

---

**Methodendokumentation  
der kriminologischen Schülerbefragung  
in Duisburg 2002 bis 2007  
– Sechs-Wellen-Panel –**

Daniela Pollich

**Schriftenreihe:  
Jugendkriminalität in der modernen Stadt – Methoden  
Nr. 16 / 2010**

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).



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
<b>2</b>	<b>Studiendesign</b>	<b>3</b>
<b>3</b>	<b>Erstellung der Paneldaten</b>	<b>7</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>21</b>
5.1	Die Panelquerschnitte . . . . .	21
5.1.1	Panelquerschnitt 2004 ( $P_{Q_{t_3}}$ ) . . . . .	21
5.1.2	Panelquerschnitt 2005 ( $P_{Q_{t_4}}$ ) . . . . .	24
5.1.3	Panelquerschnitt 2006 ( $P_{Q_{t_5}}$ ) . . . . .	26
5.1.4	Panelquerschnitt 2007 ( $P_{Q_{t_6}}$ ) . . . . .	28
5.2	Die Zwei-Wellen-Panel . . . . .	31
5.2.1	Zwei-Wellen-Panel 2005/2006 ( $P_{t_{4,5}}$ ) . . . . .	31
5.2.2	Zwei-Wellen-Panel 2006/2007 ( $P_{t_{5,6}}$ ) . . . . .	34
5.2.3	Zwei-Wellen-Panel 2004/2006 ( $P_{t_{3,5}}$ ) . . . . .	36
5.2.4	Zwei-Wellen-Panel 2005/2007 ( $P_{t_{4,6}}$ ) . . . . .	39
5.3	Die Drei-Wellen-Panel . . . . .	42
5.3.1	Drei-Wellen-Panel 2004 bis 2006 ( $P_{t_{3,4,5}}$ ) . . . . .	42
5.3.2	Drei-Wellen-Panel 2005 bis 2007 ( $P_{t_{4,5,6}}$ ) . . . . .	47
5.3.3	Drei-Wellen-Panel 2002, 2004 und 2006 ( $P_{t_{1,3,5}}$ ) . . . . .	50
5.3.4	Drei-Wellen-Panel 2003, 2004 und 2006 ( $P_{t_{2,3,5}}$ ) . . . . .	55
5.3.5	Drei-Wellen-Panel 2003, 2005 und 2006 ( $P_{t_{2,4,5}}$ ) . . . . .	59
5.3.6	Drei-Wellen-Panel 2003, 2005 und 2007 ( $P_{t_{2,4,6}}$ ) . . . . .	62
5.3.7	Drei-Wellen-Panel 2004, 2005 und 2007 ( $P_{t_{3,4,6}}$ ) . . . . .	66
5.3.8	Drei-Wellen-Panel 2004, 2006 und 2007 ( $P_{t_{3,5,6}}$ ) . . . . .	70
5.3.9	Drei-Wellen-Panel 2002, 2005 und 2006 ( $P_{t_{1,4,5}}$ ) . . . . .	73
5.3.10	Drei-Wellen-Panel 2002, 2005 und 2007 ( $P_{t_{1,4,6}}$ ) . . . . .	78
5.4	Die Vier-Wellen-Panel . . . . .	81
5.4.1	Vier-Wellen-Panel 2003 bis 2006 ( $P_{t_{2,3,4,5}}$ ) . . . . .	81
5.4.2	Vier-Wellen-Panel 2004 bis 2007 ( $P_{t_{3,4,5,6}}$ ) . . . . .	87

5.4.3	Vier-Wellen-Panel 2002, 2003, 2004 und 2006 ( $P_{t_{1,2,3,5}}$ ) . . . . .	92
5.4.4	Vier-Wellen-Panel 2002, 2003, 2005 und 2006 ( $P_{t_{1,2,4,5}}$ ) . . . . .	97
5.4.5	Vier-Wellen-Panel 2002, 2003, 2005 und 2007 ( $P_{t_{1,2,4,6}}$ ) . . . . .	102
5.4.6	Vier-Wellen-Panel 2002, 2004, 2005 und 2006 ( $P_{t_{1,3,4,5}}$ ) . . . . .	107
5.4.7	Vier-Wellen-Panel 2002, 2004, 2005 und 2007 ( $P_{t_{1,3,4,6}}$ ) . . . . .	112
5.4.8	Vier-Wellen-Panel 2002, 2004, 2006 und 2007 ( $P_{t_{1,3,5,6}}$ ) . . . . .	117
5.4.9	Vier-Wellen-Panel 2003, 2004, 2005 und 2007 ( $P_{t_{2,3,4,6}}$ ) . . . . .	122
5.4.10	Vier-Wellen-Panel 2003, 2004, 2006 und 2007 ( $P_{t_{2,3,5,6}}$ ) . . . . .	127
5.4.11	Vier-Wellen-Panel 2003, 2005, 2006 und 2007 ( $P_{t_{2,4,5,6}}$ ) . . . . .	132
5.4.12	Vier-Wellen-Panel 2002, 2005, 2006 und 2007 ( $P_{t_{1,4,5,6}}$ ) . . . . .	137
5.5	Die Fünf-Wellen-Panel . . . . .	142
5.5.1	Fünf-Wellen-Panel 2002 bis 2006 ( $P_{t_{1,2,3,4,5}}$ ) . . . . .	142
5.5.2	Fünf-Wellen-Panel 2003 bis 2007 ( $P_{t_{2,3,4,5,6}}$ ) . . . . .	147
5.5.3	Fünf-Wellen-Panel 2002, 2003, 2004, 2005 und 2007 ( $P_{t_{1,2,3,4,6}}$ ) .	152
5.5.4	Fünf-Wellen-Panel 2002, 2003, 2004, 2006 und 2007 ( $P_{t_{1,2,3,5,6}}$ ) .	157
5.5.5	Fünf-Wellen-Panel 2002, 2003, 2005, 2006 und 2007 ( $P_{t_{1,2,4,5,6}}$ ) .	162
5.5.6	Fünf-Wellen-Panel 2002, 2004, 2005, 2006 und 2007 ( $P_{t_{1,3,4,5,6}}$ ) .	167
5.6	Das Sechs-Wellen-Panel . . . . .	172
<b>6</b>	<b>Zusammenfassung und Erklärung</b>	<b>179</b>
	<b>Literaturverzeichnis</b>	<b>181</b>
	<b>Anhang</b>	<b>183</b>

## Tabellenverzeichnis

3.1	Durchgeführte Kontrollschritte 2005/2006 . . . . .	11
3.2	Durchgeführte Kontrollschritte 2006/2007 . . . . .	11
3.3	Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2005/2006	13
3.4	Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2006/2007	14
4.1	Bezeichnungen und Fallzahlen aller Paneldatensätze 2002 bis 2007 . . .	17
4.2	Ausschöpfungen der durchgehenden Paneldatensätze . . . . .	19
5.1	Zusammensetzung nach Geschlecht $P_{Q_{t_3}}$ . . . . .	22
5.2	Altersdurchschnitt $P_{Q_{t_3}}$ . . . . .	22
5.3	Zusammensetzung $P_{Q_{t_3}}$ nach Geschlecht und Schulform (Zeilenproz.)	23
5.4	Zusammensetzung $P_{Q_{t_3}}$ nach Geschlecht und Schulform (Spaltenproz.)	23
5.5	Zusammensetzung nach Geschlecht $P_{Q_{t_4}}$ . . . . .	24
5.6	Altersdurchschnitt $P_{Q_{t_4}}$ . . . . .	24
5.7	Zusammensetzung $P_{Q_{t_4}}$ nach Geschlecht und Schulform (Zeilenproz.)	25
5.8	Zusammensetzung $P_{Q_{t_4}}$ nach Geschlecht und Schulform (Spaltenproz.)	25
5.9	Zusammensetzung nach Geschlecht $P_{Q_{t_5}}$ . . . . .	27
5.10	Altersdurchschnitt $P_{Q_{t_5}}$ . . . . .	27
5.11	Zusammensetzung $P_{Q_{t_5}}$ nach Geschlecht und Schulform (Zeilenproz.)	27
5.12	Zusammensetzung $P_{Q_{t_5}}$ nach Geschlecht und Schulform (Spaltenproz.)	28
5.13	Zusammensetzung nach Geschlecht $P_{Q_{t_6}}$ . . . . .	29
5.14	Altersdurchschnitt $P_{Q_{t_6}}$ . . . . .	29
5.15	Zusammensetzung $P_{Q_{t_6}}$ nach Geschlecht und Schulform (Zeilenproz.)	30
5.16	Zusammensetzung $P_{Q_{t_6}}$ nach Geschlecht und Schulform (Spaltenproz.)	30
5.17	Zusammensetzung nach Geschlecht $P_{t_{4,5}}$ . . . . .	31
5.18	Altersdurchschnitt $P_{t_{4,5}}$ . . . . .	32
5.19	Zusammensetzung $P_{t_{4,5}}$ nach Geschlecht und Schulform (Zeilenproz.)	32
5.20	Zusammensetzung $P_{t_{4,5}}$ nach Geschlecht und Schulform (Spaltenproz.)	33
5.21	Zusammensetzung nach Geschlecht $P_{t_{5,6}}$ . . . . .	34
5.22	Altersdurchschnitt $P_{t_{5,6}}$ . . . . .	34
5.23	Zusammensetzung $P_{t_{5,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	35
5.24	Zusammensetzung $P_{t_{5,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	36
5.25	Zusammensetzung nach Geschlecht $P_{t_{3,5}}$ . . . . .	37
5.26	Altersdurchschnitt $P_{t_{3,5}}$ . . . . .	37



5.27	Zusammensetzung $P_{t_{3,5}}$ nach Geschlecht und Schulform (Zeilenproz.)	38
5.28	Zusammensetzung $P_{t_{3,5}}$ nach Geschlecht und Schulform (Spaltenproz.)	39
5.29	Zusammensetzung nach Geschlecht $P_{t_{4,6}}$	40
5.30	Altersdurchschnitt $P_{t_{4,6}}$	40
5.31	Zusammensetzung $P_{t_{4,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	41
5.32	Zusammensetzung $P_{t_{4,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	42
5.33	Zusammensetzung nach Geschlecht $P_{t_{3,4,5}}$	43
5.34	Altersdurchschnitt $P_{t_{3,4,5}}$	43
5.35	Zusammensetzung $P_{t_{3,4,5}}$ nach Geschlecht und Schulform (Zeilenproz.)	44
5.36	Zusammensetzung $P_{t_{3,4,5}}$ nach Geschlecht und Schulform (Spaltenproz.)	46
5.37	Zusammensetzung nach Geschlecht $P_{t_{4,5,6}}$	47
5.38	Altersdurchschnitt $P_{t_{4,5,6}}$	47
5.39	Zusammensetzung $P_{t_{4,5,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	48
5.40	Zusammensetzung $P_{t_{4,5,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	50
5.41	Zusammensetzung nach Geschlecht $P_{t_{1,3,5}}$	50
5.42	Altersdurchschnitt $P_{t_{1,3,5}}$	51
5.43	Zusammensetzung $P_{t_{1,3,5}}$ nach Geschlecht und Schulform (Zeilenproz.)	53
5.44	Zusammensetzung $P_{t_{1,3,5}}$ nach Geschlecht und Schulform (Spaltenproz.)	54
5.45	Zusammensetzung nach Geschlecht $P_{t_{2,3,5}}$	55
5.46	Altersdurchschnitt $P_{t_{2,3,5}}$	55
5.47	Zusammensetzung $P_{t_{2,3,5}}$ nach Geschlecht und Schulform (Zeilenproz.)	56
5.48	Zusammensetzung $P_{t_{2,3,5}}$ nach Geschlecht und Schulform (Spaltenproz.)	58
5.49	Zusammensetzung nach Geschlecht $P_{t_{2,4,5}}$	59
5.50	Altersdurchschnitt $P_{t_{2,4,5}}$	59
5.51	Zusammensetzung $P_{t_{2,4,5}}$ nach Geschlecht und Schulform (Zeilenproz.)	60
5.52	Zusammensetzung $P_{t_{2,4,5}}$ nach Geschlecht und Schulform (Spaltenproz.)	62
5.53	Zusammensetzung nach Geschlecht $P_{t_{2,4,6}}$	63
5.54	Altersdurchschnitt $P_{t_{2,4,6}}$	63
5.55	Zusammensetzung $P_{t_{2,4,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	64
5.56	Zusammensetzung $P_{t_{2,4,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	66
5.57	Zusammensetzung nach Geschlecht $P_{t_{3,4,6}}$	67
5.58	Altersdurchschnitt $P_{t_{3,4,6}}$	67
5.59	Zusammensetzung $P_{t_{3,4,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	68
5.60	Zusammensetzung $P_{t_{3,4,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	69
5.61	Zusammensetzung nach Geschlecht $P_{t_{3,5,6}}$	70
5.62	Altersdurchschnitt $P_{t_{3,5,6}}$	71
5.63	Zusammensetzung $P_{t_{3,5,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	72
5.64	Zusammensetzung $P_{t_{3,5,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	73
5.65	Zusammensetzung nach Geschlecht $P_{t_{1,4,5}}$	74
5.66	Altersdurchschnitt $P_{t_{1,4,5}}$	74

5.67	Zusammensetzung $P_{1,4,5}$ nach Geschlecht und Schulform (Zeilenproz.)	75
5.68	Zusammensetzung $P_{1,4,5}$ nach Geschlecht und Schulform (Spaltenproz.)	77
5.69	Zusammensetzung nach Geschlecht $P_{1,4,6}$	78
5.70	Altersdurchschnitt $P_{1,4,6}$	78
5.71	Zusammensetzung $P_{1,4,6}$ nach Geschlecht und Schulform (Zeilenproz.)	79
5.72	Zusammensetzung $P_{1,4,6}$ nach Geschlecht und Schulform (Spaltenproz.)	81
5.73	Zusammensetzung nach Geschlecht $P_{2,3,4,5}$	82
5.74	Altersdurchschnitt $P_{2,3,4,5}$	82
5.75	Zusammensetzung $P_{2,3,4,5}$ nach Geschlecht und Schulform (Zeilenproz.)	85
5.76	Zusammensetzung $P_{2,3,4,5}$ nach Geschlecht und Schulform (Spaltenproz.)	86
5.77	Zusammensetzung nach Geschlecht $P_{3,4,5,6}$	87
5.78	Altersdurchschnitt $P_{3,4,5,6}$	88
5.79	Zusammensetzung $P_{3,4,5,6}$ nach Geschlecht und Schulform (Zeilenproz.)	90
5.80	Zusammensetzung $P_{3,4,5,6}$ nach Geschlecht und Schulform (Spaltenproz.)	91
5.81	Zusammensetzung nach Geschlecht $P_{1,2,3,5}$	92
5.82	Altersdurchschnitt $P_{1,2,3,5}$	93
5.83	Zusammensetzung $P_{1,2,3,5}$ nach Geschlecht und Schulform (Zeilenproz.)	95
5.84	Zusammensetzung $P_{1,2,3,5}$ nach Geschlecht und Schulform (Spaltenproz.)	96
5.85	Zusammensetzung nach Geschlecht $P_{1,2,4,5}$	97
5.86	Altersdurchschnitt $P_{1,2,4,5}$	98
5.87	Zusammensetzung $P_{1,2,4,5}$ nach Geschlecht und Schulform (Zeilenproz.)	100
5.88	Zusammensetzung $P_{1,2,4,5}$ nach Geschlecht und Schulform (Spaltenproz.)	101
5.89	Zusammensetzung nach Geschlecht $P_{1,2,4,6}$	102
5.90	Altersdurchschnitt $P_{1,2,4,6}$	103
5.91	Zusammensetzung $P_{1,2,4,6}$ nach Geschlecht und Schulform (Zeilenproz.)	105
5.92	Zusammensetzung $P_{1,2,4,6}$ nach Geschlecht und Schulform (Spaltenproz.)	106
5.93	Zusammensetzung nach Geschlecht $P_{1,3,4,5}$	107
5.94	Altersdurchschnitt $P_{1,3,4,5}$	108
5.95	Zusammensetzung $P_{1,3,4,5}$ nach Geschlecht und Schulform (Zeilenproz.)	110
5.96	Zusammensetzung $P_{1,3,4,5}$ nach Geschlecht und Schulform (Spaltenproz.)	111
5.97	Zusammensetzung nach Geschlecht $P_{1,3,4,6}$	112
5.98	Altersdurchschnitt $P_{1,3,4,6}$	113
5.99	Zusammensetzung $P_{1,3,4,6}$ nach Geschlecht und Schulform (Zeilenproz.)	115
5.100	Zusammensetzung $P_{1,3,4,6}$ nach Geschlecht und Schulform (Spaltenproz.)	116
5.101	Zusammensetzung nach Geschlecht $P_{1,3,5,6}$	117
5.102	Altersdurchschnitt $P_{1,3,5,6}$	118
5.103	Zusammensetzung $P_{1,3,5,6}$ nach Geschlecht und Schulform (Zeilenproz.)	120
5.104	Zusammensetzung $P_{1,3,5,6}$ nach Geschlecht und Schulform (Spaltenproz.)	121
5.105	Zusammensetzung nach Geschlecht $P_{2,3,4,6}$	122
5.106	Altersdurchschnitt $P_{2,3,4,6}$	123

5.107	Zusammensetzung $P_{t_{2,3,4,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	125
5.108	Zusammensetzung $P_{t_{2,3,4,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	126
5.109	Zusammensetzung nach Geschlecht $P_{t_{2,3,5,6}}$ . . . . .	127
5.110	Altersdurchschnitt $P_{t_{2,3,5,6}}$ . . . . .	128
5.111	Zusammensetzung $P_{t_{2,3,5,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	130
5.112	Zusammensetzung $P_{t_{2,3,5,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	131
5.113	Zusammensetzung nach Geschlecht $P_{t_{2,4,5,6}}$ . . . . .	132
5.114	Altersdurchschnitt $P_{t_{2,4,5,6}}$ . . . . .	133
5.115	Zusammensetzung $P_{t_{2,4,5,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	135
5.116	Zusammensetzung $P_{t_{2,4,5,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	136
5.117	Zusammensetzung nach Geschlecht $P_{t_{1,4,5,6}}$ . . . . .	137
5.118	Altersdurchschnitt $P_{t_{1,4,5,6}}$ . . . . .	138
5.119	Zusammensetzung $P_{t_{1,4,5,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	140
5.120	Zusammensetzung $P_{t_{1,4,5,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	141
5.121	Zusammensetzung nach Geschlecht $P_{t_{1,2,3,4,5}}$ . . . . .	142
5.122	Altersdurchschnitt $P_{t_{1,2,3,4,5}}$ . . . . .	143
5.123	Zusammensetzung $P_{t_{1,2,3,4,5}}$ nach Geschlecht und Schulform (Zeilenproz.)	145
5.124	Zusammensetzung $P_{t_{1,2,3,4,5}}$ nach Geschlecht und Schulform (Spaltenproz.)	146
5.125	Zusammensetzung nach Geschlecht $P_{t_{2,3,4,5,6}}$ . . . . .	147
5.126	Altersdurchschnitt $P_{t_{2,3,4,5,6}}$ . . . . .	148
5.127	Zusammensetzung $P_{t_{2,3,4,5,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	150
5.128	Zusammensetzung $P_{t_{2,3,4,5,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	151
5.129	Zusammensetzung nach Geschlecht $P_{t_{1,2,3,4,6}}$ . . . . .	152
5.130	Altersdurchschnitt $P_{t_{1,2,3,4,6}}$ . . . . .	153
5.131	Zusammensetzung $P_{t_{1,2,3,4,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	155
5.132	Zusammensetzung $P_{t_{1,2,3,4,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	156
5.133	Zusammensetzung nach Geschlecht $P_{t_{1,2,3,5,6}}$ . . . . .	157
5.134	Altersdurchschnitt $P_{t_{1,2,3,5,6}}$ . . . . .	158
5.135	Zusammensetzung $P_{t_{1,2,3,5,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	160
5.136	Zusammensetzung $P_{t_{1,2,3,5,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	161
5.137	Zusammensetzung nach Geschlecht $P_{t_{1,2,4,5,6}}$ . . . . .	162
5.138	Altersdurchschnitt $P_{t_{1,2,4,5,6}}$ . . . . .	163
5.139	Zusammensetzung $P_{t_{1,2,4,5,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	165
5.140	Zusammensetzung $P_{t_{1,2,4,5,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	166
5.141	Zusammensetzung nach Geschlecht $P_{t_{1,3,4,5,6}}$ . . . . .	167
5.142	Altersdurchschnitt $P_{t_{1,3,4,5,6}}$ . . . . .	168
5.143	Zusammensetzung $P_{t_{1,3,4,5,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	170
5.144	Zusammensetzung $P_{t_{1,3,4,5,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	171
5.145	Zusammensetzung nach Geschlecht $P_{t_{1,2,3,4,5,6}}$ . . . . .	172
5.146	Altersdurchschnitt $P_{t_{1,2,3,4,5,6}}$ . . . . .	173

5.147	Zusammensetzung $P_{t_{1,2,3,4,5,6}}$ nach Geschlecht und Schulform (Zeilenproz.)	176
5.148	Zusammensetzung $P_{t_{1,2,3,4,5,6}}$ nach Geschlecht und Schulform (Spaltenproz.)	177
A.1	Häufigkeiten von Co001 (erster Buchstabe Vorname Vater) . . . . .	185
A.2	Häufigkeiten von Co002 (erster Buchstabe Vorname Mutter) . . . . .	186
A.3	Häufigkeiten von Co003 (erster Buchstabe eigener Vorname) . . . . .	187
A.4	Häufigkeiten von Co004 (Geburtstag) . . . . .	188
A.5	Häufigkeiten von Co005 (Haarfarbe) . . . . .	189
A.6	Häufigkeiten von Co006 (Augenfarbe) . . . . .	190
A.7	Häufigkeiten von Co007 2006 (Befragungsteilnahme im Vorjahr ja/nein)	191
A.8	Häufigkeiten von Co007 2007 (Befragungsteilnahme im Vorjahr ja/nein)	191
A.9	Häufigkeiten von Co009 2006 (Sitzengeblieben ja/nein) . . . . .	191
A.10	Häufigkeiten von Co009 2007 (Sitzengeblieben ja/nein) . . . . .	191
A.11	Häufigkeiten der Gesamtcodes 2006 . . . . .	192
A.12	Häufigkeiten der Gesamtcodes 2007 . . . . .	192



# Abbildungsverzeichnis

2.1	Erhebungsdesign der Längsschnittstudie 2000–2009 . . . . .	4
A.1	Codeblatt 2006 . . . . .	183
A.2	Codeblatt 2007 . . . . .	184



# 1 Einleitung

Die Ziele des kriminologisch und kriminalsoziologisch ausgerichteten DFG-Projekts *Kriminalität in der modernen Stadt*<sup>1</sup> liegen darin, Entstehung sowie Verläufe devianter bzw. delinquenter Handlungsstile näher zu beleuchten. Das theoretische Analysemodell, in welches die Untersuchungen eingebettet sind, berücksichtigt dabei sowohl die Makro- als auch die Meso- und Mikroebene sozialer Phänomene (vgl. Boers & Reinecke, 2007). Die erhobenen Daten stammen aus drei hinsichtlich ihrer Größe und Sozialstruktur recht unterschiedlichen deutschen Städten, nämlich Bocholt, Münster und Duisburg.

Besondere Möglichkeiten der Auswertung kausaler Zusammenhänge ergeben sich durch das kombinierte *Panel-Kohorten-Design* der Studie (vgl. hierzu genauer Kapitel 2).

Der eingesetzte Fragebogen zu Opferwerdung, selbstberichteter Delinquenz, Erziehungsstilen, Konfliktverhalten, Kriminalitätseinstellungen, Lebens-, Freizeit-, und Konsumstilen sowie weiteren Themenbereichen wurde für eine schriftliche, anonyme Befragung konzipiert. So weit wie möglich wurden Fragen verwendet, die bereits in anderen Jugendstudien eingesetzt wurden. Die Fragebögen der einzelnen Jahre sind aufgrund des Panelansatzes der Studie weitgehend identisch. Detaillierte Darstellungen des Erhebungsinstruments, der verwendeten Skalen oder eine Validierung eben dieser sind im vorliegenden Bericht nicht vorgesehen. Ebenfalls außen vor gelassen werden genaue Beschreibungen der einbezogenen Querschnittsdatensätze und deren Erhebung sowie weitergehende inhaltliche Analysen und Ergebnisse.

Der vorliegende Bericht befasst sich ausschließlich mit der jüngeren Kohorte der Duisburger Erhebung, bzw. dem bislang daraus entstandenen Sechs-Wellen-Paneldatensatz der Jahre 2002 bis 2007.

---

<sup>1</sup> 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, 6-3, 6-4, 6-5, 6-6 und Re 832/4-1, 4-2, 4-3, 4-4, 4-5, 4-6.





## 2 Studiendesign

Beginnend mit dem Jahr 2000 wurden in Münster bis zum Jahr 2003 in Jahresabständen schulische Befragungen durchgeführt.<sup>2</sup> Die Münsteraner Schülerinnen und Schüler wurden auf diesem Wege von der siebten bis zur zehnten Jahrgangsstufe befragt. Seit dem Jahr 2002 wurden überdies in Duisburg jährlich Jugendliche befragt. Dabei umfasste die Erhebung zum ersten Befragungszeitpunkt sowohl Schülerinnen und Schüler der siebten als auch der neunten Jahrgangsstufe. Während die Schülerinnen und Schüler der damals siebten Jahrgangsstufe inzwischen über die Beendigung ihrer Schullaufbahn hinaus (2009) jährlich befragt wurden, wurden diejenigen Jugendlichen, die 2002 die neunte Klasse besucht haben nur noch ein weiteres Mal in 2003 in die Erhebung einbezogen.

Aus den beschriebenen Erhebungen entstanden bisher ein Vier-Wellen-Paneldatensatz der Münsteraner Schüler,<sup>3</sup> ein Sechs-Wellen-Paneldatensatz der jüngeren Duisburger Kohorte<sup>4</sup> sowie ein Zwei-Wellen-Paneldatensatz der älteren Duisburger Kohorte.<sup>5</sup>

Neben den mehrfach befragten Kohorten wurden auch einige Querschnittsdatensätze ohne Anschlussbefragung erhoben. So wurden in Münster im Jahr 2000 auch die neunte und die elfte Jahrgangsstufe befragt, was einen Vergleich verschiedener Altersstufen gestattet. Zudem erfolgte im Jahr 2001 eine Querschnittserhebung der siebten, neunten und elften Jahrgangsstufen in Bocholt, um einen Stadt-Land-Vergleich zu ermöglichen.<sup>6</sup> Im Jahr 2003 wurden außerdem Zusatzerhebungen in Münster und Duisburg durchgeführt, anhand derer Lehrkräfte zur Durchführung von Präventionsmaßnahmen in Schulen befragt wurden.<sup>7</sup>

Der gesamte Erhebungsplan kann Abbildung 2.1 entnommen werden. Dort wird ersichtlich, dass anhand der Daten verschiedene Altersgruppen innerhalb eines Querschnitts, gleiche Altersgruppen zu verschiedenen Erhebungszeitpunkten sowie Vergleiche innerhalb von Lebensverläufen vorgenommen werden können. Nur ein solches Paneldesign, also die jährliche Wiederbefragung *derselben* Personen kann dabei die Analysierbarkeit sowohl auf der Aggregatebene als auch auf der individuellen Ebene gewährleisten (vgl. Engel & Reinecke, 1994, S. 5 f.).

---

2 Zur Dokumentation der einzelnen Querschnittserhebungen in Münster siehe genauer Motzke & Wittenberg, 2004; Wittenberg, 2004a; Wittenberg, 2004b; Wittenberg, 2004c.

3 Zur Dokumentation des Münsteraner Vier-Wellen-Panels siehe Pöge, 2005a.

4 Für die Duisburger Erhebung ist bislang eine Dokumentation des Vier-Wellen-Panels verfügbar (vgl. Pöge, 2007).

5 Siehe hierzu genauer Pollich, 2007.

6 Zur Dokumentation der Bocholter Erhebung siehe Wittenberg & Hilfert, 2004.

7 Eine Dokumentation dieser Erhebung findet sich in Brondies, 2004b.

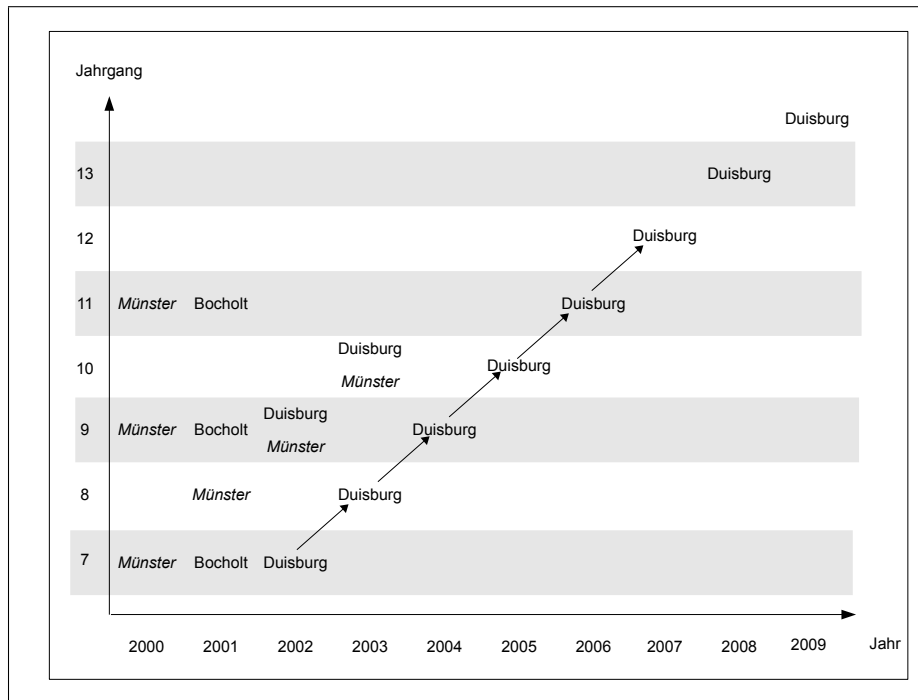


Abbildung 2.1: Erhebungsdesign der Längsschnittstudie 2000–2009

Wie bereits erwähnt, gingen in die im Folgenden beschriebenen Paneldatensätze ausschließlich Daten aus Duisburg ein. Die zur Konstruktion des Sechs-Wellen-Panels herangezogenen Datensätze sind in Abbildung 2.1 mit Pfeilen verbunden.

In der *ersten* Duisburger Schülerbefragung im Jahr 2002 wurde eine Vollerhebung der Schülerinnen und Schüler der siebten und neunten Jahrgangsstufen angestrebt, um eine möglichst umfassende Ausgangsgrundlage für den Paneldatensatz zu schaffen und Auswirkungen der Panelmortalität nach Möglichkeit einzudämmen (vgl. hierzu genauer Motzke & Brondies, 2004). Die *zweite* Duisburger Schülerbefragung im Jahr 2003 strebte wie die erste Erhebungswelle nochmals eine Vollerhebung der Schülerinnen und Schüler der inzwischen achten und zehnten Jahrgangsstufen an (vgl. hierzu genauer Brondies, 2004a). Nachdem die ältere der beiden Duisburger Kohorten nach 2003 nicht weiter befragt wurde, war es das Ziel der *dritten* und *vierten* Schülerbefragungen der Jahre 2004 und 2005 (vgl. hierzu genauer Hilfert, 2005 und Kunadt, 2006), möglichst viele derjenigen Schülerinnen und Schüler der neunten bzw. zehnten Jahrgangsstufe, die bereits an der ersten und/oder zweiten Erhebung teilgenommen hatten, wieder zu befragen.

Mit dem Erreichen der elften Jahrgangsstufe im Jahr 2006 wurde es zunehmend schwieriger, die Erhebung im Klassenverband durchzuführen, da viele Befragte bereits ihre schulische Laufbahn beendet hatten. Für die Studie bedeutete dies mit der *fünften* Erhebungswelle eine teilweise Umstellung auf eine postalische Befragungsweise (damals in Form einer Nachbefragung), um möglichst viele der bereits befragten Jugendlichen

---

wieder zu erreichen und damit den weiteren Ausbau des Paneldatensatzes der jüngeren Duisburger Kohorte nicht zu gefährden. Einmalig wurden in diesem Jahr zudem Schüler in denjenigen Kursen der Duisburger Berufskollegs im Klassenverband befragt, in denen mit hoher Wahrscheinlichkeit Abgänger aus den vorher in der Stichprobe befindlichen Haupt- und Realschulen zu finden waren. Dies geschah zu dem Zweck, die panelrelevanten Personen, die ihre Schullaufbahn bereits beendet hatten, in den Berufsschulen wieder für eine Teilnahme gewinnen zu können. Da durch diese Maßnahme jedoch eine erhebliche Vergrößerung des Querschnittsdatensatzes bei gleichzeitig fehlender Panelrelevanz eines großen Teils der befragten Berufsschüler resultierte, blieb die Zusatzerhebung einmalig (vgl. hierzu genauer Bentrup, 2007). Die *sechste* Erhebungswelle stammt aus dem Jahr 2007. Die Vorgehensweise bei der Erhebung verschob sich in diesem Jahr – mit zunehmender Beendigung der Schulausbildung in der befragten Kohorte – weiter in Richtung einer postalischen Befragung. Um eine maximale Anzahl der bereits im Paneldatensatz befindlichen Jugendlichen wieder befragen zu können, wurde zudem erstmals eine Face-to-face-Befragung derjenigen Teilnehmer durchgeführt, die weder in der Schule noch auf postalischem Wege erreichbar waren (vgl. hierzu genauer Bentrup, 2009).

Die *siebte* Duisburger Schülerbefragung aus dem Jahr 2008 – die für den im Folgenden beschriebenen Sechs-Wellen-Paneldatensatz nicht relevant ist – wurde erstmals ausschließlich auf postalischem Weg durchgeführt; ergänzt wurde sie wieder durch eine anschließende Face-to-face-Erhebung. Zwar besuchten in diesem Jahr zumindest die Gymnasiasten noch regelmäßig eine allgemeinbildende Schule, jedoch wurde aufgrund der im Erhebungszeitraum anstehenden Abiturprüfungen auf eine Befragung in den Schulen verzichtet. Die bislang aktuellste Erhebung – und damit die *achte* Erhebungswelle – aus dem Jahr 2009 umfasste ebenfalls eine postalische Befragung und eine ergänzende Face-to-face-Befragung. Im Jahr 2010 wird erstmalig eine Befragungspause stattfinden, um im Jahr 2011 die *neunte* Erhebungswelle auf postalischem Wege durchzuführen.

Der vorliegende Bericht konzentriert sich auf die Entstehung und die Repräsentativität des Sechs-Wellen-Paneldatensatzes, der sich aus Daten der Erhebungsjahre 2002 bis 2007 konstituiert. Da der Vier-Wellen-Paneldatensatz selbiger Studie (Erhebungsjahre 2002 bis 2005) bereits ausführlich dokumentiert ist (vgl. Pöge, 2007), beschränkt sich der vorliegende Bericht auf die Erweiterungen zum Sechs-Wellen-Paneldatensatz einschließlich der neu entstandenen Teildatensätze.



### 3 Erstellung der Paneldaten

Mit der Durchführung einer Panelstudie sind wegen 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 bzw. – ab der Befragung in 2003 – sechs 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, so dass 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>8</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

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 Codeblättern zum einen einfachere Fragen mit einer gleichzeitig größeren Antwortvarianz gestellt, zum anderen wurde im Jahr 2003 der Code – ebenso wie in Münster – um eine Frage erweitert und das Layout des gesamten Blattes überarbeitet. Das Layout

---

<sup>8</sup> Abdrucke der Codeblätter der Erhebungsjahre 2006 und 2007 finden sich in Abbildungen A.1 und A.2 (im Anhang). Codeblätter früherer Erhebungen finden sich bei Pöge (2007, S. 88 ff.)

wurde dergestalt geändert, dass kein handschriftliches Ausfüllen mehr erfolgen musste, sondern alle Antwortvorgaben zum Ankreuzen aufgeführt wurden.

Seit dem Befragungsjahr 2003 wurden zusätzlich auch die folgenden Fragen gestellt:

- *Co007*: Befragungsteilnahme im Vorjahr (ja/nein)
- *Co008*: Schulwechsel im vergangenen Jahr (ja/nein)
- *Co009*: Nichtversetzung im vergangenen Jahr (ja/nein)

Mit Beendigung der Schullaufbahn der meisten Befragten wurde seit der Befragungswelle  $t_5$  (2006) auf die Erhebung von *Co008* wieder verzichtet; somit war diese Information für die hier beschriebenen Zuordnungen nicht mehr verfügbar. *Co009* wurde seit 2006 nur noch für diejenigen Befragten erhoben, die ein Gymnasium oder eine Gesamtschule besuchten. Seit der Erhebungswelle  $t_8$  (2009) wird nur noch nach der Befragungsteilnahme im Vorjahr (*Co009*) gefragt.

Zusätzlich zu den sechs Codefragen und den Zusatzfragen<sup>9</sup> standen für die Fragebogenzuordnungen das Geschlecht der Befragten zur Verfügung, sowie die (zuletzt) besuchte Schule, die mit einer Kennnummer erhoben wurde.

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).

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 Tabelle A.11 und A.12 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 Prozent (berechnet nach Pöge, 2007, S. 7). Bezüglich der hier betrachteten Zuordnungsschritte 2005/2006 sowie 2006/2007 liegt dieser Wert bei 61,0 % bzw. im letzten Schritt sogar bei 68,0 % (bei einer etwas niedrigeren Zahl insgesamt gefundener Zuordnungen). Dies bedeutet, dass die Reproduktion des Codes mit steigender

---

<sup>9</sup> Die Häufigkeitsverteilungen sämtlicher verwendeter Codes und Zusatzfragen finden sich in den Tabellen A.1 bis A.10 im Anhang.

---

Laufzeit der Befragung und zunehmender Erfahrung der Befragten offensichtlich besser funktioniert.

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 und die besuchte Schule übereinstimmen. Zudem wurden die Schülerinnen und Schüler anhand von Zusatzvariablen (*Co0007*, *Co0008*, *Co0009*) zugeordnet, durch die abgefragt wurde, ob sie im letzten Jahr an der Befragung teilgenommen haben, sowie nicht die Schule gewechselt haben oder sitzen geblieben sind.<sup>10</sup> 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.

Durch die Tatsache, dass die Befragten in 2006 (falls sie nicht durchgefallen waren) bereits der elften Jahrgangsstufe zuzurechnen waren – und damit bei weitem nicht mehr durchgehend eine allgemeinbildende Schule besuchten – wurde die Bedingung des Besuchs der gleichen Schule über zwei aufeinander folgende Jahre hinweg problematisch. Auch wurden im Jahr 2006 zusätzlich die Eingangsklassen der Duisburger Berufskollegs befragt. Unterschiede in den besuchten Schulen zwischen 2005 und 2006 waren damit

---

<sup>10</sup> Eine Ausnahme bildet hier der Kontrollschritt S10 des Abgleichs der Jahre 2005 und 2006. Da die Aufgabe der Restriktion der Teilnahme im letzten Jahr hier nur wenige zusätzliche Kontrollen erforderte, wurden die Kontrollen aus Effektivitätsgründen in diesem Fall von Beginn an weniger restriktiv durchgeführt.



sogar beabsichtigt; die Restriktion des Besuchs gleicher Schulen in beiden Erhebungsjahren hätte im Falle der Berufsschüler ein Auffinden von Bögen identischer Personen sogar behindert. So wurde im Jahr 2006 zum Zwecke des Abgleichs eine Hilfsvariable gebildet: Im Falle von Gymnasiasten und Gesamtschülern, die die Sekundarstufe II besuchten, wurde die Schulinformation unverändert übernommen;<sup>11</sup> hier war nicht von einer Änderung zwischen beiden Jahren auszugehen. Für alle Berufsschüler und diejenigen, die die allgemeinbildenden Schulen bereits verlassen hatten, wurde anstatt der aktuellen Schule die *im Vorjahr* besuchte Schule (abgefragt seit der Erhebung 2006) in die Hilfsvariable integriert. Im Jahr 2007 war diese Hilfskonstruktion nicht mehr sinnvoll: Da der größte Anteil der Befragten bereits vor *mehr* als einem Jahr die allgemeinbildenden Schulen verlassen hatte, wurde die für 2007 aktuelle Schulinformation nicht mehr durch die vor einem Jahr besuchte Schule ersetzt.

Insgesamt ist durch den teilweisen Wegfall der sehr zuverlässigen Schulinformation zur Identifikation identischer Personen die Anzahl an belastbaren Zuordnungskriterien zurückgegangen. Insbesondere beim Zulassen mehrerer Fehler in den Codes ergab sich dadurch eine deutliche Erhöhung der Anzahl erforderlicher Handschriftenkontrollen, da durch dieses fehlende Kriterium die Anzahl an Kombinationsmöglichkeiten möglicherweise passender Bogenpaare deutlich größer wurde.<sup>12</sup> Verschärfung erfuhr dieses Problem insbesondere im Befragungsjahr 2006 dadurch, dass im Zuge der einmaligen Erhebung an den Berufskollegs etwa 1 000 junge Erwachsene zusätzlich befragt wurden.<sup>13</sup> Im Vergleich zu den ersten vier Erhebungsjahren hat sich durch diese Umstände die Anzahl nötig gewordener Kontrollen deutlich erhöht. Deshalb wurde es erforderlich, die Zusatzvariablen nicht mehr gleichzeitig, sondern stufenweise freizusetzen.<sup>14</sup> Um den hierarchischen Aufbau des Verfahrens beizubehalten, wurden zunächst Restriktionen gewählt, die ein Zusammenpassen der Bogenpaare wahrscheinlicher machen (z. B. Teilnahme im letzten Jahr: ja), eine völlig restriktionsfreie Zuordnung wurde nur dann vorgenommen, falls der Kontrollaufwand und die zu erwartende Trefferquote den Arbeitsaufwand vertretbar erscheinen ließ. Eine Übersicht über die durchgeführten Zuordnungsschritte und die Bedingungen, die der Code erfüllen musste, findet sich in Tabellen 3.1 und 3.2.

---

11 Die Angaben sind besonders zuverlässig, da sie nicht vom Befragten selbst angegeben werden sollten, sondern im Zuge der Erhebung an den Schulen durch die Interviewer vermerkt wurden.

12 In der Befragung des Jahres 2009 wurde deshalb eine zusätzliche Codefrage eingeführt, die in Bezug auf dieses Problem mittelfristig Abhilfe schaffen soll. Jedoch kann erst im Zuordnungsschritt 2009/2011 von der neuen Information Gebrauch gemacht werden, da die betreffende Frage in beiden zu kontrollierenden Datensätzen vorhanden sein muss, um gewinnbringend eingesetzt werden zu können.

13 Zum Vergleich die Fallzahlen der einzelnen Querschnittsdatsätze 2002: 3 411, 2003: 3 392, 2004: 3 339, 2005: 3 405, 2006: 4 548, 2007: 3 336

14 Der Schritt S6 der Kontrollen für 2005/2006 musste aufgrund eines sehr hohen Kontrollaufwandes bei gleichzeitig extrem niedriger Trefferquote vorzeitig abgebrochen werden. Aus demselben Grund

**Tabelle 3.1:** Durchgeführte Kontrollschritte 2005/2006

	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 <sup>a</sup>
	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

<sup>a</sup> vorzeitiger Abbruch der Kontrollen

**Tabelle 3.2:** Durchgeführte Kontrollschritte 2006/2007

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

In Tabellen 3.3 und 3.4 sind jeweils die Anzahlen der in den einzelnen Schritten durchgeführten Handschriftenkontrollen sowie die jeweiligen Anteile passender und nicht passender Bogenpaare aufgeführt. Es wird ersichtlich, dass durch die gestiegene Gesamtzahl durchgeführter Kontrollen der Anteil *passender* Zuordnungen an allen kontrollierten Zuordnungsmöglichkeiten zurückgegangen ist. Während in den Zuordnungsschritten 2002/2003 bis 2004/2005 die Quoten dieser Ausschöpfungen noch zwischen 60,3 % und 68,3 % liegen (berechnet auf Basis von Pöge, 2007, S. 7), beträgt der Anteil passender Zuordnungen an allen durchgeführten Kontrollen für die Jahre 2005 und 2006 43,2 % und für die Jahre 2006 und 2007 48,2 %. Dadurch zeigt sich deutlich, dass durch die vergrößerte Stichprobe aus 2006 zum Auffinden ähnlicher absoluter Zuordnungszahlen

musste in diesem Zuordnungszyklus auf ein völliges Freisetzen der Zusatzvariablen bei zwei Fehlern im Code verzichtet werden.

über zwei aufeinander folgende Erhebungswellen hinweg ein erhöhter Kontrollaufwand betrieben werden musste.

Eine weitere Erhöhung des Kontrollaufwands resultierte aus der sukzessiven Umstellung auf eine postalische und persönliche Befragung seit dem Jahr 2006. Durch die Tatsache, dass aus befragungsorganisatorischen Gründen einige Befragte in der Übergangsphase zweier Erhebungsrounden möglicherweise doppelt zur Teilnahme aufgefordert wurden und dieser Aufforderung nachgekommen sind, konnte es passieren, dass ein Befragter zwei Mal innerhalb eines Jahres an der Befragung teilnahm. Deshalb wurden vor der eigentlichen Zuordnung über zwei Erhebungsjahre hinweg Handschriftenkontrollen nötig, die es zum Ziel hatten, innerhalb eines Befragungsjahres doppelt teilnehmende Schüler zu identifizieren. Das technische Vorgehen ähnelte hier stark dem oben beschriebenen, jedoch wurden nicht Fragebögen aus verschiedenen Jahren abgeglichen, sondern solche, die aus verschiedenen Erhebungsrounden desselben Jahres stammten. Insgesamt belief sich der Aufwand dieser Vorkontrollen im Jahr 2006 auf 704 und im Jahr 2007 auf 505 durchgeführte Abgleiche.

