1098
$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	12,6	25,0	31,6	30,4	0,2	—	0,3	—	1097
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	0,3	0,5	22,5	36,4	40,4	—	—	—	1023
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	—	—	19,1	31,9	29,8	—	10,8	8,3	1079
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	—	—	17,1	30,8	25,5	—	12,8	13,8	1075
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{t_{2,3,4,5,6,7,8,9,10}}$	—	—	4,1	4,6	22,3	—	19,2	49,9	1065

<sup>a</sup>HS=Hauptschule, RS=Realschule, GS=Gesamtschule, GY=Gymnasium, BK=Berufskolleg, SO=Sonderschule

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

Alle (nach den Plausibilitätskontrollen) gültigen Fälle, die neun weiteren Fällen im Datensatz zugeordnet werden konnten – also entsprechend in allen zehn Erhebungswellen vorhanden sind – bilden das vollständige Zehn-Wellen-Panel. In diesem Panel sind also ausschließlich Personen enthalten, die an jedem der zehn Befragungszeitpunkte von 2002 bis 2013 teilgenommen haben und anhand ihres persönlichen Codes zugeordnet werden konnten. Dieses Zehn-Wellen-Panel ist der umfassendste Datensatz des in diesem Methodenbericht beschriebenen Erhebungszeitraums von 2002 bis 2013.

Unter Betrachtung der in Tabelle 5.106 gezeigten Zusammensetzung dieses Panels nach Geschlecht wird im Vergleich zu den Referenzdaten des Querschnitts deutlich, dass eine Verschiebung des Panels  $P_{t_{1,2,3,4,5,6,7,8,9,10}}$  vorliegt.

Die Schulstatistiken der Zeitpunkte  $t_1$  bis  $t_3$  sowie die Querschnitte von  $t_1$  bis  $t_5$  sagen aus, 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. Noch deutlicher spiegelt sich dieses Ungleichgewicht im Zehn-Wellen-Panel wider: Zu allen Zeitpunkten liegt der Anteil an Personen weiblichen Geschlechts bei rund 66 Prozent. der Paneldatensatz ist also in Richtung einer Überrepräsentanz weiblicher Befragter verzerrt.

**Tabelle 5.106:** Zusammensetzung nach Geschlecht  $P_{t_1,2,3,4,5,6,7,8,9,10}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_1$	$S_{t_1}$	2003	51,2	1907	48,8	3910
	$Q_{t_1}$	1728	50,7	1679	49,3	3407
	$P_{t_1,2,3,4,5,6,7,8,9,10}$	309	33,9	602	66,1	911
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{t_1,2,3,4,5,6,7,8,9,10}$	314	34,5	597	65,5	911
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_1,2,3,4,5,6,7,8,9,10}$	309	33,9	602	66,1	911
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_1,2,3,4,5,6,7,8,9,10}$	310	34,0	601	66,0	911
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_1,2,3,4,5,6,7,8,9,10}$	312	34,2	599	65,8	911
$t_6$	$Q_{t_6}$	1567	47,0	1768	53,0	3335
	$P_{t_1,2,3,4,5,6,7,8,9,10}$	311	34,1	600	65,9	911
$t_7$	$Q_{t_7}$	1452	47,1	1634	52,9	3086
	$P_{t_1,2,3,4,5,6,7,8,9,10}$	313	34,4	598	65,6	911
$t_8$	$Q_{t_8}$	1446	46,8	1644	53,2	3090
	$P_{t_1,2,3,4,5,6,7,8,9,10}$	312	34,2	599	65,8	911
$t_9$	$Q_{t_9}$	1423	46,8	1620	53,2	3043
	$P_{t_1,2,3,4,5,6,7,8,9,10}$	311	34,1	600	65,9	911
$t_{10}$	$Q_{t_{10}}$	1302	45,7	1546	54,3	2848
	$P_{t_1,2,3,4,5,6,7,8,9,10}$	311	34,1	600	65,9	911

Tabelle 5.107 zeigt das durchschnittliche Alter im Zehn-Wellen-Panel sowie vergleichend die Werte der Querschnittserhebungen. Zu allen zehn Zeitpunkten liegt das Durchschnittsalter im Querschnitt leicht über dem im Zehn-Wellen-Panel. Diese Abwei-

chung beläuft sich dabei auf 0,2 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. Dabei weisen die Querschnitte der Zeitpunkte  $t_5$ ,  $t_6$  und  $t_{10}$  die geringste Homogenität der Altersverteilung auf – zu diesen Zeitpunkten bestehen die größten Abweichungen im Alter der Befragten untereinander.

