1 Introduction

In many European countries the educational achievements of immigrants' children trail those of natives' children.[1] A recent OECD (2009) study compares educational and labor market outcomes of children of immigrants and finds that for the classical immigration countries Australia, Canada, New Zealand and the United States "the children of migrants have education and labor market outcomes that tend to be at least at par with those of the children of natives." Similarly, Dustmann *et al*. (2010) demonstrate that in the United Kingdom immigrants' children close initial test score gaps over the course of their compulsory schooling. A very different picture emerges for Austria, Germany and Belgium. Test scores from the OECD's Programme for International Student Assessment (PISA) reveal that test score gaps between children whose parents were born in Germany and native children of immigrants amount to the equivalent of about at least two years of schooling. Part of the gap can be explained by the children's socio-economic background, notably the education of parents, but even after controlling for those factors a substantial and significant gap remains. The study confirms the results by Riphahn (2003) on the educational attainment of 2nd generation immigrants in Germany. Riphahn finds that substantial gaps in education relative to natives' children exist and moreover that they do not seem to shrink over cohorts. The children of immigrants from guest-worker nations, notably Turkey, the former Yugoslavia, and Italy, are particularly disadvantaged.

Since immigrants to Continental Europe tended to be less-skilled and less-educated compared to Anglo-Saxon countries, the concept of ethnic capital as introduced by Borjas (1992) emerges as a possible explanation for these different experiences with the second generation. The idea put forward by Borjas is that the average human capital endowment of a certain immigrant group exerts a positive externality on the human capital accumulation of a child belonging to that group. Consequently, a child of immigrants belonging to an immigrant community with high average education should find it easier to acquire human capital compared to an otherwise equal child of immigrants belonging to a less educated community. In a follow-up paper Borjas (1995) specifies the transmission mechanism of ethnic capital as neighborhoods and reports some evidence for the importance of

[1] See Dustmann and Glitz (2011) for a comprehensive overview of the literature on education and migration, including the intergenerational transmission of human capital among immigrants.

this channel.

The innovations of the present paper are 1) the theoretical and empirical extension of the human-capital model in Borjas (1992) to include both ethnic and native capital, 2) allowing for the marginal effects of those variables to vary with the ethnic concentration in the region of a 2nd generation immigrant's upbringing, and 3) linking educational outcomes of immigrants directly to completed educational outcomes of their children using a 25-year running panel study. The last point requires me to confront the substantial panel attrition (partly due to return migration) that inevitably occurs over such an extended period of time. Furthermore, the possibility that immigrants sort along unobserved characteristics into different regions has to be accounted for. To check for the robustness of results to attrition bias I exclude the 10% of the sample who were most likely to drop out of the panel (but did not). To correct for endogeneity of ethnic concentration I use ethnic concentration measured in the year of the mother's year of arrival in Germany or ethnic concentration in 1975 if the mother had arrived before that year (because the ethnic concentration measure is not available for earlier years). As a robustness check I also conduct a generalized propensity score analysis taking into account the continuous nature of the "treatment variable".

The main findings of this paper are:

1. The effect of the parents' schooling in the educational attainment of immigrants' children is much smaller (about a fifth) than for natives' children. This finding confirms results of earlier work by Gang and Zimmermann (2000) and Dustmann (2008). Parental schooling is more important for women's educational attainment while it seems to play no role for men.

2. Immigrants' children who lived in regions with high ethnic concentrations attain lower levels of education.

3. Ethnic capital is economically and statistically insignificant. An additional year of schooling of the *native* population is associated with 0.3 years of additional schooling for immigrants' children (albeit with a high standard error).

4. The mother's education is an important determinant for women's education, and ethnic concentration's effect on education is stronger for men.

The absence of any ethnic capital effect is a remarkable finding and echoes similar results for Denmark in Jakobsen and Smith (2003). The exploration of native capital externalities on 2nd generation immigrants is a novelty in this paper. While the literature has considered the effect of ethnic capital and neighborhood effects in educational or employment outcomes of second generation immigrants, the role that native capital might play in this context has not been explored theoretically nor empirically, with the exception of Borjas (1995) who found that a small number of neighborhood characteristics including the educational composition of the neighborhood accounts for most of the ethnic capital effect.

The question addressed in this paper lies at the nexus of several research areas. It is related to the vast literatures of intergenerational human capital mobility both in general – recent examples including Black *et al*. (2005) or Bleakley and Chin (2008), and for the German case Heineck and Riphahn (2007) or Casey and Dustmann (2005) – and for immigrants in particular.[2] The paper is further related to the neighborhood or, more generally, the social capital literature, in particular studies of children's outcomes. An early and influential paper in economics is Case and Katz (1991), who find substantial neighborhood effects for disadvantaged youth in Boston. Ginther *et al*. (2000) review some of the earlier literature and demonstrate the sensitivity of neighborhood effects on children's outcomes (including high-school graduation) to the inclusion of household variables. More recent studies finding significant effects of peers and neighborhoods on children's educational outcomes include Goux and Maurin (2006) and Rury (2004). Raaum *et al*. (2006) find negligible, if any, neighborhood effects for Norway. Bobonis and Finan (2009) provide evidence of peer effects on secondary school enrollment in Mexican villages using a randomized controlled trial conducted to evaluate a policy intervention.

A closer relation exists to studies which have explicitly looked at the role of ethnic capital or network effects in the context of the intergenerational transmission of education, Borjas (1992) and Borjas (1995) being the pioneering papers. Gang and Zimmermann (2000) study the schooling attainment of second generation immigrants who were born in Germany or arrived before the age of 16. The authors find that immigrants' education has no effect on the educational attainment of their children (contrary to natives' education on their children) and that there is a positive group size effect. However, the latter is measured as the number of immigrants belonging to

[2]For example Van Ours and Veenman (2003), Belzil and Poinas (2010), and Riphahn (2003) for Germany.

the group in Germany and must be considered a rather crude measure of network effects. Åslund *et al*. (2011) find a positive externality of ethnic capital on children's GPA rank. Their study enjoys the advantage of an exogenous initial allocation of refugee immigrants to Swedish municipalities. A comparison of estimates based on initial locations with later residence locations reveal that less educated immigrants sort into ethnically concentrated neighborhoods, creating a severe downward bias of ethnic concentration coefficients on educational attainment. Nielsen *et al*. (2003) consider the role of ethnic capital and neighborhood characteristics in various school-to-work transition measures of second generation immigrants with a rather mixed set of results: ethnic capital seems to prolong the waiting time till entering the first job, but only for females, and seems to play no discernible role other than that. Ethnic concentration increases the same duration for males, and decreases it for females, and is associated with a longer first-job tenure for females, and a shorter first-job tenure for males. The authors control for several neighborhood characteristics, but do not have a measure for native capital. Jakobsen and Smith (2003) find that the concentration of 1st and 2nd generation immigrants in the childhood municipality reduces the probability of finishing a qualifying education, whereas ethnic capital has no impact.

