

Tobias Linberg & Sebastian E. Wenz

## **Ausmaß und Verteilung sozioökonomischer und migrationspezifischer Ungleichheiten im Sprachstand fünfjähriger Kindergartenkinder**

### **Zusammenfassung**

*Ungleichheiten im sprachlichen Kompetenzstand können nicht nur in Abhängigkeit des jeweils betrachteten Indikators der sozioökonomischen und migrationspezifischen Herkunft unterschiedlich groß ausfallen, sondern auch an unterschiedlichen Stellen der Kompetenzverteilung mehr oder weniger stark ausgeprägt sein. Dieser Beitrag untersucht dazu drei Forschungsfragen anhand einer bundesweiten Stichprobe von etwa fünfjährigen Kindergartenkindern: (1) In welchem Ausmaß zeigen sich Ungleichheiten im Wortschatz von Kindergartenkindern entlang sozialer Klasse, Bildungshintergrund, Haushaltseinkommen und Migrationshintergrund? (2) Lassen sich bestimmte Muster von Ungleichheiten entlang der Kompetenzverteilung erkennen? (3) Inwiefern lassen sich etwaige Muster durch kumulative Effekte erklären? Diese Fragen werden mittels Quantilsregressionen anhand der Daten des Nationalen Bildungspanels (NEPS) analysiert. Die Ergebnisse zeigen, dass Ungleichheiten nach Bildungs- und Migrationshintergrund am größten ausfallen und sich vor allem auf den unteren Bereich der Leistungsverteilung konzentrieren. Dagegen unterscheidet sich der Wortschatz insbesondere der kompetentesten Kinder nur gering in Abhängigkeit der sozialen und migrationspezifischen Herkunft. Diese Muster bleiben auch unter Kontrolle von Kovariaten weitestgehend erhalten.*

### **Schlagwörter**

*Soziale Ungleichheit; sprachliche Kompetenzen; Kindergartenkinder; Quantilsregression; Nationales Bildungspanel (NEPS)*

---

Dr. Tobias Linberg (corresponding author), LIfBi – Leibniz-Institut für Bildungsverläufe e.V., Wilhelmsplatz 3, 96047 Bamberg, Deutschland  
E-Mail: tobias.linberg@lifbi.de

Sebastian E. Wenz, GESIS – Leibniz-Institut für Sozialwissenschaften, Unter Sachsenhausen 6–8, 50667 Köln, Deutschland  
E-Mail: sebastian.wenz@gesis.org

## **Extent and distribution of socio-economic and immigration-related inequalities in the language abilities of 5-year-old preschool children**

### **Abstract**

*Inequalities in language competences may be more or less pronounced depending not only on the social or immigration background indicator under consideration, but can also differ in size when looking at different positions of the competence distribution. Based on a nationwide sample of about 5-year-old preschool children this paper examines three research questions in this context: (1) To what extent are differences by social class, educational background, income, and immigrant background visible in children's vocabulary? (2) Do certain patterns of inequality exist along the distribution of competences? (3) To what extent can these patterns be explained by cumulative effects? These questions were investigated using quantile regression on data from the National Educational Panel Study (NEPS). The results show that differences are most pronounced with regard to parents' levels of formal education and their immigration background and concentrate in the lower areas of the competence distribution. In contrast, children of different socioeconomic and immigrant background differ only little in higher quantiles of language competence. These patterns remain largely intact when controlling for covariates.*

### **Keywords**

*Social inequalities; Language competences; Preschool children; Quantile regression; National Educational Panel Study (NEPS)*

## **1. Einleitung**

Die Beziehung von Herkunft mit Kompetenzen und Bildungserfolg ist seit langem Gegenstand empirischer Untersuchungen und es existiert eine Vielzahl von Studien, die belegen, dass Indikatoren des sozioökonomischen Status und des Migrationshintergrunds mit den sprachlichen Kompetenzen von Kindern im vorschulischen Alter assoziiert sind (z. B. Lee & Burckham, 2002; Niklas, Möllers & Schneider, 2013; Sylva, Melhuish, Sammons, Siraj-Blatchford & Taggart, 2008; Weinert & Ebert, 2013). Für Deutschland liegen bislang jedoch nur Untersuchungen mit regional begrenzten Stichproben vor, die zumeist die Beantwortung eingegrenzter Forschungsfragen – etwa die Einflüsse bestimmter Lernumwelten oder Mechanismen zur Erklärung von migrationsspezifischen Ungleichheiten – fokussieren (z. B. Niklas, Segerer, Schmiedeler & Schneider, 2012; Niklas et al., 2013; Weinert & Ebert, 2013). Ein vergleichend angelegter Überblick über soziale und migrationsspezifische Ungleichheiten in sprachlichen Kompetenzen von Kindern im Vorschulalter fehlt dagegen für Deutschland.

Zudem wendeten die genannten Studien in der Regel auf Mittelwerten basierende Auswertungsverfahren an (z. B. lineare Regressionen). Damit wird zumindest implizit angenommen, dass die Differenzen an allen Stellen der (konditionalen) Kompetenzverteilung in etwa gleich groß ausfallen bzw. bleibt es unentdeckt, wenn dies nicht der Fall ist.

Daher werden in der vorliegenden Studie soziale und migrationsspezifische Ungleichheiten im Wortschatz fünfjähriger Kindergartenkinder anhand einer bundesweit gezogenen Stichprobe untersucht und hierbei auf die in der deutschen Bildungsforschung bislang kaum verwendete Methode der Quantilsregression zurückgegriffen.

## **2. Hintergrund und Forschungsfragen**

Die Kindheit ist eine Phase intensiver Entwicklungen in kognitiven und nicht-kognitiven Bereichen. Die in der Kindheit erlernten Fähigkeiten sind für das weitere Lernen und den weiteren Bildungsweg sowie für gesellschaftliche Teilhabe von großer Bedeutung (Cunha, Heckman, Lochner & Masterov, 2005). Von besonderer Relevanz für Bildungsprozesse sind dabei sprachliche Kompetenzen, da Vermittlung und Abruf von Wissen im Wesentlichen sprachlich erfolgen. Darüber hinaus wurden Zusammenhänge nicht nur mit dem Leseerwerb und Leseverständnis gezeigt, sondern auch mit Gedächtnisleistungen und Gedächtnisentwicklung, so dass frühe Unterschiede in sprachlichen Kompetenzen mit Folgen für ganz unterschiedliche Entwicklungs- und Leistungsbereiche verbunden sein können (vgl. Weinert & Ebert, 2013). Die Entwicklung des Wortschatzes wiederum hängt wesentlich von sozioökonomischen Ressourcen und den in der häuslichen Lernumwelt erfahrenen sprachlichen Anregungen ab (Hart & Risley, 1995; Hoff, 2003; Niklas et al., 2013). Eine genauere Betrachtung dieser Zusammenhänge wird im Folgenden entlang theoretischer Bezüge und des Forschungsstands beschrieben.

### **2.1 Theoretische Bezüge und Forschungsstand**

#### **2.1.1 Theoretische Bezüge**

Die Beziehungen zwischen Kompetenzen und Kontextmerkmalen lassen sich mit verschiedenen Ansätzen theoretisieren. Ökologische Entwicklungstheorien heben unter anderem hervor, wie das angelegte Potential kindlicher Entwicklung durch Merkmale unmittelbarer Lernumwelten hervorgebracht und beeinflusst werden kann (z. B. Bronfenbrenner & Morris, 2006). Die Familie stellt hierbei die erste und wichtigste Lernumwelt dar (z. B. Roßbach, 2011), ihr Anregungsgehalt kann sich jedoch in Abhängigkeit der sozialen Lage der Familien unterscheiden. So verfügen Familien unterschiedlicher sozialer Lagen beispielsweise in unterschiedlich starkem Ausmaß über Ressourcen zur Ausgestaltung dieser Lernumwelten (Betz,

2008). Familiäre Investitionsmodelle gehen davon aus, dass Eltern mit größeren *finanziellen Ressourcen* durch unterschiedliche Unterstützungsmöglichkeiten in der Lage sind, substantiell in die Entwicklung ihrer Kinder zu investieren, während Familien mit geringem Einkommen gezwungen sind, ihre Ressourcen für die Befriedigung von unmittelbaren Bedürfnissen der Familie einzusetzen und zudem armutsbedingter Stress mit negativen Konsequenzen für familiäre Prozesse und das kindliche Verhalten assoziiert ist (z. B. Duncan & Magnuson, 2003). Mit dem erweiterten Investitionsmodell wird weiterhin angenommen, dass auch der *elterliche Bildungshintergrund* sowie deren berufliche Position – also die *soziale Klassenlage* – in vergleichbarer Art und Weise mit den Möglichkeiten, die Entwicklung der Kinder zu unterstützen, in Zusammenhang stehen. Von Eltern mit höherer Bildung wird eher erwartet, dass sie solche Aktivitäten und Materialien priorisieren, die akademische und soziale Kompetenzen ihrer Kinder fördern (Conger & Dogan, 2007). Kompetenzunterschiede nach *Migrationshintergrund* können auch als Spezialfall sozialstruktureller Herkunftseffekte bezeichnet werden (Kalter, 2003). Die auch unter Kontrolle von sozialstrukturellen Ressourcen verbleibenden Unterschiede zwischen Kindern mit und ohne Migrationshintergrund werden dabei häufig als *ethnische Residuen* oder *ethnic penalties* bezeichnet (z. B. Kalter, 2006). Als naheliegendes Beispiel sind Nachteile zu nennen, die daraus resultieren, dass die Verkehrssprache von den Eltern nicht beherrscht oder angewendet wird.