**Tabelle 5.107:** Altersdurchschnitt  $P_{t_{1,2,3,4,5,6,7,8,9,10}}$

		Alter	Std.-abw.	gesamt
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{t_{1,2,3,4,5,6,7,8,9,10}}$	12,8	0,6	908
$t_2$	$Q_{t_2}$	14,0	0,8	3 270
	$P_{t_{1,2,3,4,5,6,7,8,9,10}}$	13,8	0,6	902
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_{1,2,3,4,5,6,7,8,9,10}}$	14,9	0,7	907
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_{1,2,3,4,5,6,7,8,9,10}}$	15,8	0,7	906
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_{1,2,3,4,5,6,7,8,9,10}}$	16,9	0,7	908
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_{1,2,3,4,5,6,7,8,9,10}}$	17,8	0,7	910
$t_7$	$Q_{t_7}$	19,1	0,8	3 081
	$P_{t_{1,2,3,4,5,6,7,8,9,10}}$	18,8	0,7	911
$t_8$	$Q_{t_8}$	20,1	0,8	3 073
	$P_{t_{1,2,3,4,5,6,7,8,9,10}}$	19,8	0,7	911
$t_9$	$Q_{t_9}$	22,1	0,9	3 021
	$P_{t_{1,2,3,4,5,6,7,8,9,10}}$	21,8	0,7	911
$t_{10}$	$Q_{t_{10}}$	24,1	1,2	2 839
	$P_{t_{1,2,3,4,5,6,7,8,9,10}}$	23,8	0,8	909

Bei der Zusammensetzung nach Schulform (s. Tabelle 5.108) 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 Zehn-Wellen-Wellen unterrepräsentiert sind. Umgekehrt befinden sich in diesem Panel vier Prozent mehr Realschülerinnen und -schüler sowie fast zehn Prozent mehr Gymnasiastinnen und Gymnasiasten als der Schulstatistik zu entnehmen ist. Eine Verschiebung dieser Art lässt sich ebenso für die weiteren Zeitpunkte und auch beim Vergleich von Zehn-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 Hauptschule, 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 die selben Tendenzen der Überrepräsentanz von Gymnasien und auch Gesamtschulen abzeichnen.

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 zehn-Wellen-Panel feststellen. Auf eine Betrachtung der Zeitpunkte  $t_9$  und  $t_{10}$  wird auch an dieser Stelle wie gehabt verzichtet, da in diesen Erhebungsjahren – das durchschnittliche Alter der Befragten liegt hier bei 22 bzw. 24 Jahren – so gut wie keine Befragten mehr eine schulische Einrichtung besuchen.

**Tabelle 5.108:** Zusammensetzung  $P_{1,2,3,4,5,6,7,8,9,10}$  nach Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							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
	$P_{1,2,3,4,5,6,7,8,9,10}$	12,2	26,1	30,7	31,0	—	—	—	—	911
$t_2$	$S_{t_2}$	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
	$P_{1,2,3,4,5,6,7,8,9,10}$	12,6	26,3	30,7	30,3	—	—	—	—	911
$t_3$	$S_{t_3}$	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
	$P_{1,2,3,4,5,6,7,8,9,10}$	12,7	26,6	30,7	30,0	—	—	—	—	911
$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{1,2,3,4,5,6,7,8,9,10}$	12,2	26,5	30,6	30,3	0,1	—	0,3	—	910
$t_5$	$Q_{t_5}$	0,4	0,3	16,2	20,2	62,7	0,1	—	—	4253
	$P_{1,2,3,4,5,6,7,8,9,10}$	0,2	0,5	22,6	36,7	40,0	—	—	—	848
$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
	$P_{1,2,3,4,5,6,7,8,9,10}$	—	—	18,7	31,8	29,7	—	11,5	8,3	895
$t_7$	$Q_{t_7}$	—	—	13,8	22,8	28,1	—	14,7	20,6	2975
	$P_{1,2,3,4,5,6,7,8,9,10}$	—	—	16,7	30,5	25,2	—	13,5	14,1	894
$t_8$	$Q_{t_8}$	—	—	4,1	3,4	22,9	—	18,3	51,2	2934
	$P_{1,2,3,4,5,6,7,8,9,10}$	—	—	4,3	4,5	21,3	—	19,4	50,5	882

<sup>a</sup>HS=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.

