

Swiss Leading House

Economics of Education • Firm Behaviour • Training Policies

Working Paper No. 19

Schulleistungen von Mädchen und Jungen. Gleichberechtigung als Bildungsmotor?

Florian Birkenfeld



Universität Zürich

ISU – Institut für Strategie und Unternehmensökonomik

u^b

b
**UNIVERSITÄT
BERN**

Leading House Working Paper No. 19

Schulleistungen von Mädchen und Jungen. Gleichberechtigung als Bildungsmotor?

Florian Birkenfeld

February 2008

Die Discussion Papers dienen einer möglichst schnellen Verbreitung von neueren Forschungsarbeiten des Leading House. Die Beiträge liegen in alleiniger Verantwortung der Autoren und stellen nicht notwendigerweise die Meinung des Leading House dar.

Discussion Papers are intended to make results of the Leading House research promptly available to other economists in order to encourage discussion and suggestions for revisions. The authors are solely responsible for the contents which do not necessarily represent the opinion of the Leading House.

The Swiss Leading House on Economics of Education, Firm Behavior and Training Policies is a Research Programme of the Swiss Federal Office for Professional Education and Technology (OPET).

www.economics-of-education.ch

Schulleistungen von Mädchen und Jungen. Gleichberechtigung als Bildungsmotor?*

Florian Birkenfeld[†]

Universität Passau

22. Februar 2008

Diese Arbeit untersucht Unterschiede in den Schulleistungen zwischen Jungen und Mädchen sowie mögliche Ursachen, die sich aus Charakteristika der Eltern und institutionellen Rahmenbedingungen ergeben. Hierzu werden die Daten der Schulleistungsstudie PISA 2000 mit einer Drei-Ebenenanalyse untersucht. Dabei wird auf den Global Gender Gap Index des World Economic Forum als Maß für Gleichberechtigung zurückgegriffen. Die Bildung der Eltern hat eine deutliche Wirkung auf die Schulleistungen der Kinder. Das Abitur der Mütter ist für Töchter wichtiger als das der Väter, welches wiederum für die Söhne wichtiger ist. Gleichzeitig unterliegt der Einfluss des Vaters einer größeren Varianz zwischen den 37 Staaten. Größere Anstrengungen zur Gleichberechtigung von Frauen können durchaus ein Schlüssel für zukünftigen Bildungserfolg sein. Dies gilt insbesondere für die Länder die bei PISA 2000 speziell beim Leseverständnis unterdurchschnittlich abgeschnitten hatten.

JEL Klassifikation: I21; I28; J16

Schlagworte: Schulleistung; Familie; Gleichberechtigung; Mehrebenenanalyse

*Mein Dank gilt den Teilnehmern des Kurses *Economics of Education and the Labour Market* des *Swiss Leading House „Economics of Education – Firm Behaviour – Training Policies“* im April 2007 an der Universität Zürich, insbesondere Barbara Müller für kritische Anmerkungen und Paul Bingley für wertvolle Hinweise zur Mehrebenenanalyse. Ferner gilt mein Dank den Teilnehmern des Dokrorandenkolloquiums der Universität Passau vom November 2007, hier insbesondere Gerald Csipek. Die möglicherweise verbliebenen Fehler gehen dennoch zu meinen Lasten.

[†]Florian Birkenfeld, Wirtschaftswissenschaftliche Fakultät, Universität Passau, Innstraße 27, 94032 Passau. Tel +49 (0) 851 509-2544, e-mail: florian.birkenfeld@uni-passau.de.

1 Theorie und Literatur

Individuelle Lebenschancen wie auch die Wohlfahrt einer Volkswirtschaft hängen in großem Maße von Bildung und Ausbildung ab.¹ Beides kann – im Sinne der Startchancengerechtigkeit bzw. der Internalisierung externer Effekte – einen Eingriff des Staates rechtfertigen.

Dort, wo einzelne Gruppen einen größeren Nutzen aus dem Bildungssystem ziehen als andere oder die Bildungsrenditen hinter denen anderer Staaten zurückbleiben, gilt es, die konkrete Ausgestaltung des Systems zu überdenken. Andernfalls beraubt sich eine Gesellschaft ihrer Zukunftschancen. Gruppen, die häufig mit geringeren Schulleistungen in Verbindung gebracht werden sind Migrantinnen- und Arbeiterkinder, aber insbesondere in weniger entwickelten Volkswirtschaften auch Mädchen.

Ziel der Analyse ist es, geschlechtsspezifische Unterschiede in der Schulleistung aufzuzeigen und mittels geschlechtsspezifischer Einflüsse vor dem Hintergrund unterschiedlicher Gleichberechtigung in den einzelnen Staaten zu erklären.

1.1 Schulleistungen von Mädchen und Jungen

Mädchen und Jungen haben aufgrund unterschiedlich verlaufender Entwicklungen im Kindesalter² zum Zeitpunkt des PISA-Tests im Durchschnitt recht unterschiedlich entwickelte Fähigkeiten in den Bereichen Leseverständnis, Mathematik und Naturwissenschaften: „[G]irls perform much better in reading than boys, who tend to do better in mathematics although the difference is less pronounced. .. In less-developed countries, the over-performance of girls is weaker or often absent.“ (UNESCO, 2003, 11)

1.2 Einfluss von Mutter bzw. Vater

Der familiäre Hintergrund ist eine der zentralen Determinanten für späteren schulischen Erfolg. Dieser Zusammenhang gilt nicht für alle Schulsysteme gleichermaßen, ist aber doch in nahezu allen Staaten zu beobachten (Wößmann (2004), Gundlach und

¹Zu privaten und gesellschaftlichen Bildungsrenditen siehe u.a. Acemoglu und Angrist (1999), Wolter und Weber (1999) und OECD (2005, 156-158). Zu Bildung als Wachstumsmotor siehe bspw. Hanushek und Kimko (2000) sowie Barro (2001).

²Eine Übersicht findet der interessierte Leser u.a bei Eliot (2001, Kap. 16) und Plomin (1990, Kap. 4 u. 5)

Wößmann (2004)). Dies reicht über die Familienstruktur, die Bildungsabschlüsse und die Berufstätigkeit der Eltern bis zur Hausaufgabenunterstützung.

In welchem Ausmaß die Gene (*nature*) bzw. die Umgebung (*nurture*) die beobachteten kognitiven und nicht-kognitiven Fähigkeiten beeinflussen, ist nach wie vor umstritten. Der Einfluss von Genen und Umgebung scheint danach jeweils etwa zur Hälfte für die Intelligenz verantwortlich zu sein. Der Einfluss der Gene nimmt aber mit dem Alter zu. (Eliot (2001, 599-610), Plomin (1990, Kap. 5))

Behrmann und Taubmann (1989) finden in Zwillingstudien, dass Unterschiede in der genetischen Ausstattung 81 Prozent der Varianz der Dauer des Schulbesuchs erklären können. Behrmann und Rosenzweig (2002) finden dagegen in einer Studie mit eineiigen Zwillingen, dass der landläufig unterstellte Zusammenhang zwischen der Schulausbildung der Mutter und der ihrer Kinder durch vererbte Begabung – „more able women, who have more schooling, have more able children, who obtain more schooling“ – und „assortative mating“ in großem Ausmaß nach oben verzerrt ist. Tatsächlich zeigten ihre Untersuchungen keinen positiven Effekt einer besseren Schulausbildung von Müttern auf ihre Kinder.

Bohrhardt (2000, 196) vermutet: „Eltern mit niederen Bildungsabschlüssen werden auch über weniger Einkommen verfügen, was die Bildungschancen der Kinder verkleinert etwa durch die Nichtfinanzierbarkeit von Nachhilfestunden, privatem Musikunterricht, Sprachreisen etc.“ und findet weiter, dass die Bildung des Vaters in Deutschland kaum eine Rolle dafür spielt, wenn man auch die Bildung der Mutter in die Schätzung mit aufnimmt, ob ein Kind einen Schulabschluss erreicht oder nicht.

De Haan und Plug (2006) finden in einem Methodenüberblick ebenfalls kaum einen Einfluss der Schulbildung der Mutter, wenn entsprechende Kontrollvariablen verwendet werden.

Black et al. (2005) nutzen eine Reform, die in den 1960er Jahren in Norwegen die Dauer der Schulpflicht von sieben auf neun Jahre angehoben hat. Die Umsetzung in den einzelnen Landkreisen fand zu unterschiedlichen Zeitpunkten statt (1960 bis 1972). Insgesamt ergibt sich so – in einem gewissen Rahmen – eine exogene Dauer der Schulausbildung der Eltern. Der einzige kausale Zusammenhang den Black et al. aufdecken können, ist der zwischen der Schulausbildung von Müttern und Söhnen. Darüber hinaus zeigen sie, dass besser ausgebildete Frauen nicht notwendigerweise besser ausgebildete oder reichere Männer heiraten und dass sie nicht weniger Kinder im Sinne eines Qualität-Quantitäts-Tradeoffs (Becker und Lewis, 1973) haben. Der Unterschied in der Schulausbildung auf-

grund der untersuchten Reform ist zudem eher gering.

Es scheint also durchaus lohnend, Variablen die den familiären Hintergrund abdecken in die Untersuchung aufzunehmen. Wie stark der Einfluss des Bildungshintergrundes der Mütter auf die Töchter bzw. Söhne jeweils sein kann, hängt dabei sicher auch von den politischen Rahmenbedingungen ab.

