Yaman, F. (2011). Ethnic externalities and 2nd generation immigrants (Report No. 11/08). London, UK: Department of Economics, City University London.

CITY UNIVERSITY LONDON

City Research Online

School of Social Sciences
CITY UNIVERSITY LONDON

Department of Economics

Ethnic Externalities and 2nd Generation Immigrants

Firat Yaman[1]
City University London

Department of Economics
Discussion Paper Series
No. 11/08

[1] Department of Economics, City University London, Social Sciences Bldg, Northampton Square, London EC1V 0HB, UK.
Email: firat.yaman.1@city.ac.uk. F: +44 (0)20 7040 8580.

Ethnic Externalities and 2nd Generation Immigrants

Firat Yaman[1]

I analyze the role of regional ethnic capital – defined as the average years of schooling of ethnic groups – in the educational attainment of young second generation immigrants in Germany and whether results are sensitive to regional aggregation. I find evidence for externalities of ethnic capital for ethnic groups at the regional level. A higher average education of ethnics makes attendance of higher-quality secondary schools more likely. Moreover, the marginal effect of the externality is increasing in the ethnic concentration in the region. However, if higher than regional aggregates are used for the measurement of ethnic capital, no externalities are detected.

JEL classification: I20, J15, R23

Keywords: 2nd generation immigrants, ethnic capital, ethnic concentration

[1]City University London. This paper was written as part of the Marie Curie Research Training Network on 'Transnationality of Migrants' TOM, which is funded by the European Commission through the Human Resources and Mobility action of its Sixth Framework Programme (EC Contract No. MRTN-CT-2006-035873). I am indebted to my thesis supervisors Jason Abrevaya and Daniel Hamermesh.

1 Introduction

The assimilation of immigrants to natives in their host countries has been an active research field since the publication of Chiswick's (1978) paper on the "Americanization" of the earnings of foreign-born men, but in the subsequent literature the conclusion that immigrants assimilate to the native population has been questioned (see for example Borjas (1985), and for the case of Germany Pischke (1992)). While the question of immigrant assimilation remains inconclusive, it would hardly be a surprise to find a lasting gap between natives and immigrants even after long durations of stay. Some human capital components (or prejudices) might never be fully attained (or overcome), such as a full command of the language of the host country.

The prospects for 2nd generation immigrants, that is the children of those who immigrated to the country of interest, should not be hampered by the "shock" that some of the human capital accumulated earlier in life is lost or rendered useless and that new human capital specific to the new environment needs to be acquired. This optimistic view (or hope) has not been met in many European countries. 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." 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 "autochthon" 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.

Ethnic capital, introduced by Borjas (1992) could provide the missing explanation for the apparently persistent gap (albeit not for the difference between the European and the classical immigration countries). The idea put forward by Borjas is that the average human capital endowment of a certain immigrant group exerts an externality on the human capital accumulation of a child belonging to that group. In a follow-up paper Borjas (1995) specifies the transmission mechanism of ethnic capital as neighborhoods and reports some evidence for the importance of this channel, but relies on the assumption that ethnic capital is the same for all neighborhoods, calling it the "skill-invariance" assumption.

In this paper I relax the skill-invariance assumption and study the effects of ethnic capital –

measured as average years of schooling of an immigrant group within a region – on the educational attainment of immigrants' children. I then estimate the same model with ethnic capital measured at more aggregated geographical units, namely at the state, and at the country level. The latter corresponds to skill-invariance as in Borjas (1995).

The main findings of this paper are:

1. A higher average education of migrant households in the region has a positive effect on schooling outcomes of their children, and the effect is amplified through the ethnic concentration in the region.

2. Results *are* sensitive to the regional aggregation. Inference on ethnic capital results based on national averages need to be interpreted with caution.

The idea of ethnic capital is taken up by Gang and Zimmermann (2000) in their study of schooling attainment of children of immigrants who were born in Germany or arrived before the age of 16. The authors find that migrants' 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 the group in Germany and must be considered a very crude measure. This, taken together with the limited number of observations of immigrant children (ranging from 116 Greeks to 295 Turks) make the

results unconvincing. A similar study for Denmark by Nielsen et al. (2001) echoes the finding of no influence of parental education for whether 2nd generation immigrants finish a qualifying education, whereas the opposite is found by Jakobsen and Smith (2003), possibly due to using a different data set and using only mothers' education. The concentration of 1st and 2nd generation immigrants in the childhood municipality reduces the probability of finishing a qualifying education, whereas ethnic capital - including the average years of education of 1st generation immigrants of the relevant immigrant group - has no impact.

While I also consider the intergenerational transmission of human capital, I do so only to purge the regressions from potential omitted variable bias. My actual focus is on the regional composition of different ethnic groups and to what extent immigrants' children's exposure to their reference group helps or inhibits their educational attainment. Thus, this paper is loosely related to the abundant intergenerational human capital mobility literature, 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).

Of more importance is the link 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.

The distinction between ethnic capital, immigrants, and 2nd generation immigrants can be a source of confusion. I introduce the following terminology: An *immigrant* is a person who was born outside of Germany. Children of immigrants born in Germany I call *2nd generation immigrants* or simply *immigrants' children*. Somebody who is or was a non-German citizen OR (non-exclusive) who is an immigrant is an *ethnic* and the average years of schooling of ethnics older than 30 are used to form measures of *ethnic capital*. This is done separately for every region and every ethnic group. Every ethnic, and consequently every immigrant, belongs to a *group*, defined by country of birth and/or initial citizenship. A *native* is somebody born in Germany and holding German citizenship, but NOT through naturalization. With these definitions, 14% of the West-German population are immigrants. 27% belong to a household with one immigrant parent belonging to one of the groups considered in this paper.

2 Theory

I follow the group membership and conformity literature in that I assume that children of immigrants identify with children of the same ethnic background, thus introducing the ethnic externality. Predecessors of this line of work are Currarini et al. (2009), Akerlof (1997), and Akerlof and Kranton (2002).[2] Suppose that individuals choose their human capital h, and denote utility associated with human capital $U(h)$. This should be understood as the present value of all monetary and non-monetary benefits of human capital. There are diminishing marginal returns to education. In particular, let utility be given by a quadratic function of human capital[3]:

$$U = \alpha_0 h - \frac{\alpha_1}{2} h^2 \quad \alpha_0 > 0, \alpha_1 > 0, h \leq \frac{\alpha_0}{\alpha_1}$$

There are several costs to human capital investment. First, there is a direct cost, $C_d(h, X, \varepsilon)$, which is increasing and convex in h, and which is allowed to vary along observable demographic variables X and ability ε. There are also three different costs of deviation, reflecting the agent's preference for conformity: first, the agent wants to conform to his parents, second, to the children of native Germans in his generation, and third, to the children of immigrants from the same country of origin. Denote the human capital of the parents by h_p, the average human capital of the children

[2] Alternatively, ethnic capital could enter as an externality into the production of human capital as in Borjas (1992), yielding the same reduced form.

