

IZA DP No. 9800

The Educational Consequences of Language Proficiency for Young Children

Yuxin Yao
Asako Ohinata
Jan C. van Ours

March 2016

The Educational Consequences of Language Proficiency for Young Children

Yuxin Yao

CentER, Tilburg University

Asako Ohinata

*University of Leicester
and CentER, Tilburg University*

Jan C. van Ours

*CentER, Tilburg University,
University of Melbourne, CEPR, CESifo, CREAM and IZA*

Discussion Paper No. 9800
March 2016

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Educational Consequences of Language Proficiency for Young Children*

This paper studies the educational consequences of language proficiency by investigating the relationship between dialect-speaking and academic performance of 5-6 year old children in the Netherlands. We find that dialect-speaking has a modestly negative effect on boys' language test scores. In addition, we study whether there are spillover effects of peers' dialect-speaking on test scores. We find no evidence for spillover effect of peers' dialect-speaking. The test scores of neither Dutch-speaking children nor dialect-speaking children are affected by the share of dialect-speaking peers in the classroom.

JEL Classification: J24, I2

Keywords: dialect-speaking, test scores, spillover effects, language, academic performance

Corresponding author:

Jan C. van Ours
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands
E-mail: vanours@uvt.nl

* We are grateful to DANS for making the PRIMA data available for this paper. We would also like to thank Chris Wallace, Olivier Marie, Jan Kabatek, Antonio Di Paolo and participants at seminars in Tilburg, Maastricht and Tinbergen Institute, SOLE-EALE 2015, ESPE 2015, ENTER jamboree 2015, EEA 2015 and the Hitotsubashi Summer Institute on Labor Economics and the editors and anonymous reviewers for helpful comments.

1 Introduction

The economic consequences of language proficiency have received increasing attention in recent years. Language skills are viewed as part of human capital and play an important role in labor market performance, schooling, health care, consumption and investment (see an overview in Chiswick and Miller (2014)). The existing literature predominantly studies the topic in the context of immigration. These studies focus on how the proficiency in local languages contributes to adult immigrants' labor market performance and social integration in host countries (Chiswick and Miller, 1995; Dustmann and van Soest, 2001; Dustmann and Fabbri, 2003; Bleakley and Chin, 2004, 2010; Yao and van Ours, 2015). Other papers investigate the educational performance of immigrant children (Dustmann et al., 2010; Geay et al., 2013).

This paper contributes to the literature on the educational consequences of language skills but instead of studying immigrant students, it investigates the effects of speaking dialects at home on standardized test scores. We refer dialects as variations of the standard language across regions and regional languages in a country. Since immigrants speak different languages from natives, they are the obvious choice of group for studying the effects of language. However, these immigrant students do not only differ from native students in terms of the spoken language but they also have different socio-economic and cultural backgrounds. As a result, the estimated effects in previous papers are likely to reflect the combined influences of linguistic as well as cultural differences. In contrast, dialect-speaking students share a relatively homogeneous background to those who speak the standard language of the country. In this sense, our estimates are likely to capture purer language effects.

To our knowledge, our paper is the first to study the effects of speaking dialects on academic performance. Aside from the advantage of investigating the impacts of language proficiency exploiting dialect-speaking behavior, studying the economic consequences of dialect is important on its own. Dialects are an integral part of daily communication and widespread in many countries. Nonetheless, the existing economic literature on dialect is scarce. Grogger (2011, 2014) reports that non-standard speech patterns like African American and Southern American dialects are associated with lower wages in the US labor market. Similarly, Gao and Smyth (2011) find a significant wage premium associated with fluency in standard Mandarin for dialect-speaking migrating workers in

China. These papers present evidence on labor market performance; a natural extension to the literature is to study how dialect-speaking affects students' academic performance.

Besides labor market performance at adulthood, the return to language skills can be traced back to the accumulation of human capital at early stages of life. A few recent studies investigate how skills in local languages are related to academic performance of immigrant students although the evidence is still limited. Using the UK National Pupil Database and the Millennium Cohort Study, Dustmann et al. (2010) find that immigrant students in the UK lag behind native students at the beginning of primary school. This gap is smaller for students whose mother tongue is English. Moreover, the gap diminishes throughout primary and secondary schooling process and this is particularly prominent again among immigrants whose mother tongue is English. Their data, however, do not have information on parental education and therefore they cannot rule out the possibility that those who already speak English fluently come from highly educated families. Geay et al. (2013), on the other hand, study whether non-English speaking students affect native students' academic performance. They also use the National Pupil Database and present findings that non-English speaking immigrants often sort themselves into schools with more academically disadvantaged native students. Once they control for self-selection into schools, they report that there is no negative spillover effect from immigrants to native students.

As a case study, we choose the Netherlands to investigate the effects of dialect-speaking on education. There are three main reasons for our choice of country. Firstly, there exist multiple regional dialects in the Netherlands with varying degrees of linguistic distances to Standard Dutch. This variation allows us to study effectively the impact of language. Secondly, compared with other countries, native residents in the Netherlands are relatively more homogeneous in terms of ethnicity, culture, and even economic wealth. As a result, we are more likely to be able to pick out the effect of language without the influences of other socio-economic and demographic characteristics. Thirdly, despite the existence of various dialects in the Netherlands, Standard Dutch is predominantly used in school teaching, even in regions where the position of the local dialect is strong (see Cheshire et al. (1989) for an overview). In such a learning environment, dialect-speaking may pose negative effects on academic performance, especially on children's linguistic development. Furthermore, it may be more costly for these dialect-speaking children to interact with Standard Dutch speakers and teachers.

Our data from the PRIMA survey for Dutch primary schools provide us with a unique set of information collected from 5 to 6 year-old primary school children, their parents, as well as the school directors. In particular, the data include crucial information identifying those students who speak dialects at home. In addition, we also observe their test scores and the classroom and school level characteristics.

In order to identify the effects of dialect-speaking on test scores, we estimate a linear function with individual and classroom variables. In addition, we control for school fixed effects to take account of potentially endogenous selection of students into schools. We separately estimate the dialect effects on language and math test scores. This is because linguistic disadvantages faced by dialect-speaking students may affect language scores more and we may, as a result, find heterogeneous effects across subjects. Indeed, we only find a negative impact of dialect-speaking on language test scores. In addition, dialect-speaking boys suffer more from learning in a language that is different from their daily language at home. Our point estimates may be affected by omitted variables such as the ability of some dialect-speaking students to adopt Standard Dutch faster. Such an ability is likely to affect not only the rate of learning Standard Dutch, but also his/her academic performance. Therefore, we conduct a sensitivity analysis by including math test scores in our regression as a measure of individual ability. Our results are robust to this alternative specification.

The fact that dialect-speaking students, who suffer academically, share the learning environment with those who speak Standard Dutch raises a further question: would classmates' speaking patterns affect academic performance? Although spillover effects in the classroom have gained much attention among social scientists in the last decade, we are the first to explicitly focus on speaking patterns. Spillover effects in the dialect-speaking setting can occur through several channels. First, there could be a negative spillover from dialect speakers to other students. Students may learn non-standard pronunciation and/or grammar from dialect speakers. The negative effects may be more evident in language skills than in other subjects if this were the case. Second, although the use of Standard Dutch is encouraged by teachers and used by the majority of students in Dutch primary schools, informal verbal interaction in dialects can occur between dialect speakers, thus segregating students into a dialect-speaking group and a Dutch-speaking group. Insufficient in-class interaction may harm the academic performance of all students across all subjects. Therefore, these channels suggest negative consequences of studying with

dialect-speaking peers for all students. However, having more dialect-speaking peers in the same class may be beneficial for dialect speakers. That is, with more peers speaking the same dialect, dialect speakers can have interactions with peer and conduct discussions at a lower cost. To sum up, we conjecture that more dialect-speaking classmates will impede learning for Dutch speakers, but the effects will be ambiguous for dialect speakers. Therefore, spillover effects must be investigated separately for each group.

It is well known that the identification of spillover effects suffers from a number of econometric difficulties. There may be parental selection as to where their children attend schools, so that peers' mother tongue is endogenously determined by school choice. Studies using non-experimental data attempt to mitigate the bias from self-selection by exploiting exogenous variation in the composition of peers across classes or schools (Hanushek et al., 2003; Ammermueller and Pischke, 2009; Lavy et al., 2012). The study on spillover effects in European primary schools by Ammermueller and Pischke (2009) is close to our paper in terms of identification strategy. They control school fixed effects and exploit exogenous variation in class composition in the same grade. In order to test whether classrooms are formed randomly with respect to a particular student characteristic, they perform a Pearson χ^2 test. In addition, Ohinata and van Ours (2013) introduce a check for random allocation for schools with two classes in one grade. Based on the difference in the number of immigrants between two classes, they can compare the observed distribution of schools with the simulated distribution obtained from random allocation. We adopt the two tests, which suggest that dialect-speaking students are randomly allocated in schools. Relying on the variation in the share of dialect-speaking peers across classrooms in one grade and across cohorts in the same school, we find no evidence of spillover effects from peers' dialect-speaking on test scores. We conduct a range of sensitivity checks, which all suggest that our findings are robust.

To summarize, the contribution of this paper to the literature is threefold. Firstly, it adds to the scarce literature on the educational consequences of language skills. We focus on the effects of speaking Dutch dialects on individual test scores as well as the spillover effects of speaking dialects on classmates' test scores. Secondly, our paper is the first to investigate the socio-economic consequences of dialect-speaking in the context of education. Last, but not the least, we present a purer estimate of the impact of language proficiency, which is less likely to be affected by other factors such as cultural and socioeconomic differences.

The rest of the paper is organized as follows: Section 2 introduces language usage in the Netherlands. Section 3 describes the PRIMA data and presents some stylized facts. Section 4 examines the effect of dialect-speaking on test scores. Section 5 presents the baseline results for spillover effects from peers' dialect-speaking and a wide range of sensitivity checks for these spillover effects. Section 6 concludes.

2 Languages and Dialects in the Netherlands

The predominantly spoken language of the Netherlands is Standard Dutch, originating in the urban areas of Noord-Holland, Zuid-Holland and Utrecht. Besides Standard Dutch, the regional languages and dialects spoken in the Netherlands are remarkably diverse, including Frisian, Limburgish, and Low Saxon. Frisian, mostly spoken in the province of Friesland, is recognized as a separate language and promoted by the local government. In Friesland both Standard Dutch and Frisian are considered official languages and more than 80% of the adult inhabitants understand verbal Frisian. It is also an official subject and can be an instruction medium in the primary schools except for some exempted schools in Friesland. Other regional languages include Limburgish and Low Saxon, which enjoy the status as “official regional languages” in related regions although there is no clear regulation regarding government support. Limburgish is spoken in the province of Limburg by about 75% of the inhabitants and Low Saxon is spoken in the provinces of Groningen, Drenthe, Overijssel and Gelderland by approximately 60% of the inhabitants. However, neither Limburgish nor Low Saxon are taught as a school subject. Other provinces also have dialects such as Brabantish, spoken in Noord-Brabant or Zeelandic in Zeeland (see an overview in Driessen (2005) and Cheshire et al. (1989)). Standard Dutch is the only instruction medium in most provinces in the Netherlands.

