RUHR
ECONOMIC PAPERS

David Kiss

Are Immigrants Graded Worse in Primary and Secondary Education?

Evidence for German Schools

#223

Ruhr Economic Papers

Published by

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #223

Responsible Editor: Christoph M. Schmidt

ISSN 1864-4872 (online) – ISBN 978-3-86788-255-2

Ruhr Economic Papers #223

David Kiss

Are Immigrants Graded Worse in Primary and Secondary Education?

Evidence for German Schools

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen National-bibliografie; detaillierte bibliografische Daten sind im Internet über: *http//dnb.ddb.de* abrufbar.

ISSN 1864-4872 (online)
ISBN 978-3-86788-255-2

David Kiss[1]

Are Immigrants Graded Worse in Primary and Secondary Education? – Evidence for German schools

Abstract

Using PIRLS 2001 and PISA 2003 data for Germany, this paper examines whether immigrants attending primary and secondary school are graded worse in math than comparable natives. Controlling for differences in math skills, class fixed effects regressions and results of a matching approach suggest that immigrants have grade disadvantages in primary education. In Germany, track choice after primary education is mainly determined by the average of grades obtained in math and German. Hence, grade disadvantages could lead to lower level track choice. Immigrants who attend the most common secondary school tracks are not graded differently from natives.

JEL Classification: C40, I21, J15

Keywords: Grading; educational system; migration background; matching

November 2010

1 University of Erlangen-Nuremberg, former affiliation: RWI. – *I am grateful to Guido Schwerdt for motivating my research on educational outcomes of immigrants. I also thank Leilanie Basilio, Thomas K. Bauer, Alfredo R. Paloyo, Regina T. Riphahn, Marcus Tamm and participants of the international workshop on applied economics of education (IWAEE 2010) for helpful comments and suggestions. Financial support by the WGL ("Initiative for Research", Project leader: Andrea Weber) is gratefully acknowledged. The data were produced by IFS Dortmund (IGLU/PIRLS 2001) and provided to me by the research data center (FDZ) at IQB Berlin, whose staff I thank for their help and support. All remaining errors are my own. –* *All correspondence to David Kiss, University of Erlangen-Nuremberg, Lange Gasse 20, 90403 Nuremberg, Germany, E-Mail: david.kiss@wiso.uni-erlangen.de.*

1. Introduction

In Germany, about 30% of pupils have a migration background (Census Bureau 2009). Riphan (2001) and Schnepf (2007) found that immigrants perform worse than natives in educational achievement surveys. As a consequence, immigrants are more likely to attend lower level tracks. However, immigrants should not be graded differently from natives with similar skills, as measured by test scores. Using two representative cross-sectional surveys for Germany, "Progress in International Reading Literacy Study 2001" (PIRLS 2001) and "Programme for International Student Assessment 2003" (PISA 2003), this paper examines whether immigrant children are graded worse in math at the primary and secondary schooling levels after accounting for differences in math test scores.

Track choice after primary education is mainly determined by the average of grades obtained in math and German. The final degree from secondary school largely determines future labor market outcomes like income or the probability of becoming unemployed (Dustmann 2004). Hence, grade disadvantages at the end of primary education could have serious long-term effects on a pupil's economic situation.

Only a small number of studies examined for Germany whether immigrant pupils are graded worse than natives once differences in skills are controlled for. Kristen (2006) finds no evidence as to whether immigrant children have a grade disadvantage in math or German at the end of primary education. In Ditton (2005), immigrants are somewhat graded worse in the third grade. However, both studies are not representative for Germany, since the analyzed data have been collected for single federal states.[1] The paper that is closest to this analysis is Lüdemann & Schwerdt (2010) who primarily focus on track choice and potential earnings profiles of immigrants. Using PIRLS 2001 data, which were restricted to Western Germany, they find that second generation immigrants are graded worse in Math at the end of their primary education, although differences in cognitive abilities are additionally taken into account. The setting in Sprietsma (2009) is somewhat different: Using identical sets of essays written by fourth-graders from German schools, she randomly assigns typical German and Turkish names to test the effect of teacher expectations regarding immigrant background on grades. She finds that essays bearing Turkish names are significantly graded worse.

[1] The data in Kristen (2006) were collected for the city Mannheim, which is located in Baden-Württemberg. The data in Ditton (2005) are representative for the fedral state Bavaria.

The analysis of the data leads to three main findings: (i) Class fixed effects estimates and results of a matching approach suggest that second generation immigrants attending primary school are graded worse than comparable natives. (ii) Regarding first generation immigrants in primary education, empirical evidence is not clear. Depending on the chosen specification, evidence for the existence as well as the non-existence of grade disadvantages can be given. (iii) After track choice, immigrants are not graded differently from comparable natives.

This study contributes to the literature in several ways: (i) The focus on math grades is motivated by the assumption that grading-relevant characteristics are sufficiently captured by math test scores.[2] The problem of biased estimates due to omitted variables is broadly discussed here. (ii) Regarding Germany, other studies solely focused on primary education. This study also evaluates the impact of immigrant background on grades at the most common secondary school tracks. (iii) Results of a non-parametric matching approach from the labor market literature allow a more precise interpretation of class fixed effects estimates.

The paper is organized as follows. Section 2 briefly describes the data and derives the estimation strategy. In section 3, the main results are presented and potential omitted variable biases are discussed. Section 4 concludes.

2. Empirical strategy

2.1 PIRLS 2001 and PISA 2003

Two educational achievement surveys are used in this study: PIRLS 2001 focuses on pupils at the end of their primary education while PISA 2003 covers ninth-graders in secondary school.[3] Primary education generally ends after the fourth grade when children are 10 years old. The most common tracks in secondary education are Gymnasium (which is seen as the highest track), Realschule (intermediate) and Hauptschule (lowest secondary track level).[4]

Math tests used multiple choice and open-ended questions. Figure 1a and Figure 1b show some example exercises from PIRLS and PISA, respectively. These exercises also control for

[2] For instance, the data do not contain an objective measure for spoken language skills. Thus, estimating the impact of migration background on grades in German is potentially biased if only differences in reading skills are controlled for.

[3] PISA 2003 consists of three parts: PISA-international (PISA-I) is the smallest data set. It were collected for international comparisons. In Germany, this data set has been extended twice: PISA-I-PLUS (which is used here) and PISA-E.

[4] There exist additional secondary school forms, like comprehensive schools (Gesamtschulen). These are left out from the analysis, since the final degree obtained from these schools cannot be clearly ranked with degrees from Gymnasium, Realschule or Hauptschule. In PISA, 20% of the pupils attend alternative secondary school tracks.

language skills, since pupils with a bad command of the German language may not fully understand the exercises which would result to lower test scores.

Apart from test scores, additional data were collected from questionnaires completed by students, parents, teachers and principals. Most information about the pupil, namely its math grades in the last school report and its migration background are obtained from the student questionnaire.[5] In both surveys, schools are the primary sampling unit. Within each school, classes have been randomly chosen. All pupils attending a drawn class participated in the tests on a voluntary basis.

2.2 Class fixed effects model

The impact of immigrant background on math grades is estimated with the following model:

$$y_{ic} = \alpha + \beta tsmat_{ic} + \gamma_1 mig_{ic}^1 + \gamma_2 mig_{ic}^2 + \lambda_c + \varepsilon_{ic}. \qquad (1)$$

y_{ic} is the self-reported math grade of student i attending class c. $tsmat$ is the test score in math. Within a schooling level or track, $tsmat$ has mean 0 and standard deviation 1. λ controls for class fixed effects, while errors ε are clustered on class level. Migration background is captured by the dummy variables mig^1 and mig^2, where the superscript distinguishes between first and second generation immigrants. A pupil is a first generation immigrant if it was born abroad. Second (or higher) generation immigrants are born in Germany and satisfy at least one of the following conditions: (i) at least one parent was born abroad or (ii) languages other than German are (also) spoken at home. A child is treated as a native, if it is born in Germany, as well as its parents, and German is the only language spoken at home.

In Germany, one subject is usually taught by the same teacher during the whole school year. Hence, class fixed effects λ control for the average teacher effect on math grades. Grading practices usually differ among teachers. For instance, they can be influenced by average skills on class level (Himmler & Schwager 2007 and Dardanoni, Modica & Pennsi 2009), the (controversial) impact of class size and its determinants on student's performance (Angrist & Lavy 1999 and Wößmann & West 2006) or the teacher's aspiration level (Iacus & Porro 2008).[6]

---

[5] PIRLS took place in May 2001 while PISA was conducted from March to August 2003. Pupils receive their half-yearly school reports by the end of February.

[6] The usage of class fixed effects (instead of teacher fixed effects) makes sense even if two classes are taught by the same teacher because these classes could be different e.g. in terms of shares of skilled pupils or in the progress in the syllabus, which could lead to differences in applied grading practices in those classes.