Der Global Gender Gap Report des World Economic Forum (2006) versucht, den Grad der Benachteiligung von Frauen in 115 Ländern zu messen. Dazu werden in den vier Bereichen

- „economic participation and opportunity“,
- „educational attainment“,
- „health“ und
- „survival and political empowerment“

jeweils mehrere Variablen herangezogen. Der Report orientiert sich dabei allein an Ergebnissen (Unterschied in der Lebenserwartung, Frauenanteil im Parlament etc.) und nicht am politischen (guten) Willen. Eine Darstellung der Variablen findet sich bei World Economic Forum (2006, 5f).

Es ist zu erwarten, dass Mädchen in Staaten mit einem größeren Ausmaß an Gleichberechtigung bessere Schulleistungen erreichen und dass der Einfluss der Mütter ebenfalls stärker ausfällt.

Die Angaben zum Bruttonationaleinkommen pro Kopf³ in Abbildung 1 stammen aus der Datenbank der Weltbank. Der (hoch signifikante) Korrelationskoeffizient des GGG-Index mit dem BNE pro Kopf beträgt 0,457, es ist also durchaus nachvollziehbar, dass „Nations with the highest disparities ... tend to be the most disadvantaged in economic terms“. (UNESCO, 2003, 6)

Multikollinearität sollte nach den von Gujarati (1995, 335-339) angeführten Faustregeln ($R^2 < 0,36$, nicht zu viele insignifikante Variablen) für die verwendete Stichprobe kein größeres Problem darstellen. Die von Goldberger (1991, 249) eingeführte *micronumerosity* ist hier wohl bedeutender. Allerdings führen beide Phänomene zu unterschätzten

³Geht man davon aus, dass Kinder am Wohnort der Eltern zur Schule gehen, ist das Inländerkonzept des BNE näher an der Einkommenssituation der Eltern als das Inlandskonzept des BIP und somit die geeignetere Variable.

Standardfehlern, daher werden knapp signifikante Ergebnisse nur sehr vorsichtig interpretiert werden können.



Abbildung 1: Bruttonationaleinkommen pro Kopf in Euro (rechte Achse, Linie) und Global-Gender-Gap-Index (linke Achse, Säulen)

1.3 Die Schulen

Schulen unterscheiden sich in einer Vielzahl von Merkmalen. Über die offensichtlichen Unterschiede in einem gegliederten Schulsystem hinaus mag die Lage einer Schule im Zentrum einer Großstadt oder in einem Dorf sowie die Ausstattung mit qualifiziertem Personal und Lehrmitteln einen Einfluss auf die Schülerleistungen haben.

Von besonderem Interesse für eine Arbeit, die die Unterschiede in der Schulleistung von Mädchen und Jungen aufzeigen möchte, ist natürlich die Zusammensetzung der Schülerschaft einer Schule nach Geschlechtern. Einen Überblick zur Entsprechenden Diskussion liefert Salomone (2003).

Die Literatur zu sogenannten Peer-Effects hat sich diesen Phänomenen gewidmet. Ein schönes Beispiel für die auftretenden Probleme liefern Vigdor und Nechyba (2007) die auch einen schwachen positiven Effekt – auf die Lese- und Mathematikleistung – einer höheren Anzahl von Mädchen in einer Klasse finden.

1.4 Produktionsfunktion

„Education is a service that transforms fixed quantities of inputs (that is, individuals) into individuals with different qualities.“ (Hanushek, 1986, 1150) Produktionsfunktionen im Bildungsbereich können – abhängig von der Fragestellung – grundsätzlich die Leistung zu einem bestimmten Zeitpunkt (Level) oder die Leistungsverbesserung in einer Periode (Value-added) erklären (vgl. u.a. Hanushek (1986, 1156) und den Überblicksaufsatz von Todd und Wolpin (2003)). Da für die Unterschiede der Schulleistungen von Mädchen und Jungen auch Variablen außerhalb der zu einem bestimmten Zeitpunkt besuchten Schulklasse herangezogen werden sollen, bietet sich eine Produktionsfunktion in Level-Form an. Da ihre genaue Form unbekannt ist, müssen die Parameter mit ökonometrischen Verfahren geschätzt werden (Hanushek, 1986, 1149).

Die Fokussierung auf kognitive Leistungen abstrahiert von anderen Aufträgen, die Schulen und verschiedene Schulformen in unterschiedlichen Staaten haben mögen. „Nevertheless, performance on tests is being used to evaluate educational programs, and even to allocate funds, and there are some pragmatic arguments for the use of test scores as output measures.“ (Hanushek, 1986, 1154)

1.5 Mehrebenenanalyse

Für die Schätzungen kommt die Software HLM 6.02 (Raudenbush et al., 2004) zum Einsatz. Untersucht werden Schüler, Schulen und Staaten. Es werden also „Objekte verschiedener Ordnung gleichzeitig zum Gegenstand der Untersuchung“ (Hummell, 1972, 13). Für solche Datensätze mit hierarchischer Struktur ist die Mehrebenenanalyse prädestiniert.

Bei gleichzeitigem Vorliegen von Daten auf Individual- und Aggregatebene bieten sich intuitiv zwei Vorgehensweisen an, um diese Daten mit ökonometrischen Standardverfahren zu untersuchen. Zunächst könnte man versucht sein, die Individualdaten zu aggregieren, um dann auf der Aggregatebene eine Regression durchzuführen. So könnte man beispielsweise die durchschnittlichen PISA-Ergebnisse in den Staaten mit Charakteristika des Schulwesens (Aggregatvariable) und durchschnittlichen Schülereigenschaften (aggregierte Variable) erklären wollen. Fehlerfrei ist dieses Verfahren jedoch nur, wenn die Aggregate in sich homogen sind (Ditton, 1998, 31), da es sonst zu einem so genannten *aggregation bias* kommen kann. Diese Verzerrung tritt auf, wenn aggregierte Daten andere Koeffizienten liefern als nicht aggregierte Daten und das Ergebnis der Untersuchung daher stark vom Untersuchungsdesign abhängt (Sellin (1990), sowie Kreft und de Leeuw (1988)). Betrachtet man lediglich Aggregate, so können Interaktionseffekte von Individualvariablen nicht mehr untersucht werden. Da außerdem in bildungsökonomischen Untersuchungen der einzelne Schüler und nicht die Schule oder die Region die relevante Ebene ist (Wößmann, 2003, 124), sollte diese Herangehensweise verworfen werden.

Alternativ könnte man die Individualdaten um Aggregatdaten (Modellierung über sog. *fixed effects*) ergänzen, beispielsweise jedem Schüler Daten über seine Schule oder über den Staat, in dem er lebt, zuordnen. Die Anzahl der Freiheitsgrade (hier etwa die Anzahl der Schüler abzüglich der Anzahl der verwendeten Variablen) ist dann jedoch nicht mehr korrekt, da ein Teil der Daten jedes Schülers ja aus einer Aggregatebene mit – meist deutlich – weniger Beobachtungen stammt. Dadurch sind dann auch die Signifikanztests verfälscht. Da die Beobachtungen nicht unabhängig von einander sind, werden Fehlerterme korreliert sein. (Ditton (1998, 32), Blien und Wiedenbeck (2002, 313))

Untersucht man hierarchische Daten ohne die Mehrebenenstruktur zu berücksichtigen, kann die Analyse also fehlerhaft oder sogar gänzlich unbrauchbar sein. Es handelt sich weder bei den Schülern noch bei den Schulen um jeweils voneinander unabhängige Beobachtungen, sondern vielmehr um ausgewählte Cluster. Die Schüler innerhalb einer Schule

sind einander ähnlicher als bei einer tatsächlichen, nicht geclusterten Zufallsstichprobe. Gleiches gilt für die Schulen innerhalb eines Staates. Die Standardfehler, die mit statistischen Routineverfahren ermittelt würden, wären also wegen dieser Gruppierung mit großer Wahrscheinlichkeit fehlerhaft. (Vgl. zur Problematik u.a. Hox (2002, 5f) Ditton (1998, 13-15), Cronbach (1976), sowie Maas und Hox (2002, 2))

Die Mehrebenenanalyse erlaubt nun, gleichzeitig den Einfluss von Individual- (z.B. Geschlecht, Alter) und Aggregatvariablen (z.B. Charakteristika der Schulklasse und der Schule bzw. des Staates) sowie das Zusammenwirken von Individual- und Aggregatvariablen zu untersuchen (Ditton, 1998, 16).

Für eine Zwei-Ebenen-Analyse liefert eine Stichprobe „involving more schools and fewer students per school“ bessere Ergebnisse als eine Stichprobe „involving fewer schools and more students per school.“ (Mok (1995, 15), vgl. auch Maas und Hox (2002, 6)) Das Stichprobendesign ist für die vorliegende Untersuchung jedoch exogen.

Allerdings findet Mok (1995, 13) auch, dass ab einer Stichprobengröße von 800 Schülern die Schätzungen – unabhängig davon, ob nun mehr Schüler pro Schule oder mehr Schulen vorliegen – bereits recht gut werden.