Table 1 summarizes the linguistic distances between Standard Dutch and various dialects and regional languages (Van Bezooijen and Heeringa, 2006).¹ We use the Levenshtein distance which is based on a comparison of the pronunciation of typical words in one dialect and in Standard Dutch.² The distance is determined by the minimum

¹From here on, for simplicity, we refer to all dialects and separate regional languages used in the Netherlands as dialects. Frisian, Limburgish and Low Saxon are recognized as official regional languages by the European Charter for Regional or Minority Languages. Frisian has the highest level of recognition, which means that the Frisian provincial government may be more active in promoting the use of Frisian. Low Saxon and Limburgish have a lower level of recognition. See Driessen (2005) for more details.

²The Levenshtein distance is introduced by Kessler (1995). For an overview of linguistic distances

number of steps to change a particular word in one language to the same word in another languages. The more steps it requires, the farther away these two languages are considered to be. As shown in Table 1, Frisian stands out from the other dialects by having the largest Levenshtein distance, i.e. 37, followed by Limburgish. As a point of comparison, the Levenshtein distance between Standard Dutch and English is reported to be 63 (Ispording and Otten, 2013). Therefore, the distances between some Dutch dialects and Standard Dutch are likely to be non-negligible.

3 Data and Background

3.1 PRIMA data

Our analysis employs data from PRIMA, a large-scale biannual longitudinal survey for primary schools in the Netherlands. The project was conducted by the Institute for Applied Social Sciences in Nijmegen and the SCO-Kohnstamm Institute in Amsterdam from 1994 to 2005. The survey enrolls students in the second, fourth, sixth and eighth grade from 6 cohorts and over 600 schools, covering 10% of the relevant age population. It provides rich information on Dutch primary education, documenting test scores, school and class characteristics, and demographic information.

We select a cross-sectional sample of native students in the second grade. We drop immigrants from our sample and treat all non dialect-speaking students as Dutch-speaking.³ Language information is only collected from parents of second graders. We are not able to follow later grades because of attrition in survey sampling at individual level. Moreover, we exclude the first two cohorts from the sample, because math scores and teacher characteristics are made comparable across cohorts only after the 1998/1999 survey.

In the questionnaires, parents indicate in what language their child speaks to his/her mother, father, siblings and friends. Our data also reports the language spoken between the two parents. Each parent was asked to report one of three categories: Standard Dutch, dialects or Frisian, and other foreign languages. We consider a student to be dialect-speaking if he or she speaks dialect or Frisian to either mother or father. The main independent variable is whether a student speaks a Dutch dialect or Standard Dutch at

and their use in economics see Ginsburgh and Weber (2015). Refer to Appendix A for details.

³We define students as immigrants if at least one parent was born outside of the Netherlands.

home to parents.⁴ Our main variables of interest are the measures for educational outcomes. The PRIMA survey provides standardized tests on language and math for second grade children.⁵ The language test for children in second grade measures understanding of words and concepts such as first, last, many and few. The math test focuses on the sorting of objects, comparing numbers and counting.

After exploring the direct effect of speaking dialect on students' test scores, we investigate the spillover effects in classrooms. We calculate the share of dialect-speaking peers at the class level excluding oneself, based on the unrestricted sample including immigrant students and missing observations. This indicator measures how intensely students interact with dialect-speaking peers. We will impose further restrictions on the sample for the spillover effects analysis. For example, we drop classes with less than five students to obtain more accurate shares and remove schools without any dialect speaker in the relevant grade as outliers. The resulting sample consists of 411 schools from four cohorts.

3.2 Summary statistics

Standard Dutch is predominantly used in Dutch primary schools. According to our sample from the PRIMA survey, 10.1% of students in the second grade are reported to speak dialects to parents at home nationally. The survey provides location information of schools by 12 Dutch provinces. Table 1 summarizes the share of dialect-speaking students by province. The distribution of dialect-speaking students is very heterogeneous across provinces. Limburg is the province with the largest share of dialect-speaking students, 53%. It is followed by Friesland where as many as 39% of the students speak the regional language with parents. In Drenthe and Zeeland, there are around 18% of dialect-speaking students. On the other hand, less than 2% of students speak dialects at home in the provinces of Noord-Holland, Utrecht and Zuid-Holland where modern Standard Dutch originated (see also the percentage of dialect-speaking students by province in the Appendix, Figure A2).⁶

⁴We do not take into account the language spoken between siblings or friends because it may be affected by school education. The group of students speaking dialects with fathers highly overlaps with those who speak dialects with their mothers or siblings.

⁵To enable a comparison between language scores and math scores, we also normalize test scores for each subject. We divide the differences between each of the PRIMA test scores and the average test score by its standard deviation. The normalized scores, therefore, have a mean of 0 and standard deviation of 1.

⁶The share of dialect-speaking children in PRIMA survey is lower than that of dialect-speaking population. There is a considerable decrease in the use of dialects between generations and across years

Table 2 presents the means of our variables separately by language group and gender. To begin with, dialect speakers have lower test scores on both language and math compared to Dutch speakers, although the gap between the two groups is modest. Girls have higher test scores than boys regardless of whether they speak dialects or not. Dutch-speaking girls, therefore, are the most advantaged group, while dialect-speaking boys have the lowest average scores. Secondly, there is not much difference between dialect speakers and Dutch speakers in individual characteristics, such as gender, age, family composition, and whether one has always stayed in the Netherlands. However, dialect speakers are much more likely to have dialect-speaking parents than Dutch speakers. Parents of around 89% of the dialect-speakers and only around 10% of the Dutch speakers use dialects at home, indicating that the language spoken by students in the second grade is predominantly determined by their parents' language usage. Also, dialect speakers are more likely to have parents with lower educational attainment than Dutch speakers. The proportion of parents with university or higher degrees is around 10 percentage points higher for Dutch speakers than for dialect speakers. Thirdly, we find no significant difference in teacher and school characteristics between the two groups, except that dialect-speaking students are much more likely to attend schools in less urbanized areas. Intuitively, dialects play a more important role in daily interaction in less urbanized areas where population mobility is low. Finally, girls and boys have very similar characteristics in both language groups. Summarizing Table 1 and Table 2, speaking dialects associates with lower test scores and lower education level of parents; dialect speakers are mainly from less urbanized areas and certain provinces.⁷

Before we move on to our main analysis, Figure 1 compares kernel density distributions of the test scores of Dutch-speaking students and dialect-speaking students. The top and bottom graphs present language scores and math scores, respectively. Graph (a) presents the distribution of language scores by gender and language group. For both boys and girls, there is not much difference for the two language groups in the shape of distribution. But obviously dialect speakers have lower language scores than Standard Dutch speakers, especially for boys. From Graph (b), we find that the distributions of math scores for two language groups are more overlapping than those of the language scores. There is

(Driessen, 2005).

⁷Note that immigrant children have parents with much lower education attainment. For example, while only one-third of Standard-Dutch speaking children have fathers with a degree of lower secondary or lower, this is about 50% for dialect-speaking children and 75% for immigrant children.

hardly any difference between the two language groups, and this is true for both genders. Figure 1 only suggests a weak association between dialect-speaking and language scores at the individual level.

We present further similar kernel density distribution graphs of test scores in Figure 2. In these figures, we compare the test score distribution of classes with a high versus a low share of dialect-speaking students. The “high share” classes are defined to be those with the average share, i.e. 22%, or more. Similarly, “low share” classes are those with less than 22%. It is clear from the upper graph that differences in Dutch speakers’ language scores are limited with respect to the share of dialect speakers. The lower graph presents the same pattern for math scores. Figure 2 suggests that the share of dialect speakers at the class level does not seem to explain the differences in test scores for Dutch speakers. The figure also presents the distribution of test scores for dialect speakers. For both language and math scores, the distributions are similar between classes with a high share of dialect speakers and those with a low share. However, dialect speakers from high-share classes seem to have slightly higher math scores.⁸

4 Dialect-speaking and Test Scores

In this section we examine the relationship between dialect-speaking and academic performance. As a baseline, we estimate the following model using OLS:

$$Y_{ics,t} = X_{ics,t}^T \beta + \delta D_{ics,t} + \alpha_s + \gamma_t + \varepsilon_{ics,t} \quad (1)$$

where $Y_{ics,t}$ denotes the normalized test scores for student i in class c and school s at year t . $D_{ics,t}$ equals one if the student speaks a dialect to parents. $X_{ics,t}$ is a vector of all individual characteristics and teacher characteristics. α_s denotes the school fixed effects and γ_t indicates the year fixed effects. Finally, $\varepsilon_{ics,t}$ is the error term.

The key variable of interest is the dummy variable for speaking dialects at home, $D_{ics,t}$. A negative coefficient implies that dialect-speaking students perform worse in

⁸In the Appendix, we also plot average test scores against the share of dialect speakers at the class level. Considering students’ self-selection into schools, we demean the average scores and the shares at the class level relative to those at the school level. In Figure A3, the scatter plots and fitted lines indicate how test scores are correlated to the share of dialect speakers. For both Dutch speakers and dialect speakers, we find the fitted lines are flat. There is no correlation between the share of dialect speakers and average language or math scores.

the respective test. We control for a set of individual characteristics (age in months, age squared, gender, a dummy for presence of both parents, a dummy for always having stayed in the Netherlands, dummies for number of children at home, dummies for father’s and mother’s level of education) and a set of class characteristics reported by teachers (teacher’s gender, teacher’s year of experience, number of students in a class, a dummy for whether the class supports teaching combined with other grades, a dummy for whether the class supports remedial teachers, and dummies for different shares of immigrant students in a class). Students may choose schools based on their socio-economic status as well as their spoken language. Therefore, we control for school fixed effects in order to correct for the potential bias that arises from such self-sorting into schools. As indicated in Table 2, speaking dialect is correlated with parents’ education. In particular, dialect-speaking students typically have less educated parents, which in turn may affect students’ test scores. Therefore, we control for educational attainment of mothers as well as fathers.

Table 3 presents the OLS estimates for the effects of speaking dialects on language and math test scores. The parameter estimates are reported separately for boys and girls to account for gender-specific effects. In the first column, only year fixed effects are included. In subsequent columns, we gradually include individual characteristics, teacher characteristics, and school fixed effects in order to investigate how these variables affect our estimates.

