# University of Groningen

## Ethnic capital and intergenerational transmission of educational attainment

Postepska, Agnieszka

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*
Final author's version (accepted by publisher, after peer review)

[Link to publication in University of Groningen/UMCG research database](#)

# Ethnic capital and intergenerational transmission of educational attainment[*]

Agnieszka Postepska[†]

## Abstract

This paper studies the role of ethnicity in the intergenerational transmission of educational attainment within the framework outlined in Borjas (1992). Relying on heteroskedasticity to identify parameters in the presence of endogenous regressors, I find evidence that the OLS estimates of the effect of ethnic capital on intergenerational transmission of education are biased upwards due to the transfer of unobserved ability. I also find that while the role of parental capital has declined over time, ethnic capital has a relatively constant effect on intergenerational transmission of educational attainment.

[†]University of Groningen; Nettelbosje 2, 9700 Groningen, The Netherlands, +31503637164; a.postepska@rug.nl

# Summary

This paper studies the role of ethnicity in the intergenerational transmission of educational attainment. Relying on heteroskedasticity to identify parameters in the presence of endogenous regressors, I find evidence that while the OLS estimate of the effect of ethnic capital on intergenerational transmission of education is significantly biased upwards due to the transfer of unobserved ability, the effect of ethnic capital on human capital process formation remains significant.

Keywords: intergenerational transmission of education, human capital formation, ethnic capital, identification through conditional second moments

# 1 Introduction

Borjas (1992) first pointed to the distinct feature of intergenerational transmission which he referred to as the transfer of *ethnic capital*. The overall human capital gained by the group as a whole is expected to have an effect on members of a group. The skills of the next generation depend on parental human capital and on the quality of ethnic environment in which parents make their investment decisions. Borjas (1992) finds a strong and significant effect of the ethnic capital on intergenerational transmission of education. Children's educational attainment, occupational standing and earnings are affected not only by parents education, occupational prestige or earnings but also by the average education or earnings of their corresponding ethnic group. However, Bauer and Riphahn (2007) found no evidence supporting Borjas's hypothesis using 2000 Swiss census data. Similarly, Aydemir, Chen, and Corak (2013) did not confirm the importance of ethnic capital for earnings mobility among children of immigrants in Canada and Nielsen, Rosholm, Smith, and Husted (2003) does not find a convincing evidence in Denmark.

This study contributes to the literature in two ways. (i) I apply the framework developed in Borjas (1992) to a more recent data set which allows for the analysis of the changes in the role of ethnic capital over time. (ii) I improve on the estimation strategy employed in Borjas (1992) by accounting for endogeneity of both parental and ethnic capital. I apply Klein and Vella (2010) constant correlation estimation procedure, which allows for estimation of the effect of parental and ethnic capital on educational attainment in absence of exclusion restrictions.[1]

The paper is organized as follows. The next section provides basic summary statistics. Section 3 explains the estimation method and identification. Section 4 follows with empirical results and discussion. Section 5 concludes.

---

[1] This method has been successfully applied to estimate the intergenerational transmission of education in the US (Farré, Klein, & Vella, 2013), to estimate returns to schooling in the US (Farré, Klein, & Vella, 2012) and in Germany (Saniter et al., 2012), and also to estimate the occupational mobility in China (Holmlund, Lindahl, & Plug, 2011; Emran & Sun, 1988).

## 2 Data and summary statistics

I use the 1977-2014 General Social Survey data. The sample consist of 15390 individuals aged 18-64 born in the United States. I exclude individuals born abroad as well as native Americans and African Americans. Individuals who did not grow up with both parents or for whom information about their own or their parents education attainment is not available are omitted from the sample. Also, only individuals for whom there is at least 30 other individuals in the same cohort of the same ethnic origin are included.[2] Individuals in the sample were born between 1913 and 1992 and they are divided into 5 cohorts. I measure parental capital with father's education and ethnic capital as the average years of schooling of the fathers within the cohort in given region.[3]

The final sample comprises individuals of 26 different origins. Descendants of German, English, Welsh and Irish immigrants are most represented in the sample. Table 1 presents the summary statistics for all variables used in this analysis. Females represent 54 percent of the sample. An average individual is about 46 years old, has about 3 siblings and has completed 14 years of schooling. The average parental and ethnic capital are approximately the same at 11 years of schooling. Forty one percent of all individuals lived in urban setting at the age of 16 and 25 percent lived in the South at the age of 16. Only 10 percent of individuals have at least one parent born abroad. Average educational attainment as well as the average schooling of the ethnic groups increased by around one year between the earlier and more recent samples. For the sake of brevity, more detailed description of the sample and key data characteristics are presented in the Online Appendix section 1.