Ebenfalls nicht unerwähnt bleiben sollte, dass auch Kontrollen zwischen nicht direkt aufeinander folgenden Erhebungswellen durchgeführt wurden. Nicht im Panel zugeordnete Bögen eines Querschnittsdatensatzes wurden stets auch mit den nicht zugeordneten Bögen aus 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 3.3:** Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2005/2006

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

		Kontrollen gesamt	passt	passt nicht	fehlt <sup>b</sup>
ohne Fehler	S1	958	893	60	5
	S2	816	759	48	9
	gesamt	1 774	1 652	108	14
ein Fehler	S3	358	311	44	3
	S4	839	354	476	9
	gesamt	1 197	665	520	12
zwei Fehler	S5	119	91	26	2
	S6	609 <sup>a</sup>	42	522	3
	S7	751	132	606	13
	S8	114	15	94	5
	S9	1 586	91	1 482	13
	gesamt	3 179	371	2 756	36
drei Fehler	S10	115	19	90	6
Summe		6 265	2 707	3 474	

<sup>a</sup> vorzeitiger Abbruch der Kontrollen

<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
ohne Fehler	S1	100,0	93,2	6,3	0,5
	S2	100,0	93,0	5,9	1,1
	gesamt	100,0	93,1	6,1	0,8
ein Fehler	S3	100,0	86,9	12,3	0,8
	S4	100,0	42,2	56,7	1,1
	gesamt	100,0	55,6	43,4	1,0
zwei Fehler	S5	100,0	76,5	21,8	1,7
	S6	100,0	6,9	85,7	0,5
	S7	100,0	17,6	80,7	1,7
	S8	100,0	13,2	82,5	4,4
	S9	100,0	5,7	93,4	0,8
	gesamt	100,0	11,7	86,7	1,1
drei Fehler	S10	100,0	16,5	78,3	5,2
Summe		100,0	43,2	55,5	

**Tabelle 3.4:** Anzahl der handschriftlichen Kontrollen und deren Ergebnisse 2006/2007

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

		Kontrollen gesamt	passt	passt nicht	fehlt <sup>b</sup>
ohne Fehler	S1	833	769	46	18
	S2	1 039	934	81	24
	S3	38	17	6	15
	gesamt	1 910	1 720	133	57
ein Fehler	S4	249	227	15	7
	S5	652	366	251	35
	S6	140	13	103	24
	gesamt	1 041	606	369	66
zwei Fehler	S7	128	71	50	7
	S8	1 968	125	1 696	147
	gesamt	2 096	196	1 746	154
drei Fehler	S9	329	6	298	25
Summe		5 248	2 528	2 546	

<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
ohne Fehler	S1	100,0	92,3	5,5	2,2
	S2	100,0	89,9	7,8	2,3
	S3	100,0	44,7	15,8	39,5
	gesamt	100,0	90,1	7,0	3,0
ein Fehler	S4	100,0	91,2	6,0	2,8
	S5	100,0	56,1	38,5	5,4
	S6	100,0	9,3	73,6	17,1
	gesamt	100,0	58,2	35,4	6,3
zwei Fehler	S7	100,0	55,5	39,1	5,5
	S8	100,0	6,4	86,2	7,5
	gesamt	100,0	9,4	83,3	7,3
drei Fehler	S9	100,0	1,8	90,6	7,6
Summe		100,0	48,2	48,5	

## 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- und Fünf-Wellen-Panels sicherlich am bedeutsamsten. Das vorläufige Resultat der jährlichen Erhebungen in Duisburg ist ein Sechs-Wellen-Paneldatensatz der Jahre 2002 bis 2007 ( $t_{1,2,3,4,5,6}$ ) mit 1 307 durchgehend<sup>15</sup> und 4 195 insgesamt (mit Lücken) enthaltenen Fällen.

### 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).

Ab dem Zeitpunkt  $t_4$  werden daher ausschließlich die Querschnittsdaten zur Beurteilung der Repräsentativität der einzelnen Paneldatensätze herangezogen. Zwar ist diese

---

<sup>15</sup> Berücksichtigt sind hier nur die gültigen Fälle, die auch den Plausibilitätskontrollen der einzelnen Querschnittsdatensätze standhielten.

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.  $Q_{t_5}$  beinhaltet zum Beispiel alle 4 548 gültigen Fälle des Querschnittsdatensatzes aus dem Jahr 2006.

Ein » $P_{Q_{t_x}}$ « mit nur einem Zeitpunkt als Index bezeichnet einen so genannten Panelquerschnitt. Diese Panelquerschnitte beinhalten *alle* vorhandenen Fälle des gesamten Sechs-Wellen-Paneldatensatzes (mit Lücken) aus *einem* Erhebungszeitpunkt. So enthält zum Beispiel  $P_{Q_{t_5}}$  alle 3 149 gültigen Fälle die aus  $t_5$  im Gesamtpaneldatensatz enthalten sind. Dies sind im Unterschied zum Querschnitt  $Q_{t_5}$  alle Fälle, die zum Zeitpunkt  $t_5$  befragt wurden und im Zuge der Erstellung des Sechs-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  $P_{t_5, 6}$  für den Zwei-Wellen-Paneldatensatz mit allen 2 514 *gültigen* Fällen, die zwischen  $t_5$  und  $t_6$  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 im Zuge der Erstellung des Sechs-Wellen-Panels sozusagen als Nebenprodukte entstanden sind. Jedoch werden manche dieser Datensätze im vorliegenden Bericht nicht im Detail dargestellt, da sie bereits bei der Dokumentation des Vier-Wellen-Panels in dieser Form existierten und seitdem nicht mehr verändert wurden. Diese Datensätze sind bereits ausführlich von Pöge (2007) beschrieben worden. Gesondert dargestellt werden hier all diejenigen Datensätze, die seit dem Bericht zum Vier-Wellen-Panel durch weitere Kontrollen und die Hinzunahme weiterer Fälle im Zuge der Konstruktion des Sechs-Wellen-Panels verändert wurden. Diese sind in Tabelle 4.1 **fett gedruckt**.

#### 4.1 Die verwendeten Datensätze und ihre Kurzbezeichnungen

**Tabelle 4.1:** Bezeichnungen und Fallzahlen aller Paneldatensätze 2002 bis 2007

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
$P_{Q_{t_1}}$	2 750	Panelquerschnittsdaten des Jahres 2002
$P_{Q_{t_2}}$	3 132	Panelquerschnittsdaten des Jahres 2003
$P_{Q_{t_3}}$	<b>3 177</b>	<b>Panelquerschnittsdaten des Jahres 2004</b>
$P_{Q_{t_4}}$	<b>3 206</b>	<b>Panelquerschnittsdaten des Jahres 2005</b>
$P_{Q_{t_5}}$	<b>3 149</b>	<b>Panelquerschnittsdaten des Jahres 2006</b>
$P_{Q_{t_6}}$	<b>2 702</b>	<b>Panelquerschnittsdaten des Jahres 2007</b>
$P_{t_{1,2}}$	2 472	Paneldaten der Jahre 2002 und 2003
$P_{t_{2,3}}$	2 596	Paneldaten der Jahre 2003 und 2004
$P_{t_{3,4}}$	2 750	Paneldaten der Jahre 2004 und 2005
$P_{t_{4,5}}$	<b>2 696</b>	<b>Paneldaten der Jahre 2005 und 2006</b>
$P_{t_{5,6}}$	<b>2 514</b>	<b>Paneldaten der Jahre 2006 und 2007</b>
$P_{t_{1,3}}$	2 202	Paneldaten der Jahre 2002 und 2004
$P_{t_{2,4}}$	2 493	Paneldaten der Jahre 2003 und 2005
$P_{t_{3,5}}$	<b>2 552</b>	<b>Paneldaten der Jahre 2004 und 2006</b>
$P_{t_{4,6}}$	<b>2 367</b>	<b>Paneldaten der Jahre 2005 und 2007</b>
$P_{t_{1,4}}$	2 165	Paneldaten der Jahre 2002 und 2005
$P_{t_{1,2,3}}$	2 012	Paneldaten der Jahre 2002, 2003, 2004
$P_{t_{2,3,4}}$	2 260	Paneldaten der Jahre 2003, 2004, 2005
$P_{t_{3,4,5}}$	<b>2 354</b>	<b>Paneldaten der Jahre 2004, 2005, 2006</b>
$P_{t_{4,5,6}}$	<b>2 181</b>	<b>Paneldaten der Jahre 2005, 2006, 2007</b>
$P_{t_{1,2,4}}$	1 930	Paneldaten der Jahre 2002, 2003, 2005
$P_{t_{1,3,4}}$	1 922	Paneldaten der Jahre 2002, 2004, 2005
$P_{t_{1,3,5}}$	<b>1 793</b>	<b>Paneldaten der Jahre 2002, 2004, 2006</b>
$P_{t_{2,3,5}}$	<b>2 082</b>	<b>Paneldaten der Jahre 2003, 2004, 2006</b>
$P_{t_{2,4,5}}$	<b>2 112</b>	<b>Paneldaten der Jahre 2003, 2005, 2006</b>
$P_{t_{2,4,6}}$	<b>1 873</b>	<b>Paneldaten der Jahre 2003, 2006, 2007</b>
$P_{t_{3,4,6}}$	<b>2 068</b>	<b>Paneldaten der Jahre 2004, 2005, 2007</b>
$P_{t_{3,5,6}}$	<b>2 007</b>	<b>Paneldaten der Jahre 2004, 2006, 2007</b>
$P_{t_{1,4,5}}$	<b>1 861</b>	<b>Paneldaten der Jahre 2002, 2005, 2006</b>
$P_{t_{1,4,6}}$	<b>1 666</b>	<b>Paneldaten der Jahre 2002, 2005, 2007</b>



**Tabelle 4.1:** Bezeichnungen und Fallzahlen aller Paneldatensätze 2002 bis 2007

$P_{1,2,3,4}$	1 769	Paneldaten der Jahre 2002, 2003, 2004, 2005
$P_{1,2,3,4,5}$	1 950	Paneldaten der Jahre 2003, 2004, 2005, 2006
$P_{1,2,3,4,5,6}$	1 926	Paneldaten der Jahre 2004, 2005, 2006, 2007
$P_{1,2,3,5}$	1 646	Paneldaten der Jahre 2002, 2003, 2004, 2006
$P_{1,2,4,5}$	1 667	Paneldaten der Jahre 2002, 2003, 2005, 2006
$P_{1,2,4,6}$	1 500	Paneldaten der Jahre 2002, 2003, 2005, 2007
$P_{1,3,4,5}$	1 685	Paneldaten der Jahre 2002, 2004, 2005, 2006
$P_{1,3,4,6}$	1 508	Paneldaten der Jahre 2002, 2004, 2005, 2007
$P_{1,3,5,6}$	1 455	Paneldaten der Jahre 2002, 2004, 2006, 2007
$P_{1,2,3,4,6}$	1 733	Paneldaten der Jahre 2003, 2004, 2005, 2007
$P_{1,2,3,5,6}$	1 675	Paneldaten der Jahre 2003, 2004, 2006, 2007
$P_{1,2,4,5,6}$	1 739	Paneldaten der Jahre 2003, 2005, 2006, 2007
$P_{1,4,5,6}$	1 550	Paneldaten der Jahre 2002, 2005, 2006, 2007
$P_{1,2,3,4,5}$	1 552	Paneldaten der Jahre 2002, 2003, 2004, 2005, 2006
$P_{1,2,3,4,5,6}$	1 620	Paneldaten der Jahre 2003, 2004, 2005, 2006, 2007
$P_{1,2,3,4,6}$	1 396	Paneldaten der Jahre 2002, 2003, 2004, 2005, 2007
$P_{1,2,3,5,6}$	1 346	Paneldaten der Jahre 2002, 2003, 2004, 2006, 2007
$P_{1,2,4,5,6}$	1 398	Paneldaten der Jahre 2002, 2003, 2005, 2006, 2007
$P_{1,3,4,5,6}$	1 411	Paneldaten der Jahre 2002, 2004, 2005, 2006, 2007
$P_{1,2,3,4,5,6}$	1 307	Paneldaten der Jahre 2002, 2003, 2004, 2005, 2006, 2007

<sup>a</sup> nur teilnehmende Schulen

## 4.2 Abschätzung und Bewertung der Zuordnungsgüte

Die Bewertung der Zuordnungsgüte nach dem Verfahren, das im Zusammenhang mit dem Vier-Wellen-Panel verwendet wurde ist für das Sechs-Wellen-Panel nicht mehr realisierbar. Dieses Verfahren basierte in erster Linie auf in den Schulen erreichbaren Schülerzahlen bzw. der amtlichen Schulstatistik (vgl. beispielsweise Pöge, 2007, S. 14 f.). Beides ist seit dem Erhebungsjahr 2006 nicht mehr für alle Befragten zu ermitteln.

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.

Bei Betrachtung der Tabelle 4.2 wird ersichtlich, dass mit längerer Laufzeit der Panelerhebung die Größe des Datensatzes zunehmend stabiler wird, d. h. immer mehr Befragte aus dem Vorjahres-Datensatz können einerseits wieder erreicht, andererseits (und vor allem) auch zugeordnet werden. Dies mag zum einen am steigenden Alter der Befragten und der möglicherweise damit verbundenen höheren Fähigkeit, die Codefragen richtig zu beantworten bzw. sich zu konzentrieren liegen, zum anderen an einem Übungs- oder Gewöhnungseffekt durch die langjährige Teilnahme an der Befragung.

**Tabelle 4.2:** Ausschöpfungen der durchgehenden Paneldatensätze

Datensatz	$P_{t_1}$	$P_{t_{1,2}}$	$P_{t_{1,2,3}}$	$P_{t_{1,2,3,4}}$	$P_{t_{1,2,3,4,5}}$	$P_{t_{1,2,3,4,5,6}}$
Fallzahl	3411	2479	2012	1769	1552	1307
Anteil der Fallzahl des durchgehenden Datensatzes aus dem Vorjahr in %		73	81	88	88	84

Lediglich bei der Stabilität des bislang umfangreichsten Datensatzes, des Sechswellen-Panels, ist im Vergleich zum Fünf-Wellen-Panel ein leichter Einbruch zu beobachten. Dieser ist sehr wahrscheinlich der Umstellung des Befragungsmodus von einer Paper-und-Pencil-Befragung im Klassenverband auf eine postalische Befragung geschuldet. Betrachtet man die tiefgreifenden Veränderungen in der Vorgehensweise der Erhebung, erscheint dieser Einbruch tolerabel.



## 5 Zusammensetzung der Paneldaten

Im folgenden Kapitel wird die Zusammensetzung der einzelnen Paneldatensätze nach Geschlecht, Alter und Schulform beschrieben. Zum Vergleich werden dabei in jedem Erhebungsjahr die Daten der Querschnitte aus den entsprechenden Jahren hinzugezogen. Daten der Schulstatistik können – wie bereits weiter oben im Text dargestellt – nur für die drei ersten Erhebungswellen zu einem Vergleich herangezogen werden.

Bei den vergleichenden Darstellungen werden zunächst die zentralen Variablen Geschlecht und Alter berücksichtigt; die Verteilung der Jugendlichen auf verschiedene Schulformen wird ebenfalls in die Darstellungen einbezogen, obwohl die Bedeutung der besuchten Schulform mit steigendem Alter der Befragten immer mehr abnimmt und zur Einschätzung der Repräsentativität der Paneldaten immer unzuverlässiger wird. Dennoch soll die diesbezügliche Diskrepanz zwischen Schulstatistiken, Querschnitts- und Paneldaten in den einzelnen Jahren hier dokumentiert werden.

Bei der Darstellung der Zusammensetzung der Teilpanels die mehrere Wellen umfassen, können beim Geschlecht und beim Alter leichte Inkonsistenzen über die Zeit auftreten. Der Grund hierfür ist, dass bei *jeder* Frage fehlerhafte bzw. inkonsistente Antworten auftreten. So natürlich auch bei so scheinbar eindeutig zu beantwortenden Fragen wie denen nach Alter und 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>16</sup>

#### 5.1.1 Panelquerschnitt 2004 ( $P_{Q_{t_3}}$ )

Der Panelquerschnitt des Jahres 2004 ist der erste, der hier in die Betrachtung einfließt. Zwar wurde er bereits durch Pöge (2007, S. 23ff.) im Zuge der Dokumentation des Duisburger Vier-Wellen-Panels dargestellt, jedoch hat sich der Umfang dieses Pa-

---

<sup>16</sup> Wenn also genau die Population der überhaupt zuordenbaren Fälle von Interesse ist.

nelquerschnitts seitdem vergrößert. Die zusätzlichen Fälle resultieren aus dem neu durchgeführten Abgleich der Daten aus 2004 ( $t_3$ ) und 2006 ( $t_5$ ). Offensichtlich sind im Zuge dieses Abgleichs Fälle hinzugekommen, die in 2004 erstmals (zuordenbar), danach jedoch mindestens noch ein weiteres Mal (zuordenbar) in 2006 – nicht jedoch in 2005 – an der Befragung teilgenommen haben. Der Panelquerschnitt  $P_{Q_{t_3}}$  wurde in diesem Zuge von 3 125 Fällen auf 3 177 Fälle erweitert.<sup>17</sup>

**Tabelle 5.1:** Zusammensetzung nach Geschlecht  $P_{Q_{t_3}}$

	männlich		weiblich		gesamt (100 %)
	Hfk	%	Hfk	%	
$S_{t_3}$	1999	50,9	1926	49,1	3925
$Q_{t_3}$	1635	49,0	1704	51,0	3339
$P_{Q_{t_3}}$	1528	48,1	1649	51,9	3177

Aus Tabelle 5.1 geht hervor, dass im Panelquerschnitt  $P_{Q_{t_3}}$  hinsichtlich des Geschlechts der Befragten eine leichte Verschiebung vorliegt: Vergleicht man die Geschlechterverteilung im Panelquerschnitt mit der Schulstatistik des entsprechenden Erhebungsjahres, so wird deutlich, dass der Anteil der Schülerinnen im Panelquerschnitt knapp drei Prozentpunkte höher liegt. Im Vergleich zum entsprechenden Querschnittsdatensatz  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Panelquerschnitt um knapp einen Prozentpunkt erhöht. Dabei weisen die Querschnittsdaten des Jahres 2004 im Vergleich zur Schulstatistik ebenfalls eine geringe Verschiebung auf. Der erhöhte Anteil weiblicher Befragter verstärkt sich demnach schrittweise über den Querschnittsdatensatz hin zum Panelquerschnitt.

**Tabelle 5.2:** Altersdurchschnitt  $P_{Q_{t_3}}$

	Alter	Std.-abw.	gesamt
$Q_{t_3}$	15,1	0,8	3270
$P_{Q_{t_3}}$	15,1	0,8	3127

Das durchschnittliche Alter der Befragten in den Panelquerschnitten kann jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdaten  $Q_{t_3}$  und dem Panelquerschnitt zeigt, dass sich das Durchschnittsalter der Befragten in beiden Datensätzen nicht unterscheidet (siehe Tabelle 5.2).

<sup>17</sup> Damit beträgt der Anteil der Befragten aus dem Querschnittsdatensatz, der sich auch im Panelquerschnitt befindet, zu  $t_3$  95,1 %.

## 5.1 Die Panelquerschnitte

**Tabelle 5.3:** Zusammensetzung  $P_{Q_{t_3}}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
männlich	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999
	$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635
	$P_{Q_{t_3}}$	26,0	22,3	32,0	19,1	—	0,5	—	—	1528
weiblich	$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926
	$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704
	$P_{Q_{t_3}}$	18,7	23,8	31,5	25,8	—	0,2	—	—	1649
gesamt	$S_{t_3}$	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
	$P_{Q_{t_3}}$	22,2	23,1	31,7	22,6	—	0,3	—	—	3177

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

Tabelle 5.3 zeigt, dass der Panelquerschnitt  $P_{Q_{t_3}}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Panelquerschnitt  $P_{Q_{t_3}}$  über zweieinhalb Prozentpunkte unter denen der Schulstatistik. Weiterhin sind im Panelquerschnitt etwa ein halber Prozentpunkt weniger Gesamtschülerinnen und -schüler und ein Prozentpunkt weniger Sonderschülerinnen und -schüler enthalten, als in der Schulstatistik verzeichnet sind.

Im Gegensatz dazu liegt der Anteil der Gymnasiastinnen und Gymnasiasten im Panelquerschnitt um etwa zweieinhalb, der Anteil der Realschülerinnen und -schüler etwa eineinhalb Prozentpunkte höher als in der Schulstatistik. Jedoch sind diese Abweichungen in erster Linie auf Diskrepanzen zwischen der Schulstatistik und dem Querschnittsdatensatz  $Q_{t_3}$  zurückzuführen; die Anteilsunterschiede hinsichtlich der besuchten Schulform zwischen Querschnittsdatensatz und Panelquerschnitt sind recht gering.

**Tabelle 5.4:** Zusammensetzung  $P_{Q_{t_3}}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform <sup>a</sup> (%)								gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(%)
männlich	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9
	$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0
	$P_{Q_{t_3}}$	56,4	46,5	48,5	40,7	—	72,7	—	—	48,1
weiblich	$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1
	$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0
	$P_{Q_{t_3}}$	43,6	53,5	51,5	59,3	—	27,3	—	—	51,9

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

Tabelle 5.4 verdeutlicht weiterhin, dass über alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Panelquerschnitt  $P_{Q_{t_3}}$  enthalten sind als in der Schulstatistik. Dieser Trend wird bereits beim Vergleich zwischen der Schulstatistik und dem Querschnittsdatensatz deutlich, ist jedoch in Bezug auf den Panelquerschnitt in noch verstärkter Form zu beobachten.

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

Auch der Panelquerschnitt des Jahres 2005 wurde bereits durch Pöge (2007, S. 25 ff.) im Zuge der Dokumentation des Duisburger Vier-Wellen-Panels beschrieben, jedoch hat sich der Umfang dieses Panelquerschnitts seitdem ebenfalls vergrößert. Hinzugekommene Fälle resultieren aus dem neu durchgeführten Abgleich mit den Querschnittsdatensätzen der Jahre 2006 und 2007. Offensichtlich sind in diesen Jahren noch Fälle hinzugekommen, die in 2005 erstmals (zuordenbar), danach jedoch mindestens noch ein weiteres Mal (zuordenbar) an der Befragung teilgenommen haben. Der Panelquerschnitt  $P_{Q_{t_4}}$  wurde in diesem Zuge von 3 074 Fällen auf 3 206 Fälle erweitert.<sup>18</sup>

**Tabelle 5.5:** 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
$P_{Q_{t_4}}$	1 565	48,8	1 641	51,2	3 206

Aus Tabelle 5.5 geht hervor, dass im Panelquerschnitt  $P_{Q_{t_4}}$  hinsichtlich des Geschlechts der Befragten eine leichte Verschiebung vorliegt: Vergleicht man die Geschlechterverteilung im Panelquerschnitt mit dem Querschnittsdatensatz des entsprechenden Erhebungsjahres (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, vgl. hierzu die Ausführungen in Abschnitt 4.1), so wird deutlich, dass der Anteil der Schülerinnen im Panelquerschnitt um über eineinhalb Prozentpunkte höher liegt.

**Tabelle 5.6:** 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	3 149

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

<sup>18</sup> Damit beträgt der Anteil der Befragten aus dem Querschnittsdatensatz, der sich auch im Panelquerschnitt befindet, zu  $t_4$  94,1 %.

den: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdaten  $Q_{t_4}$  und dem Panelquerschnitt zeigt, dass sich das Durchschnittsalter der Befragten in beiden Datensätzen kaum unterscheidet (siehe Tabelle 5.6).

**Tabelle 5.7:** Zusammensetzung  $P_{Q_{t_4}}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
männlich	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704
	$P_{Q_{t_4}}$	26,4	20,8	31,8	18,8	1,0	0,6	0,7	0,0	1555
weiblich	$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682
	$P_{Q_{t_4}}$	17,5	22,6	32,7	25,9	0,7	0,1	0,3	0,2	1636
gesamt	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
	$P_{Q_{t_4}}$	21,8	21,7	32,2	22,5	0,8	0,3	0,5	0,1	3191

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

Tabelle 5.7 zeigt, dass der Panelquerschnitt  $P_{Q_{t_4}}$  insgesamt unwesentlich nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Panelquerschnitt  $P_{Q_{t_4}}$  etwa gleichauf mit denen des Querschnittsdatensatzes. Weiterhin sind im Panelquerschnitt etwa ein halber Prozentpunkt weniger Gesamtschülerinnen und -schüler enthalten als im Querschnittsdatensatz.

Im Gegensatz dazu liegt der Anteil der Gymnasiastinnen und Gymnasiasten im Panelquerschnitt um einen halben Prozentpunkt höher als in den Querschnittsdaten. Die Anteile der Realschülerinnen und Realschüler, sowie derjenigen Befragten, die Berufskollegs, Sonderschulen, anderweitige oder gar keine Schulen besuchen, unterscheiden sich ebenfalls – bei recht geringen Fallzahlen – höchstens unwesentlich, vergleicht man Querschnittsdatensatz und Panelquerschnitt des Erhebungszeitpunktes  $t_4$ .

**Tabelle 5.8:** Zusammensetzung  $P_{Q_{t_4}}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform <sup>a</sup> (%)								gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(%)
männlich	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3
	$P_{Q_{t_4}}$	58,9	46,7	48,0	40,9	55,6	90,0	68,8	0,0	48,7
weiblich	$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7
	$P_{Q_{t_4}}$	41,1	53,3	52,0	59,1	44,4	10,0	31,2	100,0	51,3

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



Tabelle 5.8 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Panelquerschnitt  $P_{Q_{t_4}}$  enthalten sind als in den Querschnittsdaten.

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

Der Panelquerschnitt des Jahres 2006 ist der erste hier betrachtete Datensatz, der noch nicht in einem vorangehenden Bericht beschrieben wurde. Der Panelquerschnitt enthält insgesamt 3 149 Fälle.<sup>19</sup> Dies sind alle Fälle aus der fünften Erhebungswelle, die mindestens einem weiteren Fall im Paneldatensatz zugeordnet werden konnten.

Die deutliche Diskrepanz der Fallzahlen von Querschnittsdatsatz und Panelquerschnitt in diesem Erhebungsjahr, sowie die teilweise deutlichen Anteilsunterschiede, die bei der Betrachtung der besuchten Schulform evident werden, sind einem deutlich veränderten Vorgehen bei der Datenerhebung in diesem Jahr geschuldet.

Da sich die Befragten zu  $t_5$  (überwiegend) in der elften Jahrgangsstufe befanden, war davon auszugehen, dass zumindest die Haupt-, Sonder und Realschülerinnen und -schüler ihre Schullaufbahn zu diesem Zeitpunkt bereits beendet hatten und deshalb für eine – wie bisher durchgeführte – Befragung im Klassenverband nicht mehr zur Verfügung standen. Um damit sicherlich verbundenen Ausfällen im Panel entgegenzuwirken, entschied man sich in der fünften Erhebungswelle zusätzlich für eine Befragung der Einstiegsklassen an sämtlichen Duisburger Berufskollegs. Eine offizielle Schulstatistik, die nach dem Geschlecht der Berufskollegiaten differenziert, existiert hierfür nicht (eine genaue Erläuterung des Vorgehens bei der Erhebung der Querschnittsdaten im Jahr 2006 findet sich bei Bentrup, 2007).

Diese Vorgehensweise hat im Vergleich zu den Querschnittsdatsätzen der übrigen Erhebungsjahre zu einer Vergrößerung der Stichprobe um rund 1 000 Befragte geführt. Durch den bewusst erhobenen Überschuss ist die Stichprobe zu  $t_5$  jedoch in ihrer Zusammensetzung nach Schulform nur noch bedingt mit den Stichproben der übrigen Erhebungsjahre vergleichbar: Der Anteil der Berufsschülerinnen und -schüler ist 2006 deutlich erhöht; nicht alle sind jedoch für den Paneldatensatz relevant. Die resultierende Verschiebung der Anteilswerte der besuchten Schulformen in der Stichprobe zu  $t_5$  führt dazu, dass streng genommen keine Aussage über die Qualität der Abbildung der Querschnittsdaten in den Paneldaten getroffen werden kann. Dennoch wird die erhobene Stichprobe des Jahres 2006 hier zum Vergleich herangezogen – einerseits mangels Alternativen, andererseits zur Beurteilung der panelbezogenen Ausschöpfung der Befragung an den Berufskollegs, für welche ein deutlich erhöhter Erhebungs- und Kontrollaufwand in Kauf genommen wurde.

---

<sup>19</sup> Damit beträgt der Anteil der Befragten aus dem Querschnittsdatsatz, der sich auch im Panelquerschnitt befindet, zu  $t_5$  69,2 %.

**Tabelle 5.9:** Zusammensetzung nach Geschlecht  $P_{Q_{t_5}}$

	männlich		weiblich		gesamt (100 %)
	Hfk	%	Hfk	%	
$Q_{t_5}$	2265	49,8	2282	50,2	4547
$P_{Q_{t_5}}$	1462	46,4	1687	53,6	3149

**Tabelle 5.10:** 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	3111

Aus Tabelle 5.9 geht hervor, dass im Panelquerschnitt  $P_{Q_{t_5}}$  hinsichtlich des Geschlechts der Befragten eine leichte Verschiebung vorliegt: Vergleicht man die Geschlechterverteilung im Panelquerschnitt mit dem Querschnittsdatensatz des entsprechenden Erhebungsjahres (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, vgl. hierzu die Ausführungen in Abschnitt 4.1), so wird deutlich, dass der Anteil der Schülerinnen im Panelquerschnitt um knapp dreieinhalb Prozentpunkte höher liegt.

Ein Vergleich des durchschnittlichen Alters der Befragten in den Querschnittsdaten  $Q_{t_5}$  und dem Panelquerschnitt zeigt, dass das Durchschnittsalter im Panelquerschnitt niedriger und die Altersstruktur homogener ist (siehe Tabelle 5.10). Diese recht deutlichen Altersunterschiede lassen sich auf die oben beschriebene, erweiterte Stichprobe zurückführen.

**Tabelle 5.11:** Zusammensetzung  $P_{Q_{t_5}}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)							gesamt (100 %)	
		HS	RS	GS	GY	BK	SO	andere		keine
männlich	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181
	$P_{Q_{t_5}}$	0,8	0,5	19,3	22,9	50,6	0,3	2,0	3,6	1392
weiblich	$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226
	$P_{Q_{t_5}}$	0,2	0,4	19,9	30,0	45,3	0,1	2,3	1,9	1634
gesamt	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407
	$P_{Q_{t_5}}$	0,5	0,5	19,6	26,7	47,8	0,2	2,1	2,7	3026

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

Tabelle 5.11 zeigt, dass der Panelquerschnitt  $P_{Q_{t_5}}$  insgesamt nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Panelquerschnitt  $P_{Q_{t_5}}$  etwa gleichauf mit denen des Querschnittsdatensatzes, allerdings sind sie so ge-

ring, dass keine gesicherten Aussagen getroffen werden können. Weiterhin sind im Panelquerschnitt knapp vier Prozentpunkte mehr Gesamtschülerinnen und -schüler enthalten als im Querschnittsdatensatz. Genauso liegt der Anteil der Gymnasiastinnen und Gymnasiasten im Panelquerschnitt etwas über sieben Prozentpunkte höher als in den Querschnittsdaten. Diese große Diskrepanz resultiert wohl nicht zuletzt aus der bereits erläuterten, anteilig deutlich verschiedenen Zusammensetzung von Querschnittsdatensatz und Panelquerschnitt. Die Anteile der Realschülerinnen und Realschüler, sowie derjenigen Befragten, die Sonderschulen besuchen, unterscheiden sich ebenfalls – bei recht geringen Fallzahlen – höchstens unwesentlich, vergleicht man Querschnittsdatensatz und Panelquerschnitt des Erhebungszeitpunktes  $t_5$ . Befragte, die anderweitige oder gar keine Schulen besuchen, sind mit etwas über einem halben Prozentpunkt im Panelquerschnitt überrepräsentiert.

Besonders auffällig ist zu  $t_5$  die Diskrepanz zwischen dem Anteil der Berufsschülerinnen und -schüler im Querschnittsdatensatz und im Panelquerschnitt: Mit knapp dreizehn Prozentpunkten sind die Berufskollegiaten im Panelquerschnitt aus den oben dargestellten Gründen unterrepräsentiert.

**Tabelle 5.12:** Zusammensetzung  $P_{Q_{t_5}}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform <sup>a</sup> (%)							gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(%)
männlich	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5
	$P_{Q_{t_5}}$	78,6	50,0	45,2	39,4	48,8	80,0	43,1	61,7	46,0
weiblich	$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5
	$P_{Q_{t_5}}$	21,4	50,0	54,8	60,6	51,2	20,0	56,9	38,3	54,0

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

Tabelle 5.12 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Panelquerschnitt  $P_{Q_{t_5}}$  enthalten sind als in den Querschnittsdaten (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

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

Im Jahr 2007 sind im Panelquerschnitt insgesamt 2 702 Fälle enthalten.<sup>20</sup> Dies sind alle Fälle aus diesem Jahr, die bislang mindestens einem weiteren Fall im Paneldatensatz zugeordnet werden konnten. Vergleicht man die Fallzahl dieses Panelquerschnitts mit

<sup>20</sup> Damit beträgt der Anteil der Befragten aus dem Querschnittsdatensatz, der sich auch im Panelquerschnitt befindet, zu  $t_6$  81,0 %.

denen der vorangehend beschriebenen Panelquerschnitte, erscheint sie zunächst recht niedrig. Hierbei muss jedoch bedacht werden, dass es sich bei der Erhebungswelle von 2007 um die letzte in den Sechs-Wellen-Paneldatensatz einbezogene handelt. Das bedeutet, dass – im Gegensatz zu den anderen Panelquerschnitten – bislang lediglich eine Zuordnung in Bezug auf vorangehende Befragungen stattgefunden hat. Nachdem im Zuge der Erstellung der Sieben- und Acht-Wellen-Paneldatensätze auch die Abgleiche mit den später erhobenen Wellen ( $t_7$  und  $t_8$ ) erfolgt sind (und damit auch *nachfolgende* Datensätze zugeordnet wurden), wird sich die Fallzahl dieses Panelquerschnitts voraussichtlich noch erhöhen.

**Tabelle 5.13:** Zusammensetzung nach Geschlecht  $P_{Q_{t_6}}$

	männlich		weiblich		gesamt (100 %)
	Hfk	%	Hfk	%	
$Q_{t_6}$	1 567	47,0	1 769	53,0	3 335
$P_{Q_{t_6}}$	1 191	44,1	1 511	55,9	2 702

Aus Tabelle 5.13 geht hervor, dass im Panelquerschnitt  $P_{Q_{t_6}}$  hinsichtlich des Geschlechts der Befragten eine leichte Verschiebung vorliegt: Vergleicht man die Geschlechterverteilung im Panelquerschnitt mit dem Querschnittsdatensatz des entsprechenden Erhebungsjahres (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, vgl. hierzu die Ausführungen in Abschnitt 4.1), so wird deutlich, dass der Anteil der Schülerinnen im Panelquerschnitt um knapp drei Prozentpunkte höher liegt.

**Tabelle 5.14:** Altersdurchschnitt  $P_{Q_{t_6}}$

	Alter	Std.-abw.	gesamt
$Q_{t_6}$	18,1	1,1	3 320
$P_{Q_{t_6}}$	18,1	0,8	2 690

Ein Vergleich des durchschnittlichen Alters der Befragten in den Querschnittsdaten  $Q_{t_6}$  und dem Panelquerschnitt zeigt, dass das Durchschnittsalter gleich hoch, im Panelquerschnitt die Altersstruktur aber homogener ist (siehe Tabelle 5.14).

Tabelle 5.15 zeigt, dass der Panelquerschnitt  $P_{Q_{t_6}}$  insgesamt nur leicht nach besuchter Schulform verzerrt ist: Hauptschülerinnen und -schüler, sowie Realschülerinnen und -schüler sind im Querschnittsdatensatz und im Panelquerschnitt  $P_{Q_{t_6}}$  nicht mehr enthalten. Weiterhin sind im Panelquerschnitt anteilig etwa genauso viele Gesamtschülerinnen und -schüler enthalten wie im Querschnittsdatensatz. Dagegen liegt der Anteil der Gymnasiastinnen und Gymnasiasten im Panelquerschnitt um etwas über einen Prozentpunkt höher als in den Querschnittsdaten. Befragte, die anderweitige Schulen besuchen, sind

**Tabelle 5.15:** Zusammensetzung  $P_{Q_{t_6}}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
männlich	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1 504
	$P_{Q_{t_6}}$	—	—	24,0	17,6	33,1	—	11,5	13,8	1 151
weiblich	$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1 715
	$P_{Q_{t_6}}$	—	—	17,8	30,2	30,2	—	10,2	11,5	1 474
gesamt	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3 219
	$P_{Q_{t_6}}$	—	—	17,8	27,5	31,5	—	10,8	12,5	2 625

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

im Querschnittsdatensatz und im Panelquerschnitt etwa zu gleichen Anteilen enthalten; Befragte, die keine Schule besuchen, sind mit etwas über eineinhalb Prozentpunkten im Panelquerschnitt unterrepräsentiert. Der Anteil der Berufsschülerinnen und -schüler im Querschnittsdatensatz entspricht ebenfalls in etwa dem im Panelquerschnitt: Mit einem halben Prozentpunkt sind die Berufskollegiaten im Panelquerschnitt nur sehr schwach überrepräsentiert.

**Tabelle 5.16:** Zusammensetzung  $P_{Q_{t_6}}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform <sup>a</sup> (%)								gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(%)
männlich	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7
	$P_{Q_{t_6}}$	—	—	43,6	38,3	46,1	—	46,6	48,3	43,8
weiblich	$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3
	$P_{Q_{t_6}}$	—	—	56,4	61,7	53,9	—	53,4	51,7	56,2

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

Tabelle 5.16 verdeutlicht weiterhin, dass über alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Panelquerschnitt  $P_{Q_{t_6}}$  enthalten sind als in den Querschnittsdaten (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

## 5.2 Die Zwei-Wellen-Panel

Die Datensätze der Zwei-Wellen-Panel enthalten alle Fälle, deren zugehörige (und nach Plausibilitätskontrollen gültige) Fragebögen einem (ebenfalls gültigen) Bogen mindestens eines weiteren Zeitpunktes zugeordnet werden konnten. In diesem Bericht werden die Zwei-Wellen-Paneldatensätze aller Kombinationen der sechs betrachteten Zeitpunkte dargestellt, die nicht in der Methodendokumentation zum Vier-Wellen-Panel  $t_1$  bis  $t_4$  (vgl. Pöge, 2007) bereits abgearbeitet wurden.

In der Praxis werden vermutlich die Datensätze, die zwei direkt aufeinander folgende Zeitpunkte umfassen, am bedeutsamsten sein. Diese werden zuerst dargestellt, im Anschluss werden die restlichen in der Tabelle 4.1 fett gedruckten Zwei-Wellen-Paneldatensätze einzeln beschrieben.

### 5.2.1 Zwei-Wellen-Panel 2005/2006 ( $P_{t_4,5}$ )

Das Zwei-Wellen-Panel der Jahre 2005 und 2006 enthält 2 696 Fälle. Aus Tabelle 5.17 geht hervor, dass im Zwei-Wellen-Paneldatensatz  $P_{t_4,5}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.17:** Zusammensetzung nach Geschlecht  $P_{t_4,5}$

		männlich		weiblich		gesamt
		<i>N</i>	%	<i>N</i>	%	(100 %)
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_4,5}$	1236	45,8	1460	54,2	2696
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_4,5}$	1228	45,5	1469	54,5	2696

Die Gegenüberstellung der Geschlechterverteilung im Zwei-Wellen-Paneldatensatz mit den Querschnittsdatsätzen der Erhebungszeitpunkte  $t_4$  und  $t_5$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Zwei-Wellen-Paneldatensatz etwa viereinhalb bzw. etwas über vier Prozentpunkte höher liegt.

Der Vergleich zwischen den Querschnittsdatsätzen und dem Zwei-Wellen-Paneldatensatz (siehe Tabelle 5.18) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatsätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatsätzen höher ist.

Tabelle 5.19 zeigt, dass der Zwei-Wellen-Paneldatensatz  $P_{t_4,5}$  insgesamt systematisch nach besuchter Schulform verzerrt ist (ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch

**Tabelle 5.18:** Altersdurchschnitt  $P_{t_4,5}$

		Alter	Std.-abw.	gesamt
$t_4$	$Q_{t_4}$	16,1	0,8	3313
	$P_{t_4,5}$	16,0	0,8	2658
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_4,5}$	17,1	0,9	2673

**Tabelle 5.19:** Zusammensetzung  $P_{t_4,5}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704
		$P_{t_4,5}$	23,8	19,9	32,5	21,4	0,9	0,7	0,1	0,7	1229
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181
		$P_{t_4,5}$	0,9	0,5	18,5	23,5	50,5	0,3	2,1	3,6	1168
weiblich	$t_4$	$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682
		$P_{t_4,5}$	15,7	22,8	32,7	27,8	0,5	0,1	0,2	0,2	1456
	$t_5$	$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226
		$P_{t_4,5}$	0,2	0,4	19,1	31,3	44,7	0,1	2,4	1,8	1421
gesamt	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
		$P_{t_4,5}$	19,4	21,5	32,6	24,9	0,7	0,3	0,4	0,1	2685
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407
		$P_{t_4,5}$	0,5	0,4	18,8	27,8	47,3	0,2	2,3	2,6	2589

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

die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung): Während zu  $t_4$  mit einer Abweichung von etwas über zwei Prozentpunkten die Hauptschülerinnen und -schüler im Zwei-Wellen-Paneldatensatz unterrepräsentiert sind, sind zu  $t_5$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte besuchen zu diesem Befragungszeitpunkt noch die Hauptschule. Weiterhin sind im Zwei-Wellen-Paneldatensatz zu  $t_4$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie im Querschnittsdatensatz. Dagegen sind die Gesamtschüler zu  $t_5$  im Zwei-Wellen-Paneldatensatz mit einem Unterschied von über drei Prozentpunkten im Gegensatz zu den Querschnittsdaten überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage

ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. So fällt zu  $t_5$  mit über dreizehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3).

In Bezug auf die Anteile der Gymnasiastinnen und Gymnasiasten liegt im Zwei-Wellen-Paneldatensatz zu  $t_4$  – im Vergleich zu den entsprechenden Querschnittsdatensätzen – mit knapp drei Prozentpunkten eine Überrepräsentation vor, die zu  $t_5$  auf über acht Prozentpunkte ansteigt (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler minimal unterrepräsentiert sind, sind in der folgenden Welle die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Zwei-Wellen-Paneldatensatz  $P_{t_{4,5}}$  – verglichen mit den Querschnittsdaten – überrepräsentiert. Mit etwas über einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel ebenfalls sehr leicht überrepräsentiert. (Ehemalige) Sonderschülerinnen und -schüler sind sowohl zu  $t_4$  als auch zu  $t_5$  zu sehr geringen Anteilen im Zwei-Wellen-Paneldatensatz enthalten; bedeutende Anteilsunterschiede im Vergleich zu den jeweiligen Querschnittsdaten sind nicht erkennbar.

**Tabelle 5.20:** Zusammensetzung  $P_{t_{4,5}}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(%)
männlich	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3
		$P_{t_{4,5}}$	56,1	42,5	45,7	39,4	61,1	88,9	75,0	0,0	45,8
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5
		$P_{t_{4,5}}$	78,6	54,5	44,3	38,2	48,2	75,0	42,4	61,8	45,1
weiblich	$t_4$	$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7
		$P_{t_{4,5}}$	43,9	57,5	54,3	60,6	38,9	11,1	25,0	100,0	54,2
	$t_5$	$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5
		$P_{t_{4,5}}$	21,4	45,5	55,7	61,8	51,8	25,0	57,6	38,2	54,9

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

Tabelle 5.20 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Zwei-Wellen-Paneldatensatz  $P_{t_{4,5}}$  enthalten sind als in den Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).



5.2.2 Zwei-Wellen-Panel 2006/2007 ( $P_{t_{5,6}}$ )

Das Zwei-Wellen-Panel der Jahre 2006 und 2007 enthält 2 514 Fälle. Aus Tabelle 5.21 geht hervor, dass im Zwei-Wellen-Paneldatensatz  $P_{t_{5,6}}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.21:** Zusammensetzung nach Geschlecht  $P_{t_{5,6}}$

		männlich		weiblich		gesamt
		<i>N</i>	%	<i>N</i>	%	(100 %)
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_{5,6}}$	1086	43,2	1428	56,8	2514
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3335
	$P_{t_{5,6}}$	1087	43,2	1427	56,8	2514

Die Gegenüberstellung der Geschlechterverteilung im Zwei-Wellen-Paneldatensatz mit den Querschnittsdatsätzen der Erhebungszeitpunkte  $t_5$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Zwei-Wellen-Paneldatensatz über sechseinhalb bzw. knapp vier Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatsatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

**Tabelle 5.22:** Altersdurchschnitt  $P_{t_{5,6}}$

		Alter	Std.-abw.	gesamt
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_{5,6}}$	17,1	0,9	2489
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{t_{5,6}}$	18,0	0,8	2502

Der Vergleich zwischen den Querschnittsdatsätzen und dem Zwei-Wellen-Paneldatensatz (siehe Tabelle 5.22) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatsätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatsätzen höher ist.