In the first column of Panel (a), we find a significantly negative effect of speaking dialects on language scores for both boys and girls. When we add more control variables, the relevant parameter estimates decrease because dialect-speaking is correlated to background variables. Moreover, when we introduce school fixed effects to remove the endogeneity of school choice, we find that speaking dialect with parents significantly decreases boys’ language scores by 0.079 standard deviations but has no effect on girls’ language scores. The gender-specific effects imply that boys and girls may have different trajectories of language development. At the age of 5, girls may be better at adapting to the new language environment than boys.⁹ Panel (b) presents the dialect-speaking effects on math scores. Irrespective of whether we add school fixed effects or not, we do

⁹In fact, linguists have long pointed out the superiority of girls in language performance over boys. Murray et al. (1990) present evidence that girls begin talking earlier whilst Roulstone et al. (2002) show that girls accumulate vocabulary faster than boys. The differences in language performance are observed as early as 2-3 years old and into school years (e.g. Lynn (1992); Bornstein et al. (2000); Dionne et al. (2003)). Based on these findings, it may not be so far-fetched to argue that girls adapt better to a new language at the age of 5 or 6 and outperform boys in language tests.

not find any significant effect once we control for individual characteristics. Since we use the normalized scores, we can compare the magnitude of the estimated effects between subjects. Clearly, the effect of dialect-speaking on math scores is smaller and less significant than the effect on language scores. Moreover, the negative effects of speaking dialects are present only for boys. To conclude, we find a penalty of dialect speaking on academic performance, but this only applies to boys' language scores, which is consistent with our preliminary findings in Figure 1.

It is possible that our results so far suffer from an omitted variable problem, since we may have failed to control for individual unobserved abilities. A student's own ability is likely to affect not only the rate of learning Standard Dutch, but also his academic performance. Therefore, we conduct a sensitivity analysis to investigate this issue by presenting additional results in which we include students' math test score as an explanatory variable in the language score regression. The assumption is that there is no direct effect of dialect-speaking on math scores. The math test is targeted at basic mathematical concepts and reasoning and reflects the innate ability of the children in the second grade. Panel (c) of Table 3 shows that there is a significantly negative effect of dialect-speaking on boys' language performance even after we take account of their math scores on top of other covariates. The parameter estimate of -0.066 is not substantially different from the parameter estimate in the first row of Panel (a). Introducing the math score as an explanatory variable does not influence the magnitude of the effect of dialect-speaking on language scores but increases the precision of the estimate. Also similarly to Panel (a), the effect on girls is insignificantly different from zero. All in all, this robustness check supports the idea of a possible causal effect of dialect-speaking on the language skills of children.

It is important to note that we have so far assumed that children who do not speak dialects at home can speak Standard Dutch properly. However, this is not necessarily the case. Children who are considered to speak Standard Dutch are still likely to be affected by dialects spoken by their peers in the region. In such a case, the estimated gap in language test scores between dialect speakers and Dutch speakers may be underestimated. As a result, our estimates may present a lower-bound of the actual effects.

Table 4 presents parameter estimates for sensitivity analysis. As a point of reference, Panel (a) repeats the baseline estimates of Table 3 column (4) when all characteristics and fixed effects are controlled. Panel (b) shows the effect of children's dialect-speaking

if we include a dummy variable for whether parents speak dialects between each other. The negative effects of dialect-speaking for boys' language scores increase in size while the other parameter estimates are not affected. Apparently, after we control for parental dialect speaking, the negative effects of dialect-speaking for boys' language scores are stronger. In Panel (c) we use an alternative measure of dialect-speaking status, whether children speak dialects at home to both parents, rather than to either of them. This is to test whether a stronger exposure to a dialect at home affects our results. However, we find effects that are in line with the baseline estimates. Panel (d) shows the parameter estimates if we exclude the observations from Friesland, the province with a regional language that is the farthest from Standard Dutch and enjoys high social status. Compared to Panel (a) the negative effect on boys' language scores also exist in other provinces. As discussed in Section 2, the linguistic distance to Standard Dutch differs across Dutch dialects. If it is the case that speaking dialects indeed affects students' language performance, we should observe that the negative effect of dialect-speaking is more prevalent among students whose dialect is farther away from Standard Dutch. In Panel (e), we use the interaction term between dialect dummy and the linguistic distance measure as the explanatory variable. The parameter estimates in Panel (e) suggest that an increase of 10 units in the linguistic distance will cause a decrease in dialect-speaking boys' language scores by 0.025 standard deviations. Our results indicate that the penalty of speaking dialects on boys' language scores increases with linguistic distance. Linguistic distance, however, does not affect girls' language scores or math scores in general.

5 Spillover Effects of Dialect-speaking

5.1 Set-up of the analysis

In this section we investigate the spillover effects of speaking dialect on peers' performance. As a proxy for the intensity of students' communication in dialects within a classroom, we calculate the share of dialect-speaking peers relative to the total number of peers in each class.¹⁰ We rescale the shares in the range from 0 to 1 for convenience

¹⁰Around 30% of the observations did not report whether they speak dialects. This is because parents of these students did not report the language information. In Section 5 we use the non-missing sample to calculate the share of dialect-speaking peers, assuming that the share of dialect speakers is identical between the missing sample and the non-missing sample. For further checks, we alternatively defined the share by treating all missing observations either as dialect speakers or as Dutch speakers. However,

of interpretation. As discussed in Section 3, we refine the sample to schools with at least one dialect speaker in one grade. We also drop observations from classes with fewer than 5 students in order to obtain a more precise measure of the shares. These procedures result in the native sample of 9,411 individuals from 411 schools and 1,091 classes.

As discussed in Section 1, we want to investigate whether peers' speaking behavior affects test scores of Dutch-speaking students differently from dialect-speaking students. Motivated by the literature on peer effects, we apply the linear-in-mean model. Our main explanatory variable is the share of dialect-speaking peers at the class level, calculated over all students in the class and excluding the student of interest. We estimate using the following OLS equation:

$$Y_{ics,t} = X_{ics,t}^T \beta + \lambda \overline{D_{(-i)cs,t}} + \alpha_s + \gamma_t + \varepsilon_{ics,t} \quad (2)$$

where $Y_{ics,t}$ denotes test scores for student i in class c and school s at year t . $\overline{D_{(-i)cs,t}}$ denotes the share of dialect-speaking peers at the class level, excluding individual i . $X_{ics,t}$ is a vector of all individual and teacher characteristics. We control for school fixed effects, α_s to remove the variation explained by school choice. Furthermore, γ_t are year fixed effects, indicating that we control for the cohort-specific shocks in test scores. $\varepsilon_{ics,t}$ is the error term.

In order to obtain unbiased and consistent OLS estimates in Equation (2), we rely on two assumptions. The first assumption is that the dialect-speaking students are randomly allocated across cohorts within a school. The second assumption requires that dialect speakers are randomly allocated into classes if there are two or more classes within a grade in a school. For example, if school directors intentionally allocate more dialect speakers to classes with students of disadvantaged background, the negative spillover effect of dialect speakers will be overestimated. The validity of the tests for these assumptions are discussed in Section 5.2.

A remaining concern regarding the consistency of OLS estimates is that students may change their language usage at home after attending primary schools, so that the share of dialect-speaking peers is endogenous. For example, dialect-speaking students of higher ability may start using Standard Dutch at home because of exposure to Dutch-speaking classmates and teachers. We assume that the language spoken between children and

our results are robust regardless of the definition used. Results are available upon request.

parents is persistent for young children. We believe that our assumption is valid, since it is unlikely that parents switch from speaking dialects to Standard Dutch when their children have attended primary schools for only one year.

5.2 Random allocation of dialect speakers across classes

We argue that the share of dialect-speaking peers in a class is exogenously determined if dialect speakers are randomly assigned into different classes in a grade as well as across cohorts within a school. This implies that the OLS estimate of λ in Equation (2) represents a causal spillover effect of dialect-speaking after we control for school fixed effects.

For schools with one single class in the second grade, we rely on the idiosyncratic variation in the share of dialect-speaking peers across cohorts in the particular school to identify the causal spillover effects. For schools where there are more than one class in the second grade, we rely on the random allocation of students between classes as well as across cohorts.¹¹ We assume that students are randomly allocated to each cohort within the same school. In addition, we assume that the allocation of students to classes within the same cohort in a school is also randomly determined. There are two reasons why we believe that the former assumption is likely to hold. Firstly, educational disadvantages experienced by dialect-speaking students have not drawn wide attention in the Netherlands. As a result, it is very unlikely that Dutch parents would explicitly avoid schools with high shares of dialect-speaking students in previous cohorts or change the decision when to attend primary school. Secondly, school administrators and parents usually have no information about the dialect usage of peer students prior to enrollment. Therefore, admission of students should not depend on whether they speak a dialect or Standard Dutch. Since each school director can identify whether a particular student in his or her school is speaking dialects once the student is admitted, allocation of students into second-grade classes may depend on students' spoken language. Therefore, we conduct the following two tests to examine the latter assumption, that students are randomly allocated into classes within the same cohort and school.

First, we perform the Pearson χ^2 test suggested by Ammermueller and Pischke (2009).

¹¹In our sample, we have 640 observations from 411 schools and 4 cohorts. Out of these schools, 378 school observations have a single class, 155 have two classes, 54 have three classes, 38 have four classes and 15 have five or more classes.

If the allocation of students is random, the characteristics of each student should be independent of the characteristics of other students in the class. For each school s , we define n_{cj} as the actual number of students in classroom $c = 1, 2, \dots, C_s$ with the subscript $j = 0, 1$, denoting Dutch-speaking and dialect-speaking respectively. We compute the predicted number \hat{n}_{cj} of dialect speakers and Dutch speakers in any classroom as $\frac{n_c \cdot n_j}{\sum_c \sum_j n_{cj}}$. Then the Pearson test statistic for any school is given as $\sum_c \sum_j \frac{(n_{cj} - \hat{n}_{cj})^2}{\hat{n}_{cj}}$ and follows a χ^2 distribution with $(C_s - 1)(J - 1)$ degrees of freedom. When further assuming that the allocation of students to each school is independent of any other, we can aggregate the Pearson test statistics. This aggregate follows a χ^2 distribution with $\sum_s (C_s - 1)(J - 1)$ degrees of freedom. Based on our sample of schools with multiple classes, the aggregate Pearson test statistic is 405.99 with 453 degrees of freedom. Given the p-value of 0.95, we cannot reject the null hypothesis of random allocation.