Eine der vorliegenden Verzerrungen betrifft 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 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 des Weiteren in den Paneldatensätzen nochmals besonders deutlich: Dort liegt der Frauenanteil stets wesentlich über den Referenzwerten aus Querschnitt und Schulstatistik. Dieser Anteil bzw. diese Verzerrung ist außerdem umso größer, je umfangreicher das Panel wird. So erreicht der Anteil an weiblichen Personen im bislang umfassendsten Paneldatensatz – dem Zehn-Wellen-Panel – einen Wert von 66 Prozent.

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 Verschiebungen aufweisen. 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  $t_8$  (Durchschnittsalter 20 Jahre) 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  $P_{t_{1,2,3,4,5,6}}$  bereits 61 Prozent. Im Zehn-Wellen-Panel  $P_{t_{1,2,3,4,5,6,7,8,9,10}}$ , welches Gegenstand des vorliegenden Methodenberichts ist, liegt dieser Anteil schließlich bei 66 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 Mädchen weniger Schwierigkeiten haben, ihr Codeblatt fehlerfrei auszufüllen als Jungen. 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.
- Boers, K. & Reinecke, J. (2007). Strukturdynamisches Analysemodell und Forschungshypothesen. In K. Boers & J. Reinecke (Hrsg.), *Delinquenz im Jugendalter: Erkenntnisse einer Münsteraner Längsschnittstudie* (S. 41–55). Münster and New York: Waxmann.
- Boers, K., Reinecke, J., Bentrup, C., Daniel, A., Kanz, K.-M., Schulte, P., Seddig, D., Theimann, M., Verneuer, L. & 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.
- Boers, K., Reinecke, J., Mariotti, L. & Seddig, D. (2010). Explaining the Development of Adolescent Violent Delinquency. *European Journal of Criminology*, 7 (6), 499–520.
- Brondies, M. (2004). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2003* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 6). Münster und Trier.
- Kunadt, S. (2006). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2005* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 11). 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. (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

**Erstellung des persönlichen Codes**

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

**Hier nun die sechs Fragen zur Erstellung Ihres persönlichen Codes:**

<b>1</b>	Bitte kreuzen Sie den <b>ersten</b> Buchstaben des Vornamens Ihres <b>Vaters</b> (oder einer Person, die für Sie einem Vater am nächsten kommt) an. (z. B. <input type="checkbox"/> Anton, <input checked="" 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
<b>2</b>	Bitte kreuzen Sie den <b>ersten</b> Buchstaben des Vornamens Ihrer <b>Mutter</b> (oder einer Person, die für Sie einer Mutter am nächsten kommt) an. (z. B. <input type="checkbox"/> Anna, <input checked="" type="checkbox"/> Beate, <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
<b>3</b>	Bitte kreuzen Sie den <b>ersten</b> Buchstaben Ihres <b>Vornamens</b> an (z. B. <input checked="" 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
<b>4</b>	Bitte kreuzen Sie den <b>Tag</b> Ihres <b>Geburtsdatums</b> an (z. B. Geburtstag am 7. Januar = <input checked="" type="checkbox"/> 7, am 12. Mai = <input type="checkbox"/> 12, am 31. Oktober = <input checked="" 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
<b>5</b>	Bitte kreuzen Sie den <b>letzten</b> Buchstaben Ihrer natürlichen <b>Haarfarbe</b> an. (z. B. braun <input checked="" type="checkbox"/> , Glatz <input type="checkbox"/> , schwarz <input type="checkbox"/> , 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
<b>6</b>	Bitte kreuzen Sie den <b>letzten</b> Buchstaben Ihrer <b>Augenfarbe</b> an. (z. B. braun <input type="checkbox"/> , grün <input type="checkbox"/> , grau <input type="checkbox"/> , 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

Haben Sie im letzten Jahr an der Befragung teilgenommen?  ja  nein

Sind Sie derzeit im 12. oder 13. Jahrgang eines Gymnasiums oder einer Gesamtschule?  ja  nein

**Wenn ja:** Sind Sie im letzten Jahr sitzen geblieben?  ja  nein

Seite 2 Westfälische Wilhelms-Universität Münster, Kriminalwissenschaften / Universität Bielefeld, Soziologie – Duisburger Jugendbefragung 2008

