Intergenerational Transmission of Language Capital and Economic Outcomes

Teresa Casey
Christian Dustmann

ABSTRACT

This paper investigates the intergenerational transmission of language capital among immigrants, and the effect of language deficiencies on the economic performance of second-generation immigrants. Using a long panel that oversamples immigrants, we can follow their children after they have left the parental home. Our results show a sizeable significant association between parents' and children's fluency, conditional on parental and family characteristics. We find that language deficiencies of the second generation are associated with poorer labor market outcomes for females only. Finally, we find a strong relationship between parental fluency and female labor market outcomes, which works through the child's language proficiency.

I. Introduction

A large and growing literature is concerned with the intergenerational transmission of income and wealth. The process by which wealth is transmitted from one generation to the next is an important component in understanding inequality and its evolution over time, as emphasized in work by Becker and Tomes (1986). Recent empirical work finds sizeable differences in intergenerational mobility across countries, with studies for the United States and the United Kingdom (see for example, Solon 1992; Zimmerman 1992; Dearden, Machin, and Reed 1997)

Teresa Casey is a research student and Christian Dustmann is a professor in the department of economics at the University College London, CReAM (Centre for Research and Analysis of Migration) and Centre for Economic Performance, London School of Economics. *The authors would like to thank two anonymous referees for comments. The Economic and Social Research Council (grant RES-000-23-0332) funded this research. The data on which this analysis is based are available from the DIW in Berlin. Users must register and sign an agreement with the data provider. The authors will help other researchers in pursuing these data.*
[Submitted March 2006; accepted July 2007]
ISSN 022-166X E-ISSN 1548-8004

establishing higher-income immobility than studies for European countries like Sweden and Germany (see for example, Bjorklund and Jantti 1997; Wiegand 1997).

But even within the same country, the intergenerational transmission of income differs across populations. Immigrants and their children are an important subgroup that highlights this. Work by Borjas (1993) and Hammarstedt and Palme (2004) illustrates substantial differences in the intergenerational earnings correlation between different immigrant groups. There are several reasons why immigrants should exhibit different patterns of intergenerational transmission. Borjas (1992) emphasizes that the socioeconomic performance of the next generation depends not only on parental skills, but also on the skills of the ethnic group of the parent's generation. Borjas (1995) argues that part of this ethnic capital may be due to selection of immigrants into particular neighbourhoods—a point that has been reemphasized in work by Nielsen, Rosholm, Smith, and Husted (2001), and Rooth and Ekberg (2003).

This paper investigates one particular aspect of human capital where parental endowment may affect transmission to the next generation: language capital. Language proficiency has long been understood to be a key factor in the process of economic assimilation of immigrants and their labor market performance. Earlier papers (see for example, Carliner 1981; McManus, Gould, and Welch 1983; Chiswick 1991; Dustmann 1994; Chiswick and Miller 1995; Shields and Wheatley-Price 2002) find a strong and significant effect of language proficiency on earnings and other economic outcomes. Work by Dustmann and van Soest (2001) shows that simple regression analysis may underestimate this effect due to measurement error—a conclusion that is supported by evidence reported in Bleakley and Chin (2004). Hardly any work exists on how language proficiency affects educational and labor market outcomes of second-generation immigrants.

Language is often cited as the principle initial barrier confronting recent immigrants (see for example, Portes and Rumbaut 1996). As language is significant in determining economic outcomes, it is important to assess the extent to which parental language proficiency affects future generations and is transmitted from one generation to the next. There are many reasons to believe that language proficiency of second-generation immigrants is related to the language proficiency of their parents. Chiswick, Lee, and Miller (2005) investigate what determines the parent/child proficiency relationship. Children of immigrants may experience a monolingual environment in the home country language in the parental home, thus hindering the development of fluency in the host-country language. Lack of exposure to a correct form of the host-country language at early stages of the child's life may have long-term consequences, affecting the child's entire educational chain and accumulation of human capital, and more directly, her labor market opportunities. It is a well-known fact among cognitive scientists that languages are learned more easily at very young ages (see for example, Johnson and Newport 1989). Thus, parental proficiency during the child's formative years in the home may be a critical determinant of the child's host-country language fluency level.

Using a long panel that oversamples immigrants and that allows their children to be followed even after they have left the parental home, we analyze the intergenerational transmission of language capital, and examine how language proficiency of second-generation immigrants affects their labor market outcomes. We contribute

to the literature in several ways. First, we study the effect of parental characteristics, and in particular parental proficiency in the host-country language, on the language proficiency of their children. Bleakley and Chin (2008) analyze the relationship between parents' language proficiency and that of their children for the United States. Their analysis is based on data from the 2000 U.S. Census, which provides self-reported language proficiency of parents as well as their children. Chiswick, Lee, and Miller (2005) analyze the language correlation between parents and their children using the 1996 Australian Census. Our data are quite different. Unlike Bleakley and Chin (2008) and Chiswick, Lee, and Miller (2005), we have repeated information on both parents and their children. This allows us to address the problem of measurement error—which is serious in self-reported data on language (see Dustmann and van Soest 2001)—by using a measure that exploits the repeated information on language proficiency available to us. Also, while Bleakley and Chin (2008) and Chiswick, Lee, and Miller (2005) observe children only before the age of 17, and if they have not yet left the parental household, our information on children's language proficiency is collected after the age of 16, and is independent of the child leaving the parental household. This avoids selection and allows us to investigate not alone the association between parental language proficiency and that of their children, but also the children's later economic outcomes. Our sample is based on survey data and is accordingly smaller than the samples used in Bleakley and Chin (2008) and Chiswick, Lee, and Miller (2005).

Parental language proficiency may be correlated with their child's language proficiency for reasons like education, transmission of ability, and cultural attitudes. We make use of the rich family background information in our data and condition on variables like parental education, age, and origin as well as permanent parental income to eliminate or reduce channels other than language exposure at childhood that may lead to parental language proficiency being correlated with that of their child. As our data allow us to follow immigrant children even after they have left the parental household, our analysis is to our knowledge the first to investigate the effect deficiencies in language proficiency of second-generation immigrants have on their economic outcomes (we investigate earnings, labor force participation, employment, and unemployment), and how this relates to the language proficiency of their parents.

Our analysis distinguishes between males and females, and the children of immigrants who are born in the host country, and those who are born abroad, but arrived in the host country before the age of ten. Our results show a significant and sizeable association between parental language fluency and that of their children. Language deficiencies of the children of immigrants are associated with poorer labor market outcomes for females, but not for males. For females, we establish a clear relationship between parental language fluency and labor market outcomes that works through their language proficiency.

The structure of the paper is as follows. In the next section, we give some theoretical considerations and explain our empirical strategy. In Section III, we discuss our data and samples, and provide some descriptive statistics. Section IV presents results on intergenerational transmission of language proficiency, and Section V analyzes how language deficiencies of second-generation immigrants affect their labor market outcomes. Section VI discusses our findings and concludes.

II. The Transmission of Language Capital

A. Theoretical Considerations

In the intergenerational permanent income model with parental investment and perfect capital markets (see Becker and Tomes 1986), earnings regress to the mean according to the intergenerational correlation of ability. This framework is a useful starting point for empirical work, and Solon (1999, 2004) provides a structural interpretation of coefficients frequently estimated in studies of intergenerational income mobility based on the Becker-Tomes model.