[3] A working paper version of this paper considers a general, non-parametric model and arrives at the same comparative statics results.

of native Germans by $\bar{h}_n$ (the subscript standing for *native*) and the average human capital of the children of immigrants to which the agent belongs by $\bar{h}_r$ (the subscript standing for *reference*). Let s_r be the share of the reference group in the region's population or the schooling cohort. Let the non-conformity costs be given by quadratic functions:

$$\begin{aligned}
C_p &= \frac{\alpha_p}{2}(h-h_p)^2 \quad \alpha_p > 0 \\
C_n &= \frac{\alpha_n}{2}(h-\bar{h}_n)^2 + \frac{\tau_n(1-s_r)}{2}(h-\bar{h}_n)^2 \quad \alpha_n \geq 0, \tau_n \geq 0 \\
C_r &= \frac{\alpha_r}{2}(h-\bar{h}_r)^2 + \pi s_r + \frac{\tau_r s_r}{2}(h-\bar{h}_r)^2 \quad \alpha_r \geq 0, \tau_r \geq 0 \\
C_d &= (\alpha_2 \delta X + \varepsilon)h
\end{aligned}$$

The marginal cost of acquiring human capital is decreasing in the human capital that the agent is conforming to, or, put differently, the cross-derivatives of the cost-components are negative: $\frac{\partial^2 C_p}{\partial h \partial h_p} < 0, \frac{\partial^2 C_n}{\partial h \partial \bar{h}_n} < 0, \frac{\partial^2 C_r}{\partial h \partial \bar{h}_r} < 0$. The last argument in the native and reference group cost is the share of the group in the region of interest. This allows for a stronger impact of ethnic capital (and possibly weaker impact of natives' capital) in regions with high concentrations of certain immigrant groups. The ability distribution, parental human capital, the share of the ethnic group s_r, and the average human capital of non-migrant peers (children of natives) are exogenous, but the average human capital of immigrant's children of a certain group $\bar{h}_r$ is determined in equilibrium and needs to be the average of chosen levels of human capital, if the model is to be internally consistent. The above formulation makes some assumptions for analytical convenience.

First, ability enters only in the direct cost component. In principle ability could be part of any or all of the cost and return components. None of them are directly observable, and thus where exactly ability enters the human capital investment decision is not known to the researcher. He can not identify the relative role of ability in, say, C_d from its role in C_p. The same holds for the variables included in X. Second, the decision-maker in this model is the child, not the parent(s), but of course the parents influence the decision through h_p and possibly components of X.[4]

Third, the conformity assumption might seem arbitrary, and in some sense it is. One might as well argue that agents want to be "special". However, a model of distancing oneself from others typically results in the overproduction of the good of interest (human capital) unless one indulges in underperforming with respect to others (see Akerlof (1997) for a discussion). The conformity assumption is in keeping with most of the peer-effect literature, examples including Falk and Ichino (2006), and Bandiera et al. (2010) for effects in work-productivity, Lyle (2009), and Sacerdote (2001) for college success, and among others Bobonis and Finan (2009) for secondary school enrollment. The peer-effect literature relies on (quasi-) experimental designs which confines it to

[4]The model differs in this respect from Borjas (1992). Borjas models the parent's decision as a trade-off between investing in the child and own consumption. In this paper I am focusing on the secondary education (not college) of teenagers in Germany. Thinking of schooling attainment of a teenager in a system with basically free education as foregone consumption seems to be less justified than, for example, the choice for higher education in the United States. Even for the latter case the role of credit constraints is called into question by Cameron and Heckman (2001).

small and selected samples, so that the generality of its results might be questioned. Ultimately, the empirical analysis will determine whether we are justified in believing in the dominance of a conformity effect.

The problem of the agent is to maximize utility net of all costs over the choice of human capital:

$$\max_h U(h) - C_d - C_p - C_n - C_r$$

The corresponding first-order condition is:

$$h = -\frac{1}{\beta}\left(\alpha_0 + \alpha_p h_p + \{\tau_n(1-s_r) + \alpha_n\}\bar{h}_n + \{\tau_r s_r + \alpha_r\}\bar{h}_r - \alpha_2 \delta X - \varepsilon\right) \tag{1}$$

and $\beta = -\alpha_1 - \alpha_p - \alpha_n - \tau_n(1-s_r) - \alpha_r - \tau_r s_r$.

Averaging h over all immigrants' children of one immigrant group and solving for $\bar{h}_r$, we find:

$$\bar{h}_r = \frac{-\alpha_0 + \alpha_2 \delta \bar{X} - \alpha_p \bar{h}_{p,r} - \{\alpha_n + \tau_n(1-s_r)\}\bar{h}_n - \bar{\varepsilon}}{-\alpha_1 - \alpha_p - \tau_n(1-s_r) - \alpha_n} \tag{2}$$

Equation (2) shows that the equilibrium in this model is unique, which is a feature of the quadratic cost functions and consequently the linear first-order condition. Plugging (2) into (1) and denoting $\gamma = -\alpha_1 - \alpha_p - \alpha_n - \tau_n(1-s_r)$ we can write h as:

$$\begin{aligned} h \;=\; & -\frac{1}{\beta}\Bigg(\alpha_0 + \alpha_p h_p + \{\alpha_n + \tau_n(1-s_r)\}\bar{h}_n + \frac{\alpha_r + \tau_r s_r}{\gamma} \\ & \left[-\alpha_0 - \alpha_p \bar{h}_{p,r} - \{\alpha_n + \tau_n(1-s_r)\}\bar{h}_n + \alpha_2 \delta \bar{X} + \bar{\varepsilon}\right] - \alpha_2 \delta X - \varepsilon\Bigg) \end{aligned} \tag{3}$$

One can verify that

$$\begin{aligned} dh/dh_p &> 0 \\ dh/d\bar{h}_n &> 0 \\ dh/d\bar{h}_{p,r} &> 0 \\ \frac{\partial^2 h}{\partial \bar{h}_{p,r} \partial s_r} &> 0 \end{aligned}$$

but $\frac{\partial^2 h}{\partial \bar{h}_n \partial s_r}$ is ambiguous due to the following effect: Imagine an increase in $\bar{h}_n$. This exerts a positive externality on h through the cost reduction in C_n – call this the primary externality. The positive externality is *dampened* if there is an increase in s_r, making the cost component C_n relatively unimportant. An increase in the share s_r decreases the positive effect of increasing $\bar{h}_n$ directly. However, the primary externality increases $\bar{h}_r$, and this exerts a secondary externality through C_r. The second externality is *amplified* by an increase in s_r.