Tabelle 5.23 zeigt, dass der Zwei-Wellen-Paneldatensatz  $P_{t_{5,6}}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: In Bezug auf die Hauptschülerinnen und -schüler sind im Zwei-Wellen-Paneldatensatz keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte bzw. keine Befragten besuchen zu  $t_5$  und  $t_6$  noch die

**Tabelle 5.23:** Zusammensetzung  $P_{t_{5,6}}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)									gesamt
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181
		$P_{t_{5,6}}$	1,0	0,7	22,2	38,0	43,2	0,3	1,9	3,4	1036
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504
		$P_{t_{5,6}}$	—	—	18,6	25,2	32,8	—	10,5	12,9	1055
weiblich	$t_5$	$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226
		$P_{t_{5,6}}$	0,2	0,5	21,2	33,3	40,4	0,1	2,3	2,0	1385
	$t_6$	$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715
		$P_{t_{5,6}}$	—	—	18,1	31,1	30,1	—	10,1	10,5	1395
gesamt	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407
		$P_{t_{5,6}}$	0,5	0,6	21,6	30,7	41,6	0,2	2,1	2,6	2421
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
		$P_{t_{5,6}}$	—	—	18,3	28,6	31,3	—	10,3	11,6	2450

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

Hauptschule. Während sich zu  $t_6$  der Anteil der Gesamtschüler vom entsprechenden Querschnittsdatensatz  $Q_{t_6}$  mit weniger als einem halben Prozentpunkt unterscheidet, sind die Gesamtschüler zu  $t_5$  im Zwei-Wellen-Paneldatensatz mit einem Unterschied von knapp sechs Prozentpunkten im Gegensatz zu den Querschnittsdaten deutlich überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zu  $t_5$  fällt mit knapp neunzehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3). Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Zwei-Wellen-Paneldatensatz  $P_{t_{5,6}}$  ähnlich, mit weniger als einem halben Prozentpunkt sind Befragte, die ein Berufskolleg besuchen nur leicht im Zwei-Wellen-Paneldatensatz überrepräsentiert. Vergleichsgrundlage sind hier stets die Querschnittsdatensätze der betreffenden Erhebungsjahre.

In Bezug auf die Anteile der Gymnasiastinnen und Gymnasiasten liegt im Zwei-Wellen-Paneldatensatz zu  $t_5$  – im Vergleich zu den entsprechenden Querschnittsdatensätzen – mit über elf Prozentpunkten eine deutliche Überrepräsentation vor (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3), die zu  $t_6$  auf knapp zweieinhalb Prozentpunkte absinkt. Die Anteile der Realschülerinnen und -schüler sind zu  $t_5$  und  $t_6$  zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  sind diejenigen Befragten, die eine anderweitige Schule besuchen, mit etwa einem halben Prozentpunkt leicht überrepräsentiert, zu  $t_6$  sind sie im Zwei-Wellen-Paneldatensatz  $P_{t_{5,6}}$  – verglichen mit den jeweiligen Querschnittsdatsätzen – mit knapp einem halben Prozentpunkt nicht nennenswert unterrepräsentiert. Mit etwas über einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel sehr leicht überrepräsentiert. Dagegen zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Zwei-Wellen-Paneldatensatz mit zweieinhalb Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nur zu  $t_5$  im Zwei-Wellen-Paneldatensatz enthalten; ihr Anteil ist zu gering, um Aussagen hinsichtlich der Repräsentativität zuzulassen.

**Tabelle 5.24:** Zusammensetzung  $P_{t_{5,6}}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(%)
männlich	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5
		$P_{t_{5,6}}$	76,9	50,0	43,9	38,0	44,5	75,0	38,5	55,6	42,8
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7
		$P_{t_{5,6}}$	—	—	43,7	38,0	45,2	—	44,0	48,1	43,1
weiblich	$t_5$	$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5
		$P_{t_{5,6}}$	23,1	50,0	56,1	62,0	55,5	25,0	61,5	44,4	57,2
	$t_6$	$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3
		$P_{t_{5,6}}$	—	—	56,3	62,0	54,8	—	56,0	51,9	56,9

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

Tabelle 5.24 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Zwei-Wellen-Paneldatensatz  $P_{t_{5,6}}$  enthalten sind als in den Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

### 5.2.3 Zwei-Wellen-Panel 2004/2006 ( $P_{t_{3,5}}$ )

Das Zwei-Wellen-Panel der Jahre 2004 und 2006 enthält 2 552 Fälle. Aus Tabelle 5.25 geht hervor, dass im Zwei-Wellen-Paneldatensatz  $P_{t_{3,5}}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

Vergleicht man die Geschlechterverteilung zum Erhebungszeitpunkt  $t_3$  im Zwei-Wellen-Paneldatensatz mit der Schulstatistik der Erhebungswelle  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Zwei-Wellen-Paneldatensatz in diesem Jahr etwas über sechs Prozentpunkte höher liegt. Im Vergleich zum entsprechenden Querschnittsdatsatz  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Zwei-Wellen-Paneldatensatz um

**Tabelle 5.25:** Zusammensetzung nach Geschlecht  $P_{t_{3,5}}$

		männlich		weiblich		gesamt
		<i>N</i>	%	<i>N</i>	%	(100 %)
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_{3,5}}$	1142	44,7	1410	55,3	2552
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_{3,5}}$	1155	45,3	1397	54,7	2552

knapp viereinhalb Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Zwei-Wellen-Paneldatensatz mit dem Querschnittsdatensatz des Erhebungszeitpunkts  $t_5$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Zwei-Wellen-Paneldatensatz viereinhalb Prozentpunkte höher liegt.

**Tabelle 5.26:** Altersdurchschnitt  $P_{t_{3,5}}$

		Alter	Std.-abw.	gesamt
$t_3$	$Q_{t_3}$	15,1	0,8	3270
	$P_{t_{3,5}}$	15,0	0,8	2519
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_{3,5}}$	17,1	0,9	2529

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Zwei-Wellen-Paneldatensatz (siehe Tabelle 5.26) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel zu  $t_5$  im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.27 zeigt, dass der Zwei-Wellen-Paneldatensatz  $P_{t_{3,5}}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Der Anteil der Hauptschülerinnen und -schüler liegt im Zwei-Wellen-Paneldatensatz zum Erhebungszeitpunkt  $t_3$  etwas über viereinhalb Prozentpunkte unter denen der Schulstatistik. Zu  $t_5$  sind keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte besuchen zu diesem Befragungszeitpunkt noch die Hauptschule. Weiterhin sind im Zwei-Wellen-Paneldatensatz zu  $t_3$  ähnlich

**Tabelle 5.27:** Zusammensetzung  $P_{t_{3,5}}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
männlich	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635
		$P_{t_{3,5}}$	23,8	20,5	33,7	21,6	—	0,4	—	—	1142
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181
		$P_{t_{3,5}}$	1,0	0,6	18,1	22,8	51,9	0,1	2,3	3,3	1107
	weiblich	$t_3$	$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—
$Q_{t_3}$			19,3	23,5	31,3	25,6	—	0,3	—	—	1704
$P_{t_{3,5}}$			17,4	24,0	31,4	27,0	—	0,2	—	—	1410
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226
		$P_{t_{3,5}}$	0,1	0,4	18,4	30,4	46,3	0,0	2,3	2,1	1350
gesamt		$t_3$	$S_{t_3}$	24,9	21,5	32,3	20,0	—	1,3	—	—
	$Q_{t_3}$		22,6	22,9	31,9	22,2	—	0,5	—	—	3339
	$P_{t_{3,5}}$		20,3	22,4	32,4	24,6	—	0,3	—	—	2552
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407
		$P_{t_{3,5}}$	0,5	0,5	18,2	26,9	48,8	0,0	2,3	2,6	2457

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

viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik. Zu  $t_5$  sind die Gesamtschüler im Zwei-Wellen-Paneldatensatz mit einem Unterschied von zweieinhalb Prozentpunkten im Gegensatz zu den Querschnittsdaten überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Gesicherte Aussagen über die Anteile derjenigen Jugendlichen, die ein Berufskolleg besuchen, können erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. Zu  $t_5$  fällt mit über elfeinhalb Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3).

Im Gegensatz dazu liegt der Anteil der Gymnasiastinnen und Gymnasiasten im Zwei-Wellen-Paneldatensatz zum Zeitpunkt  $t_3$  etwa viereinhalb Prozentpunkte über dem der Schulstatistik. Im Anschluss wird diese Überrepräsentation mit knapp siebeneinhalb Prozentpunkten zu  $t_5$  – jedoch diesmal im Vergleich zum entsprechenden Querschnittsdatensatz – noch stärker (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Der Anteil der Realschülerinnen und -schüler liegt im Zwei-Wellen-Paneldatensatz im Zeitraum  $t_3$  knapp einen Prozentpunkt höher als in der Schulstatistik

dieses Erhebungsjahres. Zu  $t_5$  sind die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Auch diejenigen Befragten, die eine anderweitige Schule besuchen, sind zu  $t_5$  mit weniger als einem Prozentpunkt im Zwei-Wellen-Paneldatensatz  $P_{t_{3,5}}$  – verglichen mit dem zugehörigen Querschnittsdatensatz – überrepräsentiert. Mit etwas über einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel ebenfalls sehr leicht überrepräsentiert. (Ehemalige) Sonderschülerinnen und -schüler sind zu  $t_3$  und  $t_5$  im Zwei-Wellen-Paneldatensatz enthalten, besonders zu  $t_3$  fällt eine Unterrepräsentation im Panel mit einem Prozentpunkt auf.

**Tabelle 5.28:** Zusammensetzung  $P_{t_{3,5}}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(%)
männlich	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0
		$P_{t_{3,5}}$	52,6	40,9	46,5	39,3	—	57,1	—	—	44,7
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5
		$P_{t_{3,5}}$	84,6	58,3	44,6	38,1	47,9	100,0	44,6	55,4	45,1
	weiblich	$t_3$	$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—
$Q_{t_3}$			43,6	52,4	50,2	58,8	—	31,3	—	—	51,0
$P_{t_{3,5}}$			47,4	59,1	53,5	60,7	—	42,9	—	—	55,3
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5
		$P_{t_{3,5}}$	15,4	41,7	55,4	61,9	52,1	0,0	55,4	44,6	54,9

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

Tabelle 5.28 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Zwei-Wellen-Paneldatensatz  $P_{t_{3,5}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

#### 5.2.4 Zwei-Wellen-Panel 2005/2007 ( $P_{t_{4,6}}$ )

Das Zwei-Wellen-Panel der Jahre 2005 und 2007 enthält 2 367 Fälle. Aus Tabelle 5.29 geht hervor, dass im Zwei-Wellen-Paneldatensatz  $P_{t_{4,6}}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

Die Gegenüberstellung der Geschlechterverteilung im Zwei-Wellen-Paneldatensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in

**Tabelle 5.29:** Zusammensetzung nach Geschlecht  $P_{t_4,6}$

		männlich		weiblich		gesamt
		$N$	%	$N$	%	(100 %)
$t_4$	$Q_{t_4}$	1 717	50,4	1 688	49,6	3 405
	$P_{t_4,6}$	1 029	43,5	1 338	56,5	2 367
$t_6$	$Q_{t_6}$	1 567	47,0	1 769	53,0	3 335
	$P_{t_4,6}$	1 019	43,1	1 348	56,9	2 367

Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Zwei-Wellen-Paneldatensatz zu  $t_4$  zunächst knapp sieben, zu  $t_6$  knapp vier Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatsatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

**Tabelle 5.30:** Altersdurchschnitt  $P_{t_4,6}$

		Alter	Std.-abw.	gesamt
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_4,6}$	16,0	0,8	2 337
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_4,6}$	18,0	0,8	2 358

Der Vergleich zwischen den Querschnittsdatsätzen und dem Zwei-Wellen-Paneldatensatz (siehe Tabelle 5.30) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatsätze liegt. Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel zu  $t_6$  im Vergleich zu den Querschnittsdatsätzen höher ist.

Tabelle 5.31 zeigt, dass der Zwei-Wellen-Paneldatensatz  $P_{t_4,6}$  insgesamt systematisch nach besuchter Schulform verzerrt ist (ab dem Erhebungszeitpunkt  $t_4$  stehen wiederum nur noch die Querschnittsdatsätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung): Während zu  $t_4$  mit einer Abweichung von etwas über drei Prozentpunkten die Hauptschülerinnen und -schüler im Zwei-Wellen-Paneldatensatz unterrepräsentiert sind, sind zu  $t_6$  keine gesicherten Aussagen mehr möglich: Kein Befragter besucht zu diesem Befragungszeitpunkt mehr die Hauptschule. Zu  $t_4$  und  $t_6$  sind die Anteile der Gesamtschülerinnen und -schüler von den entsprechenden Querschnittsdatsätzen  $Q_{t_4}$  und  $Q_{t_6}$  mit weniger als einem Prozentpunkt bzw. genau einem Prozentpunkt verschieden, sie sind in beiden Jahren leicht unterrepräsentiert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_6$  getroffen werden. Ver-

**Tabelle 5.31:** Zusammensetzung  $P_{t_{4,6}}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
männlich	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704
		$P_{t_{4,6}}$	22,8	20,0	31,6	23,5	1,0	0,7	0,5	0,0	1022
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504
		$P_{t_{4,6}}$	—	—	16,8	23,8	33,9	—	11,9	13,6	987
weiblich	$t_4$	$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682
		$P_{t_{4,6}}$	15,1	23,1	32,1	28,3	0,7	0,1	0,2	0,3	1335
	$t_6$	$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715
		$P_{t_{4,6}}$	—	—	17,2	30,3	30,4	—	10,8	11,4	1311
gesamt	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
		$P_{t_{4,6}}$	18,5	21,7	31,9	26,2	0,8	0,3	0,3	0,2	2357
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
		$P_{t_{4,6}}$	—	—	17,0	27,5	31,9	—	11,2	12,4	2298

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

gleichsgrundlage ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Zwei-Wellen-Paneldatensatz  $P_{t_{4,6}}$  ähnlich, mit knapp einem Prozentpunkt sind Befragte, die ein Berufskolleg besuchen nur leicht im Zwei-Wellen-Paneldatensatz überrepräsentiert.

Die Anteile der Gymnasiastinnen und Gymnasiasten liegen im Zwei-Wellen-Paneldatensatz zum Zeitpunkt  $t_4$  – im Vergleich zu den entsprechenden Querschnittsdatensätzen – über vier Prozentpunkte höher als in den Querschnittsdaten. Zu  $t_6$  sind die Gymnasiastinnen und Gymnasiasten verglichen mit dem Querschnittsdatensatz mit etwas über einem Prozentpunkt in geringerem Ausmaß überrepräsentiert. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler im Zwei-Wellen-Paneldatensatz und in den Querschnittsdaten etwa gleiche Anteile aufweisen, sind zu  $t_6$  die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Zwei-Wellen-Paneldatensatz  $P_{t_{4,6}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Während im vorliegenden Zwei-Wellen-Panel bis zur Erhebungswelle  $t_6$  keine Befragten enthalten sind, die keine Schule besuchen, zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Zwei-Wellen-Paneldatensatz mit etwas mehr als eineinhalb Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind zu  $t_4$  im Zwei-Wellen-Paneldatensatz



enthalten, Anteilsunterschiede zwischen Querschnitts- und Paneldaten lassen sich hier kaum erkennen.

**Tabelle 5.32:** Zusammensetzung  $P_{t_4,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform <sup>a</sup> (%)								gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3
		$P_{t_4,6}$	53,6	39,8	43,0	38,8	50,0	87,5	62,5	0,0	43,3
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7
		$P_{t_4,6}$	—	—	42,5	37,2	45,7	—	45,3	47,2	43,0
weiblich	$t_4$	$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7
		$P_{t_4,6}$	46,4	60,2	57,0	61,2	50,0	12,5	37,5	100,0	56,7
	$t_6$	$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3
		$P_{t_4,6}$	—	—	57,5	62,8	54,3	—	54,7	52,8	57,0

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

Tabelle 5.32 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Zwei-Wellen-Paneldatensatz  $P_{t_4,6}$  enthalten sind als in den Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

### 5.3 Die Drei-Wellen-Panel

Die Datensätze der Drei-Wellen-Panel enthalten alle Fälle, deren zugehörige – und nach Plausibilitätskontrollen gültige – Fragebögen mindestens zwei (ebenfalls gültigen) Bögen zwei weiterer Zeitpunkte zugeordnet werden konnten. In diesem Bericht werden sämtliche Drei-Wellen-Datensätze dargestellt, die nicht schon bei der Beschreibung des Vier-Wellen-Panels  $t_1$  bis  $t_4$  (vgl. Pöge, 2007) endgültig fertig gestellt waren. In der Praxis werden vermutlich die beiden Datensätze, die drei direkt aufeinander folgende Erhebungszeitpunkte umfassen, am bedeutsamsten sein. Diese werden zuerst dargestellt, im Anschluss werden die restlichen in der Tabelle 4.1 fett gedruckten Drei-Wellen-Paneldatensätze einzeln beschrieben.

#### 5.3.1 Drei-Wellen-Panel 2004 bis 2006 ( $P_{t_3,4,5}$ )

Das Drei-Wellen-Panel der Jahre 2004 bis 2006 enthält 2 354 Fälle. Aus Tabelle 5.33 geht hervor, dass im Drei-Wellen-Paneldatensatz  $P_{t_3,4,5}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.33:** Zusammensetzung nach Geschlecht  $P_{t_3,4,5}$

		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
	$P_{t_3,4,5}$	1042	44,3	1312	55,7	2354
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_3,4,5}$	1054	44,8	1300	55,2	2354
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_3,4,5}$	1050	44,6	1304	55,4	2354

Vergleicht man die Geschlechterverteilung des Erhebungszeitpunkts  $t_3$  im Drei-Wellen-Panel Datensatz mit der Schulstatistik der Erhebungswelle  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Drei-Wellen-Panel Datensatz in diesem Jahr um etwa sechseinhalb Prozentpunkte höher liegt. Im Vergleich zum entsprechenden Querschnittsdatensatz  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Drei-Wellen-Panel Datensatz etwas über viereinhalb Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Drei-Wellen-Panel Datensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$  und  $t_5$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Drei-Wellen-Panel Datensatz in diesen Jahren etwa fünfeinhalb Prozentpunkte höher liegt.

**Tabelle 5.34:** Altersdurchschnitt  $P_{t_3,4,5}$

		Alter	Std.-abw.	gesamt
$t_3$	$Q_{t_3}$	15,1	0,8	3270
	$P_{t_3,4,5}$	15,0	0,8	2325
$t_4$	$Q_{t_4}$	16,1	0,8	3313
	$P_{t_3,4,5}$	16,0	0,7	2326
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_3,4,5}$	17,1	0,9	2336

Das durchschnittliche Alter der Befragten in den Panel Datensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Drei-Wellen-Panel Datensatz (siehe Tabelle 5.34) zeigt, dass das Durchschnittsalter der Befragten im Panel Datensatz durchgehend geringfügig unter dem der einzelnen

Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

**Tabelle 5.35:** Zusammensetzung  $P_{t_3,4,5}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635	
		$P_{t_3,4,5}$	24,0	19,9	33,7	22,1	—	0,4	—	—	1042	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{t_3,4,5}$	23,6	19,8	32,9	21,8	1,0	0,3	0,7	0,0	1048	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{t_3,4,5}$	1,1	0,6	18,5	23,8	50,6	—	2,3	3,1	1004	
	weiblich	$t_3$	$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926
			$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704
$P_{t_3,4,5}$			16,8	23,9	31,8	27,5	—	0,1	—	—	1312	
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{t_3,4,5}$	15,2	23,9	32,2	27,8	0,5	0,0	0,2	0,2	1297	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{t_3,4,5}$	0,2	0,4	18,7	31,6	44,6	—	2,5	2,1	1260	
gesamt		$t_3$	$S_{t_3}$	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
	$P_{t_3,4,5}$		20,0	22,1	32,6	25,1	—	0,2	—	—	2354	
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{t_3,4,5}$	18,9	22,0	32,5	25,1	0,8	0,1	0,4	0,1	2345	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{t_3,4,5}$	0,6	0,5	18,6	28,1	47,3	—	2,4	2,5	2264	

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

Tabelle 5.35 zeigt, dass der Drei-Wellen-Paneldatensatz  $P_{t_3,4,5}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Der Anteil der Hauptschülerinnen und -schüler liegt im Drei-Wellen-Paneldatensatz zum Erhebungszeitpunkt  $t_3$  knapp fünf Prozentpunkte unter dem der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von knapp drei Prozentpunkten die Hauptschülerinnen und -schüler im Drei-Wellen-Paneldatensatz ebenfalls unterrepräsentiert sind, sind zu  $t_5$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte besuchen zu diesem Befragungszeitpunkt noch die Hauptschule. Weiterhin sind im Drei-Wellen-

Paneldatensatz zu  $t_3$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik. Während sich zu  $t_4$  der Anteil der Gesamtschüler vom entsprechenden Querschnittsdatsatz  $Q_{t_4}$  mit weit weniger als einem Prozentpunkt unterscheidet, sind die Gesamtschüler zu  $t_5$  im Drei-Wellen-Paneldatensatz mit einem Unterschied von knapp drei Prozentpunkten im Gegensatz zu den Querschnittsdaten deutlicher überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatsatz des betreffenden Erhebungsjahres. So fällt zu  $t_5$  mit über dreizehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten im Panel auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3).

Im Gegensatz dazu liegt der Anteil der Gymnasiastinnen und Gymnasiasten im Drei-Wellen-Paneldatensatz zum Zeitpunkt  $t_3$  über fünf Prozentpunkte über dem der Schulstatistik. Im Anschluss wird diese Überrepräsentation mit knapp über drei Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatsätzen – zunächst geringer, um zu  $t_5$  auf über achteinhalb Prozentpunkte anzusteigen (zu den Besonderheiten des Querschnittsdatsatzes zu  $t_5$  siehe Abschnitt 5.1.3). Der Anteil der Realschülerinnen und -schüler liegt im Drei-Wellen-Paneldatensatz im Zeitraum  $t_3$  etwa einen halben Prozentpunkt höher als in der Schulstatistik dieses Erhebungsjahres. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler mit weit unter einem Prozentpunkt kaum als überrepräsentiert gelten können, sind in den folgenden Wellen die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Drei-Wellen-Paneldatensatz  $P_{t_{3,4,5}}$  – verglichen mit dem Querschnittsdatsatz – überrepräsentiert. Mit einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel ebenfalls sehr leicht überrepräsentiert. (Ehemalige) Sonderschülerinnen und -schüler sind bis zum Zeitpunkt  $t_4$  im Drei-Wellen-Paneldatensatz enthalten, besonders zu  $t_3$  sind sie mit über einem Prozentpunkt nur leicht unterrepräsentiert.

Tabelle 5.36 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Drei-Wellen-Paneldatensatz  $P_{t_{3,4,5}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

**Tabelle 5.36:** Zusammensetzung  $P_{t_3,4,5}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9	
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0	
		$P_{t_3,4,5}$	53,2	39,8	45,7	38,9	—	80,0	—	—	44,3	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_3,4,5}$	55,6	40,0	45,3	38,8	—	100,0	—	—	44,5	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{t_3,4,5}$	84,6	54,5	44,1	37,5	47,5	—	42,6	54,4	44,3	
	weiblich	$t_3$	$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1
			$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0
$P_{t_3,4,5}$			46,8	60,2	54,3	61,1	—	20,0	—	—	55,7	
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_3,4,5}$	44,4	60,0	54,7	61,2	—	0,0	—	—	55,5	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{t_3,4,5}$	15,4	45,5	55,9	62,5	52,5	—	57,4	45,6	55,7	

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

5.3.2 Drei-Wellen-Panel 2005 bis 2007 ( $P_{t_4,5,6}$ )

Das Drei-Wellen-Panel der Jahre 2005 bis 2007 enthält 2 181 Fälle. Aus Tabelle 5.37 geht hervor, dass im Drei-Wellen-Panel Datensatz  $P_{t_4,5,6}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt.

**Tabelle 5.37:** Zusammensetzung nach Geschlecht  $P_{t_4,5,6}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_4$	$Q_{t_4}$	1 717	50,4	1 688	49,6	3 405
	$P_{t_4,5,6}$	924	42,4	1 257	57,6	2 181
$t_5$	$Q_{t_5}$	2 265	49,8	2 282	50,2	4 547
	$P_{t_4,5,6}$	919	42,1	1 262	57,9	2 181
$t_6$	$Q_{t_6}$	1 567	47,0	1 769	53,0	3 335
	$P_{t_4,5,6}$	917	42,0	1 264	58,0	2 181

Die Gegenüberstellung der Geschlechterverteilung im Drei-Wellen-Panel Datensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$ ,  $t_5$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Drei-Wellen-Panel Datensatz zunächst etwa acht ( $t_4$  und  $t_5$ ), zu  $t_6$  fünf Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatensatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

**Tabelle 5.38:** Altersdurchschnitt  $P_{t_4,5,6}$

		Alter	Std.-abw.	gesamt
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_3,4,6}$	16,0	0,8	2 156
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_3,4,6}$	17,0	0,8	2 168
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_3,4,6}$	18,0	0,8	2 172

Der Vergleich zwischen den Querschnittsdatensätzen und dem Drei-Wellen-Panel Datensatz (siehe Tabelle 5.38) zeigt, dass das Durchschnittsalter der Befragten im Panel Datensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

**Tabelle 5.39:** Zusammensetzung  $P_{t_4,5,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
männlich	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704
		$P_{t_4,5,6}$	21,1	19,7	32,1	25,1	0,9	0,7	0,4	0,0	919
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181
		$P_{t_4,5,6}$	1,1	0,7	21,0	27,6	44,1	0,3	2,1	3,2	878
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504
		$P_{t_4,5,6}$	—	—	17,9	25,2	33,7	—	10,8	12,4	893
weiblich	$t_4$	$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682
		$P_{t_4,5,6}$	14,3	23,7	31,7	29,2	0,5	0,1	0,2	0,2	1255
	$t_5$	$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226
		$P_{t_4,5,6}$	0,2	0,4	20,0	34,1	40,7	0,1	2,4	2,0	1225
	$t_6$	$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715
		$P_{t_4,5,6}$	—	—	17,5	31,3	30,3	—	10,6	10,3	1232
gesamt	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
		$P_{t_4,5,6}$	17,2	22,0	31,9	27,5	0,6	0,3	0,3	0,1	2174
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407
		$P_{t_4,5,6}$	0,6	0,5	20,4	31,4	42,1	0,2	2,2	2,5	2103
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
		$P_{t_4,5,6}$	—	—	17,6	28,8	31,7	—	10,7	11,2	2125

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

Tabelle 5.39 zeigt, dass der Drei-Wellen-Paneldatensatz  $P_{t_4,5,6}$  insgesamt systematisch verzerrt ist (ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatsätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung): Während zu  $t_4$  mit einer Abweichung von viereinhalb Prozentpunkten die Hauptschülerinnen und -schüler im Drei-Wellen-Paneldatensatz deutlich unterrepräsentiert sind, sind zu  $t_5$  und  $t_6$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte bzw. keine Befragten besuchen zu diesen Befragungszeitpunkten noch die Hauptschule. Weiterhin unterscheiden sich im Drei-Wellen-Paneldatensatz zu  $t_4$  und  $t_6$  die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatsätzen  $Q_{t_4}$  und  $Q_{t_6}$  mit weniger als einem Prozentpunkt, zu  $t_5$  sind diese im Drei-Wellen-Paneldatensatz mit einem Unterschied von über viereinhalb Prozentpunkten im Gegensatz zu den Querschnittsdaten dagegen überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage

sind stets die Querschnittsdatensätze der betreffenden Erhebungsjahre. So fällt zu  $t_5$  mit über achtzehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3). Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Drei-Wellen-Paneldatensatz  $P_{t_{4,5,6}}$  ähnlich, mit unter einem Prozentpunkt sind Befragte, die ein Berufskolleg besuchen nur leicht im Drei-Wellen-Paneldatensatz überrepräsentiert.

Die Anteile der Gymnasiastinnen und Gymnasiasten liegen im Drei-Wellen-Paneldatensatz zu  $t_4$  – im Vergleich zu den entsprechenden Querschnittsdatensätzen – etwa fünfeinhalb Prozentpunkte höher, zu  $t_5$  knapp zwölf Prozentpunkte (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Zu  $t_6$  sind die Gymnasiastinnen und Gymnasiasten verglichen mit dem Querschnittsdatensatz mit knapp über zweieinhalb Prozentpunkten wiederum in geringerem Ausmaß überrepräsentiert. Der Anteil der Realschülerinnen und -schüler liegt im Drei-Wellen-Paneldatensatz im Zeitraum  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  in etwa gleichauf, in den folgenden Wellen sind die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Drei-Wellen-Paneldatensatz  $P_{t_{4,5,6}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert, zu  $t_6$  liegen die Anteilswerte genau gleichauf. Mit einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel ebenfalls sehr leicht überrepräsentiert. Dagegen zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Drei-Wellen-Paneldatensatz mit knapp drei Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind bis zum Zeitpunkt  $t_5$  im Drei-Wellen-Paneldatensatz enthalten, jedoch sind ihre Anteile zu gering, um Aussagen zur Repräsentativität zuzulassen.

Tabelle 5.40 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Drei-Wellen-Paneldatensatz  $P_{t_{4,5,6}}$  enthalten sind als in den Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).



**Tabelle 5.40:** Zusammensetzung  $P_{t_{4,5,6}}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(%)
männlich	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3
		$P_{t_{4,5,6}}$	51,9	37,8	42,6	38,6	57,1	85,7	66,7	0,0	42,2
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5
		$P_{t_{4,5,6}}$	76,9	54,5	42,9	36,7	43,7	75,0	38,3	52,8	41,7
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7
		$P_{t_{4,5,6}}$	—	—	42,7	36,8	44,7	—	42,3	46,6	42,0
weiblich	$t_4$	$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7
		$P_{t_{4,5,6}}$	48,1	62,2	57,4	61,4	42,9	14,3	33,3	100,0	57,8
	$t_5$	$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5
		$P_{t_{4,5,6}}$	23,1	45,5	57,1	63,3	56,3	25,0	61,7	47,2	58,3
	$t_6$	$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3
		$P_{t_{4,5,6}}$	—	—	57,3	63,2	55,3	—	57,7	53,4	58,0

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

### 5.3.3 Drei-Wellen-Panel 2002, 2004 und 2006 ( $P_{t_{1,3,5}}$ )

Das Drei-Wellen-Panel der Jahre 2002, 2004 und 2006 enthält 1 793 Fälle. Aus Tabelle 5.41 geht hervor, dass im Drei-Wellen-Paneldatensatz  $P_{t_{1,3,5}}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.41:** Zusammensetzung nach Geschlecht  $P_{t_{1,3,5}}$

		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
	$P_{t_{1,3,5}}$	771	43,0	1022	57,0	1793
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_{1,3,5}}$	772	43,1	1021	56,9	1793
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_{1,3,5}}$	777	43,3	1016	56,7	1793

Vergleicht man die Geschlechterverteilung der Erhebungszeitpunkte  $t_1$  und  $t_3$  im Drei-Wellen-Paneldatensatz mit den Schulstatistiken der Erhebungswellen  $t_1$  und  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Drei-Wellen-Paneldatensatz in jedem

dieser Jahre um etwa acht Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatensätzen  $Q_{t_1}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Drei-Wellen-Paneldatensatz mehr als siebeneinhalb bzw. knapp sieben Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Drei-Wellen-Paneldatensatz mit dem Querschnittsdatensatz des Erhebungszeitpunktes  $t_5$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Drei-Wellen-Paneldatensatz sechseinhalb Prozentpunkte höher liegt.

**Tabelle 5.42:** Altersdurchschnitt  $P_{t_{1,3,5}}$

		Alter	Std.-abw.	gesamt
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{t_{1,3,5}}$	12,9	0,7	1 776
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_{1,3,5}}$	14,9	0,7	1 781
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_{1,3,5}}$	17,0	0,9	1 781

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Drei-Wellen-Paneldatensatz (siehe Tabelle 5.42) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.43 zeigt, dass der Drei-Wellen-Paneldatensatz  $P_{t_{1,3,5}}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Drei-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_1$  und  $t_3$  sieben bzw. siebeneinhalb Prozentpunkte unter denen der Schulstatistik. Zu  $t_5$  sind keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte besuchen zu diesem Befragungszeitpunkt noch die Hauptschule. Weiterhin sind im Drei-Wellen-Paneldatensatz zu  $t_1$  und  $t_3$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik; die Anteilsunterschiede betragen weit weniger als einen Prozentpunkt. Zu  $t_5$  sind die Gesamtschüler im Drei-Wellen-Paneldatensatz mit einem Unterschied von knapp drei Prozentpunkten im Gegensatz zu den Querschnittsdaten überrepräsentiert. Warum die-

ser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Gesicherte Aussagen über die Anteile derjenigen Befragten, die ein Berufskolleg besuchen, können erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. So fällt zu  $t_5$  mit über dreizehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3).

Im Gegensatz dazu liegen die Anteile der Gymnasiastinnen und Gymnasiasten im Drei-Wellen-Paneldatensatz zu den Zeitpunkten  $t_1$  und  $t_3$  etwas über sechs Prozentpunkte über denen der jeweiligen Schulstatistik. Im Anschluss steigt diese Überrepräsentation auf knapp zehn Prozentpunkte zu  $t_5$  an (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Der Anteil der Realschülerinnen und -schüler liegt im Drei-Wellen-Paneldatensatz im Zeitraum  $t_1$  und  $t_3$  knapp eineinhalb bzw. zweieinhalb Prozentpunkte höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. In der Erhebungswelle  $t_5$  sind die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit weniger als einem Prozentpunkt im Drei-Wellen-Paneldatensatz  $P_{t_{1,3,5}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Mit weit unter einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel höchstens minimal unterrepräsentiert. (Ehemalige) Sonderschülerinnen und -schüler sind zu  $t_3$  im Drei-Wellen-Paneldatensatz enthalten, sie sind mit über einem Prozentpunkt unterrepräsentiert.

Tabelle 5.44 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Drei-Wellen-Paneldatensatz  $P_{t_{1,3,5}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

**Tabelle 5.43:** Zusammensetzung  $P_{t_{1,3,5}}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,3	18,1	—	1,3	—	—	2003
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728
		$P_{t_{1,3,5}}$	20,8	20,4	34,1	24,8	—	—	—	—	771
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635
		$P_{t_{1,3,5}}$	21,0	21,2	33,8	23,8	—	0,1	—	—	772
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181
		$P_{t_{1,3,5}}$	1,1	0,9	18,6	25,6	49,5	—	2,1	2,1	749
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—
$Q_{t_1}$			17,9	24,1	31,8	26,3	—	0,0	—	—	1679
$P_{t_{1,3,5}}$			14,0	25,5	30,9	29,5	—	—	—	—	1022
$t_3$		$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704
		$P_{t_{1,3,5}}$	14,7	26,1	31,3	27,9	—	0,0	—	—	1021
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226
		$P_{t_{1,3,5}}$	0,1	0,2	18,7	32,0	45,3	—	2,0	1,6	980
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—
	$Q_{t_1}$		22,0	23,7	31,1	22,8	—	0,4	—	—	3407
	$P_{t_{1,3,5}}$		16,9	23,3	32,3	27,5	—	—	—	—	1793
	$t_3$	$S_{t_3}$	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
		$P_{t_{1,3,5}}$	17,4	24,0	32,4	26,2	—	0,1	—	—	1793
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407
		$P_{t_{1,3,5}}$	0,5	0,5	18,6	29,3	47,1	—	2,1	1,9	1729

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

**Tabelle 5.44:** Zusammensetzung  $P_{1,3,5}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(%)
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7
		$P_{1,3,5}$	52,8	37,6	45,4	38,7	—	—	—	—	43,0
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0
		$P_{1,3,5}$	51,9	38,1	44,9	39,2	—	100,0	—	—	43,1
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5
		$P_{1,3,5}$	88,9	77,8	43,2	37,9	45,5	—	44,4	50,0	43,3
	weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	56,4	—	12,9	—	—
$Q_{t_1}$			40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
$P_{1,3,5}$			47,2	62,4	54,6	61,3	—	—	—	—	57,0
$t_3$		$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1
		$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0
		$P_{1,3,5}$	48,1	61,9	55,1	60,8	—	0,0	—	—	56,9
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5
		$P_{1,3,5}$	11,1	22,2	56,8	62,1	54,5	—	55,6	50,0	56,7

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

5.3.4 Drei-Wellen-Panel 2003, 2004 und 2006 ( $P_{t_2,3,5}$ )

Das Drei-Wellen-Panel der Jahre 2003, 2004 und 2006 enthält 2 082 Fälle. Aus Tabelle 5.45 geht hervor, dass im Drei-Wellen-Panel Datensatz  $P_{t_2,3,5}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.45:** Zusammensetzung nach Geschlecht  $P_{t_2,3,5}$

		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
	$P_{t_2,3,5}$	905	43,5	1177	56,5	2082
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_2,3,5}$	896	43,0	1186	57,0	2082
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_2,3,5}$	900	43,2	1182	56,8	2082

Vergleicht man die Geschlechterverteilung der Erhebungszeitpunkte  $t_2$  und  $t_3$  im Drei-Wellen-Panel Datensatz mit den Schulstatistiken der Erhebungswellen  $t_2$  und  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Drei-Wellen-Panel Datensatz in jedem der Jahre um knapp acht Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatensätzen  $Q_{t_2}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Drei-Wellen-Panel Datensatz etwas über sechseinhalb bzw. sechs Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Drei-Wellen-Panel Datensatz mit dem Querschnittsdatensatz des Erhebungszeitpunkts  $t_5$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Drei-Wellen-Panel Datensatz knapp über sechseinhalb Prozentpunkte höher liegt.

**Tabelle 5.46:** Altersdurchschnitt  $P_{t_2,3,5}$

		Alter	Std.-abw.	gesamt
$t_2$	$Q_{t_2}$	14,0	0,8	3275
	$P_{t_2,3,5}$	13,9	0,7	2053
$t_3$	$Q_{t_3}$	15,1	0,8	3270
	$P_{t_1,3,5}$	15,0	0,7	2062
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_1,3,5}$	17,0	0,9	2063

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Drei-Wellen-Paneldatensatz (siehe Tabelle 5.46) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

**Tabelle 5.47:** Zusammensetzung  $P_{2,3,5}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927	
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703	
		$P_{t_2,3,5}$	23,5	20,1	35,7	20,7	—	—	—	—	905	
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635	
		$P_{t_2,3,5}$	23,7	20,6	35,9	19,8	—	—	—	—	896	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{t_2,3,5}$	1,1	0,7	19,2	22,1	51,7	—	2,1	3,1	870	
	weiblich	$t_2$	$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836
$Q_{t_2}$			20,1	24,9	31,5	23,5	—	0,0	—	—	1687	
$P_{t_2,3,5}$			15,7	25,0	32,5	26,8	—	—	—	—	1177	
$t_3$		$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926	
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704	
		$P_{t_2,3,5}$	16,2	25,4	32,6	25,8	—	—	—	—	1186	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{t_2,3,5}$	0,2	0,3	19,3	30,1	46,5	—	1,9	1,7	1146	
gesamt		$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
			$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390
			$P_{t_2,3,5}$	19,1	22,9	33,9	24,2	—	—	—	—	2082
		$t_3$	$S_{t_3}$	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	
	$P_{t_2,3,5}$		19,4	23,3	34,1	23,2	—	—	—	—	2082	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{t_2,3,5}$	0,6	0,5	19,2	26,6	48,8	—	2,0	2,3	2016	

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

Tabelle 5.47 zeigt, dass der Drei-Wellen-Panel Datensatz  $P_{t_2,3,5}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Drei-Wellen-Panel Datensatz zu den Erhebungszeitpunkten  $t_2$  und  $t_3$  knapp achteinhalb bzw. fünfeinhalb Prozentpunkte unter denen der Schulstatistik. Zu  $t_5$  sind keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte besuchen zu diesem Befragungszeitpunkt noch die Hauptschule. Weiterhin sind im Drei-Wellen-Panel Datensatz zu  $t_2$  und  $t_3$  etwas mehr Gesamtschülerinnen und -schüler vertreten als in der Schulstatistik; sie sind mit dreieinhalb bzw. knapp zwei Prozentpunkten überrepräsentiert. Zu  $t_5$  sind die Gesamtschüler im Drei-Wellen-Panel Datensatz mit einem Unterschied von dreieinhalb Prozentpunkten im Gegensatz zu den Querschnittsdaten ebenfalls überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Gesicherte Aussagen über die Anteile derjenigen Befragten, die ein Berufskolleg besuchen, können erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. So fällt zu  $t_5$  mit über elfeinhalb Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3).

Im Gegensatz dazu liegen die Anteile der Gymnasiastinnen und Gymnasiasten im Drei-Wellen-Panel Datensatz zu den Zeitpunkten  $t_2$  und  $t_3$  etwa fünfeinhalb bzw. drei Prozentpunkte über denen der jeweiligen Schulstatistik. Im Anschluss steigt diese Überrepräsentation zu  $t_5$  – diesmal im Vergleich zum entsprechenden Querschnittsdatensatz – auf über sieben Prozentpunkte an (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Der Anteil der Realschülerinnen und -schüler liegt im Drei-Wellen-Panel Datensatz im Zeitraum  $t_2$  und  $t_3$  weit unter einen bzw. knapp zwei Prozentpunkte höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. In der folgenden Welle sind die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit einem halben Prozentpunkt im Drei-Wellen-Panel Datensatz  $P_{t_2,3,5}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – leicht überrepräsentiert. Mit etwas unter einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel ebenfalls sehr leicht überrepräsentiert. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Drei-Wellen-Panel Datensatz enthalten.

Tabelle 5.48 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Drei-Wellen-Panel Datensatz  $P_{t_2,3,5}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen



**Tabelle 5.48:** Zusammensetzung  $P_{t_2,3,5}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(%)
männlich	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2
		$P_{t_2,3,5}$	53,5	38,2	45,8	37,2	—	—	—	—	43,5
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0
		$P_{t_2,3,5}$	52,5	38,1	45,4	36,6	—	—	—	—	43,0
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5
		$P_{t_2,3,5}$	83,3	60,0	43,0	35,8	45,8	—	45,0	58,7	43,2
	weiblich	$t_2$	$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—
$Q_{t_2}$			40,7	52,5	49,8	58,4	—	0,0	—	—	49,8
$P_{t_2,3,5}$			46,5	61,8	54,2	62,8	—	—	—	—	56,5
$t_3$		$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1
		$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0
		$P_{t_2,3,5}$	47,5	61,9	54,6	63,4	—	—	—	—	57,0
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5
		$P_{t_2,3,5}$	16,7	40,0	57,0	64,2	54,2	—	55,0	41,3	56,8

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

Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

5.3.5 Drei-Wellen-Panel 2003, 2005 und 2006 ( $P_{t_2,4,5}$ )

Das Drei-Wellen-Panel der Jahre 2003, 2005 und 2006 enthält 2 112 Fälle. Aus Tabelle 5.49 geht hervor, dass im Drei-Wellen-Panel Datensatz  $P_{t_2,4,5}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.49:** Zusammensetzung nach Geschlecht  $P_{t_2,4,5}$

		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
	$P_{t_2,4,5}$	920	43,6	1192	56,4	2112
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_2,4,5}$	917	43,4	1195	56,6	2112
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_2,4,5}$	912	43,2	1200	56,8	2112

Vergleicht man die Geschlechterverteilung der Welle  $t_2$  im Drei-Wellen-Panel Datensatz mit der Schulstatistik der Erhebungswelle  $t_2$ , so wird deutlich, dass der Anteil der Schülerinnen im Drei-Wellen-Panel Datensatz in diesem Jahr über siebeneinhalb Prozentpunkte höher liegt. Im Vergleich zum entsprechenden Querschnittsdatensatz  $Q_{t_2}$  ist der Anteil der weiblichen Befragten im Drei-Wellen-Panel Datensatz etwa sechseinhalb Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Drei-Wellen-Panel Datensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$  und  $t_5$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Drei-Wellen-Panel Datensatz sieben bzw. gut sechseinhalb Prozentpunkte höher liegt.