Second, as suggested by Ohinata and van Ours (2013), we can formally test whether dialect speakers are randomly assigned to different classes in schools with 2 classes in the grade. Some schools allocate equal number of dialect-speaking students to each class whilst other schools allocate different numbers of dialect speaking students to each class. Let us indicate the difference in the number of dialect-speaking students across classes within the same school as n , where $n = 0, 1, 2, \dots, 13$. We then plot a frequency distribution of the schools with difference n in the number of dialect-speaking students from our sample. Next, we plot a simulated frequency distribution of schools with difference n in the number of dialect-speaking students under the assumption that these students are randomly allocated to classes (See more details in Ohinata and van Ours (2013)). If students are indeed randomly allocated, we should observe that the actual distribution from our data to be statistically comparable to the simulated distribution. This is shown in Figure 3. The p-value for the F-test on the difference between the two distributions is 0.83, indicating that we cannot reject the null hypothesis.

In addition to the random allocation of dialect speakers into classes, we also need to ensure that the allocation of teaching resources are uncorrelated with the share of dialect speakers in a class. That is, the share of dialect speakers should not be correlated with other class level variables which may determine academic performance. In Table 5, we regress the share of dialect speakers on teacher characteristics and average background characteristics at the class level. We also add school fixed effects and year fixed effects. The first column presents estimates based on the entire sample. The second column

includes estimates based on schools with multiple classrooms per cohort. Irrespective of the samples, we find that all control variables except for the share of girls are uncorrelated with the share of dialect speakers. Either average characteristics or teaching resources are jointly significant, suggesting that these control variables cannot explain the share of dialect speakers at the class level.

5.3 Baseline results

Table 6 presents the estimated effects of peers' dialect-speaking on the test scores of Dutch speakers and dialect speakers separately. This is to investigate the potentially heterogeneous spillover effects, which may depend on the language spoken by the affected students. We include the individual characteristics, teacher characteristics, school fixed effects and year fixed effects as we did in the previous section. The dependent variable in each regression is the test scores after normalization with zero mean and standard deviation of 1. The independent variable of interest is the share of dialect-speaking peers with a range from 0 to 1. From Column 1 to 4 in Panel (a), we find that the share of dialect-speaking peers has no significant effect on Dutch speakers' language scores. Irrespective of including control variables and school fixed effects, all parameter estimates are very small and insignificant. This suggests that the share of dialect-speaking peers in a class does not have significant spillover effects among Standard Dutch-speaking students. Similarly in Panel (b), we find that there is no significant spillover effect on Dutch speakers' math scores. In conclusion, our results show that providing a common learning environment for dialect as well as Standard Dutch speakers does not negatively affect the academic performance of students who speak Standard Dutch.

It is also of interest to investigate whether having more dialect-speaking peers in the same class/cohort benefits or harms dialect speakers themselves. In Table 6, we also report the estimated spillover effects for dialect speakers. In Column 1 of Panel (a), we regress language scores only on the share of dialect-speaking peers and find modestly positive effects. When the share of dialect-speaking peers increases by 10 percentage points, the language score of dialect speakers will increase by 0.019 standard deviations. However, when we control for individual characteristics, teacher characteristics, school fixed effects and year fixed effects, the significant spillover effects disappear. This is most likely because the positive correlation between the share and the test score of dialect-

speaking students is explained by individual background and school choice. Similarly, Panel (b) indicates that there is no spillover effect on dialect speakers' math scores when more control variables are added. Interaction with more classmates speaking the same dialect neither benefits nor harms dialect speakers.

As discussed before, the spillover effects of dialect speakers on the majority Dutch speakers are expected to be negative for at least two reasons. First, dialect speakers are usually from more disadvantaged background and they may have difficulty in studying the standard language. According to the peer effects literature, disadvantaged students may have negative effects on classmates' performance in general. Second, with many dialect speakers there could be linguistic segregation between groups of students in a class. This would make in-class interaction costly. However, the spillover effects on the minority dialect speakers are ambiguous. Besides the potentially negative effects discussed above, more peers speaking the same language can help dialect speakers to integrate in class, leading to positive effects on performance. Our findings, however, support none of these arguments. We find no spillover effects for both groups and both subjects. The negative spillover from dialect speakers can be more or less offset by curriculum in Dutch and sufficient exposure with Dutch speakers.

5.4 Sensitivity checks

Table 7 presents a series of sensitivity checks for the causal spillover effects of peers' dialect-speaking. Panel (a) of Table 7 reports regression results separately for schools with multiple classes and schools with one single class in the second grade. For multiple-class schools, we are using variation both across classes and across cohorts. For these schools, we find that the relevant parameter estimates are larger in size than the baseline estimates but still insignificant. The parameters presented in the second row in Panel (a) are estimated using a sample of schools with a single classroom and therefore based on across cohort variation. We find that the parameter estimates change substantially due to sample selection, but the estimates are still small and only marginally significant.

In Panel (b), we control for the peers' average individual background characteristics. In the language of Manski (1993), this allows us to account for the contextual effects. This is also to control for the correlation between peers' speaking pattern and their characteristics. According to the estimates, peers' speaking pattern has no effect on

academic performance.

Panel (c) reports the spillover effects separately between classes with less or more than 22% dialect-speaking students. Our hypothesis is that negative spillover effects may dominate in classes with more dialect speakers, since there is more linguistic segregation and less exposure to Standard Dutch. However, in both high-share and low-share dialect-speaking classes, dialect-speaking peers have little influence on academic performance, except for a negative spillover effect on dialect speakers' language scores at the 10% level.

We also investigate whether spillover effects differ between boys and girls. We regress the test scores of two groups separately on the share of dialect-speaking peers in the whole class. As shown in Panel (d), the spillover effects are small and insignificant for both boys and girls, although the point estimates are generally larger for boys.

In Panel (e), we check heterogeneous spillover effects across provinces. We use the interaction of the share of dialect-speaking peers and the linguistic distance of dialect (divided by 10) as the main explanatory variable, and find no spillover effect no matter how far a dialect is from Standard Dutch.

Finally, the number of dialect-speaking peers in a class, instead of the share, is used as a measure for exposure to dialect speakers. The estimated parameters in Panel (f) also suggest that there are no significant spillover effects.

6 Conclusion

This paper investigates the importance of language skills on academic performance among young children in the Netherlands. In contrast to the existing literature which study this issue by focusing exclusively on immigrant children, we explore the impacts of dialect-speaking on educational achievement. To our knowledge, this is the first paper that addresses this issue by focusing on non-immigrant children. In addition, it is the first to study the consequences of dialect-speaking in the context of education.

Using the unique Dutch dataset PRIMA, we study the effect of dialect-speaking on individual test scores. In addition, we examine the potentially negative spillover effects of peers' dialect-speaking on the academic performance of classmates.

We find that dialect-speaking at home is strongly correlated with parental usage of dialects. It also correlates with province of residence, family background and urbanization level of the location of schools. We find that speaking dialect has a modest penalty on

language skills for boys and this penalty increases as the linguistic distance between the students' dialect and Standard Dutch increases. We find no significant effect on language skills for girls. Dialect-speaking does not seem to affect math skills either for boys or for girls. Assuming that dialect-speaking does not affect math skills and including math scores to control for individual ability, we still find a negative effect of dialect-speaking on language scores for boys. For girls, we find no such effect.

We also study the spillover effects of peers' dialect-speaking on academic performance of Standard Dutch and dialect-speaking children. In order to deal with the endogeneity problem of spillover effects, we rely on random allocation of dialect speakers across classes in one grade and idiosyncratic variation of the share of dialect speakers across cohorts in one school. In the linear-in-mean model, we estimate individual test scores on the share of dialect-speaking peers at the class level. Neither of the two groups experienced spillover effects and this finding is robust to several sensitivity checks.

The fact that boys are negatively affected by speaking dialect can be related to different trajectories of language development between young boys and girls. We hypothesize that the dialect penalty on boys' academic performance would disappear at later stages of life. In fact, the findings from the existing literature indicate that dialect-speaking may even have positive effects in the long run. For example, the existing literature on bilingual systems suggest that there exists a wage premium for being proficient in an official regional language or a dialect (Christofides and Swidinsky, 2010; Di Paolo and Raymond, 2012; Chen et al., 2014). Speaking a regional language can signal social identity, increase search efficiency and decrease transaction cost in the local labor market. Unfortunately, we cannot investigate whether or when boys catch up with girls in language performance, since our dataset suffers from severe attrition across grades. Nor can we examine the long-run impact on their labor market performance. In fact, it is important to keep in mind that our finding of negative dialect-speaking effects only applies to the performances of Standard Dutch and math in the short run. However, this is an interesting and important question that deserves further attention in the future. Our results also suggest that it is unlikely that speaking dialects would be detrimental to classmates. The finding indicates that imposing interventions such as discouraging children to learn dialects or track students by mother tongue is likely to be unnecessary.

In order to put our findings in the broader context of the educational consequences of language skills, we draw findings from a comparable study on immigrant students in the

Netherlands (Ohinata and van Ours, 2012). Whilst we find that the dialect-speaking boys lag behind their counterparts in language by 0.08 standard deviations, first-generation immigrant students in the Netherlands perform worse in reading tests in comparison to native students by 0.28 standard deviations, nearly 4 times more than that of dialect-speaking students. The difference in the magnitude is likely partially a result of the lack of cultural differences experienced by dialect-speaking students. It is also possible that the linguistic barriers that dialect-speaking students experience are much less than those experienced by first-generation immigrant students.

Given this, it is probably safe to say that it is not at all surprising that we do not find any spillover effect from dialect-speaking students to Dutch-speaking students. Even though immigrant students are in a much more disadvantaged position compared to dialect-speaking students, Ohinata and van Ours (2013) find no spillover effect from immigrants to natives. Of course, our study cannot rule out the possibility that the educational consequences of language would be non-negligible in countries where the linguistic barriers experienced by dialect speakers or non-native speakers in general are larger. We leave this question for future research.