---

[2]This is an arbitrary threshold and a higher threshold would be more desired. However, higher thresholds resulted in significant sample size loss and more importantly fewer ethnic groups.

[3]Note that Farré et al. (2012) find that the high correlation between parents education makes it difficult to disentangle the effects of mother's and father's schooling. Inclusion of mothers education does not change the main qualitative results.

| Female | 0.54 (0.49) | Parental capital | 11.28 (4.01) |
|---|---|---|---|
| Age | 45.89 (16.62) | Ethnic capital | 11.14 (1.98) |
| Number of siblings | 3.15 (2.30) | Living in a city at 16 | 0.41 (0.49) |
| Years of schooling | 13.86 (2.73) | Living in a Southern state at 16 | 0.25 (0.43) |
| At least one parent born abroad | 0.10 (0.30) | | |

**Notes:** Standard deviations in parenthesis. Number of observations: 15390

# 3  Model and Identification

I follow Farré et al. (2012, 2013) to describe the identification strategy and its interpretation in this framework. In absence of exclusion restrictions, identification of the parameters relies on assumptions about the structure of the error term and heteroskedasticity in the model (see Klein and Vella (2010) for details). Let $edu$ denote the individual's education, $edup$ parental education (parental capital) and $eduav$ the average education of the ethnic group (ethnic capital). The model consist of three equations:

$$edu_i = \gamma_1 edup_i + \gamma_2 eduav_i + \delta_0 X_i + u_i$$

$$edup_i = \delta_2 X_i + v_i^p \text{ and } eduav_i = \delta_3 X_i + v_i^{av} \tag{1}$$

I assume that all variables in $X$ are exogenous and that there are no instruments available for the two endogenous regressors. Exogeneity of $X$ implies that $E(u_i|X_i) = E(v_i^p|X_i) = E(v_i^{av}|X_i) = 0$. Since there are no variables that provide exogenous variation to identify the $\gamma's$, assume for simplicity that the same $X's$ appear in all three equations.

Furthermore, assume that the errors are heteroskedastic and can be defined as:

$$u_i = H_u(X_i)u_i^*, \quad v_i^p = H_v^p(X_i)v_i^{p*} \text{ and } v_i^{av} = H_v^{av}(X_i)v_i^{av*} \tag{2}$$

where $u_i^*, v_i^{p*}, v_i^{av*}$ are correlated homoskedastic error terms and $H_u^2(X_i)$, $H_v^{p2}(X_i)$ and $H_v^{av2}(X_i)$ denote the conditional variance functions for $u_i$, $v_i^p$ and $v_i^{av}$, respectively. The homoskedastic part reflects the transfer of unobserved ability, $u_i^*, v_i^{p*}, v_i^{av*}$ which is independent of the parent's, and child's environment as implied by equation 2. However, the heteroskedasticity implies that once we condition on the vector of exogenous variables X, the transfer of ability contributes

differently to human capital accumulation depending on respective socioeconomic background.[4] Identification in the model is achieved through this variation. Without this variation the mapping from $u^*$'s and $v^{p*}$'s or $v^{av*}$'s is identical to the mapping between $u$'s and $v^p$'s or $v^{av}$'s and therefore we cannot estimate the relationship between the $u^*$'s and $v^{p*}$'s or $v^{av*}$'s. In addition to the assumption of heteroskedastcity, the following constant correlation conditions are necessary for identification:

$$E[u_i^* v_i^{p*}|X_i] = E[u_i^* v_i^{p*}] = \rho^p \text{ and } E[u_i^* v_i^{av*}|X_i] = E[u_i^* v_i^{av*}] = \rho^{av} \tag{3}$$