The concept of ethnic – or social – capital is admittedly vague. It is thus important to stress what this paper does and does not do. What this paper does is to analyze the presence, direction and magnitude of externalities exerted by the educational quality of immigrants and natives on the second generation of immigrants. It does not attempt a dissection of the concept of ethnic capital. In other words, it cannot distinguish between different channels through which these externalities might operate on an individual level (e.g. peer effects, exogenous effects, or endogenous effects in the sense of Manski (1993)), which is the main limitation of this study. The next section extends the Borjas (1992) model to include native capital and ethnic concentrations. The third section discusses the data and the construction of the variables of interest, the fourth section describes the estimation and the strategy to deal with panel attrition and selection into regions. The fifth section presents results before I conclude with the sixth section.

2 Theory

Consider the ethnic capital model of Borjas (1992). An ethnic household consists of a parent and a child. The parent has a CES utility function defined over the child's human capital h_{t+1} and his own consumption C_t

$$U = [\delta_1 h_{t+1}^{\rho} + \delta_2 C_t^{\rho}]^{1/\rho} \tag{1}$$

with $\rho < 1$ and $\sigma = 1/(1-\rho)$ is the elasticity of substitution between consumption and child quality. The parent can divide his time between "renting out" his own human capital at a rate R and devoting it to his child's human capital production. With the price of the consumption good normalized to 1, the budget constraint is

$$R(1-s_t)h_t = C_t \tag{2}$$

where s_t is the fraction of time devoted to the child, C_t is the parent's consumption, and h_t his/her human capital (which I will call *parental human capital*). I modify the production function for the child's human capital production in Borjas (1992)

$$h_{t+1} = \beta_0 (s_t h_t)^{\beta_1} \bar{h}_{t,e}^{\beta_2(\theta)} \bar{h}_{t,n}^{\beta_3(\theta)} \tag{3}$$

The child's human capital is produced with parental input, consisting of time devoted and the parent's own human capital, with *ethnic* capital $\bar{h}_{t,e}$, and with *native* capital $\bar{h}_{t,n}$. The relative weights of ethnic and native capital are allowed to vary according to the share of ethnics in a pre-defined locality (neighborhood, school district, city, etc), denoted by θ.[3] Thus, the parent's problem is characterized by

$$\max_{s_t} U = \left[\delta_1 \left(\beta_0 (s_t h_t)^{\beta_1} \bar{h}_{t,e}^{\beta_2(\theta)} \bar{h}_{t,n}^{\beta_3(\theta)}\right)^{\rho} + \delta_2 (R(1-s_t)h_t)^{\rho}\right]^{1/\rho} \tag{4}$$

The inclusion of native capital and the role of the share of ethnics is motivated by the possibility that natives and ethnics can be substitutes (or complements) in the constitution of the relevant externality in human capital production. Imagine two ethnic households with equal parental human capital, one living in an ethnic enclave, the other in a mainly native region. It is conceivable that ethnic capital plays a larger role in the education of the former child (e.g. through peer effects, role models, etc) while for the latter child the natives in the region form

[3] There are different ways to introduce ethnic and/or native capital externalities into the education of a child. An earlier version of this paper has used a model in which children had a preference over conforming to other children of the same ethnicity.

the only reference group (see Currarini *et al*. (2009) for a theoretical motivation of the interaction between group sizes and network formation). The dependence of the weight of the two forms of capital on the share of ethnics is left unspecified, but it seems reasonable to expect a priori a larger role for ethnic capital when the exposure of the child to it is higher.

Solving the parents maximization problem results in the following equations:

$$\begin{aligned}
\frac{\partial \log s_t}{\partial \log h_t} &= \frac{\rho(1-s_t)(\beta_1-1)}{(1-s_t)(1-\beta_1\rho)+s_t(1-\rho)} \\
\frac{\partial \log s_t}{\partial \log \bar{h}_{t,e}} &= \frac{\rho(1-s_t)\beta_2(\theta)}{(1-s_t)(1-\beta_1\rho)+s_t(1-\rho)} \\
\frac{\partial \log s_t}{\partial \log \bar{h}_{t,n}} &= \frac{\rho(1-s_t)\beta_3(\theta)}{(1-s_t)(1-\beta_1\rho)+s_t(1-\rho)}
\end{aligned}$$

Taking logs of the human capital production equation, differentiating, and using the preceding results give:

$$\frac{\partial \log h_{t+1}}{\partial \log h_t} = \frac{\beta_1(1-\rho)}{(1-s_t)(1-\rho\beta_1)+s_t(1-\rho)} > 0 \tag{5}$$

$$\frac{\partial \log h_{t+1}}{\partial \log \bar{h}_{t,e}} = \frac{\beta_2(\theta)(1-\rho s_t)}{(1-s_t)(1-\rho\beta_1)+s_t(1-\rho)} > 0 \tag{6}$$

$$\frac{\partial \log h_{t+1}}{\partial \log \bar{h}_{t,n}} = \frac{\beta_3(\theta)(1-\rho s_t)}{(1-s_t)(1-\rho\beta_1)+s_t(1-\rho)} > 0 \tag{7}$$

Thus, the child's human capital is increasing in its parent's human capital, in ethnic capital, and in native capital. The signs of

$$\frac{\partial^2 \log h_{t+1}}{\partial \log \bar{h}_{t,e} \partial \theta}$$

and

$$\frac{\partial^2 \log h_{t+1}}{\partial \log \bar{h}_{t,n} \partial \theta}$$

are an empirical question. I conjecture that the former is positive, and the latter negative. Intuitively, if ethnic concentration is high, ethnic capital should exert a stronger and native capital a weaker externality.