References

- Ammermueller, A. and J. S. Pischke (2009). Peer effects in european primary schools: Evidence from the progress in international reading literacy study. *Journal of Labor Economics* 27, 315–348.
- Bleakley, H. and A. Chin (2004). Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and Statistics* 86(2), 481–496.
- Bleakley, H. and A. Chin (2010). Age at arrival, english proficiency, and social assimilation among U.S. immigrants. *American Economic Journal: Applied Economics* 2(1), 165–192.
- Bornstein, M. H., O. Haynes, K. M. Painter, and J. L. Genevro (2000). Child language with mother and with stranger at home and in the laboratory: A methodological study. *Journal of Child Language* 27(02), 407–420.
- Chen, Z., M. Lu, and L. Xu (2014). Returns to dialect identity exposure through language in the chinese labor market. *China Economic Review* 30, 27–43.
- Cheshire, J., V. Edwards, H. Munstermann, and B. Weltens (1989). *Dialect and Education: Some European Perspectives*. Multilingual Matters Ltd.
- Chiswick, B. R. and P. W. Miller (1995). The endogeneity between language and earnings: International analyses. *Journal of Labor Economics* 13(2), 246–288.
- Chiswick, B. R. and P. W. Miller (2014). International migration and the economics of language. In B. R. Chiswick and P. W. Miller (Eds.), *Handbook of the Economics of Immigration*, pp. 211–269. Elsevier.
- Christofides, L. N. and R. Swidinsky (2010). The economic returns to the knowledge and use of a second official language: English in quebec and french in the rest-of-canada. *Canadian Public Policy* 36, 137–158.
- Di Paolo, A. and J. L. Raymond (2012). Language knowledge and earnings in Catalonia. *Journal of Applied Economics* 15(1), 89–118.
- Dionne, G., P. S. Dale, M. Boivin, and R. Plomin (2003). Genetic evidence for bidirectional effects of early lexical and grammatical development. *Child Development* 74(2), pp. 394–412.
- Driessen, G. (2005). In Dutch? Usage of Dutch regional languages and dialects. *Language, Culture and Curriculum* 18, 271–285.
- Dustmann, C. and F. Fabbri (2003). Language proficiency and labour market performance of immigrants in the UK. *Economic Journal* 113(489), 695–717.
- Dustmann, C., S. Machin, and U. Schonberg (2010). Ethnicity and educational achievement in compulsory schooling. *Economic Journal* 120, 272–297.
- Dustmann, C. and A. van Soest (2001). Language fluency and earnings: Estimations with misspecified indicators. *Review of Economics and Statistics* 83(4), 663–674.
- Gao, W. and R. Smyth (2011). Economic returns to speaking 'standard Mandarin' among migrants in China's urban labour market. *Economic of Education Review* 30, 342–352.

- Geay, C., S. McNally, and S. Telhaj (2013). Non-native speakers of english in the classroom: What are the effects on pupil performance? *Economic Journal* 123, 281–307.
- Ginsburgh, V. and S. Weber (2015). Linguistic distances and their use in economics. CEPR Discussion Paper 10640.
- Gooskens, C. and W. Heeringa (2004). Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data. *Language Variation and Change* 16, 189–207.
- Grogger, J. (2011). Speech patterns and racial wage inequality. *Journal of Human Resources* 46, 1–25.
- Grogger, J. (2014). Speech and wages. *University of Chicago*.
- Hanushek, E. A., J. F. Kain, J. M. Markman, and S. G. Rivkin (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics* 18, 527–544.
- Heeringa, W. and J. Nerbonne (2001). Dialect areas and dialect continua. *Language Variation and Change* 13, 375–400.
- Isphording, I. E. and S. Otten (2013). The costs of Babylon? Linguistic distance in applied economics. *Review of International Economics* 21(2), 354–369.
- Kessler, B. (1995). Computational dialectology in Irish Gaelic. In *Proceedings of the 7th Conference of the European Chapter of the Association for Computational Linguistics*.
- Lavy, V., M. D. Paserman, and A. Schlosser (2012). Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *Economic Journal* 122, 208–237.
- Levenshtein, V. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Cybernetics and Control Theory* 10, 707–710.
- Lynn, R. (1992). Sex differences on the differential aptitude test in british and american adolescents. *Educational Psychology* 12(2), 101–102.
- Manski, C. F. (1993). Identification of social endogenous effects: the reflection problem. *Review of Economic Studies* 60, 531–542.
- Murray, A. D., J. Johnson, and J. Peters (1990). Fine-tuning of utterance length to preverbal infants: Effects on later language development. *Journal of Child Language* 17(03), 511–525.
- Nerbonne, J., W. Heeringa, E. Hout, P. van den Kooi, S. Otten, and W. van de Vis (1996). Phonetic distance between Dutch dialects. In G. Durieux, W. Daelemans, and S. Gillis (Eds.), *Papers from the Sixth CLIN Meeting*, pp. 185–202.
- Ohinata, A. and J. C. van Ours (2012). Young immigrant children and their educational attainment. *Economics Letters* 116(3), 288–290.
- Ohinata, A. and J. C. van Ours (2013). How immigrant children affect the academic achievement of native Dutch children? *Economic Journal* 123, 308–326.
- Roulstone, S., S. Loader, K. Northstone, and M. Beveridge (2002). The speech and language of children aged 25 months: Descriptive data from the avon longitudinal study of parents and children. *Early Child Development and Care* 172(3), 259–268.

- Smakman, D. (2012). The definition of the standard language: a survey in seven countries. *International Journal of the Sociology of Language* 218, 25–58.
- Swadesh, S. (1952). Lexico-statistic dating of prehistoric ethnic contacts. *Proceedings of the American Philosophical Society* 96, 121–137.
- Valls, E., J. Prokic, E. Clua, and M. Lloret (2012). Applying the levenshtein distance to catalan dialects: a brief comparison of two dialectometric approaches. *Verba* 39, 35–61.
- Van Bezooijen, R. and W. Heeringa (2006). Intuitions on linguistic distance: geographically or linguistically based? In T. Koole, J. Nortier, and B. Tahitu (Eds.), *Vijfde sociolinguïstische conferentie*, pp. 77–87. Eburon Uitgeverij BV.
- Wieling, A., S. Montemagni, J. Nerbonne, and R. H. Baayen (2014). Lexical differences between Tuscan dialects and standard Italian: Accounting for geographic and sociodemographic variation using generalized additive mixed modeling. *Language* 90, 669–692.
- Yao, Y. and J. C. van Ours (2015). Language skills and labor market performance of immigrants in the Netherlands. *Labour Economics* 34, 76–85.

TABLE 1: LINGUISTIC DISTANCES AND THE SHARE OF DIALECT-SPEAKING STUDENTS IN PRIMA

Province	Linguistic Distance	Dialect Speakers (%)	Number of Students
Drenthe	19	18.8	756
Flevoland	12	3.3	332
Friesland	37	39.0	1,346
Gelderland	28	2.8	2,801
Groningen	28	10.0	807
Limburg	32	53.0	1,816
Noord-Brabant	28	3.0	3,811
Noord-Holland	12	1.4	4,104
Overijssel	29	7.2	1,276
Utrecht	18	1.4	767
Zeeland	29	18.6	825
Zuid-Holland	12	0.4	3,573
Total	22.5	10.1	22,214

Source: Van Bezooijen and Heeringa (2006) and PRIMA Survey.

Note: In Column 1, the linguistic distances between various dialects spoken in each province and Standard Dutch are shown. The larger the value of the index, the more distant a dialect is from Standard Dutch. The weighted average linguistic distance is 22.5. In Column 2 and 3, we present the share of dialect-speaking students based on our sample for analysis. The shares are very similar based on the combined sample of native students and immigrants in the second grade.

TABLE 2: MEANS OF VARIABLES BY LANGUAGE GROUP AND GENDER

	Dutch speakers		Dialect speakers	
	Boys	Girls	Boys	Girls
Test scores				
Language scores (standard deviation)	-0.086 (0.99)	0.124 (0.99)	-0.219 (0.95)	0.015 (0.99)
Math scores (standard deviation)	-0.022 (1.01)	0.051 (0.99)	-0.152 (0.93)	0.013 (1.05)
Individual Characteristics				
Complete family (%)	97.2	98.1	99.1	99.5
Age in months	69.2	68.6	69.3	69.0
Always stay in the Netherlands (%)	98.1	98.1	99.1	99.5
Number of children at home	2.4	2.4	2.5	2.4
Dialect-speaking between parents (%)	9.9	11.0	89.3	90.6
Father's education (%)				
Lower secondary school or lower	33.2	32.4	47.8	48.5
Upper secondary school	35.8	36.3	33.8	34.7
University or higher	25.8	25.4	15.7	13.9
Not available	6.2	5.9	2.8	2.9
Mother's education (%)				
Lower secondary school or lower	28.5	28.8	41.6	44.1
Upper secondary school	43.4	43.4	44.7	42.6
University or higher	21.9	21.8	11.2	10.5
Not available	6.2	6.0	2.6	2.8
Teacher and school characteristics				
Female teacher (%)	97.9	98.1	96.5	96.5
Year of teaching	16.3	16.3	17.8	18.4
Combining class (%)	75.0	75.2	66.0	65.4
Remedial class (%)	76.9	75.8	72.5	74.0
Number of students	15.9	15.8	17.5	17.3
Share of immigrants in a class	18.0	17.7	15.7	16.4
Urbanization of location of school (%)				
Not urban	21.9	21.8	34.2	32.3
Little urban	24.8	25.9	36.0	37.4
Moderately urban	23.2	21.7	16.5	16.7
Very urban	23.3	24.3	13.2	13.5
Extremely urban	6.8	6.3	0.0	0.0
Number of Obs.	10,607	9,942	1,225	1,045

Note: The table presents the average statistics based on the sample of native students from 4 cohorts in the second grade. The test scores are normalized such that for the full sample the mean is 0 and the standard deviation is 1.