**Tabelle 5.50:** Altersdurchschnitt  $P_{t_2,4,5}$

		Alter	Std.-abw.	gesamt
$t_2$	$Q_{t_2}$	14,0	0,8	3275
	$P_{t_2,4,5}$	14,0	0,7	2080
$t_4$	$Q_{t_4}$	16,1	0,8	3313
	$P_{t_2,4,5}$	16,0	0,7	2082
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_2,4,5}$	17,0	0,9	2096

Das durchschnittliche Alter der Befragten in den Panel Datensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden

Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Drei-Wellen-Paneldatensatz (siehe Tabelle 5.50) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz ab  $t_4$  geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

**Tabelle 5.51:** Zusammensetzung  $P_{t_2,4,5}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927	
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703	
		$P_{t_2,4,5}$	24,7	20,2	34,5	20,7	—	—	—	—	920	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{t_2,4,5}$	23,9	20,4	34,6	20,1	0,5	—	0,5	—	914	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{t_2,4,5}$	1,1	0,6	19,3	22,9	51,1	—	2,2	2,9	877	
	weiblich	$t_2$	$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836
			$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1687
$P_{t_2,4,5}$			16,8	24,2	32,6	26,4	—	—	—	—	1192	
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{t_2,4,5}$	16,2	24,2	32,7	26,4	0,4	—	0,1	—	1194	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{t_2,4,5}$	0,3	0,3	19,7	30,9	45,3	—	2,1	1,5	1165	
gesamt		$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
			$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390
	$P_{t_2,4,5}$		20,2	22,5	33,4	23,9	—	—	—	—	2112	
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{t_2,4,5}$	19,5	22,5	33,5	23,7	0,5	—	0,3	—	2108	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{t_2,4,5}$	0,6	0,4	19,5	27,5	47,8	—	2,1	2,1	2042	

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

Tabelle 5.51 zeigt, dass der Drei-Wellen-Paneldatensatz  $P_{t_2,4,5}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Drei-Wellen-Paneldatensatz zum Erhebungszeitpunkt  $t_2$  etwas mehr als sieben Prozentpunkte unter denen der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur

Verfügung. Während zu  $t_4$  mit einer Abweichung von etwas über zwei Prozentpunkten die Hauptschülerinnen und -schüler im Drei-Wellen-Panel datensatz ebenfalls unterrepräsentiert sind, sind zu  $t_5$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte besuchen zu diesem Befragungszeitpunkt noch die Hauptschule. Weiterhin sind im Drei-Wellen-Panel datensatz zu  $t_2$  die Gesamtschülerinnen und -schüler mit drei Prozentpunkten überrepräsentiert. Während sich zu  $t_4$  der Anteil der Gesamtschüler vom entsprechenden Querschnittsdatensatz  $Q_{t_4}$  mit weniger als einem Prozentpunkt unterscheidet, sind die Gesamtschüler zu  $t_5$  im Drei-Wellen-Panel datensatz mit einem Unterschied von knapp vier Prozentpunkten im Gegensatz zu den Querschnittsdaten deutlicher überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. So fällt zu  $t_5$  mit mehr als zwölf einhalb Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3).

Im Gegensatz dazu liegen die Anteile der Gymnasiastinnen und Gymnasiasten im Drei-Wellen-Panel datensatz zum Zeitpunkt  $t_2$  mehr als fünf Prozentpunkten über denen der entsprechenden Schulstatistik. Im Anschluss wird diese Überrepräsentation mit etwas über eineinhalb Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – zunächst geringer, um zu  $t_5$  auf acht Prozentpunkte anzusteigen (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Der Anteil der Realschülerinnen und -schüler liegt im Drei-Wellen-Panel datensatz im Zeitraum  $t_2$  in etwa gleichauf mit der jeweiligen Schulstatistik dieses Erhebungsjahres. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler mit unter einem Prozentpunkt nur sehr leicht überrepräsentiert sind, sind in der folgenden Welle die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Drei-Wellen-Panel datensatz  $P_{t_2,4,5}$  – verglichen mit dem zugehörigen Querschnittsdatensatz – überrepräsentiert. Mit weit unter einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel höchstens minimal überrepräsentiert. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Drei-Wellen-Panel datensatz enthalten.

Tabelle 5.52 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Drei-Wellen-Panel datensatz  $P_{t_2,4,5}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schul-

**Tabelle 5.52:** Zusammensetzung  $P_{t_2,4,5}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2	
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2	
		$P_{t_2,4,5}$	53,2	39,2	45,0	37,6	—	—	—	—	43,6	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_2,4,5}$	53,0	39,2	44,7	36,9	50,0	—	83,3	—	43,4	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{t_2,4,5}$	76,9	55,6	42,5	35,8	45,9	—	44,2	59,5	42,9	
	weiblich	$t_2$	$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8
			$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8
$P_{t_2,4,5}$			46,8	60,8	55,0	62,4	—	—	—	—	56,4	
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_2,4,5}$	47,0	60,8	55,3	63,1	50,0	—	16,7	—	56,6	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{t_2,4,5}$	23,1	44,4	57,5	64,2	54,1	—	55,8	40,5	57,1	

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

formen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

### 5.3.6 Drei-Wellen-Panel 2003, 2005 und 2007 ( $P_{t_2,4,6}$ )

Das Drei-Wellen-Panel der Jahre 2003, 2005 und 2007 enthält 1 873 Fälle. Aus Tabelle 5.53 geht hervor, dass im Drei-Wellen-Paneldatensatz  $P_{t_2,4,6}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

Vergleicht man die Geschlechterverteilung zum Befragungszeitpunkt  $t_2$  im Drei-Wellen-Paneldatensatz mit der Schulstatistik der Erhebungswelle  $t_2$ , so wird deutlich, dass der Anteil der Schülerinnen im Drei-Wellen-Paneldatensatz in diesem Jahr über neun Prozentpunkte höher liegt. Im Vergleich zum entsprechenden Querschnittsdatensatz  $Q_{t_2}$  ist der Anteil der weiblichen Befragten im Drei-Wellen-Paneldatensatz mehr als acht Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Drei-Wellen-Paneldatensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Drei-Wellen-Paneldatensatz zu  $t_4$  zunächst knapp achteinhalb, zu  $t_6$  fünfeinhalb

**Tabelle 5.53:** Zusammensetzung nach Geschlecht  $P_{t_2,4,6}$

		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
	$P_{t_2,4,6}$	784	41,9	1089	58,1	1873
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_2,4,6}$	786	42,0	1087	58,0	1873
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3335
	$P_{t_2,4,6}$	777	41,5	1096	58,5	1873

Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatensatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

**Tabelle 5.54:** Altersdurchschnitt  $P_{t_2,4,6}$

		Alter	Std.-abw.	gesamt
$t_2$	$Q_{t_2}$	14,0	0,8	3275
	$P_{t_2,4,6}$	13,9	0,7	1844
$t_4$	$Q_{t_4}$	16,1	0,8	3313
	$P_{t_2,4,5}$	15,9	0,7	1848
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{t_2,4,5}$	18,0	0,7	1867

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Drei-Wellen-Paneldatensatz (siehe Tabelle 5.54) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.55 zeigt, dass der Drei-Wellen-Paneldatensatz  $P_{t_2,4,6}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Der Anteil der Hauptschülerinnen und -schüler liegt im Drei-Wellen-Paneldatensatz zum Erhebungszeitpunkt  $t_2$  mehr als acht Prozentpunkte unter dem der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von über drei Prozentpunkten die Hauptschülerinnen und -schüler im Drei-Wellen-Paneldatensatz ebenfalls unterrepräsentiert sind, sind

**Tabelle 5.55:** Zusammensetzung  $P_{t_2,4,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
männlich	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703
		$P_{t_2,4,6}$	23,9	19,9	33,3	23,0	—	—	—	—	784
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704
		$P_{t_2,4,6}$	22,7	20,3	33,6	22,1	0,6	0,1	0,5	—	783
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504
$P_{t_2,4,6}$		—	—	17,2	23,2	33,9	—	12,6	13,1	749	
weiblich	$t_2$	$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836
		$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1687
		$P_{t_2,4,6}$	16,2	24,8	31,7	27,4	—	—	—	—	1089
	$t_4$	$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682
		$P_{t_2,4,6}$	15,3	24,5	32,3	27,2	0,6	0,0	0,2	—	1087
	$t_6$	$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715
$P_{t_2,4,6}$		—	—	18,1	29,8	30,4	—	11,0	10,8	1068	
gesamt	$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
		$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390
		$P_{t_2,4,6}$	19,4	22,7	32,4	25,5	—	—	—	—	1873
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
		$P_{t_2,4,6}$	18,4	22,7	32,8	25,1	0,6	0,1	0,3	—	1870
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
$P_{t_2,4,6}$		—	—	17,7	27,1	31,9	—	11,6	11,7	1817	

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

zu  $t_5$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte besuchen zu diesem Befragungszeitpunkt noch die Hauptschule. Weiterhin sind im Drei-Wellen-Paneldatensatz zu  $t_2$  die Gesamtschülerinnen und -schüler mit zwei Prozentpunkten überrepräsentiert. Zu  $t_4$  und  $t_6$  unterscheiden sich die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatsätzen  $Q_{t_4}$  und  $Q_{t_6}$  mit weniger als einem halben Prozentpunkt – allerdings in abweichende Richtungen.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_6$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatsatz des betreffenden Erhebungsjahres. Zu  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatsatz  $Q_{t_6}$  und im Drei-Wellen-Paneldatensatz  $P_{t_2,4,6}$  ähnlich, mit weniger als einem Prozentpunkt

sind Befragte, die ein Berufskolleg besuchen nur leicht im Drei-Wellen-Paneldatensatz überrepräsentiert.

Der Anteil der Gymnasiastinnen und Gymnasiasten liegt im Drei-Wellen-Paneldatensatz zum Zeitpunkt  $t_2$  knapp sieben Prozentpunkte über dem der Schulstatistik. Im Anschluss wird diese Überrepräsentation – diesmal im Vergleich zu den jeweiligen Querschnittsdatensätzen – mit etwas über drei Prozentpunkten zu  $t_4$  und unter einem Prozentpunkt zu  $t_6$  geringer. Der Anteil der Realschülerinnen und -schüler liegt im Drei-Wellen-Paneldatensatz im Zeitraum  $t_2$  genau gleichauf mit der Schulstatistik dieses Erhebungsjahres. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler mit knapp einem Prozentpunkt leicht überrepräsentiert sind, sind in der folgenden Welle die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Drei-Wellen-Paneldatensatz  $P_{2,4,6}$  – verglichen mit dem betreffenden Querschnittsdatensatz – überrepräsentiert. Während im vorliegenden Drei-Wellen-Panel bis zur Erhebungswelle  $t_6$  keine Befragten enthalten sind, die keine Schule besuchen, zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Drei-Wellen-Paneldatensatz mit knapp zweieinhalb Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nur zu  $t_4$  im Drei-Wellen-Paneldatensatz enthalten, jedoch sind ihre Anteile zu gering, um Aussagen zur Repräsentativität zuzulassen.

Tabelle 5.56 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Drei-Wellen-Paneldatensatz  $P_{2,4,6}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).



**Tabelle 5.56:** Zusammensetzung  $P_{t_2,4,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(%)
männlich	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2
		$P_{t_2,4,6}$	51,5	36,6	43,1	37,7	—	—	—	—	41,9
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3
		$P_{t_2,4,6}$	51,7	37,4	42,8	36,9	45,5	100,0	66,7	—	41,9
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7
$P_{t_2,4,6}$		—	—	40,1	35,4	43,9	—	44,5	46,0	41,2	
weiblich	$t_2$	$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8
		$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8
		$P_{t_2,4,6}$	48,3	63,4	56,9	62,3	—	—	—	—	58,1
	$t_4$	$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7
		$P_{t_2,4,6}$	48,3	62,6	57,2	63,1	54,5	0,0	33,3	—	58,1
	$t_6$	$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3
		$P_{t_2,4,6}$	—	—	59,9	64,6	56,1	—	55,5	54,0	58,8

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

### 5.3.7 Drei-Wellen-Panel 2004, 2005 und 2007 ( $P_{t_3,4,6}$ )

Das Drei-Wellen-Panel der Jahre 2004, 2005 und 2007 enthält 2 068 Fälle. Aus Tabelle 5.57 geht hervor, dass im Drei-Wellen-Paneldatensatz  $P_{t_3,4,6}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

Vergleicht man die Geschlechterverteilung zum Erhebungszeitpunkt  $t_3$  im Drei-Wellen-Paneldatensatz mit der Schulstatistik der Erhebungswelle  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Drei-Wellen-Paneldatensatz in diesem Jahr um über neun Prozentpunkte höher liegt. Im Vergleich zum entsprechenden Querschnittsdatensatz  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Drei-Wellen-Paneldatensatz knapp sieben-einhalb Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Drei-Wellen-Paneldatensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Drei-Wellen-Paneldatensatz zu  $t_4$  zunächst über acht, zu  $t_6$  genau fünf Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatensatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schul-

**Tabelle 5.57:** Zusammensetzung nach Geschlecht  $P_{t_3,4,6}$

		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
	$P_{t_3,4,6}$	861	41,6	1207	58,4	2068
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_3,4,6}$	875	42,3	1193	57,7	2068
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3335
	$P_{t_3,4,6}$	868	42,0	1200	58,0	2068

**Tabelle 5.58:** Altersdurchschnitt  $P_{t_3,4,6}$

		Alter	Std.-abw.	gesamt
$t_3$	$Q_{t_3}$	15,1	0,8	3270
	$P_{t_3,4,6}$	15,0	0,7	2046
$t_4$	$Q_{t_4}$	16,1	0,8	3313
	$P_{t_3,4,6}$	16,0	0,7	2045
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{t_3,4,6}$	18,0	0,7	2061

statistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Drei-Wellen-Paneldatensatz (siehe Tabelle 5.58) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.59 zeigt, dass der Drei-Wellen-Paneldatensatz  $P_{t_3,4,6}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Der Anteil der Hauptschülerinnen und -schüler liegt im Drei-Wellen-Paneldatensatz zum Erhebungszeitpunkt  $t_3$  knapp sechs Prozentpunkte unter dem der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von knapp vier Prozentpunkten die Hauptschülerinnen und -schüler im Drei-Wellen-Paneldatensatz ebenfalls unterrepräsentiert sind, sind zu  $t_6$  keine gesicherten Aussagen mehr möglich: Kein Befragter besucht zu diesem Befragungszeitpunkt mehr die Hauptschule. Weiterhin sind im Drei-Wellen-Paneldatensatz zu  $t_3$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik, der Anteilsunterschied liegt hier bei unter einem Prozentpunkt. Zu  $t_4$  und  $t_6$  sind die

**Tabelle 5.59:** Zusammensetzung  $P_{t_3,4,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
männlich	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635
		$P_{t_3,4,6}$	23,0	20,4	32,1	24,5	—	0,0	—	—	861
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704
		$P_{t_3,4,6}$	22,2	20,6	31,8	23,8	1,0	0,1	0,6	0,0	871
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504
$P_{t_3,4,6}$		—	—	16,9	23,9	34,2	—	12,0	13,0	840	
weiblich	$t_3$	$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704
		$P_{t_3,4,6}$	16,2	24,0	31,2	28,4	—	0,1	—	—	1207
	$t_4$	$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682
		$P_{t_3,4,6}$	14,5	23,9	31,4	28,7	0,8	0,0	0,3	0,3	1190
	$t_6$	$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715
$P_{t_3,4,6}$		—	—	16,7	30,6	30,1	—	11,1	11,6	1168	
gesamt	$t_3$	$S_{t_3}$	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
		$P_{t_3,4,6}$	19,1	22,5	31,6	26,8	—	0,0	—	—	2068
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386
		$P_{t_3,4,6}$	17,8	22,5	31,6	26,6	0,9	0,0	0,4	0,2	2061
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
$P_{t_3,4,6}$		—	—	16,8	27,8	31,8	—	11,5	12,2	2008	

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

Gesamtschüler im Panel im Vergleich zu den jeweiligen Querschnittsdatensätzen mit etwas mehr als einem Prozentpunkt unterrepräsentiert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_6$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. Zu  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Drei-Wellen-Paneldatensatz  $P_{t_3,4,6}$  ähnlich, mit knapp einem Prozentpunkt sind Befragte, die ein Berufskolleg besuchen nur leicht im Drei-Wellen-Paneldatensatz überrepräsentiert.

Der Anteil der Gymnasiastinnen und Gymnasiasten liegt im Drei-Wellen-Paneldatensatz zum Zeitpunkt  $t_3$  knapp sieben Prozentpunkte über dem der Schulstatistik. Im Anschluss wird diese Überrepräsentation mit über viereinhalb Prozentpunkten zu  $t_4$  und mehr als eineinhalb Prozentpunkten zu  $t_6$  – im Vergleich zu den jeweiligen

Querschnittsdaten – geringer. Der Anteil der Realschülerinnen und -schüler liegt im Drei-Wellen-Paneldatensatz im Zeitraum  $t_3$  genau einen Prozentpunkt höher als in der Schulstatistik dieses Erhebungsjahres. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit weniger als einem Prozentpunkt nur leicht überrepräsentiert sind, sind in der folgenden Welle die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Drei-Wellen-Paneldatensatz  $P_{t_3,t_4,t_6}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Während im vorliegenden Drei-Wellen-Paneldatensatz bis zur Erhebungswelle  $t_6$  keine Befragten enthalten sind, die keine Schule besuchen, zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Panel mit knapp zwei Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind bis  $t_4$  im Drei-Wellen-Paneldatensatz enthalten; besonders zu  $t_3$  fällt auf, dass sie im Vergleich zur Schulstatistik mit über einem Prozentpunkt unterrepräsentiert sind.

**Tabelle 5.60:** Zusammensetzung  $P_{t_3,t_4,t_6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9	
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0	
		$P_{t_3,t_4,t_6}$	50,3	37,8	42,3	38,1	—	0,0	—	—	41,6	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_3,t_4,t_6}$	52,7	38,6	42,5	37,7	50,0	100,0	62,5	0,0	42,3	
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7	
		$P_{t_3,t_4,t_6}$	—	—	42,1	36,0	45,0	—	43,7	44,7	41,8	
	weiblich	$t_3$	$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1
			$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0
$P_{t_3,t_4,t_6}$			49,7	62,2	57,7	61,9	—	100,0	—	—	58,4	
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_3,t_4,t_6}$	47,3	61,4	57,5	62,3	50,0	0,0	37,5	100,0	57,7	
$t_6$		$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3	
		$P_{t_3,t_4,t_6}$	—	—	57,9	64,0	55,0	—	56,3	55,3	58,2	

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

Tabelle 5.60 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Drei-Wellen-Paneldatensatz  $P_{t_3,t_4,t_6}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen

Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

### 5.3.8 Drei-Wellen-Panel 2004, 2006 und 2007 ( $P_{t_3,5,6}$ )

Das Drei-Wellen-Panel der Jahre 2004, 2006 und 2007 enthält 2 007 Fälle. Aus Tabelle 5.61 geht hervor, dass im Drei-Wellen-Paneldatensatz  $P_{t_3,5,6}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.61:** Zusammensetzung nach Geschlecht  $P_{t_3,5,6}$

		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
	$P_{t_3,5,6}$	822	41,0	1185	59,0	2007
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_3,5,6}$	834	41,6	1173	58,4	2007
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3335
	$P_{t_3,5,6}$	833	41,5	1174	58,5	2007

Vergleicht man die Geschlechterverteilung zum Erhebungszeitpunkt  $t_3$  im Drei-Wellen-Paneldatensatz mit der Schulstatistik der Erhebungswelle  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Drei-Wellen-Paneldatensatz in diesem Jahr um knapp zehn Prozentpunkte höher liegt. Im Vergleich zum entsprechenden Querschnittsdatensatz  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Drei-Wellen-Paneldatensatz um acht Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Drei-Wellen-Paneldatensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_5$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Drei-Wellen-Paneldatensatz zu  $t_5$  zunächst über acht, zu  $t_6$  fünfeinhalb Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatensatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Drei-Wellen-Paneldatensatz (siehe Tabelle 5.62) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass

**Tabelle 5.62:** Altersdurchschnitt  $P_{3,5,6}$

		Alter	Std.-abw.	gesamt
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{3,5,6}$	15,0	0,8	1 985
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{3,5,6}$	17,0	0,8	1 994
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{3,5,6}$	18,0	0,7	2 000

die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.63 zeigt, dass der Drei-Wellen-Paneldatensatz  $P_{3,5,6}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Der Anteil der Hauptschülerinnen und -schüler liegt im Drei-Wellen-Paneldatensatz zum Erhebungszeitpunkt  $t_3$  knapp sieben Prozentpunkte unter dem der Schulstatistik. Zu  $t_5$  und  $t_6$  sind keine gesicherten Aussagen zur Repräsentativität mehr möglich: Nur sehr wenige Befragte bzw. keine Befragten besuchen zu diesen Befragungszeitpunkten noch die Hauptschule. Weiterhin sind im Drei-Wellen-Paneldatensatz zu  $t_3$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik; es liegt mit einem Prozentpunkt nur eine leichte Unterrepräsentation vor. Während sich zu  $t_6$  der Anteil der Gesamtschüler vom entsprechenden Querschnittsdatensatz  $Q_{t_6}$  ebenfalls mit weniger als einem Prozentpunkt unterscheidet, sind die Gesamtschüler zu  $t_5$  im Drei-Wellen-Paneldatensatz mit einem Unterschied von knapp viereinhalb Prozentpunkten im Gegensatz zu den Querschnittsdaten deutlich überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Gesicherte Aussagen über die Anteile derjenigen Befragten, die ein Berufskolleg besuchen, können erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage sind stets die Querschnittsdatensätze der betreffenden Erhebungsjahre. So fällt zu  $t_5$  mit über achtzehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3). Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Drei-Wellen-Paneldatensatz  $P_{3,5,6}$  ähnlich, mit unter einem Prozentpunkt sind Befragte, die ein Berufskolleg besuchen nur sehr schwach im Drei-Wellen-Paneldatensatz überrepräsentiert.

Der Anteil der Gymnasiastinnen und Gymnasiasten liegt im Drei-Wellen-Paneldatensatz zum Zeitpunkt  $t_3$  über siebeneinhalb Prozentpunkte über dem der Schulstatistik. Im Anschluss steigt diese Überrepräsentation zu  $t_5$  – im Vergleich zu den

**Tabelle 5.63:** Zusammensetzung  $P_{t_3,5,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
männlich	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635
		$P_{t_3,5,6}$	21,9	20,3	32,1	25,7	—	0,0	—	—	822
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181
		$P_{t_3,5,6}$	1,2	0,9	20,7	27,4	44,3	—	2,1	3,4	803
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504
$P_{t_3,5,6}$		—	—	17,8	25,3	33,6	—	11,1	12,2	810	
weiblich	$t_3$	$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704
		$P_{t_3,5,6}$	15,5	24,7	30,7	28,9	—	0,2	—	—	1185
	$t_5$	$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226
		$P_{t_3,5,6}$	0,2	0,4	19,6	34,2	41,0	—	2,3	2,3	1134
	$t_6$	$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715
$P_{t_3,5,6}$		—	—	16,8	31,4	30,3	—	10,9	10,6	1146	
gesamt	$t_3$	$S_{t_3}$	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
		$P_{t_3,5,6}$	18,1	22,9	31,3	27,6	—	0,1	—	—	2007
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407
		$P_{t_3,5,6}$	0,6	0,6	20,0	31,4	42,4	—	2,2	2,7	1937
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219
$P_{t_3,5,6}$		—	—	17,2	28,9	31,6	—	11,0	11,2	1956	

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

Querschnittsdatensätzen – auf knapp zwölf Prozentpunkte an (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Zu  $t_6$  sind die Gymnasias-tinnen und Gymnasiasten verglichen mit dem Querschnittsdatensatz mit etwas mehr als zweieinhalb Prozentpunkten wiederum in geringerem Ausmaß überrepräsentiert. Der Anteil der Realschülerinnen und -schüler liegt im Drei-Wellen-Paneldatensatz im Zeitraum  $t_3$  knapp eineinhalb Prozentpunkte höher als in der Schulstatistik dieses Erhebungsjahres. In den folgenden Wellen sind die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  und  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit weniger als einem Prozentpunkt im Drei-Wellen-Paneldatensatz  $P_{t_3,5,6}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – sehr leicht überrepräsentiert. Mit etwas über einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel ebenfalls sehr leicht überrepräsentiert. Dagegen zeigt sich zum

letzten Erhebungszeitpunkt, dass diese Befragten im Drei-Wellen-Panel Datensatz mit knapp drei Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nur zu  $t_3$  im Drei-Wellen-Panel Datensatz enthalten; zu diesem Zeitpunkt sind sie im Panel Datensatz mit über einem Prozentpunkt unterrepräsentiert.

**Tabelle 5.64:** Zusammensetzung  $P_{t_3,5,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9	
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0	
		$P_{t_3,5,6}$	49,5	36,3	42,0	38,2	—	0,0	—	—	41,0	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{t_3,5,6}$	83,3	58,3	42,8	36,2	43,4	—	39,5	50,9	41,5	
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7	
		$P_{t_3,5,6}$	—	—	42,7	36,3	43,9	—	41,9	45,0	41,4	
	weiblich	$t_3$	$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1
			$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0
$P_{t_3,5,6}$			50,5	63,7	58,0	61,8	—	100,0	—	—	59,0	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{t_3,5,6}$	16,7	41,7	57,2	63,8	56,6	—	60,5	49,1	58,5	
$t_6$		$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3	
		$P_{t_3,5,6}$	—	—	57,3	63,7	56,1	—	58,1	55,0	58,6	

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

Tabelle 5.64 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Drei-Wellen-Panel Datensatz  $P_{t_3,5,6}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

### 5.3.9 Drei-Wellen-Panel 2002, 2005 und 2006 ( $P_{t_1,4,5}$ )

Das Drei-Wellen-Panel der Jahre 2002, 2005 und 2006 enthält 1 861 Fälle. Aus Tabelle 5.65 geht hervor, dass im Drei-Wellen-Panel Datensatz  $P_{t_1,4,5}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:



**Tabelle 5.65:** Zusammensetzung nach Geschlecht  $P_{t_1,4,5}$

		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
	$P_{t_1,4,5}$	817	43,9	1044	56,1	1861
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_1,4,5}$	818	44,0	1043	56,0	1861
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_1,4,5}$	817	43,9	1044	56,1	1861

Vergleicht man die Geschlechterverteilung der ersten Welle im Drei-Wellen-Paneldatensatz mit der Schulstatistik der Erhebungswelle  $t_1$ , so wird deutlich, dass der Anteil der Schülerinnen im Drei-Wellen-Paneldatensatz in diesem Jahr über sieben Prozentpunkte höher liegt. Im Vergleich zum entsprechenden Querschnittsdatensatz  $Q_{t_1}$  ist der Anteil der weiblichen Befragten im Drei-Wellen-Paneldatensatz um knapp sieben Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Drei-Wellen-Paneldatensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$  und  $t_5$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Drei-Wellen-Paneldatensatz etwa sechseinhalb bzw. knapp sechs Prozentpunkte höher liegt.

**Tabelle 5.66:** Altersdurchschnitt  $P_{t_1,4,5}$

		Alter	Std.-abw.	gesamt
$t_1$	$Q_{t_1}$	13,0	0,8	3346
	$P_{t_1,4,5}$	12,9	0,7	1841
$t_4$	$Q_{t_4}$	16,1	0,8	3313
	$P_{t_1,4,5}$	15,9	0,7	1839
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_1,4,5}$	17,0	0,8	1848

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Drei-Wellen-Paneldatensatz (siehe Tabelle 5.66) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen

Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

**Tabelle 5.67:** Zusammensetzung  $P_{t_1,4,5}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,6	18,1	—	1,3	—	—	2003	
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728	
		$P_{t_1,4,5}$	22,3	20,0	32,3	25,1	—	0,4	—	—	817	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{t_1,4,5}$	21,2	20,5	32,9	24,0	0,4	0,4	0,6	—	815	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{t_1,4,5}$	1,0	0,8	19,3	27,1	47,0	0,3	2,3	2,3	783	
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—	1907
			$Q_{t_1}$	17,9	24,1	31,8	26,3	—	0,0	—	—	1679
$P_{t_1,4,5}$			14,5	24,9	31,0	29,6	—	0,0	—	—	1044	
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{t_1,4,5}$	14,3	25,1	31,1	28,7	0,5	0,0	0,3	—	1041	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{t_1,4,5}$	0,2	0,2	19,1	33,2	43,7	0,0	2,2	1,4	1006	
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
			$Q_{t_1}$	22,0	23,7	31,1	22,8	—	0,4	—	—	3407
	$P_{t_1,4,5}$		17,9	22,7	31,6	27,6	—	0,2	—	—	1861	
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{t_1,4,5}$	17,3	23,1	31,9	26,7	0,4	0,2	0,4	—	1856	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{t_1,4,5}$	0,6	0,4	19,2	30,5	45,2	0,1	2,2	1,8	1789	

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

Tabelle 5.67 zeigt, dass der Drei-Wellen-Paneldatensatz  $P_{t_1,4,5}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Der Anteil der Hauptschülerinnen und -schüler liegt im Drei-Wellen-Paneldatensatz zum Erhebungszeitpunkt  $t_1$  sechs Prozentpunkte unter dem der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von etwa viereinhalb Prozentpunkten die Hauptschülerinnen und -schüler im Drei-Wellen-Paneldatensatz ebenfalls deutlich unterrepräsentiert sind, sind zu  $t_5$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte besuchen zu diesem Befragungszeitpunkt noch die Hauptschule. Weiterhin sind

im Drei-Wellen-Paneldatensatz zu  $t_1$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik; der Anteilsunterschied liegt bei unter einem halben Prozentpunkt. Während zu  $t_4$  die Gesamtschüler im Vergleich zum entsprechenden Querschnittsdatensatz  $Q_{t_4}$  mit weniger als einem Prozentpunkt unterrepräsentiert sind, sind die Gesamtschüler zu  $t_5$  im Drei-Wellen-Paneldatensatz mit einem Unterschied von dreieinhalb Prozentpunkten im Gegensatz zu den Querschnittsdaten überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. So fällt zu  $t_5$  mit über fünfzehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten im Panel auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3).

Im Gegensatz dazu liegt der Anteil der Gymnasiastinnen und Gymnasiasten im Drei-Wellen-Paneldatensatz zum Zeitpunkt  $t_1$  über sechs Prozentpunkte über dem der betreffenden Schulstatistik. Im Anschluss wird diese Überrepräsentation mit über viereinhalb Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – zunächst geringer, um zu  $t_5$  auf elf Prozentpunkte anzusteigen (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Der Anteil der Realschülerinnen und -schüler liegt im Drei-Wellen-Paneldatensatz im Zeitraum  $t_1$  etwas über einen halben Prozentpunkt höher als in der Schulstatistik dieses Erhebungsjahres. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit knapp eineinhalb Prozentpunkten überrepräsentiert sind, sind in den folgenden Wellen die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Drei-Wellen-Paneldatensatz  $P_{t_{1,4,5}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Mit weit unter einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel höchstens minimal unterrepräsentiert. (Ehemalige) Sonderschülerinnen und -schüler sind durchgehend im Drei-Wellen-Paneldatensatz enthalten, besonders zu  $t_1$  fällt eine Unterrepräsentation von über einem halben Prozentpunkt auf.

Tabelle 5.68 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Drei-Wellen-Paneldatensatz  $P_{t_{1,4,5}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen

### 5.3 Die Drei-Wellen-Panel

**Tabelle 5.68:** Zusammensetzung  $P_{1,4,5}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2	
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7	
		$P_{t_1,4,5}$	54,7	38,5	44,9	39,9	—	100,0	—	—	43,9	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_1,4,5}$	53,7	39,0	45,3	39,6	—	100,0	—	—	43,9	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{t_1,4,5}$	80,0	75,0	44,0	38,8	45,5	100,0	45,0	56,3	43,8	
	weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	54,6	—	12,9	—	—	48,8
			$Q_{t_1}$	40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
$P_{t_1,4,5}$			45,3	61,5	55,1	60,1	—	0,0	—	—	56,1	
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_1,4,5}$	46,3	61,0	54,7	60,4	—	0,0	—	—	56,1	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{t_1,4,5}$	20,0	25,0	56,0	61,2	54,5	0,0	55,0	43,8	56,2	

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

Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

5.3.10 Drei-Wellen-Panel 2002, 2005 und 2007 ( $P_{t_1,4,6}$ )

Das Drei-Wellen-Panel der Jahre 2002, 2005 und 2007 enthält 1 666 Fälle. Aus Tabelle 5.69 geht hervor, dass im Drei-Wellen-Paneldatensatz  $P_{t_1,4,6}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.69:** Zusammensetzung nach Geschlecht  $P_{t_1,4,6}$

		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
	$P_{t_1,4,6}$	699	42,0	967	58,0	1666
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_1,4,6}$	706	42,4	960	57,6	1666
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3335
	$P_{t_1,4,6}$	701	42,1	965	57,9	1666

Vergleicht man die Geschlechterverteilung der ersten Welle im Drei-Wellen-Paneldatensatz mit der Schulstatistik der Erhebungswelle  $t_1$ , so wird deutlich, dass der Anteil der Schülerinnen im Drei-Wellen-Paneldatensatz in diesem Jahr um über neun Prozentpunkte höher liegt. Im Vergleich zum entsprechenden Querschnittsdatensatz  $Q_{t_1}$  ist der Anteil der weiblichen Befragten im Drei-Wellen-Paneldatensatz über achteinhalb Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Drei-Wellen-Paneldatensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Drei-Wellen-Paneldatensatz zu  $t_4$  zunächst acht, zu  $t_6$  knapp fünf Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatensatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

**Tabelle 5.70:** Altersdurchschnitt  $P_{t_1,4,6}$

		Alter	Std.-abw.	gesamt
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{t_1,4,6}$	12,9	0,7	1 649
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_1,4,6}$	15,9	0,7	1 649
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_1,4,6}$	17,9	0,7	1 660

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Drei-Wellen-Paneldatensatz (siehe Tabelle 5.70) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

**Tabelle 5.71:** Zusammensetzung  $P_{t_{1,4,6}}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,3	18,1	—	1,3	—	—	2003	
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728	
		$P_{t_{1,4,6}}$	20,7	19,7	31,8	27,3	—	0,4	—	—	699	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{t_{1,4,6}}$	19,8	21,1	32,1	25,6	0,6	0,6	0,3	—	702	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504	
		$P_{t_{1,4,6}}$	—	—	17,1	26,7	33,1	—	11,8	11,3	679	
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—	1907
			$Q_{t_1}$	17,9	24,1	31,8	26,3	—	0,0	—	—	1679
$P_{t_{1,4,6}}$			14,2	25,0	31,0	29,8	—	0,0	—	—	967	
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{t_{1,4,6}}$	13,9	24,9	31,3	29,1	0,6	0,0	0,3	—	960	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715	
		$P_{t_{1,4,6}}$	—	—	17,5	31,5	29,3	—	11,4	10,2	937	
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
			$Q_{t_1}$	22,0	23,7	31,1	22,8	—	0,4	—	—	3407
	$P_{t_{1,4,6}}$		16,9	22,8	31,3	28,8	—	0,2	—	—	1666	
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{t_{1,4,6}}$	16,4	23,3	31,6	27,6	0,6	0,2	0,3	—	1662	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
		$P_{t_{1,4,6}}$	—	—	17,3	29,5	30,9	—	11,6	10,7	1616	

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

Tabelle 5.71 zeigt, dass der Drei-Wellen-Paneldatensatz  $P_{t_{1,4,6}}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Der Anteil der Hauptschülerinnen und -schüler liegt im Drei-Wellen-Paneldatensatz zum Erhebungszeitpunkt  $t_1$  sieben Prozentpunkte unter dem der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfü-

gung. Während zu  $t_4$  mit einer Abweichung von etwas über fünf Prozentpunkten die Hauptschülerinnen und -schüler im Drei-Wellen-Paneldatensatz ebenfalls deutlich unterrepräsentiert sind, sind zu  $t_6$  keine gesicherten Aussagen mehr möglich: Kein Befragter besucht zu diesem Befragungszeitpunkt mehr die Hauptschule. Weiterhin sind im Drei-Wellen-Paneldatensatz zu  $t_1$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik; diese sind mit weniger als einem Prozentpunkt unterrepräsentiert. Zu  $t_4$  und  $t_6$  unterscheiden sich die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatensätzen  $Q_{t_4}$  und  $Q_{t_6}$  mit etwa einem Prozentpunkt.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_6$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. So fällt zu  $t_6$  auf, dass sie Anteile derjenigen Befragten, die ein Berufskolleg besuchen, im Querschnitts- und im Paneldatensatz nahezu gleich sind.

Der Anteil der Gymnasiastinnen und Gymnasiasten liegt im Drei-Wellen-Paneldatensatz zum Zeitpunkt  $t_1$  siebeneinhalb Prozentpunkte über dem der jeweiligen Schulstatistik. Im Anschluss wird diese Überrepräsentation mit über fünfeinhalb Prozentpunkten zu  $t_4$  und knapp dreieinhalb Prozentpunkten zu  $t_6$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – geringer. Der Anteil der Realschülerinnen und -schüler liegt im Drei-Wellen-Paneldatensatz im Zeitraum  $t_1$  knapp einen Prozentpunkt höher als in der Schulstatistik dieses Erhebungsjahres. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit eineinhalb Prozentpunkten überrepräsentiert sind, sind in den folgenden Wellen die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit knapp einem Prozentpunkt im Drei-Wellen-Paneldatensatz  $P_{1,4,6}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Während im vorliegenden Drei-Wellen-Paneldatensatz bis zur Erhebungswelle  $t_6$  keine Befragten enthalten sind, die keine Schule besuchen, zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Panel mit knapp dreieinhalb Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind bis  $t_4$  im Drei-Wellen-Paneldatensatz enthalten, besonders zu  $t_1$  fällt mit etwas über einem halben Prozentpunkt eine leichte Unterrepräsentation auf.

Tabelle 5.72 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Drei-Wellen-Paneldatensatz  $P_{1,4,6}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

**Tabelle 5.72:** Zusammensetzung  $P_{1,4,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2	
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7	
		$P_{t_1,4,6}$	51,4	36,3	42,5	39,9	—	100,0	—	—	42,0	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_1,4,6}$	51,1	38,2	42,9	39,2	40,0	100,0	40,0	—	42,2	
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7	
		$P_{t_1,4,6}$	—	—	41,4	38,0	45,0	—	42,8	44,5	42,0	
	weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	54,6	—	12,9	—	—	48,8
			$Q_{t_1}$	40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
$P_{t_1,4,6}$			48,6	63,7	57,5	60,1	—	0,0	—	—	58,0	
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_1,4,6}$	48,9	61,8	57,1	60,8	60,0	0,0	60,0	—	57,8	
$t_6$		$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3	
		$P_{t_1,4,6}$	—	—	58,6	62,0	55,0	—	57,2	55,5	58,0	

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

## 5.4 Die Vier-Wellen-Panel

Die Datensätze der Vier-Wellen-Panel enthalten alle Fälle, deren zugehörige – und nach Plausibilitätskontrollen gültige – Fragebögen mindestens drei (ebenfalls gültigen) Bögen weiterer Zeitpunkte zugeordnet werden konnten. In diesem Bericht werden sämtliche Vier-Wellen-Datensätze dargestellt, die nicht schon bei der Beschreibung des Vier-Wellen-Panels  $t_1$  bis  $t_4$  (vgl. Pöge, 2007) endgültig fertig gestellt waren. In der Praxis werden vermutlich die beiden Datensätze, die vier direkt aufeinander folgende Erhebungszeitpunkte umfassen, am bedeutsamsten sein. Diese werden zuerst dargestellt, im Anschluss werden die restlichen in der Tabelle 4.1 fett gedruckten Vier-Wellen-Paneldatensätze einzeln beschrieben.

### 5.4.1 Vier-Wellen-Panel 2003 bis 2006 ( $P_{2,3,4,5}$ )

Das Vier-Wellen-Panel der Jahre 2003 bis 2006 enthält 1 950 Fälle. Aus Tabelle 5.73 geht hervor, dass im Vier-Wellen-Paneldatensatz  $P_{2,3,4,5}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

Vergleicht man die Geschlechterverteilung der Erhebungszeitpunkte  $t_2$  und  $t_3$  im Vier-Wellen-Paneldatensatz mit den Schulstatistiken der Erhebungswellen  $t_2$  und  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz in jedem der



**Tabelle 5.73:** Zusammensetzung nach Geschlecht  $P_{t_2,3,4,5}$

		männlich		weiblich		gesamt Hfk (100 %)
		Hfk	%	Hfk	%	
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{t_2,3,4,5}$	836	42,9	1114	57,1	1950
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_2,3,4,5}$	828	42,5	1122	57,5	1950
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_2,3,4,5}$	833	42,7	1117	57,3	1950
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_2,3,4,5}$	830	42,6	1120	57,4	1950

beiden Jahre um etwa achteinhalb Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatensätzen  $Q_{t_2}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Vier-Wellen-Paneldatensatz knapp siebeneinhalb bzw. sechseinhalb Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Vier-Wellen-Paneldatensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$  und  $t_5$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz zu  $t_4$  zunächst über siebeneinhalb, zu  $t_5$  knapp über sieben Prozentpunkte höher liegt.

**Tabelle 5.74:** Altersdurchschnitt  $P_{t_2,3,4,5}$

		Alter	Std.-abw.	gesamt Hfk
$t_2$	$Q_{t_2}$	14,0	0,8	3275
	$P_{t_2,3,4,5}$	13,9	0,7	1924
$t_3$	$Q_{t_3}$	15,1	0,8	3270
	$P_{t_2,3,4,5}$	15,0	0,7	1932
$t_4$	$Q_{t_4}$	16,1	0,8	3313
	$P_{t_2,3,4,5}$	15,9	0,7	1926
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_2,3,4,5}$	17,0	0,9	1935

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und

dem Vier-Wellen-Paneldatensatz (siehe Tabelle 5.74) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.75 zeigt, dass der Vier-Wellen-Paneldatensatz  $P_{t_{2,3,4,5}}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Vier-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_2$  und  $t_3$  etwa achteinhalb bzw. fünfeinhalb Prozentpunkte unter denen der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von drei Prozentpunkten die Hauptschülerinnen und -schüler im Vier-Wellen-Paneldatensatz ebenfalls unterrepräsentiert sind, sind zu  $t_5$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte besuchen zu diesem Befragungszeitpunkt noch die Hauptschule. Weiterhin sind im Vier-Wellen-Paneldatensatz zu  $t_2$  und  $t_3$  etwas mehr Gesamtschülerinnen und -schüler vertreten als in der Schulstatistik; zu  $t_2$  sind diese im Vier-Wellen-Paneldatensatz mit über drei Prozentpunkten überrepräsentiert, zu  $t_3$  mit knapp zwei Prozentpunkten. Während sich zu  $t_4$  die Anteile der Gesamtschüler vom entsprechenden Querschnittsdatensatz  $Q_{t_4}$  mit etwa einem Prozentpunkt unterscheiden, sind die Gesamtschüler zu  $t_5$  im Vier-Wellen-Paneldatensatz mit einem Unterschied von vier Prozentpunkten im Gegensatz zu den Querschnittsdaten wieder deutlicher überrepräsentiert. Warum dieser Anteilsunterschied zu  $t_5$  nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. So fällt zu  $t_5$  mit über dreizehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten im Panel auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3).

Im Gegensatz dazu liegen die Anteile der Gymnasiastinnen und Gymnasiasten im Vier-Wellen-Paneldatensatz zu den Zeitpunkten  $t_2$  und  $t_3$  knapp sechs bzw. etwa dreieinhalb Prozentpunkte über denen der jeweiligen Schulstatistik. Im Anschluss wird diese Überrepräsentation mit knapp zwei Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – zunächst geringer, um zu  $t_5$  auf über acht Prozentpunkte anzusteigen (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Der Anteil der Realschülerinnen und -schüler liegt im Vier-Wellen-Paneldatensatz im Zeitraum  $t_2$  in etwa gleichauf mit der Schulstatistik

dieses Erhebungsjahres. Zu  $t_3$  ist er im Vergleich zur jeweiligen Schulstatistik um etwa eineinhalb Prozentpunkte erhöht. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit knapp über einem Prozentpunkt überrepräsentiert sind, sind in den folgenden Wellen die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  ist der Anteil diejenigen Befragten, die eine anderweitige Schule besuchen im Vier-Wellen-Paneldatensatz  $P_{t_{2,3,4,5}}$  etwa einen halben Prozentpunkt höher als im betreffenden Querschnittsdatensatz. Mit weit unter einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel höchstens minimal überrepräsentiert. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Vier-Wellen-Paneldatensatz enthalten.