3 Data

The sample of immigrants, their children, and their individual characteristics come from the German Socio-Economic Panel (SOEP). The earliest wave of the SOEP that allows identification of the county of residence of

households is the 1985 wave. I include all children aged 0 to 16 (some of whom were not born in Germany) in the 1985 wave and, tracking those individuals over all waves up to and including 2010, extract their years of schooling. The most numerous immigrant groups in the SOEP will be considered in this paper. Those are immigrants from Turkey, the former Yugoslavia, Italy, Spain, and Greece. Regional aggregates for the years 1975-1995 are projected from the Employee Sample (IABS) – a 2% sample of the employee population in Germany subject to social security payments. See Bender *et al*. (2000) for a detailed description of this dataset.

3.1 Individual and household characteristics

The outcome variable of interest is educational attainment of immigrants' children, measured as years of schooling. Explanatory variables at the individual or household level are a gender dummy, the mother's years of schooling[4], monthly net household income[5], a dummy for the mother having had any schooling in Germany, a measure of the mother's German speaking ability, the mother's age at migration, the number of adult household members, the number of children in the household, a dummy for the child not being born in Germany, and the year of birth.

The inclusion of the language variable is due to the evidence that the language spoken at home is an important determinant in the educational achievements of immigrants' children presented in Dustmann *et al*. (2010) and Dustmann and Glitz (2011). It is measured on an integer scale from 1 (speaking very well) to 5 (not speaking at all). I use the average speaking ability reported in the survey years 1985-1995 to minimize issues of measurement error. The year of birth is intended to capture a secular increase in the educational attainment in Germany over the last decades, documented by Heineck and Riphahn (2007) and Bildungsbericht (2008). Younger cohorts are more likely to finish high-school and to attend higher education.

[4]Regressions with the father's characteristics are run as a robustness check. See results section.

[5]Averaged over the reported incomes when the child was aged 6 through 10, and deflated to 1985 values by assuming an annual inflation rate of 2%.

3.2 Regional variables

Regional variables are ethnic capital, native capital, ethnic concentration, and interactions between ethnic capital and ethnic concentration as well as native capital and ethnic concentration. For a given year these aggregates are defined for 142 different regions (every region being tied to a local office of the German Employment Agency) in Germany, including West Berlin, and the five aforementioned ethnicities using the IABS. These regions are metropolitan areas and their suburbs, or groups of economically tied counties. The population of West Germany in 1990 was 63.7 million, making the average population of a region approximately 450,000. The choice of the number of regions is a compromise between choosing small regions, thus more accurately capturing the geography in which most interactions of its inhabitants take place, and larger regions, thus preserving larger samples per region to more accurately measure regional variables and decrease attenuation bias.

Ethnic concentration is defined as the share of employees aged 18 to 64 of an ethnicity in a region in all employees aged 18 to 64 of that region. Ethnic capital is defined as the average years of schooling of all employees of an ethnicity in a region. The construction of these regional variables has some shortcomings. First, the variables are based on region of employment rather than region of residence (though the regions are rather large, so that the two are often likely to coincide). Second, the fragmentation of the sample into 142 regions and 5 ethnicities introduces considerable measurement error due to small cell-sample sizes. Third, measurement error is also due to the sample including only employees or transfer recipients. Individuals not in the labor force are not in the population from which the sample is drawn. This might be an important omission if different ethnicities have different participation rates. Finally, the education variable in the IABS exhibits some inconsistencies over time (such as a regression in the educational attainment of an individual) and a high number of missing values. I use the imputation and correction procedures from Fitzenberger *et al*. (2005) to construct the years of schooling.[6]

The regional variables enter the regressions through the degree of exposure that the immigrants' child had to them at the ages 6 through 10.[7] Every year is equally weighted. For example, if child i of ethnicity j is

[6]The results reported here use the imputation rule 1. Results with imputation rules 2 and 3 are similar and can be requested from the author.

[7]I have also tried exposure windows of 6 to 12 and 6 to 15. The results remained unchanged.

8 in 1985 and lives in region 1 in 1985 and 1986, and in region 2 in 1987, then the following ethnic capital value is assigned to it: $C_{i,j} = (1/3) * C_{r=1,j,t=1985} + (1/3) * C_{r=1,j,t=1986} + (1/3) * C_{r=2,j,t=1987}$ where r denotes the region, t the year, and C denotes ethnic capital. The same weighting is applied to ethnic concentration and to native capital. Children whose residence at or before age 10 is not known are assigned 0 for the regional variables, but are marked by a dummy variable (denoted *after10*). The ages 6 through 10 are chosen as the relevant exposure time because of the structure of the German educational system in the 1980s and early 1990s. Children typically were enrolled in primary school at the age of 6. During fourth grade teachers recommended one of several secondary education school types for the child, depending on what they believed to be the child's educational potential. Traditionally there have been three different school types, the *Hauptschule* providing a basic secondary education, the *Realschule* being more advanced than the former, and the *Gymnasium* intended to prepare schoolchildren for university. Few school switches occurred after this initial assignment.[8]

3.3 Descriptive Statistics

Summary statistics for individual characteristics are presented in table 1. The table reports statistics on the variables of interest separately for natives' and immigrants' children separated by ethnicity. Children of immigrants attain lower levels of education than their native peers. The ranking of their years of schooling corresponds roughly to the ranking in education among their mothers, suggesting an important role for the intergenerational transmission of education. Turkish mothers have considerably worse German skills than other ethnicities, and Greek households have the highest income levels. 36% of the Turkish children were not born in Germany. We also see that Turkish and Greek children had exposure to lower degrees of ethnic capital, and Turkish and Yugoslavian children had the highest exposure to their respective ethnic concentrations, reflecting the fact that these immigrant groups were the largest in the 1980s.