2.1 Identification

Comparing equations (1) and (3) reveals the infamous reflection problem in Manski (1993). Suppose we could estimate the term $\frac{\alpha_r + \tau_r s_r}{\beta}$ in equation (1). Would we be able to interpret these coefficients as evidence of peer-effects (what Manski calls *endogenous* effects)? Or are we measuring the effect of average human capital of the parent generation and average demographics of

the region, as would be suggested by (3) (*exogenous* effects)? Identification of these effects is possible under some circumstances. Brock and Durlauf (2001a), (2001b), and Durlauf (2004) discuss models of social interactions and provide sufficient conditions for identification of endogenous effects. The conditions are not innocuous however. The condition of unbounded support for the regional variables will rarely be met (unless one is willing to omit bounded variables), and clearly in this paper one of the key variables, s_r, is bounded between zero and one. Lee (2007) proposes identification through variation in the group sizes, a method which is taken up by Davezies et al. (2009) and applied to test scores of school-children in Canada by Boucher et al. (2010). In this paper I do not distinguish between endogenous and contextual effects. Identification would be weak at best, and the focus of this study is (necessarily) the presence or absence of ethnic capital effects, rather than through which channel they are operationalized.

The equations derived so far are not estimable, since human capital is not directly observed. Rather, we see a categorical outcome for teenagers: whether they have dropped out of school, have obtained a degree or attend a grammar school (*Gymnasium*), or have finished another secondary school. Details on the schooling system are described in the data section. For now suffice it to say that we can write the human capital model as a latent variable model. That is, we see the grammar school outcome if and only if human capital is above some threshold. Then we have the probability of attending grammar school:

$$P(G = 1) = P(h > \eta_g)$$

I include a region-specific constant in X, so that without loss of generality I can set the average ability $\bar{\varepsilon}$ to zero. Substituting in h from equation (3) and rearranging, the probability can be written approximately as:

$$P(G=1) = P(a_0 + a_1 s_r + a_2 h_p + a_3 \bar{h}_{p,r} + a_4 s_r \bar{h}_{p,r} + a_5 X + a_6 \bar{X} + a_7 s_r \bar{X} + a_8 \bar{h}_n + a_9 s_r \bar{h}_n > \varepsilon) \quad (4)$$

where the a are functions of the structural parameters α and τ, and I have already verified that a_2, a_3, a_4 and a_8 are positive, and a_9 could not be signed unambiguously. However, if externalities of native peers are weakened by stronger ethnic concentrations, a_9 should be negative. The region of residence and the parent are exogenous for the child in a behavioral sense (the child doesn't choose either). Yet, statistical endogeneity is still present if ability of children and of parents correlate (through h_p and X) and parents sort into regions of similar average human capital (through $\bar{h}_p$ and $\bar{X}$). Intergenerational transmission of ability, if present, possibly obeys a regression to the mean, but evidence for its importance is given by Black et al. (2005).

'Table 1 here.'

The bias from selection into regions is attenuated by the high correlations that concentrations of different ethnic groups exhibit. Table 1 reports those correlations across regions. They range from 0.80 between Russians and Arabs to 0.95 between Turks and Polish. If the choice of residence depends on ability, it seems fair to assume and it is supported by the high correlation of ethnic shares that the choice is similar for the different ethnic groups. Thus, I can compare two ob-

servations with the same (assumed) expected ability in the same region, but with different degrees of exposure to the ethnic group, since the ethnic groups differ in their numbers for Germany as a whole. To further deal with selection bias I use the information contained in observing a parent in one region (and not another). I follow a non-parametric approach due to Dahl (2002), circumventing distributional assumptions on and the modeling of tastes for locations. The procedure is: in a first step immigrants of a certain education are grouped together and to every immigrant I ascribe a measure ρ, corresponding to the share of immigrants of this education being observed in the observation's region. A polynomial of ρ is then included as a term in the main regression. The variable ρ or a function of it can be regarded as a measure of how typical the immigrant is with respect to its residence of choice.[5] Suppose that from observable characteristics we would have expected an immigrant to move to a region with a high share of other immigrants, but that he has not done so, because his "ability" was high and he did not feel the need to live close to other immigrants. Then ρ would be low, and some inverse function of it could be used as a proxy for his ability.

[5] I thank Simone Bertoli for suggesting this method. A parametric version of the same method has been developed by Lee (1983), which is itself a generalization of the Heckman selection correction to multiple first stages.

3 Secondary education in Germany

The sixteen states of Germany enjoy a high degree of autonomy in the design of their schooling systems. Primary education runs from first to fourth grade in all states except for Berlin and Brandenburg, where children finish primary school after sixth grade. Traditionally secondary education has been divided into three kinds of schools, which is in stark contrast to the high-school system in the USA. The Hauptschule was designed to prepare students for a manual profession and graduation is after 9^{th} grade. The Realschule lasts one year longer and is supposed to prepare students for administrative and lower white-collar jobs. The *Gymnasium* is the most prestigious of the traditional schooling forms and is the only one which runs through grades 11, 12, and 13, these being called *Oberstufe* and preparing students for higher education. The *Gymnasium* is now undergoing substantial reforms, including cutting the 13^{th} grade. While this is already the case for some states of East Germany, only the state of "Saarland" in West Germany had made the transition to cutting the last year by 2009. In 2006 all West-German states were still under the 13^{th} grade regime. The *Gymnasium* finishes with a graduation exam, the *Abitur*, the completion of which is a sufficient (but not necessary) condition to apply for and attend a university. While other paths to higher education are possible, the Abitur is still the most common one. Mixed school types exist, but even in those the grades 11 through 13 are considered part of the *Gymnasium*. The conditions for attending the *Gymnasium* differ across states. Some states are more selective in admitting children to

the *Gymnasium* than others and in most states parents can decide to send their children to a school against the recommendation of school teachers. In the empirical analysis this will be accounted for by the inclusion of state- (or region-) fixed effects, which in the latent variable is equivalent to state-specific threshold values η_g.

4 Data

All estimations are performed on the 2006 cross-section of the German Microcensus, an annual representative household survey of 1% of the German population. I have excluded East Germany (except for Berlin) from the analysis, because of the virtual absence of a history of migration to East Germany. The relevant population are household members below the age of 20, living with at least one immigrant parent or guardian and who fall into one of three categories: 1) They attend the Oberstufe of a *Gymnasium* (grades 11, 12, or 13), or they have graduated from a *Gymnasium*, 2) They do not belong to the first category, but have obtained another schooling diploma (such as Real- or Hauptschule, and possibly had some vocational training), 3) They are not in school, but also have no schooling degree (drop-outs). These are the three possible outcomes of all regressions. Some observations reported having obtained a secondary degree, but not from which school, so that I excluded them. Since graduation from a *Gymnasium* typically occurs at the age of 19 and is often followed by attending college and a related relocation, sample selection is a potential

problem. I find that 86% of all non-married observations of age 19 still live with at least one parent, compared to 93% of all 18-year-olds. For ethnics the numbers are not different from the overall population, so that sample selection affects some 7-15% of the relevant population. 13% (11%) of those 19-year old ethnics who live without parents, and 29% (7%) of those who live with a parent attend or graduated from *Gymnasium* (are school-dropouts), so that teenagers with a high educational attainment are over-represented in the sample. Unfortunately, I do not observe the residence of those who live without parents before they moved out of their parents' household, nor can I determine whether they are children of immigrants. If those who have left their parents' household have stayed in the same region (or moved to regions with similar regional characteristics as their parents' home) sample selection should not be a problem. The estimates will be biased only if those who obtained a higher degree have moved to regions with different characteristics compared to those with lower degrees.