Die gerade genannten Ansätze beschreiben also Zusammenhänge zwischen Herkunft und Kompetenzen der Kinder durch Unterschiede in ihrem Sozialisations- und Entwicklungskontext, die sich wiederum in Indikatoren des sozio-ökonomischen Status und des Migrationshintergrunds widerspiegeln können. Ob die zu erwartenden Ungleichheiten an verschiedenen Stellen der Kompetenzverteilungen unterschiedlich groß ausfallen – also spezifische Ungleichheitsmuster zu erwarten sind – lässt sich allerdings weder aus dem bestehenden Forschungsstand noch aus den gängigen Theorien und Modellen unmittelbar ableiten. Vielmehr folgen auch die beschriebenen theoretischen Ansätze der Logik einfacher Mittelwertbeziehungen: Je mehr Ressourcen vorhanden sind, umso höher ist der erwartete Kompetenzstand. Damit wird jedoch der Blick auf Gruppenunterschiede dies- und jenseits des arithmetischen Mittels verstellt und angenommen, dass Ungleichheiten im unteren Leistungsbereich gleich groß wie im mittleren und im oberen Bereich ausfallen. Variieren jedoch nicht nur die Mittelwerte im Leistungsstand nach sozialer oder migrationsspezifischer Herkunft, sondern auch die Streuungen, könnten sich Muster von Bildungseliten oder abgehängten Gruppen zeigen, die sich jedoch weder mit den üblichen mittelwertbasierten Verfahren aufdecken und untersuchen lassen, noch in gängigen theoretischen Modellen explizit Berücksichtigung finden.

### 2.1.2 Forschungsstand

*Sozioökonomischer Status.* Auf Basis der Daten der BiKS-Studie<sup>1</sup> zeigten Weinert und Ebert (2013), dass der Wortschatz von Kindergartenkindern mit dem Wert des *Highest International Socio-Economic Index of Occupational Status* (HISEI) im Haushalt kovariiert und insbesondere Kinder der niedrigsten Statusgruppe (HISEI 16–33) in ihren sprachlichen Kompetenzen eingeschränkt sind. Im Vergleich zur Klassenposition fallen Unterschiede nach *elterlichem Bildungshintergrund* etwas stärker aus. Entsprechend hielt auch die Autorengruppe Bildungsberichterstattung (2014) fest, dass „bereits vor dem Schulstart manifeste sprachliche Kompetenzunterschiede in Abhängigkeit vom Schulabschluss der Eltern ... bestehen“ (S. 61). Auch in Großbritannien fallen die Rückstände in schriftsprachlichen Vorläuferfähigkeiten von Kindern mit niedrigem Bildungshintergrund etwas größer aus, als die von Kindern aus niedrigen sozialen Klassen. Sullivan, Ketende und Joshi (2013) zeigten, dass Unterschiede nach Klasse und Bildungshintergrund in den kognitiven Fähigkeiten von sieben- bzw. elfjährigen Kindern auch nach Kontrolle verschiedener Merkmale der häuslichen Lernumwelten signifikant bleiben. Hinsichtlich des *Haushaltseinkommens* fällt der deutsche Forschungsstand vergleichsweise knapp aus. Biedinger (2009) analysierte den Einfluss relativer Einkommensarmut auf Basis der Daten der ESKOM-V-Studie<sup>2</sup> und fand signifikante Zusammenhänge mit dem kognitiven Entwicklungsstand und dem Wortschatz von Kindergartenkindern. Auch Weinert und Ebert (2013) zeigten, dass das Einkommen Varianz im Wortschatz deutscher Kindergartenkinder aufklärt und vor allem Kinder aus Familien mit sehr geringem Einkommen zurückliegen. Im Vergleich mit anderen Indikatoren des sozioökonomischen Status zeigt auch der US-amerikanisch geprägte internationale Forschungsstand, dass Einkommenseffekte in der Regel eher gering ausfallen (vgl. Blau, 1999). National und international lassen sich also in Bezug auf Indikatoren des sozioökonomischen Status Hinweise dafür finden, dass sich in jungen Jahren Ungleichheiten in sprachlichen und kognitiven Fähigkeiten stärker am elterlichen Bildungshintergrund manifestieren, als an der Zugehörigkeit zu einer sozialen Klasse und dem Haushaltseinkommen.

*Migrationshintergrund.* Migrationsspezifische Unterschiede in sprachlichen Kompetenzen sind vergleichsweise gut dokumentiert. Sowohl beim Wortschatz als auch in Hinsicht auf Fähigkeiten im Bereich Grammatik haben Kinder mit Migrationshintergrund bzw. mit einer anderen Haushaltssprache als Deutsch im Kindergartenalter einen deutlichen Nachteil (Tietze, Becker-Stoll,

1 Die Studie *Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter* (BiKS-3-10) startete mit 550 Kindern an ausgewählten Orten in Bayern und Hessen. Die Ergebnisse von Weinert und Ebert (2013) auf die hier verwiesen wird, basieren auf  $N = 279$  bis  $N = 387$  Probanden (siehe S. 314).

2 ESKOM-V ist das Akronym der Längsschnittstudie *Erwerb von sprachlichen und kulturellen Kompetenzen von Migrantenkindern in der Vorschulzeit*, die im Raum Rhein-Neckar mit ca. 1281 Personen startete. Die berichteten Auswertungen basieren auf  $N = 1061$  Probanden (siehe Biedinger, 2009, S. 201).

Bensel, Eckhardt, Haug-Schnabel, Kalicki, Keller & Leyendecker, 2013; Autorengruppe Bildungsberichterstattung, 2014; Dubowy, Ebert, von Maurice & Weinert, 2008). Dieser nimmt im Verlauf der Kindergartenzeit ab, verschwindet aber nicht vollständig (Becker, Klein & Biedlinger, 2013; Niklas et al., 2012). Die Größe der Disparitäten variiert dabei zwischen den Migrantengruppen und nach Generationenstatus (Relikowski, Schneider & Linberg, 2015), wobei die Nachteile von Kindern mit zwei im Ausland geborenen (Niklas et al., 2012) bzw. zwei nicht deutsch-muttersprachlichen Elternteilen (Weinert & Ebert, 2013) deutlich größer ausfallen, als wenn das nur auf ein Elternteil zutrifft. Die Ergebnisse deuten zudem darauf hin, dass die Unterschiede im Wortschatz von Kindern mit zwei im Ausland geborenen bzw. nicht deutsch-muttersprachlichen Elternteilen größer ausfallen als Ungleichheiten nach sozialer Herkunft (z. B. Weinert & Ebert, 2013). Bei Berücksichtigung der sozialen Herkunft sowie der zuhause gesprochenen Sprache kann ein substantieller Anteil migrationspezifischer Kompetenzunterschiede erklärt werden, allerdings lassen sich auch nach Kontrolle diverser Merkmale der häuslichen und institutionellen Lernumwelt für einige Migrantengruppen nicht alle Differenzen erklären (Becker et al., 2013; Tietze et al., 2013; Relikowski et al., 2015).

## 2.2 Forschungsfragen

Sowohl aufgrund theoretischer Argumente als auch aufgrund der empirischen Befundlage sind in der frühen Kindheit soziale und migrationspezifische Bildungsungleichheiten zu erwarten. Da aber das Ausmaß dieser Ungleichheiten bislang noch nicht entlang verschiedener Herkunftsindikatoren vergleichend an einer bundesweiten Stichprobe untersucht wurde, stellt sich zunächst die Frage: Wie groß fallen soziale und migrationspezifische Ungleichheiten in sprachlichen Kompetenzen bei fünfjährigen Kindergartenkindern in Deutschland aus? Der Forschungsstand deutet darauf hin, dass der Bildungshintergrund ein zentraler Indikator für soziale Ungleichheiten in sprachlichen Kompetenzen ist. Hier sind größere Differenzen zu erwarten als beispielsweise in Hinsicht auf das Haushalteinkommen (vgl. Blau, 1999). Gleichzeitig kann auf Basis des Forschungsstands erwartet werden, dass die Kompetenzunterschiede zwischen den Kindern aus den untersten sozialen Gruppen und denen der mittleren Gruppen – zumindest in Bezug auf Einkommensgruppen und der sozialen Klasse – größer ausfallen als die Kompetenzunterschiede zwischen den Kindern der mittleren und der oberen Gruppen.

Sowohl die theoretische Herleitung als auch die empirische Überprüfung von Bildungsungleichheiten orientiert sich üblicherweise an mittleren Gruppenunterschieden. Was dies- und jenseits des Mittelwerts geschieht, bleibt dabei unberücksichtigt. Das ist insofern erstaunlich, da sich zum einen Kinder privilegierterer Familien im oberen Kompetenzbereich absetzen, oder aber Kinder weniger privilegierter Familien überproportional häufig im unteren Kompetenzbereich verbleiben könnten. Die zweite Forschungsfrage lautet mithin: Welches Muster sozia-

ler und migrationspezifischer Ungleichheiten zeigt sich entlang der gruppenspezifischen Kompetenzverteilungen – eine Gleichverteilung oder höhere Differenzen in den oberen oder unteren Quantilen? Spezifische Erwartungen oder Hypothesen, an welchen Quantilen mehr oder weniger große Ungleichheiten zu erwarten sind, lassen sich aus Theorie und Forschungsstand nur bedingt ableiten.