This error structure implies that the unobservables affecting educational attainment are positively correlated with both parental and ethnic capital. This is consistent with ability being responsible for the confounding effect of parental education and average educational attainment within the ethnic group. However, there is also a possibility that this correlation is negative. It would be the case if there were other unobserved factors that are not captured by ability. It is possible to extend the error structure to accommodate this case without compromising any of the identification in the model (Klein & Vella, 2010; Farré et al., 2012). However, since I find positive correlations in this application, I will refer to the simple structure as defined in 3. Notice however, that the identification fails if there are factors that are related to the exogenous variables in the model and to the correlations between the unobserved factors that are not controlled for. In the context of this paper, the conditional constant correlation assumption implies that, after controlling for all the exogenous variables in the model, the correlation between the unobserved factors affecting individual's educational attainment and parental educational attainment or average educational attainment in the ethnic group remains constant. Therefore, the identification would fail if the correlation between the transfer of unobservables was affected by individual's behavior or environment. The heteroskedasticity implies that the contribution of the unobservables to the formation of educational attainment differ depending

---

[4]This is one of the possible error structure. Klein and Vella (2010) show that other structures are consistent with the constant correlation coefficient assumption.

on characteristics.

To summarize, both heteroskedasticity and constant correlation between the homoskedastic error term in the child's educational attainment equation and the parental schooling or the ethnic capital equation are necessary for identification. Consider the latter condition first. If unobserved ability is transferred genetically, than this assumption is clearly satisfied. This approach was successfully applied in Farré et al. (2012, 2013). Since systematic differences in the "quality" of cohorts of immigrants have been found (Borjas, 1987, 2006), individual unobserved ability is likely to be correlated with the average unobserved ability of the ethnic group. Moreover, ethnic features are passed on from the parents to the children in the form of *cultural capital* (Bourdieu, 2011) which comprises formal education attainment as well as norms, beliefs, attitudes and skills that originate in culture that is shared by an ethnic group (Portes, 2000; Rosen, 1959). This form of capital is internalized during socialization process through exposures to role models within the family and is enacted regardless of the presence of social interactions with other group members (Bourdieu, 2011; Portes, 2000). A relevant example of an expression of such capital is the above average performance of Asian students in the US which is often attributed to the heavy weight put on education as a vehicle for upward mobility in Confucian cultures (Portes, 2000; Kao & Thompson, 2003). Moreover, Borjas (1995) finds that neighborhood effects cannot account for the entire impact of ethnicity on intergenerational transmission of education which can be interpreted as an existence of a constant element of the ethnic capital transmission.

Nevertheless, since parents can shape children contacts with ethnic group members and time and origin of migration correlates with the cohort quality, the effect of the transfer of the unobserved ability and cultural capital differs depending on observed characteristics - second condition required for identification. Depending on when and which country are the parents migrating from, they will be either positively or negatively selected and therefore the $H_v^{p2}(X_i)$ and $H_v^{av2}(X_i)$ will not be the same across individuals. In addition to peer effects, heteroskedacticty

in the child's schooling equation is granted by the fact that parents will invest less effort in child's education in favorable ethnic environment and more in less favorable (Bisin & Verdier, 2001; Patacchini & Zenou, 2011). Also, since parents apply different ethnic socialization models to sons and daughters (Suárez-Orozco & Qin, 2006; Dion & Dion, 2001) gender introduces another source of heteroskedasticity. Moreover, heteroskedasticity in the child and parental educational attainment also arises due to regional differences in access to educational institutions (Farré et al., 2012) while place of birth of the parents can affect attachment to the ethnic community which may influence the transfer of unobservables.[5]

The above error structure allows construction of control functions which inclusion in the main equation makes estimation of the unknown parameters $\gamma = \{\gamma_1, \gamma_2\}$ feasible. This is obtained by including the consistent estimates of $v_i^{av}$ and $v_i^p$ in the child's education equation. Let $\lambda_1 = \frac{Cov(u_i, v_i^p)}{Var(v_i^p)}$ and $\lambda_2 = \frac{Cov(u_i, v_i^{av})}{Var(v_i^{av})}$. Then we can rewrite the error term $u$ as $u_i = \epsilon_i + \lambda_1 v_i^p + \lambda_2 v_i^{av}$ which explicitly shows why heteroskedasticity is necessary for identification. If all errors are homoskedastic, the control function has the same impact across all individuals, i.e. $\lambda_1$ and $\lambda_2$ are constant. Let $A_1(X_i) = \rho^p \frac{H_u(X_i)}{H_v^p(x_i)}$ and $A_2(X_i) = \rho^{av} \frac{H_u(X_i)}{H_v^{av}(X_i)}$. Then, under the conditional correlation assumption in equation 3, we can rewrite the above error term as $u_i = \epsilon_i + A_1(X_i)v_i^p + A_2(X_i)v_i^{av}$. The fact that both $A_1(X_i)$ and $A_2(X_i)$ are non linear in $X_i's$ grants identification of the parameters of the child's education equation:

$$ edu_i \quad = \quad \delta_0 X_i + \gamma_1 edup_i + \gamma_2 eduav_i + \rho^p \frac{H_u(X_i)}{H_v^p(X_i)}v_i^p + \rho^{av}\frac{H_u(X_i)}{H_v^{av}(X_i)}v_i^{av} + \epsilon_i \qquad (4) $$