Summing up, β captures the impact of grading-relevant characteristics, while γ_1 and γ_2 are expected to be estimated insignificantly. Calculations are carried out separately on the four subsamples primary school, Gymnasium, Realschule and Hauptschule.[7]

In Germany, 1 is the best grade and 6 the worst. For the purpose of this study, math grades have been recoded to range from 0 (fail) to 5 (very good). As shown in Table 1, natives in primary school and Gymnasium obtain better grades than immigrants. Differences in mean grades are relatively small in secondary education. First generation immigrants are graded best in Realschule and Hauptschule.

The higher a pupil's math test score the higher its math proficiency. Mean math test scores are similarly distributed within each schooling level and track: natives perform best, as followed by second generation immigrants who perform better than first generation immigrants. Immigrants are more likely to attend classes with low levels of average math proficiency since, compared to natives, average math test scores on class level are significantly lower in classes that are attended by immigrants.

2.3 Matching approach

The matching approach originates from the labor market literature. In Ñopo (2008), a non-parametric alternative to the classical Blinder-Oaxaca decomposition (see Blinder 1973) is used to estimate gender wage gaps.[8] It depends on matching men and women that are similar in terms of relevant labor market characteristics like education or potential work experience and calculating the wage differentials of matched individuals. For the purpose of this study, the wage gap between men and women can be transformed into a math grade gap between natives and immigrants with same math test scores.

The approach depends on constructing "statistical twin couples" which always consist of an immigrant and a native sibling. The grade difference between siblings is the variable of interest. The procedure can be split up into two steps: First, each immigrant attending a class c is matched on a native child with similar math test score x_{ic}. Therefore, immigrants are left out if no suitable native child attends the same class. However, sometimes an immigrant child can be matched on more than one native from the same class. In these cases, $J-1$ clones of the immigrant child are created which are indexed by j. In the second step, the grade difference of

[7] Lüdemann & Schwerdt (2010) also report results of a similar regression. They additionally control for cognitive abilities and a broad set of individual and socio-economic background variables.

[8] Ñopo (2008) is mainly motivated by Barsky et al. 2002 who analyzed the black-white wealth gap in the United States.

each statistical twin couple is calculated and weighted by J^{-1}, the inverse of the number of suitable native siblings. Grade gaps are denoted by $\tilde{\Delta}^0_{ijc}$.[9] By definition, positive values of $\tilde{\Delta}^0_{ijc}$ imply a grade advantage for immigrants and vice versa. Consequently, $\tilde{\Delta}^0_{ijc} = 0$ means that the immigrant and native sibling are not graded differently. The approach is derived in more detail in the appendix.

Compared to OLS, matching has two advantages: (i) For each class, OLS assumes the same linear relationship between grades and test scores. The matching approach, however, requires no assumption regarding the functional relationship between grades and relevant characteristics. Hence, matching allows for differences in applied grading practices. (ii) Grade differentials obtained by matching solely depend on pupils that are comparable in terms of relevant characteristics. In OLS, immigrants and natives that remain unmatched also contribute to the estimation of the parameters.

However, matching makes sense only for groups which do not differ too much in relevant characteristics. Figure 2 illustrates the distributions of math test scores of immigrants and natives. In most cases, the common area between the distributions of second generation immigrants and natives is larger than the common area between first generation immigrants and natives. Thus, a larger share of second generation immigrants can be matched on natives.

3. Results and discussion

3.1 Main results

Table 2 contains class fixed effects estimates of migration background on math grades. As expected, the highly significant estimates for math test scores imply that high test scores are correlated with better grades. Compared to natives, first and second generation immigrants attending primary school are graded worse 0.23 grades on average. Figure 3a depicts the distribution of grade gaps $\tilde{\Delta}^0_{ijc}$ of matched first generation immigrants attending primary school. 380 first generation immigrants from 199 classes were matched on 540 natives.[10] For test scores between -2 and 2, the 90% confidence band of the smoothed regression line runs below the zero-grade-gap line. Therefore, Figure 3a confirms the findings in Table 2.

---

[9] As already mentioned, c is the index of the class, while i (with $i = 1, \ldots, I_c$) is the index of matching variable that constitutes a statistical twin couple in that class. j is the index of the (cloned) immigrant for matching variable x_{ic}.

[10] Since the number of grade gaps (540) exceeds the number of first generation immigrants (380), some of them have been "cloned".

Figure 3b plots the smoothed regression line for second generation immigrants attending primary school. Here, 297 immigrants have been matched on 429 natives. For test scores smaller than zero, the 90% confidence band of the smoothed regression line overlaps with the zero-grade-gap line. This implies that second generation immigrants with a test score smaller than zero are not graded differently from natives. Compared to natives, however, second generation immigrants with test scores above zero are graded worse since the confidence band of the smoothed regression line runs below the zero-grade-gap line. The results from the matching approach suggest that the negative estimate for second generation immigrant status in Table 2 is caused by well-performing second generation immigrants.

Regarding Gymnasium, class fixed effects estimates imply that immigrants and natives are graded the same way. The matching approach confirms this finding since the smoothed regression lines in Figure 4a and Figure 4b do not significantly deviate from the zero-grade-gap line. For Realschule, it can be inferred from Figure 5a that the positive estimate for first generation immigrant status in Table 2 is caused by first generation immigrants with test scores below zero. The insignificant estimate for second generation immigrant status is consistent with the related smoothed regression in Figure 5b. Regarding Hauptschule, class fixed effects estimates suggest that first generation immigrants have a grade advantage. More precisely, first generation immigrants with test scores between −1 and 1 are graded better as shown in Figure 6a. Second generation immigrants attending Hauptschule are not graded differently from natives as indicated by the related estimate and Figure 6b.

3.2 Omitted variable bias?

Only in primary education, first and second immigrants are systematically graded worse. On the other hand, first generation immigrants attending Realschule or Hauptschule are graded better than comparable natives. Second generation immigrants in secondary education are not graded differently from natives.

However, results in Table 2 could be biased due to unobserved heterogeneity for at least three reasons. (i) Documents from the German educational ministry state that grading should solely depend on the student's performance in written exams *and* its oral participation (KMK 2009: 130). It is likely that math skills and oral participation in math are strongly correlated. However, differences in oral participation are not taken into account in Table 2. (ii) Although estimates of test scores are highly significant, there is a possibility that math proficiency is insufficiently captured by math test scores. (iii) Other pupil characteristics, like sex or educational background of parents, could additionally affect grading although math skills are controlled for.

In Table 3, different sets of controls are taken into account: In the first columns math skills are considered as the only grading-relevant characteristic. Consequently, these estimates are identical to those in Table 2. In the second columns of Table 3, differences in self-reported oral participation in math are additionally controlled for. In both surveys, pupils were asked to which extent they agree with the statement "I frequently participate in math". The four possible outcomes of that variable have been decomposed into four dummy variables.

Comparing columns (1) and (2) in Table 3 shows that oral participation is strongly correlated with math grades. Regarding the impact of immigrant background on grades, however, estimates for immigrant status remain almost unaffected in primary education and Gymnasium. In Realschule and Hauptschule, estimates for first generation immigrant status become insignificant. Thus one could draw the conclusion that first generation immigrants attending Realschule are not graded differently from natives if differences in math skills and oral participation are controlled for. From that point of view, the positive grade gaps in Figure 5a and Figure 6a are caused by higher participation rates of (a sufficiently large share of) first generation immigrants.

Instead of oral participation, reading skills are additionally controlled for in the third columns of Table 3. Comparing columns (3) with columns (1) for all secondary school tracks, the point-estimate for reading skills is close to zero and insignificant. Estimates for the impact of immigrant background on math grades are almost identical in columns (3) and (1). In primary education, however, controlling additionally for reading skills yields insignificant point estimates for first generation immigrant status. Second generation immigrants in primary education are still graded worse but the absolute value of the estimate decreases by about 50%.

Depending on the schooling level, once differences in oral participation or reading skills are controlled for, estimates for the impact of immigrant status remain almost unchanged if oral participation and reading skills (columns 4) as well as other pupil characteristics (column 5) are additionally taken into account. Regarding columns (5), it is noteworthy that controlling amongst others for the highest educational background of parents does not change the coefficients for immigrant status. On average, immigrants' parents hold lower level school degrees.

However, differences in grading practices applied on immigrants and natives are driven by differences in socio-economic background variables.[11]

3.3 Do first generation immigrants in primary education suffer from grade disadvantages?

It can be inferred from Table 3 that immigrants in secondary school are not graded differently from natives. Regarding primary education, no clear answer can be given to the question whether column (1) or (3) should be considered as relevant for the research question.