Table 2 reports unweighted averages of the regional variables (averaged over all regions). Ethnic capital averages are reflected closely in the sample statistics in table 1. Ethnic concentrations in the sample are higher

[8] According to the Bildungsbericht 2008, a report on education in Germany published by the state ministries of education and the federal ministry of education and research, 3% of children in grades 7-9 switched school types in 2006/2007.

because immigrants are more likely to live in and be sampled from regions with high ethnic concentrations. The table also illustrates that regional variables will be measured with considerable error, especially for smaller immigrant groups such as Greeks and Spaniards. There are 1,196 Spaniards in the IABS in 1985, observed in 121 regions. Thus, for Spaniards regional variables will be based on average on only 10 Spanish observations per region.

Table 3 reports the inter-regional correlations of ethnic concentrations. All correlations are positive and highly significant, pointing to the fact that immigrants of different ethnicities cluster in the same regions. This reflects that these groups of immigrants were recruited for similar work tasks in similar industries.

4 Estimation

I estimate the following equation:

$$y_i = \beta_0 + \beta_1 y_{i,mother} + \gamma' X_i + \delta' Y_i + \varepsilon_i \tag{8}$$

with y_i being years of schooling of observation i, and $y_{i,mother}$ the years of schooling of the mother of i. X_i is a vector of other control variables, Y_i is a vector of regional variables (ethnic and native capital, ethnic concentration, and interactions) and ε_i is the residual. To handle multiple children belonging to the same household all regressions will be clustered at the 1985 household level. Results for this benchmark model are reported in table 4 and will be discussed in the results section.

4.1 Attrition and Endogeneity

To observe the completed years of education of a new-born child in 1985, the child would have to be followed for at least 16 years, and in some cases for as long as 25 to 30 years. Panel attrition is thus substantial, and it is more so for immigrant households since return migration is an additional source of attrition. There are 1,133 observations of immigrants' children with information on all explanatory variables, but only 687 for whom completed years of schooling are observed (61% retention). For natives 1,389 out of 1,935 observations have

information on years of schooling (72% retention). Furthermore, the residence of an immigrant is clearly a choice variable and it is likely – at least for some immigrants – that the educational "environment" for their children plays some role in that decision. If immigrants with stronger preferences towards their children's education move to "better" regions (with higher ethnic and/or native capital), then a simple regression might attribute higher educational attainments of these children to ethnic and/or native capital, even though it might to some extent be due to more parental investment.

To assess the impact of these issues I exclude those 10% of the sample who had the highest predicted probability to attrite. Probabilities were predicted from a probit regression with the dependent variable being an indicator for whether the observation has information on completed years of schooling. This information is missing for observations which have discontinued participation in the panel (through refusal, return migration, death, etc). The explanatory variables are all variables included in the regression of equation 8 and additionally a full set of year-of-birth dummies and an indicator equalling one if in the 1985 survey the parents stated an intention to return to their native country within a few years. The coefficient on the intention to return was -0.241 with a standard error of 0.1 (Pseudo-R2 in this regression was 0.23, see table 8 in the appendix).

Even though all regional variables come as a "packet" with the choice of residence, for lack of good instruments I treat only ethnic concentration as endogenous given its prominence in the regression results and the high cross-regional correlations of ethnic concentration among different immigrant groups. The instrument in the regression of ethnic concentration is the mother's year-of-immigration ethnic concentration in the region of residence in 1985 if the mother immigrated in or after 1975, and the 1975 ethnic concentration in the region of residence in 1985 if the mother immigrated before 1975 (ethnic concentrations before 1975 are not available). Call this the initial ethnic concentration. The initial employer region of residence prior to 1973 was usually determined by Germany's federal employment agency.

The recruitment practice of the federal employment agency is described in Danzer and Yaman (2014) as follows:[9] "The recruitment offices assigned workers from an application pool to specific firms. Employers received almost no information about the characteristics of the guest-worker they hired. While a basic questionnaire de-

[9]This paper is unpublished yet and can be requested from the author.

tailed gender, age and place of origin, no useable information was given regarding previous work experience, training, health conditions, the ability to manage the future job or the willingness to adapt in Germany (general fitness was, however, screened) (...). On the migrant side, the German Labour Office advertised the guest-worker program in the source countries through native-language leaflets and brief video clips in cinemas. Individuals interested in working abroad applied at recruitment offices. Prospective guest-workers provided some basic personal information and were screened to exclude illiterate applicants. Applicants could not state a preference for a destination region or employer but retained the right to refuse an allocated working place. As of 1966, 72% of the foreign workforce comprised unskilled workers because of high labor demand in low-skilled occupations and because several sending countries preferred emigration from underdeveloped regions or disaster areas (...). In Germany, most guest-workers became employed in manufacturing, notably in the construction, mining, metal and ferrous industries."

I cannot rule out that the recruitment office assigned workers to regions based on characteristics that are unobserved in the data, but important for the human capital acquisition of the worker's children, but given the little information that was requested from the applicants and their relative skill homogeneity, it seems unlikely that the assigned region was related to worker characteristics that would also influence their children's educational attainment. The guest-worker recruitment was stopped in 1973, and immigration from the countries considered in this paper continued in the 1970s mainly for family reunification. Ethnic concentrations in the year of immigration and in the initial period after immigration are thus assumed exogenous to immigrants' ability and inclination to invest in their children's education, but exhibit strong correlations with subsequent ethnic concentrations.

To assess the robustness of the IV approach I also employ a propensity score estimation with continuous treatment (ethnic concentration being the treatment) following the steps outlined in Hirano and Imbens (2004) and using the Stata routines written by and described in Bia and Mattei (2008). This substitutes the generalized propensity score (GPS) for the propensity score in the binary treatment case. It is defined as the conditional density of the treatment given the covariates. Importantly, this procedure still makes the weak unconfoundedness assumption, namely that the potential outcome is independent of the actual treatment given the covariates, the equivalent to the exogeneity assumption in linear regression. Thus, in the presence of relevant omitted variables,

this approach might reduce, but cannot remedy omitted variable bias.

