



WESTFÄLISCHE
WILHELMS-UNIVERSITÄT
MÜNSTER



Universität Bielefeld

**Methodendokumentation
der kriminologischen Schülerbefragung
in Duisburg 2002 bis 2005**

– Vier-Wellen-Panel –

Andreas Pöge

**Schriftenreihe:
Jugendkriminalität in der modernen Stadt – Methoden
Nr. 13 / 2007**

ISSN 1610-2819

Informationen zur Schriftenreihe
Jugendkriminalität in der modernen Stadt – Methoden

Herausgeber:

Prof. Dr. Klaus Boers

Institut für Kriminalwissenschaften,
Abteilung Kriminologie,
Bispinghof 24/25
48143 Münster

und

Prof. Dr. Jost Reinecke

Fakultät für Soziologie,
Universität Bielefeld,
Postfach 10 01 31,
33501 Bielefeld

Internet-Adressen

<http://www.jura.uni-muenster.de/go/organisation/institute/strafrecht/kr4.html>

<http://www.uni-bielefeld.de/soz/personen/reinecke.html>

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).**

Inhaltsverzeichnis

1	Einleitung	1
2	Studiendesign	3
3	Erstellung der Paneldaten	5
4	Realisierte Zuordnungen	11
4.1	Die verwendeten Datensätze und ihre Kurzbezeichnungen	11
4.2	Abschätzung und Bewertung der Zuordnungsgüte	14
5	Zusammensetzung der Paneldaten	19
5.1	Die Panelquerschnitte	19
5.1.1	Panelquerschnitt 2002 ($P_{Q_{t_1}}$)	19
5.1.2	Panelquerschnitt 2003 ($P_{Q_{t_2}}$)	21
5.1.3	Panelquerschnitt 2004 ($P_{Q_{t_3}}$)	23
5.1.4	Panelquerschnitt 2005 ($P_{Q_{t_4}}^*$ und $P_{Q_{t_4}}$)	25
5.2	Die Zwei-Wellen-Panel	29
5.2.1	Zwei-Wellen-Panel 2002/2003 ($P_{t_{1,2}}$)	29
5.2.2	Zwei-Wellen-Panel 2003/2004 ($P_{t_{2,3}}$)	31
5.2.3	Zwei-Wellen-Panel 2004/2005 ($P_{t_{3,4}}^*$ und $P_{t_{3,4}}$)	33
5.2.4	Zwei-Wellen-Panel 2002/2004 ($P_{t_{1,3}}$)	39
5.2.5	Zwei-Wellen-Panel 2002/2005 ($P_{t_{1,4}}^*$ und $P_{t_{1,4}}$)	41
5.2.6	Zwei-Wellen-Panel 2003/2005 ($P_{t_{2,4}}^*$ und $P_{t_{2,4}}$)	46
5.3	Die Drei-Wellen-Panel	51
5.3.1	Drei-Wellen-Panel 2002 bis 2004 ($P_{t_{1,2,3}}$)	51
5.3.2	Drei-Wellen-Panel 2003 bis 2005 ($P_{t_{2,3,4}}^*$ und $P_{t_{2,3,4}}$)	54
5.3.3	Drei-Wellen-Panel 2002, 2003, 2005 ($P_{t_{1,2,4}}^*$ und $P_{t_{1,2,4}}$)	60
5.3.4	Drei-Wellen-Panel 2002, 2004, 2005 ($P_{t_{1,3,4}}^*$ und $P_{t_{1,3,4}}$)	66
5.4	Das Vier-Wellen-Panel 2002 bis 2005 ($P_{t_{1,2,3,4}}^*$ und $P_{t_{1,2,3,4}}$)	72
6	Zusammenfassung und Erklärung	81
6.1	Individualebene	81
6.2	Schulebene	83
	Literaturverzeichnis	85

A	Anhang	87
A.1	Codeblätter	87
A.2	Häufigkeiten der Codevariablen	92

Tabellenverzeichnis

3.1	Anzahl der handschriftlichen Kontrollen und deren Ergebnisse (vor Plausibilitätskontrollen)	7
3.2	Kontrollschritte Duisburg	9
4.1	Bezeichnungen und Fallzahlen der Duisburger Datensätze 2002 bis 2005	13
4.2	Zuordnungsgüte der Paneldaten in Duisburg und Münster im Vergleich .	16
5.1	Zusammensetzung nach Geschlecht $P_{Q_{t_1}}$	20
5.2	Altersdurchschnitt $P_{Q_{t_1}}$	20
5.3	Zusammensetzung $P_{Q_{t_1}}$ nach Geschlecht und Schulform (Zeilenproz.) .	21
5.4	Zusammensetzung $P_{Q_{t_1}}$ nach Geschlecht und Schulform (Spaltenproz.) .	21
5.5	Zusammensetzung nach Geschlecht $P_{Q_{t_2}}$	22
5.6	Altersdurchschnitt $P_{Q_{t_2}}$	22
5.7	Zusammensetzung $P_{Q_{t_2}}$ nach Geschlecht und Schulform (Zeilenproz.) .	22
5.8	Zusammensetzung $P_{Q_{t_2}}$ nach Geschlecht und Schulform (Spaltenproz.) .	23
5.9	Zusammensetzung nach Geschlecht $P_{Q_{t_3}}$	23
5.10	Altersdurchschnitt $P_{Q_{t_3}}$	24
5.11	Zusammensetzung $P_{Q_{t_3}}$ nach Geschlecht und Schulform (Zeilenproz.) .	24
5.12	Zusammensetzung $P_{Q_{t_3}}$ nach Geschlecht und Schulform (Spaltenproz.) .	25
5.13	Zusammensetzung nach Geschlecht $P_{Q_{t_4}}^*$	25
5.14	Altersdurchschnitt $P_{Q_{t_4}}^*$	26
5.15	Zusammensetzung $P_{Q_{t_4}}^*$ nach Geschlecht und Schulform (Zeilenproz.) .	26
5.16	Zusammensetzung $P_{Q_{t_4}}^*$ nach Geschlecht und Schulform (Spaltenproz.) .	27
5.17	Zusammensetzung nach Geschlecht $P_{Q_{t_4}}$	27
5.18	Altersdurchschnitt $P_{Q_{t_4}}$	28
5.19	Zusammensetzung $P_{Q_{t_4}}$ nach Geschlecht und Schulform (Zeilenproz.) .	28
5.20	Zusammensetzung $P_{Q_{t_4}}$ nach Geschlecht und Schulform (Spaltenproz.) .	28
5.21	Zusammensetzung nach Geschlecht $P_{I_{1,2}}$	29
5.22	Altersdurchschnitt $P_{I_{1,2}}$	30
5.23	Zusammensetzung $P_{I_{1,2}}$ nach Geschlecht und Schulform (Zeilenproz.) .	30
5.24	Zusammensetzung $P_{I_{1,2}}$ nach Geschlecht und Schulform (Spaltenproz.) .	31
5.25	Zusammensetzung nach Geschlecht $P_{I_{2,3}}$	31
5.26	Altersdurchschnitt $P_{I_{2,3}}$	32
5.27	Zusammensetzung $P_{I_{2,3}}$ nach Geschlecht und Schulform (Zeilenproz.) .	32

5.28 Zusammensetzung $P_{t_{2,3}}$ nach Geschlecht und Schulform (Spaltenproz.) .	33
5.29 Zusammensetzung nach Geschlecht $P_{t_{3,4}}^*$	34
5.30 Altersdurchschnitt $P_{t_{3,4}}^*$	34
5.31 Zusammensetzung $P_{t_{3,4}}^*$ nach Geschlecht und Schulform (Zeilenproz.) .	35
5.32 Zusammensetzung $P_{t_{3,4}}^*$ nach Geschlecht und Schulform (Spaltenproz.) .	36
5.33 Zusammensetzung nach Geschlecht $P_{t_{3,4}}$	36
5.34 Altersdurchschnitt $P_{t_{3,4}}$	37
5.35 Zusammensetzung $P_{t_{3,4}}$ nach Geschlecht und Schulform (Zeilenproz.) .	38
5.36 Zusammensetzung $P_{t_{3,4}}$ nach Geschlecht und Schulform (Spaltenproz.) .	38
5.37 Zusammensetzung nach Geschlecht $P_{t_{1,3}}$	39
5.38 Altersdurchschnitt $P_{t_{1,3}}$	39
5.39 Zusammensetzung $P_{t_{1,3}}$ nach Geschlecht und Schulform (Zeilenproz.) .	40
5.40 Zusammensetzung $P_{t_{1,3}}$ nach Geschlecht und Schulform (Spaltenproz.) .	41
5.41 Zusammensetzung nach Geschlecht $P_{t_{1,4}}^*$	41
5.42 Altersdurchschnitt $P_{t_{1,4}}^*$	42
5.43 Zusammensetzung $P_{t_{1,4}}^*$ nach Geschlecht und Schulform (Zeilenproz.) .	42
5.44 Zusammensetzung $P_{t_{1,4}}^*$ nach Geschlecht und Schulform (Spaltenproz.) .	43
5.45 Zusammensetzung nach Geschlecht $P_{t_{1,4}}$	44
5.46 Altersdurchschnitt $P_{t_{1,4}}$	44
5.47 Zusammensetzung $P_{t_{1,4}}$ nach Geschlecht und Schulform (Zeilenproz.) .	45
5.48 Zusammensetzung $P_{t_{1,4}}$ nach Geschlecht und Schulform (Spaltenproz.) .	46
5.49 Zusammensetzung nach Geschlecht $P_{t_{2,4}}^*$	46
5.50 Altersdurchschnitt $P_{t_{2,4}}^*$	47
5.51 Zusammensetzung $P_{t_{2,4}}^*$ nach Geschlecht und Schulform (Zeilenproz.) .	47
5.52 Zusammensetzung $P_{t_{2,4}}^*$ nach Geschlecht und Schulform (Spaltenproz.) .	48
5.53 Zusammensetzung nach Geschlecht $P_{t_{2,4}}$	49
5.54 Altersdurchschnitt $P_{t_{2,4}}$	49
5.55 Zusammensetzung $P_{t_{2,4}}$ nach Geschlecht und Schulform (Zeilenproz.) .	50
5.56 Zusammensetzung $P_{t_{2,4}}$ nach Geschlecht und Schulform (Spaltenproz.) .	51
5.57 Zusammensetzung nach Geschlecht $P_{t_{1,2,3}}$	52
5.58 Altersdurchschnitt $P_{t_{1,2,3}}$	52
5.59 Zusammensetzung $P_{t_{1,2,3}}$ nach Geschlecht und Schulform (Zeilenproz.) .	53
5.60 Zusammensetzung $P_{t_{1,2,3}}$ nach Geschlecht und Schulform (Spaltenproz.)	54
5.61 Zusammensetzung nach Geschlecht $P_{t_{2,3,4}}^*$	55
5.62 Altersdurchschnitt $P_{t_{2,3,4}}^*$	55
5.63 Zusammensetzung $P_{t_{2,3,4}}^*$ nach Geschlecht und Schulform (Zeilenproz.) .	56
5.64 Zusammensetzung $P_{t_{2,3,4}}^*$ nach Geschlecht und Schulform (Spaltenproz.)	57
5.65 Zusammensetzung nach Geschlecht $P_{t_{2,3,4}}$	58
5.66 Altersdurchschnitt $P_{t_{2,3,4}}$	58
5.67 Zusammensetzung $P_{t_{2,3,4}}$ nach Geschlecht und Schulform (Zeilenproz.) .	59

5.68 Zusammensetzung $P_{t_2,3,4}$ nach Geschlecht und Schulform (Spaltenproz.)	60
5.69 Zusammensetzung nach Geschlecht $P_{t_1,2,4}^*$	61
5.70 Altersdurchschnitt $P^*t_{1,2,4}$	61
5.71 Zusammensetzung $P_{t_1,2,4}^*$ nach Geschlecht und Schulform (Zeilenproz.)	62
5.72 Zusammensetzung $P_{t_1,2,4}^*$ nach Geschlecht und Schulform (Spaltenproz.)	63
5.73 Zusammensetzung nach Geschlecht $P_{t_1,2,4}$	64
5.74 Altersdurchschnitt $P_{t_1,2,4}$	64
5.75 Zusammensetzung $P_{t_1,2,4}$ nach Geschlecht und Schulform (Zeilenproz.)	65
5.76 Zusammensetzung $P_{t_1,2,4}$ nach Geschlecht und Schulform (Spaltenproz.)	66
5.77 Zusammensetzung nach Geschlecht $P_{t_1,3,4}^*$	67
5.78 Altersdurchschnitt $P^*t_{1,3,4}$	67
5.79 Zusammensetzung $P_{t_1,3,4}^*$ nach Geschlecht und Schulform (Zeilenproz.)	68
5.80 Zusammensetzung $P_{t_1,3,4}^*$ nach Geschlecht und Schulform (Spaltenproz.)	69
5.81 Zusammensetzung nach Geschlecht $P_{t_1,3,4}$	70
5.82 Altersdurchschnitt $P_{t_1,3,4}$	70
5.83 Zusammensetzung $P_{t_1,3,4}$ nach Geschlecht und Schulform (Zeilenproz.)	71
5.84 Zusammensetzung $P_{t_1,3,4}$ nach Geschlecht und Schulform (Spaltenproz.)	72
5.85 Zusammensetzung nach Geschlecht $P_{t_1,2,3,4}^*$	73
5.86 Altersdurchschnitt $P_{t_1,2,3,4}^*$	73
5.87 Zusammensetzung $P_{t_1,2,3,4}^*$ nach Geschlecht und Schulform (Zeilenproz.)	75
5.88 Zusammensetzung $P_{t_1,2,3,4}^*$ nach Geschlecht und Schulform (Spaltenproz.)	76
5.89 Zusammensetzung nach Geschlecht $P_{t_1,2,3,4}$	77
5.90 Altersdurchschnitt $P_{t_1,2,3,4}$	78
5.91 Zusammensetzung $P_{t_1,2,3,4}$ nach Geschlecht und Schulform (Zeilenproz.)	79
5.92 Zusammensetzung $P_{t_1,2,3,4}$ nach Geschlecht und Schulform (Spaltenproz.)	80
6.1 Verzerrung der Paneldaten ($P_{t_3,4}$ zu t_3) nach Geschlecht und Anzahl der Fehler in der Codezuordnung	
	82
6.2 Verzerrung der Paneldaten ($P_{t_3,4}$ zu t_3) nach Schulform und Anzahl der Fehler in der Codezuordnung	
	82
6.3 Zusammensetzung $P_{t_1,2,3,4}$ nach Geschlecht und Schulform (Zeilenproz.)	
	84
A.1 Häufigkeiten von Co001 (erster Buchstabe Vorname Vater)	
	92
A.2 Häufigkeiten von Co002 (erster Buchstabe Vorname Mutter)	
	93
A.3 Häufigkeiten von Co003 (erster Buchstabe eigener Vorname)	
	94
A.4 Häufigkeiten von Co004 (Geburtstag)	
	95
A.5 Häufigkeiten von Co005 (Haarfarbe)	
	96
A.6 Häufigkeiten von Co006 (Augenfarbe)	
	97
A.7 Häufigkeiten von Co007 in Duisburg 2003 (Befragungsteilnahme im Vorjahr ja/nein)	
	98

A.8 Häufigkeiten von Co007 in Duisburg 2004 (Befragungsteilnahme im Vorjahr ja/nein)	98
A.9 Häufigkeiten von Co007 in Duisburg 2005 (Befragungsteilnahme im Vorjahr ja/nein)	98
A.10 Häufigkeiten von Co008 in Duisburg 2003 (Schulwechsel ja/nein) . . .	99
A.11 Häufigkeiten von Co008 in Duisburg 2004 (Schulwechsel ja/nein) . . .	99
A.12 Häufigkeiten von Co008 in Duisburg 2005 (Schulwechsel ja/nein) . . .	99
A.13 Häufigkeiten von Co009 in Duisburg 2003 (Sitzengeblieben ja/nein) . .	100
A.14 Häufigkeiten von Co009 in Duisburg 2004 (Sitzengeblieben ja/nein) . .	100
A.15 Häufigkeiten von Co009 in Duisburg 2005 (Sitzengeblieben ja/nein) . .	100
A.16 Häufigkeiten der Gesamtcodes in Duisburg 2002	101
A.17 Häufigkeiten der Gesamtcodes in Duisburg 2003	101
A.18 Häufigkeiten der Gesamtcodes in Duisburg 2004	101
A.19 Häufigkeiten der Gesamtcodes in Duisburg 2005	101

Abbildungsverzeichnis

2.1	Erhebungsdesign	4
A.1	Codeblatt Duisburg 2002	88
A.2	Codeblatt Duisburg 2003	89
A.3	Codeblatt Duisburg 2004	90
A.4	Codeblatt Duisburg 2005	91

1 Einleitung

Die Ziele des DFG-Projektes *Kriminalität in der modernen Stadt*¹ liegen unter anderem in der Untersuchung der Entstehung und Entwicklung delinquenter bzw. devianter Handlungsstile, sowie deren Kontrollbedingungen bzw. Kontrollprozesse im Längsschnitt. Der methodische Zugang zu diesen Problemkomplexen erforderte ein für diesen theoretischen und inhaltlichen Kontext in Deutschland zum ersten Mal geplantes kohortenspezifisches Längsschnittdesign, das in zwei westdeutschen Städten unterschiedlicher Größe und Sozialstruktur, Münster und Duisburg, umgesetzt wurde. Die durch wiederholte Schülerbefragungen erhobenen Längsschnittinformationen (Paneldaten) sollen sowohl die Analyse von Ursache-Wirkungsbeziehungen zwischen den zentralen Untersuchungsvariablen (einschließlich möglicher Rückkopplungsprozesse) als auch die Analyse von Stabilität und Veränderung der individuellen Kriminalität ermöglichen.

Diesem Ansatz folgend wurden in den Jahren 2000 bis 2003 kriminologische Jugendbefragungen in der *Stadt Münster* durchgeführt. Die bei den Befragungen erhobenen Daten wurden im Jahr 2003 zu einem Vier-Wellen-Panel zusammengefügt (vgl. Pöge 2005a). Parallel wurde ab dem Jahr 2002 begonnen, in der *Stadt Duisburg* zwei Paneldatensätze aufzubauen. Zum einen beginnend mit Schülerinnen und Schülern der siebten, zum anderen der neunten Jahrgangsstufe (siehe Abbildung 2.1). Dieser Bericht dokumentiert zentrale Punkte der methodischen Vorgehensweise der Duisburger Panelbefragungen, die mit der siebten Jahrgangsstufe im Jahr 2002 begannen. Insbesondere werden das Studiendesign, das Zuordnungsverfahren und die realisierte Stichprobe beschrieben. Eine weitere Analyse der Daten und Interpretationen der Ergebnisse sind nicht Gegenstand dieses als Methodendokumentation angelegten Berichtes. Ebenfalls ausgeklammert sind die Bereiche Instrumentenentwicklung, Fragebogengestaltung, Skalendokumentation und Validierung. Zeitvergleiche zwischen den Erhebungswellen werden nur auf der Ebene der Stichprobenszusammensetzung angestellt.

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

2 Studiendesign

Seit dem Jahr 2000 wurden in einjährigen Abständen an Münsteraner, seit dem Jahr 2002 auch an Duisburger Schulen Fragebogeninterviews durchgeführt. Entsprechend der methodischen Konzeption des Forschungsprojektes als Panelstudie wurden dabei dieselben Personen wiederholt befragt. Der genaue Erhebungsplan kann der Abbildung 2.1 entnommen werden. Da für das hier dokumentierte Duisburger Vier-Wellen-Panel die Daten der Duisburger Befragungen der Jahre 2002 bis 2005 die Grundlage bilden, wird im Folgenden nur auf diese eingegangen.

Im Jahr 2002 begann die Untersuchung in Duisburg mit dem Ziel, parallel für zwei Alterskohorten je einen Paneldatensatz aufzubauen. Wie in Münster wurden Schülerinnen und Schüler einbezogen, die eine Sonder-, Haupt-, Real-, Gesamtschule oder ein Gymnasium besuchten. Unter den Sonderschulen wurden ausschließlich die Schulen für Erziehungshilfe ausgewählt. Um eine möglichst breite Datengrundlage für die nachfolgenden Panel-Wellen zu schaffen und spätere Panel-Ausfälle kompensieren zu können, wurde in dieser *ersten* Welle eine Vollerhebung der siebten und neunten Klassen angestrebt.² Auch die *zweite* Duisburger Schülerbefragung im Januar und Februar 2003 strebte eine Vollerhebung der nunmehr achten und zehnten Klassen der Haupt-, Real-, Gesamtschulen, Gymnasien und Sonderschulen für Erziehungshilfe an.³ Im Jahr 2004 wurde die zweite Kohorte (neunte Klasse von 2002) nicht weitergeführt, der Fokus der Erhebung lag in dieser *dritten* Welle dementsprechend auf den nunmehr neunte Klassen der Siebenerkohorte von 2002 (vgl. Abbildung 2.1), wobei vor allem versucht wurde, alle Schülerinnen und Schüler aus den vorigen Erhebungswellen wieder zu befragen.⁴ Im Jahr 2005 wurde während der *vierten* Befragung der zehnten Jahrgangsstufe im Anschluss an die Schul-Haupterhebungen erstmals Nachholbefragungen in den Schulen, in Jugendzentren und auf dem Postweg durchgeführt. Ebenfalls erstmals wurden Schülerinnen und Schüler befragt, die die neunte Jahrgangsstufe wiederholten.⁵

Der bei den Befragungen eingesetzte Fragebogen zu Opferwerdung, selbstberichteter Delinquenz, Erziehungsstilen, Konfliktverhalten, Kriminalitätseinstellungen sowie zu Lebens-, Freizeit-, und Konsumstilen war für eine schriftliche, anonyme Befragung der Schülerinnen und Schüler konzipiert. So weit wie möglich wurden Fragen verwendet,

2 Zur Dokumentation der Duisburger Studie aus dem Jahr 2002 siehe Motzke und Brondies (2004).

3 Zur Dokumentation der Duisburger Studie aus dem Jahr 2003 siehe Brondies (2004).

4 Zur Dokumentation der Duisburger Studie aus dem Jahr 2004 siehe Hilfert (2005).

5 Zur Dokumentation der Duisburger Studie aus dem Jahr 2005 siehe Kunadt (2006).

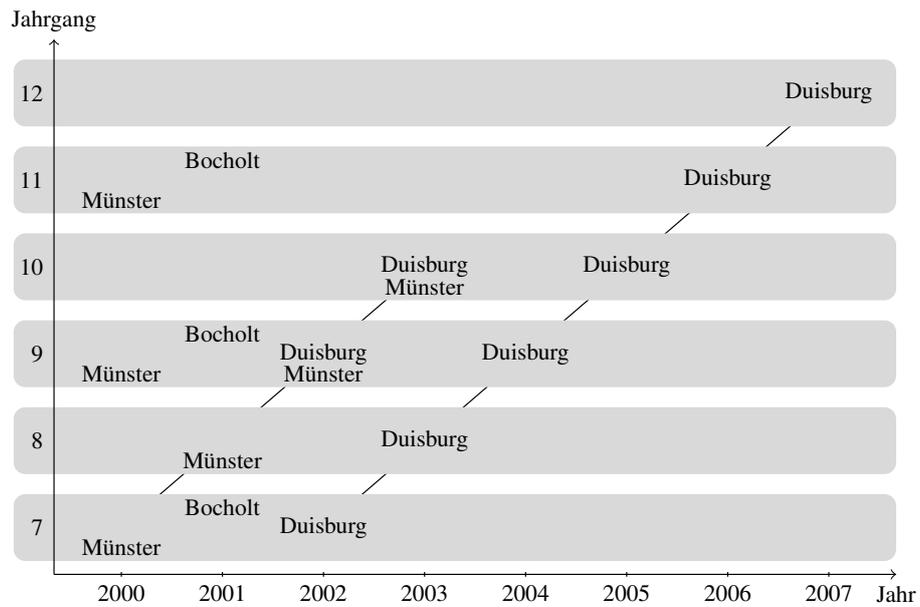


Abbildung 2.1: Erhebungsdesign

die bereits in anderen Jugendstudien eingesetzt wurden. Die Fragebögen der einzelnen Jahre waren aufgrund des Panelansatzes der Studie weitgehend identisch.

Vor den jeweiligen Haupterhebungen wurden die Eltern und Schüler über die Untersuchungsziele informiert und auf die Freiwilligkeit der Teilnahme hingewiesen. Die schriftlichen Befragungen fanden im Klassenverband statt. Der zeitliche Rahmen erstreckte sich über eine Schuldoppelstunde. In wenigen Einzelfällen kam es jedoch vor, dass die Beantwortung der Fragen mehr Zeit in Anspruch nahm, insbesondere wenn sprachliche Schwierigkeiten die Beantwortung beeinträchtigten.

3 Erstellung der Paneldaten

Mit der Durchführung einer Panelstudie sind wegen der 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? In diesem Forschungsprojekt entschied man sich für ein Verfahren, welches mit Hilfe eines persönlichen Codes die Fragebögenzuordnungen zwischen den unterschiedlichen Jahren leisten sollte.

Um eine Zuordnung der Fragebögen aus den verschiedenen Erhebungswellen zu ermöglichen, wurden Codes verwendet, die über Codeblätter abgefragt wurden. Jeder befragte Schüler sollte beim Interview ein Codeblatt ausfüllen, auf dem fünf bzw. ab der 2003er Befragung sechs persönliche Fragen gestellt wurden, aus deren Antwort jeweils ein bestimmter Buchstabe oder eine Zahl notiert werden musste. Die Fragen bezogen sich auf unveränderliche Merkmale des Probanden oder seines Umfeldes (natürliche Haarfarbe, Name des Vaters etc.). Diese Buchstaben-Zahlenkombination sollte den gesamten Code bilden. In jeder Erhebungswelle wurden die Codeblätter den Schülern zum Ausfüllen wieder vorgelegt, so dass bei einem stabilen Antwortverhalten der Schüler die Codeblätter, die von ein und derselben Person in den verschiedenen Wellen ausgefüllt wurden, denselben Code aufweisen mussten.