One should rely on the results in column (3) if grading-relevant characteristics are additionally captured by reading skills. Hence, one must try to explain why reading skills have a large impact in primary education but are irrelevant in secondary education at the same time. There are at least three explanations for this pattern: (i) Math skills could be insufficiently measured in PIRLS since the math test took 20 minutes (Bos et al. 2003: 196). In PISA, pupils were tested for a total of 210 minutes (Baumert et al. 2004: 29). From that point of view, one should rely more on estimates in column (3). On the other hand, 20 minutes should be enough to evaluate basic math skills of fourth graders. Furthermore, math tests in PIRLS are mainly in accordance with the German curriculum (Bos et al. 2003: 193). (ii) Main subjects in primary school, namely math and German, are usually taught by the same teacher. Subject teachers are common in secondary school. If teachers in primary education do not clearly distinguish between a student's performance in math and German when grading, math grades could be correlated with skills that are relevant for German. This reasoning would be consistent with the finding that reading skills have no impact on math grades in secondary education once math skills are controlled for. If this is the case, reading skills may not be considered as grading-relevant for math. (iii) Depending on the schooling level and age of a child, reading skills could capture different aspects of proficiency: In primary education, reading skills could measure grading-relevant characteristics like the scope of vocabulary or capability of understanding spoken German. In PISA, pupils are five years older than in PIRLS. Consequently, the scope of vocabulary as well as the capability of understanding could have strongly improved meanwhile. Hence, one should rely on the results in column (3). On the other hand, math tests in PIRLS control indirectly for differences in basic language skills since most exercises consist of short texts (see Figure 1a).

[11] The only exception are first generation immigrants in Hauptschule. The related estimate becomes insignificant only in column (5).

Regarding primary education, there are convincing arguments which support and refute the hypothesis that reading skills do additionally capture grading-relevant characteristics in math once differences in math skills are taken into account. Hence, the scope of the grade disadvantage in primary education ranges from zero (insignificancy) to 0.23 for first generation immigrants and 0.12 to 0.23 for second generation immigrants.

4. Conclusion

Assuming that math test scores, as measured in PIRLS and PISA, sufficiently capture all grading-relevant characteristics in the subject math, this paper examined whether immigrants are graded worse in primary and secondary school if differences in math skills are controlled for. Class fixed effects regressions and results from a non-parametric matching approach suggest that immigrants attending the most common secondary school tracks are not graded differently from natives.

Regarding primary school, there is robust empirical evidence that second generation immigrants are graded worse than natives although differences in math skills are taken into account. More precisely, the related matching estimates imply that second generation immigrants who performed above-average in the PIRLS math tests are the group which suffers from grade disadvantages. Only if math *and* reading skills are considered as grading-relevant, first generation immigrants attending primary school are graded like natives.

In Germany, track choice after primary education mainly depends on the average of grades obtained in math and German. Consequently, grade disadvantages for immigrants in primary education could result to biased track choice which is supported by the results of Lüdemann and Schwerdt (2010) who find an educational disadvantage for second-generation immigrants at the transition to secondary schools.[12] This is a serious problem since (i) immigrants already account for 30% of the student body in primary education and (ii) track choice has a lifelong impact on the future economic situation of a pupil.

---

[12] Most students do not change tracks after their initial choice, although it is generally possible (Schnepf 2002).

Appendix

Derivation and properties of the matching approach

Let X_c^k be the set of observed math test scores of group k ($k = n, m$) in class c. k distinguishes between natives n and immigrants m. In the following, the subscript c in X_c^k will be omitted, since the results of the theoretical concept are the same for any class. X^k consists of two subsets:

$$X^k = \bar{S}^k \cup S^k, \tag{2}$$

with

$$\bar{S}^k = X^k \setminus X^{-k} \text{ and } S^k = X^k \cap X^{-k}, \text{ where } -k = m \text{ if } k = n \text{ and vice versa.}$$

S^k is the subset of X^k which overlaps with X^{-k}, so $S^k = S^{-k}$. $\bar{S}^k$ is the subset of X^k with $\bar{S}^k \cap X^{-k} = \{\}$.

Let $g^n(x) = E(y|n, x)$ and $g^m(x) = E(y|m, x)$ be the expected values of math grades y, conditional on migration background and test scores $x \in X^k$. $g^k(x)$ may be considered as the grading practices applied on pupils with identical test scores and migration background. The difference in mean grades Δ can be expressed as

$$\Delta = E(y|n) - E(y|m) = \int_{X^n} g^n(x) f^n(x) dx - \int_{X^m} g^m(x) f^m(x) dx, \tag{3}$$

where $f^k(x)$ are the corresponding probability densities. Using equation (2) yields

$$\Delta = \left[\int_{\bar{S}^n} g^n(x) f^n(x) dx + \int_{S^n} g^n(x) f^n(x) dx\right] - \left[\int_{S^m} g^m(x) f^m(x) dx + \int_{\bar{S}^m} g^m(x) f^m(x) dx\right]. \tag{4}$$

Since $X^k \cap X^{-k} = S^n = S^m$, (4) can be rewritten as

$$\Delta = \int_{\bar{S}^n} g^n(x) f^n(x) dx - \int_{\bar{S}^m} g^m(x) f^m(x) dx + \Delta^M = \Delta^U + \Delta^M, \tag{5}$$

with

$$\Delta^M = \int_{X^k \cap X^{-k}} g^n(x) f^n(x) dx - \int_{X^k \cap X^{-k}} g^m(x) f^m(x) dx. \tag{6}$$

Δ^M measures the part of the grade gap Δ that is caused by matched natives and immigrants. Adding $0 = \int_{X^k \cap X^{-k}} g^n(x) f^m(x) dx - \int_{X^k \cap X^{-k}} g^n(x) f^m(x) dx$ results in

$$\Delta^M = \int_{X^k \cap X^{-k}} g^n(x)[f^n(x) - f^m(x)]dx + \int_{X^k \cap X^{-k}} f^m(x)[g^n(x) - g^m(x)]dx = \Delta^E + \Delta^0. \quad (7)$$

Substituting (7) into (5) yields

$$\Delta = \Delta^U + \Delta^E + \Delta^0. \quad (8)$$

Δ can be decomposed into 3 components: Δ^U is the part of the grade gap which is caused by unmatched pupils. Hence, Δ^U disappears if there are no immigrants with characteristics that remain unmatched by natives. Δ^M consists of two parts: $\Delta^E = \int g^n(x)[f^n(x) - f^m(x)]dx$ explains the gap in mean grades due to differences in the distributions of test scores (of matched individuals). The second part,

$$\Delta^0 = \int_{X^k \cap X^{-k}} f^m(x)[g^n(x) - g^m(x)]dx, \quad (9)$$

is the "unexplained component" of the grade gap.[13] Since Δ^0 can be derived for any class c,

$$\Delta_c^0 = \int_{X_c^k \cap X_c^{-k}} f_c^m(x)[g_c^n(x) - g_c^m(x)]dx. \quad (10)$$

Fair grading practices imply $g_c^n(x) = g_c^m(x)$ for any level of x, since migration background should not be relevant for grading. It is noteworthy that equation (10) has a great flexibility because different grading practices may be applied in different classes since equation (10) allows $g_c^k(x) \neq g_{-c}^k(x)$ if $c \neq -c$. Moreover, $g_c^k(x)$ does not require any assumptions regarding the functional relationship between y and x.[14]

In this study, the following estimator of Δ_c^0 is proposed and implemented:

$$\hat{\Delta}_c^0 = I_c^{-1} \sum_{i=1}^{I_c} J_{ic}^{-1} \sum_{j=1}^{J_{ic}} g_c^n\left(x_{ic}^{2,j}\right) - g_c^m\left(x_{ic}^{1,j}\right) = I_c^{-1} \sum_{i=1}^{I_c} \sum_{j=1}^{J_{ic}} -J_{ic}^{-1} \Delta_{ijc}^0. \quad (11)$$

---

[13] In the classical Blinder-Oaxaca setup, $\beta^n(x^n - x^m)$ is the counterpart of Δ^E, while $x^m(\beta^n - \beta^m)$ is the equivalent of the unexplained component Δ^0.

[14] For instance, the Blinder-Oaxaca approach assumes a linear relationship between y and x.

I_c is the number of suitable matching variables x_{ic} in class c. Within each class a twin couple is constituted, if $x_{ic}^{1,j} = x_{ic}^{2,j}$, where $j = 1, \ldots, J_{ic}$ is the index of the twin couple which has been constituted on x_{ic}. The first sibling is always the immigrant child which is denoted by the superscript 1. In some classes an immigrant child with test score x_{ic} can be matched on more than one native sibling. In these cases $(J_{ic} - 1)$ "clones" of the first sibling are created which are indexed by j. Hence ic is the index of the common matching variable and J_{ic} is the number of twin couples that are constituted on x_{ic}. The inverse of J_{ic} serves as a weight. $g_c^m(x_{ic}^{1,j})$ and $g_c^n(x_{ic}^{2,j})$ is the math grade of the first and second sibling, respectively. The difference

$$\Delta_{ijc}^0 = g_c^m\left(x_{ic}^{1,j}\right) - g_c^n\left(x_{ic}^{2,j}\right)$$

is the grade gap of the twin couple i, clone j in class c. Positive values of Δ_{ijc}^0 indicate that the immigrant child has an advantage relatively to its native statistical sibling, while negative values imply a grade disadvantage.