5 Results

A first set of results is reported in table 4. The first four columns are results for immigrants' children, the last two columns are results for natives' children. The difference in the determinants of years of schooling between natives' and immigrants' children is striking. Intergenerational transmission of education is much stronger for natives, confirming the finding in Gang and Zimmermann (2000) and Dustmann (2008). The reason for this difference is an unresolved issue. Measurement error could be a contributing factor because of differences in schooling systems and qualities across the sending countries and the difficulty to assign years of education to school and training diplomas obtained in a foreign country. To assess the impact of measurement error I have run the regression in column 1 (together with the ethnic concentration variable) separately for the different immigrant groups, which should at least rule out measurement error due to cross-country differences. The coefficient for mother's education ranged from 0.119 to 0.181 for Turks, Yugoslavians, and Greeks, but was virtually zero for Italians and negative for Spaniards (the number of Spanish observations was only 52). Excluding observations who were not born in Germany resulted in an even smaller coefficient.[10] Thus, measurement error seems to introduce attenuation bias, but cannot alone explain the gap between natives and immigrants. Dustmann (2008) considers the possibility that it is the parents' permanent income rather than their education that determines the child's educational attainment. The results in table 4 do not support this since household income in native households is a strong predictor for educational attainment, but not in immigrant households. It is possible however that household income could be a better proxy for permanent income for native households. Daughters of natives attain more schooling than sons, but not significantly so. Again, the effect is absent for immigrants' children. Looking at determinants specific to immigrants, we see that German skills (1 speaks very well, 5 does not speak at all) increase the child's education, but not significantly so. Children who were born in Germany enjoy an educational advantage of some 7 months of schooling. Even when only household variables are included and when the same regressors are used, a significant difference between mother's education coefficients between

[10] Detailed results are available upon request from the author.

the native and the immigrant samples remains.[11]

For the regional variables – ethnic and native capital, ethnic concentration (native concentration for natives), and an interaction term of the two – we observe a nuanced picture. If only ethnic capital and ethnic concentration are included (column 2), ethnic concentration has a significantly negative effect. The coefficient of ethnic capital is virtually zero. For natives (column 5), the effects of both of those variables are small and insignificant. The effect of native capital (in column 3) is statistically insignificant but much larger in magnitude compared to the coefficient of ethnic capital.[12] Including interaction terms (column 4 for immigrants' children, and column 6 for natives' children) predictably changes the coefficients of the other regional variables, but they are neither individually nor jointly significant. Both are negative which is contrary to our expectation with regard to ethnic capital, but in line with our expectation with regard to native capital (any positive effect of native capital is diminished as the ethnic concentration increases). The individual and household determinants are mostly unchanged by the inclusion of regional variables (except for the dummy indicating whether the mother has had any education in Germany).

Table 5 shows results for the regional variables obtained after excluding the observations who were most likely to attrite, for the 2SLS results with ethnic concentration instrumented by initial ethnic concentration, and for the generalized propensity score (GPS) regression. We see that the coefficient of ethnic concentration becomes smaller in magnitude in the restricted sample suggesting that attrition could be biasing the concentration coefficient downward. We also observe that the magnitude of the coefficient on ethnic concentration is 10% smaller in the 2SLS regression than in the OLS regression. The instrument exhibits a very strong first stage. The upward correction of the ethnic concentration effect is in line with the selection of regions with low ethnic concentrations by immigrants who have stronger preferences over their children's education, and who might be moving to those regions in the expectation that it will benefit their children. The last column reports the coefficient on ethnic concentration in a regression of the child's educational attainment on ethnic concentration and the generalized propensity score. The coefficient is considerably smaller.[13] I would like to caution the

[11] These results are available upon request from the author.

[12] The correlation coefficient between ethnic and native capital among those observations for whom the variables are measured – that is those observations whose region of residence is observed for at least some time when they are aged 6 to 10 – is 0.28.

[13] The number of observations is smaller because cases with ethnic concentrations equal to 0 were excluded. This was required to

reader that even after adjusting for the GPS the balancing property - that is insignificant differences of covariate means between groups of different treatment intensity - could not be obtained. This was mainly due to surviving significant differences in the number of some ethnicities in higher and other ethnicities in lower concentration sub-samples. Figure 1 plots the *dose response function* for the GPS defined as the expected outcome at a given treatment level. The plot confirms the regression coefficient. We see that educational attainment is highest at relatively low levels of ethnic concentration and drops linearly as ethnic concentration increases.

Table 6 reports four sets of results. The first two columns use the *mother's* years of schooling as the measure of parental human capital, while the last two columns use the *father's* years of schooling ($y_{i,mother}$ in equation 8 being replaced by $y_{i,father}$). Columns 1 and 3 restrict the sample to males only, and columns 2 and 4 to females only. For example, the first row in column 2 is the mother's schooling's effect on the daughter's educational attainment, and the first row in column 3 is the father's schooling's effect on the son's educational attainment. We see that sons are not influenced by either parent's human capital (columns 1 and 3), and that ethnic concentration effects them more adversely than it does the daughters. The effect of ethnic concentration on the daughters is negative, albeit insignificant, while the mother's years of schooling exerts a positive influence on the daughter's education (column 2). Still, the effect of mother's education is only about two thirds of the effect for natives (table 6, first row of columns 5 and 6). The results suggest that boys and girls have different weights for parental human capital and ethnic concentration in their production of human capital, with girls being more influenced by the parents (especially the mother) and boys more - and negatively - influenced by the ethnic concentration. The variables native capital and ethnic capital are neither individually nor jointly significant.

Table 7 reports some robustness and sub-sample results. Column 2 reports results when including two additional regional variables – the share of employees working in mining, manufacturing, or construction, and the median wage in the region. The magnitude of the ethnic concentration coefficient increases by 40% from -0.203 to -0.281 while the other coefficients remain largely unchanged relative to the benchmark model (column 1). Including a set of state dummies defined over the state of residence in 1985 (column 3) has no discernible effect. The fourth column reports results from a sample using only observations for whom information on their residence

transform the treatment variable to a normal distribution.

at any point between ages 6 and 10 is available. Results are similar to the benchmark model (column 1). Looking at observations who were born in Germany only, few differences arise. First, mother's education becomes completely unimportant, and second, the magnitudes for native capital and ethnic concentration increase but remain insignificant, with the familiar result that native capital exhibits a positive, and ethnic concentration a negative sign.

6 Discussion and Conclusion

This paper has analyzed the determinants of educational achievement of immigrants' children in Germany, with a focus on the role played by regional aggregates such as ethnic concentration, ethnic capital, and native capital. Results show that 1) household variables – and in particular the parents' education – play a minor role, especially in comparison to natives' children, 2) ethnic concentrations have a negative effect, 3) ethnic capital has no effect, 4) native capital has a positive (though insignificant) effect, and 5) there are important differences in the relative importance of those variables between men and women.