Tabelle 5.76 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Vier-Wellen-Paneldatensatz  $P_{t_{2,3,4,5}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

## 5.4 Die Vier-Wellen-Panel

**Tabelle 5.75:** Zusammensetzung  $P_{2,3,4,5}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927	
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703	
		$P_{2,3,4,5}$	23,9	19,4	35,5	21,2	—	—	—	—	836	
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635	
		$P_{2,3,4,5}$	24,0	20,0	35,7	20,2	—	—	—	—	828	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{2,3,4,5}$	23,3	19,9	35,3	20,4	0,6	—	0,6	—	830	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{2,3,4,5}$	1,2	0,6	19,7	23,1	50,3	—	2,2	2,7	801	
	weiblich	$t_2$	$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836
			$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1687
$P_{2,3,4,5}$			15,4	24,9	32,6	27,1	—	—	—	—	1114	
$t_3$		$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926	
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704	
		$P_{2,3,4,5}$	15,8	25,3	32,8	26,1	—	—	—	—	1122	
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{2,3,4,5}$	15,3	25,1	32,6	26,4	0,4	—	0,1	—	1116	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{2,3,4,5}$	0,2	0,4	19,6	30,9	45,3	—	2,0	1,6	1086	
gesamt		$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
			$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390
	$P_{2,3,4,5}$		19,1	22,5	33,8	24,6	—	—	—	—	1950	
	$t_3$	$S_{t_3}$	24,9	21,5	32,3	20,0	—	1,3	—	—	3925	
		$Q_{t_3}$	22,6	22,9	31,9	22,2	—	0,5	—	—	3339	
		$P_{2,3,4,5}$	19,3	23,1	34,1	23,6	—	—	—	—	1950	
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{2,3,4,5}$	18,7	22,9	33,8	23,8	0,5	—	0,3	—	1946	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{2,3,4,5}$	0,6	0,5	19,7	27,6	47,4	—	2,1	2,1	1887	

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

**Tabelle 5.76:** Zusammensetzung  $P_{t_2,3,4,5}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2	
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2	
		$P_{t_2,3,4,5}$	53,8	36,9	45,0	37,0	—	—	—	—	42,9	
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9	
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0	
		$P_{t_2,3,4,5}$	52,9	36,9	44,6	36,3	—	—	—	—	42,5	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_2,3,4,5}$	53,0	37,1	44,6	36,4	50,0	—	83,3	—	42,7	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{t_2,3,4,5}$	83,3	55,6	42,6	35,5	45,0	—	45,0	56,4	42,4	
	weiblich	$t_2$	$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8
			$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8
			$P_{t_2,3,4,5}$	46,2	63,1	55,0	63,0	—	—	—	—	57,1
		$t_3$	$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1
			$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0
$P_{t_2,3,4,5}$			47,1	63,1	55,4	63,7	—	—	—	—	57,5	
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_2,3,4,5}$	47,0	62,9	55,4	63,6	50,0	—	16,7	—	57,3	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{t_2,3,4,5}$	16,7	44,4	57,4	64,5	55,0	—	55,0	43,6	57,6	

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

5.4.2 Vier-Wellen-Panel 2004 bis 2007 ( $P_{t_{3,4,5,6}}$ )

Das Vier-Wellen-Panel der Jahre 2004 bis 2007 enthält 1 926 Fälle. Aus Tabelle 5.77 geht hervor, dass im Vier-Wellen-Paneldatensatz  $P_{t_{3,4,5,6}}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.77:** Zusammensetzung nach Geschlecht  $P_{t_{3,4,5,6}}$

		männlich		weiblich		gesamt Hfk (100 %)
		Hfk	%	Hfk	%	
$t_3$	$S_{t_3}$	1 999	50,9	1 926	49,1	3 925
	$Q_{t_3}$	1 635	49,0	1 704	51,0	3 339
	$P_{t_{3,4,5,6}}$	783	40,7	1 143	59,3	1 926
$t_4$	$Q_{t_4}$	1 717	50,4	1 688	49,6	3 405
	$P_{t_{3,4,5,6}}$	795	41,3	1 131	58,7	1 926
$t_5$	$Q_{t_5}$	2 265	49,8	2 282	50,2	4 547
	$P_{t_{3,4,5,6}}$	793	41,2	1 133	58,8	1 926
$t_6$	$Q_{t_6}$	1 567	47,0	1 769	53,0	3 335
	$P_{t_{3,4,5,6}}$	791	41,1	1 135	58,9	1 926

Vergleicht man die Geschlechterverteilung der Erhebungswelle  $t_3$  im Vier-Wellen-Paneldatensatz mit der Schulstatistik der Erhebungswelle  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz in diesem Jahr um etwa zehn Prozentpunkte höher liegt. Im Vergleich zum entsprechenden Querschnittsdatensatz  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Vier-Wellen-Paneldatensatz um über acht Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Vier-Wellen-Paneldatensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$ ,  $t_5$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz zunächst etwa neun bzw. achteinhalb ( $t_4$  und  $t_5$ ), zu  $t_6$  knapp sechs Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatensatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Vier-Wellen-Paneldatensatz (siehe Tabelle 5.78) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass

**Tabelle 5.78:** Altersdurchschnitt  $P_{t_{3,4,5,6}}$

		Alter	Std.-abw.	gesamt Hfk
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_{3,4,5,6}}$	15,0	0,7	1 906
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_{3,4,5,6}}$	16,0	0,7	1 907
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_{3,4,5,6}}$	17,0	0,8	1 915
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_{3,4,5,6}}$	18,0	0,7	1 919

die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.79 zeigt, dass der Vier-Wellen-Paneldatensatz  $P_{t_{3,4,5,6}}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Vier-Wellen-Paneldatensatz zum Erhebungszeitpunkt  $t_3$  knapp sieben Prozentpunkte unter denen der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von knapp fünf Prozentpunkten die Hauptschülerinnen und -schüler im Vier-Wellen-Paneldatensatz ebenfalls deutlich unterrepräsentiert sind, sind zu  $t_5$  und  $t_6$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte bzw. keine Befragten besuchen zu diesen Befragungszeitpunkten noch die Hauptschule. Weiterhin sind im Vier-Wellen-Paneldatensatz zu  $t_3$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik; die Abweichung der Anteile liegt hier bei knapp einem Prozentpunkt. Während sich zu  $t_4$  und  $t_6$  die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatensätzen  $Q_{t_4}$  und  $Q_{t_6}$  mit weniger als einem Prozentpunkt unterscheiden, sind die Gesamtschüler zu  $t_5$  im Vier-Wellen-Paneldatensatz mit einem Unterschied von knapp viereinhalb Prozentpunkten im Gegensatz zu den Querschnittsdaten deutlich überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage sind stets die Querschnittsdatensätze der betreffenden Erhebungsjahre. So fällt zu  $t_5$  mit über achtzehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer

Abschnitt 5.1.3). Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Vier-Wellen-Paneldatensatz  $P_{t_{3,4,5,6}}$  ähnlich, mit unter einem Prozentpunkt sind Befragte, die ein Berufskolleg besuchen nur leicht im Vier-Wellen-Paneldatensatz überrepräsentiert.

Der Anteil der Gymnasiastinnen und Gymnasiasten im Vier-Wellen-Paneldatensatz liegt zum Zeitpunkt  $t_3$  knapp acht Prozentpunkte über dem der betreffenden Schulstatistik. Im Anschluss wird diese Überrepräsentation mit knapp sechs Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – zunächst geringer, um zu  $t_5$  auf knapp über zwölf Prozentpunkte anzusteigen (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Zu  $t_6$  sind die Gymnasiastinnen und Gymnasiasten verglichen mit dem Querschnittsdatensatz mit knapp drei Prozentpunkten wiederum in geringerem Ausmaß überrepräsentiert. Der Anteil der Realschülerinnen und -schüler liegt im Vier-Wellen-Paneldatensatz zu  $t_3$  etwas über einen Prozentpunkt höher als in der Schulstatistik dieses Erhebungsjahres. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler mit knapp einem Prozentpunkt nur schwach überrepräsentiert sind, sind in den folgenden Wellen die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  und  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Vier-Wellen-Paneldatensatz  $P_{t_{3,4,5,6}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – leicht überrepräsentiert. Mit etwas über einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel ebenfalls sehr leicht überrepräsentiert. Dagegen zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Vier-Wellen-Paneldatensatz mit etwa drei Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind im Vier-Wellen-Paneldatensatz zu  $t_3$  und  $t_4$  enthalten; besonders zu  $t_3$  fällt mit etwas über einem Prozentpunkt eine Unterrepräsentation auf.

Tabelle 5.80 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Vier-Wellen-Paneldatensatz  $P_{t_{3,4,5,6}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).



**Tabelle 5.79:** Zusammensetzung  $P_{t_3,4,5,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635	
		$P_{t_3,4,5,6}$	21,8	19,9	32,3	25,9	—	0,0	—	—	783	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{t_3,4,5,6}$	21,1	19,8	32,1	25,3	1,0	0,1	0,5	0,0	791	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{t_3,4,5,6}$	1,3	0,8	20,7	27,5	44,6	—	2,1	3,0	763	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504	
		$P_{t_3,4,5,6}$	—	—	18,0	25,3	34,1	—	10,8	11,8	768	
	weiblich	$t_3$	$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926
			$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704
			$P_{t_3,4,5,6}$	21,8	19,9	32,3	37,8	—	0,1	—	—	1143
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{t_3,4,5,6}$	13,8	24,6	31,0	29,6	0,5	0,0	0,2	0,3	1129	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{t_3,4,5,6}$	0,2	0,5	19,5	34,5	40,7	—	2,4	2,3	1097	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715	
		$P_{t_3,4,5,6}$	—	—	16,9	31,6	29,9	—	11,1	10,5	1107	
gesamt		$t_3$	$S_{t_3}$	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
			$P_{t_3,4,5,6}$	18,0	22,7	31,4	27,9	—	0,1	—	—	1926
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{t_3,4,5,6}$	16,8	22,7	31,5	27,8	0,7	0,1	0,3	0,2	1920	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{t_3,4,5,6}$	0,6	0,6	20,0	31,6	42,3	—	2,3	2,6	1860	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
		$P_{t_3,4,5,6}$	—	—	17,3	29,0	31,6	—	11,0	11,0	1875	

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

**Tabelle 5.80:** Zusammensetzung  $P_{t_3,4,5,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9	
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0	
		$P_{t_3,4,5,6}$	49,4	35,7	41,8	37,8	—	0,0	—	—	40,7	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_3,4,5,6}$	51,7	36,1	42,1	37,5	57,1	100,0	66,7	0,0	41,1	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{t_3,4,5,6}$	83,3	54,5	42,5	35,7	43,2	—	38,1	47,9	41,0	
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7	
		$P_{t_3,4,5,6}$	—	—	42,5	35,7	44,2	—	40,3	44,0	41,0	
	weiblich	$t_3$	$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1
			$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0
			$P_{t_3,4,5,6}$	50,6	64,3	58,2	62,2	—	100,0	—	—	59,3
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_3,4,5,6}$	48,3	63,9	57,9	62,5	42,9	0,0	33,3	100,0	58,9	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{t_3,4,5,6}$	16,7	45,5	57,5	64,3	56,8	—	61,9	52,1	59,0	
$t_6$		$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3	
		$P_{t_3,4,5,6}$	—	—	57,5	64,3	55,8	—	59,7	56,0	59,0	

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

5.4.3 Vier-Wellen-Panel 2002, 2003, 2004 und 2006 ( $P_{t_1,2,3,5}$ )

Das Vier-Wellen-Panel der Jahre 2002, 2003, 2004 und 2006 enthält 1 646 Fälle. Aus Tabelle 5.81 geht hervor, dass im Vier-Wellen-Paneldatensatz  $P_{t_1,2,3,5}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.81:** Zusammensetzung nach Geschlecht  $P_{t_1,2,3,5}$

		männlich		weiblich		gesamt Hfk
		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
	$P_{t_1,2,3,5}$	688	41,8	958	58,2	1646
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{t_1,2,3,5}$	697	42,3	949	57,7	1646
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_1,2,3,5}$	689	41,9	957	58,1	1646
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_1,2,3,5}$	694	42,2	952	57,8	1646

Vergleicht man die Geschlechterverteilung der ersten drei Wellen im Vier-Wellen-Paneldatensatz mit den Schulstatistiken der Erhebungswellen  $t_1$  bis  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz in jedem der Jahre um etwa neun Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatensätzen  $Q_{t_1}$ ,  $Q_{t_2}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Vier-Wellen-Paneldatensatz zwischen knapp neun und sieben Prozentpunkten erhöht. Die Gegenüberstellung der Geschlechterverteilung im Vier-Wellen-Paneldatensatz mit dem Querschnittsdatensatz des Erhebungszeitpunkts  $t_5$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz über siebeneinhalb Prozentpunkte höher liegt.

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Vier-Wellen-Paneldatensatz (siehe Tabelle 5.82) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass

**Tabelle 5.82:** Altersdurchschnitt  $P_{1,2,3,5}$ 

		Alter	Std.-abw.	gesamt Hfk
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{1,2,3,5}$	12,9	0,7	1 630
$t_2$	$Q_{t_2}$	14,0	0,8	3 275
	$P_{1,2,3,5}$	13,9	0,7	1 626
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{1,2,3,5}$	14,9	0,7	1 635
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{1,2,3,5}$	17,0	0,9	1 634

die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.83 zeigt, dass der Vier-Wellen-Paneldatensatz  $P_{1,2,3,5}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Vier-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_1$  bis  $t_3$  zwischen etwa sechseinhalb und knapp zehn Prozentpunkten unter denen der Schulstatistik. Zu  $t_5$  sind keine gesicherten Aussagen zur Repräsentativität mehr möglich: Nur sehr wenige Befragte besuchen zu diesem Befragungszeitpunkt noch die Hauptschule. Weiterhin sind im Vier-Wellen-Paneldatensatz zu  $t_1$ ,  $t_2$  und  $t_3$  etwas mehr Gesamtschülerinnen und Gesamtschüler vertreten wie in der Schulstatistik; zu  $t_2$  sind diese im Vier-Wellen-Paneldatensatz mit etwa drei Prozentpunkten überrepräsentiert, zu  $t_1$  und  $t_3$  jeweils mit etwas über einem Prozentpunkt. Zu  $t_5$  sind die Gesamtschüler im Vier-Wellen-Paneldatensatz mit einem Unterschied von knapp vier Prozentpunkten im Gegensatz zu den Querschnittsdaten überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Gesicherte Aussagen über die Anteile derer, die ein Berufskolleg besuchen, können erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. So fällt zu  $t_5$  mit über zwölfteinhalf Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten im Panel auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3).

Im Gegensatz dazu liegen die Anteile der Gymnasiastinnen und Gymnasiasten im Vier-Wellen-Paneldatensatz zu den Zeitpunkten  $t_1$ ,  $t_2$  und  $t_3$  etwa zwischen dreieinhalb und fünfeinhalb Prozentpunkten über denen der jeweiligen Schulstatistik. Im Anschluss steigt diese Überrepräsentation zu  $t_5$  – diesmal im Vergleich zum Querschnittsdatensatz – auf acht Prozentpunkte an (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe

Abschnitt 5.1.3). Der Anteil der Realschülerinnen und -schüler liegt im Vier-Wellen-Paneldatensatz im Zeitraum  $t_1$  bis  $t_3$  zwischen eineinhalb und dreieinhalb Prozentpunkten höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. Zum Zeitpunkt  $t_5$  sind die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  sind außerdem diejenigen Befragten, die eine anderweitige Schule besuchen, mit etwa einem Prozentpunkt im Vier-Wellen-Paneldatensatz  $P_{t_{1,2,3,5}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Mit weit unter einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel höchstens minimal unterrepräsentiert. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Vier-Wellen-Paneldatensatz enthalten.

Tabelle 5.84 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Vier-Wellen-Paneldatensatz  $P_{t_{1,2,3,5}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

## 5.4 Die Vier-Wellen-Panel

**Tabelle 5.83:** Zusammensetzung  $P_{t_1,2,3,5}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,6	18,1	—	1,3	—	—	2003
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728
		$P_{t_1,2,3,5}$	21,2	21,1	36,3	21,4	—	—	—	—	688
	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703
		$P_{t_1,2,3,5}$	21,8	21,1	36,3	20,8	—	—	—	—	697
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635
		$P_{t_1,2,3,5}$	21,6	21,9	36,0	20,5	—	—	—	—	689
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181
		$P_{t_1,2,3,5}$	1,2	0,9	19,7	22,9	50,8	—	2,2	2,2	669
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—
$Q_{t_1}$			17,9	24,1	31,8	26,3	—	0,0	—	—	1679
$P_{t_1,2,3,5}$			14,5	26,6	31,3	27,6	—	—	—	—	958
$t_2$		$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836
		$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1687
		$P_{t_1,2,3,5}$	14,8	26,9	31,6	26,8	—	—	—	—	949
$t_3$		$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704
		$P_{t_1,2,3,5}$	14,9	27,3	31,9	25,9	—	—	—	—	957
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226
		$P_{t_1,2,3,5}$	0,1	0,2	19,3	30,9	45,8	—	2,1	1,5	921
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—
	$Q_{t_1}$		22,0	23,7	31,1	22,8	—	0,4	—	—	3407
	$P_{t_1,2,3,5}$		17,3	24,3	33,4	25,0	—	—	—	—	1646
	$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
		$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390
		$P_{t_1,2,3,5}$	17,7	24,4	33,6	24,2	—	—	—	—	1646
	$t_3$	$S_{t_3}$	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
		$P_{t_1,2,3,5}$	17,7	25,0	33,6	23,6	—	—	—	—	1646
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407
		$P_{t_1,2,3,5}$	0,6	0,5	19,5	27,5	47,9	—	2,1	1,8	1590

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

**Tabelle 5.84:** Zusammensetzung  $P_{1,2,3,5}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt	
			HS	RS	GS	GY	BK	SO	andere	keine	(%)
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7
		$P_{1,2,3,5}$	51,2	36,3	45,5	35,8	—	—	—	—	41,8
	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2
		$P_{1,2,3,5}$	52,1	36,6	45,8	36,3	—	—	—	—	42,3
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0
		$P_{1,2,3,5}$	51,0	36,7	44,8	36,2	—	—	—	—	41,9
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5
		$P_{1,2,3,5}$	88,9	75,0	42,6	34,9	44,6	—	44,1	51,7	42,1
	weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	54,6	—	12,9	—	—
$Q_{t_1}$			40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
$P_{1,2,3,5}$			48,8	63,8	54,5	64,2	—	—	—	—	58,2
$t_2$		$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8
		$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8
		$P_{1,2,3,5}$	47,9	63,4	54,2	63,7	—	—	—	—	57,7
$t_3$		$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1
		$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0
		$P_{1,2,3,5}$	49,0	63,3	55,2	63,8	—	—	—	—	58,1
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5
		$P_{1,2,3,5}$	11,1	25,0	57,4	65,1	55,4	—	55,9	48,3	57,9

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

5.4.4 Vier-Wellen-Panel 2002, 2003, 2005 und 2006 ( $P_{t_1,2,4,5}$ )

Das Vier-Wellen-Panel der Jahre 2002, 2003, 2005 und 2006 enthält 1 667 Fälle. Aus Tabelle 5.85 geht hervor, dass im Vier-Wellen-Paneldatensatz  $P_{t_1,2,4,5}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.85:** Zusammensetzung nach Geschlecht  $P_{t_1,2,4,5}$

		männlich		weiblich		gesamt Hfk (100 %)
		Hfk	%	Hfk	%	
$t_1$	$S_{t_1}$	2003	51,2	1907	48,8	3910
	$Q_{t_1}$	1728	50,7	1679	49,3	3407
	$P_{t_1,2,4,5}$	709	42,5	958	57,5	1667
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{t_1,2,4,5}$	716	43,0	951	57,0	1667
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_1,2,4,5}$	712	42,7	955	57,3	1667
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_1,2,4,5}$	711	42,7	956	57,3	1667

Vergleicht man die Geschlechterverteilung der ersten zwei Wellen im Vier-Wellen-Paneldatensatz mit den Schulstatistiken der Erhebungswellen  $t_1$  und  $t_2$ , so wird deutlich, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz in jedem der Jahre um etwa achteinhalb Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatensätzen  $Q_{t_1}$  und  $Q_{t_2}$  ist der Anteil der weiblichen Befragten im Vier-Wellen-Paneldatensatz etwa acht bzw. sieben Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Vier-Wellen-Paneldatensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$  und  $t_5$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz in beiden Jahren etwa sieben Prozentpunkte höher liegt.

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Vier-Wellen-Paneldatensatz (siehe Tabelle 5.86) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass



**Tabelle 5.86:** Altersdurchschnitt  $P_{t_1,2,4,5}$

		Alter	Std.-abw.	gesamt Hfk
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{t_1,2,4,5}$	12,9	0,7	1 649
$t_2$	$Q_{t_2}$	14,0	0,8	3 275
	$P_{t_1,2,4,5}$	13,9	0,7	1 644
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_1,2,4,5}$	15,9	0,7	1 646
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_1,2,4,5}$	17,0	0,8	1 656

die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatsätzen höher ist.

Tabelle 5.87 zeigt, dass der Vier-Wellen-Paneldatensatz  $P_{t_1,2,4,5}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Vier-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_1$  und  $t_2$  fünfeinhalb bzw. etwa achteinhalb Prozentpunkte unter denen der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatsätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von etwa dreieinhalb Prozentpunkten die Hauptschülerinnen und -schüler im Vier-Wellen-Paneldatensatz ebenfalls unterrepräsentiert sind, sind zu  $t_5$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte besuchen zu diesem Befragungszeitpunkt noch die Hauptschule. Weiterhin sind im Vier-Wellen-Paneldatensatz zu  $t_1$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik, die Abweichung liegt bei unter einem Prozentpunkt. Zu  $t_2$  sind die Gesamtschülerinnen und -schüler im Vier-Wellen-Paneldatensatz mit zweieinhalb Prozentpunkten überrepräsentiert. Während sich zu  $t_4$  der Anteil der Gesamtschüler vom Querschnittsdatsatz  $Q_{t_4}$  kaum unterscheidet, sind die Gesamtschüler zu  $t_5$  im Vier-Wellen-Paneldatensatz mit einem Unterschied von über vier Prozentpunkten im Gegensatz zu den Querschnittsdaten deutlich überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatsatz des betreffenden Erhebungsjahres. So fällt zu  $t_5$  mit über vierzehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten im Panel auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie

dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3).

Im Gegensatz dazu liegen die Anteile der Gymnasiastinnen und Gymnasiasten im Vier-Wellen-Paneldatensatz zu den Zeitpunkten  $t_1$  und  $t_2$  etwa dreieinhalb bzw. fünfeinhalb Prozentpunkte über denen der jeweiligen Schulstatistik. Im Anschluss wird diese Überrepräsentation mit mehr als zwei Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – zunächst geringer, um zu  $t_5$  auf über neun Prozentpunkte anzusteigen (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Der Anteil der Realschülerinnen und -schüler liegt im Vier-Wellen-Paneldatensatz im Zeitraum  $t_1$  und  $t_2$  etwa zwei Prozentpunkte höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler im Panel ebenfalls mit etwa zweieinhalb Prozentpunkten überrepräsentiert sind, sind in der folgenden Welle die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit knapp unter einem Prozentpunkt im Vier-Wellen-Paneldatensatz  $P_{t_1,2,4,5}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Mit knapp einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel sehr leicht unterrepräsentiert. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Vier-Wellen-Paneldatensatz enthalten.

Tabelle 5.88 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Vier-Wellen-Paneldatensatz  $P_{t_1,2,4,5}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

**Tabelle 5.87:** Zusammensetzung  $P_{t_1,2,4,5}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,3	18,1	—	1,3	—	—	2003	
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728	
		$P_{t_1,2,4,5}$	22,7	21,0	34,4	21,9	—	—	—	—	709	
	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927	
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703	
		$P_{t_1,2,4,5}$	23,0	21,1	34,5	21,4	—	—	—	—	716	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{t_1,2,4,5}$	22,0	21,4	34,4	21,2	0,4	—	0,6	—	709	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{t_1,2,4,5}$	1,2	0,7	19,9	24,2	49,3	—	2,3	2,3	683	
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—	1907
			$Q_{t_1}$	17,9	24,1	31,8	26,3	—	0,0	—	—	1679
$P_{t_1,2,4,5}$			15,1	26,1	31,4	27,3	—	—	—	—	958	
$t_2$		$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836	
		$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1687	
		$P_{t_1,2,4,5}$	15,6	26,1	31,7	26,7	—	—	—	—	951	
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{t_1,2,4,5}$	15,1	26,4	31,6	26,4	0,4	—	0,1	—	954	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{t_1,2,4,5}$	0,2	0,2	19,9	31,9	44,2	—	2,3	1,3	925	
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
			$Q_{t_1}$	22,0	23,7	31,1	22,8	—	0,4	—	—	3407
	$P_{t_1,2,4,5}$		18,4	23,9	32,7	25,0	—	—	—	—	1667	
	$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763	
		$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390	
		$P_{t_1,2,4,5}$	18,8	23,9	32,9	24,4	—	—	—	—	1667	
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{t_1,2,4,5}$	18,0	24,3	32,8	24,2	0,4	—	0,3	—	1663	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{t_1,2,4,5}$	0,6	0,4	19,9	28,6	46,4	—	2,3	1,7	1608	

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

**Tabelle 5.88:** Zusammensetzung  $P_{t_1,2,4,5}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2	
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7	
		$P_{t_1,2,4,5}$	52,6	37,3	44,8	37,2	—	—	—	—	42,5	
	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2	
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2	
		$P_{t_1,2,4,5}$	52,7	37,8	45,1	37,6	—	—	—	—	43,0	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_1,2,4,5}$	52,0	37,6	44,8	37,3	42,9	—	80,0	—	42,6	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{t_1,2,4,5}$	80,0	71,4	42,5	35,9	45,2	—	43,2	57,1	42,5	
	weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	56,4	—	12,9	—	—	48,8
			$Q_{t_1}$	40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
			$P_{t_1,2,4,5}$	47,4	62,7	55,2	62,8	—	—	—	—	57,5
		$t_2$	$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8
			$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8
$P_{t_1,2,4,5}$			47,3	62,2	54,9	62,4	—	—	—	—	57,0	
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_1,2,4,5}$	48,0	62,4	55,2	62,7	57,1	—	20,0	—	57,4	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{t_1,2,4,5}$	20,0	28,6	57,5	64,1	54,8	—	56,8	42,9	57,5	

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

5.4.5 Vier-Wellen-Panel 2002, 2003, 2005 und 2007 ( $P_{t_1,2,4,6}$ )

Das Vier-Wellen-Panel der Jahre 2002, 2003, 2005 und 2007 enthält 1 500 Fälle. Aus Tabelle 5.89 geht hervor, dass im Vier-Wellen-Paneldatensatz  $P_{t_1,2,4,6}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.89:** Zusammensetzung nach Geschlecht  $P_{t_1,2,4,6}$

		männlich		weiblich		gesamt Hfk (100 %)
		Hfk	%	Hfk	%	
$t_1$	$S_{t_1}$	2003	51,2	1907	48,8	3910
	$Q_{t_1}$	1728	50,7	1679	49,3	3407
	$P_{t_1,2,4,6}$	610	40,7	890	59,3	1500
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{t_1,2,4,6}$	617	41,1	883	58,9	1500
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_1,2,4,6}$	618	41,2	882	58,8	1500
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3335
	$P_{t_1,2,4,6}$	612	40,8	888	59,2	1500

Vergleicht man die Geschlechterverteilung der ersten beiden Wellen im Vier-Wellen-Paneldatensatz mit den Schulstatistiken der Erhebungswellen  $t_1$  und  $t_2$ , so wird deutlich, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz in jedem der beiden Jahre um mehr als zehn Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatsätzen  $Q_{t_1}$  und  $Q_{t_2}$  ist der Anteil der weiblichen Befragten im Vier-Wellen-Paneldatensatz um zehn bzw. neun Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Vier-Wellen-Paneldatensatz mit den Querschnittsdatsätzen der Erhebungszeitpunkte  $t_4$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz zu  $t_4$  zunächst über neun, zu  $t_6$  über sechs Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatsatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Paneldatsätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatsätzen und dem Vier-Wellen-Paneldatensatz (siehe Tabelle 5.90) zeigt, dass das Durchschnittsalter der Befragten im Paneldatsatz durchgehend geringfügig unter dem der einzelnen Quer-

**Tabelle 5.90:** Altersdurchschnitt  $P_{1,2,4,6}$ 

		Alter	Std.-abw.	gesamt Hfk
$t_1$	$Q_{t_1}$	13,0	0,8	3346
	$P_{1,2,4,6}$	12,9	0,7	1485
$t_2$	$Q_{t_2}$	14,0	0,8	3275
	$P_{1,2,4,6}$	13,9	0,7	1479
$t_4$	$Q_{t_4}$	16,1	0,8	3313
	$P_{1,2,4,6}$	15,9	0,7	1483
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{1,2,4,6}$	17,9	0,7	1495

schnittsdatensätze liegt. Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.91 zeigt, dass der Vier-Wellen-Paneldatensatz  $P_{1,2,4,6}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Vier-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_1$  und  $t_2$  sechseinhalb bzw. neuneinhalb Prozentpunkte unter denen der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von etwa viereinhalb Prozentpunkten die Hauptschülerinnen und -schüler im Vier-Wellen-Paneldatensatz ebenfalls deutlich unterrepräsentiert sind, sind zu  $t_6$  keine gesicherten Aussagen mehr möglich: Kein Befragter besucht zu diesem Befragungszeitpunkt mehr die Hauptschule. Weiterhin sind im Vier-Wellen-Paneldatensatz zu  $t_1$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik; zu  $t_2$  sind diese im Vier-Wellen-Paneldatensatz mit knapp zwei Prozentpunkten überrepräsentiert. Zu  $t_4$  und  $t_6$  unterscheiden sich die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatensätzen  $Q_{t_4}$  und  $Q_{t_6}$  mit weniger als einem Prozentpunkt.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_6$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. Der Anteil der Berufskollegiatinnen und -kollegiaten ist im Querschnittsdatensatz  $Q_{t_6}$  und im Vier-Wellen-Paneldatensatz  $P_{1,2,4,6}$  ähnlich, mit weniger als einem Prozentpunkt sind Befragte, die ein Berufskolleg besuchen, nur sehr schwach im Vier-Wellen-Paneldatensatz überrepräsentiert.

Die Anteile der Gymnasiastinnen und Gymnasiasten liegen im Vier-Wellen-Paneldatensatz zu den Zeitpunkten  $t_1$  und  $t_2$  etwa fünf bzw. sieben Prozentpunkte über denen der jeweiligen Schulstatistik. Im Anschluss wird diese Überrepräsentation

mit etwa dreieinhalb Prozentpunkten zu  $t_4$  und knapp über einem Prozentpunkt zu  $t_6$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – geringer. Der Anteil der Realschülerinnen und -schüler liegt im Vier-Wellen-Paneldatensatz im Zeitraum  $t_1$  und  $t_2$  knapp zwei bzw. knapp eineinhalb Prozentpunkte höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit zweieinhalb Prozentpunkten überrepräsentiert sind, sind der letzte Welle keine Realschülerinnen und Realschüler mehr enthalten.

Zu  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit eineinhalb Prozentpunkten im Vier-Wellen-Paneldatensatz  $P_{t_1,2,4,6}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Während im vorliegenden Vier-Wellen-Panel bis zur Erhebungswelle  $t_6$  keine Befragten enthalten sind, die keine Schule besuchen, zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Vier-Wellen-Paneldatensatz mit etwa drei Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nur zu  $t_4$  im Vier-Wellen-Paneldatensatz enthalten, der Anteil ist jedoch zu gering, um Rückschlüsse auf die Repräsentativität zuzulassen.

Tabelle 5.92 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Vier-Wellen-Paneldatensatz  $P_{t_1,2,4,6}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

## 5.4 Die Vier-Wellen-Panel

**Tabelle 5.91:** Zusammensetzung  $P_{1,2,4,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,6	18,1	—	1,3	—	—	2003	
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728	
		$P_{1,2,4,6}$	21,3	20,8	33,8	24,1	—	—	—	—	610	
	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927	
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703	
		$P_{1,2,4,6}$	21,7	21,1	33,7	23,5	—	—	—	—	617	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{1,2,4,6}$	20,7	21,8	33,8	22,6	0,7	0,2	0,3	—	615	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504	
		$P_{1,2,4,6}$	—	—	17,4	23,6	34,3	—	12,5	12,2	592	
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—	1907
			$Q_{t_1}$	17,9	24,1	31,8	26,3	—	0,0	—	—	1679
$P_{1,2,4,6}$			14,9	26,1	31,0	28,0	—	—	—	—	890	
$t_2$		$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836	
		$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1687	
		$P_{1,2,4,6}$	15,5	26,0	31,0	27,4	—	—	—	—	883	
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{1,2,4,6}$	14,5	26,1	31,3	27,3	0,6	0,0	0,2	—	882	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715	
		$P_{1,2,4,6}$	—	—	18,1	30,0	29,9	—	11,9	10,2	864	
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
			$Q_{t_1}$	22,0	23,7	31,1	22,8	—	0,4	—	—	3407
	$P_{1,2,4,6}$		17,5	23,9	32,1	26,4	—	—	—	—	1500	
	$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763	
		$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390	
		$P_{1,2,4,6}$	18,1	24,0	32,1	25,8	—	—	—	—	1500	
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{1,2,4,6}$	17,0	24,3	32,3	25,4	0,6	0,1	0,3	—	1497	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
		$P_{1,2,4,6}$	—	—	17,8	27,4	31,7	—	12,2	11,0	1456	

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



**Tabelle 5.92:** Zusammensetzung  $P_{t_1,2,4,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2	
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7	
		$P_{t_1,2,4,6}$	49,4	35,4	42,7	37,1	—	—	—	—	40,7	
	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2	
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2	
		$P_{t_1,2,4,6}$	49,4	36,1	43,2	37,5	—	—	—	—	41,1	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_1,2,4,6}$	49,8	36,8	43,0	36,6	44,4	100,0	50,0	—	41,1	
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7	
		$P_{t_1,2,4,6}$	—	—	39,8	35,1	44,0	—	41,8	45,0	40,7	
	weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	56,4	—	12,9	—	—	48,8
			$Q_{t_1}$	40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
$P_{t_1,2,4,6}$			50,6	64,6	57,3	62,9	—	—	—	—	59,3	
$t_2$		$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8	
		$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8	
		$P_{t_1,2,4,6}$	50,6	63,9	56,8	62,5	—	—	—	—	58,9	
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_1,2,4,6}$	50,2	63,2	57,0	63,4	55,6	0,0	50,0	—	58,9	
$t_6$		$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3	
		$P_{t_1,2,4,6}$	—	—	60,2	64,9	56,0	—	58,2	55,0	59,3	

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

5.4.6 Vier-Wellen-Panel 2002, 2004, 2005 und 2006 ( $P_{t_{1,3,4,5}}$ )

Das Vier-Wellen-Panel der Jahre 2002, 2004, 2005 und 2006 enthält 1 685 Fälle. Aus Tabelle 5.93 geht hervor, dass im Vier-Wellen-Paneldatensatz  $P_{t_{1,3,4,5}}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.93:** Zusammensetzung nach Geschlecht  $P_{t_{1,3,4,5}}$

		männlich		weiblich		gesamt Hfk
		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
	$P_{t_{1,3,4,5}}$	718	42,6	967	57,4	1685
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_{1,3,4,5}}$	720	42,7	965	57,3	1685
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_{1,3,4,5}}$	722	42,8	963	57,2	1685
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_{1,3,4,5}}$	723	42,9	962	57,1	1685

Vergleicht man die Geschlechterverteilung der Befragungszeitpunkte  $t_1$  und  $t_3$  im Vier-Wellen-Paneldatensatz mit den Schulstatistiken der Erhebungswellen  $t_1$  und  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz in jedem der Jahre um etwa achteinhalb Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatensätzen  $Q_{t_1}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Vier-Wellen-Paneldatensatz etwa acht bzw. sechseinhalb Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Vier-Wellen-Paneldatensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$  und  $t_5$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz zu  $t_4$  zunächst über siebeneinhalb, zu  $t_5$  knapp sieben Prozentpunkte höher liegt.

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Vier-Wellen-Paneldatensatz (siehe Tabelle 5.94) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass

**Tabelle 5.94:** Altersdurchschnitt  $P_{t_{1,3,4,5}}$

		Alter	Std.-abw.	gesamt Hfk
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{t_{1,3,4,5}}$	12,9	0,7	1 669
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_{1,3,4,5}}$	14,9	0,7	1 675
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_{1,3,4,5}}$	15,9	0,7	1 666
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_{1,3,4,5}}$	17,0	0,8	1 675

die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.95 zeigt, dass der Vier-Wellen-Paneldatensatz  $P_{t_{1,3,4,5}}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Vier-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_1$  und  $t_3$  sieben bzw. siebeneinhalb Prozentpunkte unter denen der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von fünf Prozentpunkten die Hauptschülerinnen und -schüler im Vier-Wellen-Paneldatensatz ebenfalls unterrepräsentiert sind, sind zu  $t_5$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte besuchen zu diesem Befragungszeitpunkt noch die Hauptschule. Weiterhin sind im Vier-Wellen-Paneldatensatz zu  $t_1$  und  $t_3$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik. Während sich zu  $t_4$  der Anteil der Gesamtschüler vom entsprechenden Querschnittsdatensatz  $Q_{t_4}$  ebenfalls mit weniger als einem Prozentpunkt unterscheidet, sind die Gesamtschüler zu  $t_5$  im Vier-Wellen-Paneldatensatz mit einem Unterschied von über drei Prozentpunkten im Gegensatz zu den Querschnittsdaten überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. So fällt zu  $t_5$  mit über vierzehneinhalb Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten im Panel auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3).

Im Gegensatz dazu liegen die Anteile der Gymnasiastinnen und Gymnasiasten im Vier-Wellen-Paneldatensatz zu den Zeitpunkten  $t_1$  und  $t_3$  etwa sechseinhalb Prozentpunkte über denen der jeweiligen Schulstatistik. Im Anschluss wird diese Überrepräsentation mit knapp viereinhalb Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – zunächst geringer, um zu  $t_5$  auf knapp elf Prozentpunkte anzusteigen (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Der Anteil der Realschülerinnen und -schüler liegt im Vier-Wellen-Paneldatensatz im Zeitraum  $t_1$  und  $t_3$  einen bzw. knapp über zwei Prozentpunkte höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler im Panel ebenfalls mit knapp zwei Prozentpunkten überrepräsentiert sind, sind in der folgenden Welle die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Vier-Wellen-Paneldatensatz  $P_{t_1,3,4,5}$  – verglichen mit dem Querschnittsdatensatz – überrepräsentiert. Mit knapp einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel sehr leicht unterrepräsentiert. (Ehemalige) Sonderschülerinnen und -schüler sind nur zu  $t_3$  im Vier-Wellen-Paneldatensatz enthalten; hier liegt mit über einem Prozentpunkt eine Unterrepräsentation im Panel vor.

Tabelle 5.96 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Vier-Wellen-Paneldatensatz  $P_{t_1,3,4,5}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

**Tabelle 5.95:** Zusammensetzung  $P_{t1,3,4,5}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,3	18,1	—	1,3	—	—	2003	
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728	
		$P_{t_1,3,4,5}$	20,8	19,5	34,1	25,6	—	—	—	—	718	
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635	
		$P_{t_1,3,4,5}$	21,1	20,6	33,8	24,4	—	0,1	—	—	720	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{t_1,3,4,5}$	20,3	20,7	33,8	24,1	0,4	—	0,7	—	719	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{t_1,3,4,5}$	1,1	0,9	19,0	26,6	48,1	—	2,3	2,0	696	
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—	1907
			$Q_{t_1}$	17,9	24,1	31,8	26,3	—	0,0	—	—	1679
$P_{t_1,3,4,5}$			14,1	25,5	31,0	29,4	—	—	—	—	967	
$t_3$		$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926	
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704	
		$P_{t_1,3,4,5}$	14,7	25,8	31,6	27,9	—	0,0	—	—	965	
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{t_1,3,4,5}$	14,0	25,9	31,1	28,2	0,5	—	0,3	—	962	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{t_1,3,4,5}$	0,1	0,2	18,9	33,0	44,1	—	2,2	1,5	927	
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
			$Q_{t_1}$	22,0	23,7	31,1	22,8	—	0,4	—	—	3407
	$P_{t_1,3,4,5}$		16,9	23,0	32,3	27,8	—	—	—	—	1685	
	$t_3$	$S_{t_3}$	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	
		$P_{t_1,3,4,5}$	17,4	23,6	32,5	26,4	—	0,1	—	—	1685	
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{t_1,3,4,5}$	16,7	23,7	32,2	26,4	0,5	—	0,5	—	1681	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{t_1,3,4,5}$	0,6	0,5	18,9	30,3	45,8	—	2,2	1,7	1623	

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

**Tabelle 5.96:** Zusammensetzung  $P_{t_1,3,4,5}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2	
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7	
		$P_{t_1,3,4,5}$	52,3	36,2	45,0	39,3	—	—	—	—	42,6	
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9	
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0	
		$P_{t_1,3,4,5}$	51,7	37,3	44,3	39,6	—	100,0	—	—	42,7	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_1,3,4,5}$	52,0	37,4	44,8	39,0	37,5	—	62,5	—	42,8	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{t_1,3,4,5}$	88,9	75,0	43,0	37,7	45,0	—	44,4	50,0	42,9	
	weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	56,4	—	12,9	—	—	48,8
			$Q_{t_1}$	40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
$P_{t_1,3,4,5}$			47,7	63,8	55,0	60,7	—	—	—	—	57,4	
$t_3$		$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1	
		$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0	
		$P_{t_1,3,4,5}$	48,3	62,7	55,7	60,4	—	0,0	—	—	57,3	
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_1,3,4,5}$	48,0	62,6	55,2	61,0	62,5	—	37,5	—	57,2	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{t_1,3,4,5}$	11,1	25,0	57,0	62,3	55,0	—	55,6	50,0	57,1	

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

5.4.7 Vier-Wellen-Panel 2002, 2004, 2005 und 2007 ( $P_{t_1,3,4,6}$ )

Das Vier-Wellen-Panel der Jahre 2002, 2004, 2005 und 2007 enthält 1 508 Fälle. Aus Tabelle 5.97 geht hervor, dass im Vier-Wellen-Paneldatensatz  $P_{t_1,3,4,6}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.97:** Zusammensetzung nach Geschlecht  $P_{t_1,3,4,6}$

		männlich		weiblich		gesamt Hfk (100 %)
		Hfk	%	Hfk	%	
$t_1$	$S_{t_1}$	2003	51,2	1907	48,8	3910
	$Q_{t_1}$	1728	50,7	1679	49,3	3407
	$P_{t_1,3,4,6}$	607	40,3	901	59,7	1508
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_1,3,4,6}$	609	40,4	899	59,6	1508
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_1,3,4,6}$	614	40,7	894	59,3	1508
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3335
	$P_{t_1,3,4,6}$	610	40,5	898	59,5	1508

Vergleicht man die Geschlechterverteilung der Erhebungszeitpunkte  $t_1$  und  $t_3$  im Vier-Wellen-Paneldatensatz mit den Schulstatistiken der Erhebungswellen  $t_1$  und  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz in jedem der Jahre um etwa zehneinhalb Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatensätzen  $Q_{t_1}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Vier-Wellen-Paneldatensatz um zehneinhalb bzw. achteinhalb Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Vier-Wellen-Paneldatensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz zu  $t_4$  zunächst über neuneinhalb, zu  $t_6$  sechseinhalf Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatensatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Vier-Wellen-Paneldatensatz (siehe Tabelle 5.98) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Quer-

**Tabelle 5.98:** Altersdurchschnitt  $P_{1,3,4,6}$ 

		Alter	Std.-abw.	gesamt Hfk
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{1,3,4,6}$	12,9	0,7	1 494
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{1,3,4,6}$	14,9	0,7	1 499
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{1,3,4,6}$	15,9	0,7	1 494
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{1,3,4,6}$	17,9	0,7	1 503

schnittsdatensätze liegt. Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.99 zeigt, dass der Vier-Wellen-Paneldatensatz  $P_{1,3,4,6}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Vier-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_1$  und  $t_3$  acht bzw. knapp achteinhalb Prozentpunkte unter denen der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von sechs Prozentpunkten die Hauptschülerinnen und -schüler im Vier-Wellen-Paneldatensatz ebenfalls unterrepräsentiert sind, sind zu  $t_6$  keine gesicherten Aussagen mehr möglich: Kein Befragter besucht zu diesem Befragungszeitpunkt mehr die Hauptschule. Weiterhin sind im Vier-Wellen-Paneldatensatz zu  $t_1$  und  $t_3$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik. Zu  $t_4$  und  $t_6$  unterscheiden sich die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatensätzen  $Q_{t_4}$  und  $Q_{t_6}$  ebenfalls mit genau einem Prozentpunkt bzw. weniger als einem Prozentpunkt; Gesamtschülerinnen und -schüler sind sehr schwach unterrepräsentiert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_6$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Vier-Wellen-Paneldatensatz  $P_{1,3,4,6}$  ähnlich, der Anteilsunterschied liegt bei weit unter einem Prozentpunkt.

Die Anteile der Gymnasiastinnen und Gymnasiasten liegen im Vier-Wellen-Paneldatensatz zu den Zeitpunkten  $t_1$  und  $t_3$  knapp acht Prozentpunkte über denen der jeweiligen Schulstatistik. Im Anschluss wird diese Überrepräsentation mit etwa fünfeinhalb Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden



Querschnittsdatensätzen – geringer, um zu  $t_6$  noch weiter auf knapp über drei Prozentpunkte abzusinken. Der Anteil der Realschülerinnen und -schüler liegt im Vier-Wellen-Paneldatensatz im Zeitraum  $t_1$  und  $t_3$  knapp eineinhalb bzw. zweieinhalb Prozentpunkte höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit etwas über zwei Prozentpunkten überrepräsentiert sind, sind in den folgenden Wellen die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit knapp über einem Prozentpunkt im Vier-Wellen-Paneldatensatz  $P_{t_{1,3,4,6}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Während im vorliegenden Vier-Wellen-Paneldatensatz bis zur Erhebungswelle  $t_6$  keine Befragten enthalten sind, die keine Schule besuchen, zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Panel mit dreieinhalb Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Vier-Wellen-Paneldatensatz enthalten.