It is less clear, however, that this framework is equally appropriate as a model for the study of other types of characteristics that are transmitted across generations, like language capital. As language proficiency is a part of human capital, part of any correlation between parental language proficiency and that of the child may be explained by the Becker-Tomes model, to the extent that parental language capacity is correlated with parental permanent earnings, and as far as intergenerational correlation in ability may result in a correlation between parent's and children's language proficiency. However, most of the correlation between children's and parent's language proficiency is likely to be driven by exposure to the host-country language at earlier stages in the child's life cycle. As suggested by the cognitive psychology literature, there is a strong relationship between the age of exposure to a foreign language and later proficiency in it.[1]

A number of further factors could be picked up by any correlation of language proficiency between parent and child. Parental language proficiency and that of their children may be related to the social and ethnic context in which children grow up. This argument is similar to that of Borjas (1992) who emphasizes the importance of ethnic capital for intergenerational mobility. Poor language capacity of the parent also may capture more intense ethnic networking, embedding the child into an environment where the host-country language is not often used, and may seem to be of less value.

Bleakley and Chin (2008) define the exposure effects of parental language proficiency on that of their children as their parameter of interest. To isolate these exposure effects, they use parental age of arrival as an instrument for parental language fluency, arguing that those parents who arrived at a young age learn the host-country language more easily. As this variable may work on their children's language proficiency through channels other than parental language proficiency, they use for identification the interaction of this variable with non-English speaking country of origin. This identification strategy therefore assumes the same nonlanguage-related age of arrival effects for children of immigrants from English-speaking and non-English-speaking countries.

1. In the neurolinguistic literature, Penfield and Roberts (1959) proposed the biologically based ''critical period'' hypothesis for second language acquisition. Later studies (Johnson and Newport 1989; Birdsong and Molis 2001; Mayberry and Lock 2003) confirm the hypothesis that there exists a strong relationship between age at exposure to second language and later proficiency in it. Newport (2002) states that decline in average proficiency in second language acquisition can begin as early as ages 4-6. While there has also been some work (Hakuta K, Bialystok E, Wiley E. 2003) which disputes the critical period hypothesis, the consensus in the cognitive psychology literature appears to be that second language attainment is negatively correlated with age of learning.

Our approach makes use of the background characteristics that we have available in our survey data. Instead of using an instrumental variable type approach, we make a ''selection on observables'' assumption, by conditioning on different sets of factors that may lead to confounding the effect of parental language on the offspring's language proficiency, and cultural and environmental factors (see Angrist and Krueger 1999 for discussion). The parameter we estimate is the change in the conditional expectation of the child's language proficiency (L^c) if parental proficiency (L^p) changes, conditional on a set of background variables $X : \frac{\partial E(L^c \mid L^p, X)}{\partial L^p}$. To the extent that the set of conditioning variables eliminates the correlation of child's language proficiency with parental language proficiency through any other confounding channels of the type we have discussed above, this measures the exposure effect. Notice that as our conditioning variables contain measures of parental permanent earnings, the mechanism that creates intergenerational mobility in the Becker-Tomes model is eliminated.

We implement this strategy by estimating regressions of the following form:

$$(1) \quad L_i^c = a_1 + \theta L_i^p + X_i'\alpha + v_i.$$

The variables L_i^c and L_i^p are measures of language proficiency, and we discuss their construction below. The vector of conditioning variables X_i includes family and background characteristics that take account of confounding factors of the sort discussed above. One such factor may be related to origin of the parent, as different home country languages may be more or less distant from the host-country language. Origin dummies (measured as origin country of the father) also pick up ethnic capital (see Borjas 1992) and networking differences across groups.[2] Correlation in genetic endowment may lead to a positive correlation between parental language proficiency and the child's language proficiency. We condition on parental education, as well as a permanent measure of father's earnings.[3] Exposure of the child to the host-country language may further depend on the extent to which parents are integrated in the host-country society and labor market, as well as their social context. To capture this, we condition on the number of years the mother and the father have been in the host country when the child is aged ten years, and survey information about contact of parents with residents of the host country.

B. Measurement of Language Proficiency

A key issue is measurement of the variables L_i^c and L_i^p. In our data, both are self-reported. Self-reported data on language proficiency suffer from measurement error. While this leads only to a loss in efficiency where the child's language measure is concerned, mismeasurement in parental language ability leads (as long as the measurement error is classical and ignoring correlations with the other regressors) to an attenuation bias in the parameter θ. Recent work by Dustmann and van Soest (2001) suggests that measurement error in language ability may lead to a downward bias by

2. In the 17 cases where fathers are absent, we use the origin country of the mother.

3. Again, in 28 cases where we have no earnings information on fathers, we use a permanent measure of mother's earnings. There are 17 children in our sample for which we have no earnings data for either father or mother.

up to a factor of three in estimation equations that regress log wages on self-reported language measures. This finding is supported by evidence provided in Bleakley and Chin (2004).

We address the measurement error problem by making use of the repeated information we have on parental language proficiency to reduce the noise in our data. We construct a time-averaged fixed measure of language proficiency for the individuals in our sample. We do this by estimating fixed effects language equations of the following form:

$$(2) \quad y_{it} = b_0 + age_{it}b_1 + age_{it}{}^2 b_2 + u_i + e_{it},$$

where y_{it} is a measure of language proficiency for individual i in period t, *age* is the individual's age, e_{it} is an idiosyncratic error term, and u_i is an individual specific fixed effect.

Our measure for an individual's language proficiency is then the prediction $\hat{b_0} + age_i\hat{b_1} + age_i^2\hat{b_2} + \hat{u_i}$, evaluated at the parent's age when the child was ten years old.[4] Age ten is chosen here as this is the age that the secondary-school track choice is made in Germany, and therefore a critical point in the child's future education and labor market outcomes.[5] We use the same estimation method to predict the language proficiency of the child (using child's age), where again we predict proficiency at age ten. Notice that choosing another age would not affect the estimates as it does not affect the variation in this variable *across* individuals. The coefficient estimates for b_0, b_1, b_2 will be unbiased and consistent as the number of individuals grows large. The estimate $\hat{u_i}$, though unbiased, will be consistent only as the number of periods grows large. Below, we will run robustness checks by increasing the minimum number of parental language observations on which we base estimation, therefore reducing the remaining measurement error.

III. The Data, the Sample, and Descriptive Evidence

The data we use for this analysis stem from 19 waves (between 1984 and 2002) of the German Socio-Economic Panel (GSOEP), which is a household-based panel survey, similar to the Panel Study of Income Dynamics in the United States or the British Household Panel Survey in the United Kingdom. The GSOEP was initiated in 1984, when it oversampled the then-resident migrant population in West Germany. In the first wave, about 4,500 households with a German-born household head, and an additional sample with about 1,500 households with a foreign-born household head were interviewed, and subsequent interviews took place on a yearly basis. The foreign-born households were collected from the five largest immigrant communities at the time: Turkey, Spain, Italy, former Yugoslavia, and Greece.

4. A similar approach has been used by Dustmann and van Soest (2002).

5. At age ten, children (or their parents) have to choose between three different tracks: lower, intermediate, or higher secondary school. While lower and intermediate secondary school usually lead to apprenticeship training in blue or white-collar professions, only higher secondary school allows access to university. See Dustmann (2004) for more details.

Table 1
Child's Country of Birth by Head of Household's Country of Origin

Head of Household Country of Origin	Children Born in Germany	Children Born Abroad	Total
Turkey	222	111	333
Ex-Yugoslavia	120	43	163
Greece	91	15	106
Italy	113	32	145
Spain	53	10	63
Total	599	211	810

Source: GSOEP, all waves 1984–2002.

Importantly, questionnaires for these households are available in the home country language. The data are thus unique in providing repeated information on a large sample of immigrants over a long period.