The first result could be a feature of low and rather homogenous educational levels of first generation immigrants, especially within, but also between immigrant groups. One has to keep in mind that the immigrant groups considered in this paper were for the most part recruited for work with little skill requirements. The second result points strongly to the possibility that ethnic neighborhoods act as an impediment to educational attainment. It seems unlikely that this is a regional effect, since the effect is absent for natives, and it persists even after controlling for other regional variables and 1985 state-of-residence dummy variables. The impeding effect of ethnic neighborhoods could be due to the children failing to learn proper German before school enrolment or to more network based alternatives in the local labor market, increasing the opportunity cost of staying in education.

The third result constitutes a refutation of the ethnic capital model, at least for the case of Germany. The native capital coefficient has the sign predicted in equation 7 and its magnitude is much larger than the ethnic capital coefficient. Furthermore, the results show that there is some interplay between those two variables. Inclusion of native capital typically decreases the coefficient on ethnic capital. While it is improbable to shed

more light into this relationship with household survey data, the results highlight the importance of considering native capital as much as ethnic capital when immigrants' educational attainments are analyzed.

References

[1] Åslund, O., Edin, P.-A., Fredriksson, P. and Grönqvist, H. (2011) Peers, Neighborhoods, and Immigrant Student Achievement: Evidence from a Placement Policy, *American Economic Journal: Applied Economics*, **3**, 67-95.

[2] Belzil, C. and Poinas, F. (2010) Education and early career outcomes of second-generation immigrants in France, *Labour Economics*, **17**, 101-110.

[3] Bender, S., Haas, A. and Klose, C. (2000) IAB Employment Subsample 1975-1995 - Opportunities for Analysis Provided by the Anonymised Subsample, Discussion Paper No. 117, Institute for the Study of Labor, Bonn, Germany.

[4] Bia, M. and Mattei, A. (2008) A Stata package for the estimation of the dose-response function through adjustment for the generalized propensity score, *The Stata Journal*, **8**, 354-373.

[5] Bildungsbericht: Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland und des Bundesministeriums für Bildung und Forschung (2008), Bildungsbericht 2008, Bertelsmann Verlag, Bielefeld, Germany.

[6] Black, S. E., Devereux, P. J. and Salvanes, K. G. (2005) Why the apple doesn't fall far: Understanding intergenerational transmission of human capital, *American Economic Review*, **95**, 437-449.

[7] Bleakley, H. and Chin, A. (2008) What holds back the second generation? The intergenerational transmission of language human capital among immigrants, *Journal of Human Resources*, **43**, 267-298.

[8] Bobonis, G. J. and Finan, F. (2009) Neighborhood peer effects in secondary school enrollment decisions, *Review of Economics and Statistics*, **91**, 695-716.

[9] Borjas, G. J. (1992) Ethnic capital and intergenerational mobility, *Quarterly Journal of Economics*, **107**, 123-150.

[10] Borjas, G. J. (1995) Ethnicity, neigborhoods, and human-capital externalities, *American Economic Review*, **85**, 365-390.

[11] Case, A. C. and Katz, L. F. (1991) The company you keep: The effects of family and neighborhood on disadvantaged youth, Working Paper No. 3705, National Bureau of Economic Research, Cambridge, MA.

[12] Casey, T. and Dustmann, C. (2005) Intergenerational transmission of language capital and economic outcomes, Discussion Paper 06/05, Centre for Research and Analysis of Migration, London, UK.

[13] Currarini, S., Jackson, M. O. and Pin, P. (2009) An economic model of friendship: homophily, minorities and segregation, *Econometrica*, **77**, 1003-1045.

[14] Danzer, A. M. and Yaman, F. (2014) Ethnic Concentration and Language Fluency of Immigrants: Evidence from the Guest-Worker Placement in Germany. *mimeo*.

[15] Dustmann, C. (2008), Return migration, investment in children, and intergenerational mobility, *Journal of Human Resources*, **43**, 299-324.

[16] Dustmann, C., Machin, S. and Schönberg, U. (2010) Ethnicity and educational achievement in compulsory schooling, *The Economic Journal*, **120**, F272-F297.

[17] Dustmann, C. and Glitz, A. (2011) Migration and education, in: Hanushek, E. A., Machin, S. and Woessmann, L. *Handbook of the Economics of Education*, vol. 4, 327-439.

[18] Fitzenberger, B., Osikominu, A. and Völter, R. (2005) Imputation rules to improve the education variable in the IAB Employment Subsample, Discussion Paper No. 05-10, Centre for European Economic Research, Mannheim, Germany.

[19] Gang, I. N. and Zimmermann, K. F. (2000) Is child like parent? Educational attainment and ethnic origin, *Journal of Human Resources*, **35**, 550-569.

[20] Ginther, D., Haveman, R. and Wolfe B. (2000) Neigborhood attributes as determinants of children's outcomes: How robust are the relationships, *Journal of Human Resources*, **35**, 603-642.

[21] Goux, D. and Maurin, E. (2006) Close neighbours matter: Neighbourhood effects on early performance at school, Discussion Paper No. 68, Centre for the Economics of Education, LSE, London, UK.

[22] Heineck, G. and Riphahn, R. T. (2007) Intergenerational transmission of educational attainment in Germany: The last five decades, Discussion Paper No. 2985, Institute for the Study of Labor, Bonn, Germany.

[23] Hirano, K. and Imbens, G. W. (2004) The propensity score with continous treatment. In *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, ed. Gelman, A. and Meng, X.-L., 73-84. West Sussex, England: Wiley InterScience.

[24] Jakobsen, V. and Smith, N. (2003) The educational attainment of the children of the danish 'guest worker' immigrants, Discussion Paper No. 749, Institute for the Study of Labor, Bonn, Germany.

[25] Manski, C. F. (1993) Identification of endogenous social effects: the reflection problem, *The Review of Economic Studies*, **60**, 531-542.

[26] Nielsen, H. S., Rosholm, M., Smith, N. and Husted, L. (2003) The school-to-work transition of 2nd generation immigrants in Denmark, *Journal of Population Economics*, **16**, 755-786.