Die betrachteten Ungleichheitsindikatoren – also Bildungshintergrund, soziale Klasse, Einkommen und Migrationshintergrund – sind nicht unabhängig voneinander. Sowohl soziale als auch migrationspezifische Ungleichheiten sollten zu einem guten Teil durch andere Merkmale des sozialen und migrationspezifischen Hintergrunds und deren kumulative Effekte auf frühkindliche Kompetenzen erklärt werden können. Zudem scheinen beispielsweise elterliche Entscheidungen über den Zeitpunkt des Eintritts in institutionelle Betreuung wiederum mit Merkmalen des sozioökonomischen Status zu kovariieren (Schober & Spiess, 2013). Die dritte Forschungsfrage lautet folglich: Inwiefern bleiben Ausmaß und Muster der Ungleichheiten bestehen, wenn sich die Kinder nicht mehr entlang der jeweils anderen Herkunftsindikatoren sowie hinsichtlich des Alters bei Eintritt in frühe institutionelle Betreuung unterscheiden? Zu erwarten ist eine Reduktion der Differenzen, wenn sich bspw. Kinder mit Migrationshintergrund nicht mehr in ihrem sozioökonomischen Status und ihrem Alter bei Eintritt in institutionelle Kindertagesbetreuung von den Kindern ohne Migrationshintergrund unterscheiden. Klare Erwartungen, an welchen Stellen der gruppenspezifischen Kompetenzverteilungen mehr oder weniger wegkontrolliert wird, lassen sich, wie bereits in Bezug auf die zweite Fragestellung eruiert, nicht unmittelbar aus Theorie und Forschungsstand ableiten.

### 3. Daten, Variablen und Auswertungsmethoden

#### 3.1 Daten

Um diese Fragen zu beantworten, wird auf Daten der ersten Welle der Kindergartenkohorte des Nationalen Bildungspanels (NEPS) zurückgegriffen (Blossfeld, Roßbach & von Maurice, 2011).<sup>3</sup> Diese wurden 2011 erhoben und enthalten Informationen zu 2995 Kindern, die zum Schuljahr 2012/13 regulär eingeschult werden sollten – bei der Erhebung also etwa fünf Jahre alt waren – und eine institutionelle Kindertagesbetreuung in Deutschland besuchten. In einem zweistufigen Ziehungsverfahren wurde zunächst eine bundesweite Zufallsstichprobe von Grundschulen gezogen und anschließend alle Kindergärten erfasst, aus de-

3 Diese Arbeit nutzt Daten des Nationalen Bildungspanels (NEPS): Startkohorte 2 (Kindergarten), doi:10.5157/NEPS:SC2:1.0.0. Die Daten des NEPS wurden von 2008 bis 2013 als Teil des Rahmenprogramms zur Förderung der empirischen Bildungsforschung erhoben, welches vom Bundesministerium für Bildung und Forschung (BMBF) finanziert wurde. Seit 2014 wird NEPS vom Leibniz-Institut für Bildungsverläufe e.V. (LIfBi) an der Otto-Friedrich-Universität Bamberg in Kooperation mit einem deutschlandweiten Netzwerk weitergeführt.

nen Kinder auf die gezogenen Grundschulen übergehen. Aus diesen Kindergärten wurden wiederum zufällig Einrichtungen gezogen und um Teilnahme gebeten, sodass von einer bundesweit repräsentativen Stichprobe gesprochen werden kann. Die Kompetenzen der Kinder wurden von geschulten Testleiterinnen in den Kindergärten erhoben und u. a. durch Angaben zur Familie im Rahmen eines telefonischen Elterninterviews ergänzt (Fey, Linberg, Schlesiger & Roßbach, 2012). Für die 2995 Teilnehmer lagen bei 110 Fällen keine gültigen Kompetenzdaten vor und für 655 keine Informationen zu Merkmalen des familiären Hintergrunds. Darüber hinaus hätten weitere Fälle mit fehlenden Werten auf den anderen verwendeten Variablen ausgeschlossen werden müssen, sodass sich die Analysestichprobe bei fallweisem Ausschluss auf 2009 Kinder reduziert hätte. In der empirischen Sozialforschung besteht jedoch inzwischen weitestgehend Konsens darüber, Fälle mit fehlenden Werten nicht nachträglich auszuschließen, sondern stattdessen Imputationsverfahren anzuwenden (z. B. Enders, 2013). Die Schätzung fehlender Werte wurde zehnfach über *Multiple Imputation with Chained Equations* (MICE) mit dem Paket *ice* (Royston, 2004) in Stata 14 durchgeführt. Die resultierende Analysestichprobe umfasst damit  $N = 2995$  Kinder aus 272 Kindergärten. Das Geschlechtsverhältnis ist ausgeglichen (50.85 % Jungen), das Alter der Kinder betrug bei Testung  $M = 59.99$  Monate und im Durchschnitt befanden sich diese Kinder seit gut zweieinhalb Jahren in einer Einrichtung der Kindertagesbetreuung (Alter bei Eintritt in Monaten:  $M = 31.87$ ). Die deskriptiven Kennwerte der Prädiktorvariablen sind in Tabelle 1 zu finden.

### 3.2 Variablen

Als Indikator für sprachliche Kompetenzen wurde ein Maß des rezeptiven Wortschatzes verwendet, das auf dem *Peabody Picture Vocabulary Test* (PPVT) von Dunn und Dunn (1981) basiert (77 Items,  $\alpha = .86$ ; siehe NEPS, 2011). Um die Interpretation der Ergebnisse möglichst anschaulich und leicht verständlich zu gestalten, wurde der Test auf den prozentualen Anteil korrekt gelöster Items an allen potentiell lösbaren Items skaliert.

Als Indikator zur Bestimmung der Klassenherkunft wird auf das Schema nach Erikson, Goldthorpe und Portocarero (EGP-Klassenschema; Erikson, Goldthorpe, & Portocarero, 1979) zurückgegriffen, mit dem die Klassenlage einer Person über deren Beschäftigungsverhältnis bzw. Art des Arbeitsvertrags bestimmt wird. Die verschiedenen Klassenpositionen sind in drei Kategorien zusammengefasst: Die Kategorie der oberen Klassen umfasst die EGP-Klassen I und II, die sogenannten Dienstklassen [1]. Die Kategorie der mittleren Klassen [2] umfasst IIIa, IV und V. Die Kategorie der unteren Klassen oder Arbeiterklassen [3] umfasst IIIb, VI und VII.

Als Indikator für den elterlichen Bildungshintergrund wurde die Klassifikation aus dem Projekt *Comparative Analysis of Social Mobility in Industrial Nations* (CASMIN) verwendet (König, Lüttinger & Müller, 1988), in die Informationen

über das Vorliegen von sowohl allgemeinbildenden als auch beruflichen Bildungszertifikaten eingehen. Auch hier sind die Werte in drei Kategorien rekodiert. Die Kategorie [3] *Niedrig* beinhaltet alle Personen, die maximal über einen Hauptschulabschluss verfügen (CASMIN 1a/b/c). Die Kategorie [2] *Mittel* umfasst alle Personen, die maximal über ein Abitur verfügen (2a/b/c). Der Kategorie [1] *Hoch* gehören schließlich alle Personen an, die über einen Fachhochschulabschluss oder einen Hochschulabschluss verfügen (3a/b). Wie auch bei den EGP-Klassen wird der höchste im Haushalt verfügbare Wert zur Eingruppierung der Kinder verwendet.

Um Kompetenzunterschiede nach Einkommensgruppen vergleichen zu können, wurde das monatliche Haushalteinkommen mit der Anzahl der im Haushalt lebenden Personen gewichtet.<sup>4</sup> Das so gebildete Nettoäquivalenzeinkommen des Haushalts ist in drei gleich große Teile – also in Terzile – aufgeteilt (niedriges, mittleres und hohes Nettoäquivalenzeinkommen).

Der Migrationshintergrund ist über das Geburtsland definiert. Hierbei werden drei Gruppen unterschieden: (a) kein Migrationshintergrund, (b) ein Elternteil ist im Ausland geboren und (c) das Kind selbst oder beide Eltern sind im Ausland geboren. Detaillierte Unterscheidungen nach Herkunftsland und Generationenstatus haben sich aufgrund der Heterogenität unterschiedlicher Migrantengruppen zwar als sinnvoll erwiesen (z. B. Relikowski et al., 2015), sind aber nicht im Fokus dieser Arbeit und würden ihren Rahmen sprengen. In dieser Stichprobe waren unter den Kindern mit Migrationshintergrund solche mit mind. einem in der Türkei und mit mind. einem in einem Land der ehemaligen Sowjetunion geborenen Elternteil am häufigsten vertreten.

Als Kontrollvariablen werden das Alter bei Testung sowie das Alter bei Kindergarteneintritt berücksichtigt. Deskriptive Kennwerte der Variablen sind in Tabelle 1 dargestellt.