From the foreign-headed households, we construct a panel of 810 children, which forms the basis of our analysis. These children come from 467 different households; 238 of these are one-child households—the structure of siblings in the household can be seen in Table A1 in the Appendix. Households in which either immigrant parent obtains German citizenship are excluded from our sample as this renders their children—39 children in our case—ineligible for inclusion in the foreigner survey that contains the language-proficiency questions. Households without children are also obviously excluded from our sample, as are households where the children are too young to complete an adult questionnaire in those years where language questions were asked.

Each individual in a relevant household and older than the age of 15 is interviewed. The household head provides information about all other individuals in the household and those below the interviewing age. Individuals who leave households and form their own households are also tracked and included in the panel.

When individuals are 16 years old, they receive their own personal identification, and pointers to their mother and their father. We construct a sample of parent-child pairs. We follow all children in the sample after the age of 15, and construct a corresponding data set of all mothers and all fathers. We define a second-generation immigrant as an individual who is born in Germany, and whose head of household is born abroad. We have 599 of these children in our sample. We also include children of foreign-born parents who are themselves foreign born, but arrived before the age of ten. There are 211 of these children in our sample. Table 1 shows the distribution of the origin country of the head of household for the 810 children in the sample. As can be seen from this, the country from which the largest share of head of households in our sample originates is Turkey, both for children born in Germany and born abroad. Their mean age at arrival is 4.2 years, and Table A2 in the Appendix shows the distribution of arrival age. In most of our analysis, we distinguish between these two groups.

Detailed language information that we use in our analysis is reported in 11 waves of the GSOEP.[6] Individuals are asked to report their fluency in German on a five-point scale. The question is, "How well do you speak the German language …?" and possible responses are, "Very well … Well … Satisfactory … Badly … Very badly." We scale this information between 0 (for very badly) and 1 (for very well).[7]

We have 810 children in our panel for which we have language observations. We also observe their parents' language proficiency over several years. Tables 2 and 3 give more information on the frequency of this language data.

Table 2 displays the number of language observations we have for the children and their parents in our panel. We have multiple language observations for 79 percent of the children, with language proficiency observed just once for the remaining 21 percent; 98 percent of mothers and fathers also have multiple language observations.

Table 3 shows the percentage of children and parents in our sample that have language observations in each wave of the GSOEP that contains questions on language proficiency. As the children mature into the adult questionnaire (which contains the questions on language), the percentage of parents that remain in the panel to answer language questions decreases. However, this does not pose a problem for us, as it is parental fluency when the child is at a young impressionable age that is most relevant to our analysis.

The kernel densities of the predicted language proficiency of all the parents and children in our sample (by gender of the child) are displayed in Figure 1. The parents' predictions are obtained as explained above; those of their children are obtained in the same way, and fixed at the child's age of ten. As we explain above, we use age ten as the reference age because the secondary-school choice is made at that age, which is important for future careers. Again, the age at which we fix the prediction only affects the location of this distribution, and not its shape.

The left panel of the figure displays densities of the language proficiency of both mothers and fathers. Parental distributions are quite dispersed, with a clear gender difference. The mean and median for fathers is 0.537 and 0.532 respectively, and for mothers 0.428 and 0.432, with corresponding standard deviations of 0.138 for fathers and 0.190 for mothers. The right panel displays corresponding distributions for the children. Although a large proportion of children are concentrated towards the upper part of the distribution, there is a sizeable fraction of the sample at the intermediate part of the unit-language fluency scale.

In Table 4 we display information about children in the sample and their parents. The last column in each panel is the p-value on the difference of the respective characteristic between individuals born in Germany and abroad. Many mean characteristics of second-generation children born in Germany and abroad as well as their parents are significantly different. German-born second-generation immigrants have more years of schooling than those born abroad, with the differential being larger for females than it is for males. The means for spoken language proficiency (predicted at the age of ten years) show that those born in the host country are more proficient in the host-country language than those born abroad. Labor force participation is higher amongst children born abroad, but in this case, the differential is larger between males than females.

6. Language data are contained in waves 1984–87 and then every second year from 1987–2001.

7. The same weight is given to each change in the scale, that is, the values that we attribute to the above 5 levels are 0 (very badly), 0.25 (badly), 0.5 (satisfactory), 0.75 (well), 1 (very well).

Table 2
Number of Times Language Observed for Children and Parents

Number of Language Observations	0	1	2	3	4	5	6	7	8	9	10	11	Total
Percent children	0	21.48	19.75	17.28	11.73	9.38	7.90	4.94	3.33	2.84	1.23	0.12	100
	(0)	(174)	(160)	(140)	(95)	(76)	(64)	(40)	(27)	(23)	(10)	(1)	(810)
Percent mothers	0.99	0.12	0.49	1.36	2.47	3.83	5.68	9.51	8.89	9.75	12.72	44.20	100
	(8)	(1)	(4)	(11)	(20)	(31)	(46)	(77)	(72)	(79)	(103)	(358)	(810)
Percent fathers	2.10	0	1.11	1.48	2.72	3.58	5.19	10.62	8.64	10.12	13.33	41.11	100
	(17)	(0)	(9)	(12)	(22)	(29)	(42)	(86)	(70)	(82)	(108)	(333)	(810)

Note: number of observations in parentheses. Source: GSOEP, all waves 1984–2002.

Table 3
Number of Children's and Parent's Language Observations in Each Wave

Year	1984	1985	1986	1987	1989	1991	1993	1995	1997	1999	2001	Total
Children percent	0.25	7.28	15.06	22.35	33.70	45.56	52.22	48.02	46.54	43.70	39.88	
Observations	(2)	(59)	(122)	(181)	(273)	(369)	(423)	(389)	(377)	(354)	(323)	2,872
Mothers percent	95.29	95.29	94.42	96.28	95.04	91.95	85.75	77.82	67.53	58.49	52.54	
Observations	(769)	(769)	(762)	(777)	(767)	(742)	(692)	(628)	(545)	(472)	(424)	7,347
Fathers percent	97.73	96.97	94.96	96.47	94.45	91.42	85.50	75.79	66.46	56.49	49.43	
Observations	(775)	(769)	(753)	(765)	(749)	(725)	(678)	(601)	(527)	(448)	(392)	7,182

Note: number of observations in parentheses. Source: GSOEP, all waves 1984–2002.

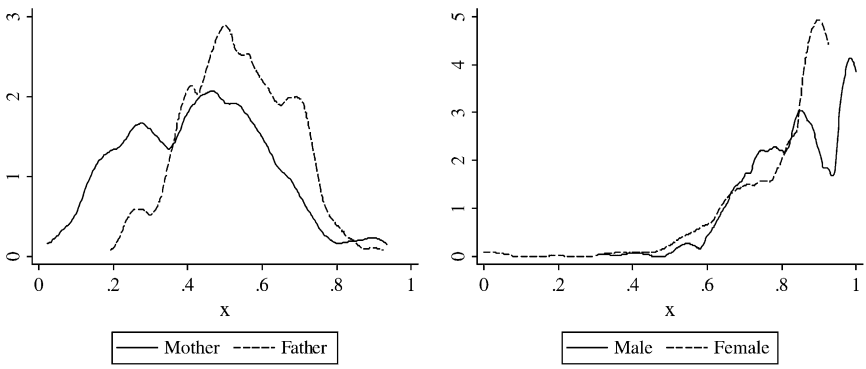

Figure 1
Kernal density of language proficiency, parents (left panel), and children (right panel).

Fathers of children who are born in Germany are slightly older than fathers of children who are born abroad, and fathers are older than mothers. This difference is larger for children who are born in the host country. The years of residence (which are computed when the child was ten years old) are higher for parents of children born in the host country, with a clear difference between males and females, which suggests the typical pattern of male migration and subsequent female migration.