Maas und Hox (2002) untersuchen anhand eines Zwei-Ebenen-Modells die für unverzerrte Ergebnisse nötigen Stichprobengrößen. Bereits ab 50 Einheiten auf der zweiten Ebene sind die geschätzten Koeffizienten verlässlich, aber auch mit nur 30 Einheiten sind die Ergebnisse bereits vorsichtig interpretierbar. Laut Browne und Draper (2000, 400-410) sollen mit restricted maximum likelihood Schätzungen sechs bis zwölf Gruppen genügen, um keine zu kleinen Standardfehler zu erzeugen, bei full maximum likelihood – wie von der Software HLM verwendet – 48 Gruppen. Leider führen Browne und Draper keine Schritte zwischen zwölf und 48 auf.

Afshartous (1995, 10) findet in einem Zwei-Ebenen-Modell bereits ab einer Zahl von 40 Schulen unverzerrte Schätzer.

Die verwendeten Daten enthalten (wie in Abschnitt 2.3 beschrieben) 5360 Schulen mit durchschnittlich 20 Schülern pro Schule sowie 38 Staaten mit durchschnittlich 141 Schulen pro Staat. Es ist also davon auszugehen, dass die Standardfehler etwas zu gering ausgewiesen werden.

| | Mittelwert | Std.abweichung | Minimum | Maximum |
|----------------|------------|----------------|---------|---------|
| gggi_score | 0,7003 | 0,045 | 0,6447 | 0,8133 |
| anteil_mädchen | 0,50 | 0,14 | 0,001 | 0,998 |
| weiblich | 0,52 | 0,50 | 0 | 1 |
| m_inländer | 0,92 | 0,28 | 0 | 1 |
| v_inländer | 0,92 | 0,28 | 0 | 1 |
| m_beruf | 0,66 | 0,47 | 0 | 1 |
| m_abitur | 0,48 | 0,50 | 0 | 1 |
| v_abitur | 0,47 | 0,50 | 0 | 1 |
| m_allein | 0,06 | 0,24 | 0 | 1 |
| v_allein | 0,01 | 0,11 | 0 | 1 |

Tabelle 1: Deskriptive Statistiken der erklärenden Variablen

| | Mittelwert | Std.abweichung | Minimum | Maximum |
|-----------------|------------|----------------|---------|---------|
| bne_kopf | 16673 | 13014 | 590 | 43560 |
| <i>scmatedu</i> | 0,29 | 1,12 | -1,90 | 3,22 |
| Privatschule | 0,15 | 0,36 | 0 | 1 |
| Schuljahr | 9,55 | 0,74 | 4 | 12 |
| Inländer | 0,96 | 0,21 | 0 | 1 |
| <i>famedsup</i> | 0,04 | 1,04 | -1,49 | 3,35 |
| <i>hedres</i> | -0,23 | 1,16 | -5,93 | 0,76 |
| Testsprache | 0,90 | 0,31 | 0 | 1 |
| Patchworkfam. | 0,08 | 0,28 | 0 | 1 |
| Kernfamilie | 0,84 | 0,37 | 0 | 1 |

Tabelle 2: Deskriptive Statistiken der Kontrollvariablen

2 Daten

Der PISA-2000-Datensatz besteht aus 228784 Schülern, die 8526 Schulen in 43 Ländern besuchen. Alle Schüler haben am Leseverständnistest teilgenommen, 127236 zusätzlich am naturwissenschaftlichen Test und 127388 statt dessen am Mathematiktest.

Die im PISA-Datensatz enthaltenen Gewichte⁴ werden nicht verwendet (Vgl. zu diesem Vorgehen z.B. Dronkers und Robert (2003)). Ergebnisse mit und ohne Gewichtungsfaktoren unterscheiden sich bei Verwendung umfangreicher Schulleistungsstudien ohnehin nur geringfügig. (Wößmann, 2003, 128)

Deskriptive Statistiken zu den verwendeten Variablen finden sich in Tabelle 1.

2.1 Staatenebene

Der GGG-Report 2006 des World Economic Forum (2006) beinhaltet alle Staaten, die auch an PISA 2000 teilnahmen mit Ausnahme von Hong Kong und Liechtenstein. Beide Staaten konnten daher nicht berücksichtigt werden.

Für Korea fehlen jeweils die Angaben für die Geburtsländer des Schülers, der Mutter und des Vaters. Für die Untersuchung ist es jedoch interessant, festzustellen ob es für die Schulleistung von Migrantenkindern einen Unterschied macht, ob der Vater und/oder die Mutter im Ausland geboren wurde.

Für Kanada fehlen Angaben zur Anzahl der Mädchen und Jungen an den Schulen. Da der Einfluss des Mädchenanteils an der Schule (der sich auch in den einzelnen Klassen widerspiegeln dürfte) auf die jeweiligen Leistungen untersucht werden soll, wird auch Kanada nicht in die Schätzung einbezogen. Ungarn wird nicht berücksichtigt, da hier für keinen Schüler bekannt ist, ob zu Hause vor allem die Testsprache gesprochen wird. Da es sich bei reinen Mädchen- bzw. Jungenschulen möglicherweise häufiger als bei gemischten Schulen um Internate und/oder Privatschulen handelt, soll der reine Privatschuleffekt mit einem entsprechenden Dummy aufgefangen werden. Dieser steht für Australien nicht zur Verfügung.

Es fehlen also insgesamt Australien, Hong Kong, Liechtenstein, Kanada, Korea und Ungarn. Der gesamte Datensatz besteht nun noch aus 37 Ländern.

⁴Zur Ausgestaltung der Gewichte und den Problemen der Gewichtung bei Verwendung aller drei Teiltests siehe OECD (2000, 17f). Die Gewichte wurden jeweils für den Lese-, den mathematischen und den naturwissenschaftlichen Test separat ermittelt da die Auswahl der Schüler nicht durchgehend repräsentativ erfolgte.

2.2 Schülerebene

Für die Variablen *weiblich*, das Geburtsland des Schülers bzw. eines Elternteils, die Familienstruktur, die Berufstätigkeit der Mutter⁵ (*m_beruf*), ob zu Hause die Testsprache gesprochen wird sowie die im PISA Datensatz bereits errechneten Indizes sind fehlende Werte (Missings), wenn überhaupt, nur sehr schwer durch geeignete Verfahren (Imputation) zu ersetzen. Entsprechende Schüler wurden daher aus dem Datensatz entfernt. Für die Variable *schuljahr* könnten Missings durch den Modus (der häufigste Wert) der Schüler der gleichen Schule ersetzt werden. Allerdings liegt der Verdacht nahe, dass Schüler, die nicht angegeben haben in welchem Schuljahr sie sind, auch sonst wenig brauchbare Angaben gemacht haben. Es wird daher auf Imputation gänzlich verzichtet. (Vgl. auch Ammermüller (2004, 4))

2.3 Schulebene

Schulen, die aufgrund dieser Missings über weniger als drei Schüler verfügen, wurden aus dem Schuldatensatz entfernt.⁶

Nicht sinnvoll imputiert werden kann die Variablen *anteil_mädchen*. Für den ebenfalls im Datensatz enthaltenen Index *scmatedu* könnte man jeden fehlenden Wert durch den Mittelwert des jeweiligen Staates ersetzen. Dies würde jedoch einigermaßen gleichwertig ausgestattete Schulen für die jeweiligen Staaten unterstellen. Um diese wenig realistische Annahme nicht halten zu müssen, wird auf eine Imputation auch für den Schuldatensatz gänzlich verzichtet.

Der gesamte Datensatz besteht nun noch aus 105036 Schülern in 5156 Schulen in 37 Ländern.

⁵Plünnecke (2003, 17) weist für die Berufstätigkeit des Vaters in allen drei Tests hohe Koeffizienten und hohe Signifikanzniveaus aus. Dies konnte mit den vorliegenden Daten nicht reproduziert werden. Vielmehr erwies sich die Berufstätigkeit des Vaters als durchgehend insignifikant. Daher wird auf diese Variable verzichtet.

⁶Dronkers und Robert (2003, 14) wählen diese Schwelle bei elf Schülern. Da die Autoren die Schule als zentralen Untersuchungsgegenstand gewählt haben und für alle Kontrollvariablen Missings durch Durchschnittswerte ersetzen ist diese größere Mindestanzahl von Schülern gerechtfertigt. Für die vorliegende Arbeit soll eine Mindestgröße von drei Schülern genügen.

2.4 PISA-Indizes

Neben den von Schülern, Eltern und Schulleitern erfragten Rohdaten enthalten die PISA-Datensätze auch eine Vielzahl von aus diesen Daten errechneten Indizes. Diese bieten sich in idealer Weise als Kontrollvariablen für die Mehrebenenanalyse an, da sie verdichtete Daten enthalten und somit die Anzahl der Variablen gering halten.

Die als Kontrollvariablen herangezogenen PISA-Indizes sind auf der Schülerebene *famedsup* und *hedres*. Auf Schulebene wird der Index *scmatedu* als Kontrollvariablen eingesetzt.

Beschrieben sind diese Indices bei OECD (2000, 31-33). Alle Indices sind auf einen Mittelwert von Null und eine Standardabweichung von 1 normiert.

3 Ergebnisse

In Anlehnung an Bryk und Raudenbush (1992, 201-203), Hox (2002, 15-21 u. 51-54) und Dronkers und Robert (2003) soll das Drei-Ebenen-Modell nach und nach aufgebaut werden. Dies erleichtert das Verständnis für die Bedeutung der einzelnen Ebenen.