Während der Befragungen sollte auf den Codeblättern⁶ 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 (ab 2003)
- *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 größeren Antwortvarianz gestellt und zum anderen im Jahr 2003 der Code – ebenso wie in Münster – um eine Frage erweitert und das Layout des gesamten Blattes überarbeitet. So wurde zum Beispiel nicht mehr nach dem *letzten* Buchstaben der Haarfarbe des Vaters, sondern dem *ersten* Buchstaben des Vornamens des Vaters gefragt. Das Layout wurde dergestalt geändert,

6 Abdrucke der Codeblätter aller Befragungszeitpunkte finden sich im Anhang Abbildungen A.1 bis A.4

dass kein handschriftliches Ausfüllen mehr erfolgen musste, sondern alle Antwortvorgaben zum Ankreuzen aufgeführt wurden. Ab dem Befragungsjahr 2003 wurden zusätzlich die folgenden Fragen gestellt:

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

Da im letzten für diesen Bericht relevanten Erhebungsjahr 2005 zusätzlich die nicht versetzten Schülerinnen und Schüler befragt wurden, musste auf dem Codeblatt eine Zusatzfrage nach der entsprechenden Jahrgangsstufe gestellt werden.

Zusätzlich zu diesen acht bzw. neun Fragen standen für die Fragebogenzuordnungen das Geschlecht der Befragten zur Verfügung, sowie die 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 bekannten Probleme auf, sie konnten durch die Modifikationen und Verbesserungen jedoch – zum Teil erheblich – abgemildert werden.

Als generelle Problembereiche des Zuordnungsverfahrens waren schon in Münster die Probleme der *Identifizierung* und *Reproduktion* der Codes aufgetreten. Zum Problembereich der Identifizierung ist festzustellen, dass dieser durch den erweiterten Code und die geänderten Codefragen selbst als gelöst anzusehen ist. In Münster waren in allen vier Wellen nur bei rund drei Vierteln der Schülerinnen und Schülern eindeutige Codes aufgetreten – bei rund einem Viertel der Befragten wiesen unterschiedliche Personen einen gleichen Code auf und waren somit ohne Hinzunahme weiterer Informationen nicht eindeutig zu identifizieren. In Duisburg hingegen weisen nun bei dem fünfstelligen Code aus dem Jahr 2002 immerhin rund 90 Prozent, bei dem sechststelligen Code, der in allen weiteren Befragungsjahren Anwendung fand, fast alle Befragten eindeutige Codes auf (siehe im Anhang Tabellen A.16 bis A.19).

Das Problem der Reproduktion, dass also Befragte in den einzelnen Welle einen nicht übereinstimmenden Code angeben, existiert auch in Duisburg noch in einem nicht zu vernachlässigendem Maße, konnte aber gleichwohl reduziert werden. In Münster fanden sich zwischen jeweils zwei Erhebungszeitpunkten lediglich rund 500 bis 600 Zuordnungen mit exakt übereinstimmendem Code, was bei allen realisierten Zuordnungen einen Prozentsatz von 40 bis 50 Prozent ausmachte (vgl. Pöge 2005a, S. 7 f; Pöge 2005b, S. 66). In Duisburg hingegen liegt die Quote der Zuordnungen mit einem exakt übereinstimmendem Code recht stabil bei rund 60 Prozent aller gefundenen Zuordnungen (siehe Tabelle 3.1). Auch die Reproduktion des Codes hat also mit Hilfe des verbesserten Layouts und der vereinfachten Fragen deutlich besser funktioniert, obwohl eine Codestelle mehr fehlerfrei angegeben werden musste als in Münster.

Tabelle 3.1: Anzahl der handschriftlichen Kontrollen und deren Ergebnisse (vor Plausibilitätskontrollen)

	passt	passt nicht	fehlt	Kontrollen
ohne Fehler	1448	195	9	1652
ein Fehler	723	422	8	1153
zwei Fehler	308	992	5	1305
Summe	2479	1609	22	4110

(a) 2002/2003

	passt	passt nicht	fehlt	Kontrollen
ohne Fehler	1528	42	16	1586
ein Fehler	735	175	7	917
zwei Fehler	292	853	5	1150
drei Fehler	67	119	1	187
Summe	2622	1189	29	3840

(b) 2003/2004

	passt	passt nicht	fehlt	Kontrollen
ohne Fehler	1675	52	8	1735
ein Fehler	728	126	7	861
zwei Fehler	326	1382	5	1713
drei Fehler	52	78	1	131
Summe	2781	1638	21	4440

(c) 2004/2005

Das zur Zuordnung der Bögen aus den unterschiedlichen Bafragungs Jahren angewendete *fehlertolerante Verfahren mit manuellem Handschriftenvergleich* bestand aus drei bzw. vier Schritten: In einem *ersten Schritt* wurden maschinell alle exakt übereinstimmenden Codes aus zwei Erhebungswellen herausgefunden. Die zusammengehörigen Fragebögen und Codeblätter wurden daraufhin einer manuellen Handschriftenkontrolle unterzogen, wobei die offensichtlich nicht passenden Zuordnungen aussortiert wurden. Die passenden Fragebogennummern wurden aus dem Datensatz genommen, so dass sie für die nachfolgenden Schritte nicht mehr zur Verfügung standen. Im *zweiten Schritt* wurde nach Codeübereinstimmungen unter Zulassung eines Fehlers und im *dritten Schritt* unter Zulassung von zwei Fehlern gesucht und die zugehörigen Bogennummern herausgeschrieben. Seit dem Jahr 2003 war es mit dem eingesetzten erweiterten Code möglich, in einem *vierten Schritt* drei Fehler zuzulassen. Auch in diesen drei Schritten wurden als Validierung der Zuordnungen Handschriftenvergleiche durchgeführt, die

offensichtlich nicht passenden Zuordnungen verworfen und vor der Durchführung des nächsten Schrittes die erfolgreich zugeordneten Nummern aus dem Datensatz entfernt. Eine Übersicht über die durchgeführten Zuordnungsschritte und die Bedingungen, die der Code erfüllen musste, findet sich in Tabelle 3.2.

Das geschilderte Verfahren ist codegeleitet und hierarchisch. Zum einen ist das erste Kriterium für eine Zuordnung nach wie vor der Code, denn es werden nur offensichtliche Fehlzuordnungen ausgesondert. Zum anderen wird durch die schrittweise Durchführung gewährleistet, dass weniger Fehler im Code bei der Zuordnung bevorzugt werden.

Tabelle 3.2: Kontrollschritte Duisburg

	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	Schule muss übereinstimmen
	S7	ein Fehler	Schulwechsel »ja« oder »fehlend«
zwei Fehler	S5	zwei Fehler	ohne Fehler
	S6	zwei Fehler	Schule muss übereinstimmen

(a) 2002/2003

	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	Schule muss übereinstimmen
	S8	zwei Fehler	Geschlecht muss übereinstimmen
drei Fehler	S7	drei Fehler	ohne Fehler

(b) 2003/2004

	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	keine Bedingung
drei Fehler	S7	drei Fehler	ohne Fehler

(c) 2004/2005

4 Realisierte Zuordnungen

Mit Hilfe des geschilderten Zuordnungsverfahrens konnten für die Paneldatensätze die in Tabelle 4.1 aufgeführten Fallzahlen realisiert werden. Für die geplanten Längsschnittanalysen sind neben den Zwei-Wellen-Panels die lückenlosen Drei-Wellen-Panels 2002 bis 2004 ($t_{1,2,3}$) mit 2 012 Fällen und 2003 bis 2005 ($t_{2,3,4}$) mit 2 260 Fällen sehr bedeutsam. Das vorläufige Resultat der Längsschnitt-Erhebungen in Duisburg stellt jedoch der Datensatz des Vier-Wellen-Panels der Jahre 2002 bis 2005 ($t_{1,2,3,4}$) mit seinen 1 769 Fällen dar.⁷

Da zum vierten Zeitpunkt (2005) zusätzlich zum zehnten Jahrgang auch Schülerinnen und Schüler der neunten Jahrgangsstufe in den Datensätzen enthalten sind und in den weiteren Ausführungen Vergleiche mit der Schulstatistik nur jeweils einer Jahrgangsstufe durchgeführt werden sollen, wird bei den Paneldatensätzen, die auch den vierten Zeitpunkt umfassen eine Differenzierung vorgenommen. So werden hierfür ausschließlich nur diejenigen Jugendlichen aus den betreffenden Panel- und Querschnittsdatsätzen betrachtet, die die zehnte Jahrgangsstufe besuchen. Zusätzlich werden allerdings mit den vollständigen Panel-Datensätzen (neunte und zehnte Jahrgangsstufe) Vergleiche mit den betreffenden Querschnittsdaten durchgeführt.

4.1 Die verwendeten Datensätze und ihre Kurzbezeichnungen

Im Folgenden sollen der Einfachheit halber Kurzbezeichnungen für die verwendeten Daten benutzt werden (siehe Tabelle 4.1).

Die Bezeichnung »S« mit indiziertem Zeitpunkt steht für die Daten der amtlichen Schulstatistik zu einem bestimmten Zeitpunkt. Hierbei sind nur die Schulen berücksichtigt, die an den Befragungen teilnahmen. S_{t_3} enthält zum Beispiel alle (5 627 Fälle) der Schulstatistik des Schuljahres 2003/2004, wobei nur die teilnehmenden Schulen berücksichtigt sind. Ein »Q« mit einem Zeitpunkt als Index steht für einen ursprünglichen Querschnittsdatsatz der Einzelbefragungen. Hierbei handelt es sich um die Querschnittsdaten der Schülerinnen und Schüler, die eine Schule besuchten, die an den Befragungen teilnahm und den Plausibilitätskontrollen standhielten. Q_{t_1} beinhaltet zum Beispiel alle (3 411) gültigen Fälle des Jahres 2002. Die einzelnen Paneldaten durch ein »P« mit einem Index, der die Zeitpunkte ausdrückt, bezeichnet. So steht $P_{t_{2,3}}$

⁷ Berücksichtigt sind hier nur die gültigen Fälle, die auch den Plausibilitätskontrollen der einzelnen Querschnitte standhielten.

für den Paneldatensatz mit allen (2 596) *gültigen* Fällen, die zwischen t_2 und t_3 einander zugeordnet werden konnten. Ein » P_Q « mit einem Zeitpunkt als Index bezeichnet einen Panelquerschnitt. Diese Panelquerschnitte beinhalten *alle* vorhandenen Fälle des gesamten Datensatzes aus *einem* Zeitpunkt. So enthält zum Beispiel $P_{Q_{t_2}}$ alle (3 132) gültigen Fälle des Datensatzes, die für t_2 vorhanden sind. Dies sind im Unterschied zum Querschnitt Q_{t_2} alle Fälle, die im Jahre 2003 befragt wurden und mindestens *einem* Fall *eines* anderen Zeitpunktes zugeordnet werden konnten. Die hochgestellten Sternchen bezeichnen diejenigen Datensätze, bei denen aus vergleichsgründen zum vierten Zeitpunkt nur Jugendliche der zehnten Jahrgangsstufe enthalten sind. So umfasst $Q_{t_4}^*$ die 3 243 gültigen Fälle des Jahres 2005, die die die zehnte Jahrgangsstufe besuchen.

4.1 Die verwendeten Datensätze und ihre Kurzbezeichnungen

Tabelle 4.1: Bezeichnungen und Fallzahlen der Duisburger Datensätze 2002 bis 2005

Abkürzung	Fallzahl (N_b)	Beschreibung
S_{t_1}	3910	Amtliche Schulstatistik, Schuljahr 2001/2002 (Stichtag: 15. 10.) ^a
S_{t_2}	3763	Amtliche Schulstatistik, Schuljahr 2002/2003 (Stichtag: 15. 10.) ^a
S_{t_3}	3925	Amtliche Schulstatistik, Schuljahr 2003/2004 (Stichtag: 15. 10.) ^a
S_{t_4}	3543	Amtliche Schulstatistik, Schuljahr 2004/2005 (Stichtag: 15. 10.) ^a
Q_{t_1}	3411	Querschnittsdaten des ersten Befragungsjahres 2002 ^a
Q_{t_2}	3392	Querschnittsdaten des zweiten Befragungsjahres 2003 ^a
Q_{t_3}	3339	Querschnittsdaten des dritten Befragungsjahres 2004 ^a
Q_{t_4}	3405	Querschnittsdaten des vierten Befragungsjahres 2005 ^a
$Q_{t_4}^*$	3243	Querschnittsdaten des vierten Befragungsjahres 2005 (nur zehnte Jahrgangsstufe) ^a
$P_{Q_{t_1}}$	2750	Panelquerschnittsdaten des Jahres 2002 ^b
$P_{Q_{t_2}}$	3132	Panelquerschnittsdaten des Jahres 2003 ^b
$P_{Q_{t_3}}$	3125	Panelquerschnittsdaten des Jahres 2004 ^b
$P_{Q_{t_4}}$	3074	Panelquerschnittsdaten des Jahres 2005 ^b
$P_{Q_{t_4}}^*$	2943	Panelquerschnittsdaten des Jahres 2005 (t_4 nur zehnte Jahrgangsstufe) ^b
$P_{t_{1,2}}$	2472	Paneldaten der Jahre 2002 und 2003 ^b
$P_{t_{2,3}}$	2596	Paneldaten der Jahre 2003 und 2004 ^b
$P_{t_{3,4}}$	2750	Paneldaten der Jahre 2004 und 2005 ^b
$P_{t_{3,4}}^*$	2640	Paneldaten der Jahre 2004 und 2005 (t_4 nur zehnte Jahrgangsstufe) ^b
$P_{t_{1,3}}$	2202	Paneldaten der Jahre 2002 und 2004 ^b
$P_{t_{1,4}}$	2165	Paneldaten der Jahre 2002 und 2005 ^b
$P_{t_{1,4}}^*$	2086	Paneldaten der Jahre 2002 und 2005 (t_4 nur zehnte Jahrgangsstufe) ^b
$P_{t_{2,4}}$	2493	Paneldaten der Jahre 2003 und 2005 ^b
$P_{t_{2,4}}^*$	2411	Paneldaten der Jahre 2003 und 2005 (t_4 nur zehnte Jahrgangsstufe) ^b
$P_{t_{1,2,3}}$	2012	Paneldaten der Jahre 2002 bis 2004 ^b
$P_{t_{2,3,4}}$	2260	Paneldaten der Jahre 2003 bis 2005 ^b
$P_{t_{2,3,4}}^*$	2186	Paneldaten der Jahre 2003 bis 2005 (t_4 nur zehnte Jahrgangsstufe) ^b
$P_{t_{1,2,4}}$	1930	Paneldaten der Jahre 2002, 2003, 2005 ^b
$P_{t_{1,2,4}}^*$	1868	Paneldaten der Jahre 2002, 2003, 2005 (t_4 nur zehnte Jahrgangsstufe) ^b
$P_{t_{1,3,4}}$	1922	Paneldaten der Jahre 2002, 2004, 2005 ^b
$P_{t_{1,3,4}}^*$	1862	Paneldaten der Jahre 2002, 2004, 2005 (t_4 nur zehnte Jahrgangsstufe) ^b
$P_{t_{1,2,3,4}}$	1769	Paneldaten der Jahre 2002 bis 2005 ^b
$P_{t_{1,2,3,4}}^*$	1715	Paneldaten der Jahre 2002 bis 2005 (t_4 nur zehnte Jahrgangsstufe) ^b

N_b bedeutet beobachtetes N

^a nur teilnehmende Schulen

^b gültige Fälle zu allen entsprechenden Befragungszeitpunkten

4.2 Abschätzung und Bewertung der Zuordnungsgüte

Die Bewertung des angewendeten Zuordnungsverfahrens nach der realisierten Zuordnungsquote gestaltet sich komplizierter als auf den ersten Blick erkennbar.

Um die Zuordnungsquote bestimmen zu können, muss die Anzahl derjenigen Schülerinnen und Schüler bestimmt bzw. geschätzt werden, die zwischen den Befragungswellen hätten zugeordnet werden *können*.⁸ Auf Schulebene beschränkt sich hierbei die Grundgesamtheit zunächst auf die Jugendlichen der befragten Jahrgänge, die eine *Schule* besuchten, die der Befragung zustimmte. Diese Anzahl kann relativ genau den jeweiligen Schulstatistiken entnommen werden (siehe Tabelle 4.1). Auch die Schulstatistiken selbst sind meist nicht exakt, da die ausgewiesenen Belegungszahlen zu den *Statistikstichtagen* durch zwischenzeitliche Ab- und Zugänge der Schulen von den tatsächlichen Belegungen an den *Befragungsstichtagen* abweichen können; diese Abweichungen scheinen jedoch vernachlässigbar klein. Darüber hinaus problematisch sind Veränderungen der Schulpopulation über den *gesamten Befragungszeitraum* (vier Jahre) im Sinne von Zu- und Wegzügen nach bzw. aus Duisburg, Schulwechslern in bzw. von Schulen, die nicht an den Befragungen teilnahmen, Sitzenbleibern etc. Durch die Wahl der Erhebungsmethode (Fragebogeninterviews im Klassenverband) ist die erreichbare Probandenzahl auf diejenigen beschränkt, die an den jeweiligen Befragungstagen anwesend waren. Da die Fragebogenzuordnungen über die oben geschilderte Methode unter Zuhilfenahme von Codeblättern vorgenommen wurden, reduziert sich die Zahl weiter auf diejenigen Befragungsteilnehmerinnen und -teilnehmer, die ihr Codeblatt verwertbar ausgefüllt hatten.⁹

Da als einzige Referenzdaten die aggregierten Schulstatistiken der Befragungsjahre zur Verfügung stehen und diese keinen befriedigenden Aufschluss über die oben genannten

8 Dazu folgendes Beispiel: Zum Zeitpunkt t_1 gingen 3 910 Schülerinnen und Schüler auf die von uns befragten Schulen, 3 411 von ihnen konnten von uns (verwertbar) befragt werden, 499 nicht. Zum Zeitpunkt t_2 waren es laut Schulstatistik 3 763 Probanden, von denen wir 3 392 befragen konnten, 371 nicht. Wie viele der Probanden aus t_1 haben wir zu t_2 wiederbefragt bzw. wie viele der Fragebögen aus t_1 können Bögen aus t_2 theoretisch zugeordnet werden? Nimmt man an, dass die Schulpopulation (relativ) stabil blieb, kann man zwei Extrempunkte formulieren. Zum einen könnten alle 3 392 wiederbefragt worden sein, die 371 Ausfälle (t_2) sind in den 499 Ausfällen (t_1) enthalten. Zum anderen könnten alle 255 Ausfälle (t_1) zum zweiten Zeitpunkt wiederbefragt worden sein, alle 266 Ausfälle (t_2) könnten allerdings im ersten Zeitpunkt schon befragt worden sein, so dass nur noch 1 329 Personen zu beiden Zeitpunkten befragt wurden. Die realisierten Zuordnungen von 1 271 ergeben somit eine Quote von 96 Prozent oder 68 Prozent. Beides ist unwahrscheinlich.

9 Lässt man diese Betrachtungen außer Acht, so sind noch 1 769 Jugendliche von den ursprünglichen 3 411 im Vier-Wellen-Panel enthalten. Dies sind Jugendliche, die im beobachteten Zeitraum dreimal versetzt wurden, nie umgezogen sind, nie auf eine Schule außerhalb der Stichprobe umgeschult wurden, nie an den Befragungstagen gefehlt haben und bei jeder Befragung einen hinreichenden Code geliefert haben. Von einem sehr konservativen Standpunkt aus kann hieraus mit einer gewissen Berechtigung eine Panelmortalität von 48 Prozent abgeleitet werden.

Probleme geben können, kann die Zahl der Probanden, die *tatsächlich mehrfach befragt wurden*, also hätten zugeordnet werden können, nur geschätzt werden.

Um die zu erwartende Zahl (N_e) der möglichen Zuordnungen zwischen den jeweiligen Zeitpunkten schätzen zu können, seien folgende Annahmen zugelassen:

1. In der größten offiziellen Schulpopulation der Befragungszeitpunkte sind alle Schülerinnen und Schüler der anderen Zeitpunkte enthalten
2. Die Differenzen zwischen den Stichproben und den jeweiligen Schulstatistiken beruhen auf zufälligen, stichprobenneutralen Ausfällen

Annahme 1 ist notwendig, um eine Referenzgröße über die Anzahl aller Probanden zu bestimmen, die überhaupt zu irgendeinem Zeitpunkt hätten teilnehmen können. Problematisch ist die Annahme deshalb, da Veränderungen der Schulpopulation stattgefunden haben, die auf Aggregatebene nicht mehr erkennbar sind. Es wird vernachlässigt, dass zu jedem Zeitpunkt ein gewisser Teil an Schülerinnen und Schülern nicht in der größten Population enthalten waren.

Annahme 2 ist problematisch, da in der Realität die Befragungsteilnahme vermutlich nicht zufällig, sondern von bestimmten Faktoren abhängig ist. So haben erste Analysen gezeigt, dass der Rücklauf in den Querschnitten unter anderem von Geschlecht und Schulform abhängt. Es kann also vermutet werden, dass auch der Anteil der Personen, die zu *mehreren* Zeitpunkten nicht teilnahmen, nicht zufällig verteilt ist, sondern von denselben Faktoren abhängt.

Trifft man die oben genannten Annahmen, kann folgender Satz formuliert werden: Die zu erwartende Anzahl der Befragungsteilnehmer (N_e) zu bestimmten Zeitpunkten (t_i) ergibt sich aus der größten Schulpopulation ($\max_{i=1, \dots, k}(N(S_{t_i}))$) multipliziert mit dem Produkt der Quotienten aus den jeweiligen Anzahlen der realisierten Querschnittstichproben ($N(Q_{t_i})$) und den Schulpopulationen ($N(S_{t_i})$).¹⁰

$$N_e = \max_{i=1, \dots, k} (N(S_{t_i})) \cdot \prod_{i=1}^k \frac{N(Q_{t_i})}{N(S_{t_i})} \quad (4.1)$$

Aus den zu erwartenden Anzahlen (N_e) und den tatsächlich beobachteten bzw. realisierten Anzahlen der Zuordnungen (N_b) lassen sich die in Tabelle 4.2 angeführten Quoten berechnen. Sie dürften, wie oben angedeutet, die »wahre« Zuordnungsquote eher unterschätzen, da auf aggregierter Ebene der Schulstatistiken nicht sichtbare Veränderungen (Sitzenbleiber, Umzüge etc., die sich pro Jahr – aber nicht im Panel – ausgleichen), nicht berücksichtigt werden. Dieser Anteil liegt schätzungsweise bei mindestens zehn Prozent. Eine Überschätzung der Quoten durch nicht berücksichtigte, systematische Ausfälle erscheint im Vergleich dazu weniger bedeutsam.

¹⁰ Anders ausgedrückt entspricht dies dem Produkt der größten Ausgangsstichprobe und den multiplizierten Rücklaufquoten der beteiligten Querschnitte.

Tabelle 4.2: Zuordnungsgüte der Paneldaten in Duisburg und Münster im Vergleich

	Duisburg			Münster			
	N_e	N_b	Quote (%)	N_e	N_b	Quote (%)	
$P_{Q_{t_1}}$	3411	2750	80,6	$P_{Q_{t_1}}$	1850	1382	74,7
$P_{Q_{t_2}}$	3392	3132	92,3	$P_{Q_{t_2}}$	1915	1668	87,1
$P_{Q_{t_3}}$	3339	3125	93,6	$P_{Q_{t_3}}$	1947	1725	88,6
$P_{Q_{t_4}}$	3405	3074	90,3	$P_{Q_{t_4}}$	1819	1490	81,9
$P_{t_{1,2}}$	3075	2472	80,4	$P_{t_{1,2}}$	1683	1271	75,5
$P_{t_{2,3}}$	3010	2596	86,2	$P_{t_{2,3}}$	1710	1373	80,3
$P_{t_{3,4}}^*$	3106	2640	85,0	$P_{t_{3,4}}$	1705	1406	82,2
$P_{t_{1,3}}$	2913	2202	75,6	$P_{t_{1,3}}$	1711	1052	61,5
$P_{t_{1,4}}^*$	3122	2086	66,8	$P_{t_{1,4}}$	1620	910	56,2
$P_{t_{2,4}}^*$	3105	2411	77,6	$P_{t_{2,4}}$	1677	1105	65,9
$P_{t_{1,2,3}}$	2626	2012	76,6	$P_{t_{1,2,3}}$	1502	997	66,4
$P_{t_{2,3,4}}^*$	2755	2186	79,3	$P_{t_{2,3,4}}$	1497	1075	71,8
$P_{t_{1,2,4}}^*$	2814	1868	66,4	$P_{t_{1,2,4}}$	1474	823	55,8
$P_{t_{1,3,4}}^*$	2666	1862	69,8	$P_{t_{1,3,4}}$	1499	851	56,8
$P_{t_{1,2,3,4}}^*$	2403	1715	71,4	$P_{t_{1,2,3,4}}$	1316	813	61,8

Die Quoten können für Paneldatensätze, die den vierten Zeitpunkt umfassen nur für den eingeschränkten Datensatz (gekennzeichnet durch *) berechnet werden, da für die erweiterten Datensätze, die beide Jahrgangsstufen enthalten Annahme 1 verletzt ist.

Um die Zuordnungsquoten der realisierten Paneldatensätze¹¹ bewerten zu können, können Quoten, die in anderen Panelstudien erreicht wurden, herangezogen werden. Ein übliches Verfahren bei Paneluntersuchungen stellt die Zuordnung der Fragebögen über erhobene Adressen der Schülerinnen und Schüler dar. Hierbei tritt erfahrungsgemäß schon bei der *ersten Adresserhebung* ein Ausfall von ca. 50 Prozent auf (exemplarisch hierzu siehe Böttger et al. 2003, S. 35 ff.), der im Laufe solcher Adresspanels deutlich größer wird.