**Abbildung A.1:** Codeblatt Duisburg 2008

**Erstellung des persönlichen Codes**

**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:**

<b>1</b>	Bitte kreuzen Sie den <b>ersten</b> Buchstaben des Vornamens Ihres <b>Vaters</b> (oder einer Person, die für Sie einem Vater am nächsten kommt) an. (z. B. <input checked="" type="checkbox"/> Anton, <input checked="" type="checkbox"/> Bernd, <input checked="" 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
<b>2</b>	Bitte kreuzen Sie den <b>ersten</b> Buchstaben des Vornamens Ihrer <b>Mutter</b> (oder einer Person, die für Sie einer Mutter am nächsten kommt) an. (z. B. <input checked="" type="checkbox"/> Anna, <input checked="" type="checkbox"/> Beate, <input checked="" type="checkbox"/> Jutta, <input checked="" 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
<b>3</b>	Bitte kreuzen Sie den <b>ersten</b> Buchstaben Ihres <b>Vornamens</b> an (z. B. <input checked="" type="checkbox"/> Michael, <input checked="" type="checkbox"/> Thomas, <input checked="" 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
<b>4</b>	Bitte kreuzen Sie den <b>Tag</b> Ihres <b>Geburtsdatums</b> an (z. B. Geburtstag am 7. Januar = <input checked="" type="checkbox"/> 7, am 12. Mai = <input checked="" type="checkbox"/> 12, am 31. Oktober = <input checked="" type="checkbox"/> 31). <input type="checkbox"/> 1 <input checked="" 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
<b>5</b>	Bitte kreuzen Sie den <b>letzten</b> Buchstaben Ihrer natürlichen <b>Haarfarbe</b> an. (z. B. braun <input checked="" type="checkbox"/> , Glatz <input checked="" type="checkbox"/> , schwarz <input checked="" type="checkbox"/> , 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
<b>6</b>	Bitte kreuzen Sie den <b>letzten</b> Buchstaben Ihrer <b>Augenfarbe</b> an. (z. B. braun <input checked="" type="checkbox"/> , grün <input checked="" type="checkbox"/> , grau <input checked="" type="checkbox"/> , 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
<b>7</b>	Bitte kreuzen Sie den <b>letzten</b> Buchstaben Ihres <b>Nachnamens</b> 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

Haben Sie im letzten Jahr an der Befragung teilgenommen? ja nein

Seite 2 Westfälische Wilhelms-Universität Münster, Kriminalwissenschaften / Universität Bielefeld, Soziologie – Duisburger Jugendbefragung 2009

Abbildung A.2: Codeblatt Duisburg 2009

**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:**

<b>1</b>	Bitte kreuzen Sie den <b>ersten</b> Buchstaben des Vornamens Ihres <b>Vaters</b> (oder einer Person, die für Sie einen 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
<b>2</b>	Bitte kreuzen Sie den <b>ersten</b> Buchstaben des Vornamens Ihrer <b>Mutter</b> (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.). <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
<b>3</b>	Bitte kreuzen Sie den <b>ersten</b> Buchstaben Ihres <b>Vornamens</b> 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
<b>4</b>	Bitte kreuzen Sie den <b>Tag</b> Ihres <b>Geburtsdatums</b> 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
<b>5</b>	Bitte kreuzen Sie den <b>letzten</b> Buchstaben Ihrer natürlichen <b>Haarfarbe</b> an (z. B. braun <sup>7</sup> , Glatz <sup>7</sup> , schwarz <sup>7</sup> , 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
<b>6</b>	Bitte kreuzen Sie den <b>letzten</b> Buchstaben Ihrer <b>Augenfarbe</b> an (z. B. braun <sup>7</sup> , grün <sup>7</sup> , grau <sup>7</sup> , 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
<b>7</b>	Bitte kreuzen Sie den <b>letzten</b> Buchstaben Ihres <b>Nachnamens</b> 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.3: Codeblatt Duisburg 2011