Parental hourly earnings are log hourly permanent earnings of the father, or, where there is no data on father's earnings (28 cases), permanent log hourly earnings of the mother. The data provides information on average monthly gross earnings in the month preceding the interview, and on hours worked for pay during that month. From that information, we compute a log hourly wage rate. We compute permanent log hourly earnings by running fixed effects regressions of log hourly earnings on the individual's age and its square (where the earnings are deflated by a CPI). Our measure of permanent log hourly earnings is the sum of the individual fixed effect and the age polynomial,[8] weighted by the estimated coefficients and evaluated when the child was aged ten.

Permanent log hourly wages are higher amongst parents of children who are born in the host country, as is the percentage of mothers employed when the child is aged 16 years. This may partly be explained by the fact that parents of children born in the host country have been in Germany about seven years longer and have slightly more years of education than parents of those born abroad.

IV. Intergenerational Transmission of Language Capital

A. Basic Estimates

Table 5 reports results of regressions of child's proficiency in the host-country language on parental language proficiency. We distinguish between children who are

8. Eighty-seven percent of sample individuals report hourly earnings in at least four years, and 70 percent in at least eight years.

Table 4
Sample Characteristics

	Females			Males		
	Born in Germany	Born Abroad	*p*-values	Born in Germany	Born Abroad	*p*-values
Children						
Age arrival	—	4.51 (2.45)		—	3.98 (2.28)	
Years of education[a]	10.69 (1.97)	9.88 (1.56)	*0.001*	10.44 (2.07)	10.12 (1.68)	*0.199*
Earnings[b]	2.20 (0.34)	2.09 (0.32)	*0.000*	2.35 (0.36)	2.36 (0.30)	*0.431*
Percent unemployed[c]	8.85	12.40		9.82	9.11	
Percent labor force participation[c]	56.32	60.12		69.71	79.16	
Percent employed[c]	50.77	52.18		62.52	71.33	
Spoken German[d]	0.84 (0.13)	0.75 (0.17)	*0.008*	0.90 (0.10)	0.73 (0.11)	*0.000*
Siblings[e]	74.33	73.68	*0.674*	75.25	71.55	*0.591*
Sample size	300	95		299	116	
Parents						
Mother's age[f]	36.79 (6.10)	36.11 (5.61)	*0.313*	37.46 (6.26)	36.45 (5.71)	*0.118*
Father's age	41.46 (6.31)	39.01 (5.09)	*0.000*	41.82 (6.16)	40.33 (5.22)	*0.015*
Mother's YSM[g]	14.65 (4.12)	7.49 (3.96)	*0.000*	14.96 (4.12)	8.00 (4.17)	*0.000*
Father's YSM	17.48 (4.63)	10.74 (3.91)	*0.000*	17.22 (4.23)	11.45 (4.52)	*0.000*
Parental YSM[h]	17.27 (4.70)	10.73 (3.83)	*0.000*	17.13 (4.27)	11.34 (4.52)	*0.000*
Mother's years of education[i]	8.93 (1.94)	8.28 (1.61)	*0.001*	8.85 (1.96)	8.10 (1.27)	*0.000*
Father's years of education	9.73 (1.96)	9.19 (1.98)	*0.024*	9.60 (1.98)	9.43 (1.84)	*0.422*
Parental earnings[j]	2.43 (0.24)	2.28 (0.25)	*0.000*	2.44 (0.22)	2.34 (0.25)	*0.000*
Mother's German[k]	0.48 (0.18)	0.36 (0.19)	*0.000*	0.47 (0.18)	0.34 (0.18)	*0.000*

(*continued*)

Table 4 (*continued*)

	Females			Males		
	Born in Germany	Born Abroad	*p*-values	Born in Germany	Born Abroad	*p*-values
Father's German	0.56 (0.14)	0.51 (0.14)	*0.007*	0.54 (0.13)	0.51 (0.13)	*0.011*
Parental German	0.52 (0.14)	0.43 (0.15)	*0.000*	0.50 (0.14)	0.43 (0.13)	*0.000*
Percent mothers employed[l]	55.17	41.76		52.41	36.94	
Percent fathers employed	86.02	90.59		83.68	92.66	
Sample size: mothers	298	95		298	116	
Sample size: fathers	293	91		294	115	

Note: In the above table: the number in the first column is the mean of the variable in question, the number in parentheses refers to the standard deviation, and the number in italics is the *p*-value from *t*-tests of mean equality of the variables between females born in Germany and born abroad, and between males born in Germany and born abroad. Sample size: 810 children of immigrant parents.

a. Years of education are the mean years of education of immigrant children who have completed education.

b. Earnings are the mean log hourly earnings of immigrant children who have completed education.

c. Labor Force Participation is as defined by the ILO. The unemployed are defined as all labor force participants who are not working. The employed are defined as all those in our sample who are working.

d. Spoken German is a measure of the child's language proficiency, predicted at age ten, as explained earlier in the paper.

e. Siblings refer to the percent of immigrant children who have at least one sibling.

f. Mother's/father's Age is the age of the parent when the child was ten years old.

g. Mother's/father's YSM refer to their years since migration, when their child was ten years old.

h. Parental YSM is the father's years since migration when the child was ten years old, or if the father is missing, the mother's years since migration when the child was ten years old.

i. Mother's/father's Years Education are the maximum years of education obtained by parents.

j. Parental earnings are the permanent log hourly earnings of fathers (or in 28 cases where missing data on fathers, the mother's permanent log hourly earnings is used), predicted when their child was ten years old.

k. Mother's/father's/Parental German is their measure of spoken German proficiency, predicted when their child was ten years old, as explained in the paper. Parental German is a linear combination of the mother's and father's language measures, or where either parent is missing, the remaining parent.

l. Percent mothers/fathers Employed refers to the employment status of the parent when their child was 16 years old.

Table 5
OLS Language Regressions; Dependent variable is Child's Language Proficiency

	Children Born in Germany			Children Born Abroad		
	(1)	(2)	(3)	(1)	(2)	(3)
Parental language	0.301	0.308	0.252	0.377	0.373	0.327
	(0.034)**	(0.036)**	(0.049)**	(0.072)**	(0.076)**	(0.110)**
Cohort		0.001	0.001		−0.001	0.008
		(0.001)	(0.002)		(0.003)	(0.005)
Male		0.071	0.073		−0.026	−0.039
		(0.009)**	(0.010)**		(0.019)	(0.018)*
Controls for						
Age at arrival				√ (0.000)	√ (0.000)	√ (0.000)
Individual background		√ (0.000)	√ (0.000)		√ (0.000)	√ (0.000)
Parental background			√ (0.000)			√ (0.000)
Observations	599	599	556	211	211	175
R-squared	0.11	0.19	0.23	0.13	0.15	0.26

Note: Robust standard errors in parentheses; * significant at 5 percent; ** significant at 1 percent. (1) No controls. (2) Controls for individual background—cohort, gender, and siblings. (3) Controls for individual and parental background—cohort, gender, siblings, father's years of education, mother's years of education, parental years since migration, parental log hourly wages, country of origin of the head of household, age arrived in Germany for those children born abroad, parental contact with Germans.

born in Germany, and children who are born abroad but who enter the host country before the age of ten. The parental language measure is a linear combination with equal weights on father's and mother's language measures, which are computed according to Equation 3 (based on estimation of models as in Equation 2), and normalized between 0 and 1.[9] The table reports three different specifications for both categories of children, with different sets of conditioning variables. The reported coefficients are the parental language measure, gender, and cohort.