4.1 Ethnic groups and their education

I classify ethnics and immigrants as Yugoslavian (Serbia/Montenegro, Croatia, Bosnia-Hercegovina, Slovenia, Former Yugoslavia), Russian (previous Soviet Republics), Polish, Turkish, Italian, and Arabs (North African country, Iraq). The classification can be criticized since the groups exhibit high degrees of heterogeneity (the different ethnicities and history of the former Yugoslavia being

the most obvious example), but members of a group share a common language. Except for Arabs, these groups above are the most numerous ones in Germany. Yugoslavia, Turkey, and Italy are the countries which sent most immigrants in the framework of the guest-worker recruitment in the sixties and early seventies. Immigrants from the former Soviet-Union and from Poland are mainly immigrants who could claim German descent, giving them the right to obtain German citizenship. Arabs are the least numerous group and are included because the public discourse on 2nd generation immigrants has often focused on Turkish and Arab youth. Since their numbers are too small in some regions to base ethnic capital variables on, I have run all regressions with and without Arab children, and obtained comparable results. Every immigrant belongs to one of the immigrant groups only. I drop households in which the two parents belong to different immigrant groups. The household is then assigned to one of the groups if one of the parents belongs to it. For example, Turkish-German and Turkish-Turkish households are categorized as Turkish,[6] but Turkish-Polish households are excluded. The final sample consists of children who fall into one of the three educational categories noted above and who live with at least one immigrant parent belonging to one of the ethnic groups.

The human capital of a parent is approximated by assigning the German equivalent years of schooling to ISCED classifications of his/her education, where I use a conversion table from OECD (2004, page 308). Ethnic capital $\bar{h}_{p,r}$ for group j in region r is the average years of schooling of

[6]Dummies for mixed households are included in all regressions.

ethnics of origin j above the age of 30 in region r.

I use several individual and household (the variables in X) and regional variables as controls in all regressions.[7] The income variable in the Microcensus is only categorical (with 24 categories), so that I approximate income with the average of the bounds of the category into which the observation falls.

4.2 Regional aggregation

The Microcensus allows the division of West-Germany including Berlin into 104 regions, with on average 5,300 observations per region. Thus, the average population of the regions amounts to 500,000. Every region is either one of the major cities with populations 500,000 or above, or medium to small size cities with the surrounding counties. Given that our interest lies in children between ages 15 and 19, these regions are too big. It seems natural to assume that teenagers would want to conform to peers and role models that are visible for them, and their day-to-day contacts will take place within city and school districts. A smaller regional unit would be desirable, but is not available. On the other hand, it is well-known that ethnic and social segregation within cities

[7] In particular parents' income, employment status, household size, sex, age indicators, ethnic dummies, years since and age at migration for both parents, the share of immigrants in the ethnic group, and the share of German citizens in the ethnic group.

is not as pronounced in Germany as it is in the United States (see Musterd (2005)), so that the bias from aggregation might not be too severe. To the extent that it is present, the bias works against the hypothesis of the presence of ethnic capital externalities (estimates will be biased towards zero), because, if the predictions hold, teenagers with less educational success will be assigned higher than "true" ethnic capital and vice versa.

For all regions I compute the shares of natives' children falling into one of the three school categories, thereby proxying for $\bar{h}_n$. I use all observations older than thirty to compute the average years of schooling of ethnics (for every group) to proxy $\bar{h}_{p,r}$, and to compute the shares of the ethnic groups in the population of the region, s_r.

Since the regional variables are constructed as averages and shares of a 1% sample, they will be measured with error. The median of observations in a region is 4,587, the mean is 5,319. Measurement error is thus greater for smaller regions and smaller immigrant groups. As a robustness check I run the regressions excluding observations for whom ethnic capital is based on less than 20 observations. For the estimates that correct for the endogeneity of the regions I form four groups of ethnics, ordered by their years of schooling. For every region I compute the share of ethnics in one educational group being observed in that region. This share is denoted by ρ. A quadratic polynomial is then included in the main regressions.

4.3 Descriptive Statistics

Summary statistics are presented in tables 2, 3, 4, and 5. Table 2 summarizes the educational and other variables of ethnics older than 30. Those observations are used to form the regional variables. Some differences are striking, such as the much lower average years of education of Turks, but also of Arabs, Italians, and Yugoslavs, compared to "Germans" (here including all omitted ethnic groups) and other ethnic groups. Polish ethnics seem to have the most favorable educational background. We also see that the average exposure to one's own ethnic group differs considerably for the immigrant groups. Average concentrations range from less than one percent (for Arabs) to almost 4%. There is considerable variation of concentrations across regions. For example, the standard deviation of the share variable is 1.9 for Turkish, and 0.6 for Arab ethnics. The second table summarizes ethnic capital for immigrants proper and mainly echoes the numbers for ethnics. The high share of ethnic Germans among immigrants from Poland and Russia is reflected by the high share of German citizens among them, since German citizenship was granted unconditionally to immigrants of German descent. The inclusion of immigrant groups which are different along the dimensions of educational background, time and purpose of migration, and linguistic and cultural distance is problematic, but important for identifying the role of regional variables in shaping educational outcomes for immigrants' children. It enables me to compare immigrants' children of different groups (with differing exposure to the own immigrant group)

within and immigrant children of the same group *between* regions. I account for the observable differences of the immigrants by controlling for ethnic- and region-fixed effects and for variables such as age at and years since migration.

'Table 2 here.'

'Table 3 here.'

Table 4 reports means and standard deviations of ethnic capital for the sample. Importantly, we see that one standard deviation of ethnic capital amounts to one year of schooling, pointing to a good degree of regional differences in ethnic capital even within ethnic groups. Table 5 shows differences in schooling outcomes for 19-year olds: 46% of natives' children attend or attended *Gymnasium*, and only 2% of them don't obtain a schooling degree. Only children of Polish immigrants attain a comparable education, while the children of all other groups lack behind to different degrees. An interesting observation in the data (not reported) is that a much higher fraction of immigrants' children start *Gymasium*, but do not graduate. To account for this phenomenon I have included a dummy for ages 18 and 19 in the regressions.