**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:**

<b>1</b>	<p>Bitte kreuzen Sie den <b>ersten</b> Buchstaben des Vornamens Ihres <b>Vaters</b> (oder einer Person, die für Sie einen Vater am nächsten kommt) an. (z. B. <input type="checkbox"/>Anton, <input type="checkbox"/>Bernd, <input type="checkbox"/>Jans-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																		
<b>2</b>	<p>Bitte kreuzen Sie den <b>ersten</b> Buchstaben des Vornamens Ihrer <b>Mutter</b> (oder einer Person, die für Sie eine 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																		
<b>3</b>	<p>Bitte kreuzen Sie den <b>ersten</b> Buchstaben Ihres <b>Vornamens</b> 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																		
<b>4</b>	<p>Bitte kreuzen Sie den <b>Tag</b> Ihres <b>Geburtsdatums</b> 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																	
<b>5</b>	<p>Bitte kreuzen Sie den <b>letzten</b> Buchstaben Ihrer natürlichen <b>Haarfarbe</b> an (z. B. braun <input type="checkbox"/>, Glatze <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																		
<b>6</b>	<p>Bitte kreuzen Sie den <b>letzten</b> Buchstaben Ihrer <b>Augenfarbe</b> 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																		
<b>7</b>	<p>Bitte kreuzen Sie den <b>letzten</b> Buchstaben Ihres <b>Nachnamens</b> 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.4: Codeblatt Duisburg 2013**

## A.2 Häufigkeiten der Codevariablen

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

	2008		2009		2011		2013	
	N	%	N	%	N	%	N	%
a	250	8,0	254	8,2	231	7,6	217	7,6
ä	—	—	—	—	—	—	—	—
b	87	2,8	87	2,8	88	2,9	84	2,9
c	62	2,0	62	2,0	62	2,0	61	2,1
d	131	4,2	129	4,1	130	4,3	116	4,1
e	97	3,1	91	2,9	95	3,1	82	2,9
f	122	3,9	127	4,1	126	4,1	111	3,9
g	139	4,5	129	4,1	120	3,9	118	4,1
h	357	11,4	347	11,2	355	11,6	337	11,8
i	64	2,1	57	1,8	58	1,9	52	1,8
j	177	5,7	174	5,6	170	5,6	159	5,6
k	153	4,9	158	5,1	154	5,0	141	4,9
l	36	1,2	41	1,3	42	1,4	41	1,4
m	325	10,4	321	10,3	319	10,4	289	10,1
n	72	2,3	77	2,5	74	2,4	71	2,5
o	40	1,3	39	1,3	37	1,2	34	1,2
ö	5	0,2	3	0,1	4	0,1	5	0,2
p	101	3,2	110	3,5	107	3,5	102	3,6
q	—	—	—	—	—	—	—	—
r	237	7,6	244	7,8	250	8,2	230	8,1
s	140	4,5	137	4,4	133	4,4	119	4,2
ß	—	—	1	0,0	—	—	—	—
t	92	2,9	94	3,0	92	3,0	88	3,1
u	125	4,0	123	4,0	121	4,0	114	4,0
ü	5	0,2	4	0,1	4	0,1	4	0,1
v	62	2,0	66	2,1	59	1,9	59	2,1
w	160	5,1	160	5,1	143	4,7	142	5,0
x	—	—	—	—	1	0,0	2	0,1
y	31	1,0	29	0,9	28	0,9	20	0,7
z	18	0,6	18	0,6	12	0,4	13	0,5
fehlend	32	1,0	30	1,0	41	1,3	44	1,5
gesamt	3 120	100,0	3 112	100,0	3 056	100,0	2 855	100,0

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

	2008		2009		2011		2013	
	N	%	N	%	N	%	N	%
a	283	9,1	279	9,0	278	9,1	253	8,9
ä	1	0,0	1	0,0	1	0,0	1	0,0
b	265	8,5	269	8,6	263	8,6	253	8,9
c	171	5,5	172	5,5	172	5,6	148	5,2
d	92	2,9	91	2,9	90	2,9	87	3,0
e	179	5,7	183	5,9	176	5,8	169	5,9
f	78	2,5	78	2,5	82	2,7	62	2,2
g	172	5,5	173	5,6	166	5,4	145	5,1
h	225	7,2	214	6,9	211	6,9	199	7,0
i	123	3,9	119	3,8	121	4,0	112	3,9
j	47	1,5	52	1,7	43	1,4	39	1,4
k	120	3,8	122	3,9	113	3,7	105	3,7
l	51	1,6	50	1,6	45	1,5	46	1,6
m	347	11,1	357	11,5	337	11,0	321	11,2
n	121	3,9	114	3,7	116	3,8	113	4,0
o	15	0,5	13	0,4	12	0,4	12	0,4
ö	2	0,1	1	0,0	1	0,0	1	0,0
p	94	3,0	99	3,2	103	3,4	90	3,2
q	—	—	—	—	1	0,0	—	—
r	147	4,7	142	4,6	134	4,4	129	4,5
s	322	10,3	316	10,2	329	10,8	307	10,8
ß	—	—	—	—	—	—	—	—
t	32	1,0	36	1,2	38	1,2	34	1,2
u	110	3,5	110	3,5	100	3,3	98	3,4
ü	4	0,1	2	0,1	2	0,1	3	0,1
v	33	1,1	35	1,1	31	1,0	33	1,2
w	11	0,4	11	0,4	11	0,4	12	0,4
x	1	0,0	2	0,1	1	0,0	2	0,1
y	14	0,4	14	0,4	15	0,5	16	0,6
z	32	1,0	33	1,1	29	0,9	31	1,1
fehlend	28	0,9	24	0,8	35	1,1	34	1,2
gesamt	3 120	100,0	3 112	100,0	3 056	100,0	2 855	100,0