3.1 Nullmodell

Es wird zunächst das Nullmodell (auch: unkonditioniertes oder leeres Modell) geschätzt. Dies ist gleichzusetzen mit dem – vielleicht zunächst etwas seltsam anmutenden – Versuch, die Schülerleistung ohne jede erklärende Variable zu erklären. Das in den Gleichungen 1 bis 3 dargestellte Nullmodell soll zunächst auch nur helfen, die Varianz des gesamten Datensatzes auf die drei untersuchten Ebenen zu verteilen (Bryk und Raudenbush (1992, 175-177), Ditton (1998, 160-162), Snijnders und Bosker (1999, 46f)). Mit dieser Ausgangsinformation lässt sich der Erklärungsgehalt der Variablen später besser beurteilen.

$$score = \pi_0 + e \quad (1)$$

$$\pi_0 = \beta_{00} + r_0 \quad (2)$$

$$\beta_{00} = \gamma_{000} + u_{00} \quad (3)$$

| | Lesen | Mathematik | Naturwissenschaft |
|------------------------------|--------------|--------------|-------------------|
| Varianz Staaten (u_{00}) | 2954 (26,8%) | 4138 (33,0%) | 2616 (23,9%) |
| Varianz Schulen (r_0) | 2895 (26,2%) | 2694 (21,5%) | 2496 (22,8%) |
| Varianz Schüler (e) | 5180 (47,0%) | 5723 (45,6%) | 5817 (53,2%) |
| Varianz gesamt | 11029 (100%) | 12555 (100%) | 10929 (100%) |
| Durchschnitt | 474,22 | 471,99 | 474,56 |
| Standardabweichung | 105,02 | 112,05 | 104,54 |
| Staaten | 37 | 37 | 37 |
| Schulen | 5156 | 5154 | 5148 |
| Schüler | 105036 | 58257 | 58508 |

Tabelle 3: Nullmodell

Baumert und Schümer (2001) zeigen anhand der Lesekompetenz, dass die in Tabelle 3 dargestellten Werte für einzelne Länder auch deutlich abweichen können. So ist die Varianz innerhalb der Schulen in Deutschland mit 39% deutlich geringer als bspw. in Norwegen (89%) und Schweden (90%). Um dem Rechnung zu tragen, soll im Weiteren für geeignete erklärende Variablen auch ihre Varianz im Modell berücksichtigt werden. Kreft (1996) gelangt in ihrem Literaturüberblick zu der Überzeugung, dass eine korrekte Schätzung der Varianzanteile in einem Zwei-Ebenen-Modell erst bei 300 Aggregateinheiten erfolgt – und nicht schon bei 100. Afshartous (1995, 11f) findet für 160 Aggregateinheiten noch stark verzerrte Ergebnisse bei der Varianzaufteilung, für 320 Einheiten sind die Ergebnisse deutlich besser. Für die vorliegende Arbeit wird daher auf eine Zuordnung der verbleibenden Varianz zu den einzelnen Ebenen über das Nullmodell hinaus verzichtet, da diese Aufteilung kaum realistisch ausfallen dürfte.

Der Mittelwert der Schüler liegt für alle drei Leistungstests bei etwa 475 Punkten, also etwa 5% niedriger als die 500 Punkte auf die die Leistungen ursprünglich normiert worden waren. Bedingt ist dieser Unterschied durch das Entfernen von Missings bzw. durch das Nichtverwenden der Gewichte.

Die Standardabweichung liegt mit 104,54 bis 112,05 dagegen über der der Gesamtstichprobe von 100.

Die Varianz der Leseergebnisse verteilt sich etwa zur Hälfte auf die Schülerebene und zu jeweils einem Viertel auf die Schul- und die Staatenebene. Staaten und auch Schulen innerhalb eines Staates sind sich also ähnlicher als Schüler. Bei den Mathematikleistungen sind sich die Schulen etwas ähnlicher, die Staaten etwas heterogener. Die Ergebnisse des naturwissenschaftlichen Tests zeigen auf Schülerebene die größte (absolute und relative)

Varianz.

Wößmann (2003, 154) findet für Mathematik und Naturwissenschaften mit TIMSS-Daten ähnliche Ergebnisse: 25,1% und 16,2% auf der Staatenebene, 23,5% und 21,8% auf der Schulebene und 51,4% und 67,0% auf der Schülerebene. In der Varianz auf Schulebene sind bei Wößmann 8,1% und 8,5% Varianz auf Klassenebene enthalten.

3.2 Schülermodell

Die Tabellen 4 und 5 stellen ein einfaches Modell dar, in dem die PISA-Leistungen allein über Variablen der Schülerebene erklärt werden sollen. Neben den dargestellten Variablen sind auch die Kontrollvariablen⁷ aus Abschnitt 2.2 im Modell enthalten.

$$\begin{aligned} score = \pi_0 + \pi_2(\textit{weiblich}) + \pi_4(\textit{m_inländer}) + \pi_5(\textit{v_inländer}) \\ + \pi_{11}(\textit{m_beruf}) + \pi_{14}(\textit{m_abitur}) + \pi_{15}(\textit{v_abitur}) \\ + \pi_{16}(\textit{m_allein}) + \pi_{17}(\textit{v_allein}) + e \end{aligned} \quad (4)$$

$$\pi_2 = \beta_{20} \quad (5)$$

$$\pi_4 = \beta_{40} \quad (6)$$

...

$$\beta_{20} = \gamma_{200} + u_{20} \quad (7)$$

$$\beta_{40} = \gamma_{400} + u_{40} \quad (8)$$

...

Die Gleichungen 4 bis 8 stellen zusammen mit den Gleichungen 2 und 3 einen Teil des einfachen Schülermodells dar. Gleichung 14 auf Seite 30 fasst das Modell zusammen.

⁷Alle Variablen in Tabelle 4 sind Dummy-Variablen. *** steht für Signifikanz auf dem 1%-Niveau, ** für das 5%-Niveau und * für das 10%-Niveau.

| | Lesen | Mathematik | Naturwissenschaft |
|-------------------|-----------|------------|-------------------|
| | 426,27*** | 450,25*** | 436,35*** |
| weiblich | 19,50*** | -18,29*** | -8,19*** |
| m_inländer | 6,13** | 5,42** | 6,26** |
| v_inländer | 7,43*** | 7,39*** | 9,09*** |
| m_beruf | 2,05** | 2,19** | 2,19* |
| m_abitur | 13,07*** | 14,53*** | 15,60*** |
| v_abitur | 11,41*** | 12,72*** | 12,70*** |
| m_allein | 1,02 | -0,86 | -0,53 |
| v_allein | -7,08*** | -7,20** | -7,47** |
| Erklärte Varianz: | 25,3% | 25,7% | 17,9% |

Tabelle 4: Schülermodell

| | Lesen | Mathematik | Naturwissenschaft |
|------------|-----------|------------|-------------------|
| weiblich | 118,35*** | 59,20*** | 65,57*** |
| m_inländer | 132,66*** | 115,74*** | 159,27*** |
| v_inländer | 63,20*** | 69,77*** | 141,67*** |
| m_abitur | 32,37*** | 39,09*** | 44,84*** |
| v_abitur | 24,27*** | 41,70*** | 46,00*** |
| m_allein | 10,35 | 18,83 | 16,16 |
| v_allein | 45,09 | 54,60 | 182,04** |

Tabelle 5: Schülermodell, Varianz einzelner Variabler

Mädchen haben im Alter von 15 Jahren deutlich bessere Leseergebnisse und deutlich schlechtere Ergebnisse im Mathematik- und im naturwissenschaftlichen Test als Jungen. Dies deckt sich mit anderen Untersuchungen.⁸

Migrantenkinder schneiden signifikant schlechter ab als Kinder, deren Eltern im Inland geboren wurden. Ob der Vater im Inland geboren ist hat dabei für alle drei Teiltests eine größere Bedeutung als das Geburtsland der Mutter.

Was die Bedeutung der Erwerbstätigkeit der Mutter angeht, so scheinen die Ergebnisse eindeutig. Kinder berufstätiger Mütter haben in allen drei Tests einen signifikanten Vorsprung von zwei Punkten. Es steht jedoch durchaus zu vermuten, dass diese Berufstätigkeit mit anderen, für die Kinder ebenfalls vorteilhaften, Eigenschaften einhergeht.

Bei der Unterscheidung nach dem höchsten Schulabschluss der Eltern ist es die Mutter, die einen etwas stärkeren Einfluss auf die Leistung des Kindes ausübt. Da die Abiturquote unter den Eltern der untersuchten Schüler für Mütter (48%) und Väter (47%) nahezu gleich ist, kann dieser Unterschied als durchaus bedeutend angesehen werden.

Kinder alleinerziehender Mütter haben keine signifikant abweichenden Testergebnisse, Kinder alleinerziehender Väter schneiden hingegen in allen drei Tests deutlich schlechter ab als Kinder, die mit beiden leiblichen Eltern zusammenleben.⁹

Tabelle 5 zeigt darüber hinaus, dass die Koeffizienten der meisten Variablen von Land zu Land durchaus unterschiedlich sind. Lediglich der Einfluss allein erziehender Eltern scheint zwischen den Staaten nicht signifikant zu variieren. Eine Ausnahme bildet jedoch der Einfluss allein erziehender Väter auf die naturwissenschaftlichen Leistungen ihrer Kinder.