---

4 Zur Berechnung wurde die modifizierte Skala der *Organisation for Economic Co-operation and Development* (OECD) angewendet, die das Nettohaushaltseinkommen mit der Anzahl der im Haushalt lebenden Personen in Verhältnis setzt und dabei den ersten Erwachsenen mit 1, jede weitere erwachsene Person mit 0.5 und Kinder unter 14 mit 0.3 gewichtet (siehe: <http://www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf>).

Tabelle 1: Mittelwerte und Standardabweichungen des Wortschatzes nach sozialer Klasse, Bildungshintergrund, Haushaltseinkommen und Migrationshintergrund

	Wortschatz		
	%	<i>M</i>	<i>SD</i>
Wortschatz	100	60.59	19.34
[1] Soziale Klasse: hoch	58.06	64.48	17.52
[2] Soziale Klasse: mittel	19.67	59.36	18.87
[3] Soziale Klasse: niedrig	22.27	51.61	21.06
[1] Bildungshintergrund: hoch	28.07	68.98	14.33
[2] Bildungshintergrund: mittel	58.74	59.58	19.03
[3] Bildungshintergrund: niedrig	13.18	47.38	21.41
[1] Haushaltseinkommen: hoch	30.58	66.06	16.83
[2] Haushaltseinkommen: mittel	35.50	62.42	18.07
[3] Haushaltseinkommen: niedrig	33.91	53.76	20.70
[1] Kein Migrationshintergrund	72.00	65.63	15.65
[2] Ein Elternteil mit Migrationshintergrund	13.05	52.95	20.92
[3] Beide Eltern mit Migrationshintergrund	14.95	43.19	21.52

Anmerkungen. *M* = arithmetisches Mittel. *SD* = Standardabweichung. *N* = 2996. Quelle: 10.5157/NEPS:SC2:3.0.0, eigene Berechnungen.

### 3.3 Auswertungsmethoden

Zur detaillierten Betrachtung der Kompetenzunterschiede zwischen den Kindern der verschiedenen Herkunftsgruppen werden sowohl lineare Regressionen via *Ordinary Least Squares* (OLS) als auch Quantilsregressionen (Koenker, 2005) geschätzt. Die Quantilsregression ist – wie die lineare Regression – ein lineares Modell, mit dem Zusammenhänge zwischen einer abhängigen Variable,  $y$ , und einer oder mehrerer unabhängiger Variablen,  $x_1, x_2, \dots, x_j$ , modelliert werden können. Im Unterschied zur linearen Regression modelliert die Quantilsregression allerdings nicht bedingte Erwartungswerte bzw. konditionale Mittelwerte,  $E(y|x) = x\beta$ , sondern bedingte bzw. konditionale Quantile,  $Quant_\tau(y|x) = x\beta_\tau$ , einschließlich des Medians<sup>5</sup>, als lineare Funktion der Kovariaten. Kompetenzunterschiede können somit an verschiedenen Stellen der gruppenspezifischen Kompetenzverteilungen geschätzt werden. Zu diesem Zweck wird statt der Fehlerquadratsumme,  $\sum_{i=1}^N (y_i - x_i\hat{\beta})^2$  oder kurz  $\sum_{i=1}^N \hat{\epsilon}_i^2$ , die Summe der gewichteten absoluten Abweichungen minimiert:

5 Der Median ist das .50-Quantil oder das 50-Prozent-Quantil oder kurz  $Q_{.50}$ . Allgemein gilt folgende Beziehung: Das  $\tau$ -Quantil ist das  $100 \times \tau$ . Perzentil; so ist das .10-Quantil das 10. Perzentil.

$$\sum_{i: y_i \geq x_i \hat{\beta}_\tau} \tau |y_i - x_i \hat{\beta}_\tau| + \sum_{i: y_i < x_i \hat{\beta}_\tau} (1 - \tau) |y_i - x_i \hat{\beta}_\tau|$$

Dabei ist  $\tau$  das interessierende Quantil und  $y_i - x_i \hat{\beta}_\tau$  sind die Residuen,  $\hat{\varepsilon}_{it}$ , hier formuliert als Differenz zwischen den beobachteten Werten der abhängigen Variable,  $y_i$ , und den vorhergesagten Werten,  $\hat{y}_{it}$  bzw.  $x_i \hat{\beta}_\tau$ . Wird der obige Ausdruck umformuliert in  $\sum_{i: \hat{\varepsilon}_{it} \geq 0} \tau |\hat{\varepsilon}_{it}| + \sum_{i: \hat{\varepsilon}_{it} < 0} (1 - \tau) |\hat{\varepsilon}_{it}|$ , lässt sich einfacher erkennen, dass die absoluten Werte der positiven Residuen mit  $\tau$  gewichtet und die absoluten Werte der negativen Residuen mit  $1 - \tau$  gewichtet werden.

Neben der Möglichkeit alle Quantile  $\tau$ ,  $0 < \tau < 1$ , modellieren zu können, sind Quantilsregressionen für mittlere Quantile – z. B. für den Median,  $\tau = .50$  – robuster gegenüber Ausreißern und anderen systematischen Abweichungen vom Ideal normalverteilter Fehler als Modellierungen des arithmetischen Mittels.<sup>6</sup> Um dieses Potential zu nutzen, werden Kompetenzunterschiede in bivariaten (Modell I) und multiplen Quantilsregressionen (Modell II) im unteren Bereich (.10- und .25-Quantil = 10. und 25. Perzentil), am Median (.50-Quantil = 50. Perzentil) und im oberen Bereich (.75- und .90-Quantil = 75. und 90. Perzentil) der konditionalen Kompetenzverteilungen modelliert.

Die Ergebnisse für Modell I beantworten die Frage nach den sozioökonomischen und migrationsspezifischen Unterschieden im Wortschatz der Kinder. Anders als viele bisherige Studien werden aber auch Gruppenunterschiede in den verschiedenen Quantilen des Wortschatzes untersucht.

In Modell II wird dann für die jeweils anderen betrachteten Herkunftsindikatoren – soziale Klasse, elterliche Bildung, Einkommen und Migrationshintergrund – sowie für das Alter bei Testung und das Alter beim Eintritt in den Kindergarten kontrolliert. Die betrachteten Differenzen sind dann – in der linearen Regression wie auch in den Quantilsregressionen – also konditional auf diesen Kovariaten zu verstehen. Damit beantwortet Modell II einerseits die Frage, inwieweit die in Modell I gefundenen Gruppenunterschiede bestehen bleiben, wenn für die kumulativen Effekte der anderen Herkunftsindikatoren ebenso kontrolliert wird, wie für die Effekte des Alters bei Eintritt in den Kindergarten, die unterschiedlichen Gruppen in unterschiedlichem Ausmaß zu Gute kommen. Darüber hinaus beantwortet Modell II die Frage, ob ein in Modell I gegebenenfalls gefundenes Muster der Unterschiede zwischen den Quantilen auch unter Kontrolle der Drittvariablen fortbesteht. Geschlecht und Alter bei Testung werden nicht als konfundierende Drittvariablen verstanden, sondern als mögliche Determinanten, die zur präziseren Schätzung kontrolliert werden. Sämtliche Analysen wurden mit Stata 14 durchgeführt.

6 Für ausführliche Einführungen in die Methode siehe Koenker (2005) sowie Hao und Naiman (2007).

## 4. Ergebnisse

### 4.1 Soziale Klasse

Die Ergebnisse der linearen OLS und Quantilsregressionen zu Unterschieden im Wortschatz zwischen Kindern unterschiedlicher Klassenherkunft zeigt Tabelle 2. Während in Modell I die OLS-Regression den Unterschied im Wortschatz zwischen Kindern aus Familien mit hoher und niedriger Klassenposition auf 12.90 Prozentpunkte ( $p < .001$ ) – immerhin etwa zwei Drittel einer Standardabweichung ( $SD = 19.34$ ; siehe Tabelle 1) – schätzt, zeigt die Ergebnisse der Quantilsregressionen, dass die Leistungsunterschiede je nach Quantil sehr unterschiedlich ausgeprägt sind. So fallen die Unterschiede im oberen Bereich ( $Q_{.75}$  und  $Q_{.90}$ ) mit 7.92 und 6.49 Prozentpunkten ( $p < .001$ ) viel geringer aus als die Unterschiede im unteren Bereich. Hier liegt der Wortschatz der Kinder aus den Arbeiterklassen 22.10 ( $Q_{.25}$ ) bzw. 18.10 ( $Q_{.10}$ ) Prozentpunkte ( $p < .001$ ) – etwa eine Standardabweichung – unter dem der Kinder der Dienstklassen.

Modell I zeigt außerdem, dass die Unterschiede im Wortschatz zwischen Kindern mit hoher und mittlerer Klassenherkunft ( $\Delta_{[1] \text{ Hoch} - [2] \text{ Mittel}}$ ) geringer ausfallen als jene zwischen Kindern der mittleren und unteren Klassen. Für den Sprachstand der Kinder ist es also vor allem dann nachteilig, wenn sie einer Arbeiterklasse angehören.