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

	2008		2009		2011		2013	
	N	%	N	%	N	%	N	%
a	278	8,9	277	8,9	269	8,8	246	8,6
ä	—	—	—	—	1	0,0	—	—
b	84	2,7	78	2,5	78	2,6	70	2,5
c	169	5,4	168	5,4	163	5,3	148	5,2
d	249	8,0	245	7,9	239	7,8	219	7,7
e	91	2,9	87	2,8	94	3,1	90	3,2
f	84	2,7	97	3,1	86	2,8	78	2,7
g	36	1,2	33	1,1	34	1,1	35	1,2
h	74	2,4	80	2,6	85	2,8	71	2,5
i	47	1,5	44	1,4	44	1,4	41	1,4
j	249	8,0	256	8,2	246	8,0	235	8,2
k	167	5,4	158	5,1	161	5,3	159	5,6
l	114	3,7	109	3,5	107	3,5	99	3,5
m	379	12,1	377	12,1	375	12,3	342	12,0
n	140	4,5	148	4,8	140	4,6	135	4,7
o	24	0,8	20	0,6	23	0,8	20	0,7
ö	13	0,4	14	0,4	16	0,5	9	0,3
p	98	3,1	107	3,4	98	3,2	90	3,2
q	—	—	—	—	—	—	—	—
r	84	2,7	94	3,0	88	2,9	89	3,1
s	420	13,5	410	13,2	398	13,0	394	13,8
ß	—	—	—	—	—	—	—	—
t	138	4,4	135	4,3	130	4,3	117	4,1
u	10	0,3	9	0,3	8	0,3	9	0,3
ü	2	0,1	2	0,1	1	0,0	—	—
v	66	2,1	62	2,0	64	2,1	64	2,2
w	14	0,4	10	0,3	11	0,4	8	0,3
x	1	0,0	—	—	—	—	—	—
y	50	1,6	50	1,6	47	1,5	36	1,3
z	15	0,5	15	0,5	17	0,6	14	0,5
fehlend	24	0,8	27	0,9	33	1,1	37	1,3
gesamt	3 120	100,0	3 112	100,0	3 056	100,0	2 855	100,0

**Tabelle A.4:** Häufigkeiten von Co004 (Geburtstag)

	2008		2009		2011		2013	
	N	%	N	%	N	%	N	%
1	104	3,3	101	3,2	92	3,0	92	3,2
2	80	2,6	83	2,7	74	2,4	71	2,5
3	96	3,1	101	3,2	96	3,1	88	3,1
4	107	3,4	100	3,2	105	3,4	97	3,4
5	106	3,4	110	3,5	92	3,0	97	3,4
6	89	2,9	92	3,0	96	3,1	87	3,0
7	103	3,3	101	3,2	98	3,2	88	3,1
8	107	3,4	115	3,7	106	3,5	106	3,7
9	125	4,0	119	3,8	114	3,7	105	3,7
10	118	3,8	113	3,6	118	3,9	105	3,7
11	103	3,3	100	3,2	97	3,2	94	3,3
12	98	3,1	102	3,3	103	3,4	91	3,2
13	102	3,3	95	3,1	93	3,0	91	3,2
14	111	3,6	108	3,5	103	3,4	96	3,4
15	116	3,7	104	3,3	110	3,6	96	3,4
16	114	3,7	108	3,5	118	3,9	103	3,6
17	93	3,0	106	3,4	106	3,5	98	3,4
18	105	3,4	110	3,5	114	3,7	106	3,7
19	118	3,8	113	3,6	106	3,5	107	3,7
20	100	3,2	102	3,3	100	3,3	88	3,1
21	91	2,9	87	2,8	82	2,7	87	3,0
22	98	3,1	95	3,1	92	3,0	86	3,0
23	98	3,1	94	3,0	91	3,0	85	3,0
24	90	2,9	93	3,0	91	3,0	90	3,2
25	102	3,3	98	3,1	106	3,5	88	3,1
26	106	3,4	104	3,3	102	3,3	94	3,3
27	84	2,7	81	2,6	82	2,7	80	2,8
28	86	2,8	88	2,8	81	2,7	78	2,7
29	91	2,9	89	2,9	91	3,0	84	2,9
30	85	2,7	93	3,0	90	2,9	82	2,9
31	56	1,8	57	1,8	60	2,0	52	1,8
fehlend	38	1,2	50	1,6	47	1,5	43	1,5
gesamt	3 120	100,0	3 112	100,0	3 056	100,0	2 855	100,0