Tabelle 5.100 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Vier-Wellen-Paneldatensatz  $P_{t_{1,3,4,6}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

## 5.4 Die Vier-Wellen-Panel

**Tabelle 5.99:** Zusammensetzung  $P_{1,3,4,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,3	18,1	—	1,3	—	—	2003	
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728	
		$P_{1,3,4,6}$	19,1	20,1	33,1	27,7	—	—	—	—	607	
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635	
		$P_{1,3,4,6}$	19,9	21,5	32,3	26,3	—	—	—	—	609	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{1,3,4,6}$	18,8	21,7	32,8	25,7	0,7	—	0,3	—	612	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504	
		$P_{1,3,4,6}$	—	—	17,1	26,3	33,2	—	12,2	11,2	590	
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—	1907
			$Q_{t_1}$	17,9	24,1	31,8	26,3	—	0,0	—	—	1679
$P_{1,3,4,6}$			13,8	25,4	30,7	30,1	—	—	—	—	901	
$t_3$		$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926	
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704	
		$P_{1,3,4,6}$	14,3	25,6	31,3	28,8	—	—	—	—	899	
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{1,3,4,6}$	13,5	25,6	30,9	29,0	0,7	—	0,3	—	894	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715	
		$P_{1,3,4,6}$	—	—	17,3	31,3	29,6	—	11,6	10,2	872	
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
			$Q_{t_1}$	22,0	23,7	31,1	22,8	—	0,4	—	—	3407
	$P_{1,3,4,6}$		15,9	23,3	31,7	29,1	—	—	—	—	1508	
	$t_3$	$S_{t_3}$	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	
		$P_{1,3,4,6}$	16,6	23,9	31,7	27,8	—	—	—	—	1508	
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{1,3,4,6}$	15,7	24,0	31,7	27,6	0,7	—	0,3	—	1506	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
		$P_{1,3,4,6}$	—	—	17,2	29,3	31,1	—	11,8	10,6	1462	

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

**Tabelle 5.100:** Zusammensetzung  $P_{t_1,3,4,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2	
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7	
		$P_{t_1,3,4,6}$	48,3	34,8	42,1	38,3	—	—	—	—	40,3	
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9	
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0	
		$P_{t_3,3,4,6}$	48,4	36,3	41,2	38,2	—	—	—	—	40,4	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_4,3,4,6}$	48,7	36,7	42,1	37,7	40,0	—	40,0	—	40,6	
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7	
		$P_{t_6,3,4,6}$	—	—	40,1	36,2	43,2	—	41,6	42,6	40,4	
	weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	56,4	—	12,9	—	—	48,8
			$Q_{t_1}$	40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
$P_{t_1,3,4,6}$			51,7	65,2	57,9	61,7	—	—	—	—	59,7	
$t_3$		$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1	
		$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0	
		$P_{t_3,3,4,6}$	51,6	63,7	58,8	61,8	—	—	—	—	59,6	
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_4,3,4,6}$	51,3	63,3	57,9	62,3	60,0	—	60,0	—	59,4	
$t_6$		$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3	
		$P_{t_6,3,4,6}$	—	—	59,9	63,8	56,8	—	58,4	57,4	59,6	

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

5.4.8 Vier-Wellen-Panel 2002, 2004, 2006 und 2007 ( $P_{t_1,3,5,6}$ )

Das Vier-Wellen-Panel der Jahre 2002, 2004, 2006 und 2007 enthält 1 455 Fälle. Aus Tabelle 5.101 geht hervor, dass im Vier-Wellen-Panel Datensatz  $P_{t_1,3,5,6}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.101:** Zusammensetzung nach Geschlecht  $P_{t_1,3,5,6}$

		männlich		weiblich		gesamt Hfk (100 %)
		Hfk	%	Hfk	%	
$t_1$	$S_{t_1}$	2003	51,2	1907	48,8	3910
	$Q_{t_1}$	1728	50,7	1679	49,3	3407
	$P_{t_1,3,5,6}$	577	39,7	878	60,3	1455
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_1,3,5,6}$	577	39,7	878	60,3	1455
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_1,3,5,6}$	582	40,0	873	60,0	1455
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3335
	$P_{t_1,3,5,6}$	582	40,0	873	60,0	1455

Vergleicht man die Geschlechterverteilung der Erhebungszeitpunkte  $t_1$  und  $t_3$  im Vier-Wellen-Panel Datensatz mit den Schulstatistiken der Erhebungswellen  $t_1$  und  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Vier-Wellen-Panel Datensatz in jedem der Jahre um etwa elfeinhalb Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatensätzen  $Q_{t_1}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Vier-Wellen-Panel Datensatz elf bzw. etwas über neun Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Vier-Wellen-Panel Datensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_5$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Vier-Wellen-Panel Datensatz zu  $t_5$  zunächst knapp zehn, zu  $t_6$  sieben Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatensatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Panel Datensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Vier-Wellen-Panel Datensatz (siehe Tabelle 5.102) zeigt, dass das Durchschnittsalter der Befragten im Panel Datensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhe-

**Tabelle 5.102:** Altersdurchschnitt  $P_{1,3,5,6}$

		Alter	Std.-abw.	gesamt Hfk
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{1,3,5,6}$	12,9	0,7	1 444
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{1,3,5,6}$	14,9	0,7	1 446
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{1,3,5,6}$	16,9	0,7	1 447
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{1,3,5,6}$	17,9	0,7	1 450

bungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatsätzen höher ist.

Tabelle 5.103 zeigt, dass der Vier-Wellen-Paneldatensatz  $P_{1,3,5,6}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Vier-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_1$  und  $t_3$  etwa neun Prozentpunkte unter denen der Schulstatistik. Zu  $t_5$  und  $t_6$  sind in Bezug auf Hauptschülerinnen und -schüler keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte bzw. keine Befragten besuchen zu diesen Befragungszeitpunkten noch die Hauptschule. Weiterhin sind im Vier-Wellen-Paneldatensatz zu  $t_1$  und  $t_3$  Gesamtschülerinnen und -schüler mit etwa einem Prozentpunkt im Panel leicht überrepräsentiert. Während sich zu  $t_6$  der Anteil der Gesamtschüler von den entsprechenden Querschnittsdatsätzen  $Q_{t_6}$  nur minimal unterscheidet, sind die Gesamtschüler zu  $t_5$  im Vier-Wellen-Paneldatensatz mit einem Unterschied von über viereinhalb Prozentpunkten im Gegensatz zu den Querschnittsdaten überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Gesicherte Aussagen über die Anteile derer, die ein Berufskolleg besuchen, können ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage sind stets die Querschnittsdatsätze der betreffenden Erhebungsjahre. So fällt zu  $t_5$  mit neunzehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3). Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatsatz  $Q_{t_6}$  und im Vier-Wellen-Paneldatensatz  $P_{1,3,5,6}$  genau übereinstimmend.

Die Anteile der Gymnasiastinnen und Gymnasiasten liegen im Vier-Wellen-Paneldatensatz zu den Zeitpunkten  $t_1$  und  $t_3$  etwa neun Prozentpunkte über denen der

jeweiligen Schulstatistik. Im Anschluss steigt diese Überrepräsentation zu  $t_5$  auf knapp vierzehn Prozentpunkte an (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Zu  $t_6$  sind die Gymnasiastinnen und Gymnasiasten verglichen mit dem Querschnittsdatensatz mit knapp viereinhalb Prozentpunkten wiederum in geringerem Ausmaß überrepräsentiert. Der Anteil der Realschülerinnen und -schüler liegt im Vier-Wellen-Paneldatensatz im Zeitraum  $t_1$  und  $t_3$  eineinhalb bzw. knapp drei Prozentpunkte höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. In den folgenden Wellen sind die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  und  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Vier-Wellen-Paneldatensatz  $P_{t_1,3,5,6}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Mit weit unter einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel höchstens minimal unterrepräsentiert. Überdies zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Vier-Wellen-Paneldatensatz mit viereinhalb Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Vier-Wellen-Paneldatensatz enthalten.

Tabelle 5.104 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Vier-Wellen-Paneldatensatz  $P_{t_1,3,5,6}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

**Tabelle 5.103:** Zusammensetzung  $P_{t_1,3,5,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,3	18,1	—	1,3	—	—	2003	
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728	
		$P_{t_1,3,5,6}$	18,2	20,1	32,4	29,3	—	—	—	—	577	
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635	
		$P_{t_3,3,5,6}$	18,9	21,3	31,9	27,9	—	—	—	—	577	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{t_5,3,5,6}$	1,2	1,2	21,0	30,1	42,5	—	2,1	1,8	562	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504	
		$P_{t_6,3,5,6}$	—	—	18,2	28,0	32,7	—	10,8	10,3	565	
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—	1907
			$Q_{t_1}$	17,9	24,1	31,8	26,3	—	0,0	—	—	1679
$P_{t_1,3,5,6}$			12,9	25,9	30,3	31,0	—	—	—	—	878	
$t_3$		$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926	
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704	
		$P_{t_3,3,5,6}$	13,6	26,3	30,6	29,5	—	—	—	—	878	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{t_5,3,5,6}$	0,1	0,2	19,8	35,5	40,8	—	1,8	1,8	844	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715	
		$P_{t_6,3,5,6}$	—	—	17,5	32,2	29,8	—	11,3	9,2	851	
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
			$Q_{t_1}$	22,0	23,7	31,1	22,8	—	0,4	—	—	3407
	$P_{t_1,3,5,6}$		15,0	23,6	31,1	30,3	—	—	—	—	1455	
	$t_3$	$S_{t_3}$	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	
		$P_{t_3,3,5,6}$	15,7	24,3	31,1	28,9	—	—	—	—	1455	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{t_5,3,5,6}$	0,6	0,6	20,3	33,4	41,5	—	1,9	1,8	1406	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
		$P_{t_6,3,5,6}$	—	—	17,8	30,5	31,0	—	11,1	9,6	1416	

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

**Tabelle 5.104:** Zusammensetzung  $P_{1,3,5,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2	
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7	
		$P_{t_{1,3,5,6}}$	48,2	33,8	41,3	38,3	—	—	—	—	39,7	
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9	
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0	
		$P_{t_{1,3,5,6}}$	47,8	34,7	40,6	38,3	—	—	—	—	39,7	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{t_{1,3,5,6}}$	87,5	77,8	41,4	36,0	41,0	—	44,4	40,0	40,0	
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7	
		$P_{t_{1,3,5,6}}$	—	—	40,9	36,6	42,1	—	38,9	42,6	39,9	
	weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	56,4	—	12,9	—	—	48,8
			$Q_{t_1}$	40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
$P_{t_{1,3,5,6}}$			51,8	66,2	58,7	61,7	—	—	—	—	60,3	
$t_3$		$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1	
		$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0	
		$P_{t_{1,3,5,6}}$	52,2	65,3	59,4	61,7	—	—	—	—	60,3	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{t_{1,3,5,6}}$	12,5	22,2	58,6	64,0	59,0	—	55,6	60,0	60,0	
$t_6$		$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3	
		$P_{t_{1,3,5,6}}$	—	—	59,1	63,4	57,9	—	61,1	57,4	60,1	

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



5.4.9 Vier-Wellen-Panel 2003, 2004, 2005 und 2007 ( $P_{t_2,3,4,6}$ )

Das Vier-Wellen-Panel der Jahre 2003, 2004, 2005 und 2007 enthält 1 733 Fälle. Aus Tabelle 5.105 geht hervor, dass im Vier-Wellen-Paneldatensatz  $P_{t_2,3,4,6}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.105:** Zusammensetzung nach Geschlecht  $P_{t_2,3,4,6}$

		männlich		weiblich		gesamt Hfk
		Hfk	%	Hfk	%	(100 %)
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3 763
	$Q_{t_2}$	1703	50,2	1687	49,8	3 390
	$P_{t_2,3,4,6}$	709	40,9	1024	59,1	1 733
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3 925
	$Q_{t_3}$	1635	49,0	1704	51,0	3 339
	$P_{t_2,3,4,6}$	700	40,4	1033	59,6	1 733
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_2,3,4,6}$	711	41,0	1022	59,0	1 733
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3 335
	$P_{t_2,3,4,6}$	704	40,6	1029	59,4	1 733

Vergleicht man die Geschlechterverteilung der Erhebungszeitpunkte  $t_2$  und  $t_3$  im Vier-Wellen-Paneldatensatz mit den Schulstatistiken der Erhebungswellen  $t_2$  und  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz in jedem der Jahre um etwa zehneinhalb Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatsätzen  $Q_{t_2}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Vier-Wellen-Paneldatensatz um knapp neuneinhalb bzw. etwa achteinhalb Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Vier-Wellen-Paneldatensatz mit den Querschnittsdatsätzen der Erhebungszeitpunkte  $t_4$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz zu  $t_4$  zunächst knapp neuneinhalb, zu  $t_6$  knapp sechseinhalb Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatsatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Paneldatsätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatsätzen und dem Vier-Wellen-Paneldatensatz (siehe Tabelle 5.106) zeigt, dass das Durchschnittsalter der Befragten im Paneldatsatz durchgehend geringfügig unter dem der einzelnen Quer-

**Tabelle 5.106:** Altersdurchschnitt  $P_{t_2,3,4,6}$ 

		Alter	Std.-abw.	gesamt Hfk
$t_2$	$Q_{t_2}$	14,0	0,8	3 275
	$P_{t_2,3,4,6}$	13,9	0,7	1 709
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_2,3,4,6}$	15,0	0,7	1 718
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_2,3,4,6}$	15,9	0,7	1 713
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_2,3,4,6}$	17,9	0,7	1 728

schnittsdatensätze liegt. Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.107 zeigt, dass der Vier-Wellen-Paneldatensatz  $P_{t_2,3,4,6}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Vier-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_2$  und  $t_3$  neuneinhalb bzw. sechseinhalb Prozentpunkte unter denen der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von etwa viereinhalb Prozentpunkten die Hauptschülerinnen und -schüler im Vier-Wellen-Paneldatensatz ebenfalls unterrepräsentiert sind, sind zu  $t_6$  keine gesicherten Aussagen mehr möglich: Kein Befragter besucht zu diesem Befragungszeitpunkt mehr die Hauptschule. Weiterhin sind im Vier-Wellen-Paneldatensatz zu  $t_2$  die Gesamtschülerinnen und -schüler mit zweieinhalb Prozentpunkten überrepräsentiert; zu  $t_3$  sind bei einer sehr schwachen Überrepräsentation von unter einem Prozentpunkt ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik. Auch zu  $t_4$  und  $t_6$  unterscheiden sich die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatensätzen  $Q_{t_4}$  und  $Q_{t_6}$  mit weit weniger als einem Prozentpunkt – allerdings in abweichende Richtungen.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_6$  getroffen werden. Vergleichsgrundlage ist der Querschnittsdatensatz des betreffenden Erhebungsjahres. So sind zu  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Vier-Wellen-Paneldatensatz  $P_{t_2,3,4,6}$  ähnlich, mit genau einem Prozentpunkt sind Befragte, die ein Berufskolleg besuchen nur leicht im Vier-Wellen-Paneldatensatz überrepräsentiert.

Im Gegensatz dazu liegen die Anteile der Gymnasiastinnen und Gymnasiasten im Vier-Wellen-Paneldatensatz zu den Zeitpunkten  $t_2$  und  $t_3$  etwa siebeneinhalb bzw. fünfeinhalb Prozentpunkte über denen der jeweiligen Schulstatistik. Im Anschluss wird diese Überrepräsentation mit dreieinhalb Prozentpunkten zu  $t_4$  und knapp über einem Prozentpunkt zu  $t_6$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – geringer. Der Anteil der Realschülerinnen und -schüler liegt im Vier-Wellen-Paneldatensatz im Zeitraum  $t_2$  und  $t_3$  weit unter einem Prozentpunkt bzw. knapp zwei Prozentpunkte höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit knapp eineinhalb Prozentpunkten überrepräsentiert sind, sind in der folgenden Welle die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Vier-Wellen-Paneldatensatz  $P_{t_{2,3,4,6}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Während im vorliegenden Vier-Wellen-Paneldatensatz bis zur Erhebungswelle  $t_6$  keine Befragten enthalten sind, die keine Schule besuchen, zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Panel mit etwas über drei Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Vier-Wellen-Paneldatensatz enthalten.

Tabelle 5.108 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Vier-Wellen-Paneldatensatz  $P_{t_{2,3,5,6}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

## 5.4 Die Vier-Wellen-Panel

**Tabelle 5.107:** Zusammensetzung  $P_{t_2,3,4,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927	
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703	
		$P_{t_2,3,4,6}$	22,6	19,6	34,6	23,3	—	—	—	—	709	
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635	
		$P_{t_2,3,4,6}$	22,6	20,3	34,6	22,6	—	—	—	—	700	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{t_2,3,4,6}$	21,6	20,3	34,4	22,4	0,7	—	0,6	—	709	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504	
		$P_{t_2,3,4,6}$	—	—	17,7	23,3	34,3	—	12,4	12,3	677	
	weiblich	$t_2$	$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836
			$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1687
$P_{t_2,3,4,6}$			14,8	25,2	31,8	28,1	—	—	—	—	1024	
$t_3$		$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926	
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704	
		$P_{t_2,3,4,6}$	15,4	25,3	32,0	27,3	—	—	—	—	1033	
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{t_2,3,4,6}$	14,5	25,1	31,9	27,7	0,6	—	0,2	—	1022	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715	
		$P_{t_2,3,4,6}$	—	—	18,0	29,9	30,5	—	11,0	10,6	1002	
gesamt		$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
			$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390
	$P_{t_2,3,4,6}$		18,0	22,9	32,9	26,1	—	—	—	—	1733	
	$t_3$	$S_{t_3}$	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	
		$P_{t_2,3,4,6}$	18,3	23,3	33,1	25,4	—	—	—	—	1733	
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{t_2,3,4,6}$	17,4	23,2	32,9	25,5	0,6	—	0,3	—	1731	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
		$P_{t_2,3,4,6}$	—	—	17,9	27,3	32,0	—	11,6	11,3	1679	

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

**Tabelle 5.108:** Zusammensetzung  $P_{t_2,3,4,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2	
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2	
		$P_{t_2,3,4,6}$	51,3	35,0	42,9	36,4	—	—	—	—	40,9	
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9	
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0	
		$P_{t_2,3,4,6}$	49,8	35,2	42,2	35,9	—	—	—	—	40,4	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_2,3,4,6}$	50,8	35,9	42,8	36,0	45,5	—	66,7	—	41,0	
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7	
		$P_{t_2,3,4,6}$	—	—	40,0	34,5	43,1	—	43,3	43,9	40,3	
	weiblich	$t_2$	$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8
			$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8
$P_{t_2,3,4,6}$			48,7	65,0	57,1	63,6	—	—	—	—	59,1	
$t_3$		$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1	
		$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0	
		$P_{t_2,3,4,6}$	50,2	64,8	57,8	64,1	—	—	—	—	59,6	
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_2,3,4,6}$	49,2	64,1	57,2	64,0	54,5	—	33,3	—	59,0	
$t_6$		$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3	
		$P_{t_2,3,4,6}$	—	—	60,0	65,5	56,9	—	56,7	56,1	59,7	

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

5.4.10 Vier-Wellen-Panel 2003, 2004, 2006 und 2007 ( $P_{t_2,3,5,6}$ )

Das Vier-Wellen-Panel der Jahre 2003, 2004, 2006 und 2007 enthält 1 675 Fälle. Aus Tabelle 5.109 geht hervor, dass im Vier-Wellen-Panel Datensatz  $P_{t_2,3,5,6}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.109:** Zusammensetzung nach Geschlecht  $P_{t_2,3,5,6}$

		männlich		weiblich		gesamt Hfk (100 %)
		Hfk	%	Hfk	%	
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{t_2,3,5,6}$	669	39,9	1006	60,1	1675
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_2,3,5,6}$	660	39,4	1015	60,6	1675
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_2,3,5,6}$	667	39,8	1008	60,2	1675
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3335
	$P_{t_2,3,5,6}$	667	39,8	1008	60,2	1675

Vergleicht man die Geschlechterverteilung der Erhebungszeitpunkte  $t_2$  und  $t_3$  im Vier-Wellen-Panel Datensatz mit den Schulstatistiken der Erhebungswellen  $t_2$  und  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Vier-Wellen-Panel Datensatz in jedem der Jahre um etwa elfeinhalb Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatsätzen  $Q_{t_2}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Vier-Wellen-Panel Datensatz etwa zehneinhalb bzw. neuneinhalb Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Vier-Wellen-Panel Datensatz mit den Querschnittsdatsätzen der Erhebungszeitpunkte  $t_5$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Vier-Wellen-Panel Datensatz zu  $t_5$  zunächst zehn, zu  $t_6$  etwas über sieben Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatsatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Panel Datensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatsätzen und dem Vier-Wellen-Panel Datensatz (siehe Tabelle 5.110) zeigt, dass das Durchschnittsalter der Befragten im Panel Datensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatsätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhe-

**Tabelle 5.110:** Altersdurchschnitt  $P_{2,3,5,6}$

		Alter	Std.-abw.	gesamt Hfk
$t_2$	$Q_{t_2}$	14,0	0,8	3 275
	$P_{2,3,5,6}$	13,9	0,7	1 653
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{2,3,5,6}$	15,0	0,7	1 661
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{2,3,5,6}$	17,0	0,8	1 663
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{2,3,5,6}$	17,9	0,7	1 670

bungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.111 zeigt, dass der Vier-Wellen-Paneldatensatz  $P_{2,3,5,6}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Vier-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_2$  und  $t_3$  etwas über zehn bzw. siebeneinhalb Prozentpunkte unter denen der Schulstatistik. Zu  $t_5$  und  $t_6$  sind keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte bzw. keine Befragten besuchen zu diesen Befragungszeitpunkten noch die Hauptschule. Weiterhin sind im Vier-Wellen-Paneldatensatz zu  $t_2$  sind die Gesamtschülerinnen und -schüler mit etwas über zwei Prozentpunkten überrepräsentiert; zu  $t_3$  sind ähnlich viele Gesamtschülerinnen und -schüler im Panel vertreten wie in der Schulstatistik. Während sich zu  $t_6$  die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatensätzen  $Q_{t_6}$  ebenfalls mit weniger als einem Prozentpunkt unterscheiden, sind die Gesamtschüler zu  $t_5$  im Vier-Wellen-Paneldatensatz mit einem Unterschied von knapp über fünf Prozentpunkten im Gegensatz zu den Querschnittsdaten deutlicher überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Gesicherte Aussagen über die Anteile derjenigen Befragten, die ein Berufskolleg besuchen, können erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage sind stets die Querschnittsdatensätze der betreffenden Erhebungsjahre. So fällt zu  $t_5$  mit über siebzehneinhalb Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3). Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Vier-Wellen-Paneldatensatz

$P_{t_{2,3,5,6}}$  ähnlich, mit knapp eineinhalb Prozentpunkten sind Befragte, die ein Berufskolleg besuchen nur leicht im Vier-Wellen-Paneldatensatz überrepräsentiert.

Die Anteile der Gymnasiastinnen und Gymnasiasten liegen im Vier-Wellen-Paneldatensatz zu den Zeitpunkten  $t_2$  und  $t_3$  etwa achteinhalb bzw. sechseinhalb Prozentpunkte über denen der jeweiligen Schulstatistik. Im Anschluss steigt diese Überrepräsentation zu  $t_5$  auf über elf Prozentpunkte an (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Zu  $t_6$  sind die Gymnasiastinnen und Gymnasiasten verglichen mit dem Querschnittsdatensatz mit knapp über zwei Prozentpunkten wiederum in geringerem Ausmaß überrepräsentiert. Der Anteil der Realschülerinnen und -schüler liegt im Vier-Wellen-Paneldatensatz im Zeitraum  $t_2$  und  $t_3$  einen halben Prozentpunkt bzw. zwei Prozentpunkte höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. In den folgenden Wellen sind die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  und  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit weit unter einem Prozentpunkt im Vier-Wellen-Paneldatensatz  $P_{t_{2,3,5,6}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – allenfalls minimal überrepräsentiert. Mit knapp unter einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel ebenfalls sehr leicht überrepräsentiert. Dagegen zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Vier-Wellen-Paneldatensatz mit knapp vier Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Vier-Wellen-Paneldatensatz enthalten.

Tabelle 5.112 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Vier-Wellen-Paneldatensatz  $P_{t_{2,3,5,6}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).



**Tabelle 5.111:** Zusammensetzung  $P_{t_2,3,5,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1 927	
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1 703	
		$P_{t_2,3,5,6}$	21,5	19,6	34,4	24,5	—	—	—	—	669	
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1 999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1 635	
		$P_{t_2,3,5,6}$	21,5	20,2	34,5	23,8	—	—	—	—	660	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2 181	
		$P_{t_2,3,5,6}$	1,4	0,9	21,7	26,6	44,4	—	2,0	2,9	646	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1 504	
		$P_{t_2,3,5,6}$	—	—	18,9	24,5	34,1	—	11,5	11,0	645	
	weiblich	$t_2$	$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1 836
			$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1 687
$P_{t_2,3,5,6}$			14,4	25,5	31,3	28,7	—	—	—	—	1 006	
$t_3$		$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1 926	
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1 704	
		$P_{t_2,3,5,6}$	14,9	25,7	31,5	27,9	—	—	—	—	1 015	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2 226	
		$P_{t_2,3,5,6}$	0,2	0,4	20,3	33,6	41,9	—	1,8	1,8	977	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1 715	
		$P_{t_2,3,5,6}$	—	—	17,9	30,8	31,2	—	10,5	9,7	984	
gesamt		$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3 763
			$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3 390
	$P_{t_2,3,5,6}$		17,3	23,2	32,5	27,0	—	—	—	—	1 675	
	$t_3$	$S_{t_3}$	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	
		$P_{t_2,3,5,6}$	17,5	23,5	32,7	26,3	—	—	—	—	1 675	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4 407	
		$P_{t_2,3,5,6}$	0,7	0,6	20,8	30,8	42,9	—	1,9	2,3	1 623	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3 219	
		$P_{t_2,3,5,6}$	—	—	18,3	28,3	32,4	—	10,9	10,2	1 629	

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

**Tabelle 5.112:** Zusammensetzung  $P_{2,3,5,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2	
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2	
		$P_{t_2,3,5,6}$	49,8	33,8	42,2	36,2	—	—	—	—	39,9	
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9	
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0	
		$P_{t_2,3,5,6}$	48,5	33,8	41,6	35,7	—	—	—	—	39,4	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{t_2,3,5,6}$	81,8	60,0	41,4	34,4	41,2	—	41,9	51,4	39,8	
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7	
		$P_{t_2,3,5,6}$	—	—	40,9	34,3	41,7	—	41,8	42,8	39,6	
	weiblich	$t_2$	$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8
			$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8
$P_{t_2,3,5,6}$			50,2	66,2	57,8	63,8	—	—	—	—	60,1	
$t_3$		$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1	
		$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0	
		$P_{t_2,3,5,6}$	51,5	66,2	58,4	64,3	—	—	—	—	60,6	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{t_2,3,5,6}$	18,2	40,0	58,6	65,6	58,8	—	58,1	48,6	60,2	
$t_6$		$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3	
		$P_{t_2,3,5,6}$	—	—	59,1	65,7	58,3	—	58,2	57,2	60,4	

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

5.4.11 Vier-Wellen-Panel 2003, 2005, 2006 und 2007 ( $P_{t_{2,4,5,6}}$ )

Das Vier-Wellen-Panel der Jahre 2003, 2005, 2006 und 2007 enthält 1 739 Fälle. Aus Tabelle 5.113 geht hervor, dass im Vier-Wellen-Paneldatensatz  $P_{t_{2,4,5,6}}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.113:** Zusammensetzung nach Geschlecht  $P_{t_{2,4,5,6}}$

		männlich		weiblich		gesamt Hfk (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
	$P_{t_{2,4,5,6}}$	703	40,4	1036	59,6	1 739
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3 405
	$P_{t_{2,4,5,6}}$	703	40,4	1036	59,6	1 739
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_{2,4,5,6}}$	699	40,2	1040	59,8	1 739
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3 335
	$P_{t_{2,4,5,6}}$	698	40,1	1041	59,9	1 739

Vergleicht man die Geschlechterverteilung zum Erhebungszeitpunkt  $t_2$  im Vier-Wellen-Paneldatensatz mit der Schulstatistik der Erhebungswelle  $t_2$ , so wird deutlich, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz um knapp elf Prozentpunkte höher liegt. Im Vergleich zum entsprechenden Querschnittsdatsatz  $Q_{t_2}$  ist der Anteil der weiblichen Befragten im Vier-Wellen-Paneldatensatz um knapp zehn Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Vier-Wellen-Paneldatensatz mit den Querschnittsdatsätzen der Erhebungszeitpunkte  $t_4$ ,  $t_5$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz zunächst etwa zehn ( $t_4$  und  $t_5$ ), zu  $t_6$  knapp sieben Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatsatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Paneldatsätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatsätzen und dem Vier-Wellen-Paneldatsatz (siehe Tabelle 5.114) zeigt, dass das Durchschnittsalter der Befragten im Paneldatsatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatsätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass

**Tabelle 5.114:** Altersdurchschnitt  $P_{t_2,4,5,6}$

		Alter	Std.-abw.	gesamt Hfk
$t_2$	$Q_{t_2}$	14,0	0,8	3275
	$P_{t_2,4,5,6}$	13,9	0,7	1715
$t_4$	$Q_{t_4}$	16,1	0,8	3313
	$P_{t_2,4,5,6}$	15,9	0,7	1718
$t_5$	$Q_{t_5}$	17,4	1,3	4468
	$P_{t_2,4,5,6}$	17,0	0,8	1728
$t_6$	$Q_{t_6}$	18,1	1,1	3320
	$P_{t_2,4,5,6}$	17,9	0,7	1733

die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.115 zeigt, dass der Vier-Wellen-Paneldatensatz  $P_{t_2,4,5,6}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Der Anteil der Hauptschülerinnen und -schüler liegt im Vier-Wellen-Paneldatensatz zum Erhebungszeitpunkt  $t_2$  über neun Prozentpunkte unter dem der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von knapp viereinhalb Prozentpunkten die Hauptschülerinnen und -schüler im Vier-Wellen-Paneldatensatz ebenfalls unterrepräsentiert sind, sind zu  $t_5$  und  $t_6$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte bzw. keine Befragten besuchen zu diesen Befragungszeitpunkten noch die Hauptschule. Weiterhin sind im Vier-Wellen-Paneldatensatz zu  $t_2$  die Gesamtschülerinnen und -schüler mit etwas über eineinhalb Prozentpunkten im Panel leicht überrepräsentiert. Während sich zu  $t_4$  und  $t_6$  die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatensätzen  $Q_{t_4}$  und  $Q_{t_6}$  mit weit unter einem Prozentpunkt – allerdings in abweichende Richtungen – sehr leicht unterscheiden, sind die Gesamtschüler zu  $t_5$  im Vier-Wellen-Paneldatensatz mit einem Unterschied von über fünf Prozentpunkten im Gegensatz zu den Querschnittsdaten deutlich überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage sind stets die Querschnittsdatensätze der betreffenden Erhebungsjahre. So fällt zu  $t_5$  mit über siebzehneinhalb Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten im Panel auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschul-

det (siehe genauer Abschnitt 5.1.3). Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Vier-Wellen-Paneldatensatz  $P_{t_{2,4,5,6}}$  ähnlich, mit knapp über einem Prozentpunkt sind Befragte, die ein Berufskolleg besuchen nur leicht im Vier-Wellen-Paneldatensatz überrepräsentiert.

Der Anteil der Gymnasiastinnen und Gymnasiasten liegt im Vier-Wellen-Paneldatensatz zum Zeitpunkt  $t_2$  acht Prozentpunkte über dem der jeweiligen Schulstatistik. Im Anschluss wird diese Überrepräsentation mit etwas über vier Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – zunächst geringer, um zu  $t_5$  auf über elf Prozentpunkte anzusteigen (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Zu  $t_6$  sind die Gymnasiastinnen und Gymnasiasten verglichen mit dem Querschnittsdatensatz mit etwas mehr als zwei Prozentpunkten wiederum in geringerem Ausmaß überrepräsentiert. Der Anteil der Realschülerinnen und -schüler liegt im Vier-Wellen-Paneldatensatz im Zeitraum  $t_2$  knapp einen halben Prozentpunkt höher als in der betreffenden Schulstatistik dieses Erhebungsjahres. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit etwas über einem Prozentpunkt überrepräsentiert sind, sind in den folgenden Wellen die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  und  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Vier-Wellen-Paneldatensatz  $P_{t_{2,4,5,6}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – leicht überrepräsentiert. Während zur Erhebungswelle  $t_5$  der Anteil derjenigen, die keine Schule besuchen, im Panel- und im Querschnittsdatensatz identisch ist, zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Vier-Wellen-Paneldatensatz mit über dreieinhalb Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Vier-Wellen-Paneldatensatz enthalten.

Tabelle 5.116 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Vier-Wellen-Paneldatensatz  $P_{t_{2,4,5,6}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

**Tabelle 5.115:** Zusammensetzung  $P_{t_2,4,5,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927	
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703	
		$P_{t_2,4,5,6}$	22,2	19,9	33,4	24,5	—	—	—	—	703	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{t_2,4,5,6}$	21,3	20,1	33,8	23,8	0,6	—	0,4	—	701	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{t_2,4,5,6}$	1,3	0,7	21,7	26,9	44,8	—	2,1	2,5	674	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504	
		$P_{t_2,4,5,6}$	—	—	18,5	24,6	34,2	—	11,4	11,4	676	
	weiblich	$t_2$	$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836
			$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1687
			$P_{t_2,4,5,6}$	15,4	25,2	31,2	28,2	—	—	—	—	1036
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{t_2,4,5,6}$	14,8	25,0	31,8	28,0	0,4	—	0,1	—	1036	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{t_2,4,5,6}$	0,3	0,4	20,5	33,6	41,6	—	2,0	1,7	1010	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715	
		$P_{t_2,4,5,6}$	—	—	18,2	30,7	30,7	—	10,6	9,7	1016	
gesamt		$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
			$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390
			$P_{t_2,4,5,6}$	18,2	23,1	32,1	26,7	—	—	—	—	1739
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{t_2,4,5,6}$	17,4	23,0	32,6	26,3	0,5	—	0,2	—	1737	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{t_2,4,5,6}$	0,7	0,5	21,0	30,9	42,9	—	2,0	2,0	1684	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
		$P_{t_2,4,5,6}$	—	—	18,3	28,3	32,1	—	10,9	10,4	1692	

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

**Tabelle 5.116:** Zusammensetzung  $P_{t_2,4,5,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2	
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2	
		$P_{t_2,4,5,6}$	49,4	34,9	42,1	37,1	—	—	—	—	40,4	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_2,4,5,6}$	49,3	35,3	41,9	36,5	50,0	—	75,0	—	40,4	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{t_2,4,5,6}$	75,0	55,6	41,4	34,8	41,8	—	41,2	50,0	40,0	
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7	
		$P_{t_2,4,5,6}$	—	—	40,3	34,7	42,5	—	41,6	43,8	40,0	
	weiblich	$t_2$	$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8
			$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8
			$P_{t_2,4,5,6}$	50,6	65,1	57,9	62,9	—	—	—	—	59,6
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_2,4,5,6}$	50,7	64,8	58,1	63,5	50,0	—	25,0	—	59,6	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{t_2,4,5,6}$	25,0	44,4	58,6	65,2	58,2	—	58,8	50,0	60,0	
$t_6$		$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3	
		$P_{t_2,4,5,6}$	—	—	59,7	65,3	57,5	—	58,4	56,3	60,0	

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

5.4.12 Vier-Wellen-Panel 2002, 2005, 2006 und 2007 ( $P_{t_1,4,5,6}$ )

Das Vier-Wellen-Panel der Jahre 2002, 2005, 2006 und 2007 enthält 1 550 Fälle. Aus Tabelle 5.117 geht hervor, dass im Vier-Wellen-Paneldatensatz  $P_{t_1,4,5,6}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.117:** Zusammensetzung nach Geschlecht  $P_{t_1,4,5,6}$

		männlich		weiblich		gesamt Hfk (100 %)
		Hfk	%	Hfk	%	
$t_1$	$S_{t_1}$	2003	51,2	1907	48,8	3910
	$Q_{t_1}$	1728	50,7	1679	49,3	3407
	$P_{t_1,4,5,6}$	635	41,0	915	59,0	1550
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_1,4,5,6}$	640	41,3	910	58,7	1550
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_1,4,5,6}$	639	41,2	911	58,8	1550
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3335
	$P_{t_1,4,5,6}$	638	41,2	912	58,8	1550

Vergleicht man die Geschlechterverteilung der ersten Welle  $t_1$  im Vier-Wellen-Paneldatensatz mit der Schulstatistik der Erhebungswelle  $t_1$ , so wird deutlich, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz in diesem Jahr mehr als zehn Prozentpunkte höher liegt. Im Vergleich zum entsprechenden Querschnittsdatensatz  $Q_{t_1}$  ist der Anteil der weiblichen Befragten im Vier-Wellen-Paneldatensatz über neun-einhalb Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Vier-Wellen-Paneldatensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$ ,  $t_5$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Vier-Wellen-Paneldatensatz zunächst etwa neun ( $t_4$  und  $t_5$ ), zu  $t_6$  knapp sechs Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatensatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Vier-Wellen-Paneldatensatz (siehe Tabelle 5.118) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass



**Tabelle 5.118:** Altersdurchschnitt  $P_{1,4,5,6}$

		Alter	Std.-abw.	gesamt Hfk
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{1,4,5,6}$	12,9	0,7	1 536
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{1,4,5,6}$	15,9	0,7	1 535
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{1,4,5,6}$	16,9	0,7	1 541
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{1,4,5,6}$	17,9	0,7	1 544

die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.119 zeigt, dass der Vier-Wellen-Paneldatensatz  $P_{1,4,5,6}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Der Anteil der Hauptschülerinnen und -schüler liegt im Vier-Wellen-Paneldatensatz zum Erhebungszeitpunkt  $t_1$  über acht Prozentpunkte unter dem der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von etwas über sechs Prozentpunkten die Hauptschülerinnen und -schüler im Vier-Wellen-Paneldatensatz ebenfalls deutlich unterrepräsentiert sind, sind zu  $t_5$  und  $t_6$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte bzw. keine Befragten besuchen zu diesen Befragungszeitpunkten noch die Hauptschule. Weiterhin sind im Vier-Wellen-Paneldatensatz zu  $t_1$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik. Während sich zu  $t_4$  und  $t_6$  die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatensätzen  $Q_{t_4}$  und  $Q_{t_6}$  mit etwa eineinhalb Prozentpunkten bzw. gar nicht unterscheiden, sind die Gesamtschüler zu  $t_5$  im Vier-Wellen-Paneldatensatz mit einem Unterschied von fünf Prozentpunkten im Gegensatz zu den Querschnittsdaten deutlich überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage sind stets die Querschnittsdatensätze der betreffenden Erhebungsjahre. So fällt zu  $t_5$  mit über zwanzig Prozentpunkten im Panel eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3). Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegia-

tinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Vier-Wellen-Paneldatensatz  $P_{t_{1,4,5,6}}$  ähnlich, mit weit unter einem Prozentpunkt sind Befragte, die ein Berufskolleg besuchen nur leicht im Vier-Wellen-Paneldatensatz unterrepräsentiert.

Der Anteil der Gymnasiastinnen und Gymnasiasten liegt im Vier-Wellen-Paneldatensatz zum Zeitpunkt  $t_1$  knapp neun Prozentpunkte über dem der zugehörigen Schulstatistik. Im Anschluss wird diese Überrepräsentation mit knapp sieben Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – zunächst etwas geringer, um zu  $t_5$  auf über vierzehn Prozentpunkte anzusteigen (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Zu  $t_6$  sind die Gymnasiastinnen und Gymnasiasten verglichen mit dem Querschnittsdatensatz mit etwas über viereinhalb Prozentpunkten wiederum in geringerem Ausmaß überrepräsentiert. Der Anteil der Realschülerinnen und -schüler liegt im Vier-Wellen-Paneldatensatz im Zeitraum  $t_1$  gut einen Prozentpunkt höher als in der Schulstatistik dieses Erhebungsjahres. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit knapp zwei Prozentpunkten überrepräsentiert sind, sind in den folgenden Wellen die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  und  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Vier-Wellen-Paneldatensatz  $P_{t_{1,4,5,6}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – leicht überrepräsentiert. Mit etwas unter einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel sehr leicht unterrepräsentiert. Überdies zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Vier-Wellen-Paneldatensatz mit knapp über viereinhalb Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind bis zum Zeitpunkt  $t_5$  im Vier-Wellen-Paneldatensatz enthalten, jedoch sind ihre Anteile zu gering, um Aussagen zur Repräsentativität zuzulassen.

Tabelle 5.120 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Vier-Wellen-Paneldatensatz  $P_{t_{1,4,5,6}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

**Tabelle 5.119:** Zusammensetzung  $P_{t_1,4,5,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,3	18,1	—	1,3	—	—	2003	
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728	
		$P_{t_1,4,5,6}$	19,4	19,7	31,3	29,1	—	0,5	—	—	635	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{t_1,4,5,6}$	18,7	20,8	31,8	27,4	0,5	0,5	0,3	—	638	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{t_1,4,5,6}$	1,1	1,0	21,7	30,8	40,9	0,3	2,3	1,8	613	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504	
		$P_{t_1,4,5,6}$	—	—	18,2	28,4	32,6	—	10,6	10,2	620	
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—	1907
			$Q_{t_1}$	17,9	24,1	31,8	26,3	—	0,0	—	—	1679
			$P_{t_1,4,5,6}$	13,2	25,7	30,4	30,7	—	0,0	—	—	915
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{t_1,4,5,6}$	13,1	25,7	30,5	30,0	0,4	0,0	0,2	—	910	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{t_1,4,5,6}$	0,2	0,2	20,1	35,9	40,0	0,0	1,9	1,6	882	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715	
		$P_{t_1,4,5,6}$	—	—	17,8	32,5	29,3	—	11,3	9,1	888	
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
			$Q_{t_1}$	22,0	23,7	31,1	22,8	—	0,4	—	—	3407
			$P_{t_1,4,5,6}$	15,7	23,2	30,8	30,1	—	0,2	—	—	1550
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{t_1,4,5,6}$	15,4	23,7	31,1	28,9	0,5	0,2	0,2	—	1548	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{t_1,4,5,6}$	0,6	0,5	20,7	33,8	40,4	0,1	2,1	1,7	1495	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
		$P_{t_1,4,5,6}$	—	—	18,0	30,8	30,6	—	11,0	9,5	1508	

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

**Tabelle 5.120:** Zusammensetzung  $P_{t_1,4,5,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2	
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7	
		$P_{t_1,4,5,6}$	50,4	34,7	41,7	39,7	—	100,0	—	—	41,0	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_1,4,5,6}$	50,0	36,2	42,2	39,1	42,9	100,0	50,0	—	41,2	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{t_1,4,5,6}$	77,8	75,0	42,9	37,4	41,6	100,0	45,2	44,0	41,0	
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7	
		$P_{t_1,4,5,6}$	—	—	41,7	37,8	43,7	—	39,8	43,8	41,1	
	weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	56,4	—	12,9	—	—	48,8
			$Q_{t_1}$	40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
			$P_{t_1,4,5,6}$	49,6	65,3	58,3	60,3	—	0,0	—	—	59,0
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_1,4,5,6}$	50,0	63,8	57,8	60,9	57,1	0,0	50,0	—	58,8	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{t_1,4,5,6}$	22,2	25,0	57,1	62,6	58,4	0,0	54,8	56,0	59,0	
$t_6$		$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3	
		$P_{t_1,4,5,6}$	—	—	58,3	62,2	56,3	—	60,2	56,3	58,9	

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

## 5.5 Die Fünf-Wellen-Panel

Die Datensätze der Fünf-Wellen-Panel enthalten alle Fälle, deren zugehörige – und nach Plausibilitätskontrollen gültige – Fragebögen mindestens vier (ebenfalls gültigen) Bögen weiterer Zeitpunkte zugeordnet werden konnten. In der Praxis werden vermutlich die beiden Datensätze, die fünf direkt aufeinander folgende Erhebungszeitpunkte umfassen, am bedeutsamsten sein. Diese werden zuerst dargestellt, im Anschluss werden die restlichen in der Tabelle 4.1 fett gedruckten Fünf-Wellen-Paneldatensätze einzeln beschrieben.

### 5.5.1 Fünf-Wellen-Panel 2002 bis 2006 ( $P_{t_{1,2,3,4,5}}$ )

Das Fünf-Wellen-Panel der Jahre 2002 bis 2006 enthält 1 552 Fälle. Aus Tabelle 5.121 geht hervor, dass im Fünf-Wellen-Paneldatensatz  $P_{t_{1,2,3,4,5}}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.121:** Zusammensetzung nach Geschlecht  $P_{t_{1,2,3,4,5}}$

		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
	$P_{t_{1,2,3,4,5}}$	642	41,4	910	58,6	1552
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{t_{1,2,3,4,5}}$	651	41,9	901	58,1	1552
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_{1,2,3,4,5}}$	644	41,5	908	58,5	1552
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_{1,2,3,4,5}}$	647	41,7	905	58,3	1552
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_{1,2,3,4,5}}$	648	41,8	904	58,2	1552

Vergleicht man die Geschlechterverteilung der ersten drei Wellen im Fünf-Wellen-Paneldatensatz mit den Schulstatistiken der Erhebungswellen  $t_1$  bis  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Fünf-Wellen-Paneldatensatz  $P_{t_{1,2,3,4,5}}$  in jedem der Jahre um etwa zehn Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatsätzen  $Q_{t_1}$ ,  $Q_{t_2}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Fünf-Wellen-Paneldatensatz zwischen siebeneinhalb und etwa neun Prozentpunkten erhöht. Die Gegenüberstellung der Geschlechterverteilung im Fünf-Wellen-Paneldatensatz mit den Querschnittsdatsätzen der Erhebungszeitpunkte  $t_4$  und  $t_5$  (ein Vergleich mit der

Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Fünf-Wellen-Paneldatensatz zu  $t_4$  acht bzw. zu  $t_5$  etwas über achteinhalb ( $t_5$ )Prozentpunkte höher liegt.

**Tabelle 5.122:** Altersdurchschnitt  $P_{1,2,3,4,5}$

		Alter	Std.-abw.	gesamt
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{1,2,3,4,5}$	12,9	0,7	1 537
$t_2$	$Q_{t_2}$	14,0	0,8	3 275
	$P_{1,2,3,4,5}$	13,9	0,7	1 534
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{1,2,3,4,5}$	14,9	0,7	1 543
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{1,2,3,4,5}$	15,9	0,7	1 534
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{1,2,3,4,5}$	17,0	0,8	1 542

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Fünf-Wellen-Paneldatensatz (siehe Tabelle 5.122) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.123 zeigt, dass der Fünf-Wellen-Paneldatensatz  $P_{1,2,3,4,5}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Fünf-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_1$  bis  $t_3$  zwischen sechseinhalb und über neuneinhalb Prozentpunkten unter denen der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von etwa vier Prozentpunkten die Hauptschülerinnen und -schüler im Fünf-Wellen-Paneldatensatz ebenfalls deutlich unterrepräsentiert sind, ist zu  $t_5$  keine gesicherten Aussage mehr möglich: Nur sehr wenige Befragte besuchen zu diesem Befragungszeitpunkt noch die Hauptschule. Weiterhin sind im Fünf-Wellen-Paneldatensatz im Vergleich zur Schulstatistik die Gesamtschülerinnen und -schüler zu  $t_1$  und  $t_3$  mit etwas über einem Prozentpunkt, zu  $t_2$  mit etwa drei Prozentpunkten überrepräsentiert.