TABLE 3: EFFECT OF DIALECT-SPEAKING ON TEST SCORES

	(1)	(2)	(3)	(4)
a. Language scores				
Boys	-0.123*** (0.034)	-0.053 (0.032)	-0.064** (0.032)	-0.079** (0.036)
Girls	-0.102*** (0.036)	-0.024 (0.034)	-0.037 (0.034)	-0.008 (0.042)
b. Math scores				
Boys	-0.128*** (0.035)	-0.053 (0.033)	-0.059* (0.033)	-0.026 (0.040)
Girls	-0.038 (0.039)	0.048 (0.038)	0.029 (0.038)	0.034 (0.043)
c. Language scores				
Boys	-0.054** (0.027)	-0.026 (0.027)	-0.035 (0.027)	-0.066** (0.031)
Girls	-0.081*** (0.028)	-0.048* (0.028)	-0.052* (0.028)	-0.024 (0.038)
Individual characteristics	N	Y	Y	Y
Teacher characteristics	N	N	Y	Y
School fixed effects	N	N	N	Y
Year fixed effects	Y	Y	Y	Y

Note: The dependent variables are normalized scores. The independent variable of interest in each regression is a dummy which equals 1 if the student speaks a dialect to his/her father or mother at home. In Panel (a) and (b), we control for individual characteristics, teacher and class characteristics, school fixed effects and year fixed effects. In Panel (c), all regressions include the math scores as an independent variable. The number of boys in our sample is 11,832 and the number of girls is 10,987. Absolute t-statistics, which are based on the clustered standard errors at the class level, are shown in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE 4: PARAMETER ESTIMATES SPEAKS DIALECT AT HOME;
SENSITIVITY ANALYSIS

	(1)	(2)	(3)	(4)
Variables	Language scores		Math scores	
	Boys	Girls	Boys	Girls
a. Baseline results				
Speaking dialects at home	-0.079** (0.036)	-0.008 (0.042)	-0.026 (0.040)	0.034 (0.043)
b. Including dialect-speaking parents as an explanatory variable				
Speaking dialects at home	-0.124*** (0.040)	-0.035 (0.047)	-0.070 (0.046)	0.011 (0.048)
c. Using speaking dialects to both parents as a measure				
Speaking dialects at home	-0.071* (0.040)	-0.006 (0.046)	-0.018 (0.043)	0.037 (0.045)
d. Excluding Friesland				
Speaking dialects at home	-0.085* (0.044)	-0.022 (0.053)	-0.043 (0.049)	0.036 (0.053)
e. Province heterogeneous effects				
Speak dialects at home \times Linguistic distance (/10)	-0.025** (0.012)	-0.005 (0.014)	-0.009 (0.013)	0.005 (0.014)

Note: The dependent variables are normalized scores. Panel (a) repeats the parameter estimates of Table 3 column (4). In Panel (c) the independent variable is speaking dialects at home to both parents rather than either parent as an alternative measure. In Panel (e), the independent variable of interest is the interaction of the dummy for speaking a dialect at home and linguistic distance of the dialect divided by 10. In all panels, we include individual characteristics, teacher and class characteristics, school fixed effects and year fixed effects. Absolute t-statistics are based on the clustered standard errors at the class level and are shown in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE 5: RANDOM ASSIGNMENT OF TEACHING RESOURCES AND DUTCH-SPEAKING STUDENTS; SHARE OF DIALECT-SPEAKING STUDENTS IN THE CLASSROOM

	(1)	(2)
	All	Multiple classes
Share of girls in a class	0.054** (0.026)	0.040 (0.027)
Average age in month	-0.180 (0.241)	-0.176 (0.243)
Share of students from complete families	-0.040 (0.042)	-0.235 (0.042)
Share of students who always stay in the Netherlands	0.015 (0.053)	0.008 (0.056)
Average number of children	0.140 (0.138)	0.034 (1.378)
Average father's education level	-0.882 (1.700)	-0.585 (1.801)
Average mother's education level	0.745 (1.700)	-0.009 (1.903)
Teacher is female	-0.185 (3.260)	0.124 (3.308)
Teacher's year of experience	0.009 (0.053)	0.001 (0.053)
Combining class	1.791 (2.189)	0.737 (3.270)
Remedial class	-1.013 (1.298)	0.151 (1.460)
Number of students	-0.001 (0.155)	-0.143 (0.192)
Share of immigrant students: 10-30%	0.326 (1.341)	0.562 (1.460)
Share of immigrant students: 30-50%	-3.164 (1.954)	-3.910 (2.060)
Share of immigrant students: 50-70%	-2.733 (3.349)	-1.852 (4.033)
Share of immigrant students: 70-100%	-1.358 (3.164)	-0.481 (3.382)
F-statistics for average characteristics	0.86	0.50
F-statistics for teaching resources	0.74	1.05
Number of classrooms	1,093	717
Number of schools	411	182

Note: The dependent variable is the share of dialect-speaking students in a class. All regressions are at the class level with year fixed effects and school fixed effects. In Column 2, we only use the sample from schools with multiple classes in the second grade. Absolute t-statistics are based on the clustered standard errors at the class level and are shown in parenthesis. All the estimates include year fixed effects. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE 6: SPILLOVER EFFECTS OF DIALECT-SPEAKING ON TEST SCORES

	(1)	(2)	(3)	(4)
a. Language scores				
Dutch speakers	0.020 (0.082)	0.031 (0.078)	-0.018 (0.077)	-0.047 (0.166)
Dialect speakers	0.191** (0.090)	0.038 (0.083)	-0.000 (0.086)	-0.192 (0.213)
b. Math scores				
Dutch speakers	0.000 (0.091)	0.014 (0.087)	-0.003 (0.088)	-0.192 (0.189)
Dialect speakers	0.195* (0.101)	0.057 (0.097)	0.017 (0.098)	-0.049 (0.202)
Individual characteristics	N	Y	Y	Y
Teacher characteristics	N	N	Y	Y
School fixed effects	N	N	N	Y
Year fixed effects	Y	Y	Y	Y

Note: The dependent variable is normalized scores. The independent variable of interest is the share of dialect-speaking peers in a class. In all panels, we include individual characteristics, teacher and class characteristics, school fixed effects and year fixed effects. The number of Dutch speakers is 7,149 and the number of dialect speakers is 2,262. Absolute t-statistics are based on the clustered standard errors at the class level and are shown in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE 7: SENSITIVITY CHECKS: ESTIMATES OF SPILLOVER EFFECTS

	(1)	(2)	(3)	(4)
	Language scores		Math scores	
	Dutch speakers	Dialect speakers	Dutch speakers	Dialect speakers
a. Multiple-class and single-class samples				
Multiple-class	-0.118 (0.187)	-0.359 (0.248)	-0.266 (0.200)	-0.275 (0.198)
Single-class	0.225 (0.338)	0.139 (0.474)	0.270 (0.353)	0.771* (0.394)
b. Controlling peers' background characteristics				
Whole sample	-0.034 (0.164)	-0.123 (0.215)	-0.197 (0.188)	-0.037 (0.203)
c. High share and low share samples				
High share class	-0.056 (0.354)	0.037 (0.188)	-0.711* (0.370)	0.142 (0.216)
Low share class	0.012 (0.285)	0.439 (2.080)	0.231 (0.273)	-0.781 (2.270)
d. Gender heterogeneous effects				
Boys	-0.088 (0.238)	-0.417 (0.306)	-0.347 (0.231)	-0.216 (0.242)
Girls	-0.130 (0.232)	-0.018 (0.288)	-0.070 (0.250)	0.150 (0.314)
e. Independent variable: Share of dialect-speaking peers interacts with linguistic distance				
Whole sample	-0.030 (0.057)	-0.066 (0.065)	-0.088 (0.066)	-0.016 (0.065)
f. Independent variable: Number of dialect-speaking peers				
Whole sample	-0.005 (0.017)	0.002 (0.019)	-0.021 (0.025)	-0.024 (0.024)

Note: The dependent variable is normalized scores, except for Panel (f). The independent variable of interest is the share of dialect-speaking peers in a class, except for Panel (e). In all panels, we include individual characteristics, teacher and class characteristics, school fixed effects and year fixed effects. Absolute t-statistics are based on the clustered standard errors at the class level and are shown in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

FIGURE 1: DISTRIBUTION OF TEST SCORES BY LANGUAGE SPOKEN AT HOME

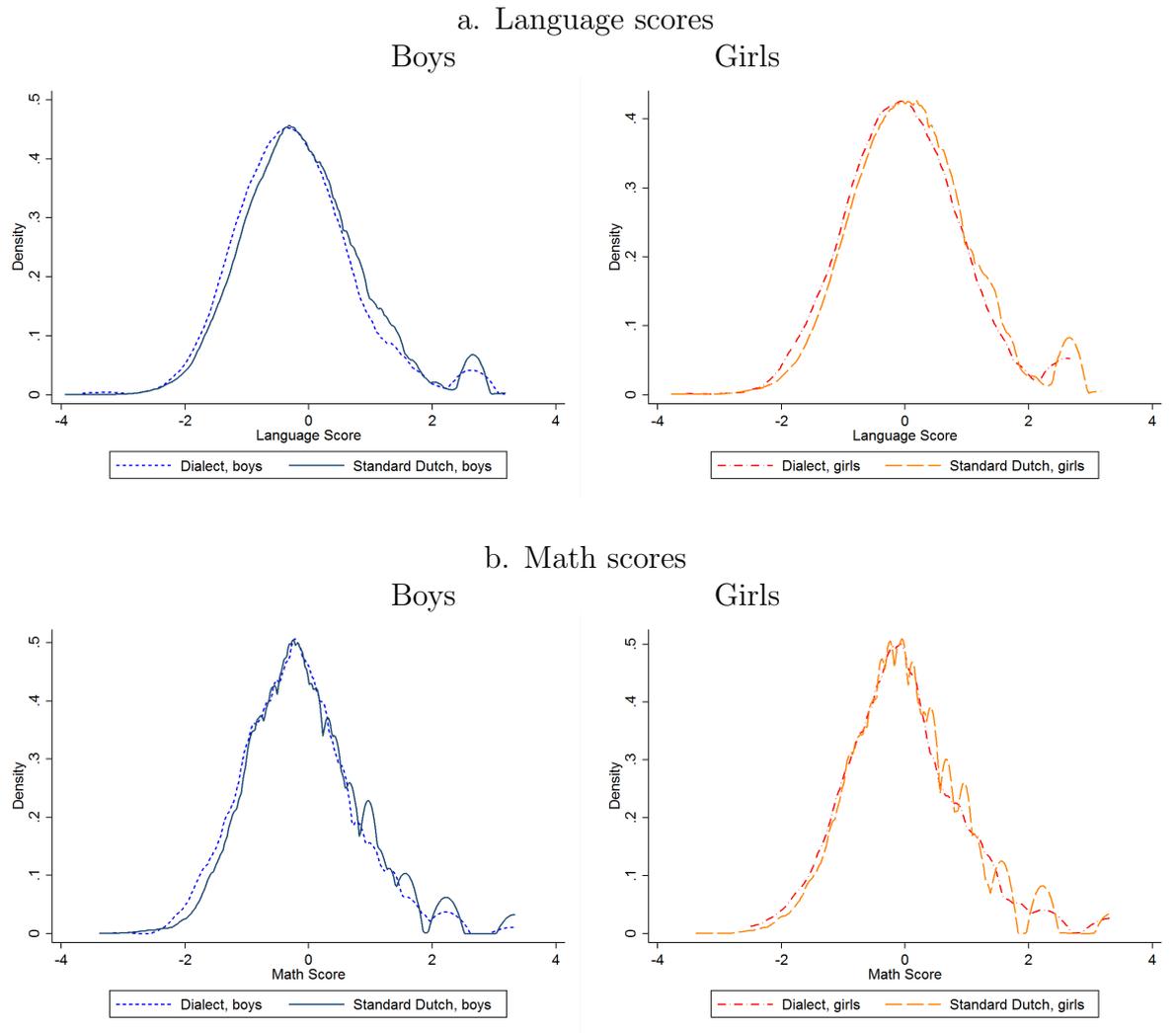
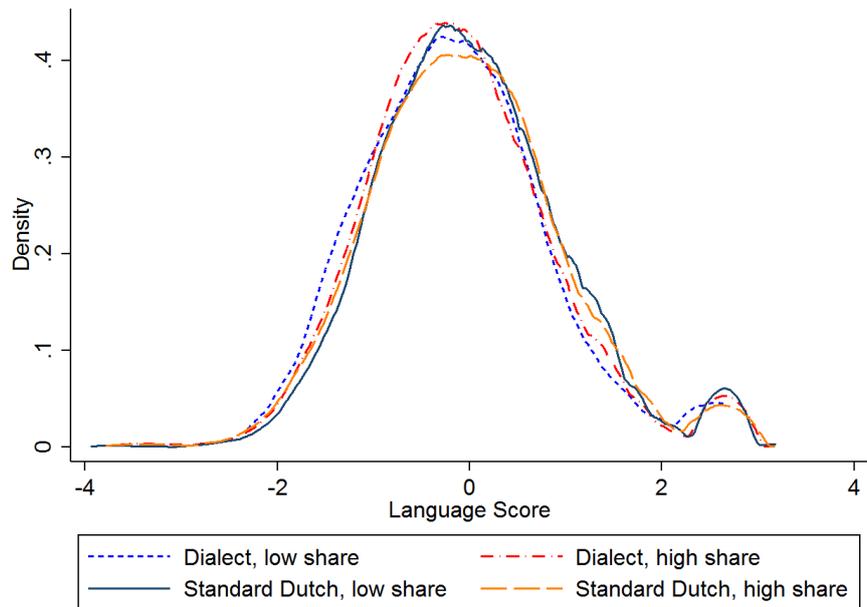