'Table 4 here.'

'Table 5 here.'

5 Results

I estimate two models. The first is OLS on a reduced form of equation (3), the second is an ordered probit model on equation (4), where the outcome for the OLS model is a binary variable (1 for attending *Gymnasium*, 0 otherwise), and the outcomes for the ordered probit are 2 for *Gymnasium*, 1 for graduate of any other school, and 0 for school drop-out. The OLS results are easy to interpret. In particular, the coefficients on the share-interaction terms are marginal effects of ethnic concentrations on ethnic capital externalities. The ordered probit model makes more use of the information on education by including drop-outs as a distinct outcome. All regressions are performed on immigrants' children younger than 20.

5.1 Linear Probability Model

'Table 6 here.'

OLS coefficients for household and individual variables are reported in table 6. The first column reports results for natives' children, column two for immigrants' children and the third adds regional fixed-effects. Results are reported only for educational and significant variables. Two things to note from this table are the significant effect of parental education (contrary to the earlier finding of Gang and Zimmermann (2000)), and the robustness of household characteristics to the

inclusion of regional fixed effects. Most signs are as expected, in particular the positive effects of income and household space per household member. Women are more likely to attend *Gymnasium* and the role of the *adult* variable has been discussed earlier. The positive effect of young household members and the negative effect of total household members is puzzling, but the net effect is negative, possibly due to fewer available resources for each household member. It is noteworthy that age at and years since migration has no significant effect for either parent, and neither does a dummy for German-Ethnic mixed households.[8]

'Table 7 here.'

Table 7 reports results from the same regressions for the regional variables. The first column reports results for natives, column 2 for immigrants but omitting the share interactions, the third column adds the interactions with immigrant shares, and the fourth column includes regional fixed effects. D1 is the share (between 0 and 1) of natives' children attending or recently graduated from *Gymnasium* and D2 is the share of natives' children with another secondary education degree. $\overline{Educ}_r$ is ethnic capital: the average years of schooling of the relevant ethnic group in the region. The coefficient on D1 suggests some positive - albeit weak - externality of other children's schooling. However, one must be cautious in interpreting this quantity because of the reflection problem. A one percentage point increase in *Gymnasium* attendance in the region and a corresponding drop

[8]To save space and focus on the ethnic capital variables I do not report household results for the remainder of this chapter. Suffice it to say that they survive all the following specifications. All results can be requested from the author.

in the drop-out rate increase the likelihood of an individual attending *Gymnasium* by six tenths of a percentage point. For immigrants' children we can compare observations of different ethnicities, so that their ethnic capital variables will differ within a region. Without allowing for share interactions, ethnic capital has a positive impact on the probability of an immigrant's child to attend *Gymnasium*. If the average schooling of ethnics could be increased by an additional year, the probability of attending *Gymnasium* for an immigrant's child would increase by two percentage points. The third and fourth columns demonstrate that this positive externality is sensitive to the share of ethnics in the region. The interaction term for $\overline{Educ_r}$ and ethnic concentration is significantly positive even after inclusion of region fixed effects, while the level effect of ethnic capital becomes essentially zero. The coefficient of $\overline{Educ_r}$ and the interaction term are jointly significant at the 1% level for the model in column (3) and at the 10% level for the model in column (4). Immigrants' children benefit from high ethnic capital more if the concentration of their ethnic group is higher. This confirms the comparative statics result in the model section. Note also that the coefficients on interactions between share and D1 and D2 are negative, but not significantly so. For immigrants' children, D1 and D2 are collinear with region fixed effects, so that the interaction coefficient might be capturing an effect on general regional characteristics.

'Table 8 here.'

Next, I repeated the same estimations, but basing ethnic capital on *state*-wide averages, re-

ducing the number of regions from 104 to 11, and on *national* averages, effectively imposing the skill-invariance assumption from Borjas (1995). The latter can not control for ethnic capital and ethnic fixed effects separately because of collinearity. Ethnic concentrations are still based on the 104 regions. Table 8 shows the main result: State-wide and national aggregation masks the effect of the ethnic capital. In particular, the coefficients on the interaction terms are essentially zero, and the independent concentration effect is negative. At least in this application, assuming skill-invariance changes results critically and might mistakenly lead one to conclude that ethnic capital does not matter, and that ethnic concentrations have an independent, negative effect on schooling outcomes.

'Table 9 here.'

Next, I run the same regressions on certain sub-samples. Table 9 reports results for boys, girls, for observations aged 18 or 19, and for observations for which ethnic capital is based on at least 20 observations in the region. The interaction effect is not present for boys, and seems absent for older observations, suggesting that ethnic capital might play a role for enrolling in *Gymnasium*, but not necessarily for graduation. The interaction effect is present for girls, and when observations with few ethnics in the region are excluded. What explains the differences between boys and girls? One explanation that one might put forth is that boys are more likely to find employment through parents' networks. If they anticipate this, their education might be more strongly determined by

this than by ethnic capital. I have checked the German Socio-Economic Panel for the years 1995 to 2006 to compare the job-finding mechanisms between men and women whose parents immigrated to Germany. While men who found employment reported more frequently having found the job through referrals of a parent or friends, the difference between men and women (26% vs. 22%) was not significantly different from zero.

'Table 10 here.'

'Table 11 here.'

5.2 Ordered Probit Model

As stated earlier, the ordered probit model exploits more information on children's educational attainment by adding school drop-out as another outcome category. Results from this model are reported in table 10. The first column does not include the share interactions, and the results suggest insignificance of ethnic capital. The second includes all interaction terms, and we see the familiar result that positive effects of high levels of ethnic capital are mainly mediated by the exposure to the ethnic group. The third column adds the selection correction regressor ρ and its square. Inclusion of these variables does not alter the main results. Indeed, a likelihood ratio test for the inclusion of ρ and ρ^2 does not reject the hypothesis of their insignificance. The choice of

region by parent seems to contain little – if any – information on unobserved characteristics of the child. Interaction terms in nonlinear models are not marginal effects, i.e. the coefficient on the interaction term is not $\frac{\partial^2 P(Gymnasium)}{\partial Educ \partial share}$. Moreover the marginal effect and its standard error will be different for every data point (see Ai and Norton (2003) for a discussion of this). I thus calculated the marginal effect at different shares for a *typical* observation by setting all characteristics to the sample average, and $\overline{Educ_r}$ to 10.5. I did this for the model where $\overline{Educ_r}$ is based on regional and on state aggregates separately. The results are reported in table 11. We see again that state aggregates fail to detect the marginal effect of ethnic concentration on the ethnic externality. We also see that this effect loses precision as the ethnic concentration increases. Compare these numbers to figure 1, which plots the marginal effect of ethnic capital on the probability of attending *Gymnasium* (y-axis) along ethnic concentration (x-axis), with ethnic capital defined on regions, states, and the country respectively. Every point is one observation in the sample. The effect of ethnic concentration is visibly positive in the first case only (upper left). With state-wide ethnic capital (upper right), one might still guess that an effect is present, but it is smaller in magnitude and has a higher variance for every ethnic concentration. Finally, the third figure (lower left) shows only a cloud.