[27] OECD (2009) Children of immigrants in the labour markets of EU and OECD countries: An overview, OECD, Paris, France.

[28] Raaum, O., Salvanes, K. G. and Sørensen, E. Ø. (2006) The neighborhood is not what it used to be, *Economic Journal*, **116**, 200-222.

[29] Riphahn, R. T. (2003) Cohort effects in the education attainment of second generation immigrants in Germany: An analysis of census data, *Journal of Population Economics*, **16**, 711-737.

[30] Rury, J. L. (2004) Social capital and secondary schooling: Interurban differences in American teenage enrollment rates in 1950, *American Journal of Education*, **110**, 293-320.

[31] Van Ours, J. C. and Veenman, J. (2003) The educational attainment of second-generation immigrants in the Netherlands, *Journal of Population Economics*, **16**, 739-753.

Figures

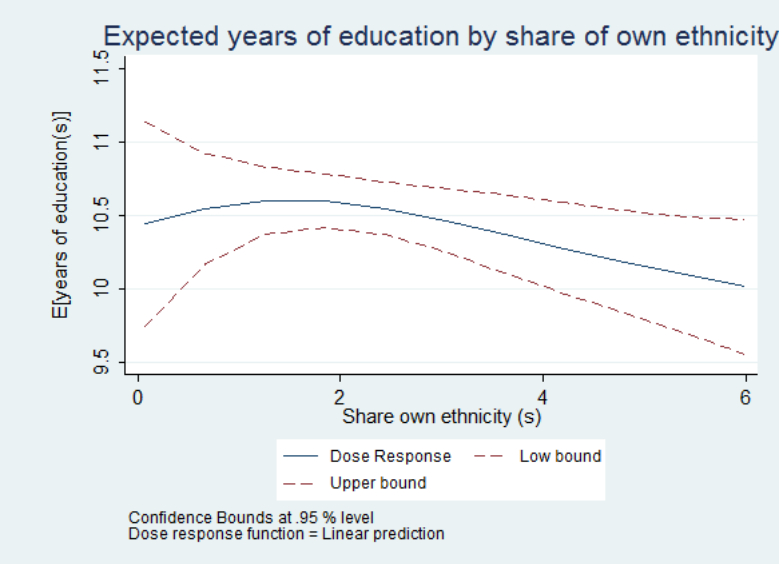

Figure 1: Mean educational attainment by ethnic share

Tables

Table 1: Summary statistics, Immigrants' children

	(1) Turkey	(2) Yugoslavia	(3) Italy	(4) Spain	(5) Greece	(6) Natives
Years of schooling (mean)	10.3	10.6	10.4	10.4	10.7	11.8
Mother's years of schooling (mean)	8.0	8.9	8.4	8.0	8.7	10.8
Mother's age at migration (mean)	26.3	22.6	22.9	22.4	21.1	n.a.
Mother's German skills (1 to 5)	3.5	2.5	3.0	2.7	2.7	n.a.
Household income (mean)	1,188	1,356	1,160	1,380	1,695	1,579
Male (%)	54	50	50	63	52	51
Immigrant (%)	36	29	24	15	11	n.a.
Adult household members (mean)	2.7	2.2	2.4	2.4	2.5	2.3
Children in household (mean)	3.0	2.5	2.2	2.1	2.0	1.9
Ethnic capital (mean)	9.9	10.5	10.2	10.5	9.9	12.2
Ethnic concentration (mean)	2.9	2.7	1.9	0.5	1.0	94.2
observations	289	155	112	52	79	1,389

Source: SOEP 1985-2010.

Table 2: Summary statistics, Regional variables 1985

	(1) Turkey	(2) Yugoslavia	(3) Italy	(4) Spain	(5) Greece	(6) Natives
Ethnic capital	9.9	10.5	10.3	10.3	10.0	
	(0.5)	(0.6)	(0.8)	(1.0)	(1.1)	
Ethnic concentration	2.1	1.4	1.0	0.4	0.5	
	(1.2)	(1.3)	(1.0)	(0.3)	(0.5)	
Native capital	12.1	12.1	12.2	12.2	12.2	12.1
	(0.2)	(0.2)	(0.2)	(0.2)	(0.2)	(0.2)
number of regions with observations	141	140	131	121	125	142
number of observations in IABS	8,275	5,937	3,644	1,196	1,925	315,588

Source: IABS. Standard deviations in parentheses.

Table 3: Regional correlations of ethnic concentrations, 1985

	Turkey	Yugoslavia	Italy	Spain	Greece
Turkey	1.00				
Yugoslavia	0.47	1.00			
Italy	0.35	0.63	1.00		
Spain	0.31	0.28	0.47	1.00	
Greece	0.43	0.59	0.53	0.43	1.00

Correlations of ethnic concentrations between over 148 regions of West-Germany, including West-Berlin. Source: IABS.

Table 4: Determinants of years of schooling

	Immigrants' children				Natives' children	
	(1)	(2)	(3)	(4)	(5)	(6)
Mother's education	0.074	0.080	0.079	0.076	0.314***	0.314***
	(0.070)	(0.070)	(0.070)	(0.070)	(0.053)	(0.053)
Household income	-0.032	-0.025	-0.024	-0.021	0.446***	0.446***
in 1,000 Euros	(0.049)	(0.045)	(0.045)	(0.045)	(0.118)	(0.118)
Male	0.011	0.016	0.023	0.023	-0.191	-0.191
	(0.169)	(0.168)	(0.168)	(0.168)	(0.138)	(0.138)
Adults in household	0.071	0.061	0.057	0.056	-0.363***	-0.363***
	(0.122)	(0.124)	(0.124)	(0.123)	(0.097)	(0.097)
Children in household	-0.183	-0.188*	-0.184*	-0.186*	0.005	0.005
	(0.116)	(0.108)	(0.110)	(0.110)	(0.103)	(0.103)
Mother's German skills	-0.200	-0.181	-0.179	-0.180		
	(0.152)	(0.149)	(0.149)	(0.149)		
Mother any education	-0.955**	-0.894*	-0.842*	-0.798*		
in Germany	(0.455)	(0.472)	(0.466)	(0.481)		
Mother's age at	-0.014	-0.016	-0.017	-0.017		
migration	(0.018)	(0.017)	(0.018)	(0.018)		
Not born in Germany	-0.605***	-0.614***	-0.621***	-0.601***		
	(0.228)	(0.229)	(0.230)	(0.228)		
Concentration own		-0.180**	-0.203***	1.104	-0.004	0.068
ethnicity		(0.075)	(0.078)	(2.035)	(0.025)	(0.734)
Ethnic capital		0.020	-0.018	0.015		
		(0.113)	(0.116)	(0.150)		
Native capital			0.266	0.488	-0.223	0.315
			(0.258)	(0.413)	(0.224)	(5.506)
Ethnic concentration*				-0.015		
Ethnic capital				(0.088)		
Ethnic concentration*				-0.093		-0.006
Native capital				(0.183)		(0.059)
Constant	-23.0	-22.5	-12.6	-16.4	13.2	6.2
	(52.2)	(51.9)	(53.7)	(53.8)	(34.5)	(80.9)
Observations	687	687	687	687	1,389	1,389
R-squared	0.079	0.087	0.089	0.089	0.106	0.106