Tabelle 2: Koeffizienten von OLS und Quantilsregressionen des Wortschatzes auf die soziale Klasse der Kinder

	OLS	$Q_{.10}$	$Q_{.25}$	$Q_{.50}$	$Q_{.75}$	$Q_{.90}$
Modell I						
$\Delta_{[1] \text{ Hoch} - [2] \text{ Mittel}}$	5.11***	6.36 <sup>+</sup>	9.09***	5.06***	3.90***	2.60***
$\Delta_{[1] \text{ Hoch} - [3] \text{ Niedrig}}$	12.90***	18.10***	22.10***	12.30***	7.92***	6.49***
$\Delta_{[2] \text{ Mittel} - [3] \text{ Niedrig}}$	7.76***	11.70**	13.00***	7.27***	4.03***	3.90***
Modell II						
$\Delta_{[1] \text{ Hoch} - [2] \text{ Mittel}}$	1.25	1.88	0.84	1.11	1.17	1.37 <sup>+</sup>
$\Delta_{[1] \text{ Hoch} - [3] \text{ Niedrig}}$	3.92**	3.79	4.17*	4.69***	3.41***	2.95**
$\Delta_{[2] \text{ Mittel} - [3] \text{ Niedrig}}$	2.67*	1.91	3.32	3.59**	2.25*	1.58

Anmerkungen. Koeffizienten geben Diff. in Prozentpunkten gelöster Items an.  $N = 2996$ . Robuste Standardfehler. Modell I: ohne Kovariaten. Modell II: kontrolliert für Geschlecht, Alter bei Testung, Alter bei Eintritt in Kindergarten, Bildungshintergrund, Haushaltseinkommen und Migrationshintergrund. Quelle: 10.5157/NEPS:SC2:3.0.0, eigene Berechnungen.

\*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$ , <sup>+</sup>  $p < .10$ .

In Modell II wird für die anderen Herkunftsindikatoren sowie das Geschlecht, Alter bei Testung und das Alter bei Eintritt in den Kindergarten kontrolliert. Die bereits in Modell I sehr kleinen Unterschiede zwischen Kindern hoher und mittlerer Klassenherkunft ( $\Delta_{[1] \text{ Hoch} - [2] \text{ Mittel}}$ ) weisen unter Kontrolle dieser Kovariaten nun an keiner Stelle mehr statistische Signifikanz auf. Der Nachteil der Kinder nied-

riger Klassenherkunft kann durch die Kontrollvariablen dagegen nur teilweise erklärt werden. Während sich die 10-Prozent-Quantile der Kinder unterer Klassen nicht mehr signifikant von denen der Kinder höherer Klassenherkunft unterscheiden, verbleiben vor allem am Quantil  $Q_{.25}$  und  $Q_{.50}$  Unterschiede im Umfang von maximal einer Viertel Standardabweichung.

## 4.2 Bildungshintergrund

Tabelle 3 zeigt die Unterschiede im Wortschatz zwischen Kindern von Eltern mit unterschiedlicher formaler Bildung. Bei Betrachtung von Modell I fällt zunächst auf, dass sich Ungleichheiten im Wortschatz deutlich stärker an der elterlichen Bildung feststellen lassen als an ihrer sozialen Klassenlage. Analog zu den Differenzen nach Klassenherkunft kann aber auch für Kinder von Eltern mit unterschiedlicher Bildung beobachtet werden, dass die Gruppenunterschiede in den unteren Quantilen  $Q_{.10}$  und  $Q_{.25}$  sehr viel höher ausfallen als am Median oder in den oberen Quantilen. Die Ungleichheit wird also mit einer einzelnen Maßzahl, wie z. B. dem Punktschätzer der OLS-Regression, nur unzureichend beschrieben, da sich das Ausmaß der Ungleichheiten je nach Quantil deutlich unterscheidet.

Tabelle 3: Koeffizienten von OLS und Quantilsregressionen des Wortschatzes auf den Bildungshintergrund der Kinder

	OLS	$Q_{.10}$	$Q_{.25}$	$Q_{.50}$	$Q_{.75}$	$Q_{.90}$
Modell I						
$\Delta_{[1] \text{ Hoch} - [2] \text{ Mittel}}$	9.40***	22.60***	13.25***	7.53***	5.19***	3.90***
$\Delta_{[1] \text{ Hoch} - [3] \text{ Niedrig}}$	21.60***	37.10***	34.40***	22.10***	13.00***	9.09***
$\Delta_{[2] \text{ Mittel} - [3] \text{ Niedrig}}$	12.20***	14.50***	21.20***	14.50***	7.79***	5.19***
Modell II						
$\Delta_{[1] \text{ Hoch} - [2] \text{ Mittel}}$	5.53***	13.30***	6.84***	4.34***	2.30***	2.41***
$\Delta_{[1] \text{ Hoch} - [3] \text{ Niedrig}}$	13.30***	22.30***	17.10***	12.00***	8.28***	6.89***
$\Delta_{[2] \text{ Mittel} - [3] \text{ Niedrig}}$	7.75***	9.03***	10.20***	7.96***	5.98***	4.48**

Anmerkungen. Koeffizienten geben Diff. in Prozentpunkten gelöster Items an.  $N = 2996$ . Robuste Standardfehler. Modell I: ohne Kovariaten. Modell II: kontrolliert für Geschlecht, Alter bei Testung, Alter bei Eintritt in Kindergarten, soziale Klasse, Haushaltseinkommen und Migrationshintergrund. Quelle: 10.5157/NEPS:SC2:3.0.0, eigene Berechnungen.

\*\*\*  $p < .001$ , \*\*  $p < .01$ .

Eindrücklich ist außerdem der Unterschied von 37.10 Prozentpunkten ( $p < .001$ ) (1.9 Standardabweichungen) der Kinder mit wenigstens einem akademisch gebildeten Elternteil vor den Kindern, deren Eltern maximal über einen Hauptschulabschluss verfügen ( $\Delta_{[1] \text{ Hoch} - [3] \text{ Niedrig}}$ ) im Quantil  $Q_{.10}$ . Dies ist die größte erfasste Differenz der vorliegenden Untersuchung. Das Ausmaß an Ungleichheit in den Quantilen fällt zwar kleiner aus, ist aber gerade im Vergleich zur sozialen Klassenlage immer noch relativ groß.

Mit Blick auf die Ergebnisse von Modell II wird erkennbar, dass die Ungleichheiten im Sprachstand nach Bildungshintergrund auch unter Kontrolle der Kovariaten nicht nur statistisch signifikant bleiben, sondern auch noch in substantiellem Ausmaß fortbestehen. So entspricht die Differenz im Wortschatz von Kindern mit niedrigem und hohem Bildungshintergrund im unteren Leistungsbereich bzw. an den unteren Quantilen noch in etwa einer Standardabweichung.

### 4.3 Haushaltseinkommen

Den Unterschied im Wortschatz zwischen Kindern aus Familien mit unterschiedlichem Einkommen zeigt Tabelle 4. Wieder wird das aus den Tabellen 2 und 3 bekannte Bild größerer Unterschiede in den unteren Quantilen erkennbar. Die in Modell I ermittelten Unterschiede sind aber allesamt kleiner als die entsprechenden Differenzen zwischen Kindern unterschiedlicher Bildungsherkunft. Der größte Unterschied ist zwischen Kindern aus Familien mit einem hohen Einkommen und solchen mit einem niedrigen Einkommen in den 10-Prozent- und 25-Prozent-Quantilen zu finden. Die Ungleichheiten fallen hier mit ca. einer Standardabweichung ebenfalls sehr groß aus.

Noch stärker als bei den anderen bislang betrachteten Herkunftsindikatoren zeigt sich, dass es für den Wortschatz der Kinder vor allem einen Unterschied bedeutet, zum untersten Drittel der Einkommensverteilung zu gehören, während die Unterschiede zwischen den mittleren und oberen Einkommensgruppen wesentlich geringer ausfallen und bereits im Modell I nur teilweise statistische Signifikanz aufwiesen.

Tabelle 4: Koeffizienten von OLS und Quantilsregressionen des Wortschatzes auf das Einkommen des Haushalts

	OLS	Q <sub>.10</sub>	Q <sub>.25</sub>	Q <sub>.50</sub>	Q <sub>.75</sub>	Q <sub>.90</sub>	
Modell I	$\Delta_{[1] \text{ Hoch} - [2] \text{ Mittel}}$	3.64**	4.81	5.32**	3.12**	1.95 <sup>+</sup>	1.43
	$\Delta_{[1] \text{ Hoch} - [3] \text{ Niedrig}}$	12.30***	19.20***	19.70***	10.80***	7.40***	5.06***
	$\Delta_{[2] \text{ Mittel} - [3] \text{ Niedrig}}$	8.66***	14.40***	14.40***	7.66***	5.45**	4.63**
Modell II	$\Delta_{[1] \text{ Hoch} - [2] \text{ Mittel}}$	-0.96	-3.89	-0.83	0.05	0.25	-0.61
	$\Delta_{[1] \text{ Hoch} - [3] \text{ Niedrig}}$	1.73	-2.25	2.65	2.37*	2.40**	1.11
	$\Delta_{[2] \text{ Mittel} - [3] \text{ Niedrig}}$	2.69*	1.64	3.48*	2.33*	2.15**	1.72*

Anmerkungen. Koeffizienten geben Diff. in Prozentpunkten gelöster Items an.  $N = 2996$ . Robuste Standardfehler. Modell I: ohne Kovariaten. Modell II: kontrolliert für Geschlecht, Alter bei Testung, Alter bei Eintritt in Kindergarten, soziale Klasse, Bildungshintergrund und Migrationshintergrund. Quelle: 10.5157/NEPS:SC2:3.0.0, eigene Berechnungen.