'Figure 1 here.'

6 Conclusion

The main variables of interest are the regional variables, particularly those associated with ethnic capital. For children of natives there is some correlation between attending *Gymnasium* and between the average attainment of other children. For immigrants' children I find significant effects of ethnic capital. Importantly, ethnic capital externalities are reinforced through ethnic concentrations: The externality is stronger the higher the share of ethnics in a region is. In other words, ethnic capital plays a stronger role if it is more "visible". Adding a variable that accounts for the choice of region does not alter the results, but measuring ethnic capital at high regional aggregates such as states or the whole country fails to detect this role of ethnic concentrations.

Ethnic capital as transmitted through immigrant shares and educational attainment in the regions has a significant influence on the educational attainment of 2nd generation immigrants. Some care has to be taken in interpreting this result. First, one has to consider that we have measured a continuous variable (human capital) by only three categories, not being able to exploit variations within the categories. Second, as described in the section on Germany's secondary education system, the separation of children to *Gymnasium* and other schools happens at a very young age, and few school switches (to or from the *Gymnasium*) occur after this initial assignment.[9] Often

[9] 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

the argument is made that the system is cementing educational careers early on, with little room for up- or downward mobility. At the age of ten or eleven, household characteristics are probably much stronger determinants than other "environmental" factors. Third, one would expect ethnic capital to operate on micro-levels such as schools, or even classes, the workplace, and neighborhoods, which relates also to the first point: We observe people in the *Gymnasium*, but can not say anything on how those influence each other (the within dimension of a category). The spatial comprehensiveness of the data set comes at the cost that such micro-layers can not be identified.

Seen in this light, the statistical presence of educational externalities is remarkable in its own right. Yet, many aspects of the effect remain in the dark. Is the effect an expression of peer-effects, or is it an exogenous/contextual effect, in the sense of Manski (1993)? At what age does it start to influence human capital investment decisions? Answering these questions would dissect the somewhat unsatisfying black-box concept of ethnic capital, but will require more information (and of a different kind) on environmental contexts people at a crucial age for their human capital investments find themselves in.

2006/2007.

References

[1] Ai, C. , and E.C. Norton (2003), "Interaction terms in logit and probit models," Economics Letters, 80, 123-129.

[2] Akerlof, George A. (1997), "Social distance and social decisions," Econometrica, 65(5), 1005-1027.

[3] Akerlof, George A. , and Rachel E. Kranton (2002), "Identity and schooling: Some lessons for the economics of education," Journal of Economic Literature, 40(4), 1167-1201.

[4] Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2010), "Social incentives in the work-place," Review of Economic Studies, 77(2), 417-458.

[5] Black, Sandra E. Paul J. Devereux, and Kjell G. Salvanes (2005), "Why the apple doesn't fall far: Understanding intergenerational transmission of human capital," American Economic Review, 95(1), 437-449.

[6] Bleakley, Hoyt, and Aimee Chin (2008), "What holds back the second generation? The inter-generational transmission of language human capital among immigrants," Journal of Human Resources, 43(2), 267-298.

[7] Bobonis, Gustavo J. and Frederico Finan (2009), "Neighborhood peer effects in secondary school enrollment decisions," Review of Economics and Statistics, 91(4), 695-716.

[8] Borjas, George J. (1985), "Assimilation, changes in cohort quality, and the earnings of immigrants," Journal of Labor Economics, 3(4), 463-489.

[9] Borjas, George J. (1992), "Ethnic capital and intergenerational mobility," Quarterly Journal of Economics, 107(1), 123-150.

[10] Borjas, George J. (1995), "Ethnicity, neigborhoods, and human-capital externalities," American Economic Review, 85(3), 365-390.

[11] Boucher, Vincent, Yann Bramoullé, Habiba Djebbari, and Bernard Fortin (2010), "Do peers affect student achievement? Evidence from Canada using group size variation," IZA Discussion Paper No. 4723.

[12] Brock, William A. and Steven N. Durlauf (2001a), "Interactions-based models," in: Heckman, James J. and E. Leamer, Handbook of Econometrics, vol. 5, 3297-3380.

[13] Brock, William A. and Steven N. Durlauf (2001b), "Discrete choice with social interactions," Review of Economic Studies, 68(2), 235-260.

[14] Cameron, Stephen V. and James J . Heckman (2001), "The dynamics of educational attainment for black, hispanic, and white males," Journal of Political Economy, 109(3), 455-499.

[15] Case, Anne C. and Lawrence F. Katz (1991), "The company you keep: The effects of family and neighborhood on disadvantaged youth," NBER Working Paper 3705.

[16] Casey, Teresa, and Christian Dustmann (2005), "Intergenerational transmission of language capital and economic outcomes," CReAM Discussion Paper 06/05.

[17] Chiswick, Barry R. (1978), "The effect of Americanization on the earnings of foreign-born men," Journal of Political Economy, 86(5), 897-921.

[18] Currarini, S., Matthew O. Jackson, and Paolo Pin (2009), "An economic model of friendship: homophily, minorities and segregation," Econometrica, 77(4), 1003-1045.

[19] Dahl, Gordon B. (2002), "Mobility and the return to education: Testing a Roy model with multiple markets," Econometrica, 70(6), 2367-2420.

[20] Davezies, Laurent, Xavier d'Haultfoeuille, Denis Fougère (2009), "Identification of peer effects using group size variation," mimeo.

[21] Durlauf, Steven N. (2004), "Neighborhood effects," in: Henderson, J.V., and J.F.Thisse, Handbook of Regional and Urban Economics, vol. 4, 2173-2242.

[22] Falk, Armin, and Andrea Ichino (2006), "Clean Evidence on Peer Effects," Journal of Labor Economics, 24(1), 39-57.

[23] Gang, Ira N. and Klaus F. Zimmermann (2000), "Is child like parent? Educational attainment and ethnic origin," Journal of Human Resources, 35(3), 550-569.

[24] Ginther, Donna, Robert Haveman, and Barbara Wolfe (2000), "Neigborhood attributes as determinants of children's outcomes: How robust are the relationships," Journal of Human Resources, 35(4), 603-642.

[25] Goux, Dominique, and Eric Maurin (2006), "Close neighbours matter: Neighbourhood effects on early performance at school," CEE Discussion Paper 68.

[26] Heineck, Guido, and Regina T. Riphahn (2007), "Intergenerational transmission of educational attainment in Germany: The last five decades," IZA Discussion Paper 2985.