Standard errors (in parentheses) are clustered at the 1985 household level.
*** p<0.01, ** p<0.05, * p<0.1

Dependent variable: Years of schooling. Other controls: Ethnic fixed effects, year of birth, dummy if residence before age 11 unobserved. Source: SOEP 1985-2010.

Table 5: Determinants of years of schooling, sample selection and endogeneity

	(1)	(2)	(3)	(4)	(5)
Model	Baseline	Attrition	Endogeneity	(2) and (3)	Propensity score
Mother's education	0.079	0.105	0.080	0.104	
	(0.070)	(0.073)	(0.069)	(0.072)	
Ethnic concentration	-0.203***	-0.175**	-0.181**	-0.146	-0.117*
	(0.078)	(0.082)	(0.091)	(0.096)	(0.066)
Ethnic capital	-0.018	0.006	0.020	0.027	
	(0.116)	(0.117)	(0.113)	(0.115)	
Native capital	0.266	0.130			
	(0.258)	(0.275)			
Excluding those with high attrition prob	no	yes	no	yes	
2^{nd} stage IV	no	no	yes	yes	
F-statistic (1,387/362)			487.6***	408.6***	
Constant	-12.622	-1.510	-22.501	-5.044	10.519
	(53.698)	(60.130)	(51.177)	(58.196)	(0.326)
Observations	687	619	687	619	597
R-squared	0.089	0.093	0.087	0.092	

Standard errors (in parentheses) are clustered at the 1985 household level.
*** p<0.01, ** p<0.05, * p<0.1

Dependent variable: Years of schooling. Other controls: Same as in table 4. The second degrees of freedom for column 3 is 387, for column 4 it is 362. Source: SOEP 1985-2010.

Table 6: Results by gender and by parent

	(1)	(2)	(3)	(4)
	Men	Women	Men	Women
	Mother's education	Mother's education	Father's education	Father's education
Parent's education	-0.076	0.240**	0.018	0.153
	(0.082)	(0.095)	(0.073)	(0.093)
Ethnic concentration	-0.197**	-0.176	-0.248**	-0.173
	(0.100)	(0.115)	(0.099)	(0.136)
Ethnic capital	0.039	-0.073	-0.002	-0.189
	(0.165)	(0.171)	(0.167)	(0.208)
Native capital	0.184	0.146	0.289	-0.049
	(0.350)	(0.367)	(0.335)	(0.396)
Constant	8.8	-47.5	2.9	-120.4
	(67.2)	(82.9)	(66.3)	(85.2)
Observations	363	324	372	330
R-squared	0.099	0.135	0.119	0.130

Standard errors (in parentheses) are clustered at the 1985 household level.
*** p<0.01, ** p<0.05, * p<0.1

Dependent variable: Years of schooling. Other controls: Same as in table 4. Source: SOEP 1985-2010.

Table 7: Robustness

	(1) Full sample	(2) Full sample	(3) Full sample	(4) measured exposure	(5) Born in Germany
Mother's education	0.079	0.093	0.081	0.107	0.008
	(0.070)	(0.070)	(0.070)	(0.074)	(0.084)
Ethnic concentration	-0.203***	-0.281***	-0.203**	-0.176**	-0.228**
	(0.078)	(0.083)	(0.094)	(0.083)	(0.092)
Ethnic capital	-0.018	-0.014	0.004	0.027	-0.045
	(0.116)	(0.116)	(0.131)	(0.121)	(0.167)
Native capital	0.266	0.248	0.286	0.243	0.422
	(0.258)	(0.538)	(0.312)	(0.260)	(0.317)
Other regional controls	no	yes	no	no	no
1985 residence dummies	no	no	yes	no	no
Constant	-12.6	53.2	0.4	-9.3	17.6
	(53.7)	(72.5)	(56.5)	(56.0)	(61.5)
Observations	687	687	687	597	493
R-squared	0.103	0.109	0.116	0.082	0.080

Standard errors (in parentheses) are clustered at the 1985 household level.
*** p<0.01, ** p<0.05, * p<0.1

Dependent variable: Years of schooling. Other controls: Same as in table 4. Source: SOEP 1985-2010.

Appendix

Table 8: Probit regression of panel attrition

Mother's education	-0.044
	(0.028)
Mother any education	-0.412**
in Germany	(0.186)
Household income	-0.015
in 1,000 Euros	(0.046)
Mother's German skills	-0.159**
	(0.068)
Male	0.037
	(0.087)
Not born in Germany	0.284**
	(0.140)
Mother's age at	-0.015*
migration	(0.008)
Adults in household	0.002
	(0.056)
Children in household	0.115***
	(0.043)
Ethnic concentration	-0.054
	(0.038)
Ethnic capital	0.005
	(0.066)
Native capital	-0.008
	(0.137)
Intends to return	-0.241**
	(0.100)
Constant	0.420
	(1.745)
Observations	1,118
Pseudo R-squared	0.231

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Dependent variable: Dummy variable for staying in panel till completion of education. Other controls: Ethnic fixed effects, year of birth dummies, dummy if residence before age 11 unobserved. Source: SOEP 1985-2010.