\*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$ , <sup>+</sup>  $p < .10$ .

Werden die in Modell II konditional auf den Kovariaten geschätzten Unterschiede betrachtet, zeigt sich, dass die Leistungsunterschiede nach Einkommen weitestgehend erklärt werden und nur wenige Schätzer ihre statistische Signifikanz behalten. Die Rückstände der Kinder aus Familien aus dem unteren Drittel der Einkommensverteilung konzentrieren sich nun auch nicht mehr an den unteren Quantilen, sondern sind eher in der oberen Mitte ( $Q_{.50}$  und  $Q_{.75}$ ) zu finden.

#### 4.4 Migrationshintergrund

Die unkonditionalen Unterschiede aus Modell I in Tabelle 5 sind – analog zu den Ergebnissen der anderen Herkunftsindikatoren – im unteren Bereich der Verteilungen deutlich größer als im oberen. Zudem liegen alle Quantile der Kinder mit nur einem im Ausland geborenen Elternteil signifikant über denen der Kinder mit zwei im Ausland geborenen Elternteilen ( $\Delta_{[2] 1}$  Elternteil –  $[3] 2$  Eltern). Insbesondere für die unteren Quantile gilt aber, dass die Unterschiede zwischen diesen zwei Gruppen substantiell geringer ausfallen als die Unterschiede zu den Kindern ohne Migrationshintergrund ( $\Delta_{[1]}$  ohne –  $[2] 1$  Elternteil;  $\Delta_{[1]}$  ohne –  $[3] 2$  Eltern).

Tabelle 5: Koeffizienten von OLS und Quantilsregressionen des Wortschatzes auf den Migrationshintergrund der Kinder

	OLS	$Q_{.10}$	$Q_{.25}$	$Q_{.50}$	$Q_{.75}$	$Q_{.90}$	
Modell I	$\Delta_{[1]}$ ohne – $[2] 1$ Elternteil	12.67***	22.99***	21.43***	10.52***	7.92***	4.94***
	$\Delta_{[1]}$ ohne – $[3] 2$ Eltern	22.40***	31.80***	33.60***	23.00***	14.90***	11.70***
	$\Delta_{[2] 1}$ Elternteil – $[3] 2$ Eltern	9.77***	8.83**	12.20***	12.50***	7.01***	6.75***
Modell II	$\Delta_{[1]}$ ohne – $[2] 1$ Elternteil	10.59***	18.10***	17.47***	8.34***	5.97***	4.57***
	$\Delta_{[1]}$ ohne – $[3] 2$ Eltern	18.10***	25.00***	27.20***	17.4***	12.40***	9.56***
	$\Delta_{[2] 1}$ Elternteil – $[3] 2$ Eltern	7.49***	6.88*	9.69***	9.09***	6.45***	4.99***

Anmerkungen. Koeffizienten geben Diff. in Prozentpunkten gelöster Items an.  $N = 2996$ . Robuste Standardfehler. Modell I: ohne Kovariaten. Modell II: kontrolliert für Geschlecht, Alter bei Testung, Alter bei Eintritt in Kindergarten, soziale Klasse, Bildungshintergrund und Haushaltseinkommen. Quelle: 10.5157/NEPS:SC2:3.0.0, eigene Berechnungen.

\*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$ .

Auch unter Kontrolle der elterlichen Klassenposition, der elterlichen Bildung und dem Haushaltseinkommen sowie dem Alter bei Testung und Eintritt in den Kindergarten sind insbesondere die Unterschiede zwischen Kindern ohne Migrationshintergrund und solchen mit zwei im Ausland geborenen Eltern noch beträchtlich und liegen für  $Q_{.10}$ ,  $Q_{.25}$  deutlich über einer, für  $Q_{.50}$  bei fast einer Standardabweichung.

Die Unterschiede zwischen den Kindern mit einem und denen mit zwei im Ausland geborenen Elternteilen ( $\Delta_{[2] 1}$  Elternteil –  $[3] 2$  Eltern) bleiben statistisch signi-

fikant, sind aber bereits in der bivariaten Betrachtung (Modell I) nicht so stark auf den unteren Leistungsbereich konzentriert.

## 5. Zusammenfassung und Diskussion

Obwohl das Kindergartenalter und kompensatorische Potentiale institutioneller Kindertagesbetreuung in den vergangenen Jahren verstärkt ins Blick- und Handlungsfeld von Wissenschaft, Politik und Praxis gerückt sind, lag bisher noch keine Studie früher Bildungsungleichheiten vor, die Ausmaß *und* Verteilung entlang verschiedener Herkunftsindikatoren betrachtet. Die erste Frage der vorliegenden Studie fokussierte daher das Ausmaß sozialer und migrationsspezifischer Ungleichheiten in sprachlichen Kompetenzen bei fünfjährigen Kindergartenkindern in Deutschland. Wie auf Basis von Theorie und Forschungsstand zu erwarten, zeigen sich auch in der hier untersuchten deutschlandweit gezogenen NEPS-Stichprobe soziale und migrationsspezifische Disparitäten im Wortschatz der Kinder. Bivariat ergeben sich die größten Unterschiede nach elterlichem Bildungsabschluss und Migrationshintergrund. Die Unterschiede zwischen Kindern aus Familien unterschiedlicher Klassenlage und unterschiedlichen Einkommens fallen geringer und in ihrem Ausmaß einander recht ähnlich aus. Weiterhin ist der Abstand der Kinder aus sozioökonomisch stärker benachteiligten Familien auf Kinder aus Familien mittlerer sozioökonomischer Position größer als der Abstand der Kinder aus sozioökonomisch privilegierten Familien auf die Kinder aus Familien mittlerer Position. Ob ein oder beide Elternteile im Ausland geboren wurden, ist hingegen weniger entscheidend als die Tatsache, überhaupt einen Migrationshintergrund zu haben (siehe Tabelle 1).

Dass die ressourcenschwächsten Gruppen weiter von den mittleren entfernt sind, als die oberen, steht in Bezug auf *ökonomische* Ressourcen der Familie in Einklang mit Theorie (Duncan & Magnuson, 2003) und Forschung (Weinert & Ebert, 2013). Auch, dass in Deutschland der *Migrationshintergrund* mit stark ausgeprägten Ungleichheiten im Sprachstand verbunden ist, stimmt mit bisherigen Befunden (z. B. Relikowski et al., 2015; Autorengruppe Bildungsberichterstattung, 2014; Tietze et al., 2013; Niklas et al., 2012; Dubowy et al., 2008) überein. Dass sich Ungleichheiten im Sprachstand der Kinder aber noch etwas stärker nach elterlichem *Bildungshintergrund* zeigen und Ungleichheiten nach Klasse und Einkommen in Deutschland ähnlich groß ausfallen, wurde bislang noch nicht dokumentiert.

Unabhängig von den jeweils betrachteten Indikatoren der sozialen oder migrationsspezifischen Herkunft können diese Ungleichheiten aber auch an unterschiedlichen Stellen der Kompetenzverteilung mehr oder weniger groß ausfallen und lassen möglicherweise eine frühe Elite oder eine früh abgehängte Gruppe erkennen. Mit der zweiten und zentralen Forschungsfrage sollte ermittelt werden, ob entsprechende Muster von Ungleichheiten in den Kompetenzverteilungen er-

kennbar sind. Die Ergebnisse der Quantilsregressionen, mit denen das Ausmaß sozialer Ungleichheiten an verschiedenen Stellen der gruppenspezifischen Kompetenzverteilungen geschätzt wurde, zeigen, dass die Ungleichheiten vor allem in den unteren Quantilen der Verteilung bestehen. Hohe Werte im Wortschatztest zu erreichen, stellt für eine Vielzahl der Kinder aus sozioökonomisch schlechter gestellten Familien und für Kinder mit Migrationshintergrund keinesfalls ein Problem dar. Ihr Anteil im Bereich hoher Kompetenzen ist nur geringfügig kleiner als der Anteil der Kinder privilegierter Elternhäuser. Vielmehr verfügt eine überproportional große Zahl dieser Kinder über einen nur sehr geringen Wortschatz, sodass die Befunde durchaus das Bild einer früh abgehängten Gruppe zeigen.