Die Zuordnungsquoten für die Zwei-Wellen-Panels aus jeweils zwei aufeinanderfolgenden Zeitpunkten liegen zwischen 80 und 86 Prozent. Es ist ersichtlich, dass die Quoten im Verlauf der Befragungen besser wurden. So wurde die schlechteste Zuordnungsquote zwischen den Jahren 2002 (t_1) und 2003 (t_2) verwirklicht; auch die absolute Fallzahl stieg entsprechend an. Nimmt man die Fallzahlen des Panels 2004/2005, wel-

11 Im Falle der *Panelquerschnitte* (P_Q) bezeichnet die Zuordnungsquote den Anteil derjenigen Bögen, zu denen in irgendeinem weiteren Jahr ein passender gefunden werden konnte. Diese Quote ist in den späteren Jahren deutlich höher als zu Beginn der Befragungen. Die niedrigere Quote des Jahres 2002 (t_1) liegt unter anderem an der Tatsache, dass ein vorhergehender Datensatz fehlt.

ches auch die neunte Jahrgangsstufe umfasst, so sie hier am höchsten. 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 mag auch ein positiver Gewöhnungseffekt bei den Probanden eine Rolle spielen. Die Quoten der beiden lückenlosen Drei-Wellen-Panels liegen bei 77 und 79 Prozent, die des Vier-Wellen-Panels erwartungsgemäß niedriger bei 71 Prozent.

Die Zuordnungsquoten der Paneldatensätze, die ein oder zwei fehlende Zeitpunkte aufweisen, sind naturgemäß niedriger. Dies liegt zum Teil daran, dass beim oben beschriebenen Zuordnungsverfahren die Validierung per Handschriftenvergleich durch veränderte Handschriften immer schwieriger wird, je größer der dazwischenliegende Zeitraum ist, zum Teil auch an den oben genannten Annahmen bei der Schätzung der zu erwartenden Zuordnungsfallzahlen. Die nicht »sichtbaren« Veränderungen der Grundgesamtheit auf der Aggregatebene, die von der oben beschriebenen Formel 4.1 nicht exakt berücksichtigt werden können, werden natürlich im Laufe der Zeit immer gravierender.

Im Vergleich zu den Quoten aus der Münsteraner Panelstudie zeigen sich auch hier deutlich bessere Ergebnisse. Die Quoten der aufeinanderfolgenden Zwei-Wellen-Panels liegen in Duisburg um rund fünf bis acht Prozent höher als in Münster. Die der Drei-Wellen- und des Vier-Wellen-Panels um rund zehn Prozent. Es zeigt sich also, dass durch den verbesserten Code und das Zuordnungsverfahren ein deutlich besseres Ergebnis bei der Ausschöpfung erzielt werden konnte.

5 Zusammensetzung der Paneldaten

In diesem Abschnitt soll die Zusammensetzung der einzelnen Paneldatensätze beschrieben werden. Zum Vergleich werden dabei die Daten der Querschnitte und der Schulstatistiken aus den entsprechenden Jahren hinzugezogen. Bei den vergleichenden Darstellungen werden die zentralen Variablen Geschlecht, Alter und Schulform verwendet. Bei der Darstellung der Zusammensetzung der »echten« Paneldaten (also ohne die Panelquerschnitte), können hierbei beim Geschlecht und Alter Inkonsistenzen über die Zeit auftreten. Dies liegt darin begründet, dass bei *jeder* Frage fehlerhafte bzw. inkonsistente Antworten auftreten. So natürlich auch bei so »eindeutig« zu beantwortenden Fragen wie denen nach Alter und Geschlecht. Einzig die Informationen über Schule und somit auch die Schulform ist in unserem Fall nahezu fehlerlos, da sie extern erhoben wurde und nicht vom Antwortverhalten der Probanden abhängt.

5.1 Die Panelquerschnitte

Wie oben 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 Jahres gefunden werden konnte. Bei diesen Datensätzen handelt es sich also um stark konstruierte Daten, die nur unter bestimmten Analysegesichtspunkten wichtig sein können.¹²

5.1.1 Panelquerschnitt 2002 ($P_{Q_{t_1}}$)

Der Panelquerschnitt des Jahres 2002 enthält 2 750 Fälle.¹³ Dies ist insofern von Bedeutung, als dass diese Anzahl die Obergrenze aller anschließenden Paneldatensätze darstellt, die das Jahr 2002 und weitere Zeitpunkte enthalten, da, wie oben beschrieben, in diesen Fällen alle überhaupt zugeordneten Fälle des Jahres 2002 enthalten sind.

Wie in Tabelle 5.1 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine leichte Verschiebung vor. Schon im Querschnitt des Jahres 2002 sind im Vergleich zur

¹² Wenn also genau die Population der überhaupt zuzuordnenden Fälle von Interesse ist.

¹³ In den Daten des Jahres 2002 existieren vier Fälle, in denen das Geschlecht der befragten Person unbekannt ist, also einen fehlenden Wert darstellt. Deshalb liegt die Fallzahl in den Auswertungen, in

Tabelle 5.1: Zusammensetzung nach Geschlecht $P_{Q_{t_1}}$

	männlich		weiblich		gesamt N (100 %)
	N	%	N	%	
S_{t_1}	2003	51,2	1907	48,8	3910
Q_{t_1}	1728	50,7	1679	49,3	3407
$P_{Q_{t_1}}$	1351	49,1	1398	50,9	2749

Tabelle 5.2: Altersdurchschnitt $P_{Q_{t_1}}$

	Alter	Std.-abw.	gesamt N
Q_{t_1}	13,0	0,8	3346
$P_{Q_{t_1}}$	13,0	0,7	2709

Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Panelquerschnitt ist dieser Unterschied noch etwas deutlicher: Es sind rund zwei Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen. Das Durchschnittsalter der Jugendlichen, deren Fälle im Panelquerschnitt enthalten sind, weicht nicht vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.2). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik bei nur einem Schuljahrgang kaum mit den vorhandenen Daten zur Deckung gebracht werden kann. Aus Tabelle 5.3 wird ersichtlich, dass der Panelquerschnitt 2002 im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen systematisch verzerrt ist. Im Paneldatensatz befinden sich – bei sehr niedrigen Basisraten – etwas weniger Sonderschülerinnen und -schüler, rund drei Prozent weniger Hauptschülerinnen und -schüler, ein Prozent weniger Gesamtschülerinnen und -schüler, dafür rund zwei Prozent mehr Realschülerinnen und -schüler und rund zweieinhalb Prozent mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Über alle Schulformen hinweg sind – mit Ausnahme der Sonderschulen – prozentual etwas mehr Schülerinnen und etwas weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.4). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

denen das Geschlecht eine Rolle spielt, in den Querschnittsdaten bei 3 407, in den Paneldaten ist noch ein Fall mit unbekanntem Geschlecht enthalten, die Fallzahl liegt also bei 2 749.

5.1 Die Panelquerschnitte

Tabelle 5.3: Zusammensetzung $P_{Q_{t_1}}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform ^a (%)					gesamt N (100 %)
		HS	RS	GS	GY	SO	
männlich	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
	$P_{Q_{t_1}}$	24,7	23,7	30,8	20,4	0,4	1351
weiblich	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
	$P_{Q_{t_1}}$	16,6	25,0	30,8	27,7	0,0	1398
gesamt	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
	$P_{Q_{t_1}}$	20,6	24,3	30,8	24,1	0,2	2749

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

Tabelle 5.4: Zusammensetzung $P_{Q_{t_1}}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform ^a (%)					gesamt (%)
		HS	RS	GS	GY	SO	
männlich	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
	$P_{Q_{t_1}}$	59,0	47,8	49,2	41,6	100,0	49,1
weiblich	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
	$P_{Q_{t_1}}$	41,0	52,2	50,8	58,4	0,0	50,9

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

5.1.2 Panelquerschnitt 2003 ($P_{Q_{t_2}}$)

Der Panelquerschnitt des Jahres 2003 enthält 3 132 Fälle.¹⁴ Dies sind alle überhaupt zuordenbaren Fälle des Jahres 2003. Dieser Panelquerschnitt stellt damit den größten Panelquerschnitts-Datensatz in den erhobenen vier Zeitpunkten dar. Wie in Tabelle 5.5 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine leichte Verschiebung vor. Schon im Querschnitt des Jahres 2003 sind im Vergleich zur Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Panelquerschnitt ist dieser Unterschied noch etwas deutlicher: Es sind rund zwei Prozent mehr Schülerinnen vorhanden als in der

¹⁴ In den Daten des Jahres 2003 existieren zwei Fälle, in denen das Geschlecht der befragten Person unbekannt ist, also einen fehlenden Wert darstellt. Deshalb liegt die Fallzahl in den Auswertungen, in denen das Geschlecht eine Rolle spielt, in den Querschnittsdaten bei 3 390.

Tabelle 5.5: Zusammensetzung nach Geschlecht $P_{Q_{t_2}}$

	männlich		weiblich		gesamt N (100 %)
	N	%	N	%	
S_{t_2}	1927	51,2	1836	48,8	3763
Q_{t_2}	1703	50,2	1687	49,8	3390
$P_{Q_{t_2}}$	1538	49,1	1594	50,9	3132

Tabelle 5.6: Altersdurchschnitt $P_{Q_{t_2}}$

	Alter	Std.-abw.	gesamt N
Q_{t_2}	14,0	0,8	3275
$P_{Q_{t_2}}$	14,0	0,8	3059

Schulstatistik ausgewiesen. Das Durchschnittsalter der Jugendlichen, deren Fälle im Panelquerschnitt enthalten sind, weicht nicht vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.6). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik bei nur einem Schuljahrgang kaum mit den vorhandenen Daten zur Deckung gebracht werden kann. Aus Tabelle 5.7 wird ersichtlich, dass der Panelquerschnitt 2003 im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen systematisch verzerrt ist. Im Paneldatensatz befinden sich keine Sonderschülerinnen und -schüler mehr, eine Schulform, die in Duisburg aufgrund der speziellen Schullandschaft allerdings sowieso

Tabelle 5.7: Zusammensetzung $P_{Q_{t_2}}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform ^a (%)					gesamt N (100 %)
		HS	RS	GS	GY	SO	
männlich	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
	$P_{Q_{t_2}}$	27,7	23,1	31,9	17,3	—	1538
weiblich	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
	$P_{Q_{t_2}}$	19,3	25,0	31,7	24,1	—	1594
gesamt	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
	$P_{Q_{t_2}}$	23,4	24,0	31,8	20,8	—	3132

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

Tabelle 5.8: Zusammensetzung $P_{Q_{t_2}}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform ^a (%)					gesamt (%)
		HS	RS	GS	GY	SO	
männlich	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
	$P_{Q_{t_2}}$	58,1	47,1	49,3	40,9	—	49,1
weiblich	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
	$P_{Q_{t_2}}$	41,9	52,9	50,7	59,1	—	50,9

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

sehr selten vertreten ist. Im Panelquerschnitt sind rund vier Prozent weniger Hauptschülerinnen und -schüler, dafür etwa ein Prozent mehr Realschülerinnen und -schüler, rund eineinhalb Prozent mehr Gesamtschülerinnen und -schüler und rund zwei Prozent mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Über alle Schulformen hinweg sind prozentual etwas mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.8). Der Unterschied zur Schulstatistik ist hierbei in der Realschule am größten und im Gymnasium am geringsten, wenn man vom Ausnahmefall der Sonderschulen absieht. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

5.1.3 Panelquerschnitt 2004 ($P_{Q_{t_3}}$)

Der Panelquerschnitt des Jahres 2004 enthält 3 125 Fälle. Dies sind alle überhaupt zurechenbaren Fälle des Jahres 2004. Wie in Tabelle 5.9 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine leichte Verschiebung vor. Schon im Querschnitt des Jahres 2004 sind im Vergleich zur Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Panelquerschnitt ist dieser Unterschied noch etwas deutlicher: Es sind fast zwei Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen. Das Durchschnittsalter der Jugendlichen, deren Fälle im Panelquerschnitt enthalten

Tabelle 5.9: Zusammensetzung nach Geschlecht $P_{Q_{t_3}}$

	männlich		weiblich		gesamt N (100 %)
	N	%	N	%	
S_{t_3}	1999	50,9	1926	49,1	3925
Q_{t_3}	1635	49,0	1704	51,0	3339
$P_{Q_{t_3}}$	1502	48,1	1623	51,9	3125

Tabelle 5.10: Altersdurchschnitt $P_{Q_{t_3}}$

	Alter	Std.-abw.	gesamt N
Q_{t_3}	15,1	0,8	3270
$P_{Q_{t_3}}$	15,1	0,8	3077

Tabelle 5.11: Zusammensetzung $P_{Q_{t_3}}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform ^a (%)					gesamt N
		HS	RS	GS	GY	SO	(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	19,2	0,5	1502
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
	$P_{Q_{t_3}}$	18,5	23,8	31,5	26,1	0,1	1623
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
	$P_{Q_{t_3}}$	22,1	23,1	31,8	22,8	0,3	3125

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

sind, weicht nicht vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.10). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik bei nur einem Schuljahrgang kaum mit den vorhandenen Daten zur Deckung gebracht werden kann. Aus Tabelle 5.11 wird ersichtlich, dass der Panelquerschnitt 2004 im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen systematisch verzerrt ist. Im Paneldatensatz befinden sich insgesamt ein Prozent weniger Sonderschülerinnen und -schüler, fast drei Prozent weniger Hauptschülerinnen und -schüler und ein halbes Prozent weniger Gesamtschülerinnen und -schüler, dafür rund eineinhalb Prozent mehr Realschülerinnen und -schüler und rund zweieinhalb Prozent mehr Gymnasiasten und Gymnasiastinnen als in der Schulstatistik ausgewiesen. Über alle Schulformen hinweg sind prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten, sieht man vom Ausnahmefall der Sonderschulen ab, die insgesamt nur von fünfzig Schülerinnen und Schülern besucht werden (siehe Tabelle 5.12). Der Unterschied zur Schulstatistik ist hierbei in der Gesamtschule am größten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Tabelle 5.12: Zusammensetzung $P_{Q_{t_3}}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform ^a (%)					gesamt (%)
		HS	RS	GS	GY	SO	
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
	$P_{Q_{t_3}}$	56,6	46,4	48,4	40,5	87,5	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
	$P_{Q_{t_3}}$	43,4	53,6	51,6	59,5	12,5	51,9

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

5.1.4 Panelquerschnitt 2005 ($P_{Q_{t_4}}^*$ und $P_{Q_{t_4}}$)

Im Jahr 2005 (t_4) wurden im Gegensatz zu den vorherigen Erhebungsjahren nicht nur Schülerinnen und Schüler einer Jahrgangsstufe befragt. Zusätzlich zu den Zehntklässlerinnen und -klässlern wurden, um die Panelausschöpfung möglichst hoch zu halten, diejenigen Jugendlichen interviewt, die aufgrund von Nicht-Versetzungen die neunte Jahrgangsstufe wiederholen mussten. Um eine Analyse der Verzerrungen im Vergleich zur Schulstatistik, die nur Jugendliche des Jahr 2005 aus den zehnten Jahrgangsstufen umfasst, zu ermöglichen, wird eine Unterscheidung vorgenommen. Für den Vergleich mit der Schulstatistik werden auch im Querschnitts- und Paneldatensatz des Jahres 2005 nur die Schülerinnen und Schüler der zehnten Jahrgangsstufe berücksichtigt ($Q_{t_4}^*$ und $P_{Q_{t_4}}^*$). Zusätzlich wird ein Vergleich des Querschnitts- mit dem Paneldatensatz (Q_{t_4} und $P_{Q_{t_4}}$), der jeweils alle Jugendlichen enthält, durchgeführt. Allerdings muss hier der Vergleich mit der Schulstatistik aus dem genannten Grund entfallen.

Nur zehnte Jahrgangsstufe ($P_{Q_{t_4}}^*$)

Der Panelquerschnitt des Jahres 2005 mit Jugendlichen, die ausschließlich die zehnte Jahrgangsstufe besuchen, enthält 2 943 Fälle. Wie in Tabelle 5.13 ersichtlich, liegt in

Tabelle 5.13: Zusammensetzung nach Geschlecht $P_{Q_{t_4}}^*$

	männlich		weiblich		gesamt N (100 %)
	N	%	N	%	
S_{t_4}	1 760	49,7	1 783	50,3	3 543
$Q_{t_4}^*$	1 627	50,2	1 616	49,8	3 243
$P_{Q_{t_4}}^*$	1 429	48,6	1 514	51,4	2 943

Tabelle 5.14: Altersdurchschnitt $P_{Q_{t_4}}^*$

	Alter	Std.-abw.	gesamt N
$Q_{t_4}^*$	16,0	0,8	3 154
$P_{Q_{t_4}}^*$	16,0	0,7	2 890

Tabelle 5.15: Zusammensetzung $P_{Q_{t_4}}^*$ nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform ^a (%)					gesamt N
		HS	RS	GS	GY	SO	(100 %)
männlich	S_{t_4}	26,3	21,8	32,8	18,8	0,3	1760
	$Q_{t_4}^*$	26,2	22,1	32,6	18,7	0,2	1627
	$P_{Q_{t_4}}^*$	26,8	21,6	32,1	19,3	0,2	1429
weiblich	S_{t_4}	26,3	21,8	32,8	18,8	0,3	1783
	$Q_{t_4}^*$	17,5	22,7	33,4	26,4	0,0	1616
	$P_{Q_{t_4}}^*$	17,4	23,2	32,6	26,8	0,0	1514
gesamt	S_{t_4}	22,5	22,0	33,2	22,1	0,2	3543
	$Q_{t_4}^*$	21,9	22,4	33,0	22,5	0,1	3243
	$P_{Q_{t_4}}^*$	22,0	22,4	32,3	23,2	0,1	2943

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

Bezug auf das Geschlecht der Probanden eine leichte Verschiebung vor. Während im Zehnte-Klasse-Querschnitt 2005 im Vergleich zur Schulstatistik Schülerinnen etwas unterrepräsentiert sind, enthält der Zehnte-Klasse-Panelquerschnitt rund ein Prozent mehr Schülerinnen als in der Schulstatistik ausgewiesen.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panelquerschnitt enthalten sind, weicht nicht vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.14). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik bei nur einem Schuljahrgang kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Aus Tabelle 5.15 wird ersichtlich, dass der Panelquerschnitt 2005 im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen nur leicht verzerrt ist. Im Paneldatensatz befinden sich insgesamt rund ein Prozent weniger Gesamtschülerinnen und -schüler, dafür rund ein Prozent mehr Gymnasiastinnen und Gymnasiasten als in der Schulstatistik ausgewiesen. Die übrigen Anteile sind dabei annähernd gleich groß.

Im Panelquerschnitt sind im Vergleich zur Schulstatistik etwas weniger Hauptschülerinnen und keine Sonderschülerinnen (bei einer schon in der Schulstatistik sehr kleinen Basisrate) mehr vertreten, dagegen weniger Realschüler, Gymnasiasten und Gesamt-

Tabelle 5.16: Zusammensetzung $P_{Q_{t_4}}^*$ nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform ^a (%)					gesamt (%)
		HS	RS	GS	GY	SO	
männlich	S_{t_4}	57,7	49,3	51,5	41,6	76,5	50,9
	$Q_{t_4}^*$	60,1	49,5	49,6	41,7	100,0	50,2
	$P_{Q_{t_4}}^*$	59,2	46,8	48,2	40,5	100,0	48,6
weiblich	S_{t_4}	42,3	50,7	48,5	58,4	23,5	49,1
	$Q_{t_4}^*$	39,9	50,5	50,4	58,3	0,0	49,8
	$P_{Q_{t_4}}^*$	40,8	53,2	51,8	59,5	0,0	51,4

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

schüler (siehe Tabelle 5.16). Der Unterschied zur Schulstatistik ist hierbei – sieht man von der Sonderschule ab – in der Gesamtschule am größten und im Gymnasium am kleinsten. Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Erweiterter Datensatz ($P_{Q_{t_4}}$)

Der Panelquerschnitt des Jahres 2005, der alle befragten Jugendlichen (also Neunt- und Zehntklässler) umfasst, enthält 3 074 Fälle. Dies sind alle überhaupt zuordenbaren Fälle des Jahres 2005. Wie in Tabelle 5.17 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine leichte Verschiebung vor. Im Vergleich zum Querschnitt sind im Panelquerschnitt rund eineinhalb Prozent mehr Schülerinnen vorhanden.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panelquerschnitt enthalten sind, weicht so gut wie nicht vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.18).

Da der erweiterte Paneldatensatz im Gegensatz zur Schulstatistik nicht ausschließlich Jugendliche der neunten Jahrgangsstufe enthält, kann er in seiner Zusammensetzung nach Schulform nur mit dem Querschnittsdatensatz verglichen werden. Aus Tabelle 5.19 wird ersichtlich, dass der Panelquerschnitt 2005 im Vergleich zum Querschnitt im Hinblick

Tabelle 5.17: Zusammensetzung nach Geschlecht $P_{Q_{t_4}}$

	männlich		weiblich		gesamt N (100 %)
	N	%	N	%	
Q_{t_4}	1717	50,4	1688	49,6	3405
$P_{Q_{t_4}}$	1496	48,7	1578	51,3	3074

Tabelle 5.18: Altersdurchschnitt $P_{Q_{t_4}}$

	Alter	Std.-abw.	gesamt N
Q_{t_4}	16,1	0,8	3 313
$P_{Q_{t_4}}$	16,0	0,8	3 020

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

		Schulform ^a (%)							gesamt N
		HS	RS	GS	GY	SO	andere	keine	(100 %)
männlich	Q_{t_4}	25,8	21,2	32,2	18,4	0,7	1,6	0,0	1 704
	$P_{Q_{t_4}}$	26,4	20,8	31,7	19,0	0,5	1,6	0,0	1 487
weiblich	Q_{t_4}	17,5	22,4	33,1	25,6	0,1	1,1	0,2	1 682
	$P_{Q_{t_4}}$	17,5	22,9	32,2	26,1	0,0	1,0	0,3	1 573
gesamt	Q_{t_4}	21,7	21,8	32,7	22,0	0,4	1,4	0,1	3 386
	$P_{Q_{t_4}}$	21,8	21,9	32,0	22,6	0,2	1,3	0,1	3 060

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

auf die Schulform der Jugendlichen nur leicht verzerrt ist. Die Abweichungen der Anteile für alle Schulformen liegen durchweg bei unter einem Prozent.

Im Panelquerschnitt sind im Vergleich zum Querschnitt über alle Schulformen hinweg prozentual mehr Schülerinnen und weniger Schüler vertreten, sieht man von den Sonderschulen, den anderen Schulformen und denjenigen, die auf keine Schule gehen ab, da sie insgesamt nur einen Anteil von unter zwei Prozent ausmachen (siehe Tabelle 5.20). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Tabelle 5.20: Zusammensetzung $P_{Q_{t_4}}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

		Schulform ^a (%)							gesamt
		HS	RS	GS	GY	SO	andere	keine	(%)
männlich	Q_{t_4}	59,8	49,1	49,6	42,2	92,3	59,6	0,0	50,3
	$P_{Q_{t_4}}$	58,7	46,3	48,2	40,8	100,0	60,0	0,0	48,6
weiblich	Q_{t_4}	40,2	50,9	50,4	57,8	7,7	40,4	100,0	49,7
	$P_{Q_{t_4}}$	41,3	53,7	51,8	59,2	0,0	40,0	100,0	51,4

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

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 eines weiteren Zeitpunktes zugeordnet werden konnten. In diesem Bericht werden die Datensätze aller Kombinationen der vier Zeitpunkte vorgestellt. In der Praxis werden allerdings vermutlich die Datensätze zweier benachbarter Zeitpunkte am bedeutsamsten sein.

5.2.1 Zwei-Wellen-Panel 2002/2003 ($P_{t_{1,2}}$)

Das Zwei-Wellen-Panel der Jahre 2002 und 2003 enthält 2 472 Fälle. Wie in Tabelle 5.21 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine leichte Verschiebung vor. Schon in den Querschnitten der Jahre 2002 und 2003 sind im Vergleich zur Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Paneldatensatz ist dieser Unterschied zu beiden Zeitpunkten noch etwas deutlicher: Es sind jeweils rund drei Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen.

Tabelle 5.21: Zusammensetzung nach Geschlecht $P_{t_{1,2}}$

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_1	S_{t_1}	2003	51,2	1907	48,8	3910
	Q_{t_1}	1728	50,7	1679	49,3	3407
	$P_{t_{1,2}}$	1191	48,2	1280	51,8	2471
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_{1,2}}$	1195	48,3	1277	51,7	2472

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht nicht vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.22). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Aus Tabelle 5.23 wird ersichtlich, dass das Zwei-Wellen-Panel im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen systematisch verzerrt ist. Im Paneldatensatz befinden sich zu beiden Zeitpunkten keine Sonderschülerinnen und -schüler mehr. Im Vergleich zur Schulstatistik ist zum Zeitpunkt t_1 der Anteil der Gymnasiastinnen und Gymnasiasten und der Gesamtschülerinnen und -schüler ungefähr gleich, der Anteil der Realschülerinnen und -schüler rund drei Prozent höher, der Hauptschülerinnen und -schüler dagegen rund drei Prozent niedriger. Zum Zeitpunkt

Tabelle 5.22: Altersdurchschnitt $P_{t_{1,2}}$

		Alter	Std.-abw.	gesamt N
t_1	Q_{t_1}	13,0	0,8	3346
	$P_{t_{1,2}}$	13,0	0,7	2434
t_2	Q_{t_2}	14,0	0,8	3275
	$P_{t_{1,2}}$	14,0	0,7	2416

t_2 sind rund zwei Prozent mehr Gymnasiastinnen und Gymnasiasten, knapp eineinhalb Prozent mehr Gesamtschülerinnen und -schüler und rund zweieinhalb Prozent mehr Realschülerinnen und -schüler aber fünfeinhalb Prozent weniger Hauptschülerinnen und -schüler vorhanden als in der Schulstatistik.