Während sich zu  $t_4$  die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatensätzen  $Q_{t_4}$  mit weniger als einem halben Prozentpunkt unterscheiden, sind die Gesamtschüler zu  $t_5$  im Fünf-Wellen-Paneldatensatz mit einem Unterschied von mehr als vier Prozentpunkten im Gegensatz zu den Querschnittsdaten deutlich überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage sind stets die Querschnittsdatensätze der betreffenden Erhebungsjahre. So fällt zu  $t_5$  mit knapp vierzehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3).

Im Gegensatz dazu liegen die Anteile der Gymnasiastinnen und Gymnasiasten im Fünf-Wellen-Paneldatensatz zu den Zeitpunkten  $t_1$ ,  $t_2$  und  $t_3$  zwischen vier und knapp sechs Prozentpunkten über denen der jeweiligen Schulstatistik. Im Anschluss wird diese Überrepräsentation mit mehr als zwei Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – zunächst geringer, um zu  $t_5$  auf neun Prozentpunkte anzusteigen (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Der Anteil der Realschülerinnen und -schüler liegt im Fünf-Wellen-Paneldatensatz im Zeitraum  $t_1$  bis  $t_3$  zwischen über einem und vier Prozentpunkten höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit etwa drei Prozentpunkten überrepräsentiert sind, sind zu  $t_5$  die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  sind diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Fünf-Wellen-Paneldatensatz  $P_{t_{1,2,3,4,5}}$  – verglichen mit dem Querschnittsdatensatz – überrepräsentiert. Mit etwas weniger als einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel sehr leicht unterrepräsentiert. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Fünf-Wellen-Paneldatensatz  $P_{t_{1,2,3,4,5}}$  enthalten.

Tabelle 5.124 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Fünf-Wellen-Paneldatensatz  $P_{t_{1,2,3,4,5}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

## 5.5 Die Fünf-Wellen-Panel

**Tabelle 5.123:** Zusammensetzung  $P_{t_1,2,3,4,5}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,3	18,1	—	1,3	—	—	2003	
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728	
		$P_{t_1,2,3,4,5}$	21,5	20,2	36,1	22,1	—	—	—	—	642	
	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927	
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703	
		$P_{t_1,2,3,4,5}$	22,1	20,3	36,1	21,5	—	—	—	—	651	
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635	
		$P_{t_1,2,3,4,5}$	21,9	21,3	35,7	21,1	—	—	—	—	644	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{t_1,2,3,4,5}$	21,0	21,3	35,4	21,3	0,5	—	0,6	—	644	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{t_1,2,3,4,5}$	1,3	0,8	20,0	23,9	49,5	—	2,4	2,1	624	
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—	1907
			$Q_{t_1}$	17,9	24,1	31,8	26,3	—	0,0	—	—	1679
$P_{t_1,2,3,4,5}$			14,5	26,7	31,3	27,5	—	—	—	—	910	
$t_2$		$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836	
		$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1687	
		$P_{t_1,2,3,4,5}$	14,8	26,7	31,6	26,9	—	—	—	—	901	
$t_3$		$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926	
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704	
		$P_{t_1,2,3,4,5}$	15,0	27,1	31,9	26,0	—	—	—	—	908	
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{t_1,2,3,4,5}$	14,7	27,2	31,2	26,3	0,4	—	0,1	—	904	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{t_1,2,3,4,5}$	0,1	0,2	19,6	31,8	44,7	—	2,2	1,4	874	
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
			$Q_{t_1}$	22,0	23,7	31,1	22,8	—	0,4	—	—	3407
	$P_{t_1,2,3,4,5}$		17,4	24,0	33,3	25,3	—	—	—	—	1552	
	$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763	
		$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390	
		$P_{t_1,2,3,4,5}$	17,8	24,0	33,5	24,6	—	—	—	—	1552	
	$t_3$	$S_{t_3}$	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	
		$P_{t_1,2,3,4,5}$	17,8	24,7	33,5	24,0	—	—	—	—	1552	
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{t_1,2,3,4,5}$	17,3	24,7	32,9	24,2	0,5	—	0,1	—	1548	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{t_1,2,3,4,5}$	0,6	0,5	19,8	28,5	46,7	—	2,3	1,7	1498	

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



**Tabelle 5.124:** Zusammensetzung  $P_{t_1,2,3,4,5}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2	
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7	
		$P_{t_1,2,3,4,5}$	51,1	34,9	44,9	36,2	—	—	—	—	41,4	
	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2	
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2	
		$P_{t_1,2,3,4,5}$	52,0	35,4	45,2	36,6	—	—	—	—	41,9	
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9	
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0	
		$P_{t_1,2,3,4,5}$	50,9	35,8	44,2	36,6	—	—	—	—	41,5	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_1,2,3,4,5}$	50,4	35,8	44,7	36,5	42,9	—	80,0	—	41,6	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{t_1,2,3,4,5}$	88,9	71,4	42,2	34,9	44,1	—	44,1	52,0	41,7	
	weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	56,4	—	12,9	—	—	48,8
			$Q_{t_1}$	40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
$P_{t_1,2,3,4,5}$			48,9	65,1	55,1	63,8	—	—	—	—	58,6	
$t_2$		$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8	
		$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8	
		$P_{t_1,2,3,4,5}$	48,0	64,6	54,8	63,4	—	—	—	—	58,1	
$t_3$		$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1	
		$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0	
		$P_{t_1,2,3,4,5}$	49,1	64,2	55,8	63,4	—	—	—	—	58,5	
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_1,2,3,4,5}$	49,6	64,2	55,3	63,5	57,1	—	20,0	—	58,4	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{t_1,2,3,4,5}$	11,1	28,6	57,8	65,1	55,9	—	55,9	48,0	58,3	

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

5.5.2 Fünf-Wellen-Panel 2003 bis 2007 ( $P_{t_{2,3,4,5,6}}$ )

Das durchgehende Fünf-Wellen-Panel der Jahre 2003 bis 2007 enthält 1 620 Fälle. Aus Tabelle 5.125 geht hervor, dass im Fünf-Wellen-Panel Datensatz  $P_{t_{2,3,4,5,6}}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.125:** Zusammensetzung nach Geschlecht  $P_{t_{2,3,4,5,6}}$

		männlich		weiblich		gesamt
		Hfk	%	Hfk	%	(100 %)
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3 763
	$Q_{t_2}$	1703	50,2	1687	49,8	3 390
	$P_{t_{2,3,4,5,6}}$	643	39,7	977	60,3	1 620
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3 925
	$Q_{t_3}$	1635	49,0	1704	51,0	3 339
	$P_{t_{2,3,4,5,6}}$	635	39,2	985	60,8	1 620
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3 405
	$P_{t_{2,3,4,5,6}}$	643	39,7	977	60,3	1 620
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4 547
	$P_{t_{2,3,4,5,6}}$	640	39,5	980	60,5	1 620
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3 335
	$P_{t_{2,3,4,5,6}}$	639	39,4	981	60,6	1 620

Vergleicht man die Geschlechterverteilung der Wellen  $t_2$  und  $t_3$  im Fünf-Wellen-Panel Datensatz mit den Schulstatistiken der Erhebungswellen  $t_2$  und  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Fünf-Wellen-Panel Datensatz in jedem dieser Jahre um etwa elfeinhalb Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatensätzen  $Q_{t_2}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Fünf-Wellen-Panel Datensatz um etwa zehn bzw. zehneinhalb Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Fünf-Wellen-Panel Datensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$ ,  $t_5$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Fünf-Wellen-Panel Datensatz zunächst über zehn ( $t_4$  und  $t_5$ ), zu  $t_6$  über siebeneinhalb Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatensatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Panel Datensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Fünf-Wellen-Panel Datensatz (siehe Tabelle 5.126) zeigt, dass das Durchschnittsalter

**Tabelle 5.126:** Altersdurchschnitt  $P_{t_2,3,4,5,6}$ 

		Alter	Std.-abw.	gesamt
$t_2$	$Q_{t_2}$	14,0	0,8	3 275
	$P_{t_2,3,4,5,6}$	13,9	0,7	1 600
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_2,3,4,5,6}$	14,9	0,7	1 607
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_2,3,4,5,6}$	15,9	0,7	1 603
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_2,3,4,5,6}$	16,9	0,8	1 610
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_2,3,4,5,6}$	17,9	0,7	1 615

der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatsätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatsätzen höher ist.

Tabelle 5.127 zeigt, dass der Fünf-Wellen-Paneldatensatz  $P_{t_2,3,4,5,6}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Fünf-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_2$  und  $t_3$  über zehn bzw. über siebeneinhalb Prozentpunkte unter denen der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatsätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von etwa fünf Prozentpunkten die Hauptschülerinnen und -schüler im Fünf-Wellen-Paneldatensatz ebenfalls deutlich unterrepräsentiert sind, sind zu  $t_5$  und  $t_6$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte bzw. keine Befragten besuchen zu diesen Befragungszeitpunkten noch die Hauptschule. Weiterhin sind im Fünf-Wellen-Paneldatensatz zu  $t_2$  die Gesamtschülerinnen und -schüler im Fünf-Wellen-Paneldatensatz mit etwa zwei Prozentpunkten überrepräsentiert; zu  $t_3$  sind ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik. Während sich zu  $t_4$  und  $t_6$  die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatsätzen  $Q_{t_4}$  und  $Q_{t_6}$  ebenfalls mit weit weniger als einem Prozentpunkt – allerdings in abweichende Richtungen – nur geringfügig unterscheiden, sind die Gesamtschüler zu  $t_5$  im Fünf-Wellen-Paneldatensatz mit einem Unterschied von über fünf Prozentpunkten im Gegensatz zu den Querschnittsdaten deutlich überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage sind stets die Querschnittsdatensätze der betreffenden Erhebungsjahre. So fällt zu  $t_5$  mit knapp achtzehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3). Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Fünf-Wellen-Paneldatensatz  $P_{t_{2,3,4,5,6}}$  ähnlich, mit etwa einem Prozentpunkt sind Befragte, die ein Berufskolleg besuchen nur leicht im Fünf-Wellen-Paneldatensatz überrepräsentiert.

Die Anteile der Gymnasiastinnen und Gymnasiasten liegen im Fünf-Wellen-Paneldatensatz zu den Zeitpunkten  $t_2$  und  $t_3$  etwa achteinhalb bzw. sechseinhalb Prozentpunkte über denen der jeweiligen Schulstatistik. Im Anschluss wird diese Überrepräsentation mit etwa viereinhalb Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – zunächst geringer, um zu  $t_5$  auf elfeinhalb Prozentpunkte anzusteigen (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Zu  $t_6$  sind die Gymnasiastinnen und Gymnasiasten verglichen mit dem Querschnittsdatensatz mit etwa zwei Prozentpunkten wiederum in geringerem Ausmaß überrepräsentiert. Der Anteil der Realschülerinnen und -schüler liegt im Fünf-Wellen-Paneldatensatz im Zeitraum  $t_2$  und  $t_3$  etwa einen halben bzw. zwei Prozentpunkte höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit etwas über eineinhalb Prozentpunkten überrepräsentiert sind, sind in den folgenden Wellen die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  und  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Fünf-Wellen-Paneldatensatz  $P_{t_{2,3,4,5,6}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Während zur Erhebungswelle  $t_5$  der Anteil derjenigen, die keine Schule besuchen, im Panel- und im Querschnittsdatensatz identisch ist, zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Fünf-Wellen-Paneldatensatz mit etwa vier Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Fünf-Wellen-Paneldatensatz enthalten.

Tabelle 5.128 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Fünf-Wellen-Paneldatensatz  $P_{t_{2,3,4,5,6}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

**Tabelle 5.127:** Zusammensetzung  $P_{2,3,4,5,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927	
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703	
		$P_{2,3,4,5,6}$	21,6	19,3	34,4	24,7	—	—	—	—	643	
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635	
		$P_{2,3,4,5,6}$	21,6	20,0	34,5	23,9	—	—	—	—	635	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{2,3,4,5,6}$	20,6	19,8	34,5	24,0	0,6	—	0,5	—	641	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{2,3,4,5,6}$	1,5	0,8	22,0	26,8	44,4	—	2,1	2,4	619	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504	
		$P_{2,3,4,5,6}$	—	—	19,0	24,6	34,5	—	11,2	10,7	617	
	weiblich	$t_2$	$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836
			$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1687
			$P_{2,3,4,5,6}$	14,2	25,6	31,3	28,9	—	—	—	—	977
		$t_3$	$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926
			$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704
			$P_{2,3,4,5,6}$	14,6	25,8	31,6	28,0	—	—	—	—	985
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{2,3,4,5,6}$	14,0	25,7	31,4	28,4	0,4	—	0,1	—	977	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{2,3,4,5,6}$	0,2	0,4	20,4	33,6	41,6	—	1,9	1,8	951	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715	
		$P_{2,3,4,5,6}$	—	—	18,1	30,8	30,7	—	10,8	9,6	957	
gesamt		$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
			$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390
			$P_{2,3,4,5,6}$	17,2	23,1	32,5	27,2	—	—	—	—	1620
		$t_3$	$S_{t_3}$	24,9	21,5	32,3	20,0	—	1,3	—	—	3925
			$Q_{t_3}$	22,6	22,9	31,9	22,2	—	0,5	—	—	3339
			$P_{2,3,4,5,6}$	17,3	23,5	32,7	26,4	—	—	—	—	1620
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{2,3,4,5,6}$	16,6	23,4	32,6	26,6	0,5	—	0,2	—	1618	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{2,3,4,5,6}$	0,7	0,6	21,0	31,0	42,7	—	2,0	2,0	1570	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
		$P_{2,3,4,5,6}$	—	—	18,4	28,4	32,2	—	10,9	10,0	1574	

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

Tabelle 5.128: Zusammensetzung  $P_{12,3,4,5,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform <sup>a</sup> (%)								gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2
		$P_{12,3,4,5,6}$	50,0	33,2	41,9	36,1	—	—	—	—	39,7
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0
		$P_{12,3,4,5,6}$	48,8	33,3	41,3	35,5	—	—	—	—	39,2
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3
		$P_{12,3,4,5,6}$	49,1	33,6	41,9	35,7	50,0	—	75,0	—	39,6
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5
		$P_{12,3,4,5,6}$	81,8	55,6	41,2	34,2	41,0	—	41,9	46,9	39,4
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7
		$P_{12,3,4,5,6}$	—	—	40,3	34,0	42,0	—	40,1	41,8	39,2
weiblich	$t_2$	$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8
		$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8
		$P_{12,3,4,5,6}$	50,0	66,8	58,1	63,9	—	—	—	—	60,3
	$t_3$	$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1
		$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0
		$P_{12,3,4,5,6}$	51,2	66,7	58,7	64,5	—	—	—	—	60,8
	$t_4$	$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7
		$P_{12,3,4,5,6}$	50,9	66,4	58,1	64,3	50,0	—	25,0	—	60,4
	$t_5$	$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5
		$P_{12,3,4,5,6}$	18,2	44,4	58,8	65,8	59,0	—	58,1	53,1	60,6
	$t_6$	$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3
		$P_{12,3,4,5,6}$	—	—	59,7	66,0	58,0	—	59,9	58,2	60,8

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

5.5.3 Fünf-Wellen-Panel 2002, 2003, 2004, 2005 und 2007 ( $P_{t_{1,2,3,4,6}}$ )

Das Fünf-Wellen-Panel der Jahre 2002, 2003, 2004, 2005 und 2007 enthält 1 396 Fälle. Aus Tabelle 5.129 geht hervor, dass im Fünf-Wellen-Paneldatensatz  $P_{t_{1,2,3,4,6}}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.129:** Zusammensetzung nach Geschlecht  $P_{t_{1,2,3,4,6}}$

		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
	$P_{t_{1,2,3,4,6}}$	548	39,3	848	60,7	1396
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{t_{1,2,3,4,6}}$	557	39,9	839	60,1	1396
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_{1,2,3,4,6}}$	550	39,4	846	60,6	1396
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_{1,2,3,4,6}}$	557	39,9	839	60,1	1396
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3335
	$P_{t_{1,2,3,4,6}}$	553	39,6	843	60,4	1396

Vergleicht man die Geschlechterverteilung der ersten drei Wellen im Fünf-Wellen-Paneldatensatz mit den Schulstatistiken der Erhebungswellen  $t_1$  bis  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Fünf-Wellen-Paneldatensatz in jedem der Jahre um etwa elfeinhalb Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatsätzen  $Q_{t_1}$ ,  $Q_{t_2}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Fünf-Wellen-Paneldatensatz etwas über zehn Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Fünf-Wellen-Paneldatensatz mit den Querschnittsdatsätzen der Erhebungszeitpunkte  $t_4$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Fünf-Wellen-Paneldatensatz zu  $t_4$  zunächst zehneinhalb, zu  $t_6$  über sieben Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatsatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Paneldatsätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatsätzen und dem

**Tabelle 5.130:** Altersdurchschnitt  $P_{1,2,3,4,6}$

		Alter	Std.-abw.	gesamt
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{1,2,3,4,6}$	12,9	0,7	1 383
$t_2$	$Q_{t_2}$	14,0	0,8	3 275
	$P_{1,2,3,4,6}$	13,9	0,7	1 380
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{1,2,3,4,6}$	14,9	0,7	1 388
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{1,2,3,4,6}$	15,9	0,7	1 382
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{1,2,3,4,6}$	17,9	0,7	1 392

Fünf-Wellen-Paneldatensatz (siehe Tabelle 5.130) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.131 zeigt, dass der Fünf-Wellen-Paneldatensatz  $P_{1,2,3,4,6}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Fünf-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_1$  bis  $t_3$  zwischen siebeneinhalb und etwa zehneinhalb Prozentpunkten unter denen der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von etwa fünfeinhalb Prozentpunkten die Hauptschülerinnen und -schüler im Fünf-Wellen-Paneldatensatz ebenfalls deutlich unterrepräsentiert sind, sind zu  $t_6$  keine gesicherten Aussagen mehr möglich: Kein Befragter besucht zu diesem Befragungszeitpunkt mehr die Hauptschule. Weiterhin sind im Fünf-Wellen-Paneldatensatz zu  $t_1$  und  $t_3$  mit etwa einem halben Prozentpunkt Unterschied ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik; zu  $t_2$  sind diese im Fünf-Wellen-Paneldatensatz mit über zwei Prozentpunkten überrepräsentiert. Zu  $t_4$  und  $t_6$  unterscheiden sich die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatensätzen  $Q_{t_4}$  und  $Q_{t_6}$  mit weniger als einem Prozentpunkt; in beiden Jahren sind sie im Panel sehr leicht unterrepräsentiert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_6$  getroffen werden. Vergleichsgrundlage sind stets die Querschnittsdatensätze der betreffenden Erhebungsjahre. Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten



im Querschnittsdatensatz  $Q_{t_6}$  und im Fünf-Wellen-Paneldatensatz  $P_{t_1,2,3,4,6}$  ähnlich, mit einem Prozentpunkt sind Befragte, die ein Berufskolleg besuchen nur leicht im Fünf-Wellen-Paneldatensatz überrepräsentiert.

Die Anteile der Gymnasiastinnen und Gymnasiasten liegen im Fünf-Wellen-Paneldatensatz zu den Zeitpunkten  $t_1$ ,  $t_2$  und  $t_3$  etwa zwischen fünfeinhalb und sieben einhalb Prozentpunkten über denen der jeweiligen Schulstatistik. Im Anschluss wird diese Überrepräsentation mit etwa dreieinhalb Prozentpunkten zu  $t_4$  und über einem Prozentpunkt zu  $t_6$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – geringer. Der Anteil der Realschülerinnen und -schüler liegt im Fünf-Wellen-Paneldatensatz im Zeitraum  $t_1$  bis  $t_3$  zwischen eineinhalb und dreieinhalb Prozentpunkten höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit etwa drei Prozentpunkten überrepräsentiert sind, sind zu  $t_6$  keine Realschüler mehr im Paneldatensatz enthalten.

Zu  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit über einem Prozentpunkt im Fünf-Wellen-Paneldatensatz  $P_{t_1,2,3,4,6}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Während im vorliegenden Fünf-Wellen-Paneldatensatz bis zur Erhebungswelle  $t_6$  keine Befragten enthalten sind, die keine Schule besuchen, zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Panel mit dreieinhalb Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Fünf-Wellen-Paneldatensatz enthalten.

Tabelle 5.132 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Fünf-Wellen-Paneldatensatz  $P_{t_1,2,3,4,6}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

## 5.5 Die Fünf-Wellen-Panel

**Tabelle 5.131:** Zusammensetzung  $P_{t_1,2,3,4,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,3	18,1	—	1,3	—	—	2003	
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728	
		$P_{t_1,2,3,4,6}$	19,7	20,8	35,4	24,1	—	—	—	—	548	
	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927	
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703	
		$P_{t_1,2,3,4,6}$	20,5	21,0	35,2	23,3	—	—	—	—	557	
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635	
		$P_{t_1,2,3,4,6}$	20,5	22,0	34,5	22,9	—	—	—	—	550	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{t_1,2,3,4,6}$	19,5	22,0	34,8	22,7	0,7	—	0,4	—	555	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504	
		$P_{t_1,2,3,4,6}$	—	—	17,6	23,5	34,9	—	12,4	11,6	533	
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—	1907
			$Q_{t_1}$	17,9	24,1	31,8	26,3	—	0,0	—	—	1679
			$P_{t_1,2,3,4,6}$	14,3	26,5	30,9	28,3	—	—	—	—	848
		$t_2$	$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836
			$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1687
$P_{t_1,2,3,4,6}$			14,5	26,6	31,1	27,8	—	—	—	—	839	
$t_3$		$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926	
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704	
		$P_{t_1,2,3,4,6}$	14,8	26,7	31,4	27,1	—	—	—	—	846	
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{t_1,2,3,4,6}$	14,1	26,8	30,8	27,5	0,6	—	0,2	—	839	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715	
		$P_{t_1,2,3,4,6}$	—	—	17,9	30,1	30,1	—	11,8	10,0	820	
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
			$Q_{t_1}$	22,0	23,7	31,1	22,8	—	0,4	—	—	3407
			$P_{t_1,2,3,4,6}$	16,4	24,3	32,7	26,6	—	—	—	—	1396
		$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
			$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390
	$P_{t_1,2,3,4,6}$		16,9	24,4	32,7	26,0	—	—	—	—	1396	
	$t_3$	$S_{t_3}$	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	
		$P_{t_1,2,3,4,6}$	17,0	24,9	32,7	25,4	—	—	—	—	1396	
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{t_1,2,3,4,6}$	16,2	24,9	32,4	25,6	0,6	—	0,2	—	1394	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
		$P_{t_1,2,3,4,6}$	—	—	17,8	27,5	32,0	—	12,0	10,6	1353	

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

**Tabelle 5.132:** Zusammensetzung  $P_{t_1,2,3,4,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2	
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7	
		$P_{t_1,2,3,4,6}$	47,2	33,6	42,5	35,5	—	—	—	—	39,3	
	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2	
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2	
		$P_{t_1,2,3,4,6}$	48,3	34,4	42,9	35,8	—	—	—	—	39,9	
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9	
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0	
		$P_{t_1,2,3,4,6}$	47,5	34,9	41,7	35,5	—	—	—	—	39,4	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_1,2,3,4,6}$	47,8	35,2	42,8	35,3	44,4	—	50,0	—	39,8	
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7	
		$P_{t_1,2,3,4,6}$	—	—	39,0	33,6	43,0	—	40,5	43,1	39,4	
	weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	56,4	—	12,9	—	—	48,8
			$Q_{t_1}$	40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
$P_{t_1,2,3,4,6}$			52,8	66,4	57,5	64,5	—	—	—	—	60,7	
$t_2$		$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8	
		$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8	
		$P_{t_1,2,3,4,6}$	51,7	65,6	57,1	64,2	—	—	—	—	60,1	
$t_3$		$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1	
		$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0	
		$P_{t_1,2,3,4,6}$	52,5	65,1	58,3	64,5	—	—	—	—	60,6	
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_1,2,3,4,6}$	52,2	64,8	57,2	64,7	55,6	—	50,0	—	60,2	
$t_6$		$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3	
		$P_{t_1,2,3,4,6}$	—	—	61,0	66,4	57,0	—	59,5	56,9	60,6	

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

5.5.4 Fünf-Wellen-Panel 2002, 2003, 2004, 2006 und 2007 ( $P_{t_1,2,3,5,6}$ )

Das Fünf-Wellen-Panel der Jahre 2002, 2003, 2004, 2006 und 2007 enthält 1 346 Fälle.

Aus Tabelle 5.133 geht hervor, dass im Fünf-Wellen-Panel Datensatz  $P_{t_1,2,3,5,6}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.133:** Zusammensetzung nach Geschlecht  $P_{t_1,2,3,5,6}$

		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
	$P_{t_1,2,3,5,6}$	516	38,3	830	61,7	1346
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{t_1,2,3,5,6}$	524	38,9	822	61,1	1346
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_1,2,3,5,6}$	516	38,3	830	61,7	1346
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_1,2,3,5,6}$	522	38,8	824	61,2	1346
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3335
	$P_{t_1,2,3,5,6}$	522	38,8	824	61,2	1346

Vergleicht man die Geschlechterverteilung der ersten drei Wellen im Fünf-Wellen-Panel Datensatz mit den Schulstatistiken der Erhebungswellen  $t_1$  bis  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Fünf-Wellen-Panel Datensatz in jedem der Jahre um etwa zwölftehalb Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatensätzen  $Q_{t_1}$ ,  $Q_{t_2}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Fünf-Wellen-Panel Datensatz zwischen zehneinhalb und zwölftehalb Prozentpunkten erhöht. Die Gegenüberstellung der Geschlechterverteilung im Fünf-Wellen-Panel Datensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_5$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Fünf-Wellen-Panel Datensatz zu  $t_5$  zunächst elf, zu  $t_6$  über acht Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatensatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Panel Datensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem

**Tabelle 5.134:** Altersdurchschnitt  $P_{t_{1,2,3,5,6}}$

		Alter	Std.-abw.	gesamt
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{t_{1,2,3,5,6}}$	12,9	0,7	1 336
$t_2$	$Q_{t_2}$	14,0	0,8	3 275
	$P_{t_{1,2,3,5,6}}$	13,9	0,7	1 330
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_{1,2,3,5,6}}$	14,9	0,7	1 338
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_{1,2,3,5,6}}$	16,9	0,7	1 338
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_{1,2,3,5,6}}$	17,9	0,7	1 342

Fünf-Wellen-Paneldatensatz (siehe Tabelle 5.134) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatsätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatsätzen höher ist.

Tabelle 5.135 zeigt, dass der Fünf-Wellen-Paneldatensatz  $P_{t_{1,2,3,5,6}}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Fünf-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_1$  bis  $t_3$  zwischen achteinhalb und elfeinhalb Prozentpunkten unter denen der Schulstatistik. Zu  $t_5$  und  $t_6$  sind keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte bzw. keine Befragten besuchen zu diesen Befragungszeitpunkten noch die Hauptschule. Weiterhin sind im Fünf-Wellen-Paneldatensatz zu  $t_1$  und  $t_3$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik; lediglich zu  $t_2$  sind diese im Fünf-Wellen-Paneldatensatz mit knapp zwei Prozentpunkten überrepräsentiert. Während sich zu  $t_6$  der Anteil der Gesamtschüler vom entsprechenden Querschnittsdatsatz  $Q_{t_6}$  ebenfalls mit weniger als einem Prozentpunkt unterscheidet, sind die Gesamtschüler zu  $t_5$  im Fünf-Wellen-Paneldatensatz mit einem Unterschied von über fünf Prozentpunkten im Gegensatz zu den Querschnittsdaten deutlich überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Gesicherte Aussagen über die Anteile derjenigen Befragten, die ein Berufskolleg besuchen, können erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage sind stets die Querschnittsdatsätze der betreffenden Erhebungsjahre. So fällt zu  $t_5$  mit achtzehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf,

die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3). Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Fünf-Wellen-Paneldatensatz  $P_{t_{1,2,3,5,6}}$  ähnlich; mit etwas über einem Prozentpunkt sind Befragte, die ein Berufskolleg besuchen nur leicht im Fünf-Wellen-Paneldatensatz überrepräsentiert.

Die Anteile der Gymnasiastinnen und Gymnasiasten liegen im Fünf-Wellen-Paneldatensatz zu den Zeitpunkten  $t_1$ ,  $t_2$  und  $t_3$  etwa zwischen sechseinhalb und achteinhalb Prozentpunkten über denen der jeweiligen Schulstatistik. Zu  $t_5$  steigt diese Überrepräsentation – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – auf zwölf Prozentpunkte an (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3), um zu  $t_6$  auf etwa zweieinhalb Prozentpunkte abzusinken. Der Anteil der Realschülerinnen und -schüler liegt im Fünf-Wellen-Paneldatensatz im Zeitraum  $t_1$  bis  $t_3$  zwischen zwei und knapp vier Prozentpunkten höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. In den folgenden Wellen sind die Anteile der Realschülerinnen und -schüler zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  und  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit einem halben Prozentpunkt im Fünf-Wellen-Paneldatensatz  $P_{t_{1,2,3,5,6}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Mit etwas weniger als einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel sehr leicht unterrepräsentiert. Überdies zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Fünf-Wellen-Paneldatensatz mit viereinhalb Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Fünf-Wellen-Paneldatensatz enthalten.

Tabelle 5.136 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Fünf-Wellen-Paneldatensatz  $P_{t_{1,2,3,5,6}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

**Tabelle 5.135:** Zusammensetzung  $P_{1,2,3,5,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,3	18,1	—	1,3	—	—	2003	
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728	
		$P_{1,2,3,5,6}$	18,8	20,9	34,9	25,4	—	—	—	—	516	
	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927	
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703	
		$P_{1,2,3,5,6}$	19,7	21,0	34,7	24,6	—	—	—	—	524	
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635	
		$P_{1,2,3,5,6}$	19,6	21,9	34,3	24,2	—	—	—	—	516	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{1,2,3,5,6}$	1,4	1,2	22,3	27,0	44,1	—	2,2	1,8	503	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504	
		$P_{1,2,3,5,6}$	—	—	19,0	24,8	34,7	—	10,9	10,7	505	
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—	1907
			$Q_{t_1}$	17,9	24,1	31,8	26,3	—	0,0	—	—	1679
			$P_{1,2,3,5,6}$	13,4	26,9	30,5	29,3	—	—	—	—	830
		$t_2$	$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836
			$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1687
$P_{1,2,3,5,6}$			13,6	27,1	30,7	28,6	—	—	—	—	822	
$t_3$		$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926	
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704	
		$P_{1,2,3,5,6}$	13,9	27,3	31,0	27,8	—	—	—	—	830	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{1,2,3,5,6}$	0,1	0,3	20,3	34,3	41,5	—	1,9	1,6	798	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715	
		$P_{1,2,3,5,6}$	—	—	18,0	31,0	30,6	—	11,4	9,0	804	
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
			$Q_{t_1}$	22,0	23,7	31,1	22,8	—	0,4	—	—	3407
			$P_{1,2,3,5,6}$	15,5	24,6	32,2	27,8	—	—	—	—	1346
		$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
			$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390
	$P_{1,2,3,5,6}$		16,0	24,7	32,2	27,0	—	—	—	—	1346	
	$t_3$	$S_{t_3}$	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	
		$P_{1,2,3,5,6}$	16,0	25,3	32,2	26,4	—	—	—	—	1346	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{1,2,3,5,6}$	0,6	0,6	21,1	31,5	42,5	—	2,0	1,7	1301	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
		$P_{1,2,3,5,6}$	—	—	18,4	28,6	32,2	—	11,2	9,6	1309	

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

**Tabelle 5.136:** Zusammensetzung  $P_{1,2,3,5,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2	
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7	
		$P_{1,2,3,5,6}$	46,6	32,6	41,6	35,0	—	—	—	—	38,3	
	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2	
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2	
		$P_{1,2,3,5,6}$	47,9	33,0	41,9	35,4	—	—	—	—	38,9	
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9	
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0	
		$P_{1,2,3,5,6}$	46,8	33,2	40,8	35,1	—	—	—	—	38,3	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{1,2,3,5,6}$	87,5	75,0	40,9	33,2	40,1	—	42,3	40,9	38,7	
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7	
		$P_{1,2,3,5,6}$	—	—	39,8	33,4	41,6	—	37,4	42,9	38,6	
	weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	56,4	—	12,9	—	—	48,8
			$Q_{t_1}$	40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
			$P_{1,2,3,5,6}$	53,4	67,4	58,4	65,0	—	—	—	—	61,7
		$t_2$	$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8
			$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8
$P_{1,2,3,5,6}$			52,1	67,0	58,1	64,6	—	—	—	—	61,1	
$t_3$		$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1	
		$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0	
		$P_{1,2,3,5,6}$	53,2	66,8	59,2	64,9	—	—	—	—	61,7	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{1,2,3,5,6}$	12,5	25,0	59,1	66,8	59,9	—	57,7	59,1	61,3	
$t_6$		$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3	
		$P_{1,2,3,5,6}$	—	—	60,2	66,6	58,4	—	62,6	57,1	61,4	

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



5.5.5 Fünf-Wellen-Panel 2002, 2003, 2005, 2006 und 2007 ( $P_{t_{1,2,4,5,6}}$ )

Das Fünf-Wellen-Panel der Jahre 2002, 2003, 2005, 2006 und 2007 enthält 1 398 Fälle. Aus Tabelle 5.137 geht hervor, dass im Fünf-Wellen-Paneldatensatz  $P_{t_{1,2,4,5,6}}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.137:** Zusammensetzung nach Geschlecht  $P_{t_{1,2,4,5,6}}$

		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
	$P_{t_{1,2,4,5,6}}$	551	39,4	847	60,6	1398
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{t_{1,2,4,5,6}}$	557	39,8	841	60,2	1398
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_{1,2,4,5,6}}$	556	39,8	842	60,2	1398
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_{1,2,4,5,6}}$	555	39,7	843	60,3	1398
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3335
	$P_{t_{1,2,4,5,6}}$	554	39,6	844	60,4	1398

Vergleicht man die Geschlechterverteilung der ersten zwei Wellen im Fünf-Wellen-Paneldatensatz mit den Schulstatistiken der Erhebungswellen  $t_1$  und  $t_2$ , so wird deutlich, dass der Anteil der Schülerinnen im Fünf-Wellen-Paneldatensatz in jedem der beiden Jahre um etwa elfeinhalb Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatsätzen  $Q_{t_1}$  und  $Q_{t_2}$  ist der Anteil der weiblichen Befragten im Fünf-Wellen-Paneldatensatz etwa elfeinhalb bzw. zehneinhalb Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Fünf-Wellen-Paneldatensatz mit den Querschnittsdatsätzen der Erhebungszeitpunkte  $t_4$ ,  $t_5$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Fünf-Wellen-Paneldatensatz zunächst über zehn ( $t_4$  und  $t_5$ ), zu  $t_6$  über sieben Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatsatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Paneldatsätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatsätzen und dem Fünf-Wellen-Paneldatensatz (siehe Tabelle 5.138) zeigt, dass das Durchschnittsalter

**Tabelle 5.138:** Altersdurchschnitt  $P_{1,2,4,5,6}$

		Alter	Std.-abw.	gesamt
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{1,2,4,5,6}$	12,9	0,7	1 386
$t_2$	$Q_{t_2}$	14,0	0,8	3 275
	$P_{1,2,4,5,6}$	13,9	0,7	1 380
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{1,2,4,5,6}$	15,9	0,7	1 383
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{1,2,4,5,6}$	16,9	0,7	1 390
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{1,2,4,5,6}$	17,9	0,7	1 393

der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.139 zeigt, dass der Fünf-Wellen-Paneldatensatz  $P_{1,2,4,5,6}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Fünf-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_1$  und  $t_2$  siebeneinhalb bzw. zehneinhalb Prozentpunkte unter denen der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von etwa fünfeinhalb Prozentpunkten die Hauptschülerinnen und -schüler im Fünf-Wellen-Paneldatensatz ebenfalls deutlich unterrepräsentiert sind, sind zu  $t_5$  und  $t_6$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte bzw. keine Befragten besuchen zu diesen Befragungszeitpunkten noch die Hauptschule. Weiterhin sind im Fünf-Wellen-Paneldatensatz zu  $t_1$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik; zu  $t_2$  sind diese im Fünf-Wellen-Paneldatensatz mit knapp eineinhalb Prozentpunkten überrepräsentiert. Während sich zu  $t_4$  und  $t_6$  die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatensätzen  $Q_{t_4}$  und  $Q_{t_6}$  mit weniger als einem Prozentpunkt – allerdings in abweichende Richtungen – unterscheiden, sind die Gesamtschüler zu  $t_5$  im Fünf-Wellen-Paneldatensatz mit einem Unterschied von über fünfeinhalb Prozentpunkten im Gegensatz zu den Querschnittsdaten deutlich überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage sind stets die Querschnittsdatensätze der betreffenden Erhebungsjahre. So fällt zu  $t_5$  mit über achtzehneinhalb Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3). Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Fünf-Wellen-Paneldatensatz  $P_{t_{1,2,4,5,6}}$  ähnlich, mit unter einem Prozentpunkt sind Befragte, die ein Berufskolleg besuchen nur leicht im Fünf-Wellen-Paneldatensatz überrepräsentiert.

Die Anteile der Gymnasiastinnen und Gymnasiasten liegen im Fünf-Wellen-Paneldatensatz zu den Zeitpunkten  $t_1$  und  $t_2$  etwa sechs bzw. achteinhalb Prozentpunkte über denen der jeweiligen Schulstatistik. Im Anschluss wird diese Überrepräsentation mit viereinhalb Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – zunächst geringer, um zu  $t_5$  auf über zwölf Prozentpunkte anzusteigen (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Zu  $t_6$  sind die Gymnasiastinnen und Gymnasiasten verglichen mit dem Querschnittsdatensatz mit zweieinhalb Prozentpunkten wiederum in geringerem Ausmaß überrepräsentiert. Der Anteil der Realschülerinnen und -schüler liegt im Fünf-Wellen-Paneldatensatz im Zeitraum  $t_1$  und  $t_2$  unter zweieinhalb bzw. über eineinhalb Prozentpunkte höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit knapp unter drei Prozentpunkten überrepräsentiert sind, sind in den folgenden Wellen die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  und  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Fünf-Wellen-Paneldatensatz  $P_{t_{1,2,4,5,6}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Mit knapp einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel sehr leicht unterrepräsentiert. Überdies zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Fünf-Wellen-Paneldatensatz mit mehr als vier Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Fünf-Wellen-Paneldatensatz enthalten.

Tabelle 5.140 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Fünf-Wellen-Paneldatensatz  $P_{t_{1,2,4,5,6}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

## 5.5 Die Fünf-Wellen-Panel

**Tabelle 5.139:** Zusammensetzung  $P_{t_1,2,4,5,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,3	18,1	—	1,3	—	—	2003	
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728	
		$P_{t_1,2,4,5,6}$	20,1	20,7	33,6	25,6	—	—	—	—	551	
	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927	
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703	
		$P_{t_1,2,4,5,6}$	20,6	20,8	33,6	25,0	—	—	—	—	557	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{t_1,2,4,5,6}$	19,7	21,5	33,6	24,4	0,5	—	0,4	—	554	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{t_1,2,4,5,6}$	1,3	0,9	22,3	27,8	43,5	—	2,3	1,9	533	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504	
		$P_{t_1,2,4,5,6}$	—	—	18,6	25,3	34,1	—	11,2	10,8	537	
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—	1907
			$Q_{t_1}$	17,9	24,1	31,8	26,3	—	0,0	—	—	1679
			$P_{t_1,2,4,5,6}$	13,9	26,7	30,6	28,8	—	—	—	—	847
		$t_2$	$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836
			$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1687
			$P_{t_1,2,4,5,6}$	14,4	26,8	30,7	28,2	—	—	—	—	841
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{t_1,2,4,5,6}$	13,8	26,8	30,9	28,0	0,4	—	0,1	—	842	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{t_1,2,4,5,6}$	0,2	0,2	20,7	34,4	40,9	—	2,1	1,5	817	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715	
		$P_{t_1,2,4,5,6}$	—	—	18,3	30,9	30,0	—	11,7	9,1	823	
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
			$Q_{t_1}$	22,0	23,7	31,1	22,8	—	0,4	—	—	3407
			$P_{t_1,2,4,5,6}$	16,4	24,3	31,8	27,5	—	—	—	—	1398
		$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763
			$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390
			$P_{t_1,2,4,5,6}$	16,9	24,4	31,8	26,9	—	—	—	—	1398
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{t_1,2,4,5,6}$	16,1	24,7	31,9	26,6	0,4	—	0,2	—	1396	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{t_1,2,4,5,6}$	0,7	0,5	21,3	31,8	41,9	—	2,1	1,6	1350	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
		$P_{t_1,2,4,5,6}$	—	—	18,5	28,7	31,6	—	11,5	9,8	1360	

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

**Tabelle 5.140:** Zusammensetzung  $P_{t_{1,2,4,5,6}}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform <sup>a</sup> (%)								gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7
		$P_{t_{1,2,4,5,6}}$	48,5	33,5	41,7	36,6	—	—	—	—	39,4
	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2
		$P_{t_{1,2,4,5,6}}$	48,7	34,0	42,0	37,0	—	—	—	—	39,8
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3
		$P_{t_{1,2,4,5,6}}$	48,4	34,5	41,7	36,4	50,0	—	66,7	—	39,7
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5
		$P_{t_{1,2,4,5,6}}$	77,8	71,4	41,3	34,5	41,0	—	41,4	45,5	39,5
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7
		$P_{t_{1,2,4,5,6}}$	—	—	39,8	34,9	42,6	—	38,5	43,6	39,5
weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	56,4	—	12,9	—	—	48,8
		$Q_{t_1}$	40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
		$P_{t_{1,2,4,5,6}}$	51,5	66,5	58,3	63,4	—	—	—	—	60,6
	$t_2$	$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8
		$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8
		$P_{t_{1,2,4,5,6}}$	51,3	66,0	58,0	63,0	—	—	—	—	60,2
	$t_4$	$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7
		$P_{t_{1,2,4,5,6}}$	51,6	65,5	58,3	63,6	50,0	—	33,3	—	60,3
	$t_5$	$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5
		$P_{t_{1,2,4,5,6}}$	22,2	28,6	58,7	65,5	59,0	—	58,6	54,5	60,5
	$t_6$	$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3
		$P_{t_{1,2,4,5,6}}$	—	—	60,2	65,1	57,4	—	61,5	56,4	60,5

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

5.5.6 Fünf-Wellen-Panel 2002, 2004, 2005, 2006 und 2007 ( $P_{1,3,4,5,6}$ )

Das Fünf-Wellen-Panel der Jahre 2002, 2004, 2005, 2006 und 2007 enthält 1 411 Fälle. Aus Tabelle 5.141 geht hervor, dass im Fünf-Wellen-Panel Datensatz  $P_{1,3,4,5,6}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.141:** Zusammensetzung nach Geschlecht  $P_{1,3,4,5,6}$

		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
	$P_{1,3,4,5,6}$	557	39,5	854	60,5	1411
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{1,3,4,5,6}$	558	39,5	853	60,5	1411
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{1,3,4,5,6}$	562	39,8	849	60,2	1411
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{1,3,4,5,6}$	562	39,8	849	60,2	1411
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3335
	$P_{1,3,4,5,6}$	561	39,8	850	60,2	1411

Vergleicht man die Geschlechterverteilung der ersten und der dritten Welle im Fünf-Wellen-Panel Datensatz mit den Schulstatistiken der Erhebungswellen  $t_1$  und  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Fünf-Wellen-Panel Datensatz in beiden Jahren um etwa elfeinhalb Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatensätzen  $Q_{t_1}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Fünf-Wellen-Panel Datensatz etwa elf bzw. neuneinhalb Prozentpunkte erhöht. Die Gegenüberstellung der Geschlechterverteilung im Fünf-Wellen-Panel Datensatz mit den Querschnittsdatensätzen der Erhebungszeitpunkte  $t_4$ ,  $t_5$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Fünf-Wellen-Panel Datensatz zunächst etwa zehn ( $t_4$  und  $t_5$ ), zu  $t_6$  über sieben Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatensatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

Das durchschnittliche Alter der Befragten in den Panel Datensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Fünf-Wellen-Panel Datensatz (siehe Tabelle 5.142) zeigt, dass das Durchschnittsalter

**Tabelle 5.142:** Altersdurchschnitt  $P_{t_{1,3,4,5,6}}$

		Alter	Std.-abw.	gesamt
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{t_{1,3,4,5,6}}$	12,9	0,7	1 400
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_{1,3,4,5,6}}$	14,9	0,7	1 403
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_{1,3,4,5,6}}$	15,9	0,7	1 399
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_{1,3,4,5,6}}$	16,9	0,7	1 404
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_{1,3,4,5,6}}$	17,9	0,7	1 406

der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.143 zeigt, dass der Fünf-Wellen-Paneldatensatz  $P_{t_{1,3,4,5,6}}$  insgesamt systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Fünf-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_1$  und  $t_3$  in etwa neun Prozentpunkte unter denen der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von knapp sieben Prozentpunkten die Hauptschülerinnen und -schüler im Fünf-Wellen-Paneldatensatz ebenfalls deutlich unterrepräsentiert sind, sind zu  $t_5$  und  $t_6$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte bzw. keine Befragten besuchen zu diesen Befragungszeitpunkten noch die Hauptschule. Weiterhin sind im Fünf-Wellen-Paneldatensatz zu  $t_1$  und  $t_3$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik; die Abweichungen liegen hier bei etwa einem Prozentpunkt, was bedeutet, dass Gesamtschülerinnen und -schüler nur leicht unterrepräsentiert sind. Während sich auch zu  $t_4$  und  $t_6$  die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatensätzen  $Q_{t_4}$  und  $Q_{t_6}$  mit weniger als einem Prozentpunkt unterscheiden, sind die Gesamtschüler zu  $t_5$  im Fünf-Wellen-Paneldatensatz mit einem Unterschied von über viereinhalb Prozentpunkten im Gegensatz zu den Querschnittsdaten deutlich überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage sind stets die Querschnittsdatensätze der betreffenden Erhebungsjahre. So fällt zu  $t_5$  mit über neunzehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3). Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatensatz  $Q_{t_6}$  und im Fünf-Wellen-Paneldatensatz  $P_{t_{1,3,4,5,6}}$  genau gleich hoch.