However, there is still a high probability of observing $g_c^k(x^1) = g_c^k(x^2)$ even if the teacher is fair and $x^1 = x^2$. For instance, this problem arises when clones of an immigrant child are matched on several natives that are identical in terms of test scores but are graded differently. Apart from measurement errors there are at least two unobserved variables, which are relevant for grading: former oral participation and former math skills. In PIRLS and PISA students have been asked to report the math grade in their last school report. That grade cannot depend on test scores which have been obtained few months later. Although one can assume that former and present math skills are highly correlated, no information about the frequency and quality of former oral participation is available. Thus, one must rewrite equation (11):

$$\hat{\Delta}_c^0 = I_c^{-1} \sum_{i=1}^{I_c} -J_{ic}^{-1} \sum_{j=1}^{J_{ic}} \left[g_c^m\left(x_{ic}^{1,j}\right) - g_c^n\left(x_{ic}^{2,j}\right) + \varepsilon_{ijc}\right] = I_c^{-1} \sum_{i=1}^{I_c} \sum_{j=1}^{J_{ic}} -J_{ic}^{-1} \tilde{\Delta}_{ijc}^0. \qquad (12)$$

with $\tilde{\Delta}_{ijc}^0 = \Delta_{ijc}^0 + \varepsilon_{ijc}$ and $\varepsilon_{ijc} = \varepsilon_{ic}^{1,j} - \varepsilon_{ic}^{2,j}$. Δ_{ijc}^0 is the variable of interest, but we can only obtain $\tilde{\Delta}_{ijc}^0$ from the data. Only if $\varepsilon_{ic}^{1,j} = \varepsilon_{ic}^{2,j} = 0$ for all i, j and c, then nonzero values of $\tilde{\Delta}_{ijc}^0$ are solely caused by different grading practices that are applied to children in with identical test scores, but different migration backgrounds. Since identical pupils (in terms of test scores) might differ in other grading-relevant characteristics, like (former) oral participation, it is more realistic to assume $\varepsilon_{ic}^{1,j} \neq \varepsilon_{ic}^{2,j} \neq 0$. Therefore $\tilde{\Delta}_{ijc}^0$ consists of two "legitimate" parts

$\varepsilon_{ic}^{1,j}$ and $\varepsilon_{ic}^{2,j}$, which are caused by grading-relevant but unobserved characteristics of matched pupils, and the "causeless" part Δ_{ijc}^{0}, which measures discrimination. The smoothed regressions in Figure 3a to Figure 6b are unbiased if $E(\varepsilon_{ijc}) = 0$ and non-autocorrelation of residuals is assumed.

$\tilde{\Delta}_{ijc}^{0}$ is calculated as follows: Since the matching variable math test scores is continuous, one must allow for matching on "similar" test scores. For the purpose of this study an immigrant child with test score x becomes the first sibling of a twin couple if:

$$\hat{\beta} \cdot \left| x_{ic}^{1,j} - x_{ic}^{2,j} \right| \leq 0.05,$$

where x^1 is the value of the matching variable x_i of the immigrant child, clone j in class c and x^2 the related value of the potential native sibling. $\hat{\beta}$ is the estimate of a class fixed effects regression of test scores on math grades.[15] Therefore $\hat{\beta}|x^1 - x^2|$ is the absolute predicted grade deviation of the second sibling from the first. Since math grades range from 0 to 5, the relative magnitude of a grade deviation of 0.05 is 1%. Note that 1% is the maximum possible value of the relative magnitude, since the difference in test scores must imply a predicted grade deviation smaller or equal 0.05. Using the absolute value of the predicted grade deviation allows for matching immigrant children on natives, which perform slightly better or worse in the math test of PISA and PIRLS respectively.

[15] $\hat{\beta}$ is calculated separately for every schooling level and track. In addition, each subsample has been restricted to natives in order to obtain estimates for (relatively) homogenous groups.

References

Angrist, J. D. & Lavy, V. (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. Quarterly Journal of Economics 114(2), pp. 533-575.

Barsky, R. ; Bound, J. ; Charles, K. ; Lupton, J. (2002). Accounting for the Black-White Wealth Gap: A Nonparametric Approach. Journal of the American Statistical Association 97, pp. 663–673.

Baumert, J. et al. (Eds.) (2004): "PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland. Ergebnisse des zweiten internationalen Vergleichs, Opladen.

Blinder, A. S. (1973). Wage Discrimination. Reduced Form and Structural Estimates, Journal of Human Resources 8(4), pp. 436-455.

Bos, W. er al. (Eds.) (2003). Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Münster.

Census Bureau. (2009). Bevölkerung mit Migrationshintergrund. Ergebnisse des Mikrozensus 2007. Wiesbaden: korrigiert am 26. März 2009.

Dardanoni, V., Modica, S., & Pennsi, A. (2009). Grading Across Schools. The B.E. Journal of Economic Analysis & Policy 9(1) . Article 16.

Ditton, H. e. (2005). Bildungsungleichheit. Der Beitrag von Familie und Schule. Zeitschrift für Erziehungswissenschaft 8(2) , pp. 285-304.

Dustmann, C. (2004). Parental Background, Secondary School Track Choice, and Wages. Oxford Economic Papers 56 , pp. 209–230.

Himmler, O., & Schwager, R. (2007). Double Standards in Educational Standards ? Are Disadvantaged Students Being Graded More Leniently? ZEW Discussion Papers 7(16).

Iacus, S., & Porro, G. (2008). School Grades and Students' Achievement: How to Identify Grading Standards and Measure their Effects. UNIMI - Research Papers in Economics, Business, and Statistics no. 1061 .

KMK (Kultusministerkonferenz) (2009). Das Bildungswesen in der Bundesrepublik Deutschland 2008. Darstellung der Kompetenzen, Strukturen und bildungspolitischen Entwicklungen für den Informationsaustausch in Europa. Bonn.

Kristen, C. (2006). Ethnische Diskriminierung in der Grundschule? Die Vergabe von Noten und Bildungsempfehlungen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 58(1) , pp. 79–97.

Lüdemann, E. & Schwerdt, G. (2010). Migration Background and School Tracking: Is there a Double Disadvantage for Second-Generation Immigrants? CESifo Working Paper.

Ñopo, H. (2008). Matching as a Tool to Decompose Wage Gaps. Review of Economics and Statistics 90(2) , pp. 290-299.

Riphahn, R. (2001). Dissimilation? The Educational Attainment of Second Generation Immigrants. CEPR Discussion Paper 2903.

Schnepf, S. (2002). A Sorting that never fails? The Transition from Primary to Secondary School in Germany. Innocenti Working Papers 92.

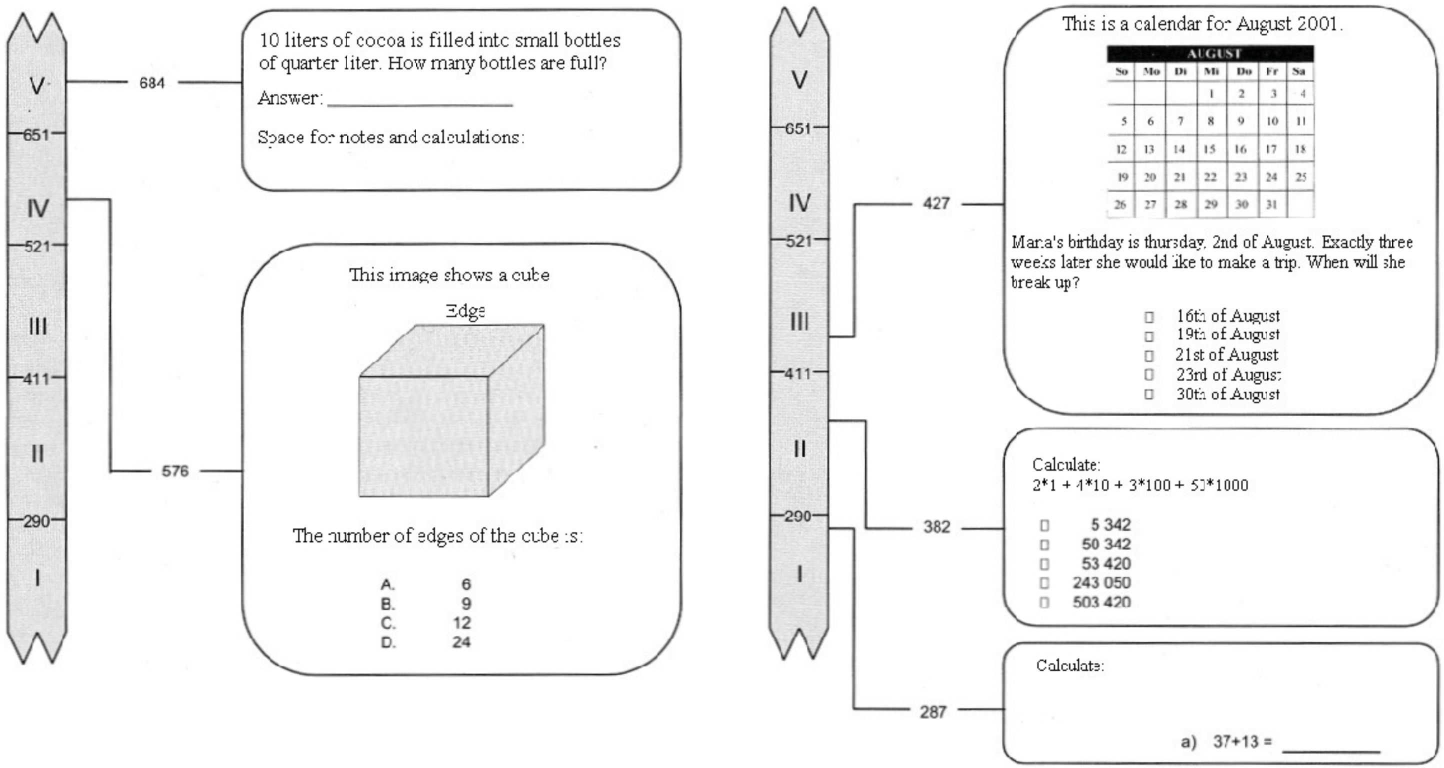
Schnepf, S. (2007). Immigrants' Educational Disadvantage: An Examination across Ten Countries and Three Surveys. Journal of Population Economics 20(3), pp. 527-545.