Columns 1 report results where we regress the child's language measure on the parental language measure only. Columns 2 include in addition the child's cohort, a gender dummy, and the number of siblings. Finally, Columns 3 add parental background variables, which include a self-reported measure for contact with other Germans,[10] the years since the father's migration when the child was aged ten years old,[11] mother's and father's years of education, and father's permanent log wage.[12] We also experimented with larger sets of conditioning variables; these resulted

9. There are 12 parent-child pairs where language information on the mother is missing, and 17 parent-child pairs where language information on the father is missing. In these cases, we use predicted language proficiency for the parent present as parental measure.
10. This is a dummy variable if the parent answers yes to the question 'Have you had contact with Germans?' when their child was aged ten or younger.
11. Mother's years since migration when the child was aged ten is used in the 30 cases where father's information on years since migration is missing.
12. Again, in 28 cases where we are missing wage information on the father, we use that of the mother.

Table 6
OLS Language Regressions; Dependent variable is Child's Language Proficiency

	Children Born in Germany			Children Born Abroad		
	(1)	(2)	(3)	(1)	(2)	(3)
Mother's language	0.174		0.145	0.180		0.104
	(0.039)**		(0.039)**	(0.095)		(0.093)
Father's language		0.170	0.108		0.293	0.251
		(0.045)**	(0.046)*		(0.104)**	(0.112)*
Observations	552	556	552	171	175	171
R-squared	0.22	0.21	0.23	0.23	0.26	0.27
P-values	0.00	0.00	0.00	0.00	0.00	0.00

Note: Robust standard errors in parentheses; * significant at 5 percent; ** significant at 1 percent. All columns control for individual and parental background—cohort, gender, siblings, father's years of education, mother's years of education, parental years since migration, parental log hourly wages, country of origin of the head of household, age arrived in Germany for those children born abroad, parental contact with Germans.

in almost identical results as those reported in Columns 3.[13] The numbers in parenthesis beside the tick ($\surd$) are *p*-values for the joint significance of the respective set of regressors. The coefficients on parental language can be interpreted as elasticities, as both the child's and parent's language measures are scaled between 0 and 1.

The estimates across the different columns decrease when we add family background characteristics. The *p*-values indicate that the additional sets of conditioning variables are jointly significant. This suggests that some of the correlation between parental language fluency and that of their children is absorbed by the set of conditioning variables, in particular by parental background. However, the estimates we obtain remain large and significant, suggesting a strong association between parental language proficiency and that of their children. For children born in the host country, the estimated standard deviation of parental language proficiency is about 0.14, so a point estimate of 0.25 implies that a one-standard deviation increase of parental language proficiency leads to an increase in language proficiency of the child of approximately 3.5 percentage points. For children born abroad, the estimated effect is about 5.0 percentage points.

For children born in the host country, there is a strong gender effect, with males being more fluent than females. This effect, if anything, is reversed for children who are born abroad. Finally, as all these estimates—both for children born in Germany and born abroad—are less than one, this suggests the occurrence of a catch-up effect for children across generations.

13. Additional conditioning variables included mother's years since migration, mother's and father's age, mother's and father's school leaving degree, and an average measure of parental education.

Table 7
OLS Language Regressions; Dependent variable is Child's Language Proficiency: Males and Females

	Males		Females	
	Born in Germany	Born Abroad	Born in Germany	Born Abroad
Parental language	0.189	0.214	0.312	0.519
	(0.058)**	(0.103)*	(0.060)**	(0.159)**
Observations	556	175	556	175
R-squared	0.24	0.29	0.24	0.29
Mother's language	0.091	0.028	0.202	0.201
	(0.046)*	(0.088)	(0.053)**	(0.133)
Father's language	0.114	0.223	0.099	0.281
	(0.057)*	(0.109)*	(0.064)	(0.164)
Observations	552	171	552	171
R-squared	0.24	0.29	0.24	0.29

Note: Robust standard errors in parentheses; * significant at 5 percent; ** significant at 1 percent. All columns control for individual and parental background—cohort, gender, siblings, father's years of education, mother's years of education, parental years since migration, parental log hourly wages, country of origin of the head of household, age arrived in Germany for those children born abroad, parental contact with Germans.

B. Which Parent is More Important?

In our analysis above, we have regressed child's language proficiency on language measures of both parents. It is not unlikely that father's and mother's language ability affects the language proficiency of their offspring differently. In Table 6, we present estimates where we regress child's language separately on mother's and father's language and then on both mother's and father's language together.

The estimates refer to the specification that includes the full set of conditioning regressors (corresponding to Columns 3 in Table 5). For children born in the host country, the effects of mother's and father's language proficiency seem to be similar in magnitude. Conditioning on both measures at the same time reveals that mother's proficiency is slightly more important than father's proficiency (although not significantly so), which may be related to children having more intensive contact with the mother in the household. For children born abroad, it seems that it is mainly father's proficiency that is associated with the child's language fluency.

C. Males and Females

In Table 7, we report estimates where we allow parental language fluency to have different effects on males and females, based on the most general specification in Table 5, where we condition on the full set of individual and parental background

Table 8a
OLS Language Regressions; Dependent variable is Child's Language proficiency. Using the Oldest or Only Child in the Household

	Children Born in Germany			Children Born Abroad		
	(1)	(2)	(3)	(1)	(2)	(3)
Parental language	0.298	0.328	0.268	0.331	0.298	0.331
	(0.050)**	(0.051)**	(0.069)**	(0.080)**	(0.092)**	(0.152)*
Observations	225	225	213	118	118	104
R-squared	0.11	0.21	0.25	0.14	0.17	0.28
P-values	0.00	0.00	0.00	0.00	0.00	0.00

Note: Robust standard errors in parentheses; * significant at 5 percent; ** significant at 1 percent
(1) No controls. (2) Controls for individual background—cohort, gender, and siblings. (3) Controls for individual and parental background—cohort, gender, siblings, father's years of education, mother's years of education, parental years since migration, parental log hourly wages, country of origin of the head of household, age arrived in Germany for those children born abroad, parental contact with Germans.

information. The upper panel of the table reports results for parental language, and the lower panel allows for different impacts of father's and mother's language.

The results in the upper panel suggest a larger impact of parental language fluency on females than on males. The coefficient is particularly large for females who are born abroad, and suggests that an increase in parental fluency by one standard deviation increases fluency of daughters who are born abroad by about eight percentage points. These differences between males and females could be explained by the hypothesis that children have more exposure to mother's than father's language in the home, and as females are more likely to spend time in their mother's company than males, then parental language will have a larger impact on females than on males.

In the lower panel of the table, we allow for different effects of father's and mother's language proficiency on male and female immigrant children. For males born in Germany, father's and mother's proficiency has roughly equal impact, while for males born abroad, it seems to be mainly father's language proficiency that affects fluency of the son. For females, there appears to be a slightly stronger association between their fluency and that of the mother.

D. Robustness Checks

As is typical in studies of intergenerational mobility using survey data of the type we use, many children in our sample have the same mother or father. In Table A1 in the Appendix, we illustrate the sibling structure in our sample. More than 70 percent of all children in our data have a brother or sister who is also in the sample.[14] To check whether this affects our estimates, we have reestimated all the models above, restricting our sample by using only the oldest or only child within a foreign-headed household.

14. We adjust standard errors in our estimates above to take account of clustering within families.

Table 8b
OLS Language Regressions; Dependent variable is Child's Language proficiency. Using the Oldest or Only Child in the Household: Males and Females

	Males		Females	
	Born in Germany	Born Abroad	Born in Germany	Born Abroad
Parental language	0.203	0.153	0.329	0.549
	(0.095)*	(0.139)	(0.091)**	(0.204)**
Observations	213	104	213	104
R-squared	0.25	0.32	0.25	0.32

Note: Robust standard errors in parentheses; * significant at 5 percent; ** significant at 1 percent. All columns control for individual and parental background—cohort, gender, siblings, father's years of education, mother's years of education, parental years since migration, parental log hourly wages, country of origin of the head of household, age arrived in Germany for those children born abroad, parental contact with Germans.