Die Anteile der Gymnasiastinnen und Gymnasiasten liegen im Fünf-Wellen-Paneldatensatz zu den Zeitpunkten  $t_1$  und  $t_3$  etwa neun Prozentpunkte über denen der jeweiligen Schulstatistik. Im Anschluss wird diese Überrepräsentation mit knapp sieben Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatensätzen – zunächst geringer, um zu  $t_5$  auf vierzehn Prozentpunkte anzusteigen (zu den Besonderheiten des Querschnittsdatensatzes zu  $t_5$  siehe Abschnitt 5.1.3). Zu  $t_6$  sind die Gymnasiastinnen und Gymnasiasten verglichen mit dem Querschnittsdatensatz mit über vier Prozentpunkten wiederum in geringerem Ausmaß überrepräsentiert. Der Anteil der Realschülerinnen und -schüler liegt im Fünf-Wellen-Paneldatensatz im Zeitraum  $t_1$  und  $t_3$  um eineinhalb bzw. mehr als zweieinhalb Prozentpunkte höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit zweieinhalb Prozentpunkten überrepräsentiert sind, sind in den folgenden Wellen die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  und  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit weniger als einem Prozentpunkt im Fünf-Wellen-Paneldatensatz  $P_{t_{1,3,4,5,6}}$  – verglichen mit den jeweiligen Querschnittsdatensätzen – überrepräsentiert. Mit etwas weniger als einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im Panel sehr leicht unterrepräsentiert. Überdies zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Fünf-Wellen-Paneldatensatz mit über viereinhalb Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Fünf-Wellen-Paneldatensatz enthalten.

Tabelle 5.144 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Fünf-Wellen-Paneldatensatz  $P_{t_{1,3,4,5,6}}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).



5 Zusammensetzung der Paneldaten

**Tabelle 5.143:** Zusammensetzung  $P_{1,3,4,5,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(100 %)	
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,3	18,1	—	1,3	—	—	2003	
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728	
		$P_{1,3,4,5,6}$	18,3	19,7	32,5	29,4	—	—	—	—	557	
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635	
		$P_{1,3,4,5,6}$	19,0	21,1	31,9	28,0	—	—	—	—	558	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{1,3,4,5,6}$	18,0	21,3	32,3	27,5	0,5	—	0,4	—	560	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{1,3,4,5,6}$	1,3	1,1	21,2	30,3	42,3	—	2,2	1,7	542	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504	
		$P_{1,3,4,5,6}$	—	—	18,2	27,9	33,1	—	10,8	9,9	544	
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—	1907
			$Q_{t_1}$	17,9	24,1	31,8	26,3	—	0,0	—	—	1679
			$P_{1,3,4,5,6}$	12,9	26,0	30,2	30,9	—	—	—	—	854
		$t_3$	$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926
			$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704
			$P_{1,3,4,5,6}$	13,5	26,3	30,7	29,5	—	—	—	—	853
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{1,3,4,5,6}$	12,8	26,4	30,3	29,8	0,5	—	0,2	—	849	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{1,3,4,5,6}$	0,1	0,2	19,7	35,7	40,7	—	1,8	1,7	821	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715	
		$P_{1,3,4,5,6}$	—	—	17,5	32,2	29,6	—	11,6	9,1	828	
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
			$Q_{t_1}$	22,0	23,7	31,1	22,8	—	0,4	—	—	3407
			$P_{1,3,4,5,6}$	15,0	23,5	31,1	30,3	—	—	—	—	1411
		$t_3$	$S_{t_3}$	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
			$P_{1,3,4,5,6}$	15,7	24,2	31,2	28,9	—	—	—	—	1411
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{1,3,4,5,6}$	14,9	24,3	31,1	28,9	0,5	—	0,3	—	1409	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{1,3,4,5,6}$	0,6	0,6	20,3	33,5	41,3	—	2,0	1,7	1363	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
		$P_{1,3,4,5,6}$	—	—	17,8	30,5	31,0	—	11,3	9,4	1372	

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

Tabelle 5.144: Zusammensetzung  $P_{1,3,4,5,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform <sup>a</sup> (%)								gesamt	
		HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7
		$P_{1,3,4,5,6}$	48,1	33,1	41,2	38,3	—	—	—	—	39,5
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0
		$P_{1,3,4,5,6}$	48,0	34,5	40,5	38,2	—	—	—	—	39,5
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3
		$P_{1,3,4,5,6}$	48,1	34,7	41,3	37,8	42,9	—	50,0	—	39,7
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5
		$P_{1,3,4,5,6}$	87,5	75,0	41,5	35,9	40,7	—	44,4	39,1	39,8
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7
		$P_{1,3,4,5,6}$	—	—	40,6	36,3	42,4	—	38,1	41,9	39,7
weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	56,4	—	12,9	—	—	48,8
		$Q_{t_1}$	40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
		$P_{1,3,4,5,6}$	51,9	66,9	58,8	61,7	—	—	—	—	60,5
	$t_3$	$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1
		$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0
		$P_{1,3,4,5,6}$	52,0	65,5	59,5	61,8	—	—	—	—	60,5
	$t_4$	$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7
		$P_{1,3,4,5,6}$	51,9	65,3	58,7	62,2	57,1	—	50,0	—	60,3
	$t_5$	$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5
		$P_{1,3,4,5,6}$	12,5	25,0	58,5	64,1	59,3	—	55,6	60,9	60,2
	$t_6$	$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3
		$P_{1,3,4,5,6}$	—	—	59,4	63,7	57,6	—	61,9	58,1	60,3

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

## 5.6 Das Sechs-Wellen-Panel

Das Sechs-Wellen-Panel enthält alle Fälle, die bislang in allen sechs Erhebungswellen an der Befragung teilgenommen haben und deren – nach Plausibilitätskontrollen gültige – Fragebögen anhand der Codefragen sechsmal einander zuordenbar waren. Dieser Sechs-Wellen-Paneldatensatz ist das vorläufige Resultat aller bisher durchgeführten Zuordnungen und der umfassendste der bislang erstellten Paneldatensätze.

Das durchgehende Sechs-Wellen-Panel der Jahre 2002 bis 2007 enthält 1 307 Fälle. Aus Tabelle 5.145 geht hervor, dass im Sechs-Wellen-Paneldatensatz  $P_{t_1,2,3,4,5,6}$  hinsichtlich des Geschlechts der Befragten eine Verschiebung vorliegt:

**Tabelle 5.145:** Zusammensetzung nach Geschlecht  $P_{t_1,2,3,4,5,6}$

		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
	$P_{t_1,2,3,4,5,6}$	499	38,2	808	61,8	1307
$t_2$	$S_{t_2}$	1927	51,2	1836	48,8	3763
	$Q_{t_2}$	1703	50,2	1687	49,8	3390
	$P_{t_1,2,3,4,5,6}$	507	38,8	800	61,2	1307
$t_3$	$S_{t_3}$	1999	50,9	1926	49,1	3925
	$Q_{t_3}$	1635	49,0	1704	51,0	3339
	$P_{t_1,2,3,4,5,6}$	500	38,3	807	61,7	1307
$t_4$	$Q_{t_4}$	1717	50,4	1688	49,6	3405
	$P_{t_1,2,3,4,5,6}$	505	38,6	802	61,4	1307
$t_5$	$Q_{t_5}$	2265	49,8	2282	50,2	4547
	$P_{t_1,2,3,4,5,6}$	505	38,6	802	61,4	1307
$t_6$	$Q_{t_6}$	1567	47,0	1769	53,0	3336
	$P_{t_1,2,3,4,5,6}$	504	38,6	803	61,4	1307

Vergleicht man die Geschlechterverteilung der ersten drei Wellen im Sechs-Wellen-Paneldatensatz mit den Schulstatistiken der Erhebungswellen  $t_1$  bis  $t_3$ , so wird deutlich, dass der Anteil der Schülerinnen im Sechs-Wellen-Paneldatensatz in jedem der Jahre um etwa dreizehn Prozentpunkte höher liegt. Im Vergleich zu den entsprechenden Querschnittsdatsätzen  $Q_{t_1}$ ,  $Q_{t_2}$  und  $Q_{t_3}$  ist der Anteil der weiblichen Befragten im Sechs-Wellen-Paneldatensatz zwischen elf und zwölf Prozentpunkten erhöht. Die Gegenüberstellung der Geschlechterverteilung im Sechs-Wellen-Paneldatensatz mit den Querschnittsdatsätzen der Erhebungszeitpunkte  $t_4$ ,  $t_5$  und  $t_6$  (ein Vergleich mit der Schulstatistik ist hier nicht mehr sinnvoll möglich, siehe hierzu die Ausführungen in Abschnitt 4.1) ergibt, dass der Anteil der Schülerinnen im Sechs-Wellen-Paneldatensatz

zunächst über elf ( $t_4$  und  $t_5$ ), zu  $t_6$  über acht Prozentpunkte höher liegt. Damit wird deutlich, dass auch im Querschnittsdatensatz der letzten Erhebungswelle  $t_6$  weibliche Befragte bereits überrepräsentiert sind.

**Tabelle 5.146:** Altersdurchschnitt  $P_{t_1,2,3,4,5,6}$

		Alter	Std.-abw.	gesamt
$t_1$	$Q_{t_1}$	13,0	0,8	3 346
	$P_{t_1,2,3,4,5,6}$	12,9	0,7	1 297
$t_2$	$Q_{t_2}$	14,0	0,8	3 275
	$P_{t_1,2,3,4,5,6}$	13,9	0,7	1 293
$t_3$	$Q_{t_3}$	15,1	0,8	3 270
	$P_{t_1,2,3,4,5,6}$	14,9	0,7	1 300
$t_4$	$Q_{t_4}$	16,1	0,8	3 313
	$P_{t_1,2,3,4,5,6}$	15,9	0,7	1 295
$t_5$	$Q_{t_5}$	17,4	1,3	4 468
	$P_{t_1,2,3,4,5,6}$	16,9	0,7	1 300
$t_6$	$Q_{t_6}$	18,1	1,1	3 320
	$P_{t_1,2,3,4,5,6}$	17,9	0,7	1 303

Das durchschnittliche Alter der Befragten in den Paneldatensätzen kann auch in den früheren Erhebungswellen jeweils nur mit den Querschnittsdaten der betreffenden Erhebungsjahre gegenübergestellt werden: Informationen zum Alter sind in der Schulstatistik nicht enthalten. Ein Vergleich zwischen den Querschnittsdatensätzen und dem Sechs-Wellen-Paneldatensatz (siehe Tabelle 5.146) zeigt, dass das Durchschnittsalter der Befragten im Paneldatensatz durchgehend geringfügig unter dem der einzelnen Querschnittsdatensätze liegt. Am auffälligsten zeigt sich dieser Unterschied zum Erhebungszeitpunkt  $t_5$  (zu den Gründen siehe Kapitel 5.1.3). Zudem lässt sich festhalten, dass die Homogenität des Alters der Befragten im Panel im Vergleich zu den Querschnittsdatensätzen höher ist.

Tabelle 5.147 zeigt, dass der Sechs-Wellen-Paneldatensatz  $P_{t_1,2,3,4,5,6}$  insgesamt im Vergleich zur Schulstatistik der Erhebungszeitpunkte  $t_1$  bis  $t_3$  systematisch nach besuchter Schulform verzerrt ist: Die Anteile der Hauptschülerinnen und -schüler liegen im Sechs-Wellen-Paneldatensatz zu den Erhebungszeitpunkten  $t_1$  bis  $t_3$  zwischen knapp achteinhalb und elf Prozentpunkten unter denen der Schulstatistik. Ab dem Erhebungszeitpunkt  $t_4$  stehen nur noch die Querschnittsdatensätze der betreffenden Jahre für einen Anteilsvergleich zur Verfügung. Während zu  $t_4$  mit einer Abweichung von etwa sechs Prozentpunkten die Hauptschülerinnen und -schüler im Sechs-Wellen-Paneldatensatz ebenfalls deutlich unterrepräsentiert sind, sind zu  $t_5$  und  $t_6$  keine gesicherten Aussagen mehr möglich: Nur sehr wenige Befragte bzw. keine Befragten besuchen zu diesen Befra-

gungszeitpunkten noch die Hauptschule. Weiterhin sind im Sechs-Wellen-Paneldatensatz zu  $t_1$  und  $t_3$  ähnlich viele Gesamtschülerinnen und -schüler vertreten wie in der Schulstatistik; lediglich zu  $t_2$  sind diese im Sechs-Wellen-Paneldatensatz mit knapp zwei Prozentpunkten überrepräsentiert. Während sich zu  $t_4$  und  $t_6$  die Anteile der Gesamtschüler von den entsprechenden Querschnittsdatsätzen  $Q_{t_4}$  und  $Q_{t_6}$  ebenfalls mit weniger als einem Prozentpunkt – allerdings in abweichende Richtungen – unterscheiden, sind die Gesamtschüler zu  $t_5$  im Sechs-Wellen-Paneldatensatz mit einem Unterschied von über fünf Prozentpunkten im Gegensatz zu den Querschnittsdaten überrepräsentiert. Warum dieser Anteilsunterschied nicht überbewertet werden sollte, wird in Abschnitt 5.1.3 näher erläutert.

Zwar geben bereits zum Erhebungszeitpunkt  $t_4$  einige Befragte an, ein Berufskolleg zu besuchen, gesicherte Aussagen über diese Anteile können aus Gründen einer ausreichenden Fallzahl jedoch erst ab dem Erhebungsjahr  $t_5$  getroffen werden. Vergleichsgrundlage sind stets die Querschnittsdatsätze der betreffenden Erhebungsjahre. So fällt zu  $t_5$  mit über achtzehn Prozentpunkten eine deutliche Unterrepräsentation derjenigen Befragten im Panel auf, die ein Berufskolleg besuchen. Diese Diskrepanz ist jedoch in erster Linie dem in diesem Jahr abweichenden Vorgehen bei der Datenerhebung geschuldet (siehe genauer Abschnitt 5.1.3). Zum Erhebungszeitpunkt  $t_6$  sind die Anteile der Berufskollegiatinnen und -kollegiaten im Querschnittsdatsatz  $Q_{t_6}$  und im Sechs-Wellen-Paneldatensatz  $P_{t_{1,2,3,4,5,6}}$  ähnlich; mit einem Prozentpunkt sind Befragte, die ein Berufskolleg besuchen nur leicht im Sechs-Wellen-Paneldatensatz überrepräsentiert.

Die Anteile der Gymnasiastinnen und Gymnasiasten liegen im Sechs-Wellen-Paneldatensatz zu den Zeitpunkten  $t_1$ ,  $t_2$  und  $t_3$  etwa zwischen sechseinhalb und achteinhalb Prozentpunkten über denen der jeweiligen Schulstatistik. Im Anschluss wird diese Überrepräsentation mit mehr als viereinhalb Prozentpunkten zu  $t_4$  – jedoch diesmal im Vergleich zu den entsprechenden Querschnittsdatsätzen – zunächst geringer, um zu  $t_5$  auf über zwölf Prozentpunkte anzusteigen (zu den Besonderheiten des Querschnittsdatsatzes zu  $t_5$  siehe Abschnitt 5.1.3). Zu  $t_6$  sind die Gymnasiastinnen und Gymnasiasten verglichen mit dem Querschnittsdatsatz mit zweieinhalb Prozentpunkten wiederum in geringerem Ausmaß überrepräsentiert. Der Anteil der Realschülerinnen und -schüler liegt im Sechs-Wellen-Paneldatensatz im Zeitraum  $t_1$  bis  $t_3$  zwischen zwei und gut dreieinhalb Prozentpunkten höher als in der jeweiligen Schulstatistik dieser Erhebungsjahre. Während zu  $t_4$  im Vergleich zu den Querschnittsdaten  $Q_{t_4}$  die Realschülerinnen und -schüler ebenfalls mit etwas über drei Prozentpunkten überrepräsentiert sind, sind in den folgenden Wellen die Anteile zu gering, um noch Aussagen zur Repräsentativität zu treffen.

Zu  $t_5$  und  $t_6$  sind auch diejenigen Befragten, die eine anderweitige Schule besuchen, mit unter einem Prozentpunkt im Sechs-Wellen-Paneldatensatz  $P_{t_{1,2,3,4,5,6}}$  – verglichen mit den jeweiligen Querschnittsdatsätzen – überrepräsentiert. Mit etwas weniger als einem halben Prozentpunkt sind zu  $t_5$  diejenigen Befragten, die keine Schule besuchen, im

Panel sehr leicht unterrepräsentiert. Überdies zeigt sich zum letzten Erhebungszeitpunkt, dass diese Befragten im Sechs-Wellen-Paneldatensatz mit knapp fünf Prozentpunkten unterrepräsentiert sind. (Ehemalige) Sonderschülerinnen und -schüler sind nicht im Sechs-Wellen-Paneldatensatz enthalten.

Tabelle 5.148 verdeutlicht weiterhin, dass über fast alle Schulformen hinweg mehr Schülerinnen und weniger Schüler im Sechs-Wellen-Paneldatensatz  $P_{1,2,3,4,5,6}$  enthalten sind als in den Schulstatistiken bzw. Querschnittsdaten der einzelnen Erhebungsjahre (Schulformen, die aus Altersgründen in bestimmten Befragungsjahren nur von sehr wenigen Befragten besucht wurden, sind in den entsprechenden Jahren von dieser Betrachtung ausgenommen).

Tabelle 5.147: Zusammensetzung  $P_{t_1,2,3,4,5,6}$  nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform <sup>a</sup> (%)								gesamt		
		HS	RS	GS	GY	BK	SO	andere	keine	(100 %)		
männlich	$t_1$	$S_{t_1}$	28,0	21,0	31,3	18,1	—	1,3	—	—	2003	
		$Q_{t_1}$	25,9	23,3	30,5	19,5	—	0,9	—	—	1728	
		$P_{1,2,3,4,5,6}$	19,0	20,4	34,9	25,7	—	—	—	—	499	
	$t_2$	$S_{t_2}$	31,8	21,8	29,9	15,1	—	1,3	—	—	1927	
		$Q_{t_2}$	29,0	22,3	31,4	16,6	—	0,6	—	—	1703	
		$P_{1,2,3,4,5,6}$	19,9	20,5	34,7	24,9	—	—	—	—	507	
	$t_3$	$S_{t_3}$	28,2	20,9	32,7	16,3	—	2,0	—	—	1999	
		$Q_{t_3}$	26,0	22,3	32,4	18,7	—	0,7	—	—	1635	
		$P_{1,2,3,4,5,6}$	19,8	21,6	34,2	24,4	—	—	—	—	500	
	$t_4$	$Q_{t_4}$	25,8	21,2	32,2	18,4	0,9	0,7	0,8	0,0	1704	
		$P_{1,2,3,4,5,6}$	18,7	21,5	34,4	24,5	0,6	—	0,4	—	503	
	$t_5$	$Q_{t_5}$	0,5	0,3	15,0	16,1	64,1	0,2	1,3	2,5	2181	
		$P_{1,2,3,4,5,6}$	1,4	1,0	22,4	27,4	43,8	—	2,3	1,6	486	
	$t_6$	$Q_{t_6}$	—	—	17,6	22,7	32,9	—	11,0	15,8	1504	
		$P_{1,2,3,4,5,6}$	—	—	18,9	25,1	34,9	—	10,9	10,3	487	
	weiblich	$t_1$	$S_{t_1}$	19,6	23,0	32,6	24,6	—	0,2	—	—	1907
			$Q_{t_1}$	17,9	24,1	31,8	26,3	—	0,0	—	—	1679
			$P_{1,2,3,4,5,6}$	13,4	27,1	30,4	29,1	—	—	—	—	808
$t_2$		$S_{t_2}$	23,1	23,5	30,9	22,4	—	0,0	—	—	1836	
		$Q_{t_2}$	20,1	24,9	31,5	23,5	—	0,0	—	—	1687	
		$P_{1,2,3,4,5,6}$	13,6	27,3	30,6	28,5	—	—	—	—	800	
$t_3$		$S_{t_3}$	21,5	22,2	31,9	23,8	—	0,6	—	—	1926	
		$Q_{t_3}$	19,3	23,5	31,3	25,6	—	0,3	—	—	1704	
		$P_{1,2,3,4,5,6}$	13,9	27,4	31,0	27,8	—	—	—	—	807	
$t_4$		$Q_{t_4}$	17,5	22,4	33,1	25,6	0,7	0,1	0,4	0,2	1682	
		$P_{1,2,3,4,5,6}$	13,5	27,6	30,3	28,2	0,4	—	0,1	—	802	
$t_5$		$Q_{t_5}$	0,2	0,3	16,3	22,8	57,1	0,0	1,7	1,5	2226	
		$P_{1,2,3,4,5,6}$	0,1	0,3	20,3	34,4	41,4	—	1,9	1,5	777	
$t_6$		$Q_{t_6}$	—	—	18,3	29,3	29,3	—	10,4	12,7	1715	
		$P_{1,2,3,4,5,6}$	—	—	18,1	30,9	30,3	—	11,7	8,9	783	
gesamt		$t_1$	$S_{t_1}$	23,9	22,0	32,0	21,3	—	0,8	—	—	3910
			$Q_{t_1}$	22,0	23,7	31,1	22,8	—	0,4	—	—	3407
			$P_{1,2,3,4,5,6}$	15,5	24,6	32,1	27,8	—	—	—	—	1307
	$t_2$	$S_{t_2}$	27,5	22,7	30,4	18,7	—	0,7	—	—	3763	
		$Q_{t_2}$	24,6	23,6	31,4	20,1	—	0,3	—	—	3390	
		$P_{1,2,3,4,5,6}$	16,1	24,6	32,2	27,1	—	—	—	—	1307	
	$t_3$	$S_{t_3}$	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	
		$P_{1,2,3,4,5,6}$	16,1	25,2	32,2	26,5	—	—	—	—	1307	
	$t_4$	$Q_{t_4}$	21,7	21,8	32,7	22,0	0,8	0,4	0,6	0,1	3386	
		$P_{1,2,3,4,5,6}$	15,5	25,2	31,9	26,7	0,5	—	0,2	—	1305	
	$t_5$	$Q_{t_5}$	0,3	0,3	15,7	19,5	60,5	0,1	1,5	2,0	4407	
		$P_{1,2,3,4,5,6}$	0,6	0,6	21,1	31,7	42,4	—	2,1	1,6	1263	
	$t_6$	$Q_{t_6}$	—	—	18,0	26,2	31,0	—	10,7	14,1	3219	
		$P_{1,2,3,4,5,6}$	—	—	18,4	28,7	32,0	—	11,4	9,4	1270	

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

## 5.6 Das Sechs-Wellen-Panel

**Tabelle 5.148:** Zusammensetzung  $P_{t_1,2,3,4,5,6}$  nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform <sup>a</sup> (%)							gesamt		
			HS	RS	GS	GY	BK	SO	andere	keine	(%)	
männlich	$t_1$	$S_{t_1}$	60,1	48,9	50,4	43,6	—	87,1	—	—	51,2	
		$Q_{t_1}$	59,8	49,9	49,7	43,3	—	100,0	—	—	50,7	
		$P_{t_1,2,3,4,5,6}$	46,8	31,8	41,4	35,3	—	—	—	—	38,2	
	$t_2$	$S_{t_2}$	59,1	49,4	50,4	41,4	—	100,0	—	—	51,2	
		$Q_{t_2}$	59,3	47,5	50,2	41,6	—	100,0	—	—	50,2	
		$P_{t_1,2,3,4,5,6}$	48,1	32,3	41,8	35,6	—	—	—	—	38,8	
	$t_3$	$S_{t_3}$	57,7	49,3	51,5	41,6	—	76,5	—	—	50,9	
		$Q_{t_3}$	56,4	47,6	49,8	41,2	—	68,8	—	—	49,0	
		$P_{t_1,2,3,4,5,6}$	46,9	32,8	40,6	35,3	—	—	—	—	38,3	
	$t_4$	$Q_{t_4}$	59,8	49,1	49,6	42,2	55,6	92,3	65,0	0,0	50,3	
		$P_{t_1,2,3,4,5,6}$	46,5	32,8	41,6	35,2	50,0	—	66,7	—	38,5	
	$t_5$	$Q_{t_5}$	73,3	50,0	47,4	40,9	52,4	80,0	42,4	61,8	49,5	
		$P_{t_1,2,3,4,5,6}$	87,5	71,4	40,8	33,3	39,8	—	42,3	40,0	38,5	
	$t_6$	$Q_{t_6}$	—	—	45,8	40,5	49,6	—	48,1	52,1	46,7	
		$P_{t_1,2,3,4,5,6}$	—	—	39,3	33,5	41,8	—	36,6	41,7	38,3	
	weiblich	$t_1$	$S_{t_1}$	39,9	51,1	49,6	56,4	—	12,9	—	—	48,8
			$Q_{t_1}$	40,2	50,1	50,3	56,7	—	0,0	—	—	49,3
			$P_{t_1,2,3,4,5,6}$	53,2	68,2	58,6	64,7	—	—	—	—	61,8
$t_2$		$S_{t_2}$	40,9	50,6	49,6	58,6	—	0,0	—	—	48,8	
		$Q_{t_2}$	40,7	52,5	49,8	58,4	—	0,0	—	—	49,8	
		$P_{t_1,2,3,4,5,6}$	51,9	67,7	58,2	64,4	—	—	—	—	61,2	
$t_3$		$S_{t_3}$	42,3	50,7	48,5	58,4	—	23,5	—	—	49,1	
		$Q_{t_3}$	43,6	52,4	50,2	58,8	—	31,3	—	—	51,0	
		$P_{t_1,2,3,4,5,6}$	53,1	67,2	59,4	64,7	—	—	—	—	61,7	
$t_4$		$Q_{t_4}$	40,2	50,9	50,4	57,8	44,4	7,7	35,0	100,0	49,7	
		$P_{t_1,2,3,4,5,6}$	53,5	67,2	58,4	64,8	50,0	—	33,3	—	61,5	
$t_5$		$Q_{t_5}$	26,7	50,0	52,6	59,1	47,6	20,0	57,6	38,2	50,5	
		$P_{t_1,2,3,4,5,6}$	12,5	28,6	59,2	66,8	60,2	—	57,7	60,0	61,5	
$t_6$		$Q_{t_6}$	—	—	54,2	59,5	50,4	—	51,9	47,9	53,3	
		$P_{t_1,2,3,4,5,6}$	—	—	60,7	66,5	58,2	—	63,4	58,3	61,7	

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





## 6 Zusammenfassung und Erklärung

Abschließend kann festgehalten werden, dass die Paneldatensätze hinsichtlich ihrer Zusammensetzung nach Geschlecht (teilweise leicht, teilweise erheblich) verzerrt sind. In sämtlichen Paneldatensätzen finden sich im Vergleich zur Schulstatistik und zu den Querschnittsdaten mehr weibliche als männliche Befragte.

Auch bei Betrachtung der besuchten Schulform fallen systematische Verzerrungen auf: Der Anteil der Hauptschülerinnen und -schüler ist zwischen  $t_1$  und  $t_4$  in den Paneldatensätzen durchweg geringer als in der Schulstatistik bzw. in den Querschnittsdaten, der Anteil der Gymnasiastinnen und Gymnasiasten durchgehend größer. Ein ebenfalls erhöhter Anteil fällt – allerdings in geringerem Ausmaß – auch in Bezug auf die Realschülerinnen und -schüler auf. Der Anteil der Gesamtschülerinnen und -schüler liegt in den Panel- und Querschnittsdatensätzen – sowie zu  $t_1$ ,  $t_2$  und  $t_3$  auch in den Schulstatistiken – meist auf vergleichbarem Niveau. Während die Berufsschülerinnen und -schüler zum Zeitpunkt  $t_5$  im Panel – verglichen mit den Querschnittsdaten – deutlich unterrepräsentiert sind, unterscheiden sich die Anteile zu  $t_6$  in Panel- und Querschnittsdaten kaum.

Der Abgleich von Querschnitts- und Paneldaten sowie die Interpretation der Anteilsunterschiede sollte zum Zeitpunkt  $t_5$  stets vor dem Hintergrund der stark veränderten Vorgehensweise bei der Stichprobenauswahl geschehen. Eine Beurteilung der Repräsentativität der Paneldatensätze ist auf Basis des Querschnittsdatensatzes der Erhebungswelle  $t_5$  streng genommen nicht möglich.

Abweichungen von den Schulstatistiken – die hier nur für die Erhebungszeitpunkte  $t_1$  bis  $t_3$  relevant sind – sind in geringerer Stärke, jedoch in gleicher Richtung, oftmals schon in den Querschnittsdaten zu erkennen. Tendenziell setzen sich die Verzerrungen der Querschnittsdaten im Vergleich zur Schulstatistik in den Paneldaten verstärkt fort. Dies lässt den Schluss zu, dass dieselben Gründe, die zu den beschriebenen Verschiebungen führten, auch (Mit-)Ursachen für die Panelverzerrungen sind. Man muss also annehmen, dass bei männlichen Befragten sowie bei Hauptschülerinnen und -schülern diese Ausfallgründe öfter auftreten als bei weiblichen Befragten, Gymnasiastinnen und Gymnasiasten sowie Realschülerinnen und -schülern. Es scheinen innerhalb dieser Personengruppen auch über die Jahre unterschiedliche Probanden wegen der genannten Ursachen auszufallen. Fielen immer dieselben Personen aus, würde der Trend im Panel nicht verstärkt.

Besonders hervorzuheben ist in diesem Zusammenhang auch der Querschnittsdatensatz  $t_6$ : Im Hinblick auf das Geschlecht der Befragten wird hier deutlich, dass – ohne

jedoch eine verlässliche offizielle Statistik hinzuziehen zu können – in diesem Querschnittsdatensatz offensichtlich eine stärkere Verschiebung der Geschlechteranteile vorliegt, als in den Querschnittsdatensätzen der vorangehenden Erhebungsjahre. Geht man zu  $t_6$  von einer ähnlichen *anvisierten* d. h. kontaktierten Stichprobe aus wie in den Vorjahren, so fällt auf, dass in diesem letzten Erhebungsjahr männliche Befragte noch etwas stärker unterrepräsentiert sind und daher offensichtlich in geringerem Maße zu einer Teilnahme bereit waren. Grund hierfür könnte die in  $t_6$  erfolgte, weitgehende Umstellung auf eine postalische Befragung sein: Wo in den Vorjahren durch die Paper-und-Pencil-Befragung im Klassenverband noch ein stärkerer Einfluss auf die Befragungsteilnahme möglich war, ist es im Zuge der postalischen Befragung und der damit einhergehenden privaten Interviewsituation nur sehr eingeschränkt möglich, zögerliche Befragte für eine Teilnahme zu gewinnen. Bei einer Entscheidung für oder gegen die Befragungsteilnahme ist dann offensichtlich ein oben bereits aufgeführtes Kriterium von Bedeutung: Weibliche Befragte scheinen eher bereit, sich an der Befragung zu beteiligen. Daher kommt es in der letzten hier einbezogenen Erhebungswelle bereits in den Querschnittsdaten zu dementsprechenden Selbstselektions-Effekten. Diese Tendenz schlägt sich auch in den Paneldatensätzen nieder: Sämtliche Teilpaneldatensätze, in denen die sechste Erhebungswelle enthalten ist, fallen durch ein etwas größeres Ausmaß der Verzerrung nach Geschlecht auf.

Eine weitere Ursache für die Verzerrungen der Paneldatensätze liegt in der Art der Zuordnung der Fragebögen aus den verschiedenen Erhebungsjahren. Das oben erläuterte Zuordnungsverfahren weist den Nachteil auf, dass das gleiche Ausfüllen der Codefragen unterschiedlichen Personengruppen unterschiedlich starke Probleme bereitete. Wie Pöge (2007, S. 82 f.) hervorhebt, existiert hier ein deutlicher Zusammenhang zwischen der Fähigkeit, den Code zu reproduzieren und dem Geschlecht sowie dem Bildungsniveau der Befragten. Dieser Zuordnungseffekt verstärkt sich nun mit jeder Erhebungswelle, um die der Paneldatensatz erweitert wird. Dies führt dazu, dass die Verzerrungen mit dem Fortschreiten der Konstruktion des Paneldatensatzes zunehmen.

## Literaturverzeichnis

- Bentrup, C. (2007). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2006*. (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 12). Münster, Bielefeld.
- Bentrup, C. (2009). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2007*. (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 15). Münster, 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: Waxmann.
- Brondies, M. (2004a). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2003*. (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 6). Münster, Trier.
- Brondies, M. (2004b). *Methodendokumentation der Lehrerbefragung an Münsteraner und Duisburger Schulen 2003. Erhebung durchgeführter Präventionsmaßnahmen* (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 8). Münster, Trier.
- Engel, U. & Reinecke, J. (1994). *Panelanalyse. Grundlagen–Techniken–Beispiele*. New York: Walter de Gruyter.
- Hilfert, N. (2005). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2004*. (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 10). Münster, Trier.
- Kunadt, S. (2006). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2005*. (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 11). Münster, Bielefeld.
- Motzke, K. & Brondies, M. (2004). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2002*. (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 5). Münster, Trier.
- Motzke, K. & Wittenberg, J. (2004). *Methodendokumentation der kriminologischen Schülerbefragung in Münster 2000*. (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 1). Münster, 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, 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, Bielefeld.
- Pöge, A. (2008). Persönliche Codes „reloaded“. *Methoden–Daten–Analysen*, 2 (1), 59–70.
- 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, Bielefeld.
- Wittenberg, J. (2004a). *Methodendokumentation der kriminologischen Schülerbefragung in Münster 2001*. (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 2). Münster, Trier.
- Wittenberg, J. (2004b). *Methodendokumentation der kriminologischen Schülerbefragung in Münster 2002*. (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 4). Münster, Trier.
- Wittenberg, J. (2004c). *Methodendokumentation der kriminologischen Schülerbefragung in Münster 2003*. (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 7). Münster, Trier.
- Wittenberg, J. & Hilfert, N. (2004). *Methodendokumentation der kriminologischen Schülerbefragung in Bocholt 2001*. (Schriftenreihe „Jugendkriminalität in der modernen Stadt – Methoden“ Nr. 3). Münster, Trier.

# Anhang

## Erstellung des persönlichen Codes

Liebe Schülerin, lieber Schüler,

da wir Ihren Fragebogen dem des letzten Jahres zuordnen wollen, bitten wir Sie, Ihren persönlichen Code zu wiederholen.

Da wir letztes Jahr den Code getrennt vom Fragebogen eingegeben, mit einem Computerprogramm verschlüsselt und danach vernichtet haben, ist es wichtig, dass Sie sich an Ihren Code erinnern. Denn nur wenn wir mit diesem Code denselben Vorgang wiederholen, können Ihre Fragebögen einander zugeordnet werden und dies, ohne dass jemand herausfinden kann, wer diesen Fragebogen ausgefüllt hat. Wichtig ist also, dass Sie denselben Code noch wissen. Aus diesem Grund haben wir letztes Jahr die nachfolgenden Fragen formuliert, die Ihnen helfen sollen, sich an dasselbe Wort 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 einen Vater am nächsten kommt) an. (z. B. <input checked="" type="checkbox"/> Anton, <input checked="" type="checkbox"/> Bernard, <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 eine 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 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 checked="" 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

Haben Sie im letzten Jahr an der Befragung teilgenommen?  ja  nein  
Haben Sie im letzten Jahr die Schule gewechselt?  ja  nein  
Sind Sie im letzten Jahr sitzen geblieben?  ja  nein  
In welcher Klasse sind Sie?  10  11

Abbildung A.1: Codeblatt 2006

### Erstellung des persönlichen Codes

Liebe Schülerin, lieber Schüler,

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 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 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.). <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 <input type="checkbox"/> , Glatzk. <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 11. oder 12. Jahrgang eines Gymnasiums oder einer Gesamtschule?  ja  nein  
**Wenn ja:** Sind Sie im letzten Jahr sitzen geblieben?  ja  nein

Abbildung A.2: Codeblatt 2007

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

	2006		2007	
	Hfk	%	Hfk	%
a	386	8,5	277	8,3
ä	2	0,0	—	—
b	138	3,0	101	3,0
c	94	2,1	63	1,9
d	185	4,1	148	4,4
e	131	2,9	104	3,1
f	171	3,8	142	4,3
g	199	4,4	139	4,2
h	510	11,2	377	11,3
i	81	1,8	59	1,8
j	279	6,1	187	5,6
k	236	5,2	159	4,8
l	53	1,2	39	1,2
m	465	10,2	332	10,0
n	106	2,3	86	2,6
o	44	1,0	40	1,2
ö	10	0,2	6	0,2
p	159	3,5	113	3,4
q	—	—	1	0,0
r	357	7,8	262	7,9
s	175	3,8	138	4,1
ß	4	0,1	3	0,1
t	118	2,6	110	3,3
u	169	3,7	132	4,0
ü	7	0,2	4	0,1
v	105	2,3	64	1,9
w	229	5,0	172	5,2
x	5	0,1	1	0,0
y	43	0,9	31	0,9
z	28	0,6	17	0,5
fehlend	59	1,3	29	0,9
gesamt	4 548	100,0	3 336	100,0



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

	2006		2007	
	Hfk	%	Hfk	%
a	413	9,1	297	8,9
ä	—	—	1	0,0
b	358	7,9	282	8,5
c	247	5,4	180	5,4
d	123	2,7	100	3,0
e	278	6,1	202	6,1
f	123	2,7	88	2,6
g	262	5,8	189	5,7
h	314	6,9	257	7,7
i	170	3,7	132	4,0
j	61	1,3	47	1,4
k	163	3,6	129	3,9
l	91	2,0	63	1,9
m	518	11,4	365	10,9
n	174	3,8	126	3,8
o	24	0,5	15	0,4
ö	3	0,1	3	0,1
p	149	3,3	98	2,9
r	183	4,0	145	4,3
s	464	10,2	341	10,2
ß	4	0,1	1	0,0
t	51	1,1	36	1,1
u	169	3,7	114	3,4
ü	7	0,2	4	0,1
v	54	1,2	38	1,1
w	26	0,6	11	0,3
x	3	0,1	2	0,1
y	26	0,6	14	0,4
z	47	1,0	31	0,9
fehlend	43	0,9	25	0,7
gesamt	4548	100,0	3336	100,0

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

	2006		2007	
	Hfk	%	Hfk	%
a	401	8,8	288	8,6
ä	2	0,0	—	—
b	130	2,9	84	2,5
c	246	5,4	180	5,4
d	352	7,7	261	7,8
e	128	2,8	100	3,0
f	132	2,9	86	2,6
g	66	1,5	42	1,3
h	125	2,7	97	2,9
i	76	1,7	46	1,4
j	346	7,6	279	8,4
k	218	4,8	165	4,9
l	142	3,1	119	3,6
m	571	12,6	409	12,3
n	200	4,4	148	4,4
o	45	1,0	27	0,8
ö	23	0,5	13	0,4
p	150	3,3	105	3,1
q	4	0,1	2	0,1
r	133	2,9	99	3,0
s	600	13,2	448	13,4
ß	1	0,0	1	0,0
t	201	4,4	144	4,3
u	14	0,3	9	0,3
ü	4	0,1	2	0,1
v	86	1,9	72	2,2
w	15	0,3	12	0,4
x	5	0,1	1	0,0
y	67	1,5	50	1,5
z	20	0,4	16	0,5
fehlend	46	1,0	31	0,9
gesamt	4 548	100,0	3 336	100,0

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

	2006		2007	
	Hfk	%	Hfk	%
1	146	3,2	114	3,4
2	128	2,8	81	2,4
3	151	3,3	99	3,0
4	135	3,0	103	3,1
5	159	3,5	105	3,1
6	153	3,4	95	2,8
7	164	3,6	100	3,0
8	153	3,4	119	3,6
9	150	3,3	136	4,1
10	174	3,8	130	3,9
11	149	3,3	109	3,3
12	131	2,9	118	3,5
13	146	3,2	105	3,1
14	145	3,2	104	3,1
15	146	3,2	119	3,6
16	153	3,4	120	3,6
17	145	3,2	113	3,4
18	161	3,5	126	3,8
19	153	3,4	122	3,7
20	138	3,0	98	2,9
21	140	3,1	100	3,0
22	141	3,1	103	3,1
23	147	3,2	104	3,1
24	139	3,1	97	2,9
25	150	3,3	99	3,0
26	161	3,5	123	3,7
27	130	2,9	87	2,6
28	131	2,9	96	2,9
29	128	2,8	101	3,0
30	132	2,9	95	2,8
31	83	1,8	66	2,0
fehlend	86	1,9	49	1,5
gesamt	4548	100,0	3336	100,0

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

	2006		2007	
	Hfk	%	Hfk	%
a	1	0,0	1	0,0
ä	—	—	1	0,0
b	172	3,8	145	4,3
c	1	0,0	—	—
d	1458	32,1	1094	32,8
e	23	0,5	11	0,3
f	1	0,0	—	—
g	8	0,2	1	0,0
h	3	0,1	1	0,0
i	—	—	2	0,1
j	2	0,0	3	0,1
k	5	0,1	1	0,0
l	5	0,1	2	0,1
m	4	0,1	4	0,1
n	2133	46,9	1603	48,1
o	1	0,0	—	—
ö	3	0,1	—	—
p	1	0,0	—	—
q	1	0,0	1	0,0
r	12	0,3	3	0,1
s	49	1,1	26	0,8
ß	2	0,0	2	0,1
t	45	1,0	30	0,9
u	6	0,1	2	0,1
ü	1	0,0	—	—
v	—	—	—	—
w	1	0,0	2	0,1
x	—	—	—	—
y	—	—	—	—
z	564	12,5	366	11,0
fehlend	46	1,0	35	1,0
Gesamt	4548	100,0	3336	100,0

Tabelle A.6: Häufigkeiten von Co006 (Augenfarbe)

	2006		2007	
	Hfk	%	Hfk	%
a	7	0,2	—	—
ä	1	0,0	—	—
b	127	2,8	81	2,4
c	—	—	—	—
d	2	0,0	4	0,1
e	2	0,0	1	0,0
f	3	0,1	1	0,0
g	39	0,9	20	0,6
h	1	0,0	2	0,1
i	1	0,0	3	0,1
j	1	0,0	3	0,1
k	1	0,0	—	—
l	7	0,2	—	—
m	6	0,1	6	0,2
n	2918	64,2	2126	63,7
o	1	0,0	—	—
ö	1	0,0	—	—
p	1	0,0	—	—
q	—	—	—	—
r	2	0,0	—	—
s	7	0,2	4	0,1
ß	1	0,0	1	0,0
t	1	0,0	1	0,0
u	1313	28,9	1007	30,2
ü	1	0,0	—	—
v	—	—	3	0,1
w	1	0,0	—	—
x	—	—	1	0,0
y	3	0,1	—	—
z	48	1,1	33	1,0
fehlend	52	1,1	39	1,2
Gesamt	4548	100,0	3336	100,0

**Tabelle A.7:** Häufigkeiten von Co007 2006 (Befragungsteilnahme im Vorjahr ja/nein)

		Hfk	%	gültige %
gültig	ja	2 805	61,7	64,6
	nein	1 537	33,8	35,4
	gesamt	4 342	95,5	100,0
k. A.		206	4,5	
gesamt		4 548	100,0	

**Tabelle A.8:** Häufigkeiten von Co007 2007 (Befragungsteilnahme im Vorjahr ja/nein)

		Hfk	%	gültige %
gültig	ja	2 589	77,6	86,0
	nein	420	12,6	14,0
	gesamt	3 009	90,2	100,0
k. A.		327	9,8	
gesamt		3 336	100,0	

**Tabelle A.9:** Häufigkeiten von Co009 2006 (Sitzengeblieben ja/nein)

		Hfk	%	gültige %
gültig	ja	157	3,5	7,1
	nein	2 045	45,0	92,9
	gesamt	2 202	48,5	100,0
k. A.		2 346	51,6	
gesamt		4 548	100,0	

**Tabelle A.10:** Häufigkeiten von Co009 2007 (Sitzengeblieben ja/nein)

		Hfk	%	gültige %
gültig	ja	273	8,2	12,5
	nein	1 907	57,2	87,5
	gesamt	2 180	65,4	100,0
k. A.		1 156	34,7	
gesamt		3 336	100,0	

**Tabelle A.11:** Häufigkeiten der Gesamtcodes 2006

		Hfk	%	gültige %
gültig	1	4457	98,0	98,5
	2	66	1,5	1,5
	gesamt	4523	99,5	100,0
k. A.		25	0,5	
gesamt		4548	100,0	

**Tabelle A.12:** Häufigkeiten der Gesamtcodes 2007

		Hfk	%	gültige %
gültig	1	3253	97,5	98,0
	2	68	2,0	2,0
	gesamt	3321	99,4	100,0
k. A.		15	0,4	
gesamt		3336	100,0	