Tabelle 5.23: Zusammensetzung $P_{t_{1,2}}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)					gesamt N
			HS	RS	GS	GY	SO	(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
		$P_{t_{1,2}}$	25,0	24,3	32,5	18,1	—	1191
	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
		$P_{t_{1,2}}$	25,7	24,3	32,6	17,5	—	1195
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
		$P_{t_{1,2}}$	17,3	26,0	31,0	25,6	—	1280
	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
		$P_{t_{1,2}}$	18,2	26,5	31,0	24,4	—	1277
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
		$P_{t_{1,2}}$	21,0	25,2	31,7	22,0	—	2471
	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
		$P_{t_{1,2}}$	21,8	25,4	31,8	21,0	—	2472

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

Über alle Schulformen hinweg sind – mit Ausnahme der Sonderschulen – zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik

Tabelle 5.24: Zusammensetzung $P_{1,2}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)					gesamt
			HS	RS	GS	GY	SO	(%)
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
		$P_{t_{1,2}}$	57,3	46,5	49,4	39,7	—	48,2
	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
		$P_{t_{1,2}}$	57,0	46,2	49,6	40,2	—	48,3
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
		$P_{t_{1,2}}$	42,7	53,5	50,6	60,3	—	51,8
	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
		$P_{t_{1,2}}$	43,0	53,8	50,4	59,8	—	51,7

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

vertreten (siehe Tabelle 5.24). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

5.2.2 Zwei-Wellen-Panel 2003/2004 ($P_{t_{2,3}}$)

Das Zwei-Wellen-Panel der Jahre 2003 und 2004 enthält 2 596 Fälle. Wie in Tabelle 5.25 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Schon in den Querschnitten der Jahre 2003 und 2004 sind im Vergleich zur Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Paneldatensatz ist dieser Unterschied zu beiden Zeitpunkten deutlicher: Es sind jeweils rund vier Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen.

Tabelle 5.25: Zusammensetzung nach Geschlecht $P_{t_{2,3}}$

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_2	S_{t_2}	1 927	51,2	1 836	48,8	3 763
	Q_{t_2}	1 703	50,2	1 687	49,8	3 390
	$P_{t_{2,3}}$	1 223	47,1	1 373	52,9	2 596
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
	$P_{t_{2,3}}$	1 213	46,7	1 383	53,3	2 596

Tabelle 5.26: Altersdurchschnitt $P_{t_2,3}$

		Alter	Std.-abw.	gesamt N
t_2	Q_{t_2}	14,0	0,8	3275
	$P_{t_2,3}$	14,0	0,7	2547
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_2,3}$	15,0	0,8	2563

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht kaum vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.26). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Tabelle 5.27: Zusammensetzung $P_{t_2,3}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)					gesamt N	
			HS	RS	GS	GY	SO	(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	
		$P_{t_2,3}$	25,5	21,8	34,4	18,2	—	1223	
	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	
		$P_{t_2,3}$	25,7	22,8	34,0	17,4	—	1213	
	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
			$P_{t_2,3}$	17,0	25,1	32,3	25,6	—	1373
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	
		$P_{t_2,3}$	17,5	25,3	32,6	24,6	—	1383	
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
			$P_{t_2,3}$	21,0	23,5	33,3	22,1	—	2596
	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	
		$P_{t_2,3}$	21,3	24,2	33,3	21,2	—	2596	

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

Aus Tabelle 5.27 wird ersichtlich, dass das Zwei-Wellen-Panel im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen systematisch verzerrt

Tabelle 5.28: Zusammensetzung $P_{t_2,3}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)					gesamt
			HS	RS	GS	GY	SO	(%)
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
		$P_{t_2,3}$	57,2	43,7	48,7	38,8	—	47,1
	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
		$P_{t_2,3}$	56,3	44,2	47,8	38,3	—	46,7
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
		$P_{t_2,3}$	42,8	56,3	51,3	61,2	—	52,9
	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
		$P_{t_2,3}$	43,7	55,8	52,2	61,7	—	53,3

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

ist. Im Paneldatensatz befinden sich zu beiden Zeitpunkten keine Sonderschülerinnen und -schüler mehr. Im Vergleich zur Schulstatistik ist zum Zeitpunkt t_2 der Anteil der Gymnasiastinnen und Gymnasiasten knapp dreieinhalb Prozent, der Anteil der Gesamtschülerinnen und -schüler knapp drei Prozent und der Anteil der Realschülerinnen und -schüler knapp ein Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler sechseinhalb Prozent niedriger als in der Schulstatistik. Zum Zeitpunkt t_3 liegt der Anteil der Gymnasiastinnen und Gymnasiasten sowie der Gesamtschülerinnen und -schüler um rund ein Prozent, der Anteil der Realschülerinnen und -schüler um rund zweieinhalb Prozent höher, der Anteil der Hauptschülerinnen und -schüler dagegen um rund dreieinhalb Prozent niedriger als in der Schulstatistik.

Über alle Schulformen hinweg sind – mit Ausnahme der Sonderschulen – zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.28). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

5.2.3 Zwei-Wellen-Panel 2004/2005 ($P_{t_3,4}^*$ und $P_{t_3,4}$)

Nur zehnte Jahrgangsstufe ($P_{t_3,4}^*$)

Das Zwei-Wellen-Panel der Jahre 2004 und 2005, welches zu t_4 nur den zehnten Jahrgang umfasst, enthält 2 640 Fälle. Wie in Tabelle 5.29 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Schon in den Querschnitten der Jahre

Tabelle 5.29: Zusammensetzung nach Geschlecht $P_{t_3,4}^*$

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_3,4}^*$	1239	46,9	1401	53,1	2640
t_4	S_{t_4}	1760	49,7	1783	50,3	3543
	$Q_{t_4}^*$	1627	50,2	1616	49,8	3243
	$P_{t_3,4}^*$	1250	47,3	1390	52,7	2640

Tabelle 5.30: Altersdurchschnitt $P_{t_3,4}^*$

		Alter	Std.-abw.	gesamt N
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_3,4}^*$	15,0	0,8	2605
t_4	$Q_{t_4}^*$	16,0	0,8	3154
	$P_{t_3,4}^*$	16,0	0,7	2600

2004 und 2005 sind im Vergleich zur Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Paneldatensatz ist dieser Unterschied zu beiden Zeitpunkten deutlicher: Es sind vier (t_3) bzw. knapp zweieinhalb (t_4) Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht kaum vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.30). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Aus Tabelle 5.31 wird ersichtlich, dass das Zwei-Wellen-Panel im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen systematisch verzerrt ist. Sieht man von dem insgesamt sehr niedrigen Anteil der Sonderschülerinnen und -schüler ab, so ist im Vergleich zur Schulstatistik zum Zeitpunkt t_3 der Anteil der Gesamtschülerinnen und -schüler gleich, der Anteil der Gymnasiastinnen und Gymnasiasten dreieinhalb Prozent und der Anteil der Realschülerinnen und -schüler eineinhalb Prozent höher und der Anteil der Hauptschülerinnen und -schüler knapp vier Prozent niedriger. Zum Zeitpunkt t_4 liegt der Anteil der Gymnasiastinnen und Gymnasiasten sowie der Realschülerinnen und -schüler rund ein Prozent höher, dagegen der Anteil der Gesamtschülerinnen und -schüler und der Anteil der Hauptschülerinnen und -schüler rund ein Prozent niedriger als in der Schulstatistik.

Tabelle 5.31: Zusammensetzung $P_{t_3,4}^*$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)					gesamt <i>N</i>
			HS	RS	GS	GY	SO	(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
		$P_{t_3,4}^*$	25,4	21,7	32,8	19,9	0,2	1239
	t_4	S_{t_4}	26,2	22,1	32,6	18,7	0,2	1627
		$Q_{t_4}^*$	26,8	21,6	32,1	19,3	0,2	1429
		$P_{t_3,4}^*$	26,2	21,6	32,3	19,7	0,2	1250
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
		$P_{t_3,4}^*$	17,2	24,2	31,9	26,7	0,0	1401
	t_4	S_{t_4}	26,3	21,8	32,8	18,8	0,3	1783
		$Q_{t_4}^*$	17,5	22,7	33,4	26,4	0,0	1616
		$P_{t_3,4}^*$	16,8	24,1	32,5	26,5	0,0	1390
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
		$P_{t_3,4}^*$	21,1	23,0	32,3	23,5	0,1	2640
	t_4	S_{t_4}	22,5	22,0	33,2	22,1	0,2	3543
		$Q_{t_4}^*$	21,9	22,4	33,0	22,5	0,1	3243
		$P_{t_3,4}^*$	21,3	22,9	32,4	23,3	0,1	2640

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

Bis auf die Hauptschulen sind über alle Schulformen hinweg – ebenfalls mit Ausnahme der Sonderschulen – zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.32). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Tabelle 5.32: Zusammensetzung $P_{t_{3,4}}^*$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)					gesamt
			HS	RS	GS	GY	SO	(%)
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
		$P_{t_{3,4}}^*$	56,7	44,2	47,6	39,8	100,0	46,9
	t_4	S_{t_4}	57,7	49,3	51,5	41,6	76,5	50,9
		$Q_{t_4}^*$	60,1	49,5	49,6	41,7	100,0	50,2
		$P_{t_{3,4}}^*$	58,4	44,6	47,2	40,0	100,0	47,3
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
		$P_{t_{3,4}}^*$	43,3	55,8	52,4	60,2	0,0	53,1
	t_4	S_{t_4}	42,3	50,7	48,5	58,4	23,5	49,1
		$Q_{t_4}^*$	39,9	50,5	50,4	58,3	0,0	49,8
		$P_{t_{3,4}}^*$	41,6	55,4	52,8	60,0	0,0	52,7

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

Erweiterter Datensatz ($P_{t_{3,4}}$)

Das Zwei-Wellen-Panel der Jahre 2004 und 2005, welches zu t_4 alle befragten Jugendlichen (also Neunt- und Zehntklässler) umfasst, enthält 2 750 Fälle. Wie in Tabelle 5.33 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Im Vergleich zur Schulstatistik (t_3) und im Vergleich zum Querschnitt (t_4) sind die Schülerinnen überrepräsentiert.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht kaum vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.34). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter

Tabelle 5.33: Zusammensetzung nach Geschlecht $P_{t_{3,4}}$

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_{3,4}}$	1289	46,9	1461	53,1	2750
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_{3,4}}$	1303	47,4	1447	52,6	2750

Tabelle 5.34: Altersdurchschnitt $P_{t_{3,4}}$

		Alter	Std.-abw.	gesamt N
t_3	Q_{t_3}	15,1	0,8	3 270
	$P_{t_{3,4}}$	15,0	0,8	2 709
t_4	Q_{t_4}	16,1	0,8	3 313
	$P_{t_{3,4}}$	16,0	0,7	2 709

differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Da der erweiterte Paneldatensatz im Gegensatz zur Schulstatistik zum Zeitpunkt t_4 nicht ausschließlich Jugendliche der neunten Jahrgangsstufe enthält, kann er in seiner Zusammensetzung nach Schulform zum Zeitpunkt t_4 nur mit dem Querschnittsdatensatz verglichen werden. Aus Tabelle 5.35 wird ersichtlich, dass das Zwei-Wellen-Panel 2004/2005 im Vergleich zur Schulstatistik (t_3) und zum Querschnitt (t_4) im Hinblick auf die Schulform der Jugendlichen leicht verzerrt ist. Für die Betrachtungen wird von den Sonderschülerinnen und -schülern, denjenigen, die eine »andere« und denjenigen, die keine Schule mehr besuchen, abgesehen, da sie insgesamt nur einen Anteil von unter zwei Prozent ausmachen. Zum Zeitpunkt t_3 ist im Vergleich zur Schulstatistik der Anteil der Gesamtschülerinnen und -schüler ungefähr gleich groß, der Anteil der Gymnasiastinnen und Gymnasiasten ein Prozent, der Anteil der Realschülerinnen und -schüler knapp ein Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler knapp drei Prozent niedriger. Zum Zeitpunkt t_4 ist im Vergleich zum Querschnittsdatensatz der Anteil der Gymnasiasten und Gymnasiastinnen knapp ein Prozent, der Anteil der Realschülerinnen und -schüler rund ein halbes Prozent höher, dagegen der Anteil der Gesamtschülerinnen und -schüler und der Hauptschülerinnen und -schüler rund ein halbes Prozent niedriger.

Im Paneldatensatz sind im Vergleich zur Schulstatistik (t_3) und zum Querschnitt (t_4) über alle Schulformen hinweg mehr Schülerinnen und weniger Schüler vertreten, sieht man ebenfalls von den Sonderschulen, den anderen Schulformen und denjenigen, die auf keine Schule gehen ab (siehe Tabelle 5.36). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Tabelle 5.35: Zusammensetzung $P_{t3,4}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)						gesamt N	
			HS	RS	GS	GY	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
		$P_{t_3,4}$	26,4	20,9	32,7	19,6	0,4	—	—	1289
	t_4	Q_{t_4}	25,8	21,2	32,2	18,4	0,7	1,6	0,0	1704
		$P_{t_3,4}$	25,8	20,9	31,9	19,4	0,2	1,8	0,0	1296
	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
$P_{t_3,4}$			18,4	23,8	31,8	25,9	0,1	—	—	1461
t_4		Q_{t_4}	17,5	22,9	32,2	26,1	0,0	1,0	0,3	1573
		$P_{t_3,4}$	16,8	23,8	32,2	25,8	0,0	1,1	0,3	1443
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
	$P_{t_3,4}$		22,1	22,4	32,2	23,0	0,2	—	—	2750
	t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,4	1,4	0,1	3386
		$P_{t_3,4}$	21,1	22,4	32,0	22,8	0,1	1,4	0,1	2739

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

Tabelle 5.36: Zusammensetzung $P_{t3,4}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)						gesamt	
			HS	RS	GS	GY	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
		$P_{t_3,4}$	55,8	43,8	47,5	40,0	83,3	—	—	46,9
	t_4	Q_{t_4}	59,8	49,1	49,6	42,2	92,3	59,6	0,0	50,3
		$P_{t_3,4}$	58,0	44,1	47,1	40,2	100,0	59,0	0,0	47,3
	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
$P_{t_3,4}$			44,2	56,2	52,5	60,0	16,7	—	—	53,1
t_4		Q_{t_4}	40,2	50,9	50,4	57,8	7,7	40,4	100,0	49,7
		$P_{t_3,4}$	42,0	55,9	52,9	59,8	0,0	41,0	100,0	52,7

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

5.2.4 Zwei-Wellen-Panel 2002/2004 ($P_{t_1,3}$)

Das Zwei-Wellen-Panel der Jahre 2002 und 2004 enthält 2 202 Fälle. Wie in Tabelle 5.37 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Schon in den Querschnitten der Jahre 2002 und 2004 sind im Vergleich zur Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Paneldatensatz ist dieser Unterschied zu beiden Zeitpunkten noch deutlicher: Es sind fünf (t_1) bzw. viereinhalb (t_3) Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen.

Tabelle 5.37: Zusammensetzung nach Geschlecht $P_{t_1,3}$

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_1	S_{t_1}	2003	51,2	1907	48,8	3910
	Q_{t_1}	1728	50,7	1679	49,3	3407
	$P_{t_1,3}$	1015	46,1	1187	53,9	2202
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_1,3}$	1017	46,2	1185	53,8	2202

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht kaum vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.38). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Aus Tabelle 5.39 wird ersichtlich, dass das Zwei-Wellen-Panel im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen systematisch verzerrt ist. Im Folgenden wird die Schulform »Sonderschule« nicht berücksichtigt, da sie insgesamt nur von unter einem Prozent der Jugendlichen besucht wird. Im Vergleich zur Schulstatistik ist zum Zeitpunkt t_1 der Anteil der Gymnasiastinnen und Gymnasiasten rund vier Prozent, der Anteil der Gesamtschülerinnen und -schüler knapp ein halbes Prozent und der Anteil der Realschülerinnen und -schüler knapp zwei Prozent höher, dagegen der

Tabelle 5.38: Altersdurchschnitt $P_{t_1,3}$

		Alter	Std.-abw.	gesamt N
t_1	Q_{t_1}	13,0	0,8	3346
	$P_{t_1,3}$	12,9	0,7	2175
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_1,3}$	15,0	0,7	2183

Tabelle 5.39: Zusammensetzung $P_{t_1,3}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)					gesamt N	
			HS	RS	GS	GY	SO	(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	
		$P_{t_1,3}$	21,8	22,5	33,8	21,8	0,2	1015	
	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	
		$P_{t_1,3}$	22,3	23,6	32,9	20,8	0,3	1017	
	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
			$P_{t_1,3}$	15,2	25,2	31,0	28,6	0,0	1187
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	
		$P_{t_1,3}$	16,1	25,5	31,6	26,8	0,0	1185	
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
			$P_{t_1,3}$	18,3	23,9	32,3	25,4	0,1	2202
	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	
		$P_{t_1,3}$	19,0	24,6	32,2	24,1	0,1	2202	

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

Anteil der Hauptschülerinnen und -schüler rund fünfeinhalb Prozent niedriger als in der Schulstatistik. Zum Zeitpunkt t_3 liegt der Anteil der Gymnasiastinnen und Gymnasiasten rund vier Prozent, der Anteil der Realschülerinnen und -schüler rund drei Prozent höher, der Anteil der Hauptschülerinnen und -schüler dagegen knapp sechs Prozent niedriger als in der Schulstatistik. Der Anteil der Gesamtschülerinnen und -schüler ist dabei ungefähr gleich groß.

Über alle Schulformen hinweg sind – ebenfalls mit Ausnahme der Sonderschulen – zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.40). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Tabelle 5.40: Zusammensetzung $P_{t_1,3}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)					gesamt	
			HS	RS	GS	GY	SO	(%)	
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	
		$P_{t_1,3}$	55,0	43,3	48,2	39,5	100,0	46,1	
	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	
		$P_{t_1,3}$	54,3	44,3	47,2	40,0	100,0	46,2	
	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
			$P_{t_1,3}$	45,0	56,7	51,8	60,5	0,0	53,9
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	
		$P_{t_1,3}$	45,7	55,7	52,8	60,0	0,0	53,8	

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

5.2.5 Zwei-Wellen-Panel 2002/2005 ($P_{t_1,4}^*$ und $P_{t_1,4}$)

Nur zehnte Jahrgangsstufe ($P_{t_1,4}^*$)

Das Zwei-Wellen-Panel der Jahre 2002 und 2005, welches zu t_4 nur den zehnten Jahrgang umfasst, enthält 2 086 Fälle. Wie in Tabelle 5.41 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Schon in den Querschnitten der Jahre 2002 und 2005 sind im Vergleich zur Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Paneldatensatz ist dieser Unterschied zu beiden Zeitpunkten noch deutlicher: Es sind rund viereinhalb (t_1) bzw. knapp drei (t_4) Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen.

Tabelle 5.41: Zusammensetzung nach Geschlecht $P_{t_1,4}^*$

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_1	S_{t_1}	2003	51,2	1907	48,8	3910
	Q_{t_1}	1728	50,7	1679	49,3	3407
	$P_{t_1,4}^*$	973	46,6	1113	53,4	2086
t_4	S_{t_4}	1760	49,7	1783	50,3	3543
	$Q_{t_4}^*$	1627	50,2	1616	49,8	3243
	$P_{t_1,4}^*$	978	46,9	1108	53,1	2086

Tabelle 5.42: Altersdurchschnitt $P_{t_1,4}^*$

		Alter	Std.-abw.	gesamt N
t_1	Q_{t_1}	13,0	0,8	3346
	$P_{t_1,4}^*$	12,9	0,7	2060
t_4	$Q_{t_4}^*$	16,0	0,8	3154
	$P_{t_1,4}^*$	15,9	0,7	2056

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht kaum vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.42). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Tabelle 5.43: Zusammensetzung $P_{t_1,4}^*$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)					gesamt N
			HS	RS	GS	GY	SO	(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
		$P_{t_1,4}^*$	23,2	22,1	32,2	22,5	—	973
	t_4	S_{t_4}	26,2	22,1	32,6	18,7	0,2	1627
		$Q_{t_4}^*$	26,8	21,6	32,1	19,3	0,2	1429
		$P_{t_1,4}^*$	23,0	22,8	32,8	21,3	0,1	978
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
		$P_{t_1,4}^*$	14,6	25,1	31,1	29,2	—	1113
	t_4	S_{t_4}	26,3	21,8	32,8	18,8	0,3	1783
		$Q_{t_4}^*$	17,5	22,7	33,4	26,4	0,0	1616
		$P_{t_1,4}^*$	15,3	25,0	31,9	27,8	0,0	1108
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
		$P_{t_1,4}^*$	18,6	23,7	31,6	26,1	—	2086
	t_4	S_{t_4}	22,5	22,0	33,2	22,1	0,2	3543
		$Q_{t_4}^*$	21,9	22,4	33,0	22,5	0,1	3243
		$P_{t_1,4}^*$	18,9	24,0	32,3	24,7	0,0	2086

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

Aus Tabelle 5.43 wird ersichtlich, dass das Zwei-Wellen-Panel im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen systematisch verzerrt ist. Sieht man von dem insgesamt sehr niedrigen Anteil der Sonderschülerinnen und -schüler ab, so ist im Vergleich zur Schulstatistik zum Zeitpunkt t_1 der Anteil der Gymnasiastinnen und Gymnasiasten fast fünf Prozent und der Anteil der Realschülerinnen und -schüler rund eineinhalb Prozent höher, dagegen der Anteil der Gesamtschülerinnen und -schüler knapp ein halbes Prozent und der Anteil der Hauptschülerinnen und -schüler knapp fünfzehn Prozent niedriger. Zum Zeitpunkt t_4 liegt der Anteil der Gymnasiastinnen und Gymnasiasten rund zweieinhalb Prozent und der Anteil der Realschülerinnen und -schüler zwei Prozent höher, dagegen der Anteil der Gesamtschülerinnen und -schüler knapp ein Prozent und der Anteil der Hauptschülerinnen und -schüler rund dreieinhalb Prozent niedriger als in der Schulstatistik.

Tabelle 5.44: Zusammensetzung $P_{t_1,4}^*$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)					gesamt
			HS	RS	GS	GY	SO	(%)
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
		$P_{t_1,4}^*$	58,1	43,5	47,5	40,3	—	46,6
	t_4	S_{t_4}	57,7	49,3	51,5	41,6	76,5	50,9
		$Q_{t_4}^*$	60,1	49,5	49,6	41,7	100,0	50,2
		$P_{t_1,4}^*$	57,0	44,6	47,6	40,3	100,0	46,9
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
		$P_{t_1,4}^*$	41,9	56,5	52,5	59,7	—	53,4
	t_4	S_{t_4}	42,3	50,7	48,5	58,4	23,5	49,1
		$Q_{t_4}^*$	39,9	50,5	50,4	58,3	0,0	49,8
		$P_{t_1,4}^*$	43,0	55,4	52,4	59,7	0,0	53,1

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

Im Paneldatensatz sind über alle Schulformen hinweg – ebenfalls mit Ausnahme der Sonderschulen – zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.44). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Erweiterter Datensatz ($P_{t_1,4}$)

Das Zwei-Wellen-Panel der Jahre 2002 und 2005, welches zu t_4 alle befragten Jugendlichen (also Neunt- und Zehntklässler) umfasst, enthält 2 165 Fälle. Wie in Tabelle 5.45

Tabelle 5.45: Zusammensetzung nach Geschlecht $P_{t_1,4}$

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_1	S_{t_1}	2003	51,2	1907	48,8	3910
	Q_{t_1}	1728	50,7	1679	49,3	3407
	$P_{t_1,4}$	1008	46,6	1157	53,4	2165
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_1,4}$	1015	46,9	1150	53,1	2165

Tabelle 5.46: Altersdurchschnitt $P_{t_1,4}$

		Alter	Std.-abw.	gesamt N
t_1	Q_{t_1}	13,0	0,8	3346
	$P_{t_1,4}$	12,9	0,7	2137
t_4	Q_{t_4}	16,1	0,8	3313
	$P_{t_1,4}$	15,9	0,7	2134

ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Im Vergleich zur Schulstatistik (t_1) und im Vergleich zum Querschnitt (t_4) sind die Schülerinnen um viereinhalb bzw. dreieinhalb Prozent überrepräsentiert.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht kaum vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.46). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Da der erweiterte Paneldatensatz im Gegensatz zur Schulstatistik zum Zeitpunkt t_4 nicht ausschließlich Jugendliche der neunten Jahrgangsstufe enthält, kann er in seiner Zusammensetzung nach Schulform zum Zeitpunkt t_4 nur mit dem Querschnittsdatensatz verglichen werden. Aus Tabelle 5.47 wird ersichtlich, dass das Zwei-Wellen-Panel 2002/2005 im Vergleich zur Schulstatistik (t_1) und zum Querschnitt (t_4) im Hinblick auf die Schulform der Jugendlichen verzerrt ist. Für die Betrachtungen wird von den Sonderschülerinnen und -schülern, denjenigen, die eine »andere« und denjenigen, die keine Schule mehr besuchen, abgesehen, da sie insgesamt nur einen Anteil von unter zwei Prozent ausmachen. Zum Zeitpunkt t_1 ist im Vergleich zur Schulstatistik der Anteil der Gesamtschülerinnen und -schüler ungefähr gleich groß, der Anteil der Gymnasiastinnen und Gymnasiasten knapp vier Prozent und der Anteil der Realschülerinnen und -schüler rund ein Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler knapp viereinhalb Prozent niedriger. Zum Zeitpunkt t_4 ist im Vergleich zum Querschnitts-

Tabelle 5.47: Zusammensetzung $P_{t_1,4}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)							gesamt N
			HS	RS	GS	GY	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
		$P_{t_1,4}$	24,1	21,6	31,9	22,0	0,3	—	—	1008
	t_4	Q_{t_4}	25,8	21,2	32,2	18,4	0,7	1,6	0,0	1704
		$P_{t_1,4}$	23,0	22,1	32,8	20,9	0,4	0,9	—	1010
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
		$P_{t_1,4}$	15,6	24,5	31,5	28,4	0,0	—	—	1157
	t_4	Q_{t_4}	17,5	22,9	32,2	26,1	0,0	1,0	0,3	1573
		$P_{t_1,4}$	15,6	24,7	31,7	27,0	0,0	1,0	—	1148
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
		$P_{t_1,4}$	19,5	23,2	31,7	25,5	0,1	—	—	2165
	t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,4	1,4	0,1	3386
		$P_{t_1,4}$	19,0	23,4	32,2	24,1	0,2	1,0	—	2158

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

datensatz der Anteil der Gymnasiasten und Gymnasiastinnen rund zwei Prozent und der Anteil der Realschülerinnen und -schüler rund eineinhalb Prozent höher, dagegen der Anteil der Gesamtschülerinnen und -schüler ein halbes Prozent und der Anteil der Hauptschülerinnen und -schüler rund zweieinhalb Prozent niedriger.