3.3 Schülermodell mit Interaktionseffekten

Die Tabellen 6 und 7 stellen die Ergebnisse aus Tabelle 4 erweitert um Interaktionseffekte zwischen Charakteristika der Eltern und dem Geschlecht des Schülers dar.

Ob Vater bzw. Mutter im Inland geboren sind, ist für Mädchen und Jungen gleichermaßen von Bedeutung. Ein entsprechender Interaktionseffekt wurde nicht in das Modell

⁸Plünnecke (2003, 18), Dronkers und Robert (2003, 53 bzw. 57), Wößmann (2003, 129-132), Fuchs und Wößmann (2004, 13). Vgl. auch Stanat und Kunter (2001), sowie zur Begründung und Herleitung Eliot (2001, ???Kap. 16).

⁹Der Datensatz erlaubt eine Unterscheidung nach alleinerziehenden Mütter, alleinerziehenden Vätern, Kernfamilien, und gemischten (ein Elternteil neu lieert) bzw. sonstigen Familien. Basiskategorie für die Regression ist die Kernfamilie.

| | Lesen | Mathematik | Naturwissenschaft |
|-------------------|-----------|------------|-------------------|
| | 426,40*** | 449,92*** | 436,06*** |
| weiblich | 19,18*** | -17,71*** | -7,69** |
| m_inländer | 6,17*** | 5,44** | 6,29** |
| v_inländer | 7,41** | 7,42*** | 9,12*** |
| m_beruf | 2,04** | 2,19** | 2,17* |
| m_abitur | 11,46*** | 13,06*** | 13,96*** |
| ... * weiblich | 3,16*** | 2,84** | 3,21* |
| v_abitur | 12,53*** | 14,79*** | 14,51*** |
| ... * weiblich | -2,15 | -3,99** | -3,49* |
| m_allein | 3,17*** | -0,27 | 2,95 |
| ... * weiblich | -4,04** | -1,08 | -6,51** |
| v_allein | -9,32** | -8,23* | -7,66 |
| ... * weiblich | 5,54 | 2,60 | 0,80 |
| Erklärte Varianz: | 25,2% | 25,7% | 17,9% |

Tabelle 6: Schülermodell mit Interaktionseffekten

| | Lesen | Mathematik | Naturwissenschaft |
|------------|-----------|------------|-------------------|
| weiblich | 117,25*** | 58,19*** | 65,69*** |
| m_inländer | 133,63*** | 115,83*** | 167,44*** |
| v_inländer | 63,15*** | 68,73*** | 173,58*** |
| m_abitur | 32,59*** | 39,44*** | 42,30*** |
| v_abitur | 24,20*** | 41,92*** | 49,29*** |
| m_allein | 10,71 | 18,75 | 16,52 |
| v_allein | 44,26 | 54,34 | 182,23* |

Tabelle 7: Schülermodell mit Interaktionseffekten, Varianz einzelner Variabler

aufgenommen, da dieser sich in Voruntersuchungen als stets insignifikant herausstellte. Gleiches gilt für die Berufstätigkeit der Mutter.

Interessant sind die Interaktionseffekte in Bezug auf die Schulbildung der Eltern. Es zeigt sich fast durchgängig, dass die Bildung der Mutter stärker auf Töchter als auf Söhne durchschlägt, und dass die Bildung des Vaters (eine Ausnahme bildet das Leseverständnis) im Gegenzug stärker auf Söhne als auf Töchter wirkt. Die Unterschiede betragen jeweils zwischen drei und vier Punkte.

Auch was die Familienstruktur angeht zeigen sich Erkenntnisgewinne aus der Einführung der Interaktionsterme. War zuvor nur eine Wirkung allein erziehender Väter auf ihre Kinder nachzuweisen, so zeigt sich nun, dass allein erziehende Mütter positiv (+3,17) auf das Leseergebnis ihrer Söhne und negativ (-0,87 bzw. -6,51) auf das Mathematik- und das naturwissenschaftliche Testergebnis ihrer Töchter wirken. Allein erziehende Väter hingegen wirken negativ auf das Lese- und das Mathematikergebnis ihrer Kinder, ohne dass ein Unterschied für Mädchen und Jungen nachweisbar wäre.

3.4 Schulmodell

Auf Schulebene steht als genderspezifische erklärende Variable nur der Anteil der Mädchen¹⁰ an der gesamten Schülerzahl (zentriert um den Mittelwert der Stichprobe)¹¹ zur Verfügung. Im Rahmen einer Mehrebenenanalyse sollen die Schulcharakteristika dazu verwendet werden, die Koeffizienten des Schülermodells zu erklären. Entsprechend wird abweichend von Gleichung 2 nun Gleichung 9 formuliert, die das Absolutglied der Schülerebene π_0 erklären soll. Das Ergebnis der Schätzungen stellt Tabelle 8 dar.

$$\begin{aligned} \pi_0 = & \beta_{00} + \dots + \beta_{01}(\text{anteil_mädchen}) + \beta_{02}(\text{anteil_mädchen})^2 \\ & + \beta_{03}(\text{anteil_mädchen})^3 + \beta_{04}(\text{anteil_mädchen})^4 \\ & + \beta_{05}(\text{anteil_mädchen})^5 + r_0 \end{aligned} \quad (9)$$

Den Einfluss des Mädchenanteils stellt Abbildung 3 dar. Die Schätzung mit elf Quantilen aus Abbildung 2 konnte also reproduziert werden.

¹⁰Eine Regression mit 11 Quantil-Dummies findet sich in Abbildung 2. Diese Analyse legte es nahe, die Variable *anteil_mädchen* auch bis zur fünften Potenz aufzunehmen.

¹¹Die Daten wurden um den Mittelwert der gesamten Schülerschaft zentriert. Eine Zentrierung um die jeweiligen Mittelwerte innerhalb der Schulen hätte die Interpretation des Modells stark verkompliziert. Zu den Problemen des Zentrierens um Gruppenmittelwerte siehe Hox (2002, 62f).

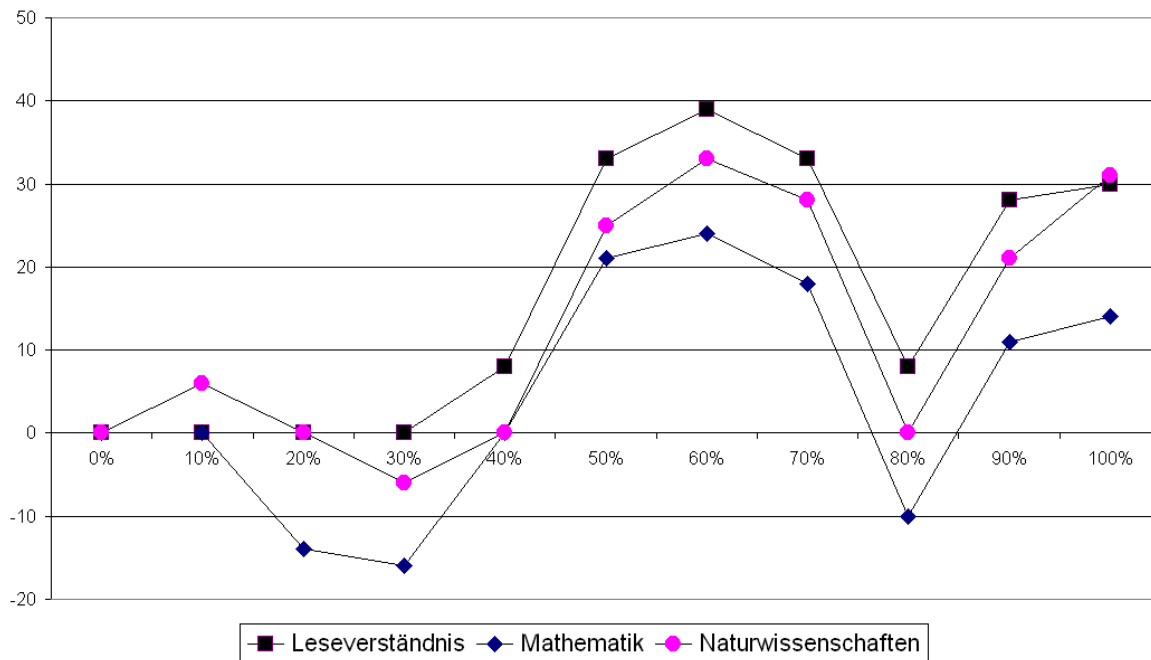


Abbildung 2: Mädchenanteil und Testergebnis

| | Lesen | Mathematik | Naturwissenschaft |
|-----------------------------|--------------|--------------|-------------------|
| | 388,33*** | 424,17*** | 399,74*** |
| anteil_mädchen | 531,98*** | 379,64* | 501,06** |
| anteil_mädchen ² | -4518,22*** | -3736,53*** | -4298,24*** |
| anteil_mädchen ³ | 13968,85*** | 12190,41*** | 13515,62*** |
| anteil_mädchen ⁴ | -17013,48*** | -15241,33*** | -16752,61*** |
| anteil_mädchen ⁵ | 7083,86*** | 6441,86*** | 7087,85*** |
| weiblich | 19,87*** | -17,50*** | -7,39** |
| m_inländer | 5,69*** | 5,32** | 6,46** |
| v_inländer | 7,53*** | 7,79*** | 8,81*** |
| m_beruf | 1,86** | 1,97* | 1,96 |
| m_abitur | 12,41*** | 13,56*** | 14,89*** |
| ... * weiblich | 3,15*** | 2,84** | 2,92* |
| v_abitur | 13,36*** | 15,05*** | 15,13*** |
| ... * weiblich | -1,86 | -3,79** | -2,85 |
| m_allein | 3,44*** | -0,02 | 2,86 |
| ... * weiblich | -4,15** | -1,67 | -6,10* |
| v_allein | -9,15*** | -8,29** | -6,75 |
| ... * weiblich | 4,00 | 1,40 | -1,89 |
| Erklärte Varianz: | 32,1% | 31,8% | 26,4% |