Sprietsma, Maresa (2009). Discrimination in Grading? Experimental Evidence from Primary School, ZEW Discussion Papers 9(74).

Wößmann, L., & West, M. R. (2006). Class-Size Effects in School Systems Around the World: Evidence from between-Grade Variation in TIMSS. European Economic Review 50(3), pp. 695-736.

Tables and figures

Figure 1a: Example excercises from PIRLS 2001

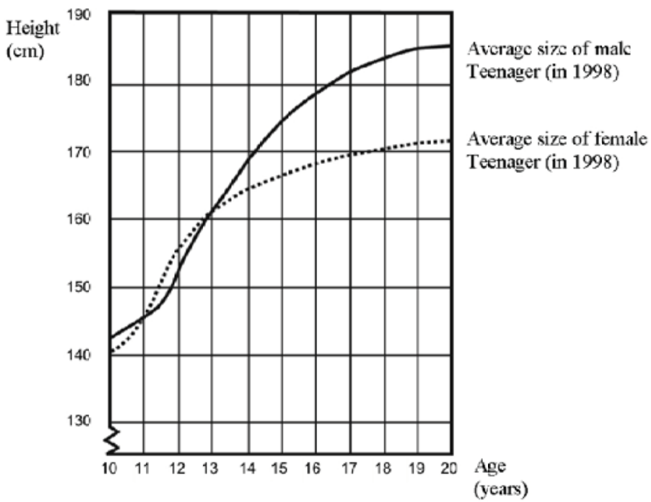


Source: Bos et al. (2003: 200f.). The more difficult the exercise, the higher test score which is measured by the scale on the left side

Figure 1b: Example exercise from PISA 2003

Teenagers are becoming taller

For 1998, the average height of male and female adolescents in the Netherlands is shown in the following graph.



Question no. 1:

Since 1980, the average size of 20-year-old women has increased by 2.3 cm to 170.6 cm. What was the average size of a 20-year-old woman in 1980?

.... cm

Question no. 2:

According to the graphs, in which period are female teenagers taller than male teenagers?

....

....

Question no. 3:

Explain how the graph shows that the growth rate for girls older than 12 years has slowed down on average.

....

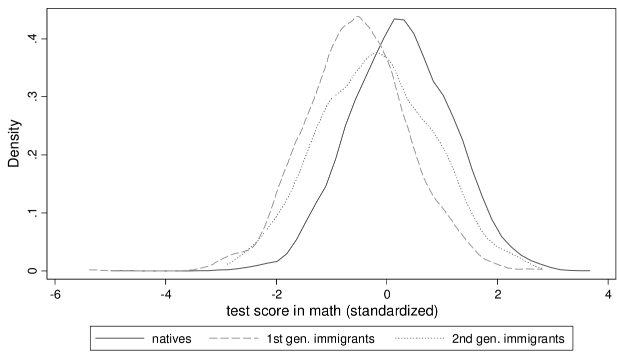
....

....

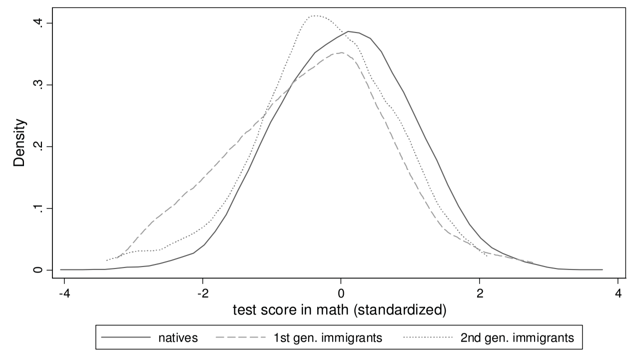
Source: http://pisa.ipn.uni-kiel.de/Aufgaben_Mathe_neu3.pdf.

Figure 2: Distribution and common support in math test scores of natives and immigrants (different schooling levels and tracks)

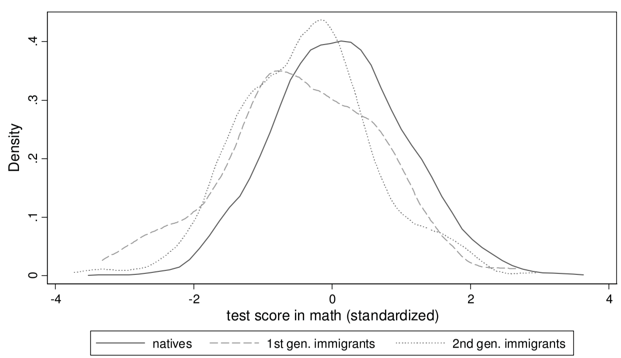
Primary Education



Gymnasium



Realschule



Hauptschule

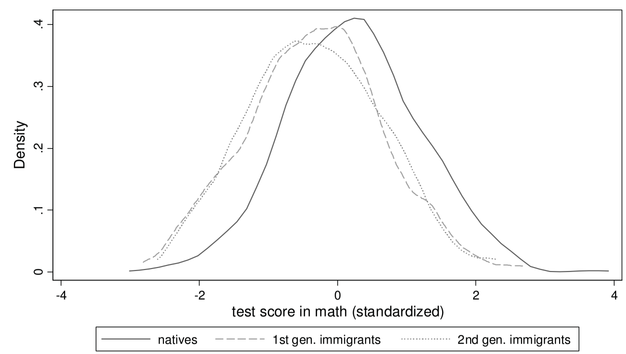
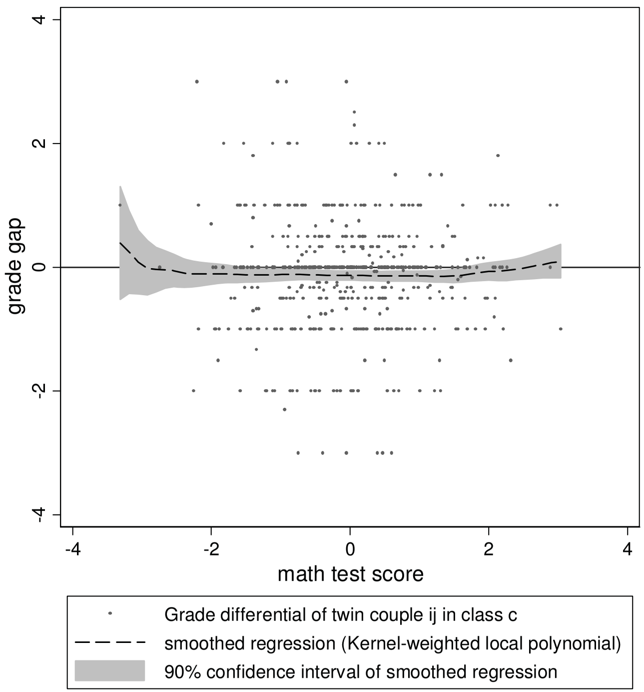
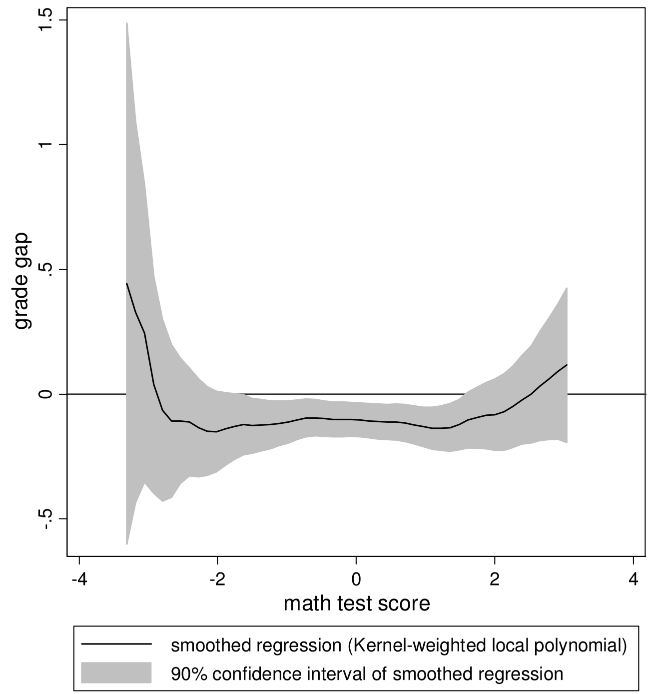