Sowohl das Ausmaß der in den bivariaten Modellen gefundenen Ungleichheiten als auch der Befund größerer Unterschiede in den unteren Quantilen könnte auf kumulative Nachteile zurückzuführen sein, die sich in der bivariaten Betrachtung alleine schon deshalb ergeben, weil die betrachteten Indikatoren miteinander korrelieren: Wer z. B. einen Migrationshintergrund hat, gehört in Deutschland auch oft zu einer unteren sozialen Klasse und einer unteren Einkommensgruppe (Stanat, Rauch & Segeritz, 2010). Weiterhin besuchen gerade Kinder mit Migrationshintergrund und geringem Einkommen den Kindergarten etwas später (Schober & Spiess, 2013), was einen zusätzlichen Nachteil bedeuten kann. Die dritte Forschungsfrage zielte daher auf kumulative Effekte und die Frage ab, inwiefern Ausmaß und Muster der bivariat gefundenen Ungleichheiten auch unter Kontrolle dieser Kovariaten nachzuweisen sind. So zeigten sich in den multiplen Modellen deutlich geringere Differenzen für die einzelnen Herkunftsindikatoren: Konditional auf den jeweils anderen Indikatoren sowie auf dem Alter bei Testung und Eintritt in den Kindergarten schrumpfen die Unterschiede zwar einerseits ganz erheblich zusammen, andererseits bestehen aber nach wie vor – insbesondere in den unteren Quantilen – substantielle Unterschiede bei elterlicher Bildung und Migrationshintergrund von bis zu einer und deutlich über einer Standardabweichung.

Die gefundenen Muster früher sozialer Ungleichheiten wurden bislang weder dokumentiert noch explizit theoretisiert. Die statistisch gesehen einfachste Erklärung erfordert lediglich, dass mit niedrigeren Kompetenzmittelwerten auch höhere Streuungen einhergehen (vgl. Tabelle 1). Dass entsprechende Ungleichheitsmuster bislang nicht thematisiert und untersucht wurden ist insofern überraschend, als auch andere Studien entsprechende Unterschiede in den Streuungen dokumentieren (siehe z. B. Biedinger, Becker & Rohling, 2008; Bos, Tarelli, Bremerich-Vos & Schwippert, 2012; Niklas et al., 2012). Die Kernfrage in Bezug auf den Befund größerer Ungleichheiten im unteren und kleinerer im oberen Leistungsbereich lautet also: Warum sollte eine abnehmende (soziale, kulturelle und ökonomische) Ressourcenausstattung – neben durchschnittlich geringeren Kompetenzen – auch mit einer höheren Streuung sprachlicher Kompetenzen einhergehen?

Aus verschiedenen theoretischen Überlegungen lassen sich durchaus plausible Erklärungen ableiten: Ökologischen Entwicklungstheorien folgend, stellt die

Familie die erste und wichtigste Lernumwelt für die sprachliche Entwicklung dar; demnach wären insbesondere elterliches Verhalten und Investitionen in eine anregungsreiche Lernumwelt für herkunftsbezogene Kompetenzunterschiede verantwortlich (z. B. Conger & Dogan, 2007; Hoff, 2006). Welche Mechanismen aber könnten die höhere Streuung bei geringer Ressourcenausstattung erklären? Geringere Potentiale könnten von sozioökonomisch bessergestellten Familien infolge ihrer besseren Möglichkeiten, die Entwicklung ihres Kindes zu unterstützen, abgemildert werden, sodass diese Kinder selbst bei einer schlechteren Ausgangslage seltener im unteren Leistungsbereich verbleiben. Gleichzeitig stehen entsprechende protektive bzw. fördernde Ressourcen für Kinder aus schlechter gestellten Familien seltener zur Verfügung, sodass sie im Falle einer schlechteren Ausgangslage häufiger im unteren Leistungsbereich verbleiben. Zudem könnten elterliche Ressourcen und Investitionen im oberen Leistungsbereich weniger ertragreich für die sprachlichen Kompetenzen sein und den vergleichsweise geringen Vorsprung in den oberen Quantilen erklären. Die höhere Bandbreite im Wortschatz von Kindern aus sozioökonomisch schlechter gestellten Familien unterstreicht aber auch, dass Annahmen zum Zusammenhang von sozialstrukturellen Merkmalen und Sprachkompetenzen nicht als deterministisch verstanden werden sollten.

Allgemein, aber insbesondere mit Blick auf die Unterschiede nach Migrationshintergrund bleibt jedoch die Frage offen, wodurch sich die verbleibenden substantiellen Unterschiede im unteren Bereich der Kompetenzverteilung erklären lassen. Zukünftige Untersuchungen könnten einerseits weitere allgemeine Indikatoren des Kindes und des Kontexts in Form der häuslichen und institutionellen Lernumwelt sowie weitere migrationsspezifische Merkmale berücksichtigen und überprüfen, inwiefern diese zur weiteren Erklärung der gefundenen Ungleichheitsmuster beitragen: Relevant könnten insbesondere Prozessmerkmale der häuslichen Lernumwelt sein, die als potentielle Mediatoren zwischen distalere Strukturmerkmalen – wie z. B. Indikatoren der sozialen Lage – und dem kindlichen Kompetenz- und Entwicklungsstand gelten. So halten bspw. Sylva et al. (2008) beim Vergleich der Bedeutung von Struktur- und Prozessmerkmalen für den kindlichen Entwicklungsstand fest, dass was Eltern im Alltag mit dem Kind unternehmen wichtiger ist als der sozioökonomische Status (ebd., S. 22). Allerdings ist nicht zu erwarten, dass die Berücksichtigung von Prozessmerkmalen für die Erklärung sozialer Ungleichheiten hinreichend ist und sich der Zusammenhang zwischen sozialem Hintergrund und Sprachstand vollständig über Merkmale der häuslichen Lernumwelt erklären ließe (vgl. Sullivan et al., 2013).

Merkmale der institutionellen Lernumwelt erklären zwar in der Regel wesentlich weniger Varianz im Entwicklungsstand (vgl. Roßbach, 2011), aber schon die Dauer des Kindergartenbesuchs ist mit der Wahrscheinlichkeit später ein Gymnasium zu besuchen verbunden (Seyda, 2009). Generell profitieren Kinder also von institutioneller vorschulischer Bildung, Betreuung und Erziehung und insbesondere von hoher Qualität, es scheinen aber intensive Bemühungen nötig zu sein um kompensatorische Effekte zu erzielen (z. B. Roßbach, Kluczniok & Kuger, 2008).

Kinder mit Migrationshintergrund zeigten besonders ausgeprägte Rückstände im sprachlichen Kompetenzstand, die auch unter Kontrolle der Merkmale des sozialen Hintergrunds und des Alters bei Eintritt in Kindertagesbetreuung fortbestanden. Um diesen auf den Grund zu gehen, ist die Differenzierung weiterer migrationspezifischer Merkmale, wie z. B. Herkunftsland, Generationenstatus und zuhause gesprochene Sprache, sowie entsprechender Theorien und Modelle nötig. So wurde nicht nur für das Schul-, sondern auch für das Kindergartenalter gezeigt, dass besonders Kinder mit türkischem Migrationshintergrund zurückliegen und diese Unterschiede auch unter Kontrolle weiterer migrationspezifischer Merkmale signifikant bleiben, während Nachteile von Kindern mit mindestens einem in einem Land der ehemaligen UdSSR geborenen Elternteil teilweise vollständig aufgeklärt werden konnten (Relikowski et al., 2015; Stanat et al., 2010; Tietze et al., 2013). Eine entsprechend tiefergehende Analyse migrationspezifischer Disparitäten entlang der (konditionalen) Kompetenzverteilung konnte hier leider nicht erfolgen. Auch darüber hinaus hat unser Beitrag einige nennenswerte Limitationen: So sind etwa sprachliche Fähigkeiten nicht unabhängig von allgemeinen kognitiven Fähigkeiten; es kann daher interessant sein, Zusammenhänge mit dem Wortschatz unter Kontrolle der kognitiven Fähigkeiten zu analysieren um soziale Ungleichheiten im Wortschatz unabhängig von allgemeinen kognitiven Fähigkeiten beziffern zu können. Da die Betrachtung von Ausmaß und Muster von frühen Ungleichheiten im Vordergrund stand, waren weiterhin spezifische, z. B. zwischen Strukturmerkmalen und sprachlichem Entwicklungsstand vermittelnde, Mechanismen nicht Gegenstand der Untersuchung. Auf Basis der hier vorgelegten Ergebnisse könnten weiterführende Analysen jedoch zwischen sozialen Hintergrund bzw. Migrationshintergrund und sprachlichen Kompetenzen vermittelnden Prozessmerkmale untersuchen und damit zum besseren Verständnis der Entstehung der Ungleichheiten beitragen. Darüber hinaus könnten längsschnittliche Analysen zur Beschreibung unterschiedlicher Entwicklungsverläufe, aber auch zur besseren kausalen Interpretierbarkeit der Ergebnisse beitragen.

Bezüglich der Auswertungsmethodik ermöglichte die hier angewendete Methode der Quantilsregression, Ungleichheiten entlang der gesamten (konditionalen) Verteilung des Sprachstandes zu ermitteln. Dieses in der deutschen Bildungsforschung bis dato kaum eingesetzte Verfahren kann auch in zukünftigen Untersuchungen zu einem besseren Verständnis von Bildungsungleichheiten beitragen. Es eignet sich sowohl zur Beschreibung unterschiedlicher Streuungen als auch zu deren kausaler Modellierung. In der Bildungsberichterstattung könnten beispielweise Kompetenzverteilungen sowohl bivariat als auch unter Kontrolle von Kovariaten auf statistische Signifikanz hin getestet werden. Ein anderes Anwendungsbeispiel sind Studien, in denen die kausalen Determinanten ungleicher Kompetenzstreuungen modelliert oder die Wirksamkeit kompensatorischer Interventionen gezielt überprüft werden sollen: In welcher Form konnten Ungleichheiten reduziert werden? Wie wirkt sich die Intervention auf die Verteilung der untersuchten Kompetenzen aus? Unter bestimmten Annahmen

(Koenker, 2005, S. 26–32; S. 48) und bei ausreichender Fallzahl<sup>7</sup> kann dann ermittelt werden, wer von der Intervention profitiert: Eher die kompetenzschwachen oder eher die kompetenzstarken Kinder? Diese und weitere Anwendungen der Quantilsregression – auch das sollte unser Beitrag zeigen – ermöglichen es, alte Fragen auf neue Art und Weise zu beantworten und neue Fragen zu stellen und zu erforschen.