Tabelle 8: Schulmodell

| | Lesen | Mathematik | Naturwissenschaft |
|------------|-----------|------------|-------------------|
| weiblich | 126,25*** | 67,19*** | 77,74*** |
| m_inländer | 124,63*** | 117,02*** | 150,11*** |
| v_inländer | 64,83*** | 65,06*** | 157,85*** |
| m_abitur | 35,49** | 40,39*** | 50,19*** |
| v_abitur | 25,37*** | 40,51** | 47,61** |
| m_allein | 9,82 | 19,62 | 17,65 |
| v_allein | 40,85 | 45,68 | 146,02** |

Tabelle 9: Schulmodell, Varianz einzelner Variabler

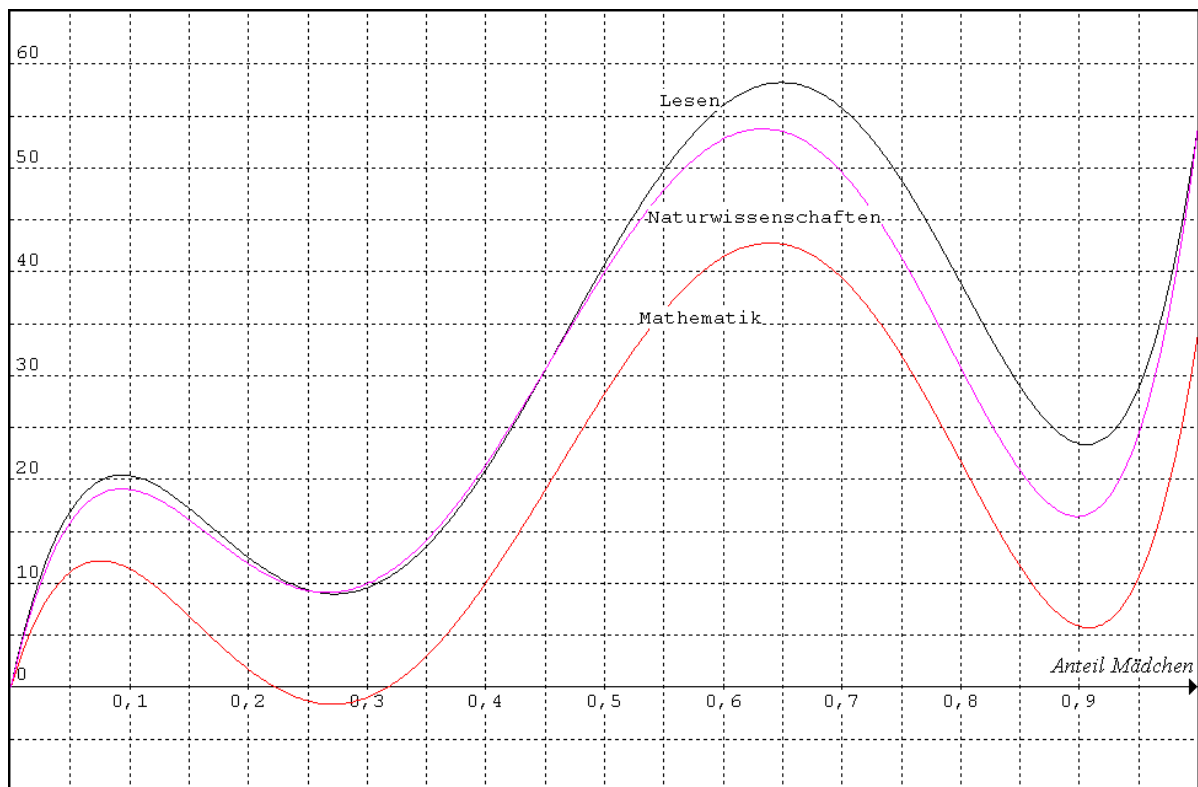


Abbildung 3: Geschätzter Einfluss des Mädchenanteils im Schulmodell

Für alle drei Tests zeigt sich, dass ein Anteil von etwa 65% Mädchen an einer Schule mit den höchsten Testergebnissen der einzelnen Schüler einhergeht. Reine Jungenschulen schneiden deutlich schlechter (43 bis 58 Punkte) ab, für reine Mädchenschulen besteht in Mathematik und beim Leseverständnis ein Nachteil (9 und 4 Punkte).

Die übrigen Werte haben sich höchstens geringfügig geändert. Zwei Koeffizienten im naturwissenschaftlichen Test (*m_beruf* und *v_abitur * weiblich*) erreichen das 10%-Signifikanzniveau dadurch nicht mehr.

Die erklärte Varianz konnte durch die zusätzlichen Variablen (die neuen Kontrollvariablen *scmatedu* und *privatschule* sind signifikant) auf 26,4 bis 32,1% erhöht werden.

3.5 Dreiebenenmodell

Der Global-Gender-Gap-Index kann Werte zwischen Null und Eins annehmen. Da eine Veränderung um einen Punkt also realistischer Weise nicht möglich ist, sind die Schätzungen jeweils für Hundertstel Indexpunkte angegeben. Dem Mittelwert am nächsten sind auf den Platz 25 Jamaika mit 0,7014 Punkten und Platz 26 die Schweiz mit 0,6997 Punkte). Eine Verbesserung um 0,01 Punkte führt zu Lettland (Platz 19, 0,7091), eine gleich große Verschlechterung führt zu Platz 34 (Botswana mit 0,6897 Punkten) von insgesamt 115 Plätzen. Japan (Platz 79 mit 0,6447 Punkten) ist das Land mit der schlechtesten GGG-Platzierung, das im Datensatz enthalten ist.

Die einzige Kontrollvariable auf Staatenebene ist das Bruttonationaleinkommen pro Kopf.

Die Gleichungen 10 bis 12 werden zusätzlich in das Modell aufgenommen.¹² Gleichung 10 beschreibt allgemein den Einfluss auf Staatenebene, Gleichung 11 prüft, ob der Einfluss auf Mädchen und Jungen unterschiedlich ist und mit Gleichung 12 soll ermittelt werden, ob der starke Einfluss des Abiturs der Mutter auf die Leistung der Tochter mit steigender Gleichberechtigung der Frauen ab- oder zunimmt.

$$\beta_{00} = \gamma_{000} + \gamma_{002}(\text{gggi_score}) + u_{00} \quad (10)$$

¹²Dass die Variable *gggi_score* nicht auch quadriert aufgenommen wird ist wieder der Kapazität der verwendeten Methode und der Software geschuldet. Um nicht einen linearen Zusammenhang zu unterstellen wo keiner ist, muss die Interpretation der Ergebnisse entsprechend vorsichtig erfolgen.

| | Lesen | Mathematik | Naturwissenschaft |
|-----------------------------|--------------|--------------|-------------------|
| | 386,77*** | 423,09*** | 398,76*** |
| gggi_score | 3,31** | 2,53 | 2,21* |
| anteil_mädchen | 533,61*** | 381,09* | 501,71*** |
| anteil_mädchen ² | -4539,55*** | -3754,92*** | -4308,95*** |
| anteil_mädchen ³ | 14038,78*** | 12245,90*** | 13540,58*** |
| anteil_mädchen ⁴ | -17093,44*** | -15299,36*** | -16767,93*** |
| anteil_mädchen ⁵ | 7113,75*** | 6461,14*** | 7087,48*** |
| weiblich | 20,00*** | -17,42*** | -7,22*** |
| ... * gggi_score | 1,42*** | 1,05*** | 0,90*** |
| m_inländer | 7,24*** | 6,25** | 6,94** |
| v_inländer | 7,73*** | 8,35*** | 9,64*** |
| m_beruf | 1,79** | 1,94* | 1,98 |
| m_abitur | 12,37*** | 13,35*** | 14,80*** |
| ... * gggi_score | 0,60*** | 0,67** | 0,69** |
| m_abitur * weiblich | 3,24*** | 2,94** | 2,65 |
| ... * gggi_score | 0,12 | -0,05 | 0,45 |
| v_abitur | 13,43*** | 15,09*** | 15,21*** |
| ... * weiblich | -1,87 | -3,84** | -2,88 |
| m_alleinerz | 3,18** | -0,16 | 3,00 |
| ... * weiblich | -4,18** | -1,74 | -6,23* |
| v_alleinerz | -9,28*** | -8,79** | -7,65 |
| ... * weiblich | 3,72 | 0,86 | -1,53 |
| Erklärte Varianz: | 31,7% | 33,9% | 25,3% |

Tabelle 10: Dreiebenenmodell

| | Lesen | Mathematik | Naturwissenschaft |
|------------|-----------|------------|-------------------|
| weiblich | 65,08*** | 45,42*** | 59,34*** |
| m_inländer | 122,54*** | 133,84*** | 156,00*** |
| v_inländer | 63,58*** | 84,65*** | 158,97*** |
| m_abitur | 23,23*** | 35,72*** | 38,43*** |
| v_abitur | 25,28*** | 40,51*** | 46,98*** |
| m_allein | 11,02 | 20,75 | 19,76 |
| v_allein | 56,82 | 44,80 | 166,26* |

Tabelle 11: Dreiebenenmodell, Varianz einzelner Variabler

$$\beta_{20} = \gamma_{200} + \gamma_{202}(gggi_score) + u_{20} \quad (11)$$

$$\beta_{140} = \gamma_{1400} + \gamma_{1401}(gggi_score) \quad (12)$$

Die Auswirkungen einer stärkeren Gleichberechtigung sind für Mädchen und Jungen unterschiedlich. Mädchen profitieren in allen drei Tests stärker davon als Jungen, am deutlichsten fällt der Vorsprung beim Leseverständnis aus. Für Jungen ist der Einfluss auf das Mathematikergebnis statistisch nicht signifikant.