Figure 3a: Distribution of grade gaps of matched pupils (schooling level: primary education, group: first generation immigrants)

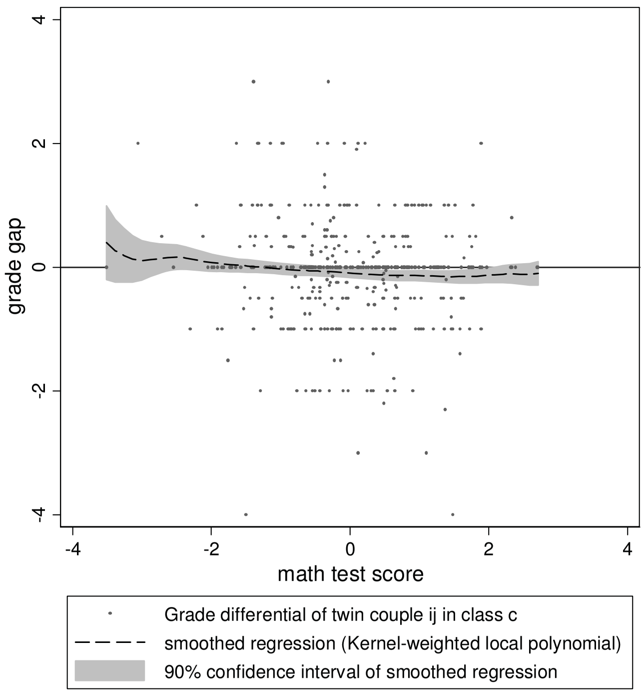




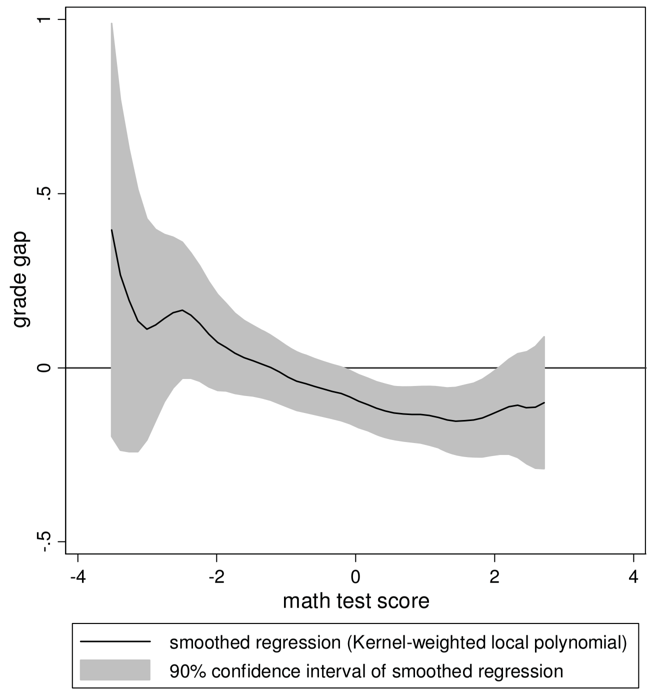
N(grade gaps): 540, N(1st gen. immigrants): 380, N(classes): 199.
(mean, sd, skewness) of grade gaps: (-0.12, 0.92, -0.06).
(5th, 25th, 50th, 75th, 95th) percentile of grade gap: (-2.00, -0.50, 0.00, 0.33, 1.00).
Used kernel: Epanechnikov, degree: 0, bandwidth: 0.59, pilot bandwidth: 0.88.

This is an enlarged section of Figure 3a. For reasons of clarity, the scatter plot has been omitted.

Figure 3b: Distribution of grade gaps of matched pupils (schooling level: primary education, group: second generation immigrants)

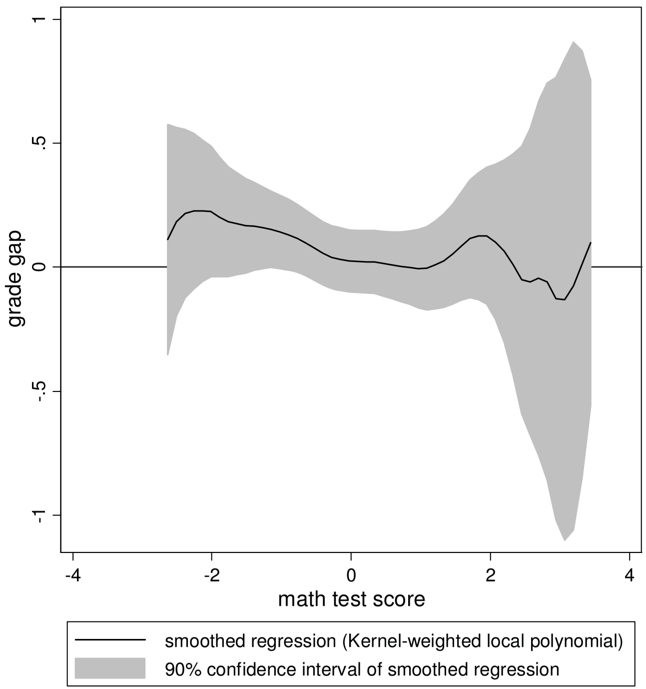


N(grade gaps): 429, N(2nd gen. immigrants): 297, N(classes): 172.
(mean, sd, skewness) of grade gaps: (-0.07, 0.88, -0.40).
(5th, 25th, 50th, 75th, 95th) percentile of grade gap: (-1.80, -0.50, 0.00, 0.33, 1.00).
Used kernel: Epanechnikov, degree: 0, bandwidth: 0.76, pilot bandwidth: 1.14.



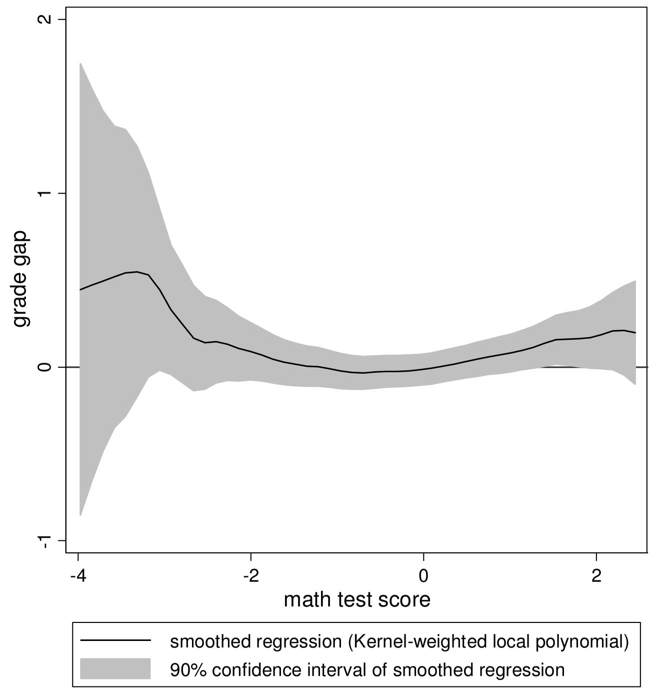
This is an enlarged section of Figure 3b. For reasons of clarity, the scatter plot has been omitted.

Figure 4a: Distribution of grade gaps of matched pupils (track: Gymnasium, group: first generation immigrants)



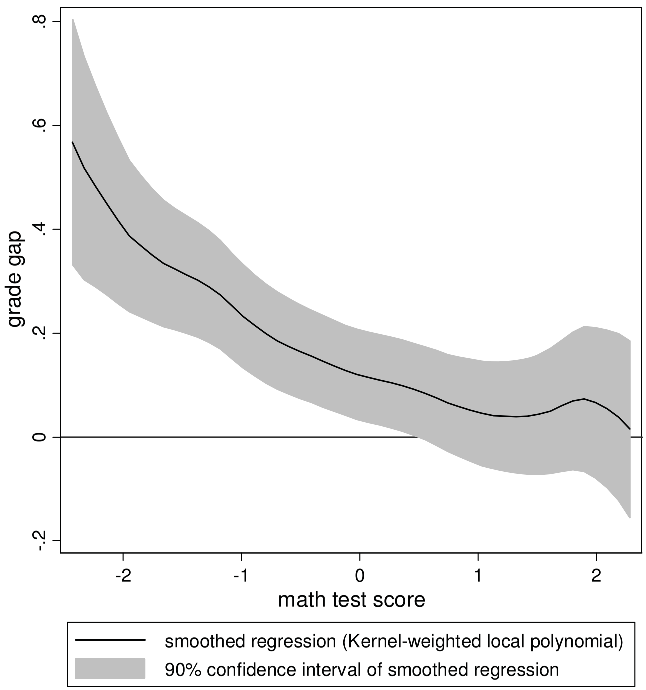
N(grade gaps): 176, N(1st gen. immigrants): 104, N(classes): 67.
(mean, sd, skewness) of grade gaps: (0.06, 0.87, -0.76).
(5th, 25th, 50th, 75th, 95th) percentile of grade gap: (-1.00, -0.33, 0.00, 0.50, 1.50).
Used kernel: Epanechnikov, degree: 0, bandwidth: 0.55, pilot bandwidth: 0.82.