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

	2008		2009		2011		2013	
	N	%	N	%	N	%	N	%
a	1	0,0	1	0,0	—	—	2	0,1
ä	—	—	—	—	—	—	—	—
b	108	3,5	100	3,2	103	3,4	83	2,9
c	—	—	1	0,0	—	—	—	—
d	1019	32,7	1022	32,7	1014	33,2	956	33,5
e	5	0,2	12	0,4	15	0,5	11	0,4
f	2	0,1	—	—	2	0,1	—	—
g	1	0,0	3	0,1	1	0,0	1	0,0
h	3	0,1	—	—	3	0,1	1	0,0
i	—	—	—	—	—	—	—	—
j	—	—	—	—	—	—	—	—
k	1	0,0	—	—	1	0,0	1	0,0
l	2	0,1	1	0,0	—	—	—	—
m	3	0,1	4	0,1	3	0,1	3	0,1
n	1514	48,5	1536	49,2	1496	49,0	1429	50,1
o	—	—	—	—	—	—	—	—
ö	—	—	—	—	—	—	—	—
p	—	—	—	—	—	—	1	0,0
q	—	—	—	—	—	—	—	—
r	3	0,1	1	0,0	4	0,1	5	0,2
s	39	1,3	23	0,7	18	0,6	18	0,6
ß	—	—	—	—	—	—	—	—
t	29	0,9	27	0,9	30	1,0	31	1,1
u	3	0,1	6	0,2	5	0,2	4	0,1
ü	—	—	—	—	2	0,1	—	—
v	—	—	1	0,0	1	0,0	—	—
w	—	—	1	0,0	—	—	1	0,0
x	—	—	—	—	1	0,0	—	—
y	—	—	1	0,0	—	—	1	0,0
z	363	11,6	351	11,2	319	10,4	261	9,1
fehlend	24	0,8	21	0,7	38	1,2	46	1,6
gesamt	3 120	100,0	3 112	100,0	3 056	100,0	2 855	100,0

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

	2008		2009		2011		2013	
	N	%	N	%	N	%	N	%
a	3	0,1	3	0,1	—	—	2	0,1
ä	—	—	—	—	—	—	—	—
b	64	2,1	58	1,9	41	1,3	47	1,6
c	—	—	—	—	—	—	—	—
d	4	0,1	2	0,1	2	0,1	2	0,1
e	—	—	—	—	1	0,0	—	—
f	1	0,0	—	—	—	—	—	—
g	22	0,7	14	0,4	12	0,4	7	0,2
h	—	—	—	—	1	0,0	—	—
i	2	0,1	—	—	—	—	—	—
j	—	—	—	—	—	—	—	—
k	1	0,0	—	—	—	—	—	—
l	4	0,1	3	0,1	1	0,0	3	0,1
m	5	0,2	3	0,1	5	0,2	5	0,2
n	2013	64,5	2037	65,5	1975	64,6	1844	64,6
o	—	—	—	—	1	0,0	—	—
ö	—	—	—	—	—	—	—	—
p	—	—	—	—	—	—	—	—
q	—	—	—	—	—	—	—	—
r	—	—	—	—	1	0,0	—	—
s	5	0,2	6	0,2	1	0,0	2	0,1
ß	—	—	—	—	1	0,0	—	—
t	—	—	1	0,0	1	0,0	1	0,0
u	935	30,0	931	29,9	945	30,9	880	30,8
ü	—	—	—	—	1	0,0	3	0,1
v	—	—	—	—	—	—	—	—
w	1	0,0	—	—	—	—	1	0,0
x	—	—	1	0,0	—	—	—	—
y	1	0,0	—	—	1	0,0	—	—
z	25	0,8	29	0,9	24	0,8	16	0,6
fehlend	34	1,1	24	0,8	42	1,4	42	1,5
gesamt	3 120	100,0	3 112	100,0	3 056	100,0	2 855	100,0