The estimates of the effects of parental language on child's language from these regressions are reported in Tables 8a and 8b (using the same specifications as in Table 5 for Table 8a, and as in Table 7 for Table 8b). Although the standard errors are slightly increased due to the decrease in sample size, the point estimates are very similar to those reported in Tables 5 and 7, and they are all statistically significant.

As we discussed above, a particular concern in studies of intergenerational transmission is measurement error. The measures of parental language proficiency we use in our analysis are in most cases based on repeated information for the same individual. Table 2 shows that 97 percent of mothers and fathers have reported their language proficiency in at least three interviews. As the construction of our parental language measure makes use of all language information reported in the sample, this suggests that any downward bias due to measurement error is significantly reduced in the estimates we report. To check whether a more radical selection would affect our estimates, we reestimate our models based on father-mother pairs where each partner reports at least five language spells. The estimates we obtain are very similar to those we report in Table 5 and we include them in the Appendix in Table A3.

V. Language Proficiency and Labor Market Outcomes

A. Language Fluency, Wages, and Employment

Having established a relationship between parental language proficiency and language proficiency of the child, we now turn to examining the effects of the child's language proficiency on their labor market outcomes as second-generation immigrants. We investigate four outcomes: labor force participation, employment, unemployment, and wages. We describe the construction of these variables in Table 4 above.

Table 9
Effect of Child's Language on Labor Market Outcomes

	Males		Females	
	Born in Germany	Born Abroad	Born in Germany	Born Abroad
		Wages		
I	−0.012	0.076	0.269	0.749
	(0.254)	(0.260)	(0.222)	(0.293)*
	628	*430*	*501*	*229*
II	0.289	0.012	0.069	0.677
	(0.311)	(0.300)	(0.262)	(0.479)
	565	*415*	*484*	*217*
		Labor Market Participation		
I	0.006	−0.195	0.467	0.804
	(0.116)	(0.147)	(0.117)**	(0.193)**
	1,471	*761*	*1,300*	*607*
II	0.041	−0.253	0.376	0.584
	(0.123)	(0.163)	(0.136)**	(0.193)**
	1,347	*738*	*1,254*	*550*
		Employment		
I	0.113	0.037	0.602	0.835
	(0.166)	(0.233)	(0.122)**	(0.231)**
	1,475	*768*	*1,311*	*610*
II	0.072	−0.051	0.537	0.640
	(0.176)	(0.263)	(0.136)**	(0.304)*
	1,351	*745*	*1,264*	*553*
		Unemployment		
I	−0.087	−0.251	−0.244	−0.397
	(0.132)	(0.186)	(0.089)**	(0.271)
	1,418	*735*	*1,104*	*438*
II	−0.012	−0.234	−0.278	−0.356
	(0.151)	(0.199)	(0.090)**	(0.330)
	1,297	*713*	*1,071*	*398*

Note: standard errors in parentheses; sample size in italics; * significant at 5 percent; ** significant at 1 percent. (I) Conditioning on child's education, child's age, number of children for females, time dummies, and age arrived in Germany for children born abroad.(II) Not conditioning on child's education; including parental controls—father's years of education, mother's years of education, parental years since migration, parental log hourly wages, country of origin of the head of household, age arrived in Germany for those children born abroad, parental contact with Germans, as well as controls for child's age, cohort, siblings, number of children for females, and time dummies.

We estimate models of the following type:

$$(3) \quad Y_{ti}^{c} = b + \delta L_{i}^{c} + Z_{it}'\beta + u_{ti}.$$

Here L_i^c is a measure of language proficiency, and we use as before the predicted language proficiency at age ten, as reported in Table 4. Z_{it} is a vector of conditioning variables, and $u_{ti} = \lambda_i + e_{it}$ is an error term, with λ_i being an individual specific random effect. For each outcome Y_{ti}^c, we estimate linear random effects models on all available observations, to take account of the covariance structure induced by repeated information on the same individual. Separate regressions are estimated for males and females, and also for those born in Germany and those born abroad. All regressions include only individuals who have finished full time education.[15] We report results in Table 9.

For each outcome we firstly report results (first row of each panel) which condition on educational attainment, age and its square, year dummies, age on arrival and its square for those children born abroad and number of children for females, and secondly, results which do not condition on education, but on family background information instead, where we add the same set of variables as in Columns 3 of Table 5 (second column of each panel), as well as the number of children for females.[16]

The results suggest no significant effect of language proficiency on any of the outcomes for males. For females however, the coefficient estimates are mostly significant, with estimates being similar for the two specifications. The parameter estimates are larger for those who are born abroad than for those who are born in the host country.

For wages, estimates are significant only for females who are born abroad. Here an increase in one standard deviation of language proficiency increases wages by about 13 percent, which is a sizeable effect.

Labor market participation for both females born in the host country as well as females born abroad is positively associated with language proficiency, with the coefficient estimate being larger for those who are born abroad. Conditional on educational achievements, an increase in language proficiency by one standard deviation increases the participation probability for females born in the host country by about six percentage points, while it increases the participation probability of females born abroad by about 14 percentage points.

Looking at employment, we again see that estimates for females born abroad are larger than they are for those born in the host country. An increase of one standard deviation in language proficiency improves employment probabilities by around 8 percentage points for females born in the host country, but by up to 14 percentage points for females born abroad, when we condition on education. This suggests that for females born abroad, language proficiency may have to compensate for a lack of social networks that assist in obtaining employment; networks which females born in Germany may be able to take advantage of in their job search.

15. We drop children who are still in full time education as their labor market outcomes are not comparable to those who have completed their education.

16. These include parental country of origin and cohort controls, and a full set of parental controls for parental years since migration, parental log hourly earnings when the child was ten years old, both parents' years of education, and parental contact with Germans when the child was ten years old or younger.

Estimation results for unemployment probabilities are similar. The probability of unemployment is significantly reduced by fluency—yet again, the effects are larger for those who are born abroad, with an increase in language proficiency by one standard deviation reducing unemployment risk by about seven percentage points, conditional on education. Overall, the effects for females are sizeable, and seem to be larger for children of immigrants who are born abroad rather than in the host country. Again, the possibility that second-generation immigrants who are born in the host country may find it easier to compensate for deficiencies in language fluency than immigrants who are born abroad could explain this finding.

These results suggest that fluency deficiencies in the host-country language are detrimental for the labor market outcomes of second-generation females, but not for males. One reason for this gender difference may be that females find employment opportunities predominantly in jobs where language fluency is quite important (for example, services), while males have more job opportunities in occupations where language is less important. Figures from the ILO (2003) suggest that in industrialized countries, females are more concentrated in the service sector, while males are more concentrated in industry. For Germany, seven of the ten most popular occupations for males (employing nearly 25 percent of all males) are either in crafts (for example, car mechanic or carpenter) or manual work (for example, warehouseman); for females, the ten most popular occupations (employing 38 percent of all females) are all in services or white-collar jobs (numbers are for 2000; own computations based on IABS administrative data). While females with language deficiencies may, in principle, take ''male'' jobs that require less communicative and language skills, they may be disadvantaged in such jobs due to the physical activity involved or employer prejudice, thus making these jobs less accessible for them.[17]

B. Parental Fluency and Child's Labor Market Outcomes

In the previous sections, we have computed the association between parental fluency and the fluency of the child, as well as the effect of the child's fluency on their labor market outcomes. In this section, we assess the importance of parental language fluency on labor market outcomes of the child. To infer the effect of parental fluency on the child's outcomes, we combine the effect of parental fluency on child's fluency with the effect of child's fluency on labor market performance, using the parameter estimates we have obtained from Equations 1 and 3. Hence, we specify exactly the mechanism by which parental fluency affects the child's labor market outcome. As this imposes restrictions on the way parental language affects the child's outcomes, our estimates are more efficient than direct estimation obtained by regressing the child's outcomes on parental language proficiency—the latter would be the reduced form effect, allowing for all other influences of parental language (other than only through the child's language).