Figure 4b: Distribution of grade gaps of matched pupils (track: Gymnasium, group: second generation immigrants)



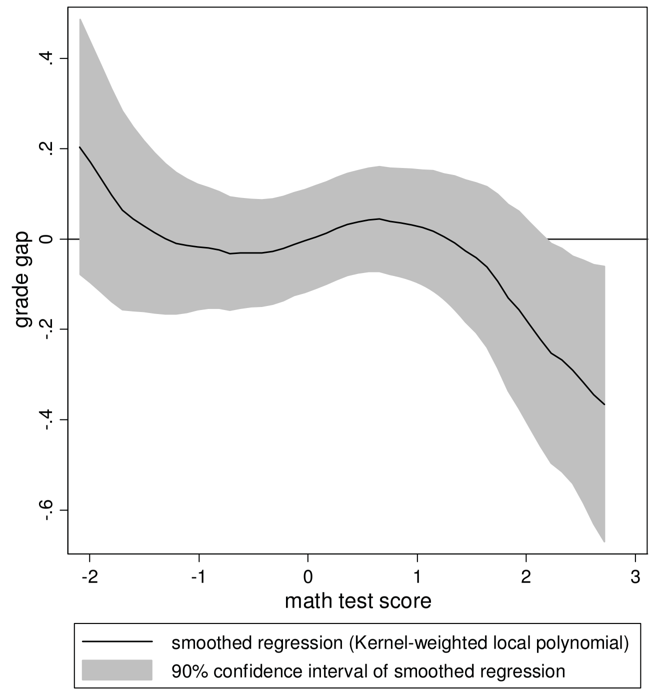
N(grade gaps): 257, N(2nd gen. immigrants): 141, N(classes): 71.
(mean, sd, skewness) of grade gaps: (0.03, 0.75, -0.11).
(5th, 25th, 50th, 75th, 95th) percentile of grade gap: (-1.00, -0.33, 0.00, 0.50, 1.00).
Used kernel: Epanechnikov, degree: 0, bandwidth: 0.60, pilot bandwidth: 0.89.

Figure 5a: Distribution of grade gaps of matched pupils (track: Realschule, group: first generation immigrants)



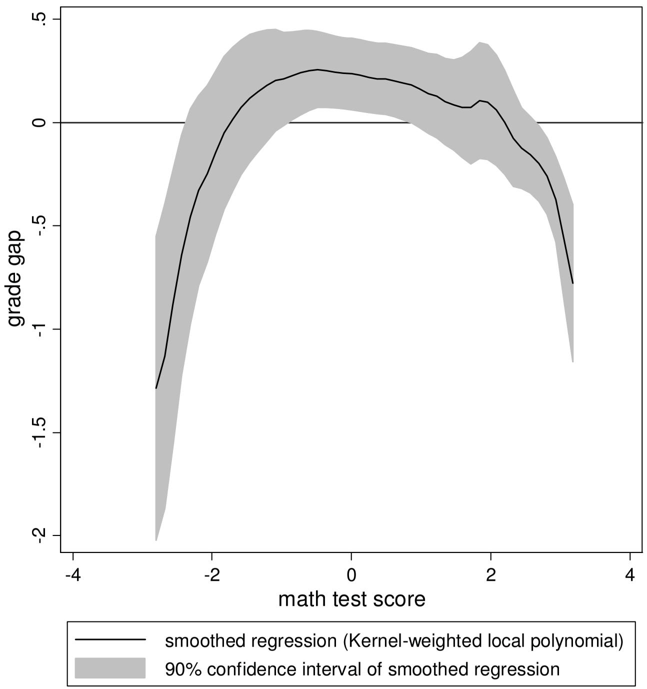
N(grade gaps): 267, N(1st gen. immigrants): 135, N(classes): 64.
(mean, sd, skewness) of grade gaps: (0.16, 0.77, 0.97).
(5th, 25th, 50th, 75th, 95th) percentile of grade gap: (-1.00, -0.33, 0.00, 0.50, 1.50).
Used kernel: Epanechnikov, degree: 0, bandwidth: 0.68, pilot bandwidth: 1.03.

Figure 5b: Distribution of grade gaps of matched pupils (track: Realschule, group: second generation immigrants)



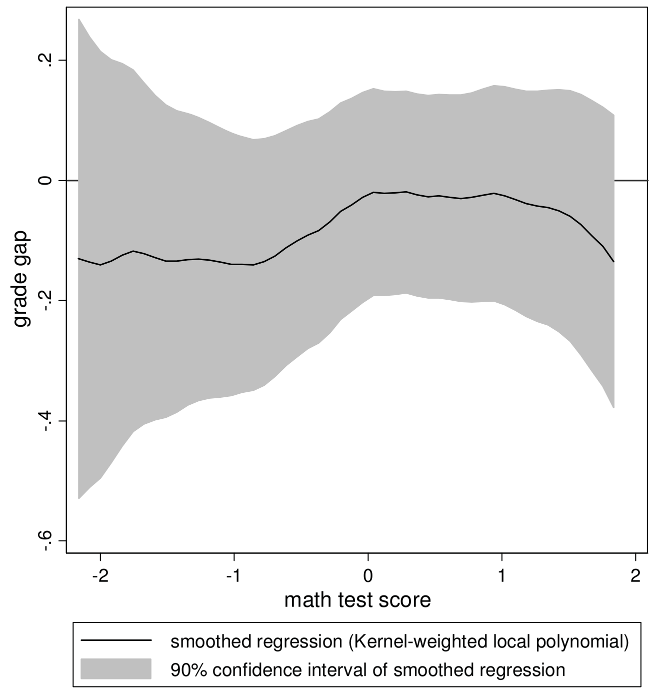
N(grade gaps): 256, N(2nd gen. immigrants): 151, N(classes): 62.
(mean, sd, skewness) of grade gaps: (0.00, 0.94, 0.06).
(5th, 25th, 50th, 75th, 95th) percentile of grade gap: (-1.00, -0.50, 0.00, 0.50, 2.00).
Used kernel: Epanechnikov, degree: 0, bandwidth: 0.56, pilot bandwidth: 0.84.

Figure 6a: Distribution of grade gaps of matched pupils (track: Hauptschule, group: first generation immigrants)



N(grade gaps): 104, N(1st gen. immigrants): 72, N(classes): 41.
(mean, sd, skewness) of grade gaps: (0.14, 0.90, -0.70).
(5th, 25th, 50th, 75th, 95th) percentile of grade gap: (-2.00, 0.00, 0.00, 0.58, 1.50).
Used kernel: Epanechnikov, degree: 0, bandwidth: 0.52, pilot bandwidth: 0.79.

Figure 6b: Distribution of grade gaps of matched pupils (track: Hauptschule, group: second generation immigrants)



N(grade gaps): 109, N(2nd gen. immigrants): 74, N(classes): 40.
(mean, sd, skewness) of grade gaps: (-0.08, 0.92, -0.17).
(5th, 25th, 50th, 75th, 95th) percentile of grade gap: (-2.00, -0.50, 0.00, 0.50, 1.50).
Used kernel: Epanechnikov, degree: 0, bandwidth: 0.60, pilot bandwidth: 0.90.