Der positive Einfluss des Abiturs der Mutter erhöht sich durch eine Verbesserung der Gleichberechtigung und zwar recht ähnlich für die drei Test (0,60 bis 0,69 Punkte). Eine stärkere Auswirkung auf Töchter als auf Söhne ist jedoch nicht nachweisbar.

Der stärkere Einfluss des Abiturs der Mutter auf ihre Tochter wird für den naturwissenschaftlichen Test insignifikant.

Weiter lässt sich festhalten, dass der Vorsprung inländischer Eltern weiter zunimmt. Die Zunahme fällt für den inländischen Vater mit bis zu 1,55 Punkte (Leseverständnis) stärker aus als für die Mutter (maximal 0,83 Punkte, naturwissenschaftlicher Test).

4 Folgerungen

Mädchen und Jungen erzielen bei PISA 2000 deutlich unterschiedliche Ergebnisse. Darüber hinaus beeinflussen Mütter und Väter ihre Kinder nicht in gleichem Maße.

Die unterschiedlichen Kompetenzen von 15-jährigen Mädchen und Jungen im Leseverständnis bzw. in Mathematik konnten im Einklang mit ähnlichen Untersuchungen identifiziert werden.

Gleichfalls wird deutlich, dass Kinder von im Inland geborenen Eltern Vorteile in allen drei Tests haben. Hier fällt zum einen auf, dass der Einfluss der Mutter sich für die drei Teilaufgaben kaum unterscheidet und zum anderen, dass der Einfluss des Vaters durchgängig größer ist als der der Mutter. Gleichzeitig unterschieden sich die Koeffizienten signifikant zwischen den einzelnen Staaten (Tabelle 11). Die schulischen Nachteile von Migrantenkinder können demnach wahrscheinlich durch unterschiedliche Politikmaßnahmen beeinflusst werden. Hier wäre ein Ansatz für weitere Forschung im Bereich der Familienpolitik.

Die Berufstätigkeit der Mutter wirkt sich auf die schulischen Leistung nicht negativ auf. Allerdings gehen mit der Berufstätigkeit eines Elternteils wohl auch andere, ebenfalls für die Kinder vorteilhafte, Charakteristika einher. Daher lässt sich aus den signifikant positiven Koeffizienten bei Leseverständnis und Mathematik keine unmittelbare Politikimplikation ableiten.

Weiter haben die Bildungsabschlüsse der Eltern grundsätzlich eine deutliche Wirkung auf die Schulleistungen der Kinder. Das Abitur der Mütter ist dabei für Töchter wichtiger als das der Väter, welches wiederum für die Söhne wichtiger ist als das der Mütter. Gleichzeitig unterliegt der Einfluss des Vaters einer größeren Varianz zwischen den 37 untersuchten Staaten. Dies wird als Hinweis darauf gewertet, dass durch unterschiedliche bildungspolitische Rahmenbedingungen der (positive wie negative) Einfluss des familiären Hintergrunds unterschiedlich stark ausgeglichen werden kann.

Bildung und Berufstätigkeit von Frauen setzt sich also unmittelbar in der Schulleistung der nächsten Generation fort und kann durchaus als Motor für eine bessere Performance eines Landes gesehen werden.

Ob Kinder bei allein erziehenden Elternteilen aufwachsen beeinflusst ihre schulischen Leistungen ebenfalls. Ein allein erziehender Vater geht jeweils mit signifikant schlechteren Leistungen bei Leseverständnis und mathematischem Test einher. Eine allein erziehende Mutter hingegen hat nur für Töchter negative Effekte bei Leseverständnis und naturwissenschaftlichem Test, für Söhne besteht dagegen ein Zusammenhang mit besseren Leistungen im Leseverständnis. Die Koeffizienten variieren zwischen den Staaten auch nicht signifikant.

Die Zusammensetzung von Schulen hat ebenfalls einen erheblichen Einfluss auf die Leistungen von Mädchen und Jungen. Der *optimale* Mädchenanteil liegt bei etwa 65%. Diesen Wert für alle Schulen zu fordern muss jedoch an genetischen Konstanten scheitern. Was jedoch festgehalten werden kann, ist, dass reine Jungenschulen wohl ein deutliche schlechteres Lernumfeld bieten als gemischte Schulen oder reine Mädchenschulen (Abbildung 3). Der Einfluss von Privatschulen ist in der Untersuchung bereits kontrolliert, weitere mögliche Besonderheiten solcher Schulformen stehen im Datensatz nicht zur Verfügung.

Auch über die bereits erwähnten Variablen wie Berufstätigkeit und höhere Bildung von Frauen hinaus hat das Ausmaß der Gleichberechtigung in einem Land signifikant positiven Einfluss auf die Schulleistung sowohl von Mädchen als auch – etwas geringer – von Jungen. Der Einfluss ist für beide Gruppen in Mathematik (für Jungen insignifikant) am

niedrigsten und beim Leseverständnis am höchsten. Darüber hinaus verstärkt sich der oben erwähnte günstige Einfluss gebildeter Mütter auf ihre Kinder mit fortschreitender Gleichberechtigung.

Größere Anstrengungen zur Gleichberechtigung von Frauen, insbesondere bei Bildungsabschlüssen und bei der Berufstätigkeit können also durchaus ein Schlüssel für zukünftigen Bildungserfolg sein. Dies gilt insbesondere für die Länder links in Abbildung 1 soweit sie bei PISA 2000 speziell beim Leseverständnis unterdurchschnittlich abgeschnitten hatten: Chile, Italien, Mexiko, Griechenland, Indonesien, Brasilien, Albanien und Peru.

Literatur

- Acemoglu, D. und Angrist, J. (1999). How large are the social returns to education? Evidence from compulsory schooling laws. NBER Working Paper 7444.
- Afshartous, D. (1995). Determination of sample size for multilevel model design. AREA Meeting, San Francisco, April 1995.
- Ammermüller, A. (2004). PISA: What makes the difference? explaining the gap in PISA test scores between finland and germany. Diskussionspapier 04-04, ZEW.
- Barro, R. J. (2001). Human capital and growth. *American Economic Review, Papers and Proceedings*, 91:12–17.
- Baumert, J. und Schümer, G. (2001). Schulformen als selektionsbedingte Lernmilieus. In: Deutsches PISA-Konsortium, (Hrsg.), *Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*, S. 454–467. Opladen.
- Becker, G. S. und Lewis, H. G. (1973). On the interaction between quantity and quality of children. *Journal of Political Economy*, 82:S279–S288.
- Behrmann, J. R. und Rosenzweig, M. K. (2002). Does increasing women's schooling raise the schooling of the next generation? *American Economic Review*, 92:323–334.
- Behrmann, J. R. und Taubmann, P. (1989). Is schooling 'mostly in the genes'? nature-nurture decomposition using data on relatives. *Journal of Political Economy*, 97:1425–1446.
- Black, S. E., Devereux, P. J., und Salvanes, K. G. (2005). Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *American Economic Review*, 95:437–449.
- Blien, U. und Wiedenbeck, M. (2002). Mehrebenenanalyse. In: Kleinhenz, G. D., (Hrsg.), *IAB-Kompodium Arbeitsmarkt- und Berufsforschung. Beiträge zur Arbeitsmarkt- und Berufsforschung, BeitrAB 250*, S. 309–324. Nürnberg.
- Bohrhardt, R. (2000). Familienstruktur und Bildungserfolg. Stimmen die alten Bilder? *Zeitschrift für Erziehungswissenschaft*, 3:189–207.