[27] Jakobsen, Vibeke, and Nina Smith (2003), "The educational attainment of the children of the danish 'guest worker' immigrants," IZA Discussion Paper 749.

[28] Lee, Lung-fei (1983), "Generalized econometric models with selectivity," Econometrica, 51(2), 507-512.

[29] Lee, Lung-fei (2007), "Identification and estimation of econometric models with group interactions, contextual factors and fixed effects," Journal of Econometrics, 140(2), 333-374.

[30] Lyle, David S. (2009), "The effects of peer group heterogeneity on the production of human capital at West Point," American Economic Journal: Applied Economics, 1(4), 69-84.

[31] Manski, Charles F. (1993), "Identification of endogenous social effects: The reflection problem," The Review of Economic Studies, 60(3), 531-542.

[32] Musterd, Sako (2005), "Social and ethnic segregation in Europe: levels, causes, and effects," Journal of Urban Affairs, 27(3), 331-348.

[33] Nielsen, Helena S. Michael Rosholm, Nina Smith, and Leif Husted (2001), "Intergenerational transmissions and the school-to-work transition of 2nd generation immigrants," IZA Discussion Paper 296.

[34] OECD (2004), "Learning for tomorrow's world – First results from PISA 2003".

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

[36] Pischke, Jörn-Steffen (1992), "Assimilation and the earnings of guestworkers in Germany," ZEW Discussion Paper No. 92-17.

[37] Raaum, Oddbjørn, Raaum, Kjell G. Salvanes, and Erik Ø. Sørensen (2006), "The neighborhood is not what it used to be," Economic Journal, 116(1), 200-222.

[38] Riphahn, Regina T. (2003), "Cohort effects in the education attainment of second generation immigrants in Germany: An analysis of census data," Journal of Population Economics, 16(4), 711-737.

[39] Rury, John L. (2004), "Social capital and secondary schooling: Interurban differences in American teenage enrollment rates in 1950," American Journal of Education, 110(4), 293-320.

[40] Sacerdote, Bruce (2001), "Peer effects with random assignment: Results for Dartmouth roommates," Quarterly Journal of Economics, 116(2), 681-704.

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

Tables and Figures

Table 1: Regional correlations of ethnic shares

	Yugo	Russ	Pole	Turk	Ital	Arab	Pop	density
Yugo	1.00							
Russ	0.86	1.00						
Pole	0.93	0.85	1.00					
Turk	0.93	0.80	0.95	1.00				
Ital	0.91	0.83	0.89	0.89	1.00			
Arab	0.89	0.80	0.87	0.87	0.81	1.00		
Pop	0.21	0.19	0.21	0.19	0.19	0.16	1.00	
density	0.33	0.20	0.34	0.38	0.24	0.43	0.20	1.00

Correlations for 104 regions of West-Germany, including Berlin. Source: Microcensus 2006.

Table 2: Summary statistics, Ethnic groups age>30

	(1) Natives	(2) Yugo	(3) Russ	(4) Pole	(5) Turk	(6) Ital	(7) Arab
educ (mean)	12.4	10.8	11.8	12.4	8.8	10.4	10.2
citizen (%)	97	17	77	77	26	11	44
bornhere (%)	89	6	2	2	8	17	3
share (mean)	n.a.	2.0	2.5	1.6	3.8	1.4	0.8
age (mean)	55.7	49.5	50.6	49.4	45.9	49.2	44.9
observations	371,529	4,428	6,544	4,260	9,546	2,992	1,147

citizen: German citizen. bornhere: Born in Germany. Source: Microcensus 2006.

Table 3: Summary statistics, Immigrants

	(1) Yugo	(2) Russ	(3) Pole	(4) Turk	(5) Ital	(6) Arab
educ (mean)	10.8	11.2	11.7	9.7	10.5	10.4
citizen (%)	35	73	69	38	30	45
age (mean)	46.6	43.4	46.3	45.3	48.3	41.2
observations	8,154	12,574	8,349	15,306	4,890	2,908

Source: Microcensus 2006.

Table 4: Ethnic capital in sample

	(1) Yugo	(2) Russ	(3) Pole	(4) Turk	(5) Ital
Educ (mean)	10.7	11.8	12.4	8.8	10.8
	(1.1)	(1.1)	(0.7)	(0.9)	(1.8)
observations	675	1,210	691	1,684	464

Source: Microcensus 2006. Standard deviations in parentheses. Arabs not reported because of small sample size.

Table 5: Schooling outcomes, 19-year olds

	(1) Native	(2) Yugo	(3) Russ	(4) Pole	(5) Turk	(6) Ital
Gymnasium (%)	46	39	35	47	38	42
Other school (%)	52	56	62	49	56	53
Drop-outs (%)	2	4	3	4	6	5
observations	3,290	322	465	279	630	189

Source: Microcensus 2006. Numbers for natives are from the 70% scientific use file.

Table 6: Determinants of attending Gymnasium, Linear Probability Model

	(1) Natives	(2) Migrants	(3) Migrants
educ parent	0.02***	0.01***	0.01***
	(0.002)	(0.003)	(0.003)
income/1,000	0.04***	0.04***	0.04***
	(0.002)	(0.005)	(0.005)
# hh members	-0.02***	-0.05***	-0.05***
	(0.007)	(0.01)	(0.01)
# young hh members	0.01*	0.03**	0.02*
	(0.008)	(0.01)	(0.01)
hh space/10	0.01***	0.01***	0.01***
	(0.002)	(0.003)	(0.003)
male	-0.12***	-0.09***	-0.09***
	(0.009)	(0.02)	(0.02)
adult	-0.08***	-0.11***	-0.11***
	(0.01)	(0.02)	(0.02)
state fixed effects	yes	yes	no
region fixed effects	no	no	yes
Observations	10,759	3,664	3,664
R-squared	0.11	0.13	0.15

Dependent variable: Binary variable of attending or having completed *Gymnasium*. educ parent is mother's education and father's education if mother is not in the household. Other controls: Ethnic fixed effects, mixed Ethnic- German household, father (mother) immigrant, age at migration father (mother), years since migration father (mother), single-parent household, all parents unemployed. *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Source: Microcensus 2006.

Table 7: Determinants Gymnasium, LPM, Ethnic Capital

	Natives	Migrants		
	(1)	(2)	(3)	(4)
D1	0.59*	0.90	1.32	
	(0.32)	(0.64)	(1.1)	
D1*share			-0.19	-0.65
			(0.37)	(0.47)
D2	0.23	0.53	1.04	
	(0.33)	(0.66)	(1.1)	
D2*share			-0.22	-0.68
			(0.36)	(0.46)
$\overline{Educ_r}$		0.02**	0.006	0.002
		(0.01)	(0.01)	(0.01)
$\overline{Educ_r}$*share			0.01**	0.01**
			(0.004)	(0.005)
share		-0.01	0.09	0.53
		(0.01)	(0.35)	(0.45)
state fixed effects	yes	yes	yes	no
region fixed effects	no	no	no	yes
Observations	10,759	3,664	3,664	3,664
R-squared	0.11	0.13	0.13	0.15

Dependent variable: Binary variable of attending or having completed *Gymnasium*. D1: Share of natives' children in *Gymnasium*, D2: Share of natives' children with secondary education other than *Gymnasium*, omitted category: school dropouts, $\overline{Educ_r}$: Average years of education of reference group. Other controls: see table 6. *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Source: Microcensus 2006.