FIGURE 2: DISTRIBUTION OF TEST SCORES BY SHARE OF DIALECT SPEAKERS (LOW-HIGH) AND LANGUAGE SPOKEN AT HOME

a. Language scores



b. Math scores

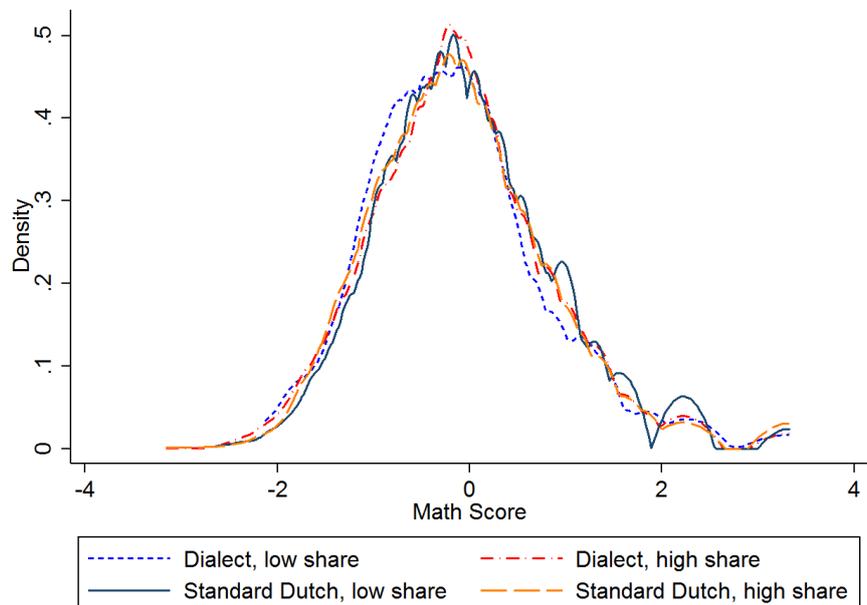
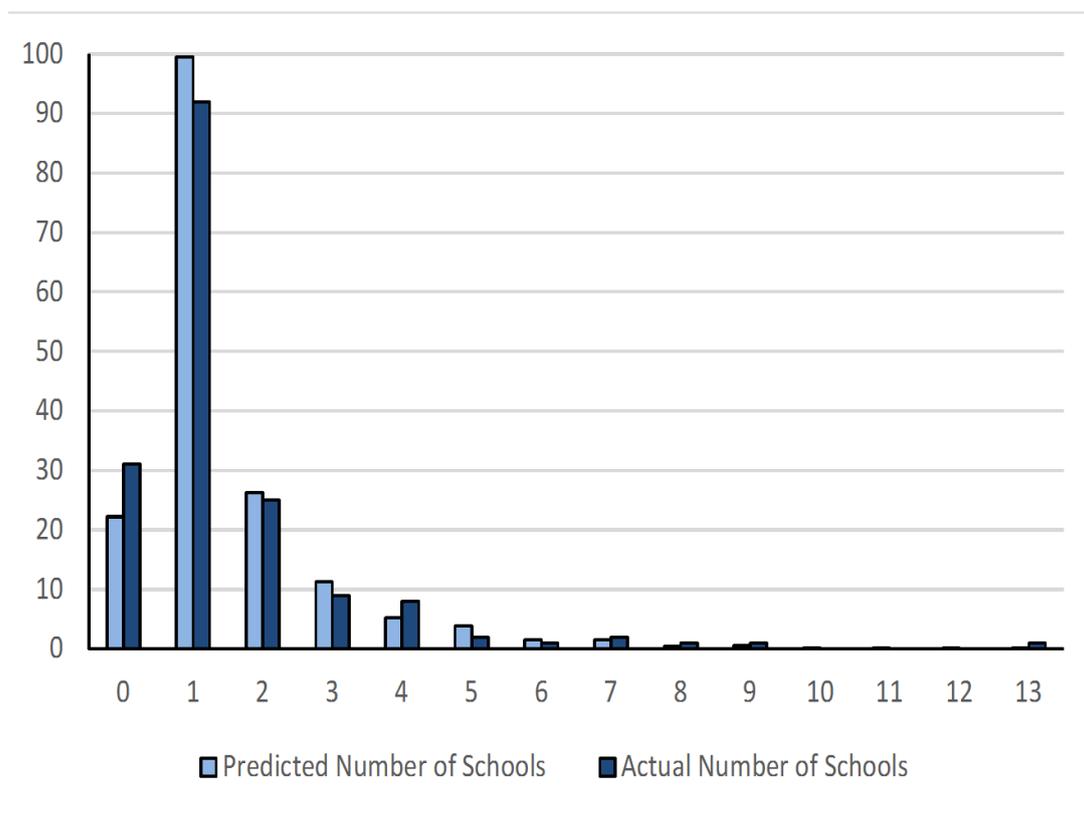


FIGURE 3: RANDOM ALLOCATION OF DIALECT-SPEAKING STUDENTS BETWEEN 2 CLASSES IN ONE GRADE; DIFFERENCE IN NUMBER OF DIALECT SPEAKERS BETWEEN TWO CLASSES IN ONE GRADE



Note: This table uses the sample of schools with two classes in the second grade. It compares the predicted distribution of school and actual distribution of school. The vertical axis is the number of schools. The horizontal axis is the difference in the number of dialect-speaking students between two classes within the same school.

Appendix A: Measuring Dutch Dialect Speaking

A1: Dialectometry - measuring linguistic distances

Ginsburgh and Weber (2015) provide an overview of linguistic distances and their use in economics. Language varies within speech communities and whether a variety is considered to be a dialect or a separate language depends on whether or not it is sufficiently similar to another variety. Nevertheless, there is no clearly defined borderline between language and dialect. Sometimes a standard language is defined as the language of the majority but it also happens that standard language is defined as the language for the elite (Smakman (2012)).

Distances between languages depend on characteristics such as vocabulary, pronunciation, syntax and grammar. To quantify distances between languages various methods are used. Levenshtein (1966) proposed an algorithm based on the minimum number of steps to change a particular word in one language to the same word in a different language. The overall distance between two languages is based on the average difference for a list of words for which often but not always the 100 words from Swadesh (1952) are used. Levenshtein's method can be based on written words but can also be based on phonetic similarities. This is especially helpful when comparing dialects as often these are spoken but not used in writing.¹²

Kessler (1995) is the first study to use the Levenshtein distance measure to study dialects. He studied differences between Irish dialects based on 51 phonetic strings of words spoken in 86 sites noticing that most languages have dialect continua where the speech in one village differs little from the speech of a neighboring village. Therefore, it is difficult to draw major dialect boundaries. Levenshtein distances between dialects have been established for only a few countries and often the results are presented on geographical maps and not as numbers. Valls et al. (2012) study Catalan dialects using the Levenshtein language distance. Gooskens and Heeringa (2004) did the same for Norwegian dialects and Wieling et al. (2014) for Tuscan dialects.

A2: Our distance measure

Our data are not informative about the detailed location of the school. The only information we have is in which province the school is located. Therefore, we use a dialect indicator that is provided at the province level. Van Bezooijen and Heeringa (2006) use two samples of Dutch dialects and apply the Levenshtein distance measure to calculate the average linguistic distances between provincial dialects and standard Dutch. In our paper, we use their distances, which are based on the New Dialect Sample. These distance measures are calculated from 100 words. Van Bezooijen and Heeringa (2006) use logarithmic distances in which small differences in pronunciation are weighted more heavily than large distances.¹³ Table A1 gives an overview of various measures of linguistic distances. The first column shows a measure based on subjective estimates, the second

¹²Note that Levenshtein distances are reported differently varying from absolute numbers to percentages based on arithmetic means or logarithmic means.

¹³The measure is defined as the logarithm of the distance plus 1, divided by the logarithm of the highest possible distance plus 1, multiplied by 100. So, the measure can be interpreted as a percentage ranging from 0 to 100.

column shows the geographical distance from the center of a province to the city Haarlem, which is considered to be the city where the language spoken is closest to Standard Dutch. The third and fourth column show the Levenshtein distance from dialects to standard Dutch based on two different dialect samples. Column (4) is the measure we present in Table 1. We consider this to be the most accurate measure as this is based on a more recent sample of dialects. There is a strong correlation between these measures. We highlight these correlations in Figure A1.

There are just a few studies to which we can compare the Dutch dialect distances.¹⁴ Gooskens and Heeringa (2004) compare 15 Norwegian dialects using 58 words and calculating percentage differences between dialects as in Van Bezooijen and Heeringa (2006). Gooskens and Heeringa (2004) find the smallest difference between two dialects of 14.4 and the largest difference as 41.7. Nerbonne et al. (1996) study the differences between 18 Dutch dialects from various parts of the Netherlands comparing the pronunciation of 100 words. They report relative distances, i.e. absolute Levenshtein differences divided by the length of the larger words. The minimum relative distance is 36, the maximum is 91. A lot of the variation between dialects may have to do with geographical distance as shown by some studies using Levenshtein distance measures. Heeringa and Nerbonne (2001) use 125 words to compare 21 Dutch dialect that are spoken in cities and villages that roughly form a straight line from northeast Netherlands to southwest Netherlands. They find a strong correlation between geographic distance and average Levenshtein distances between two dialects. Wieling et al. (2014) find that Tuscan dialects spoken more distant from Florence to be more likely to differ from standard Italian.