Table 1: Math grades and test scores of natives and immigrants (descriptive statistics), separated by schooling level and track

Schooling level	Primary education (PIRLS)			Secondary education (PISA)								
Track				Gymnasium			Realschule			Hauptschule		
Migration background	Native	1st gen. immigr.	2nd gen. immigr.	Native	1st gen. immigr.	2nd gen. immigr.	Native	1st gen. immigr.	2nd gen. immigr.	Native	1st gen. immigr.	2nd gen. immigr.
Math grades												
Mean	3.50	2.90*	3.06*	3.01	2.87	2.95	2.91	3.10*	2.81	2.93	2.95	2.78*
standard deviation	0.89	0.98	0.98	0.98	1.00	1.14	1.00	1.07	1.03	1.03	0.97	1.00
Math test scores												
Mean	0.25	-0.55*	-0.21*	0.06	-0.45*	-0.24*	0.11	-0.44*	-0.38*	0.23	-0.33*	-0.33*
standard deviation	0.92	0.93	1.03	0.98	1.13	1.00	0.96	1.12	0.98	0.95	0.98	0.96
Math test scores (classes)												
Mean	0.07	-0.14*	-0.10*	0.01	-0.07*	-0.09*	0.07	-0.29*	-0.21*	0.11	-0.15*	-0.17*
standard deviation	0.43	0.54	0.52	0.40	0.43	0.40	0.43	0.69	0.49	0.60	0.56	0.57
N(pupils)	2852	1006	762	2358	175	216	1774	227	264	666	223	244
N(classes)	259	253	230	120	82	88	99	74	73	80	64	68
N(schools)	259	253	230	61	51	51	50	46	43	43	41	39

* indicates whether mean grades or test scores of natives significantly (at least 5% confidence level) differ from mean grades or test scores of first or second generation immigrants attending the same schooling level or track. Used data sets: PIRLS 2001 for primary school, PISA 2003 (subsampled by school track) for the secondary schooling level. Gymnasium is the highest secondary school track , Hauptschule the lowest. Math grades range from 0 (fail) to 5 (very good). The higher a pupil's math test score the higher its math proficiency. Within a schooling level or track, math test scores have mean 0 and standard deviation 1. The variable math test scores (class level) captures the average test score of the class that is attended by a pupil. A pupil is a first generation immigrant if it was born abroad. Second (or higher) generation immigrants are born in Germany and satisfy at least one of the following conditions: (i) at least one parent was born abroad or (ii) languages other than German are (also) spoken at home. A child is treated as a native, if it is born in Germany, as well as its parents, and German is the only language spoken at home.

Table 2: Class fixed effects estimates of migration background on math grades, separated by schooling level and tracks

Schooling level	Primary	Secondary		
Track		GY	RS	HS
First generation immigrant (reference group: natives)	-0.232***	0.054	0.214*	0.238**
	(-6.35)	(0.62)	(2.56)	(3.12)
Second generation immigrant	-0.235***	0.022	0.006	0.040
	(-6.24)	(0.30)	(0.08)	(0.49)
Math test scores	0.536***	0.453***	0.398***	0.398***
	(36.09)	(22.26)	(16.12)	(8.91)
N(pupils)	4613	2749	2265	1133
N(first generation immigrants)	1004	175	227	223
N(second generation immigrants)	761	216	264	244
N(classes)	262	120	101	81
R^2 (adjusted)	0.31	0.18	0.13	0.12

Dependent variable: math grades. Significance levels: * 5%, ** 1%, *** 0.1%, depending on clustered (on class level) standard errors. Used data sets: PIRLS 2001 for primary education, PISA 2003 (subsampled by school tracks) for the secondary schooling level. GY means "Gymnasium", RS "Realschule", HS "Hauptschule". Gymnasium is the highest secondary school track, Hauptschule the lowest. Math grades range from 0 (fail) to 5 (very good). The higher a pupil's math test score the higher its math proficiency. Within a schooling level or track, math test scores have mean 0 and standard deviation 1. A pupil is a first generation immigrant if it was born abroad. Second (or higher) generation immigrants are born in Germany and satisfy at least one of the following conditions: (i) at least one parent was born abroad or (ii) languages other than German are (also) spoken at home. A child is treated as a native, if it is born in Germany, as well as its parents, and German is the only language spoken at home. All regressions include a constant (results not reported).

Table 3: Class fixed effects estimates of migration background on math grades (separated by schooling level, tracks and grading-relevant characteristics)

Schooling level	Primary education					Secondary education				
Track						Gymnasium				
Column	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
First generation immigrant	-0.232‡	-0.237‡	-0.067	-0.073*	-0.048	0.054	0.058	0.075	0.082	0.090
(reference group: natives)	(-6.35)	(-6.55)	(-1.79)	(-1.99)	(-1.33)	(0.62)	(0.70)	(0.85)	(0.97)	(0.99)
Second generation immigrant	-0.235‡	-0.224‡	-0.134‡	-0.125‡	-0.117†	0.022	0.053	0.026	0.058	0.041
	(-6.24)	(-6.09)	(-3.62)	(-3.46)	(-3.24)	(0.30)	(0.82)	(0.36)	(0.90)	(0.63)
Math test scores	0.536‡	0.518‡	0.384‡	0.368‡	0.353‡	0.453‡	0.376‡	0.429‡	0.348‡	0.359‡
	(36.09)	(34.14)	(22.36)	(21.64)	(20.26)	(22.26)	(19.94)	(20.63)	(18.26)	(17.55)
Reading test scores			0.309‡	0.306‡	0.279‡			0.050*	0.059†	0.024
			(18.62)	(18.64)	(17.05)			(2.53)	(3.19)	(1.25)
Very low oral participation		-0.271*		-0.252*	-0.269*		-1.070‡		-1.074‡	-1.082‡
(reference category: very high oral participation)		(-2.18)		(-2.06)	(-2.23)		(-14.46)		(-14.81)	(-14.92)
Low oral participation		-0.508‡		-0.487‡	-0.476‡		-0.728‡		-0.734‡	-0.745‡
		(-9.43)		(-9.22)	(-9.26)		(-11.07)		(-11.28)	(-11.35)
High oral participation		-0.242‡		-0.237‡	-0.224‡		-0.315‡		-0.317‡	-0.334‡
		(-5.63)		(-5.85)	(-5.61)		(-4.26)		(-4.34)	(-4.49)
Other pupil characteristics:	no	no	no	no	yes‡	no	no	no	no	yes‡
N(pupils)	4613	4613	4613	4613	4613	2749	2749	2749	2749	2749
N(classes)	262	262	262	262	262	120	120	120	120	120
R^2 (adjusted)	0.31	0.33	0.37	0.39	0.41	0.18	0.28	0.19	0.28	0.31

(continued)

Table 3 (*continued*)

Schooling level	Secondary education									
Track	Realschule					Hauptschule				
Column	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
First generation immigrant	0.214*	0.125	0.215*	0.132	0.072	0.238†	0.163*	0.229†	0.173*	0.133
(reference group: natives)	(2.56)	(1.59)	(2.57)	(1.67)	(0.83)	(3.12)	(2.10)	(2.91)	(2.20)	(1.48)
Second generation immigrant	0.006	-0.023	0.007	-0.012	-0.063	0.040	0.031	0.038	0.034	-0.032
	(0.08)	(-0.32)	(0.09)	(-0.17)	(-0.79)	(0.49)	(0.40)	(0.46)	(0.44)	(-0.42)
Math test scores	0.398‡	0.319‡	0.396‡	0.304‡	0.312‡	0.398‡	0.332‡	0.408‡	0.321‡	0.283‡
	(16.12)	(13.02)	(15.01)	(11.71)	(11.64)	(8.91)	(8.24)	(7.98)	(6.70)	(5.72)
Reading test scores			0.003	0.035	0.031			-0.022	0.026	0.039
			(0.12)	(1.51)	(1.24)			(-0.52)	(0.66)	(1.00)
Very low oral participation		-1.103‡		-1.110‡	-1.098‡		-0.772‡		-0.779‡	-0.752‡
(reference category: very high oral participation)		(-13.99)		(-14.08)	(-14.02)		(-7.78)		(-7.55)	(-7.40)
Low oral participation		-0.713‡		-0.721‡	-0.717‡		-0.608‡		-0.612‡	-0.579‡
		(-10.94)		(-10.95)	(-10.98)		(-7.34)		(-7.25)	(-6.88)
High oral participation		-0.286‡		-0.291‡	-0.287‡		-0.174*		-0.175*	-0.165
		(-4.56)		(-4.65)	(-4.59)		(-2.06)		(-2.07)	(-1.96)
Other pupil characteristics:	no	no	no	no	yes*	no	no	no	no	yes*
N(pupils)	2265	2265	2265	2265	2265	1133	1133	1133	1133	1133
N(classes)	101	101	101	101	101	81	81	81	81	81
R^2 (adjusted)	0.13	0.25	0.13	0.25	0.26	0.12	0.20	0.12	0.20	0.21

Dependent variable: math grades. Significance levels: * 5%, † 1%, ‡ 0.1%, depending on clustered (on class level) standard errors. Used data sets: PIRLS 2001 for primary education, PISA 2003 (subsampled by school tracks) for the secondary schooling level. Gymnasium is the highest secondary school track, Hauptschule the lowest. Math grades range from 0 (fail) to 5 (very good). The higher a pupil's math test score the higher its math proficiency. Within a schooling level or track, math test scores have mean 0 and standard deviation 1. A pupil is a first generation immigrant if it was born abroad. Second (or higher) generation immigrants are born in Germany and satisfy at least one of the following conditions: (i) at least one parent was born abroad or (ii) languages other than German are (also) spoken at home. A child is treated as a native, if it is born in Germany, as well as its parents, and German is the only language spoken at home. Regarding their oral participation, pupils were asked to which extent they agree with the statement "I frequently participate in math". This categorical variable has the four outcomes "I completely disagree", "I somewhat disagree", "I somewhat agree" and "I completely agree". It is decomposed into four dummy variables. In the fifth columns, other pupil characteristics (age, sex, preschool attendance and parental education) are additionally taken into account. All regressions include a constant and control for missing values in reading skills, oral participation and other pupil characteristics (results not reported).