## Literatur

- Autorengruppe Bildungsberichterstattung. (2014). *Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen*. Bielefeld: W. Bertelsmann.
- Becker, B., Klein, O. & Biedinger, N. (2013). The development of cognitive, language, and cultural skills from age 3 to 6. A comparison between children of Turkish origin and children of native-born German parents and the role of immigrant parents' acculturation to the receiving society. *American Educational Research Journal*, 50(3), 616–649.
- Betz, T. (2008). *Ungleiche Kindheiten: Theoretische und empirische Analysen zur Sozialberichterstattung über Kinder*. Weinheim: Juventa.
- Biedinger, N. (2009). Kinderarmut in Deutschland. Der Einfluss von relativer Einkommensarmut auf die kognitive, sprachliche und behavioristische Entwicklung von 3- bis 4-jährigen Kindern. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, 29(2), 197–214.
- Biedinger, N., Becker, B. & Rohling, I. (2008). Early ethnic educational inequality: The influence of duration of preschool attendance and social composition. *European Sociological Review*, 24(2), 243–256.
- Blau, D. M. (1999). The effect of income on child development. *Review of Economics and Statistics*, 81(2), 261–276.
- Blossfeld, H.-P., Roßbach, H.-G. & Maurice, J. von. (Hrsg.). (2011). *Education as a lifelong process. The German National Educational Panel Study (NEPS)* (Zeitschrift für Erziehungswissenschaft, Sonderheft 14). Wiesbaden: VS.
- Bos, W., Tarelli, I., Bremerich-Vos, A. & Schwippert, K. (Hrsg.). (2012). *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bronfenbrenner, U. & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner (Hrsg.), *Handbook of child psychology. Volume 1: Theoretical models of human development* (6. Aufl., S. 793–828). Hoboken, NJ: John Wiley & Sons.
- Conger, R. D. & Dogan, S. J. (2007). Social class and socialization in families. In J. E. Grusec & P. D. Hastings (Hrsg.), *Handbook of socialization: Theory and research* (S. 433–460). New York, NY: Guilford Press.
- Cunha, F., Heckman, J. J., Lochner, L. & Masterov, D. V. (2005). *Interpreting the evidence on life cycle skill formation* (Working Paper No. 11331). Cambridge, MA: National Bureau of Economic Research.
- Dubowy, M., Ebert, S., Maurice, J. von & Weinert, S. (2008). Sprachlich-kognitive Kompetenzen beim Eintritt in den Kindergarten: Ein Vergleich von Kindern mit

---

7 Hinsichtlich der Fallzahl muss bei Verwendung von Quantilsregressionen beachtet werden, dass – per definitionem – folgende Logik unausweichlich gilt: Je weiter ein Quantil vom Median entfernt liegt, desto weniger Beobachtungen sind für dessen Schätzung relevant.

- und ohne Migrationshintergrund. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40(3), 124–134.
- Duncan, G. J. & Magnuson, K. A. (2003). Off with hollingshead: Socioeconomic resources, parenting, and child development. In M. H. Bornstein & R. H. Bradley (Hrsg.), *Socioeconomic status, parenting, and child development* (S. 83–106). Mahwah, NJ: Lawrence Erlbaum.
- Dunn, L. M. & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test – Revised (PPVT-R)*. Circle Pines, MN: American Guidance Service.
- Enders, C. K. (2013). Dealing with missing data in developmental research. *Child Development Perspectives*, 7(1), 27–31.
- Erikson, R., Goldthorpe, J. H. & Portocarero, L. (1979). Intergenerational class mobility in three western European societies: England, France and Sweden. *The British Journal of Sociology*, 30(4), 415–441.
- Fey, D., Linberg, T., Schlesiger, C. & Roßbach, H.-G. (2012). Das Nationale Bildungspanel (NEPS) unter besonderer Berücksichtigung von Bildungsprozessen und -institutionen im vorschulischen Alter. In S. Viernickel, D. Edelmann, H. Hoffmann & A. König (Hrsg.), *Krippenforschung. Methoden, Konzepte, Beispiele* (S. 34–45). München: Ernst Reinhardt.
- Hao, L. & Naiman, D. Q. (2007). *Quantile regression*. Thousand Oaks, CA: Sage.
- Hart, B. & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H Brookes.
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74(5), 1368–1378.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26(1), 55–88.
- Kalter, F. (2003). *Chancen, Fouls und Abseitsfallen*. Wiesbaden: Springer Fachmedien.
- Kalter, F. (2006). Auf der Suche nach einer Erklärung für die spezifischen Arbeitsmarktnachteile von Jugendlichen türkischer Herkunft. *Zeitschrift für Soziologie*, 35(2), 144–160.
- Koenker, R. (2005). *Quantile regression*. Cambridge: Cambridge University Press.
- König, W., Lüttinger, P. & Müller, W. (1988). *A comparative analysis of the development and structure of educational systems: Methodological foundations and the construction of a comparative educational scale* (CASMIN-Working Paper No. 12). Mannheim: Universität Mannheim, Institut für Sozialwissenschaften.
- Lee, V. E. & Burkam, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Washington, DC: Economic Policy Institute.
- NEPS – Nationales Bildungspanel. (2011). *Startkohorte 2, Haupterhebung 2010/11 (A12), Kindergartenkinder: Informationen zum Kompetenztest*. Bamberg: Leibniz-Institut für Bildungsverläufe, Nationales Bildungspanel.
- Niklas, F., Möllers, K. & Schneider, W. (2013). Die frühe familiäre Lernumwelt als Mediator zwischen strukturellen Herkunftsmerkmalen und der basalen Lesefähigkeit am Ende der ersten Klasse. *Psychologie in Erziehung und Unterricht*, 60(2), 94–111.
- Niklas, F., Segerer, R., Schmiedeler, S. & Schneider, W. (2012). Findet sich ein „Matthäus-Effekt“ in der Kompetenzentwicklung von jungen Kindern mit oder ohne Migrationshintergrund? *Frühe Bildung*, 1(1), 26–33.
- Relikowski, I., Schneider, T. & Linberg, T. (2015). Rezeptive Wortschatz- und Grammatikkompetenzen von Fünfjährigen mit und ohne Migrationshintergrund. Eine empirische Untersuchung aus bildungssoziologischer Perspektive. *Frühe Bildung*, 4(3), 135–143.

- Roßbach, H.-G. (2011). Auswirkungen öffentlicher Kindertagesbetreuung auf Kinder. In S. Wittmann, T. Rauschenbach & H. R. Leu (Hrsg.), *Kinder in Deutschland. Eine Bilanz empirischer Studien* (S. 173–180). Weinheim: Juventa.
- Roßbach, H.-G., Kluczniok, K. & Kuger, S. (2008). Auswirkungen eines Kindergartenbesuchs auf den kognitiv-leistungsbezogenen Entwicklungsstand von Kindern. In H.-G. Roßbach & H.-P. Blossfeld (Hrsg.), *Frühpädagogische Förderung in Institutionen* (Zeitschrift für Erziehungswissenschaft, Sonderheft 11, S. 139–158). Wiesbaden: VS.
- Royston, P. (2004). Multiple imputation of missing values. *Stata Journal*, 4(3), 227–241.
- Schober, P. S. & Spiess, C. K. (2013). Early childhood education activities and care arrangements of disadvantaged children in Germany. *Child Indicators Research*, 6(4), 709–735.
- Seyda, S. (2009). Kindergartenbesuch und späterer Bildungserfolg. *Zeitschrift für Erziehungswissenschaft*, 12(2), 233–251.
- Stanat, P., Rauch, D. & Segeritz, M. (2010). Schülerinnen und Schüler mit Migrationshintergrund. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 200–230). Münster: Waxmann.
- Sullivan, A., Ketende, S. & Joshi, H. (2013). Social class and inequalities in early cognitive scores. *Sociology*, 47(6), 1187–1206.
- Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I. & Taggart, B. (2008). *Final report from the primary phase: Pre-school, school and family influences on children's development during key stage 2 (age 7–11)* (Research Report No. DCSF-RR061). London: University of London, Institute of Education.
- Tietze, W., Becker-Stoll, F., Bensel, J., Eckhardt, A. G., Haug-Schnabel, G., Kalicki, B., Keller, H. & Leyendecker, B. (Hrsg.). (2013). *Nationale Untersuchung zur Bildung, Betreuung und Erziehung in der frühen Kindheit (NUBBEK)*. Weimar: das Netz.
- Weinert, S. & Ebert, S. (2013). Spracherwerb im Vorschulalter. Soziale Disparitäten und Einflussvariablen auf den Grammatikerwerb. *Zeitschrift für Erziehungswissenschaft*, 16(2), 303–332.