- Browne, W. J. und Draper, D. (2000). Implementation and performance issues in the bayesian and likelihood fitting of multilevel models. *Computational Statistics*, 15:391–420.
- Bryk, A. S. und Raudenbush, S. W. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Parl, London, Neu Delhi.
- Cronbach, L. J. (1976). *Research on Classrooms and Schools: Formulation of Questions, Design and Analysis*. Stanford Evaluation Consortium.
- de Haan, M. und Plug, E. (2006). Estimates of the effect of parents' schooling on children's schooling using censored and uncensored samples. Diskussionspapier 2416, IZA.
- Ditton, H. (1998). *Mehrebenenanalyse*. Weinheim, München.
- Dronkers, J. und Robert, P. (2003). The effectiveness of public and private schools from a comparative perspective. Diskussionspapier, European University Institute.
- Eliot, L. (2001). *Was geht da drinnen vor? Die Gehirnentwicklung in den ersten fünf Lebensjahren*. Berlin.
- Fuchs, T. und Wößmann, L. (2004). What accounts for international differences in student performance? a re-examination using PISA data. Diskussionspapier 1287, IZA.
- Goldberger, V. (1991). *A Course in Econometrics*. Cambridge.
- Gujarati, D. N. (1995). *Basic Econometrics*. New York u.a., 3. Auflage.
- Gundlach, E. und Wößmann, L. (2004). Family background, schooling resources, and institutional features: What determines student performance in east asian countries? *East Asian Economic Perspectives*, 15:17–31.
- Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, 24:1141–1177.
- Hanushek, E. A. und Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American Economic Review*, 90:1184–1208.

- Hox, J. (2002). *Multilevel Analysis. Techniques and Applications*. Mahwah.
- Hummell, H. J. (1972). *Probleme der Mehrebenenanalyse*. Stuttgart.
- Kreft, I. G. G. (1996). Are multilevel techniques necessary? an overview, including simulation studies. Diskussionspapier, California State University, <http://www.calstatela.edu/faculty/ikreft/quarterly/quarterly.html>.
- Kreft, I. G. G. und de Leeuw, E. (1988). The see-saw effect: A multilevel problem? *Quality Quantity*, 22:127–137.
- Maas, C. J. M. und Hox, J. J. (2002). Accuracy of multilevel parameter estimates when samples are small. In: Blasius, J., Hox, J., de Leeuw, E., und Schmidt, P., (Hrsgg.), *Social Science Methodology in the New Millennium. Proceedings of the Fifth International Conference on Logic and Methodology*. Opladen.
- Mok, M. (1995). Sample size requirements for 2-level designs in educational research. *Multilevel Modelling Newsletter*, 7:11–15.
- OECD, (Hrsg.) (2000). *Manual for the PISA 2000 Database*. Paris.
- OECD, (Hrsg.) (2005). *Bildung auf einen Blick. OECD-Indikatoren*. Paris.
- Plomin, R. (1990). *Nature and Nurture. An Introduction to Human Behavioral Genetics*. Pacific Grove.
- Plünnecke, A. (2003). Bildungsreform in Deutschland. Nummer 4 in iw Positionen. Beiträge zur Ordnungspolitik, Köln.
- Raudenbush, S., Bryk, A., Cheung, Y. F., Congdon, R., und Du Toit, M. (2004). *HLM 6: Hierarchical Linear and Nonlinear Modeling*. Lincolnwood.
- Salomone, R. C. (2003). *Same, Different, Equal. Rethinking Single-Sex Schooling*. New Haven, London.
- Sellin, N. (1990). On aggregation bias. In: Cheung, K. C., Keeves, J. P., Sellin, N., und Tsoi, S., (Hrsgg.), *The Analysis of Multilevel Data in Educational Research: Studies of Problems and their Solutions*, volume 14, S. 257–268.

- Snijnders, T. A. B. und Bosker, R. J. (1999). *Multilevel Analysis, An Introduction to Basic and Advanced Multilevel Modeling*. London, Thousand Oaks, Neu Delhi.
- Stanat, P. und Kunter, M. (2001). Geschlechterunterschiede in Basiskompetenzen. In: Deutsches PISA-Konsortium, (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*, S. 249–269. Opladen.
- Todd, P. E. und Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *Economic Journal*, 113:F3–F33.
- UNESCO, (Hrsg.) (2003). *Gender and Education for All. The Leap to Equality*. Paris.
- Vigdor, J. und Nechyba, T. (2007). Peer effects in north carolina public schools. In: Wößmann, L. und Peterson, P. E., (Hrsgg.), *Schools and the Equal Opportunity Problem*. Cambridge.
- Wolter, S. C. und Weber, B. A. (1999). On the measurement of private rates of return to education. Ein Ansatz zur Messung privater Bildungsrenditen. *Jahrbücher für Nationalökonomie und Statistik*, 218:605–618.
- World Economic Forum, (Hrsg.) (2006). *The Global Gender Gap Report*. Genf.
- Wößmann, L. (2003). Schooling resources, educational institutions and student performance: The international evidence. *Oxford Bulletin of Economics and Statistics*, 65:117–170.
- Wößmann, L. (2004). How equal are educational opportunities? Family background and student achievement in europe and the united states. Diskussionspapier 1162, CESifo.

Anhang

Nullmodell

$$score = \gamma_{000} + u_{00} + r_0 + e \quad (13)$$

Schülermodell

$$\begin{aligned} score = & \gamma_{000} + \gamma_{200}(weiblich) + \gamma_{400}(m_inländer) \\ & + \gamma_{500}(v_inländer) + \gamma_{1100}(m_beruf) + \gamma_{1300}(m_abitur) \\ & + \gamma_{1500}(v_abitur) + \gamma_{1700}(m_allein) + \gamma_{1800}(v_allein) \\ & + u_{20}(weiblich) + u_{40}(m_inländer) + u_{50}(v_inländer) \\ & + u_{130}(m_abitur) + u_{150}(v_abitur) + u_{170}(m_allein) + u_{180}(v_allein) \\ & + u_{00} + r_0 + e \end{aligned} \quad (14)$$

Schülermodell mit Interaktionstermen

$$\begin{aligned} score = & \gamma_{000} + \gamma_{200}(weiblich) + \gamma_{400}(m_inländer) \\ & + \gamma_{500}(v_inländer) + \gamma_{1100}(m_beruf) + \gamma_{1300}(m_abitur) \\ & + \gamma_{1400}(m_abitur)(weiblich) + \gamma_{1500}(v_abitur) \\ & + \gamma_{1600}(v_abitur)(weiblich) + \gamma_{1700}(m_allein) + \gamma_{1800}(v_allein) \\ & + \gamma_{2000}(m_allein)(weiblich) + \gamma_{2100}(v_allein)(weiblich) \\ & + u_{20}(weiblich) + u_{40}(m_inländer) + u_{50}(v_inländer) \\ & + u_{140}(m_abitur) + u_{150}(v_abitur) + u_{160}(m_allein) + u_{170}(v_allein) \\ & + u_{00} + r_0 + e \end{aligned} \quad (15)$$

Schulmodell

$$\begin{aligned} score = & \gamma_{000} + \gamma_{010}(ant_mädchen) \\ & + \gamma_{020}(ant_mädchen)^2 + \gamma_{030}(ant_mädchen)^3 \\ & + \gamma_{040}(ant_mädchen)^4 + \gamma_{050}(ant_mädchen)^5 + \gamma_{200}(weiblich) \\ & + \gamma_{400}(m_inländer) + \gamma_{500}(v_inländer) + \gamma_{1100}(m_beruf) \\ & + \gamma_{1300}(m_abitur) + \gamma_{1400}(m_abitur)(weiblich) + \gamma_{1500}(v_abitur) \\ & + \gamma_{1600}(v_abitur)(weiblich) + \gamma_{1700}(m_allein) + \gamma_{1800}(v_allein) \\ & + \gamma_{2000}(m_allein)(weiblich) + \gamma_{2100}(v_allein)(weiblich) \\ & + u_{20}(weiblich) + u_{40}(m_inländer) + u_{50}(v_inländer) + u_{140}(m_abitur) \\ & + u_{150}(v_abitur) + u_{160}(m_allein) + u_{170}(v_allein) + u_{00} + r_0 + e \end{aligned} \quad (16)$$

Dreibenenmodell

$$\begin{aligned} score = & \gamma_{000} + \gamma_{002}(gghi_score) + \gamma_{010}(ant_mädchen) \\ & + \gamma_{020}(ant_mädchen)^2 + \gamma_{030}(ant_mädchen)^3 \\ & + \gamma_{040}(ant_mädchen)^4 + \gamma_{050}(ant_mädchen)^5 \\ & + \gamma_{200}(weiblich) + \gamma_{201}(gghi_score)(weiblich) \\ & + \gamma_{210}(ant_mädchen)(weiblich) + \gamma_{220}(ant_mädchen)^2(weiblich) \\ & + \gamma_{230}(ant_mädchen)^3(weiblich) + \gamma_{240}(ant_mädchen)^4(weiblich) \\ & + \gamma_{250}(ant_mädchen)^5(weiblich) + \gamma_{400}(m_inländer) \\ & + \gamma_{500}(v_inländer) + \gamma_{1100}(m_beruf) + \gamma_{1300}(m_abitur) \\ & + \gamma_{1301}(gghi_score)(m_abitur) + \gamma_{1400}(m_abitur)(weiblich) \\ & + \gamma_{1401}(gghi_score)(m_abitur)(weiblich) + \gamma_{1500}(v_abitur) \\ & + \gamma_{1600}(v_abitur)(weiblich) + \gamma_{1700}(m_allein) + \gamma_{1800}(v_allein) \\ & + \gamma_{2000}(m_allein)(weiblich) + \gamma_{2100}(v_allein)(weiblich) \\ & + u_{20}(weiblich) + u_{40}(m_inländer) + u_{50}(v_inländer) + u_{140}(m_abitur) \\ & + u_{150}(v_abitur) + u_{160}(m_allein) + u_{170}(v_allein) + u_{00} + r_0 + e \end{aligned} \quad (17)$$