The intergenerational language equation we estimate is the same specification as in the upper panel of Table 7. The specifications that relate child's labor market

17. In our sample, and classifying individuals who work into nine occupation groups, about 52 percent of males work in the craft and related trades occupations, compared to only 7 percent of females. In contrast, 38 percent of females work in occupation groups that consist of jobs in the service sector, shops, and market related sales, while only 7 percent of males work in these groups.

Table 10
Child's Labor Market Outcomes and Parental Language Proficiency, 2 Step Estimates

	Males		Females	
	Born in Germany	Born Abroad	Born in Germany	Born Abroad
Wages	−0.080 (0.060)	−0.159 (0.119)	−0.067 (0.089)	0.014 (0.300)
Labor force participation	−0.004 (0.030)	−0.045 (0.047)	0.123 (0.042)**	0.299 (0.139)**
Employment	0.002 (0.035)	0.003 (0.056)	0.164 (0.052)**	0.270 (0.163)
Unemployment	−0.001 (0.025)	−0.058 (0.052)	−0.092 (0.037)**	−0.075 (0.144)

Note: standard errors in parentheses; sample size in italics; * significant at 5 percent; ** significant at 1 percent. These estimates are the combination of the coefficient on the child's language measure when regressed on their different labor market outcomes, and the coefficient on the parental language measure when regressed on the child's language measure. The standard errors are derived using the Delta method. Controls for both sets of regressions include father's years of education, mother's years of education, parental years since migration, parental log hourly wages, country of origin of the head of household, age arrived in Germany for those children born abroad, parental contact with Germans, as well as controls for child's age, cohort, siblings, and number of children for females.

outcome to language are the same as Specification 2 in Table 9. Accordingly, both specifications condition on the same extensive set of background variables. The effect of parental fluency on the child's labor market outcome, which works through its effect on child's language proficiency, is therefore the product of these two parameters. Denote the estimated parameter of parental language proficiency on child's language proficiency by $\hat{\theta}$, with standard error $\hat{\sigma}_\theta$, and the estimated parameter of child's language proficiency on child's labor market outcomes by $\hat{\delta}$, with standard error $\hat{\sigma}_\delta$. The estimates we present in Table 10 are the product $\hat{\theta}\hat{\delta}$.[18]

As we reported above, language fluency is not associated with labor market outcomes for males; we should therefore expect the effect of parental fluency on males that works through the child's language fluency, to be insignificant as well. For females born abroad, we find that the labor force participation probability is positively and significantly affected by parental language proficiency. An increase in parental fluency by one standard deviation increases participation by about 4.5 percentage points. For those born in the host country, an increase in parental language proficiency by one standard deviation, increases the participation probability by 1.7 percentage points and the employment probability by 2.3 percentage points. According to these estimates, language fluency of first-generation immigrants has not only an immediate effect on labor market outcomes, as established in earlier work (see for example, Chiswick and Miller 1995; Dustmann and van Soest 2002) but, by way of affecting their proficiency in the host-country language, impacts on the labor market opportunities of their female children.

VI. Discussion and Conclusions

Language fluency of first generation immigrants is increasingly regarded as a key requirement for qualifying as an immigrant. Economic research has established a strong link between language proficiency of immigrants, and their labor market performance. This paper provides further evidence of the importance of language proficiency for first generation immigrants. It argues that poor language knowledge of immigrants may affect fluency of second-generation immigrants, and that poor fluency may in turn adversely affect labor market performance of second-generation immigrants. Thus, improvement of language fluency of first generation immigrants may not only have an immediate impact, but also may play a role in improving the performance of second-generation immigrants.

To analyze the intergenerational link between language proficiency, and the effects of language proficiency on the outcomes of second-generation immigrants requires data on parents' and their children's language fluency, as well as information about labor market outcomes of second-generation immigrants after they have entered the labor market. We use a long panel for Germany that oversamples immigrants to obtain this information. We construct measures of language proficiency where measurement error is reduced by an averaging type procedure. We find that second-generation immigrants are far more fluent in the host-country language than their

18. Under the assumption that $E(u_{ti} \mid Z_{ti}, X_i, v_i) = 0$, the covariances between δ and θ are zero, so that the standard error can be approximated by the delta method. The reported standard error is computed as $se(\hat{\theta}\hat{\delta}) = \sqrt{\hat{\theta}^2\hat{\sigma}_\delta^2 + \hat{\delta}^2\hat{\sigma}_\theta^2}$.

parents. However, a significant percentage still has language deficiencies. There is a strong and significant association between parental proficiency in the host-country language and the proficiency of their children. This effect slightly decreases, but remains strong even after conditioning on a large set of family and parental background information. The association is larger for children who are born abroad but entered the host country before the age of ten, than it is for children who are born in the host country. It is also larger for females than it is for males.

We find sizeable effects of language fluency of second-generation immigrants on their labor market outcomes for females, but not for males. One explanation is that males find it easier to switch to jobs where language proficiency is less important, while job opportunities for females are concentrated in sectors like services, where language proficiency is more important. We provide evidence that parental language proficiency—through its effect on the child's language fluency—has detrimental effects on employment and participation probabilities of females.

Our analysis is based on survey data, which are unique in the way that they contain an oversample of immigrants, and provide panel information on language fluency and labor market outcomes for both generations. Our results suggest that one reason for the poor labor market performance of second-generation immigrants is fluency deficiencies in the host-country language. Our findings emphasize the importance of fluency in the host-country language for first generation immigrants, and add further weight to policies that ensure language proficiency. Transmission of language deficiencies is found to affect females, in particular, and may therefore contribute to disadvantage female second-generation immigrants, as compared to their male counterparts. The findings provide first evidence on a link between language deficiencies of immigrants and labor market disadvantage of their female children.

Appendix I
Definition of variables used in regressions:

Parental Language: the average of mother's and father's spoken German proficiency when the child is ten years old (or just mother's/father's proficiency if either is missing), predicted using a fixed effects model, scaled on the unit interval.
Age Arrived: the age of the child when they immigrated to Germany for those born abroad.
Head of Household's Country of Origin: the father's (or if missing, mother's) country of origin—Ex-Yugoslavia, Greece, Italy, and Spain, with Turkey being the reference country.
Siblings: the number of siblings that each child in the sample has; those with no siblings are the reference group.
Parental Years Since Migration: the father's (or if missing, mother's) years since migration when the child was aged ten years old.
German Contact: dummy variable which is 1 if either parent had contact with Germans when the child was ten years or younger.
Parental Earnings: the father's (or if missing mother's) log hourly earnings, predicted using a fixed effects model, evaluated when the child was aged ten years old.
Mother's/Father's Education: the maximum years of education obtained by each parent.

Table A1
Distribution of Siblings in Sample

Number of Siblings in Family	Number of Children	Percent of Children	Number of Households	Percent of Households
0	238	29.38	238	50.96
1	292	36.05	146	31.26
2	177	21.85	59	12.63
3	76	9.38	19	4.07
4	20	2.47	4	0.86
5	0	0.00	0	0.00
6	7	0.86	1	0.21
Total	810	100.00	467	100

Source: GSOEP, all waves 1984–2002.