Table 8: Determinants Gymnasium, LPM, Different Aggregates

	Region		State		Country	
	(1)	(2)	(1)	(2)	(1)	(2)
$\overline{Educ_r}$	0.016	0.004	0.023	0.020	-0.01	-0.01
	(0.011)	0.012	(0.023)	(0.025)	(0.01)	(0.01)
$\overline{Educ_r}$*share		0.009*		0.002		-0.0
		(0.005)		(0.005)		(0.01)
share	-0.01	-0.1**	-0.02*	-0.03	-0.02**	-0.01
	(0.01)	(0.05)	(0.01)	(0.05)	(0.01)	(0.05)
Observations	3,664	3,664	4.041	4,041	4,060	4,060
R-squared	0.15	0.15	0.14	0.14	0.13	0.13

Dependent variable: Binary variable of attending or having completed *Gymnasium*. $\overline{Educ_r}$: Average years of education of reference group. Regressions control for region fixed effects. Regressions control for ethnic effects for the "Region" and "State" aggregates. Other controls: see table 6. *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Source: Microcensus 2006.

Table 9: Determinants Gymnasium, LPM, Robustness

	Boys	Girls	Age≥18	N≥20
	(1)	(2)	(3)	(4)
$\overline{Educ_r}$	0.02	-0.02	-0.001	0.003
	(0.02)	(0.02)	(0.02)	(0.02)
$\overline{Educ_r}$*share	0.001	0.02**	0.005	0.01*
	(0.007)	(0.01)	(0.007)	(0.006)
share	0.59	0.82	0.46	1.04
	(0.62)	(0.68)	(0.59)	(0.50)
Observations	1,936	1,728	2,162	3,387
R-squared	0.18	0.17	0.14	0.16

Dependent variable: Binary variable of attending or having completed *Gymnasium*. $\overline{Educ_r}$: Average years of education of reference group. Regressions control for regional fixed effects. Other controls: see table 6. *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Source: Microcensus 2006.

Table 10: Determinants Gymnasium, Ordinal Probit

	(1)	(2)	(3)
$\overline{Educ_r}$	0.04	0.01	0.01
	(0.03)	(0.03)	(0.42)
$\overline{Educ_r}$*share		0.02*	0.02*
		(0.01)	(0.01)
Share	-0.04	-0.26**	-0.26**
	(0.02)	(0.13)	(0.13)
ρ,ρ^2,	no	no	yes
Observations	3,664	3,664	3,664
LL-value	-2,845	-2,843	-2,842

Dependent variable: Ordinal schooling outcome. $\overline{Educ_r}$: Average years of education of reference group. Regressions control for regional fixed effects. Other controls: see table 6. *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Source: Microcensus 2006.

Table 11: Determinants Gymnasium, Ordinal Probit, Marginal Effect of Ethnic Capital

	Regions	States
	(1)	(2)
share=0.1	0.0091	0.0015
	(1.82)	(0.22)
share=1.2	0.0091	0.0015
	(1.83)	(0.20)
share=2.2	0.0090	0.0014
	(1.77)	(0.17)
share=5.2	0.0089	0.0012
	(1.43)	(0.13)
share=8.2	0.0087	0.0009
	(1.09)	(0.09)
Observations	3,664	4,041

Dependent variable: Ordinal schooling outcome. Regressions control for regional fixed effects. Other controls: see table 6. *** p<0.01, ** p<0.05, * p<0.1. t-statistics in parentheses are obtained with Delta method at sample means for all variables except share. Source: Microcensus 2006.

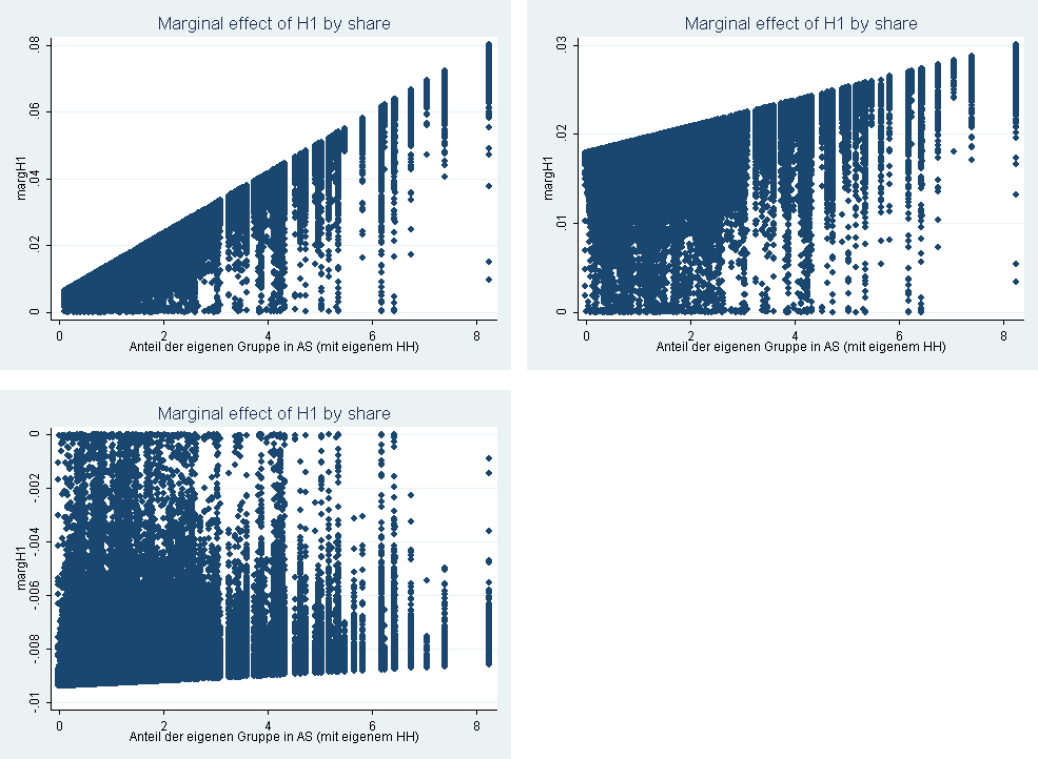

Figure 1: Marginal effect of ethnic capital, different aggregates