Im Panelquerschnitt sind im Vergleich zur Schulstatistik (t_1) und zum Querschnitt (t_4) über alle Schulformen hinweg mehr Schülerinnen und weniger Schüler vertreten, sieht man ebenfalls von den Sonderschulen, den anderen Schulformen und denjenigen, die auf keine Schule gehen, ab (siehe Tabelle 5.48). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Tabelle 5.48: Zusammensetzung $P_{t_1,4}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)							gesamt
			HS	RS	GS	GY	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
		$P_{t_1,4}$	57,4	43,4	46,9	40,3	100,0	—	—	46,6
	t_4	Q_{t_4}	59,8	49,1	49,6	42,2	92,3	59,6	0,0	50,3
		$P_{t_1,4}$	56,4	44,1	47,6	40,5	100,0	42,9	—	46,8
	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
$P_{t_1,4}$			42,6	56,6	53,1	59,7	0,0	—	—	53,4
t_4		Q_{t_4}	40,2	50,9	50,4	57,8	7,7	40,4	100,0	49,7
		$P_{t_1,4}$	43,6	55,9	52,4	59,5	0,0	57,1	—	53,2

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

5.2.6 Zwei-Wellen-Panel 2003/2005 ($P_{t_2,4}^*$ und $P_{t_2,4}$)

Nur zehnte Jahrgangsstufe ($P_{t_2,4}^*$)

Das Zwei-Wellen-Panel der Jahre 2003 und 2005, welches zu t_4 nur den zehnten Jahrgang umfasst, enthält 2 411 Fälle. Wie in Tabelle 5.49 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Schon in den Querschnitten der Jahre 2003 und 2005 sind im Vergleich zur Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Paneldatensatz ist dieser Unterschied zu beiden Zeitpunkten noch deutlicher: Es sind rund viereinhalb (t_2) bzw. knapp drei (t_4) Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht nicht vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.50). Referenz-

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

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_2,4}^*$	1129	46,8	1282	53,2	2411
t_4	S_{t_4}	1760	49,7	1783	50,3	3543
	$Q_{t_4}^*$	1627	50,2	1616	49,8	3243
	$P_{t_2,4}^*$	1131	46,9	1280	53,1	2411

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

		Alter	Std.-abw.	gesamt N
t_2	Q_{t_2}	14,0	0,8	3 275
	$P_{t_2,4}^*$	14,0	0,7	2 365
t_4	$Q_{t_4}^*$	16,0	0,8	3 154
	$P_{t_2,4}^*$	16,0	0,7	2 371

daten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

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

			Schulform ^a (%)					gesamt N (100 %)
			HS	RS	GS	GY	SO	
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
		$P_{t_2,4}^*$	26,5	21,6	33,7	18,2	—	1 129
	t_4	S_{t_4}	26,2	22,1	32,6	18,7	0,2	1 627
		$Q_{t_4}^*$	26,8	21,6	32,1	19,3	0,2	1 429
		$P_{t_2,4}^*$	26,6	21,4	34,5	17,4	0,1	1 131
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
		$P_{t_2,4}^*$	17,9	24,4	32,4	25,4	—	1 282
	t_4	S_{t_4}	26,3	21,8	32,8	18,8	0,3	1 783
		$Q_{t_4}^*$	17,5	22,7	31,4	26,4	0,0	1 616
		$P_{t_2,4}^*$	17,6	24,3	33,0	25,2	0,0	1 280
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
		$P_{t_2,4}^*$	21,9	23,1	33,0	22,0	—	2 411
	t_4	S_{t_4}	22,5	22,0	33,2	22,1	0,2	3 543
		$Q_{t_4}^*$	21,9	22,4	33,0	22,5	0,1	3 243
		$P_{t_2,4}^*$	21,8	22,9	33,7	21,5	0,0	2 411

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

Aus Tabelle 5.51 wird ersichtlich, dass das Zwei-Wellen-Panel im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen systematisch verzerrt ist. Sieht man von dem insgesamt sehr niedrigen Anteil der Sonderschülerinnen und -schüler ab,

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

			Schulform ^a (%)					gesamt
			HS	RS	GS	GY	SO	(%)
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
		$P_{t_2,4}^*$	56,6	43,8	47,9	38,7	—	46,8
	t_4	S_{t_4}	57,7	49,3	51,5	41,6	76,5	50,9
		$Q_{t_4}^*$	60,1	49,5	49,6	41,7	100,0	50,2
		$P_{t_2,4}^*$	57,2	43,8	48,0	38,0	100,0	46,9
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
		$P_{t_2,4}^*$	43,4	56,2	52,1	61,3	—	53,2
	t_4	S_{t_4}	42,3	50,7	48,5	58,4	23,5	49,1
		$Q_{t_4}^*$	39,9	50,5	50,4	58,3	0,0	49,8
		$P_{t_2,4}^*$	42,8	56,2	52,0	62,0	0,0	53,1

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

so ist im Vergleich zur Schulstatistik zum Zeitpunkt t_2 der Anteil der Gymnasiastinnen und Gymnasiasten knapp dreieinhalb Prozent, der Anteil der Gesamtschülerinnen und -schüler rund zweieinhalb Prozent und der Anteil der Realschülerinnen und -schüler knapp ein halbes Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler knapp fünfeinhalb Prozent niedriger. Zum Zeitpunkt t_4 liegt der Anteil der Gesamtschülerinnen und -schüler ein halbes Prozent und der Anteil der Realschülerinnen und -schüler knapp ein Prozent höher, dagegen der Anteil der Gymnasiastinnen und Gymnasiasten und der Anteil der Hauptschülerinnen und -schüler rund ein halbes Prozent niedriger als in der Schulstatistik.

Im Paneldatensatz sind über alle Schulformen hinweg – ebenfalls mit Ausnahme der Sonderschulen – zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.52). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Erweiterter Datensatz ($P_{t_2,4}$)

Das Zwei-Wellen-Panel der Jahre 2003 und 2005, welches zu t_4 alle befragten Jugendlichen (also Neunt- und Zehntklässler) umfasst, enthält 2 493 Fälle. Wie in Tabelle 5.53 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Im Vergleich zur Schulstatistik (t_2) und im Vergleich zum Querschnitt (t_4) sind die Schülerinnen um viereinhalb bzw. dreieinhalb Prozent überrepräsentiert.

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

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_2,4}$	1168	46,9	1325	53,1	2493
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_2,4}$	1171	47,0	1322	53,0	2493

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

		Alter	Std.-abw.	gesamt N
t_2	Q_{t_2}	14,0	0,8	3275
	$P_{t_2,4}$	14,0	0,7	2446
t_4	Q_{t_4}	16,1	0,8	3313
	$P_{t_2,4}$	16,0	0,7	2453

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht kaum vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.54). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Da der erweiterte Paneldatensatz im Gegensatz zur Schulstatistik zum Zeitpunkt t_4 nicht ausschließlich Jugendliche der neunten Jahrgangsstufe enthält, kann er in seiner Zusammensetzung nach Schulform zum Zeitpunkt t_4 nur mit dem Querschnittsdatensatz verglichen werden. Aus Tabelle 5.55 wird ersichtlich, dass das Zwei-Wellen-Panel 2003/2005 im Vergleich zur Schulstatistik (t_2) und zum Querschnitt (t_4) im Hinblick auf die Schulform der Jugendlichen verzerrt ist. Für die Betrachtungen wird von den Sonderschülerinnen und -schülern, denjenigen, die eine »andere« und denjenigen, die keine Schule mehr besuchen, abgesehen, da sie insgesamt nur einen Anteil von unter zwei Prozent ausmachen. Zum Zeitpunkt t_2 ist im Vergleich zur Schulstatistik der Anteil der Realschülerinnen und -schüler ungefähr gleich groß, der Anteil der Gymnasiastinnen und Gymnasiasten drei Prozent und der Anteil der Gesamtschülerinnen und -schüler rund zweieinhalb Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler knapp fünf Prozent niedriger. Zum Zeitpunkt t_4 ist im Vergleich zum Querschnittsdatensatz der Anteil der Hauptschülerinnen und -schüler annähernd gleichgroß, der Anteil der Gesamtschülerinnen und -schüler knapp ein Prozent und der Anteil der Realschülerinnen

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

			Schulform ^a (%)						gesamt N	
			HS	RS	GS	GY	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
		$P_{t_2,4}$	27,2	20,9	33,8	18,1	—	—	—	1168
	t_4	Q_{t_4}	25,8	21,2	32,2	18,4	0,7	1,6	0,0	1704
		$P_{t_2,4}$	26,4	20,7	34,4	17,3	0,1	1,1	—	1167
	weiblich	t_2	S_{t_2}	23,1	23,5	30,9	22,4	0,0	—	—
Q_{t_2}			20,1	24,9	31,5	23,5	0,0	—	—	1687
$P_{t_2,4}$			18,6	24,2	32,4	24,9	—	—	—	1325
t_4		Q_{t_4}	17,5	22,9	32,2	26,1	0,0	1,0	0,3	1573
		$P_{t_2,4}$	17,7	24,1	32,8	24,7	0,0	0,8	—	1321
gesamt		t_2	S_{t_2}	27,5	22,7	30,4	18,7	0,7	—	—
	Q_{t_2}		24,6	23,6	31,4	20,1	0,3	—	—	3390
	$P_{t_2,4}$		22,6	22,6	33,1	21,7	—	—	—	2493
	t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,4	1,4	0,1	3386
		$P_{t_2,4}$	21,8	22,5	33,5	21,2	0,0	0,9	—	2488

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

und -schüler rund ein halbes Prozent höher, dagegen der Anteil der Gymnasiastinnen und Gymnasiasten knapp ein Prozent niedriger.

Im Paneldatensatz sind im Vergleich zur Schulstatistik (t_1) und zum Querschnitt (t_4) über alle Schulformen hinweg mehr Schülerinnen und weniger Schüler vertreten, sieht man ebenfalls von den Sonderschulen, den anderen Schulformen und denjenigen, die auf keine Schule gehen, ab (siehe Tabelle 5.48). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Tabelle 5.56: Zusammensetzung $P_{2,4}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)							gesamt
			HS	RS	GS	GY	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
		$P_{2,4}$	56,4	43,3	47,9	39,0	—	—	—	46,9
	t_4	Q_{t_4}	59,8	49,1	49,6	42,2	92,3	59,6	0,0	50,3
		$P_{2,4}$	56,8	43,2	48,1	38,3	100,0	56,5	—	46,9
	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
$P_{2,4}$			43,6	56,7	52,1	61,0	—	—	—	53,1
t_4		Q_{t_4}	40,2	50,9	50,4	57,8	7,7	40,4	100,0	49,7
		$P_{2,4}$	43,2	56,8	51,9	61,7	0,0	43,5	—	53,1

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

5.3 Die Drei-Wellen-Panel

Die Datensätze der Drei-Wellen-Panel enthalten alle Fälle, deren zugehörige – und nach Plausibilitätskontrollen gültige – Fragebögen zwei (ebenfalls gültigen) Bögen zwei weiterer Zeitpunkte zugeordnet werden konnten. In diesem Bericht werden die Datensätze aller vier Kombinationen von drei der insgesamt vier Zeitpunkte vorgestellt. In der Praxis werden vermutlich die beiden Datensätze dreier benachbarter Zeitpunkte am bedeutsamsten sein.

5.3.1 Drei-Wellen-Panel 2002 bis 2004 ($P_{1,2,3}$)

Das Drei-Wellen-Panel der Jahre 2002 bis 2004 enthält 2 012 Fälle. Wie in Tabelle 5.57 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Schon in den Querschnitten der Jahre 2002 bis 2004 sind im Vergleich zur Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Paneldatensatz ist dieser Unterschied zu allen drei Zeitpunkten noch deutlicher: Es sind rund sechs (t_1) und rund fünfeinhalb (t_2 und t_3) Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht kaum vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.58). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Aus Tabelle 5.59 wird ersichtlich, dass das Drei-Wellen-Panel im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen systematisch verzerrt ist. Im

Tabelle 5.57: Zusammensetzung nach Geschlecht $P_{t_1,2,3}$

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_1	S_{t_1}	2003	51,2	1907	48,8	3910
	Q_{t_1}	1728	50,7	1679	49,3	3407
	$P_{t_1,2,3}$	906	45,0	1106	55,0	2012
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_1,2,3}$	916	45,5	1096	54,5	2012
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_1,2,3}$	909	45,2	1103	54,8	2012

Tabelle 5.58: Altersdurchschnitt $P_{t_1,2,3}$

		Alter	Std.-abw.	gesamt N
t_1	Q_{t_1}	13,0	0,8	3346
	$P_{t_1,2,3}$	12,9	0,7	1987
t_2	Q_{t_2}	14,0	0,8	3275
	$P_{t_1,2,3}$	13,9	0,7	1981
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_1,2,3}$	15,0	0,7	1995

Folgenden wird die Schulform »Sonderschule« nicht berücksichtigt, da sie insgesamt nur von unter einem Prozent der Jugendlichen besucht wird. Im Vergleich zur Schulstatistik ist zum Zeitpunkt t_1 der Anteil der Gymnasiastinnen und Gymnasiasten knapp zwei Prozent, der Anteil der Gesamtschülerinnen und -schüler knapp eineinhalb Prozent und der Anteil der Realschülerinnen und -schüler knapp drei Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler rund fünfeinhalb Prozent niedriger als in der Schulstatistik. Zum Zeitpunkt t_2 liegt der Anteil der Gymnasiastinnen und Gymnasiasten rund dreieinhalb Prozent, der Gesamtschülerinnen und -schüler rund drei Prozent und der Anteil der Realschülerinnen und -schüler rund zweieinhalb Prozent höher, der Anteil der Hauptschülerinnen und -schüler dagegen knapp achteinhalb Prozent niedriger als in der Schulstatistik. Zum Zeitpunkt t_3 liegt der Anteil der Gymnasiastinnen und Gymnasiasten knapp zwei Prozent, der Gesamtschülerinnen und -schüler ein Prozent und der Anteil der Realschülerinnen und -schüler rund vier Prozent höher, der Anteil der Hauptschülerinnen und -schüler dagegen rund fünfeinhalb Prozent niedriger als in der Schulstatistik.

Über alle Schulformen hinweg sind – ebenfalls mit Ausnahme der Sonderschulen – zu allen drei Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der

Tabelle 5.59: Zusammensetzung $P_{t_1,2,3}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)					gesamt N
			HS	RS	GS	GY	SO	(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
		$P_{t_1,2,3}$	22,2	23,2	35,8	18,9	—	906
	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
		$P_{t_1,2,3}$	22,6	23,1	35,9	18,3	—	916
	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
		$P_{t_1,2,3}$	22,7	24,4	34,9	18,0	—	909
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
		$P_{t_1,2,3}$	15,6	26,3	31,4	26,7	—	1106
	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
		$P_{t_1,2,3}$	16,2	26,6	31,6	25,6	—	1096
	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
		$P_{t_1,2,3}$	16,4	26,7	32,0	24,8	—	1103
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
		$P_{t_1,2,3}$	18,6	24,9	33,3	23,2	—	2012
	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
		$P_{t_1,2,3}$	19,1	25,0	33,5	22,3	—	2012
	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
		$P_{t_1,2,3}$	19,2	25,7	33,3	21,8	—	2012

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

Tabelle 5.60: Zusammensetzung $P_{t_1,2,3}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)					gesamt
			HS	RS	GS	GY	SO	(%)
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
		$P_{t_1,2,3}$	53,7	41,9	48,3	36,7	—	45,0
	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
		$P_{t_1,2,3}$	53,8	42,1	48,7	37,4	—	45,5
	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
		$P_{t_1,2,3}$	53,2	42,9	47,3	37,4	—	45,2
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
		$P_{t_1,2,3}$	46,3	58,1	51,7	63,3	—	55,0
	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
		$P_{t_1,2,3}$	46,2	57,9	51,3	62,6	—	54,5
	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
		$P_{t_1,2,3}$	46,8	57,1	52,7	62,6	—	54,8

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

Schulstatistik vertreten (siehe Tabelle 5.60). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

5.3.2 Drei-Wellen-Panel 2003 bis 2005 ($P_{t_2,3,4}^*$ und $P_{t_2,3,4}$)

Nur zehnte Jahrgangsstufe ($P_{t_2,3,4}^*$)

Das Drei-Wellen-Panel der Jahre 2003 bis 2005, welches zu t_4 nur den zehnten Jahrgang umfasst, enthält 2 186 Fälle. Wie in Tabelle 5.61 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Schon in den Querschnitten der Jahre 2003 bis 2005 sind im Vergleich zur Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Paneldatensatz ist dieser Unterschied zu allen drei Zeitpunkten noch deutlicher: Es sind rund fünfeinhalb (t_2 , t_3 und t_4) Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht nicht vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.62). Referenz-

Tabelle 5.61: Zusammensetzung nach Geschlecht $P_{t_2,3,4}^*$

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_2,3,4}^*$	998	45,7	1188	54,3	2186
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_2,3,4}^*$	990	45,3	1196	54,7	2186
t_4	S_{t_4}	2003	51,2	1907	48,8	3910
	$Q_{t_4}^*$	1627	50,2	1616	49,8	3243
	$P_{t_2,3,4}^*$	996	45,6	1190	54,4	2186

Tabelle 5.62: Altersdurchschnitt $P_{t_2,3,4}^*$

		Alter	Std.-abw.	gesamt N
t_2	Q_{t_2}	14,0	0,8	3275
	$P_{t_2,3,4}^*$	13,9	0,7	2146
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_2,3,4}$	15,0	0,7	2163
t_4	$Q_{t_4}^*$	16,0	0,8	3154
	$P_{t_2,3,4}$	15,9	0,7	2156

daten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Aus Tabelle 5.63 wird ersichtlich, dass das Drei-Wellen-Panel im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen systematisch verzerrt ist. Sieht man von dem insgesamt sehr niedrigen Anteil der Sonderschülerinnen und -schüler ab, so ist im Vergleich zur Schulstatistik zum Zeitpunkt t_2 der Anteil der Gymnasiastinnen und Gymnasiasten vier Prozent, der Anteil der Gesamtschülerinnen und -schüler dreieinhalb Prozent und der Anteil der Realschülerinnen und -schüler ein halbes Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler knapp siebeneinhalb Prozent niedriger. Zum Zeitpunkt t_3 liegt der Anteil der Gymnasiastinnen und Gymnasiasten und der Anteil der Gesamtschülerinnen und -schüler rund eineinhalb Prozent und der Anteil der Realschülerinnen und -schüler knapp zweieinhalb Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler knapp viereinhalb Prozent niedriger als in der Schulstatistik. Zum Zeitpunkt t_4 liegt der Anteil der Gymnasiastinnen

und Gymnasiasten rund ein halbes Prozent, der Anteil der Gesamtschülerinnen und -schüler knapp zwei Prozent und der Anteil der Realschülerinnen und -schüler eineinhalb Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler rund drei Prozent niedriger als in der Schulstatistik.

Tabelle 5.63: Zusammensetzung $P_{t_2,3,4}^*$ nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform ^a (%)					gesamt <i>N</i>		
		HS	RS	GS	GY	SO	(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	
		$P_{t_2,3,4}^*$	24,9	20,9	35,4	18,7	—	998	
	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	
		$P_{t_2,3,4}^*$	25,3	21,7	35,3	17,8	—	990	
	t_4	S_{t_4}	28,0	21,0	31,6	18,1	1,3	2003	
		$Q_{t_4}^*$	26,2	22,1	32,6	18,7	0,2	1627	
		$P_{t_2,3,4}^*$	25,7	21,4	35,0	17,9	—	996	
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	
		$P_{t_2,3,4}^*$	16,2	25,1	32,6	26,1	—	1188	
	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	
		$P_{t_2,3,4}^*$	16,6	25,5	32,9	25,0	—	1196	
	t_4	S_{t_4}	19,6	23,0	32,6	24,6	0,2	1907	
		$Q_{t_4}^*$	17,5	22,7	33,4	26,4	0,0	1616	
		$P_{t_2,3,4}^*$	16,6	25,3	32,9	25,2	—	1190	
	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
			$P_{t_2,3,4}^*$	20,2	23,2	33,9	22,7	—	2186
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	
		$P_{t_2,3,4}^*$	20,5	23,8	34,0	21,7	—	2186	
t_4		S_{t_4}	23,9	22,0	32,0	21,3	0,8	3910	
		$Q_{t_4}^*$	21,9	22,4	33,0	22,5	0,1	3243	
		$P_{t_2,3,4}^*$	20,8	23,5	33,9	21,9	—	2186	

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

Im Paneldatensatz sind über alle Schulformen hinweg – ebenfalls mit Ausnahme der Sonderschulen – zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger

Tabelle 5.64: Zusammensetzung $P_{t_2,3,4}^*$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)					gesamt
			HS	RS	GS	GY	SO	(%)
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
		$P_{t_2,3,4}^*$	56,3	41,2	47,7	37,6	—	45,7
	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
		$P_{t_2,3,4}^*$	55,8	41,3	47,0	37,1	—	45,3
	t_4	S_{t_4}	60,1	48,9	50,4	43,6	87,1	51,2
		$Q_{t_4}^*$	60,1	49,5	49,6	41,7	100,0	50,2
		$P_{t_2,3,4}^*$	56,4	41,4	47,2	37,2	—	45,6
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
		$P_{t_2,3,4}^*$	43,7	58,8	52,3	62,4	—	54,3
	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
		$P_{t_2,3,4}^*$	44,2	58,7	53,0	62,9	—	54,7
	t_4	S_{t_4}	39,9	51,1	49,6	56,4	12,9	48,8
		$Q_{t_4}^*$	39,9	50,5	50,4	58,3	0,0	49,8
		$P_{t_2,3,4}^*$	43,6	58,6	52,8	62,8	—	54,4

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

Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.64). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Erweiterter Datensatz ($P_{t_2,3,4}$)

Das Drei-Wellen-Panel der Jahre 2003 bis 2005, welches zu t_4 alle befragten Jugendlichen (also Neunt- und Zehntklässler) umfasst, enthält 2 260 Fälle. Wie in Tabelle 5.65 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Im Vergleich zur Schulstatistik (t_2 und t_3) und im Vergleich zum Querschnitt (t_4) sind die Schülerinnen um je fünfeinhalb bzw. knapp fünf Prozent überrepräsentiert.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht kaum vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.66). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Tabelle 5.65: Zusammensetzung nach Geschlecht $P_{t_2,3,4}$

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_2,3,4}$	1032	45,7	1228	54,3	2260
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_2,3,4}$	1023	45,3	1237	54,7	2260
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_2,3,4}$	1031	45,6	1229	54,4	2260

Tabelle 5.66: Altersdurchschnitt $P_{t_2,3,4}$

		Alter	Std.-abw.	gesamt N
t_2	Q_{t_2}	14,0	0,8	3275
	$P_{t_2,3,4}$	14,0	0,7	2220
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_2,3,4}$	15,0	0,7	2233
t_4	Q_{t_4}	16,1	0,8	3313
	$P_{t_2,3,4}$	16,0	0,7	2230

Da der erweiterte Paneldatensatz im Gegensatz zur Schulstatistik zum Zeitpunkt t_4 nicht ausschließlich Jugendliche der neunten Jahrgangsstufe enthält, kann er in seiner Zusammensetzung nach Schulform zum Zeitpunkt t_4 nur mit dem Querschnittsdatsatz verglichen werden. Aus Tabelle 5.67 wird ersichtlich, dass das Drei-Wellen-Panel 2003 bis 2005 im Vergleich zur Schulstatistik (t_2 und t_3) und zum Querschnitt (t_4) im Hinblick auf die Schulform der Jugendlichen verzerrt ist. Für die Betrachtungen wird von den Sonderschülerinnen und -schülern, denjenigen, die eine »andere« und denjenigen, die keine Schule mehr besuchen, abgesehen, da sie insgesamt nur einen Anteil von unter zwei Prozent ausmachen. Zum Zeitpunkt t_2 ist im Vergleich zur Schulstatistik der Anteil der Realschülerinnen und -schüler ungefähr gleich groß, der Anteil der Gymnasiastinnen und Gymnasiasten knapp vier Prozent und der Anteil der Gesamtschülerinnen und -schüler knapp dreieinhalb Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler sechseinhalb Prozent niedriger. Zum Zeitpunkt t_3 ist im Vergleich zur Schulstatistik der Anteil der Gymnasiastinnen und Gymnasiasten eineinhalb Prozent, der Anteil der Gesamtschülerinnen und -schüler rund eineinhalb Prozent und der Anteil der Realschülerinnen und -schüler knapp zwei Prozent höher, dagegen der Anteil der

Hauptschülerinnen und -schüler rund dreieinhalb Prozent niedriger. Zum Zeitpunkt t_4 ist im Vergleich zum Querschnittsdatensatz der Anteil der Gesamtschülerinnen und -schüler knapp ein Prozent und der Anteil der Realschülerinnen und -schüler knapp eineinhalb Prozent höher, dagegen der Anteil der Gymnasiastinnen und Gymnasiasten knapp ein halbes Prozent und der Anteil der Hauptschülerinnen und -schüler ein Prozent niedriger.