Table A2
Age at Arrival of Foreign-Born Children

	Age Arrived in Germany									
	1	2	3	4	5	6	7	8	9	Total
Percent of sample	16.39	12.67	16.39	14.89	7.32	14.59	5.40	7.98	4.38	100
Number of children	29	21	33	30	14	26	13	12	10	188[a]

a. Year of immigration is missing for 23 of foreign-born children. Source: GSOEP, all waves 1984–2002.

Table A3
OLS Language Regressions; Dependent variable is Child's Language Proficiency. Using only Parents with Five or more Language Observations

	Children Born in Germany			Children Born Abroad		
	(1)	(2)	(3)	(1)	(2)	(3)
Parental Language	0.289	0.297	0.251	0.323	0.317	0.233
	(0.036)**	(0.038)**	(0.052)**	(0.087)**	(0.093)**	(0.111)*
Observations	556	556	530	180	180	160
R-squared	0.10	0.19	0.22	0.09	0.11	0.25
P-values	0.00	0.00	0.00	0.00	0.00	0.00

Note: Robust standard errors in parentheses; * significant at 5 percent; ** significant at 1 percent
(1) No controls. (2) Controls for individual background—cohort, gender, and siblings. (3) Controls for individual and parental background—cohort, gender, siblings, father's years of education, mother's years of education, parental years since migration, parental log hourly wages, country of origin of the head of household, age arrived in Germany for those children born abroad, and parental contact with Germans. Source: GSOEP, all waves 1984–2002.

References

Angrist, Joshua, and Alan Krueger. 1999. "Empirical Strategies in Labor Economics." In *Handbook of Labor Economics*, Vol. 3A, ed. Orley Ashenfelter and David Card, 1277–1366. Amsterdam: North-Holland.

Becker, Gary, and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics* 4(3):S1–39.

Birdsong, David, and Michelle Molis. 2001. "On the Evidence for Maturational Constraints in Second-Language Acquisition." *Journal of Memory and Language* 44(2):235–49.

Bjorklund, Anders, and Markus Jantti. 1997. "Intergenerational Income Mobility in Sweden Compared to the United States." *American Economic Review* 87(5):1009–18.

Bleakley, Hoyt, and Aimee Chin. 2004. "Language Skills and Earnings: Evidence from Childhood Immigrants." *Review of Economics and Statistics* 86(2):481–96.

———. 2008. "What Holds Back the Second Generation? The Intergenerational Transmission of Language Human Capital Among Immigrants." *Journal of Human Resources* 43(2):267–98.

Borjas, George J. 1992. "Ethnic Capital and Intergenerational Mobility." *The Quarterly Journal of Economics* 107(1):123–50.

———. 1993. "The Intergenerational Mobility of Immigrants." *Journal of Labor Economics* 11(1):113–35.

———. 1995. "Ethnicity, Neighborhoods, and Human-Capital Externalities." *American Economic Review* 85(3):365–90.

Carliner, Geoffrey. 1981. "Wage Differences by Language Group and the Market for Language Skills in Canada." *Journal of Human Resources* 16(3):384–99.

Chiswick, Barry. 1991. "Speaking, Reading, and Earnings among Low-Skilled Immigrants." *Journal of Labor Economics* 9(2):149–70.

Chiswick, Barry, and Paul Miller. 1995. "The Endogeneity between Language and Earnings: International Analyses." *Journal of Labor Economics* 13(2):246–88.

Chiswick, Barry, Yew Lee, and Paul Miller. 2005. "Parents and Children Talk: English Language Proficiency Within Immigrant Families." *Review of the Economics of the Household*. 3(3):243–68.

Dearden, Lorraine, Steve Machin, and Howard Reed. 1997. "Intergenerational Mobility in Britain." *The Economic Journal* 107(440):47–66.

Dustmann, Christian. 1994. "Speaking Fluency, Writing Fluency and Earnings of Migrants." *Journal of Population Economics* 7(2):133–56.

———. 2004. "Parental Background, Secondary School Track Choice, and Wages." *Oxford Economic Papers* 56(2):209–30.

Dustmann, Christian, and Arthur van Soest. 2001. "Language Fluency and Earnings: Estimation with Misclassified Language Indicators." *Review of Economics and Statistics* 83(4):663–74.

———. 2002. "Language and the Earnings of Immigrants." *Industrial and Labor Relations Review* 55(3):473–92.

Hakuta, Kenji, Ellen Bialystok, and Edward Wiley. 2003. "A Test of the Critical-Period Hypothesis for Second-Language Acquisition." *Psychological Science* 14(1):31–38.

Hammarstedt, Mats, and Marten Palme. 2004. "Intergenerational Mobility, Human Capital Transmission and the Economic Position of Second-Generation Immigrants in Sweden." Unpublished.

ILO. 2003. *Key Indicators of the Labor Market*. International Labor Office.

Johnson, Jacqueline, and Elissa Newport. 1989. "Critical Period Effects in Second Language Learning: The Influence of Maturational State on the Acquisition of English as a Second Language." *Cognitive Psychology* 21(1):60–99.

McManus, Walter, William Gould, and Finis Welch. 1983. "Earnings of Hispanic Men: the Role of English Language Proficiency." *Journal of Labor Economics* 1(2):101–30.
Mayberry, Rachel, and Elizabeth Lock. 2003. "Age Constraints on First versus Second Language Acquisition: Evidence for Linguistic Plasticity and Epigenesis." *Brain and Language* 87(3):369–84.
Newport, Elissa. 2002. "Critical Periods in Language Development." In *Encyclopaedia of Cognitive Science*, ed. Lynn Nadel, 737–40. London: Macmillian Publishers Ltd.
Nielsen, Helena, Michael Rosholm, Nina Smith, and Leif Husted. 2001. "Intergenerational Transmissions and the School-to-Work Transition of 2nd Generation Immigrants." Working Paper 01–04, Aarhus: Centre for Labor Market and Social Research.
Penfield, Wilder, and Lamar Roberts. 1959. *Speech and Brain Mechanisms*. Princeton: Princeton University Press.
Portes, Alejandro, and Ruben G. Rumbaut. 1996. *Immigrant America*. London: University of California Press Ltd.
Riphahn, Regina T. 2001. "Dissimilation? The Educational Attainment of Second Generation Immigrants." CEPR Discussion Paper No. 2903.
Rooth, Dan-Olof, and Jan Ekberg. 2003. "Unemployment and Earnings for Second Generation Immigrants in Sweden. Ethnic Background and Parent Composition." *Journal of Population Economics* 16(4):787–814.
Shields, Michael, and Stephen Wheatley Price. 2002. "The English Language Fluency and Occupational Success of Ethnic Minority Immigrant Men Living in English Metropolitan Areas." *Journal of Population Economics* 15(1):137–60.
Solon, Gary. 1992. "Intergenerational Income Mobility in the United States." *American Economic Review* 82(3):393–408.
———. 1999. "Intergenerational Mobility in the Labor Market." In *Handbook of Labor Economics*, Vol. 3A, ed. Orley Ashenfelter and David Card, 1761–1800. Amsterdam: North-Holland.
———. 2004. "A Model of Intergenerational Mobility Variation over Time and Place." In *Generational Income Mobility in North America and Europe*, ed. Miles Corak, 38–47. Cambridge: Cambridge University Press.
Wiegand, Johannes. 1997. *Four Essays on Applied Welfare Measurement and Income Distribution Dynamics in Germany 1985–1995*. PhD Dissertation. London: University College London.
Zimmerman, David. 1992. "Regression Toward Mediocrity in Economic Stature." *American Economic Review* 82(3):409–29.