¹⁴As a point of reference, Ispording and Otten (2013) provide a comparison of languages based on the 40 words. They use a normalization of the Levenshtein distance that accounts for chance in similarities and report the following distances from Dutch: English 63.22; German: 51.50.

TABLE A1: AVERAGE LINGUISTIC DISTANCE OF DIALECTS TO STANDARD DUTCH BY PROVINCE

	Subjective (1)	Geographical (2)	Levenshtein distance	
			ODS (3)	NDS (4)
Drenthe	51	79	35	19
Flevoland	22	25	23	12
Friesland	72	62	43	37
Gelderland	34	52	30	28
Groningen	57	95	38	28
Limburg	56	85	35	32
Noord-Brabant	40	52	26	28
Noord-Holland	17	14	23	12
Overijssel	41	67	36	29
Utrecht	18	25	23	18
Zeeland	39	59	30	29
Zuid-Holland	20	23	22	12

Note: The study was done for 12 Dutch provinces and 5 Belgian provinces

(1): Subjective estimates of 69 males and 71 females from different parts in the Netherlands

(2): Geographical distance from the center of a province to Haarlem; rescaled 0-100; 100 was for a Belgian province.

(3): Levenshtein distance based on Old Dialect Sample (ODS), 125 words in 324 dialects sampled in the period 1921-1975.

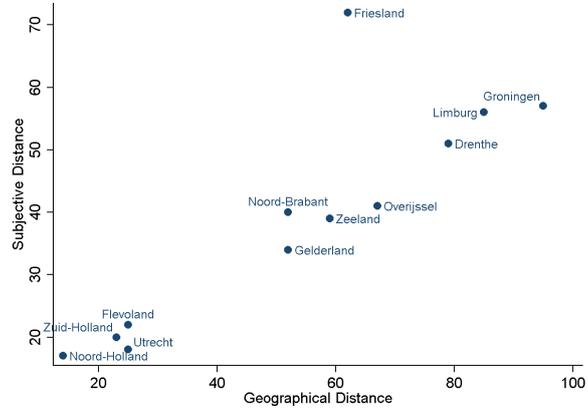
(4): Levenshtein distance based on New Dialect Sample (NDS), 100 words in 27 dialects sampled in 2001.

Note in (3) and (4) no information was available for the province of Flevoland; we assume this province to have the same distance to Standard Dutch as Noord-Holland.

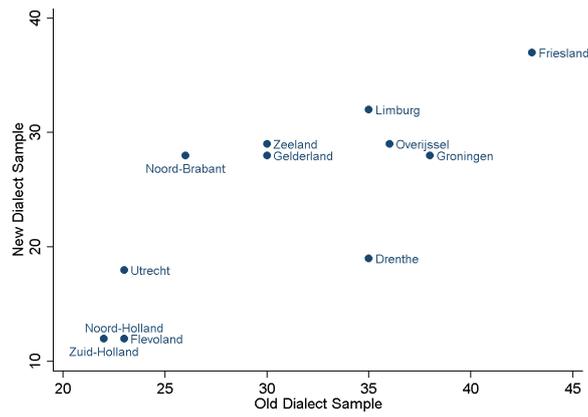
Source: Van Bezooijen and Heeringa (2006)

FIGURE A1: LANGUAGE DISTANCES

a. Geographical and subjective



b. Levenshtein; old data sample and new data sample



c. Geographical and Levenshtein new data sample

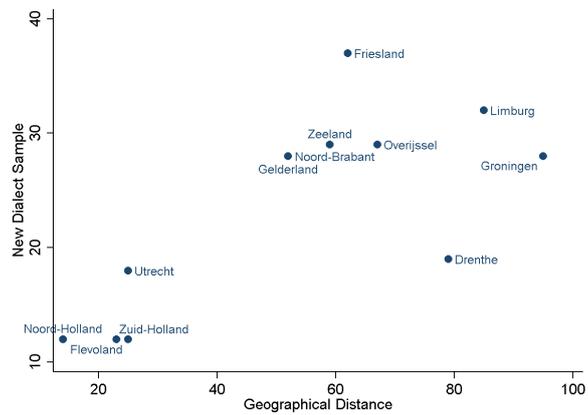
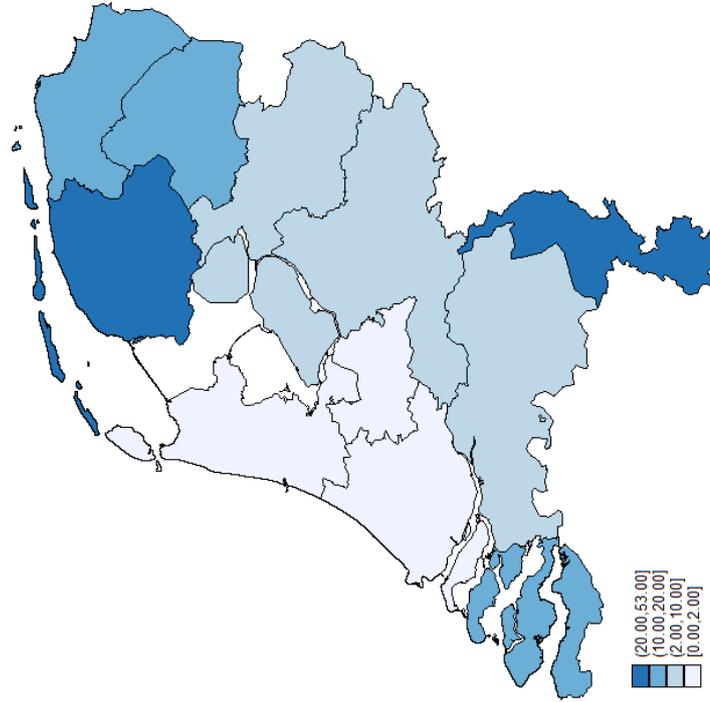


FIGURE A2: MAP OF THE NETHERLANDS BY PROVINCE

a. Dialect-speaking students (%)



b. Levenshtein linguistic distance

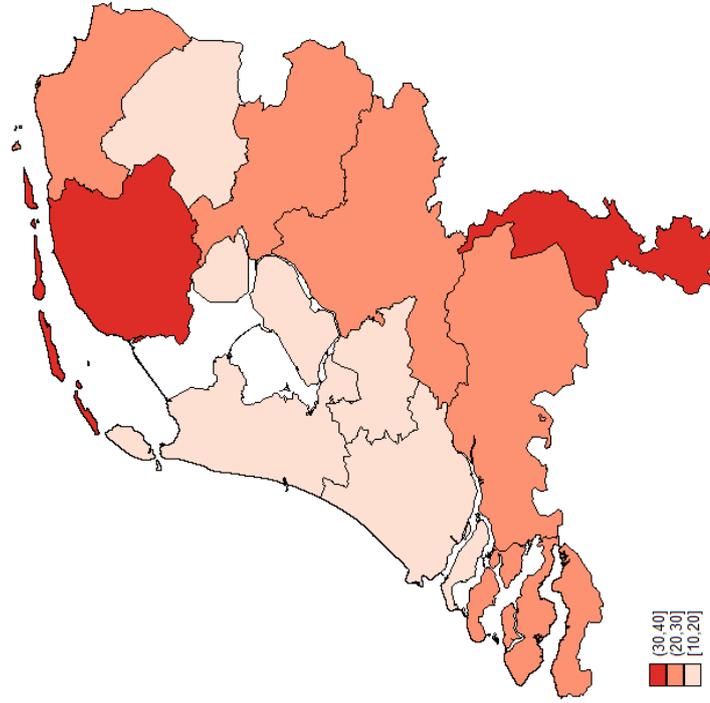
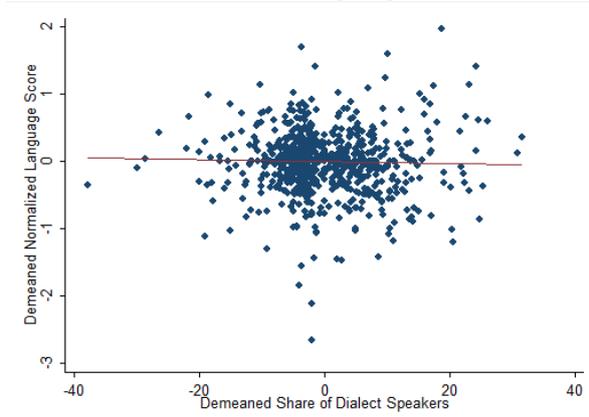


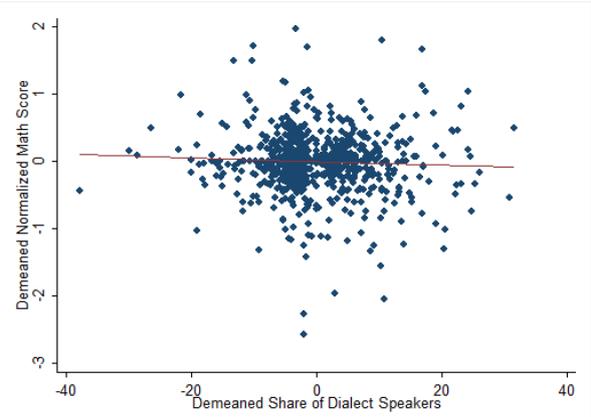
FIGURE A3: CLASS-LEVEL AVERAGE SCORE AFTER WITHIN TRANSFORMATION AND SHARE OF DIALECT-SPEAKING STUDENTS IN A CLASS

a. Dutch-speaking students

Language scores

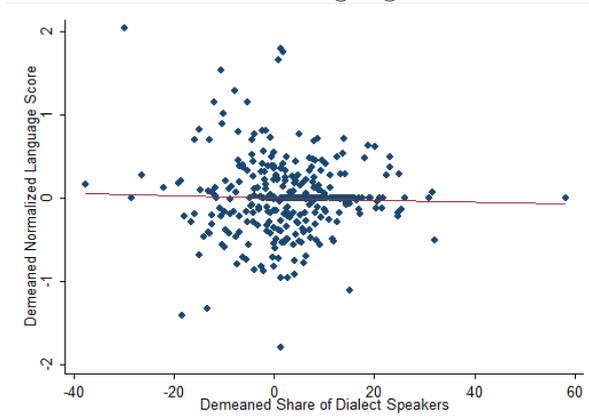


Math scores



b. Dialect-speaking students

Language scores



Math scores