Tabelle 5.67: Zusammensetzung $P_{t_2,3,4}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)						gesamt N		
			HS	RS	GS	GY	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	
		$P_{t_2,3,4}$	25,8	20,3	35,3	18,7	—	—	—	1032	
	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	
		$P_{t_2,3,4}$	26,1	21,0	35,1	17,8	—	—	—	1023	
	t_4	Q_{t_4}	25,8	21,2	32,2	18,4	0,7	1,6	0,0	1704	
		$P_{t_2,3,4}$	25,5	20,7	34,7	17,8	—	1,3	—	1028	
	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	
$P_{t_2,3,4}$			16,9	24,8	32,6	25,7	—	—	—	1228	
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	
		$P_{t_2,3,4}$	17,3	25,2	32,9	24,6	—	—	—	1237	
t_4		Q_{t_4}	17,5	22,9	32,2	26,1	0,0	1,0	0,3	1573	
		$P_{t_2,3,4}$	16,7	25,1	32,7	24,8	—	0,8	—	1228	
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
			$P_{t_2,3,4}$	21,0	22,7	33,8	22,5	—	—	—	2260
		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	
	$P_{t_2,3,4}$		21,3	23,3	33,9	21,5	—	—	—	2260	
	t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,4	1,4	0,1	3386	
		$P_{t_2,3,4}$	20,7	23,1	33,6	21,6	—	1,0	—	2256	

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

Im Paneldatensatz sind im Vergleich zur Schulstatistik (t_2 und t_3) und zum Querschnitt (t_4) über alle Schulformen hinweg mehr Schülerinnen und weniger Schüler vertreten, sieht man ebenfalls von den Sonderschulen, den anderen Schulformen und denjenigen, die auf keine Schule gehen, ab (siehe Tabelle 5.68). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Tabelle 5.68: Zusammensetzung $P_{t_2,3,4}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)						gesamt	
			HS	RS	GS	GY	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
		$P_{t_2,3,4}$	56,1	40,7	47,6	38,0	—	—	—	45,7
	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
		$P_{t_2,3,4}$	55,5	40,8	46,9	37,4	—	—	—	45,3
	t_4	Q_{t_4}	59,8	49,1	49,6	42,2	92,3	59,6	0,0	50,3
		$P_{t_2,3,4}$	56,1	40,9	47,1	37,6	—	56,5	—	45,6
	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
$P_{t_2,3,4}$			43,9	59,3	52,4	62,0	—	—	—	54,3
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
		$P_{t_2,3,4}$	44,5	59,2	53,1	62,6	—	—	—	54,7
t_4		Q_{t_4}	40,2	50,9	50,4	57,8	7,7	40,4	100,0	49,7
		$P_{t_2,3,4}$	43,9	59,1	52,9	62,4	—	43,5	—	54,4

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

5.3.3 Drei-Wellen-Panel 2002, 2003, 2005 ($P_{t_1,2,4}^*$ und $P_{t_1,2,4}$)

Nur zehnte Jahrgangsstufe ($P_{t_1,2,4}^*$)

Das Drei-Wellen-Panel der Jahre 2002, 2003 und 2005, welches zu t_4 nur den zehnten Jahrgang umfasst, enthält 1 868 Fälle. Wie in Tabelle 5.69 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Schon in den Querschnitten der Jahre 2002, 2003 und 2005 sind im Vergleich zur Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Paneldatensatz ist dieser Unterschied zu allen drei Zeitpunkten noch deutlicher: Es sind rund fünfeinhalb (t_1), rund fünf (t_3 und t_4) Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht nicht vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.70). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Aus Tabelle 5.71 wird ersichtlich, dass das Drei-Wellen-Panel im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen systematisch verzerrt ist. Sieht man von dem insgesamt sehr niedrigen Anteil der Sonderschülerinnen und -schüler ab,

Tabelle 5.69: Zusammensetzung nach Geschlecht $P_{t_1,2,4}^*$

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_1	S_{t_1}	2003	51,2	1907	48,8	3910
	Q_{t_1}	1728	50,7	1679	49,3	3407
	$P_{t_1,2,4}^*$	852	45,6	1016	54,4	1868
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_1,2,4}^*$	860	46,0	1008	54,0	1868
t_4	S_{t_4}	2003	51,2	1907	48,8	3910
	$Q_{t_4}^*$	1627	50,2	1616	49,8	3243
	$P_{t_1,2,4}^*$	859	46,0	1009	54,0	1868

Tabelle 5.70: Altersdurchschnitt $P_{t_1,2,4}^*$

		Alter	Std.-abw.	gesamt N
t_1	Q_{t_1}	13,0	0,8	3346
	$P_{t_1,2,4}^*$	12,9	0,7	1844
t_2	Q_{t_2}	14,0	0,8	3275
	$P_{t_1,2,4}^*$	13,9	0,7	1837
t_4	$Q_{t_4}^*$	16,0	0,8	3154
	$P_{t_1,2,4}^*$	15,9	0,7	1842

so ist im Vergleich zur Schulstatistik zum Zeitpunkt t_1 der Anteil der Gymnasiastinnen und Gymnasiasten rund fünfeinhalb Prozent und der Anteil der Realschülerinnen und -schüler rund vier Prozent höher, dagegen der Anteil der Gesamtschülerinnen und -schüler ein halbes Prozent und der Anteil der Hauptschülerinnen und -schüler knapp achteinhalb Prozent niedriger. Zum Zeitpunkt t_2 liegt der Anteil der Gymnasiastinnen und Gymnasiasten rund vier Prozent, der Anteil der Gesamtschülerinnen und -schüler rund zweieinhalb Prozent und der Anteil der Realschülerinnen und -schüler knapp zwei Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler knapp acht Prozent niedriger als in der Schulstatistik. Zum Zeitpunkt t_4 liegt der Anteil der Gymnasiastinnen und Gymnasiasten sowie der Anteil der Gesamtschülerinnen und -schüler rund ein Prozent und der Anteil der Realschülerinnen und -schüler rund zweieinhalb Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler rund vier Prozent niedriger als in der Schulstatistik.

Im Paneldatensatz sind über alle Schulformen hinweg – ebenfalls mit Ausnahme der Sonderschulen – zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger

Tabelle 5.71: Zusammensetzung $P_{1,2,4}^*$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)					gesamt <i>N</i>	
			HS	RS	GS	GY	SO	(100 %)	
männlich	t_3	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	
		$P_{1,2,4}^*$	23,6	22,7	34,3	19,5	—	852	
	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	
		$P_{1,2,4}^*$	23,7	22,7	34,5	19,1	—	860	
	t_4	S_{t_4}	28,0	21,0	31,6	18,1	1,3	2003	
		$Q_{t_4}^*$	26,2	22,1	32,6	18,7	0,2	1627	
		$P_{1,2,4}^*$	23,9	22,8	34,6	18,6	0,1	859	
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	
		$P_{1,2,4}^*$	15,5	26,1	31,5	27,0	—	1016	
	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	
		$P_{1,2,4}^*$	16,3	26,1	31,6	26,0	—	1008	
	t_4	S_{t_4}	19,6	23,0	32,6	24,6	0,2	1907	
		$Q_{t_4}^*$	17,5	22,7	33,4	26,4	0,0	1616	
		$P_{1,2,4}^*$	16,2	26,4	31,9	25,6	0,0	1009	
	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
			$P_{1,2,4}^*$	15,5	26,1	31,5	27,0	—	1868
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	
		$P_{1,2,4}^*$	19,7	24,5	33,0	22,8	—	1868	
t_4		S_{t_4}	23,9	22,0	32,0	21,3	0,8	3910	
		$Q_{t_4}^*$	21,9	22,4	33,0	22,5	0,1	3243	
		$P_{1,2,4}^*$	19,7	24,7	33,1	22,4	0,1	1868	

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

Tabelle 5.72: Zusammensetzung $P_{t_1,2,4}^*$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)					gesamt
			HS	RS	GS	GY	SO	(%)
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
		$P_{t_1,2,4}^*$	56,1	42,1	47,7	37,7	—	45,6
	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
		$P_{t_1,2,4}^*$	55,4	42,6	48,2	38,5	—	46,0
	t_4	S_{t_4}	60,1	48,9	50,4	43,6	87,1	51,2
		$Q_{t_4}^*$	60,1	49,5	49,6	41,7	100,0	50,2
		$P_{t_1,2,4}^*$	55,7	42,4	48,0	38,3	100,0	46,0
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
		$P_{t_1,2,4}^*$	43,9	57,9	52,3	62,3	—	54,4
	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
		$P_{t_1,2,4}^*$	44,6	57,4	51,8	61,5	—	54,0
	t_4	S_{t_4}	39,9	51,1	49,6	56,4	12,9	48,8
		$Q_{t_4}^*$	39,9	50,5	50,4	58,3	0,0	49,8
		$P_{t_1,2,4}^*$	44,3	57,6	52,0	61,7	0,0	54,0

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

Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.72). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Erweiterter Datensatz ($P_{t_1,2,4}$)

Das Drei-Wellen-Panel der Jahre 2002, 2003 und 2005, welches zu t_4 alle befragten Jugendlichen (also Neunt- und Zehntklässler) umfasst, enthält 1 930 Fälle. Wie in Tabelle 5.73 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Im Vergleich zur Schulstatistik (t_1 und t_2) und im Vergleich zum Querschnitt (t_4) sind die Schülerinnen um je fünfeinhalb bzw. viereinhalb Prozent überrepräsentiert.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht kaum vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.74). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Tabelle 5.73: Zusammensetzung nach Geschlecht $P_{t_1,2,4}$

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_1	S_{t_1}	2003	51,2	1907	48,8	3910
	Q_{t_1}	1728	50,7	1679	49,3	3407
	$P_{t_1,2,4}$	878	45,5	1052	54,5	1930
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_1,2,4}$	886	45,9	1044	54,1	1930
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_1,2,4}$	886	45,9	1044	54,1	1930

Tabelle 5.74: Altersdurchschnitt $P_{t_1,2,4}$

		Alter	Std.-abw.	gesamt N
t_1	Q_{t_1}	13,0	0,8	3346
	$P_{t_1,2,4}$	12,9	0,7	1904
t_2	Q_{t_2}	14,0	0,8	3275
	$P_{t_1,2,4}$	13,9	0,7	1898
t_4	Q_{t_4}	16,1	0,8	3313
	$P_{t_1,2,4}$	15,9	0,7	1904

Da der erweiterte Paneldatensatz im Gegensatz zur Schulstatistik zum Zeitpunkt t_4 nicht ausschließlich Jugendliche der neunten Jahrgangsstufe enthält, kann er in seiner Zusammensetzung nach Schulform zum Zeitpunkt t_4 nur mit dem Querschnittsdatsatz verglichen werden. Aus Tabelle 5.75 wird ersichtlich, dass das Drei-Wellen-Panel 2002, 2003 und 2005 im Vergleich zur Schulstatistik (t_1 und t_2) und zum Querschnitt (t_4) im Hinblick auf die Schulform der Jugendlichen verzerrt ist. Für die Betrachtungen wird von den Sonderschülerinnen und -schülern, denjenigen, die eine »andere« und denjenigen, die keine Schule mehr besuchen, abgesehen, da sie insgesamt nur einen Anteil von unter zwei Prozent ausmachen. Zum Zeitpunkt t_1 ist im Vergleich zur Schulstatistik der Anteil der Gymnasiastinnen und Gymnasiasten knapp zwei Prozent, der Anteil der Gesamtschülerinnen und -schüler knapp ein Prozent und der Anteil der Realschülerinnen und -schüler zwei Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler knapp vier Prozent niedriger. Zum Zeitpunkt t_2 ist im Vergleich zur Schulstatistik der Anteil der Gymnasiastinnen und Gymnasiasten dreieinhalb Prozent, der Anteil der Gesamtschülerinnen und -schüler rund zweieinhalb Prozent und der Anteil der Realschülerinnen und -schüler knapp eineinhalb Prozent höher, dagegen der Anteil der

Haupt Schülerinnen und -schüler knapp sieben Prozent niedriger. Zum Zeitpunkt t_4 ist im Vergleich zum Querschnittsdatensatz der Anteil der Gymnasiastinnen und Gymnasiasten ungefähr gleich groß, der Anteil der Gesamtschülerinnen und -schüler knapp ein halbes Prozent und der Anteil der Realschülerinnen und -schüler zweieinhalb Prozent höher, dagegen der Anteil der Haupt Schülerinnen und -schüler knapp zwei Prozent niedriger.

Tabelle 5.75: Zusammensetzung $P_{t_1,2,4}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)						gesamt N		
			HS	RS	GS	GY	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	
		$P_{t_1,2,4}$	24,5	22,1	34,2	19,2	—	—	—	878	
	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	
		$P_{t_1,2,4}$	24,7	22,0	34,4	18,8	—	—	—	886	
	t_4	Q_{t_4}	25,8	21,2	32,2	18,4	0,7	1,6	0,0	1704	
		$P_{t_1,2,4}$	23,8	22,2	34,6	18,4	0,1	0,9	—	882	
	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	
$P_{t_1,2,4}$			16,3	25,7	31,7	26,3	—	—	—	1052	
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	
		$P_{t_1,2,4}$	17,1	25,7	31,8	25,4	—	—	—	1044	
t_4		Q_{t_4}	17,5	22,9	32,2	26,1	0,0	1,0	0,3	1573	
		$P_{t_1,2,4}$	16,4	26,0	31,8	24,9	0,0	0,9	—	1043	
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
			$P_{t_1,2,4}$	20,0	24,0	32,8	23,1	—	—	—	1930
		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	
	$P_{t_1,2,4}$		20,6	24,0	33,0	22,4	—	—	—	1930	
	t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,4	1,4	0,1	3386	
		$P_{t_1,2,4}$	19,8	24,3	33,1	21,9	0,1	0,9	—	1925	

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

Im Paneldatensatz sind im Vergleich zur Schulstatistik (t_1 und t_2) und zum Querschnitt (t_4) über alle Schulformen hinweg mehr Schülerinnen und weniger Schüler vertreten, sieht man ebenfalls von den Sonderschulen, den anderen Schulformen und denjenigen, die auf keine Schule gehen, ab (siehe Tabelle 5.76). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Tabelle 5.76: Zusammensetzung $P_{t_1,2,4}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)							gesamt
			HS	RS	GS	GY	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
		$P_{t_1,2,4}$	55,7	41,8	47,3	37,9	—	—	—	45,5
	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
		$P_{t_1,2,4}$	55,0	42,1	47,9	38,7	—	—	—	45,9
	t_4	Q_{t_4}	59,8	49,1	49,6	42,2	92,3	59,6	0,0	50,3
		$P_{t_1,2,4}$	55,1	42,0	47,9	38,4	100,0	47,1	—	45,8
	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
$P_{t_1,2,4}$			44,3	58,2	52,7	62,1	—	—	—	54,5
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
		$P_{t_1,2,4}$	45,0	57,9	52,1	61,3	—	—	—	54,1
t_4		Q_{t_4}	40,2	50,9	50,4	57,8	7,7	40,4	100,0	49,7
		$P_{t_1,2,4}$	44,9	58,0	52,1	61,6	0,0	52,9	—	54,2

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

5.3.4 Drei-Wellen-Panel 2002, 2004, 2005 ($P_{t_1,3,4}^*$ und $P_{t_1,3,4}$)

Nur zehnte Jahrgangsstufe ($P_{t_1,3,4}^*$)

Das Drei-Wellen-Panel der Jahre 2002, 2004 und 2005, welches zu t_4 nur den zehnten Jahrgang umfasst, enthält 1 862 Fälle. Wie in Tabelle 5.77 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Schon in den Querschnitten der Jahre 2002, 2004 und 2005 sind im Vergleich zur Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Paneldatensatz ist dieser Unterschied zu allen drei Zeitpunkten noch deutlicher: Es sind rund sechseinhalb (t_1), knapp sechs (t_3) und sechs (t_4) Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht nicht vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.78). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Aus Tabelle 5.79 wird ersichtlich, dass das Drei-Wellen-Panel im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen systematisch verzerrt ist. Sieht man von dem insgesamt sehr niedrigen Anteil der Sonderschülerinnen und -schüler

Tabelle 5.77: Zusammensetzung nach Geschlecht $P_{t_1,3,4}^*$

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_1	S_{t_1}	2003	51,2	1907	48,8	3910
	Q_{t_1}	1728	50,7	1679	49,3	3407
	$P_{t_1,3,4}^*$	833	44,7	1029	55,3	1862
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_1,3,4}^*$	839	45,1	1023	54,9	1862
t_4	S_{t_4}	2003	51,2	1907	48,8	3910
	$Q_{t_4}^*$	1627	50,2	1616	49,8	3243
	$P_{t_1,3,4}^*$	841	45,2	1021	54,8	1862

Tabelle 5.78: Altersdurchschnitt $P_{t_1,3,4}^*$

		Alter	Std.-abw.	gesamt N
t_1	Q_{t_1}	13,0	0,8	3346
	$P_{t_1,3,4}^*$	12,9	0,7	1840
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_1,3,4}^*$	14,9	0,7	1853
t_4	$Q_{t_4}^*$	16,0	0,8	3154
	$P_{t_1,3,4}^*$	15,9	0,7	1838

ab, so ist im Vergleich zur Schulstatistik zum Zeitpunkt t_1 der Anteil der Gymnasiastinnen und Gymnasiasten rund fünf Prozent, der Anteil der Gesamtschülerinnen und -schüler ein halbes Prozent und der Anteil der Realschülerinnen und -schüler knapp zwei Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler rund sechseinhalb Prozent niedriger. Zum Zeitpunkt t_3 liegt der Anteil der Gymnasiastinnen und Gymnasiasten knapp fünf Prozent, der Anteil der Gesamtschülerinnen und -schüler knapp ein halbes Prozent und der Anteil der Realschülerinnen und -schüler drei Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler sieben Prozent niedriger als in der Schulstatistik. Zum Zeitpunkt t_4 liegt der Anteil der Gymnasiastinnen und Gymnasiasten knapp dreieinhalb, der Anteil der Gesamtschülerinnen und -schüler rund ein halbes Prozent und der Anteil der Realschülerinnen und -schüler rund zweieinhalb Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler knapp sechs Prozent niedriger als in der Schulstatistik.

Im Paneldatensatz sind über alle Schulformen hinweg – ebenfalls mit Ausnahme der Sonderschulen – zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger

Tabelle 5.79: Zusammensetzung $P_{1,3,4}^*$ nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform ^a (%)					gesamt N	
		HS	RS	GS	GY	SO	(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
		$P_{1,3,4}^*$	20,9	21,6	34,3	23,2	—	833
	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
		$P_{1,3,4}^*$	21,6	22,8	33,6	22,1	—	839
	t_4	S_{t_4}	28,0	21,0	31,6	18,1	1,3	2003
		$Q_{t_4}^*$	26,2	22,1	32,6	18,7	0,2	1627
		$P_{1,3,4}^*$	21,6	23,1	33,8	21,5	0,0	841
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
		$P_{1,3,4}^*$	14,4	25,7	31,0	29,0	—	1029
	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
		$P_{1,3,4}^*$	15,0	25,9	32,0	27,2	—	1023
	t_4	S_{t_4}	19,6	23,0	32,6	24,6	0,2	1907
		$Q_{t_4}^*$	17,5	22,7	33,4	26,4	0,0	1616
		$P_{1,3,4}^*$	15,2	26,0	31,6	27,2	—	1021
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
		$P_{1,3,4}^*$	17,3	23,8	32,5	26,4	—	1862
	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
		$P_{1,3,4}^*$	17,9	24,5	32,7	24,9	—	1862
	t_4	S_{t_4}	23,9	22,0	32,0	21,3	0,8	3910
		$Q_{t_4}^*$	21,9	22,4	33,0	22,5	0,1	3243
		$P_{1,3,4}^*$	18,1	24,7	32,6	24,7	—	1862

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

Tabelle 5.80: Zusammensetzung $P_{t_1,3,4}^*$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)					gesamt
			HS	RS	GS	GY	SO	(%)
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
		$P_{t_1,3,4}^*$	54,0	40,5	47,3	39,3	—	44,7
	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
		$P_{t_1,3,4}^*$	54,2	41,9	46,3	40,0	—	45,1
	t_4	S_{t_4}	60,1	48,9	50,4	43,6	87,1	51,2
		$Q_{t_4}^*$	60,1	49,5	49,6	41,7	100,0	50,2
		$P_{t_1,3,4}^*$	54,0	42,3	46,8	39,4	—	45,2
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
		$P_{t_1,3,4}^*$	46,0	59,5	52,7	60,7	—	55,3
	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
		$P_{t_1,3,4}^*$	45,8	58,1	53,7	60,0	—	54,9
	t_4	S_{t_4}	39,9	51,1	49,6	56,4	12,9	48,8
		$Q_{t_4}^*$	39,9	50,5	50,4	58,3	0,0	49,8
		$P_{t_1,3,4}^*$	46,0	57,7	53,2	60,6	—	54,8

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

Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.72). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Erweiterter Datensatz ($P_{t_1,3,4}$)

Das Drei-Wellen-Panel der Jahre 2002, 2004 und 2005, welches zu t_4 alle befragten Jugendlichen (also Neunt- und Zehntklässler) umfasst, enthält 1 922 Fälle. Wie in Tabelle 5.81 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Im Vergleich zur Schulstatistik (t_1 und t_2) und im Vergleich zum Querschnitt (t_4) sind die Schülerinnen um sechseinhalb, rund sechs und knapp fünfeneinhalb Prozent überrepräsentiert.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht kaum vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.82). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Tabelle 5.81: Zusammensetzung nach Geschlecht $P_{t_1,3,4}$

		männlich		weiblich		gesamt N
		N	%	N	%	(100 %)
t_1	S_{t_1}	2003	51,2	1907	48,8	3910
	Q_{t_1}	1728	50,7	1679	49,3	3407
	$P_{t_1,3,4}$	857	44,6	1065	55,4	1922
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_1,3,4}$	862	44,8	1060	55,2	1922
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_1,3,4}$	866	45,1	1056	54,9	1922

Tabelle 5.82: Altersdurchschnitt $P_{t_1,3,4}$

		Alter	Std.-abw.	gesamt N
t_1	Q_{t_1}	13,0	0,8	3346
	$P_{t_1,3,4}$	12,9	0,7	1898
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_1,3,4}$	14,9	0,7	1910
t_4	Q_{t_4}	16,1	0,8	3313
	$P_{t_1,3,4}$	15,9	0,7	1897

Da der erweiterte Paneldatensatz im Gegensatz zur Schulstatistik zum Zeitpunkt t_4 nicht ausschließlich Jugendliche der neunten Jahrgangsstufe enthält, kann er in seiner Zusammensetzung nach Schulform zum Zeitpunkt t_4 nur mit dem Querschnittsdatsatz verglichen werden. Aus Tabelle 5.83 wird ersichtlich, dass das Drei-Wellen-Panel 2002, 2004 und 2005 im Vergleich zur Schulstatistik (t_1 und t_3) und zum Querschnitt (t_4) im Hinblick auf die Schulform der Jugendlichen verzerrt ist. Für die Betrachtungen wird von den Sonderschülerinnen und -schülern, denjenigen, die eine »andere« und denjenigen, die keine Schule mehr besuchen, abgesehen, da sie insgesamt nur einen Anteil von unter zwei Prozent ausmachen. Zum Zeitpunkt t_1 ist im Vergleich zur Schulstatistik der Anteil der Gymnasiastinnen und Gymnasiasten rund viereinhalb Prozent, der Anteil der Gesamtschülerinnen und -schüler ein halbes Prozent und der Anteil der Realschülerinnen und -schüler knapp eineinhalb Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler rund fünfeinhalb Prozent niedriger. Zum Zeitpunkt t_3 ist im Vergleich zur Schulstatistik der Anteil der Gesamtschülerinnen und -schüler ungefähr gleichgroß, der Anteil der Gymnasiastinnen und Gymnasiasten knapp viereinhalb Prozent und der Anteil der Realschülerinnen und -schüler knapp zweieinhalb Prozent höher, dagegen der Anteil

der Hauptschülerinnen und -schüler sechs Prozent niedriger. Zum Zeitpunkt t_4 ist im Vergleich zum Querschnittsdatensatz der Anteil der Gesamtschülerinnen und -schüler ungefähr gleich groß, der Anteil der Gymnasiastinnen und Gymnasiasten rund zwei Prozent und der Anteil der Realschülerinnen und -schüler knapp zweieinhalb Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler rund dreieinhalb Prozent niedriger.

Tabelle 5.83: Zusammensetzung $P_{t_1,3,4}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)							gesamt N
			HS	RS	GS	GY	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
		$P_{t_1,3,4}$	21,9	21,1	34,1	22,9	—	—	—	857
	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
		$P_{t_1,3,4}$	22,5	22,2	33,4	21,8	0,1	—	—	862
	t_4	Q_{t_4}	25,8	21,2	32,2	18,4	0,7	1,6	0,0	1704
		$P_{t_1,3,4}$	21,7	22,5	33,6	21,2	—	1,0	—	863
	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
$P_{t_1,3,4}$			15,2	25,3	31,3	28,3	—	—	—	1065
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
		$P_{t_1,3,4}$	16,0	25,5	32,0	26,5	0,0	—	—	1060
t_4		Q_{t_4}	17,5	22,9	32,2	26,1	0,0	1,0	0,3	1573
		$P_{t_1,3,4}$	15,3	25,6	31,5	26,5	—	1,1	—	1055
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
	$P_{t_1,3,4}$		18,2	23,4	32,5	25,9	—	—	—	1922
	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
		$P_{t_1,3,4}$	18,9	24,0	32,6	24,4	0,1	—	—	1922
	t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,4	1,4	0,1	3386
		$P_{t_1,3,4}$	18,1	24,2	32,4	24,1	—	1,1	—	1918

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

Im Paneldatensatz sind im Vergleich zur Schulstatistik (t_1 und t_3) und zum Querschnitt (t_4) über alle Schulformen hinweg mehr Schülerinnen und weniger Schüler vertreten, sieht man ebenfalls von den Sonderschulen, den anderen Schulformen und denjeni-

Tabelle 5.84: Zusammensetzung $P_{t_1,3,4}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)							gesamt
			HS	RS	GS	GY	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
		$P_{t_1,3,4}$	53,7	40,2	46,7	39,4	—	—	—	44,6
	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
		$P_{t_1,3,4}$	53,3	41,4	45,9	40,1	100,0	—	—	44,8
	t_4	Q_{t_4}	59,8	49,1	49,6	42,2	92,3	59,6	0,0	50,3
		$P_{t_1,3,4}$	53,7	41,8	46,6	39,5	—	42,9	—	45,0
	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
$P_{t_1,3,4}$			46,3	59,8	53,3	60,6	—	—	—	55,4
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
		$P_{t_1,3,4}$	46,7	58,6	54,1	59,9	0,0	—	—	55,2
t_4		Q_{t_4}	40,2	50,9	50,4	57,8	7,7	40,4	100,0	49,7
		$P_{t_1,3,4}$	46,3	58,2	53,4	60,5	—	57,1	—	55,0

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

gen, die auf keine Schule gehen, ab (siehe Tabelle 5.84). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

5.4 Das Vier-Wellen-Panel 2002 bis 2005 ($P_{t_1,2,3,4}^*$ und $P_{t_1,2,3,4}$)

Der Datensatz des Vier-Wellen-Panels enthält alle Fälle, deren zugehörige (und nach Plausibilitätskontrollen gültige) Fragebögen drei (ebenfalls gültigen) Bögen drei weiterer Zeitpunkte zugeordnet werden konnten. Das Vier-Wellen-Panel der Jahre 2002 bis 2005 stellt das vorläufige Resultat der Längsschnitt-Erhebungen in Duisburg dar.