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

	2009		2011		2013	
	N	%	N	%	N	%
a	169	5,4	148	4,8	141	4,9
ä	—	—	—	—	—	—
b	18	0,6	19	0,6	13	0,5
c	76	2,4	67	2,2	61	2,1
d	66	2,1	72	2,4	64	2,2
e	178	5,7	179	5,9	169	5,9
f	34	1,1	31	1,0	35	1,2
g	94	3,0	88	2,9	79	2,8
h	137	4,4	137	4,5	131	4,6
i	170	5,5	172	5,6	161	5,6
j	14	0,4	10	0,3	7	0,2
k	216	6,9	214	7,0	204	7,1
l	147	4,7	152	5,0	139	4,9
m	55	1,8	56	1,8	50	1,8
n	467	15,0	454	14,9	430	15,1
o	41	1,3	33	1,1	35	1,2
ö	1	0,0	2	0,1	2	0,1
p	31	1,0	24	0,8	17	0,6
q	1	0,0	1	0,0	2	0,1
r	441	14,2	436	14,3	427	15,0
s	239	7,7	239	7,8	215	7,5
ß	16	0,5	16	0,5	17	0,6
t	185	5,9	171	5,6	153	5,4
u	46	1,5	50	1,6	42	1,5
ü	4	0,1	3	0,1	3	0,1
v	7	0,2	5	0,2	5	0,2
w	23	0,7	31	1,0	23	0,8
x	5	0,2	5	0,2	6	0,2
y	52	1,7	51	1,7	45	1,6
z	156	5,0	154	5,0	142	5,0
fehlend	23	0,7	36	1,2	37	1,3
gesamt	3 112	100,0	3 056	100,0	2 855	100,0

**Tabelle A.8:** Häufigkeiten von Co007 in Duisburg 2008 (Befragungsteilnahme im Vorjahr ja/nein)

	Häufigkeit	%
ja	2 680	85,9
nein	155	5,0
k. A.	285	9,1
unplausibel	—	—
Gesamt	3 120	100,0

**Tabelle A.9:** Häufigkeiten von Co007 in Duisburg 2009 (Befragungsteilnahme im Vorjahr ja/nein)

	Häufigkeit	%
ja	1 924	61,8
nein	81	2,6
k. A.	1 107	35,6
unplausibel	—	—
Gesamt	3 112	100,0

**Tabelle A.10:** Häufigkeiten von Co007 in Duisburg 2011 (Befragungsteilnahme im Vorjahr ja/nein)

	Häufigkeit	%
ja	2 589	84,7
nein	364	11,9
k. A.	102	3,3
unplausibel	1	0,0
Gesamt	3 056	100,0

**Tabelle A.11:** Häufigkeiten von Co007 in Duisburg 2013 (Befragungsteilnahme im Vorjahr ja/nein)

	Häufigkeit	%
ja	2 441	85,5
nein	341	11,9
k. A.	71	2,5
unplausibel	2	0,1
Gesamt	2 855	100,0

**Tabelle A.12:** Häufigkeiten von Co009 in Duisburg 2008 (Sitzengeblieben ja/nein)

	Häufigkeit	%
ja	217	7,0
nein	1 612	51,7
k. A.	1 291	41,4
unplausibel	—	—
Gesamt	3 120	100,0

**Tabelle A.13:** Häufigkeiten der Gesamtcodes in Duisburg 2008

	Häufigkeit	%
1	3 056	98,2
2	37	1,2
fehlend	18	0,6
Gesamt	3 120	100,0

**Tabelle A.14:** Häufigkeiten der Gesamtcodes in Duisburg 2009

	Häufigkeit	%
1	3 095	99,4
2	2	0,1
fehlend	15	0,5
Gesamt	3 112	100,0

**Tabelle A.15:** Häufigkeiten der Gesamtcodes in Duisburg 2011

	Häufigkeit	%
1	3 001	98,2
2	26	0,9
fehlend	29	0,9
Gesamt	3 056	100,0

**Tabelle A.16:** *Häufigkeiten der Gesamtcodes in Duisburg 2013*

	Häufigkeit	%
1	2810	98,4
2	14	0,5
fehlend	31	1,1
Gesamt	2855	100,0