Nur zehnte Jahrgangsstufe ($P_{t_1,2,3,4}^*$)

Das Vier-Wellen-Panel der Jahre 2002 bis 2005, welches zu t_4 nur den zehnten Jahrgang umfasst, enthält 1 715 Fälle. Wie in Tabelle 5.85 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Schon in den Querschnitten der Jahre 2002 bis 2005 sind im Vergleich zur Schulstatistik die Schülerinnen etwas überrepräsentiert. Im Paneldatensatz ist dieser Unterschied zu allen vier Zeitpunkten noch deutlicher:

Tabelle 5.85: Zusammensetzung nach Geschlecht $P_{t_{1,2,3,4}}^*$

		männlich		weiblich		gesamtN
		N	%	N	%	(100 %)
t_1	S_{t_1}	2003	51,2	1907	48,8	3910
	Q_{t_1}	1728	50,7	1679	49,3	3407
	$P_{t_{1,2,3,4}}^*$	751	43,8	964	56,2	1715
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_{1,2,3,4}}^*$	762	44,4	953	55,6	1715
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_{1,2,3,4}}^*$	756	44,1	959	55,9	1715
t_4	S_{t_4}	1760	49,7	1783	50,3	3543
	$Q_{t_4}^*$	1627	50,2	1616	49,8	3243
	$P_{t_{1,2,3,4}}^*$	760	44,3	955	55,7	1715

Es sind knapp siebeneinhalb (t_1), knapp sieben (t_2 und t_3) und knapp fünfeinhalb (t_4) Prozent mehr Schülerinnen vorhanden als in der Schulstatistik ausgewiesen.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht nicht vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.86). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Aus Tabelle 5.87 wird ersichtlich, dass das Vier-Wellen-Panel im Vergleich zur Schulstatistik im Hinblick auf die Schulform der Jugendlichen systematisch verzerrt ist. Sieht

Tabelle 5.86: Altersdurchschnitt $P_{t_{1,2,3,4}}^*$

		Alter	Std.-abw.	gesamt N
t_1	Q_{t_1}	13,0	0,8	3346
	$P_{t_{1,2,3,4}}^*$	12,9	0,7	1694
t_2	Q_{t_2}	14,0	0,8	3275
	$P_{t_{1,2,3,4}}^*$	13,9	0,7	1689
t_3	Q_{t_3}	15,1	0,8	3270
	$P_{t_{1,2,3,4}}^*$	14,9	0,7	1707
t_4	$Q_{t_4}^*$	16,0	0,8	3154
	$P_{t_{1,2,3,4}}^*$	15,9	0,7	1693

man von dem insgesamt sehr niedrigen Anteil der Sonderschülerinnen und -schüler ab, so ist im Vergleich zur Schulstatistik zum Zeitpunkt t_1 der Anteil der Gymnasiastinnen und Gymnasiasten rund zweieinhalb Prozent, der Anteil der Gesamtschülerinnen und -schüler rund eineinhalb Prozent und der Anteil der Realschülerinnen und -schüler rund zweieinhalb Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler rund sechs Prozent niedriger. Zum Zeitpunkt t_2 liegt der Anteil der Gymnasiastinnen und Gymnasiasten knapp viereinhalb Prozent, der Anteil der Gesamtschülerinnen und -schüler dreieinhalb Prozent und der Anteil der Realschülerinnen und -schüler zwei Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler rund neun Prozent niedriger als in der Schulstatistik. Zum Zeitpunkt t_3 liegt der Anteil der Gymnasiastinnen und Gymnasiasten knapp zweieinhalb Prozent, der Anteil der Gesamtschülerinnen und -schüler eineinhalb Prozent und der Anteil der Realschülerinnen und -schüler knapp vier Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler sechseinhalb Prozent niedriger als in der Schulstatistik. Zum Zeitpunkt t_4 ist der Anteil der Gesamtschülerinnen und -schüler annähernd gleichgroß, der Anteil der Gymnasiastinnen und Gymnasiasten ein halbes Prozent und der Anteil der Realschülerinnen und -schüler knapp dreieinhalb Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler knapp vier Prozent niedriger als in der Schulstatistik.

Im Paneldatensatz sind über alle Schulformen hinweg – ebenfalls mit Ausnahme der Sonderschulen – zu beiden Zeitpunkten prozentual mehr Schülerinnen und weniger Schüler als in der Schulstatistik vertreten (siehe Tabelle 5.88). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Tabelle 5.87: Zusammensetzung $P_{t_{1,2,3,4}}^*$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)					gesamt N
			HS	RS	GS	GY	SO	(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
		$P_{t_{1,2,3,4}}^*$	21,6	22,1	36,5	19,8	—	751
	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
		$P_{t_{1,2,3,4}}^*$	21,9	22,2	36,6	19,3	—	762
	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
		$P_{t_{1,2,3,4}}^*$	22,1	23,3	35,7	18,9	—	756
	t_4	S_{t_4}	26,3	21,8	32,8	18,8	0,3	1760
		$Q_{t_4}^*$	26,2	22,1	32,6	18,7	0,2	1627
		$P_{t_{1,2,3,4}}^*$	22,4	23,2	35,5	18,9	—	760
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
		$P_{t_{1,2,3,4}}^*$	14,8	26,8	31,3	27,1	—	964
	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
		$P_{t_{1,2,3,4}}^*$	15,4	26,8	31,7	26,1	—	953
	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
		$P_{t_{1,2,3,4}}^*$	15,4	27,1	32,2	25,2	—	959
	t_4	S_{t_4}	26,3	21,8	32,8	18,8	0,3	1783
		$Q_{t_4}^*$	17,5	22,7	33,4	26,4	0,0	1616
		$P_{t_{1,2,3,4}}^*$	15,8	27,2	31,5	25,4	—	955
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
		$P_{t_{1,2,3,4}}^*$	17,8	24,7	33,6	23,9	—	1715
	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
		$P_{t_{1,2,3,4}}^*$	18,3	24,7	33,9	23,1	—	1715
	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
		$P_{t_{1,2,3,4}}^*$	18,4	25,4	33,8	22,4	—	1715
	t_4	S_{t_4}	22,5	22,0	33,2	22,1	0,2	3543
		$Q_{t_4}^*$	21,9	22,4	33,0	22,5	0,1	3243
		$P_{t_{1,2,3,4}}^*$	18,7	25,4	33,3	22,6	—	1715

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

Tabelle 5.88: Zusammensetzung $P_{t_1,2,3,4}^*$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)					gesamt
			HS	RS	GS	GY	SO	(%)
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
		$P_{t_1,2,3,4}^*$	53,1	39,2	47,6	36,3	—	43,8
	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
		$P_{t_1,2,3,4}^*$	53,2	39,9	48,0	37,1	—	44,4
	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
		$P_{t_1,2,3,4}^*$	53,0	40,4	46,6	37,1	—	44,1
	t_4	S_{t_4}	57,7	49,3	51,5	41,6	76,5	50,9
		$Q_{t_4}^*$	60,1	49,5	49,6	41,7	100,0	50,2
		$P_{t_1,2,3,4}^*$	53,0	40,4	47,3	37,2	—	44,3
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
		$P_{t_1,2,3,4}^*$	46,9	60,8	52,4	63,7	—	56,2
	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
		$P_{t_1,2,3,4}^*$	46,8	60,1	52,0	62,9	—	55,6
	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
		$P_{t_1,2,3,4}^*$	47,0	59,6	53,4	62,9	—	55,9
	t_4	S_{t_4}	42,3	50,7	48,5	58,4	23,5	49,1
		$Q_{t_4}^*$	39,9	50,5	50,4	58,3	0,0	49,8
		$P_{t_1,2,3,4}^*$	47,0	59,6	52,7	62,8	—	55,7

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

Erweiterter Datensatz ($P_{t_{1,2,3,4}}$)

Das Vier-Wellen-Panel der Jahre 2002 bis 2005, welches zu t_4 alle befragten Jugendlichen (also Neunt- und Zehntklässler) umfasst, enthält 1 769 Fälle. Wie in Tabelle 5.89 ersichtlich, liegt in Bezug auf das Geschlecht der Probanden eine Verschiebung vor. Im Vergleich zur Schulstatistik (t_1 , t_2 und t_3) und im Vergleich zum Querschnitt (t_4) sind die Schülerinnen um rund siebeneinhalb, knapp sieben, sieben und rund sechs Prozent überrepräsentiert.

Das Durchschnittsalter der Jugendlichen, deren Fälle im Panel enthalten sind, weicht kaum vom Alter der Fälle in den Querschnittsdaten ab (siehe Tabelle 5.90). Referenzdaten für die Grundgesamtheit liegen nicht vor, da die Schulstatistik nicht nach Alter differenziert und die Einwohnermeldestatistik kaum mit den vorhandenen Daten zur Deckung gebracht werden kann.

Da der erweiterte Paneldatensatz im Gegensatz zur Schulstatistik zum Zeitpunkt t_4 nicht ausschließlich Jugendliche der neunten Jahrgangsstufe enthält, kann er in seiner Zusammensetzung nach Schulform zum Zeitpunkt t_4 nur mit dem Querschnittsdatensatz verglichen werden. Aus Tabelle 5.91 wird ersichtlich, dass das Vier-Wellen-Panel 2002 bis 2005 im Vergleich zur Schulstatistik (t_1 , t_2 und t_3) und zum Querschnitt (t_4) im Hinblick auf die Schulform der Jugendlichen verzerrt ist. Für die Betrachtungen wird von den Sonderschülerinnen und -schülern, denjenigen, die eine »andere« und denjenigen, die keine Schule mehr besuchen, abgesehen, da sie insgesamt nur einen Anteil von unter zwei Prozent ausmachen. Zum Zeitpunkt t_1 ist im Vergleich zur Schulstatistik der Anteil der Gymnasiastinnen und Gymnasiasten rund zwei Prozent, der Anteil der Gesamtschülerinnen und -schüler rund eineinhalb Prozent und der Anteil der

Tabelle 5.89: Zusammensetzung nach Geschlecht $P_{t_{1,2,3,4}}$

		männlich		weiblich		gesamtN (100 %)
		N	%	N	%	
t_1	S_{t_1}	2003	51,2	1907	48,8	3910
	Q_{t_1}	1728	50,7	1679	49,3	3407
	$P_{t_{1,2,3,4}}$	772	43,6	997	56,4	1769
t_2	S_{t_2}	1927	51,2	1836	48,8	3763
	Q_{t_2}	1703	50,2	1687	49,8	3390
	$P_{t_{1,2,3,4}}$	783	44,3	986	55,7	1769
t_3	S_{t_3}	1999	50,9	1926	49,1	3925
	Q_{t_3}	1635	49,0	1704	51,0	3339
	$P_{t_{1,2,3,4}}$	776	43,9	993	56,1	1769
t_4	Q_{t_4}	1717	50,4	1688	49,6	3405
	$P_{t_{1,2,3,4}}$	782	44,2	987	55,8	1769

Tabelle 5.90: Altersdurchschnitt $P_{t_1,2,3,4}$

		Alter	Std.-abw.	gesamt N
t_1	Q_{t_1}	13,0	0,8	3 346
	$P_{t_1,2,3,4}$	12,9	0,7	1 746
t_2	Q_{t_2}	14,0	0,8	3 275
	$P_{t_1,2,3,4}$	13,9	0,7	1 743
t_3	Q_{t_3}	15,1	0,8	3 270
	$P_{t_1,2,3,4}$	14,9	0,7	1 758
t_4	Q_{t_4}	16,1	0,8	3 313
	$P_{t_1,2,3,4}$	15,9	0,7	1 747

Realschülerinnen und -schüler knapp zweieinhalb Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler knapp fünfeinhalb Prozent niedriger. Zum Zeitpunkt t_2 ist im Vergleich zur Schulstatistik der Anteil der Gymnasiastinnen und Gymnasiasten vier Prozent der Anteil der Gesamtschülerinnen und -schüler knapp dreieinhalb und der Anteil der Realschülerinnen und -schüler rund eineinhalb Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler knapp achteinhalb Prozent niedriger. Zum Zeitpunkt t_3 ist im Vergleich zur Schulstatistik der Anteil der Gymnasiastinnen und Gymnasiasten rund zwei Prozent, der Anteil der Gesamtschülerinnen und -schüler knapp eineinhalb Prozent und der Anteil der Realschülerinnen und -schüler knapp dreieinhalb Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler rund fünfeinhalb Prozent niedriger. Zum Zeitpunkt t_4 ist im Vergleich zum Querschnittsdatensatz der Anteil der Gymnasiastinnen und Gymnasiasten ungefähr gleichgroß, der Anteil der Gesamtschülerinnen und -schüler knapp ein halbes Prozent und der Anteil der Realschülerinnen und -schüler rund drei Prozent höher, dagegen der Anteil der Hauptschülerinnen und -schüler knapp drei Prozent niedriger.

Im Paneldatensatz sind im Vergleich zur Schulstatistik (t_1 , t_3 und t_3) und zum Querschnitt (t_4) über alle Schulformen hinweg mehr Schülerinnen und weniger Schüler vertreten, sieht man ebenfalls von den Sonderschulen, den anderen Schulformen und denjenigen, die auf keine Schule gehen, ab (siehe Tabelle 5.92). Zu möglichen Ursachen der beschriebenen systematischen Verzerrungen siehe Abschnitt 6.

Tabelle 5.91: Zusammensetzung $P_{1,2,3,4}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

			Schulform ^a (%)							gesamt N
			HS	RS	GS	GY	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
		$P_{1,2,3,4}$	22,5	21,6	36,1	19,7	—	—	—	772
	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
		$P_{1,2,3,4}$	23,0	21,6	36,3	19,2	—	—	—	783
	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
		$P_{1,2,3,4}$	23,1	22,7	35,4	18,8	—	—	—	776
	t_4	Q_{t_4}	25,8	21,2	32,2	18,4	0,7	1,6	0,0	1704
		$P_{1,2,3,4}$	22,3	22,6	35,3	18,7	—	1,0	—	779
	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
$P_{1,2,3,4}$			15,5	26,4	31,6	26,5	—	—	—	997
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
		$P_{1,2,3,4}$	16,2	26,4	31,8	25,6	—	—	—	986
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
		$P_{1,2,3,4}$	16,3	26,7	32,3	24,7	—	—	—	993
t_4		Q_{t_4}	17,5	22,9	32,2	26,1	0,0	1,0	0,3	1573
		$P_{1,2,3,4}$	15,9	26,9	31,4	24,8	—	0,9	—	986
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
	$P_{1,2,3,4}$		18,6	24,3	33,6	23,5	—	—	—	1769
	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
		$P_{1,2,3,4}$	19,2	24,3	33,8	22,7	—	—	—	1769
	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
		$P_{1,2,3,4}$	19,3	24,9	33,7	22,1	—	—	—	1769
	t_4	Q_{t_4}	21,7	21,8	32,7	22,0	0,4	1,4	0,1	3386
		$P_{1,2,3,4}$	18,8	25,0	33,1	22,2	—	1,0	—	1765

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

Tabelle 5.92: Zusammensetzung $P_{t_1,2,3,4}$ nach Geschlecht und Schulform (Spaltenprozentuierung)

			Schulform ^a (%)							gesamt
			HS	RS	GS	GY	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
		$P_{t_1,2,3,4}$	52,9	38,8	47,0	36,5	—	—	—	43,6
	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
		$P_{t_1,2,3,4}$	52,9	39,4	47,5	37,3	—	—	—	44,3
	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
		$P_{t_1,2,3,4}$	52,5	39,9	46,1	37,3	—	—	—	43,9
	t_4	Q_{t_4}	59,8	49,1	49,6	42,2	92,3	59,6	0,0	50,3
		$P_{t_1,2,3,4}$	52,6	39,9	47,0	37,3	—	47,1	—	44,1
	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
$P_{t_1,2,3,4}$			47,1	61,2	53,0	63,5	—	—	—	56,4
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
		$P_{t_1,2,3,4}$	47,1	60,6	52,5	62,7	—	—	—	55,7
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
		$P_{t_1,2,3,4}$	47,5	60,1	53,9	62,7	—	—	—	56,1
t_4		Q_{t_4}	40,2	50,9	50,4	57,8	7,7	40,4	100,0	49,7
		$P_{t_1,2,3,4}$	47,4	60,1	53,0	62,7	—	52,9	—	55,9

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

6 Zusammenfassung und Erklärung

Bei der Analyse der oben beschriebenen Zusammensetzungen der Paneldaten im Vergleich mit den Daten der Querschnitte und der Schulstatistik ist die Unterscheidung nach Individual- und Schulebene sinnvoll. Analog zu den Methodenberichten der Querschnitterhebungen wurde auf eine Darstellung der Zusammensetzung der *gesamten* Duisburger Schulbevölkerung zugunsten einer detaillierten Darstellung der Zusammensetzung der Schulpopulationen teilnehmender Schulen verzichtet. Die oben beschriebenen Verzerrungen bezogen sich immer auf Vergleiche zwischen Panel- und Querschnittsdatensätzen und der Schulstatistik bezogen auf die teilnehmenden Schulen und sind deshalb durch Ausfälle auf Individualebene begründet. An dieser Stelle soll zusätzlich zur Zusammenfassung der Ausfälle auf der Individualebene kurz auf die Ausfälle auf der Schulebene eingegangen werden.

6.1 Individualebene

Zusammenfassend lässt sich sagen, dass die Paneldaten in Bezug auf die Zusammensetzung nach Geschlecht und Schulform (teilweise leicht, teilweise erheblich) verzerrt sind. Es finden sich im Vergleich zur Schulstatistik über alle Paneldatensätze hinweg mehr Mädchen als Jungen in den Daten. Der Anteil der Hauptschülerinnen und -schüler ist zumeist geringer als in der Schulstatistik und in den Querschnitten, der Anteil der Gymnasiastinnen und Gymnasiasten und der Realschülerinnen und -schüler dafür meist größer. Der Anteil der Gesamtschülerinnen und -schüler liegt meist auf ähnlichem Niveau.

Diese Abweichungen sind in geringerer Stärke schon in den Querschnittsdaten zu erkennen. Generell setzen sich also 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 Ursachen für die Panelverzerrungen sind. Hierzu zählen die gewollten und ungewollten Ausfälle in den Querschnitten aufgrund von Krankheit, Schwänzen, Verweigerungen etc. Man muss also annehmen, dass bei Jungen sowie bei Hauptschülerinnen und -schülern diese Ausfallgründe öfter auftreten als bei Mädchen, 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. Eine weitere Ursache für die Verzerrungen der Paneldatensätze liegt in der Art der Zuordnung

Tabelle 6.1: Verzerrung der Paneldaten ($P_{t_{3,4}}$ zu t_3) nach Geschlecht und Anzahl der Fehler in der Codezuordnung

	Anzahl der Fehler	männlich		weiblich		gesamt N (100 %)
		N	%	N	%	
$P_{t_{3,4}}$	0	694	41,6	975	58,4	1 669
	1	379	53,2	333	46,8	712
	2	180	56,8	137	43,2	317
	3	36	69,2	16	30,8	52
gesamt		1 289	46,9	1 461	53,1	2 750

Tabelle 6.2: Verzerrung der Paneldaten ($P_{t_{3,4}}$ zu t_3) nach Schulform und Anzahl der Fehler in der Codezuordnung

	Anzahl der Fehler	Schulform ^a (%)					gesamt (%)
		HS	RS	GS	GY	SO	
$P_{t_{3,4}}$	0	17,7	24,1	31,6	26,5	0,1	1 669
	1	26,8	20,8	32,0	19,8	0,6	712
	2	33,4	18,3	36,0	12,3	—	317
	3	26,5	17,3	32,7	17,3	—	52
gesamt		22,1	22,4	32,2	23,0	0,2	2 750

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

der Fragebögen. Das oben erläuterte Verfahren über wiederholte Codes weist den Nachteil auf, dass das gleiche Ausfüllen der Codefragen unterschiedlichen Personengruppen unterschiedlich stark Probleme bereitete.

Am Beispiel des Paneldatensatzes $P_{t_{3,4}}$ (Duisburg 2004/2005) soll erläutert werden, welche Auswirkungen das angewendete Zuordnungsverfahren in Bezug auf die Repräsentativität der Daten hat. In den Tabellen 6.1 und 6.2 finden sich die Paneldaten aufgeschlüsselt nach den zugelassenen Fehlern in der Codezuordnung.

Es wird deutlich, dass zwar der größte Teil ($N = 1 669$) der Schülerinnen und Schüler bei der Beantwortung des Codeblattes in 2004 und 2005 keinen Fehler machte, der Anteil derjenigen mit fehlerhaften Zuordnungen insgesamt jedoch ebenfalls bedeutsam ist. Unter diesen fehlerhaften Zuordnungen machen diejenigen mit einem Fehler ($N = 712$) die größte Gruppe aus, gefolgt von denjenigen mit zwei Fehlern ($N = 317$). Deutlich kleiner ist die Gruppe der Jugendlichen, die drei Fehler machten ($N = 52$). Es ist ersichtlich, dass die Anzahl der Fehler bei der Beantwortung des Codeblattes mit Geschlecht und Schulbildung der Befragten korrelieren. Während die Verzerrungen auf Ebene des aggregierten Paneldatensatzes, der Zuordnungen bis zu drei Fehler beinhaltet, nur leicht sind (siehe Abschnitt 5.2.3), lassen sich bei der Aufschlüsselung nach Fehlern

sehr viel deutlichere Verschiebungen erkennen. Augenscheinlich war es den weiblichen Befragungsteilnehmern eher möglich, in zwei Jahren fehlerfrei zu antworten; der Anteil der Mädchen ist bei den Zuordnungen mit keinem Fehler deutlich größer als derjenige der Jungen (Tabelle 6.1). Auch die Schulbildung hat einen (erwartungsgemäßen) Effekt. Je höher die Schulbildung, desto weniger Fehler wurden bei der Beantwortung gemacht. Vermutlich spielen kognitive Fähigkeiten – bei aller Einfachheit der Fragen – ebenso eine Rolle wie Konzentrationsfähigkeit und -wille (Tabelle 6.2).

Mit der Überlagerung der beiden Phänomene, also der schul- bzw. geschlechtsspezifischen Ausfallquote, die schon in den Querschnitten beobachtet werden konnte, wie der augenscheinlich schul- und geschlechtsspezifischen Fähigkeit, die Codeblätter richtig auszufüllen, ist die beobachtete Verzerrung der Paneldaten zu erklären.

6.2 Schulebene

Für die Teilnahme an den Befragungen war die Zustimmung der Schulkonferenzen der einzelnen Schulen notwendig, wobei schon an der ersten Befragung im Jahr 2002 nicht alle Schulen teilnahmen. Anhand der Daten des Vier-Wellen-Panels 2002 bis 2005 soll kurz die Systematik der Verzerrungen auf Schulebene dargestellt werden.

Wie in Tabelle 6.3 ersichtlich, liegt schon durch den Ausfall der Schulen im Jahr 2002 eine erhebliche Verschiebung der Daten in Bezug auf die Schulform vor. Wegen des Ausfalls auf Schulebene sind im Vergleich zwischen reduzierter Schulstatistik (nur an der Befragung teilnehmende Schulen) und gesamter Schulstatistik (alle Duisburger Schulen) zu t_1 und t_2 mehr Haupt- und Realschülerinnen und -schüler, dafür weniger Gesamtschüler und -schülerinnen sowie Gymnasiastinnen und Gymnasiasten und Sonderschülerinnen und -schüler vorhanden. Dieselbe Befundlage trifft für die Zeitpunkte t_3 und t_4 zu, jedoch ist der Anteil der Gesamtschülerinnen und -schüler hier ebenfalls erhöht. Durch die Ausfälle auf der Individualebene verringern sich die Verzerrungen jedoch wieder. So ist die Verteilung nach Schulform mit Ausnahme der Realschulen sowohl im Querschnitt als auch im Panelquerschnitt der gesamten Schulstatistik ähnlicher als der reduzierten.

Bei Längsschnittuntersuchungen liegt allerdings das Augenmerk nicht auf Repräsentativität der Stichproben, da dies für Untersuchungen von Veränderungen über die Zeit nicht relevant ist. Dennoch ist es unbestreitbar vorteilhafter, dass sich die genannten Verzerrungseffekte auf Individual- und Schulebene mit Blick auf die Repräsentativität der Daten gegenseitig eher aufheben als verstärken.

Tabelle 6.3: Zusammensetzung $P_{1,2,3,4}$ nach Geschlecht und Schulform (Zeilenprozentuierung)

		Schulform ^a (%)					gesamt N
		HS	RS	GS	GY	SO	(100 %)
t_1	S_{gt_1}	18,9	18,4	34,8	23,6	4,3	5240
	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
	$P_{1,2,3,4}^*$	17,8	24,7	33,6	23,9	—	1715
t_2	S_{gt_2}	21,3	17,6	33,8	22,3	5,0	5868
	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
	$P_{1,2,3,4}^*$	18,3	24,7	33,9	23,1	—	1715
t_3	S_{gt_3}	21,7	18,4	31,5	23,9	4,5	5778
	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
	$P_{1,2,3,4}^*$	18,4	25,4	33,8	22,4	—	1715
t_4	S_{gt_4}	20,7	18,6	31,4	24,5	4,8	5853
	S_{t_4}	22,5	22,0	33,2	22,1	0,2	3543
	$Q_{t_4}^*$	21,9	22,4	33,0	22,5	0,1	3243
	$P_{1,2,3,4}^*$	18,7	25,4	33,3	22,6	—	1715

^aHS=Hauptschule, RS=Realschule, GS=Gesamtschule,
 GY=Gymnasium, SO=Sonderschule
 S_g bezeichnet die Daten der gesamte Schulstatistik

Literaturverzeichnis

- Böttger, A., Ehret, B., Othold, F., Prein, G., Schumann, K. F. & Seus, L. (2003). Methoden der Untersuchung. In K. F. Schumann (Hg.), *Delinquenz im Lebensverlauf. Bremer Längsschnittstudie zum Übergang von der Schule in den Beruf bei ehemaligen Hauptschülern* (Bd. 2, S. 35–66). Weinheim, München: Juventa.
- Brondies, M. (2004). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2003* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 6). Münster, Trier.
- Hilfert, N. (2005). *Methodendokumentation der kriminologischen Schülerbefragung in Duisburg 2004* (Schriftenreihe »Jugendkriminalität in der modernen Stadt – Methoden« Nr. 10). Münster, Bielefeld.
- 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.
- 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.

A Anhang

A.1 Codeblätter



WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER
Prof. Dr. Klaus Boers
 Direktor des Instituts für Kriminalwissenschaften, Abteilung Kriminologie
Prof. Dr. Jost Reinecke
 Direktor der Arbeitsstelle Sozialindikatoren, Institut für Soziologie
 Bispinghof 24/25 48143 Münster Tel. 0251-83 22749
 kriminologie@uni-muenster.de

Kriminologische Befragung Duisburger Schülerinnen und Schüler
Beiblatt zur Erstellung des persönlichen Codes

Januar/Februar 2002

Liebe Schülerin, lieber Schüler,

wir wollen im nächsten Jahr die Befragung wiederholen, weil es sehr wichtig ist zu erfahren, was sich in der Zwischenzeit verändert hat. Damit nächstes Jahr deine beiden Fragebögen einander zugeordnet werden können, ohne deinen Namen zu kennen, benötigen wir ein Codewort. Dadurch kann niemand in Erfahrung bringen, dass du die beiden Fragebögen ausgefüllt hast. Wir möchten Dich deshalb bitten, auf diesem Zusatzblatt zum Fragebogen dieses Codewort zu bilden.

Wichtig ist also, dass du dasselbe Codewort nächstes Jahr noch weißt. Aus diesem Grund haben wir nachfolgend fünf Fragen formuliert, die dir helfen sollen, dich an dasselbe Wort zu erinnern.

*Wenn du eine der Fragen überhaupt nicht beantworten kannst,
 lass das entsprechende Feld bitte leer.*

Hier nun die fünf Fragen zur Erstellung deines persönlichen Codes:

	Bitte trage hier ins erste Feld den ersten Buchstaben des Vornamens deines Vaters (oder einer Person, die für dich einem Vater am nächsten kommt) ein. (z.B. Anton, Bernd, Hans-Peter, Fritz usw.)
	Bitte trage ins zweite Feld den ersten Buchstaben des Vornamens deiner Mutter (oder einer Person, die für dich einer Mutter am nächsten kommt) ein. (z.B. Anna, Beate, Jutta, Maria usw.)
	Bitte trage ins dritte und vierte Feld den Tag deines Geburtsdatums ein. (z.B. Geburtstag am 7. Januar = 0 7, am 12. Mai = 1 2, am 31. Oktober = 3 1)
	Bitte trage ins fünfte Feld den letzten Buchstaben deiner natürlichen Haarfarbe ein. (z.B. Glatze, braun, schwarz, blond, grau usw.)
	Bitte trage ins sechste Feld den letzten Buchstaben deiner Augenfarbe ein. (z.B. braun, blau, grün, grau usw.)

Abbildung A.1: Codeblatt Duisburg 2002

Beiblatt zur Erstellung des persönlichen Codes

Liebe Schülerin, lieber Schüler,

da wir deinen Fragebogen dem des letzten Jahres zuordnen wollen, bitten wir dich, deinen persönlichen Code zu wiederholen.

Da wir letztes Jahr das Codewort getrennt vom Fragebogen eingegeben, mit einem Computerprogramm verschlüsselt und danach vernichtet haben, ist es wichtig, dass du dich an dein Codewort erinnerst. Denn nur wenn wir mit diesem Codewort denselben Vorgang wiederholen, können deine Fragebögen einander zugeordnet werden und dies, ohne dass jemand herausfinden kann, wer diesen Fragebogen ausgefüllt hat.

Wichtig ist also, dass du dasselbe Codewort noch weißt. Aus diesem Grund haben wir letztes Jahr die nachfolgenden Fragen formuliert, die dir helfen sollen, dich an dasselbe Wort zu erinnern.

Wenn du eine der Fragen überhaupt nicht beantworten kannst, kreuze bitte kein Feld an!

Hier nun die sechs Fragen zur Erstellung deines persönlichen Codes:

1	Bitte kreuze den ersten Buchstaben des Vornamens deines Vaters (oder einer Person, die für dich einem Vater am nächsten kommt) an. (z. B. <u>A</u> nton, <u>B</u> ernd, <u>H</u> ans-Peter usw.) <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß
2	Bitte kreuze den ersten Buchstaben des Vornamens deiner Mutter (oder einer Person, die für dich einer Mutter am nächsten kommt) an. (z. B. <u>A</u> нна, <u>B</u> eatrice, <u>J</u> utta, <u>M</u> aria, usw.) <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß
3	Bitte kreuze den ersten Buchstaben deines Vornamens an (z. B. <u>M</u> ichael, <u>T</u> homas, <u>U</u> te usw.) <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß
4	Bitte kreuze den Tag deines Geburtsdatums an (z.B. Geburtstag am 7. Januar = <input type="checkbox"/> , am 12. Mai = <input checked="" type="checkbox"/> , am 31. Oktober = <input checked="" type="checkbox"/>) <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 <input type="checkbox"/> 7 <input type="checkbox"/> 8 <input type="checkbox"/> 9 <input type="checkbox"/> 10 <input type="checkbox"/> 11 <input type="checkbox"/> 12 <input type="checkbox"/> 13 <input type="checkbox"/> 14 <input type="checkbox"/> 15 <input type="checkbox"/> 16 <input type="checkbox"/> 17 <input type="checkbox"/> 18 <input type="checkbox"/> 19 <input type="checkbox"/> 20 <input type="checkbox"/> 21 <input type="checkbox"/> 22 <input type="checkbox"/> 23 <input type="checkbox"/> 24 <input type="checkbox"/> 25 <input type="checkbox"/> 26 <input type="checkbox"/> 27 <input type="checkbox"/> 28 <input type="checkbox"/> 29 <input type="checkbox"/> 30 <input type="checkbox"/> 31
5	Bitte kreuze den letzten Buchstaben deiner natürlichen Haarfarbe an. (z. B. braun <input type="checkbox"/> , Glatz <input type="checkbox"/> , schwarz <input type="checkbox"/> , usw.) <input type="checkbox"/> a <input type="checkbox"/> b <input type="checkbox"/> c <input type="checkbox"/> d <input type="checkbox"/> e <input type="checkbox"/> f <input type="checkbox"/> g <input type="checkbox"/> h <input type="checkbox"/> i <input type="checkbox"/> j <input type="checkbox"/> k <input type="checkbox"/> l <input type="checkbox"/> m <input type="checkbox"/> n <input type="checkbox"/> o <input type="checkbox"/> p <input type="checkbox"/> q <input type="checkbox"/> r <input type="checkbox"/> s <input type="checkbox"/> t <input type="checkbox"/> u <input type="checkbox"/> v <input type="checkbox"/> w <input type="checkbox"/> x <input type="checkbox"/> y <input type="checkbox"/> z <input type="checkbox"/> ä <input type="checkbox"/> ö <input type="checkbox"/> ü <input type="checkbox"/> ß
6	Bitte kreuze den letzten Buchstaben deiner Augenfarbe 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

Hast du im letzten Jahr an der Befragung teilgenommen? ja nein
 Hast du im letzten Jahr die Schule gewechselt? ja nein
 Bist du im letzten Jahr sitzen geblieben? ja nein

Abbildung A.2: Codeblatt Duisburg 2003

Erstellung des persönlichen Codes

Liebe Schülerin, lieber Schüler,

da wir deinen Fragebogen dem des letzten Jahres zuordnen wollen, bitten wir dich, deinen 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 du dich an deinen Code erinnerst. Denn nur wenn wir mit diesem Code denselben Vorgang wiederholen, können deine Fragebögen einander zugeordnet werden und dies, ohne dass jemand herausfinden kann, wer diesen Fragebogen ausgefüllt hat. Wichtig ist also, dass du denselben Code noch weißt. Aus diesem Grund haben wir letztes Jahr die nachfolgenden Fragen formuliert, die dir helfen sollen, dich an dasselbe Wort zu erinnern.

*Bitte kreuze bei jeder der sechs Fragen immer nur ein Feld an!
Wenn du eine der Fragen überhaupt nicht beantworten kannst, kreuze bitte kein Feld an!*

Hier nun die sechs Fragen zur Erstellung deines persönlichen Codes:

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

Hast du im letzten Jahr an der Befragung teilgenommen? ja nein
 Hast du im letzten Jahr die Schule gewechselt? ja nein
 Bist du im letzten Jahr sitzen geblieben? ja nein

Abbildung A.3: Codeblatt Duisburg 2004

Erstellung des persönlichen Codes

Liebe Schülerin, lieber Schüler,

da wir deinen Fragebogen dem des letzten Jahres zuordnen wollen, bitten wir dich, deinen 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 du dich an deinen Code erinnerst. Denn nur wenn wir mit diesem Code denselben Vorgang wiederholen, können deine Fragebögen einander zugeordnet werden und dies, ohne dass jemand herausfinden kann, wer diesen Fragebogen ausgefüllt hat. Wichtig ist also, dass du denselben Code noch weißt. Aus diesem Grund haben wir letztes Jahr die nachfolgenden Fragen formuliert, die dir helfen sollen, dich an dasselbe Wort zu erinnern.

*Bitte kreuze bei jeder der sechs Fragen immer nur ein Feld an!
Wenn du eine der Fragen überhaupt nicht beantworten kannst, kreuze bitte kein Feld an!*

Hier nun die sechs Fragen zur Erstellung deines persönlichen Codes:

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

Hast du im letzten Jahr an der Befragung teilgenommen? ja nein
 Hast du im letzten Jahr die Schule gewechselt? ja nein
 Bist du im letzten Jahr sitzen geblieben? ja nein
 In welcher Klasse bist du? 9 10

Abbildung A.4: Codeblatt Duisburg 2005

A.2 Häufigkeiten der Codevariablen

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

	2002		2003		2004		2005	
	N	%	N	%	N	%	N	%
a	279	8,1	261	7,7	266	7,8	292	8,4
ä	—	—	2	0,1	1	0,0	—	—
b	101	2,9	91	2,7	91	2,7	93	2,7
c	73	2,1	75	2,2	71	2,1	78	2,3
d	166	4,8	172	5,1	161	4,7	163	4,7
e	84	2,4	93	2,7	93	2,7	97	2,8
f	153	4,4	149	4,4	136	4,0	131	3,8
g	143	4,2	133	3,9	132	3,9	149	4,3
h	363	10,5	375	11,0	391	11,5	387	11,2
i	33	1,0	66	1,9	61	1,8	68	2,0
j	224	6,5	182	5,3	196	5,8	188	5,4
k	187	5,4	171	5,0	161	4,7	172	5,0
l	42	1,2	48	1,4	44	1,3	43	1,2
m	345	10,0	331	9,7	325	9,6	334	9,6
n	86	2,5	78	2,3	89	2,6	90	2,6
o	43	1,2	48	1,4	41	1,2	44	1,3
ö	3	0,1	4	0,1	5	0,1	10	0,3
p	124	3,6	118	3,5	118	3,5	114	3,3
q	2	0,1	2	0,1	3	0,1	—	—
r	253	7,4	263	7,7	262	7,7	243	7,0
s	151	4,4	139	4,1	154	4,5	151	4,4
ß	—	—	5	0,1	—	—	5	0,1
t	107	3,1	103	3,0	104	3,0	100	2,9
u	143	4,2	136	4,0	136	4,0	131	3,8
ü	5	0,1	5	0,1	7	0,2	5	0,1
v	71	2,1	71	2,1	70	2,0	83	2,4
w	164	4,8	165	4,8	163	4,8	160	4,6
x	3	0,1	4	0,1	7	0,2	6	0,2
y	41	1,2	36	1,1	42	1,2	37	1,1
z	15	0,4	18	0,5	12	0,4	22	0,6
fehlend	38	1,1	59	1,7	64	1,9	69	2,0
gesamt	3442	100,0	3403	100,0	3406	100,0	3465	100,0

A.2 Häufigkeiten der Codevariablen

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

	2002		2003		2004		2005	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
a	333	9,7	335	9,8	319	9,4	324	9,4
ä	1	0,0	3	0,1	2	0,1	4	0,1
b	318	9,2	287	8,4	291	8,6	286	8,3
c	180	5,2	184	5,4	185	5,4	184	5,3
d	94	2,7	85	2,5	102	3,0	101	2,9
e	190	5,5	196	5,8	198	5,8	204	5,9
f	84	2,4	83	2,4	74	2,2	83	2,4
g	196	5,7	191	5,6	185	5,4	177	5,1
h	242	7,0	245	7,2	248	7,3	241	7,0
i	73	2,1	127	3,7	134	3,9	138	4,0
j	109	3,2	58	1,7	49	1,5	59	1,7
k	135	3,9	136	4,0	125	3,7	135	3,9
l	61	1,8	63	1,9	72	2,1	56	1,6
m	386	11,2	402	11,8	376	11,0	385	11,1
n	113	3,3	120	3,5	118	3,5	125	3,6
o	15	0,4	14	0,4	12	0,4	17	0,5
ö	1	0,0	3	0,1	2	0,1	2	0,1
p	122	3,5	104	3,1	118	3,5	105	3,0
r	152	4,4	149	4,4	155	4,5	140	4,0
s	350	10,2	334	9,8	326	9,6	340	9,8
ß	—	—	1	0,0	—	—	7	0,2
t	37	1,1	37	1,1	42	1,2	48	1,4
u	116	3,4	107	3,1	115	3,4	110	3,2
ü	3	0,1	4	0,1	5	0,1	6	0,2
v	32	0,9	31	0,9	28	0,8	43	1,2
w	14	0,4	13	0,4	18	0,5	13	0,4
x	2	0,1	1	0,0	4	0,1	8	0,2
y	12	0,3	13	0,4	21	0,6	23	0,7
z	39	1,1	36	1,1	37	1,1	38	1,1
fehlend	32	0,9	41	1,2	45	1,3	63	1,8
gesamt	3442	100,0	3403	100,0	3406	100,0	3465	100,0

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

	2002		2003		2004		2005	
	N	%	N	%	N	%	N	%
a	—	—	271	8,0	275	8,1	286	8,3
ä	—	—	1	0,0	1	0,0	2	0,1
b	—	—	92	2,7	81	2,4	90	2,6
c	—	—	199	5,8	192	5,6	190	5,5
d	—	—	275	8,1	284	8,3	271	7,8
e	—	—	100	2,9	95	2,8	95	2,7
f	—	—	107	3,1	99	2,9	101	2,9
g	—	—	34	1,0	39	1,1	42	1,2
h	—	—	78	2,3	72	2,1	83	2,4
i	—	—	54	1,6	51	1,5	51	1,5
j	—	—	261	7,7	267	7,8	270	7,8
k	—	—	187	5,5	179	5,2	168	4,8
l	—	—	113	3,3	113	3,3	110	3,2
m	—	—	410	12,0	423	12,4	435	12,6
n	—	—	148	4,3	143	4,2	154	4,4
o	—	—	24	0,7	27	0,8	29	0,8
ö	—	—	15	0,4	16	0,5	17	0,5
p	—	—	113	3,3	112	3,3	111	3,2
q	—	—	4	0,1	2	0,1	2	0,1
r	—	—	106	3,1	99	2,9	96	2,8
s	—	—	452	13,3	455	13,4	450	13,0
ß	—	—	1	0,0	—	—	3	0,1
t	—	—	149	4,4	156	4,6	160	4,6
u	—	—	10	0,3	8	0,2	10	0,3
ü	—	—	3	0,1	2	0,1	4	0,1
v	—	—	72	2,1	70	2,1	76	2,2
w	—	—	10	0,3	9	0,3	19	0,5
x	—	—	5	0,1	3	0,1	4	0,1
y	—	—	43	1,3	56	1,6	52	1,5
z	—	—	14	0,4	19	0,6	18	0,5
fehlend	—	—	52	1,5	57	1,7	66	1,9
gesamt	—	—	3403	100,0	3406	100,0	3465	100,0

A.2 Häufigkeiten der Codevariablen

Tabelle A.4: Häufigkeiten von Co004 (Geburtstag)

	2002		2003		2004		2005	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
1	110	3,2	120	3,5	120	3,5	121	3,5
2	96	2,8	94	2,8	91	2,7	88	2,5
3	115	3,3	105	3,1	99	2,9	101	2,9
4	110	3,2	105	3,1	114	3,4	113	3,3
5	119	3,5	115	3,4	118	3,5	109	3,1
6	118	3,4	107	3,1	113	3,3	114	3,3
7	111	3,2	109	3,2	103	3,0	106	3,1
8	114	3,3	111	3,3	118	3,5	116	3,3
9	129	3,7	118	3,5	136	4,0	132	3,8
10	130	3,8	136	4,0	120	3,5	125	3,6
11	105	3,1	119	3,5	117	3,4	115	3,3
12	124	3,6	97	2,9	94	2,8	100	2,9
13	112	3,3	99	2,9	96	2,8	107	3,1
14	119	3,5	110	3,2	110	3,2	117	3,4
15	111	3,2	121	3,6	128	3,8	121	3,5
16	122	3,5	123	3,6	117	3,4	124	3,6
17	98	2,8	110	3,2	105	3,1	110	3,2
18	116	3,4	116	3,4	113	3,3	128	3,7
19	126	3,7	124	3,6	131	3,8	123	3,5
20	110	3,2	106	3,1	101	3,0	113	3,3
21	101	2,9	101	3,0	103	3,0	101	2,9
22	95	2,8	100	2,9	95	2,8	96	2,8
23	113	3,3	110	3,2	111	3,3	111	3,2
24	110	3,2	96	2,8	98	2,9	104	3,0
25	116	3,4	106	3,1	107	3,1	116	3,3
26	123	3,6	114	3,3	114	3,4	120	3,5
27	91	2,6	90	2,6	87	2,6	88	2,5
28	90	2,6	87	2,6	85	2,5	92	2,7
29	109	3,2	100	2,9	96	2,8	109	3,1
30	100	2,9	101	3,0	92	2,7	90	2,6
31	64	1,9	66	1,9	71	2,1	56	1,6
85	1	0,0	—	—	—	—	—	—
96	1	0,0	—	—	—	—	—	—
223	1	0,0	—	—	—	—	—	—
fehlend	32	0,9	87	2,6	101	3,0	99	2,9
gesamt	3442	100,0	3403	100,0	3406	100,0	3465	100,0

Tabelle A.5: Häufigkeiten von Co005 (Haarfarbe)

	2002		2003		2004		2005	
	N	%	N	%	N	%	N	%
a	—	—	1	0,0	3	0,1	4	0,1
ä	—	—	—	—	2	0,1	3	0,1
b	59	1,7	181	5,3	131	3,9	106	3,1
c	—	—	3	0,1	—	—	1	0,0
d	1235	35,9	1080	31,7	1115	32,7	1094	31,6
e	14	0,4	19	0,6	18	0,5	18	0,5
f	—	—	—	—	—	—	4	0,1
g	2	0,1	1	0,0	8	0,2	13	0,4
h	—	—	3	0,1	2	0,1	7	0,2
i	—	—	1	0,0	1	0,0	5	0,1
j	—	—	—	—	1	0,0	5	0,1
k	—	—	1	0,0	3	0,1	2	0,1
l	—	—	—	—	2	0,1	2	0,1
m	1	0,0	2	0,1	2	0,1	9	0,3
n	1614	46,9	1536	45,1	1557	45,7	1613	46,6
o	—	—	1	0,0	2	0,1	5	0,1
ö	—	—	1	0,0	4	0,1	4	0,1
q	—	—	—	—	1	0,0	1	0,0
r	2	0,1	3	0,1	3	0,1	5	0,1
s	12	0,3	40	1,2	36	1,1	38	1,1
ß	—	—	4	0,1	—	—	4	0,1
t	36	1,0	51	1,5	33	1,0	34	1,0
u	5	0,1	1	0,0	4	0,1	19	0,5
ü	—	—	1	0,0	2	0,1	4	0,1
v	—	—	—	—	1	0,0	3	0,1
w	1	0,0	2	0,1	1	0,0	5	0,1
x	1	0,0	—	—	3	0,1	6	0,2
y	—	—	—	—	4	0,1	4	0,1
z	429	12,5	421	12,4	414	12,2	384	11,1
fehlend	31	0,9	50	1,5	51	1,5	63	1,8
Gesamt	3442	100,0	3403	100,0	3406	100,0	3465	100,0

A.2 Häufigkeiten der Codevariablen

Tabelle A.6: Häufigkeiten von Co006 (Augenfarbe)

	2002		2003		2004		2005	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
a	1	0,0	8	0,2	9	0,3	9	0,3
ä	—	—	—	—	1	0,0	2	0,1
b	44	1,3	143	4,2	104	3,1	95	2,7
c	—	—	—	—	1	0,0	1	0,0
d	2	0,1	6	0,2	9	0,3	12	0,3
e	3	0,1	3	0,1	2	0,1	3	0,1
f	—	—	—	—	2	0,1	6	0,2
g	21	0,6	50	1,5	46	1,3	27	0,8
h	—	—	—	—	1	0,0	7	0,2
i	—	—	1	0,0	1	0,0	1	0,0
j	—	—	1	0,0	2	0,1	7	0,2
k	—	—	1	0,0	2	0,1	3	0,1
l	—	—	3	0,1	2	0,1	9	0,3
m	1	0,0	3	0,1	—	—	10	0,3
n	2244	65,2	2110	62,0	2129	62,5	2152	62,1
o	—	—	—	—	1	0,0	3	0,1
ö	—	—	1	0,0	1	0,0	7	0,2
p	—	—	—	—	—	—	1	0,0
q	—	—	1	0,0	1	0,0	—	—
r	—	—	3	0,1	—	—	—	—
s	5	0,1	8	0,2	3	0,1	7	0,2
ß	—	—	1	0,0	—	—	5	0,1
t	1	0,0	—	—	2	0,1	1	0,0
u	1070	31,1	951	27,9	982	28,8	966	27,9
ü	—	—	1	0,0	1	0,0	1	0,0
v	—	—	1	0,0	—	—	2	0,1
w	—	—	—	—	3	0,1	5	0,1
x	1	0,0	1	0,0	2	0,1	5	0,1
y	—	—	1	0,0	1	0,0	4	0,1
z	18	0,5	53	1,6	36	1,1	38	1,1
fehlend	31	0,9	52	1,5	61	1,8	76	2,2
Gesamt	3442	100,0	3403	100,0	3406	100,0	3465	100,0

Tabelle A.7: Häufigkeiten von Co007 in Duisburg 2003 (Befragungsteilnahme im Vorjahr ja/nein)

		Häufigkeit	gültige Prozent
gültig	ja	2 579	75,8
	nein	577	17,0
	k. A.	247	7,3
gesamt		3 403	100,0

Tabelle A.8: Häufigkeiten von Co007 in Duisburg 2004 (Befragungsteilnahme im Vorjahr ja/nein)

		Häufigkeit	gültige Prozent
gültig	ja	2 696	79,1
	nein	385	11,3
	k. A.	325	9,5
gesamt		3 406	100,0

Tabelle A.9: Häufigkeiten von Co007 in Duisburg 2005 (Befragungsteilnahme im Vorjahr ja/nein)

		Häufigkeit	gültige Prozent
gültig	ja	2 873	82,9
	nein	227	6,6
	k. A.	365	10,5
gesamt		3 465	100,0

A.2 Häufigkeiten der Codevariablen

Tabelle A.10: Häufigkeiten von Co008 in Duisburg 2003 (Schulwechsel ja/nein)

		Häufigkeit	gültige Prozent
gültig	ja	213	6,3
	nein	2843	83,5
	k. A.	347	10,2
gesamt		3403	100,0

Tabelle A.11: Häufigkeiten von Co008 in Duisburg 2004 (Schulwechsel ja/nein)

		Häufigkeit	gültige Prozent
gültig	ja	195	5,7
	nein	2760	81,0
	k. A.	451	13,3
gesamt		3406	100,0

Tabelle A.12: Häufigkeiten von Co008 in Duisburg 2005 (Schulwechsel ja/nein)

		Häufigkeit	gültige Prozent
gültig	ja	178	5,1
	nein	2803	80,9
	k. A.	484	14,0
gesamt		3465	100,0

Tabelle A.13: Häufigkeiten von Co009 in Duisburg 2003 (Sitzengeblieben ja/nein)

		Häufigkeit	gültige Prozent
gültig	ja	184	5,4
	nein	2 880	84,6
	k. A.	339	10,0
gesamt		3 403	100,0

Tabelle A.14: Häufigkeiten von Co009 in Duisburg 2004 (Sitzengeblieben ja/nein)

		Häufigkeit	gültige Prozent
gültig	ja	203	6,0
	nein	2 746	80,6
	k. A.	457	13,4
gesamt		3 406	100,0

Tabelle A.15: Häufigkeiten von Co009 in Duisburg 2005 (Sitzengeblieben ja/nein)

		Häufigkeit	gültige Prozent
gültig	ja	186	5,4
	nein	2 792	80,6
	k. A.	487	14,1
gesamt		3 465	100,0

Tabelle A.16: Häufigkeiten der Gesamtcodes in Duisburg 2002

		Häufigkeit	Prozent	gültige Prozent
gültig	1	3064	89,0	89,8
	2	324	9,4	9,5
	3	18	0,5	0,5
	5	5	0,1	0,1
	gesamt	3411	99,1	100,0
fehlend		31	0,9	
gesamt		3442	100,0	

Tabelle A.17: Häufigkeiten der Gesamtcodes in Duisburg 2003

		Häufigkeit	Prozent	gültige Prozent
gültig	1	3352	98,5	99,1
	2	32	0,9	0,9
	gesamt	3384	99,4	100,0
fehlend		19	0,6	
gesamt		3403	100,0	

Tabelle A.18: Häufigkeiten der Gesamtcodes in Duisburg 2004

		Häufigkeit	Prozent	gültige Prozent
gültig	1	3350	98,4	98,8
	2	40	1,2	1,2
	gesamt	3390	99,5	100,0
fehlend		16	0,5	
gesamt		3406	100,0	

Tabelle A.19: Häufigkeiten der Gesamtcodes in Duisburg 2005

		Häufigkeit	Prozent	gültige Prozent
gültig	1	3358	96,9	98,4
	2	56	1,6	1,6
	gesamt	3414	98,5	100,0
fehlend		51	1,5	
gesamt		3465	100,0	