## 4 Empirical strategy

Turning to the implementation of the identification strategy outlined in previous section, first, consider the OLS estimates of intergenerational transmission, which are presented in the first column of table 3.[6] In line with existing literature, I find that each additional year of average

---

[5]See the Online Appendix section 2 for illustrative example for the interpretation of the error structure and underlying assumptions.

[6]All results are presented in table 8 in the Online Appendix section 4.

and parental schooling increases child's education by 0.138 and 0.237 years respectively. Both coefficients are significant at 1 percent level.

I follow closely Farré et al. (2012) in the estimation strategy.[7] Since there are two endogenous regressors, parental education and ethnic capital, I first estimate these two equations using OLS. Next, the conditional variance in both equations is estimated using non linear least squares. I use exponential function to model the conditional variance in all equations. The last step involves simultaneous estimation of the heteroskedastic index and the coefficients of the main equation. This is obtained by standard iterative procedure.

## 4.1 Parental and ethnic capital equations

All results are presented separately for the whole sample, as well as for the sample used in Borjas (1992) (the pre-1989 sample) and the post-1989 sample. This allows for comparison with the results in Borjas (1992) but also reveals trends over time. The sets of variables included in the parental and ethnic capital equations are almost the same. Since the age of the parent is not known directly, I include the age of the child (and age squared) in both of the equations which, combined with the dummy variables indicating the cross section, controls for the age of the parent. Dummy variables for regions control for geographic differences in educational attainment that might result from local labor market specific needs or differential access to educational institutions. I also include dummy variables indicating whether the child was living in the South or in the city at the age of 16. Since this information is not available for the parents, I use the information for the children as proxies. In the ethnic capital equation, I also include a dummy indicating whether at least one parent is foreign born. Similarly, in the parental schooling equation, I include a dummy variable indicating whether the father is foreign born. As the results of the estimation of the conditional means are not of a prime interests and they appear to be in line with the literature, I refer the reader to table 6 in the Online Appendix section 4 for details.

---

[7]Details of estimation are explained in the Online Appendix section 3.

As one of the identifying assumptions requires the presence of heteroskedasticity in either the children's schooling or parental and ethnic capital equations, table 2 presents the results of the White and Breush-Pagan tests. The null hypothesis of homoskedastic errors is strongly rejected confirming the presence of heteroskedasticity in both equations.

Table 2: Heteroskedasticity in parental and ethnic capital equations

|  | 1977-2014 | | 1977-1989 | | 1990-2014 | |
|---|---|---|---|---|---|---|
|  | Parental capital | Ethnic capital | Parental capital | Ethnic capital | Parental capital | Ethnic capital |
| Breush-Pagan test | 544.42 | 5030.34 | 138.11 | 1540.66 | 469.63 | 3437.97 |
| White test | 1095.95 | 3126.06 | 403.03 | 1242.04 | 760.10 | 2028.23 |
| Number of observations | 15390 | | 5329 | | 10061 | |

## 4.2 Education transmission equation

The construction of the correction terms requires estimates of the heteroskesdticity indexes in all three equations. Table 7 in the Online Appendix section 4 presents the results for the heteroskedasticity indexes for parental and ethnic capital equations[8] as well as for the primary equation noting that the latter is estimated simultaneously with the coefficients of the primary equation.

Now, turn to the main results of this paper relating child's human capital accumulation to parental and ethnic capital. First two columns of table 3 present the OLS and control function (CF) estimates of the primary equation for the entire sample. All results are presented in table 8 in the Online Appendix section 4. I find that accounting for endogeneity reduces both the coefficient on parental education (from 0.24 to 0.18) and on ethnic capital (from 0.14 to 0.07). This confirms that OLS coefficients are confounded by the endogeneity of parental and ethnic capital. The coefficients on control functions are both statistically significant at 1 percent significance level confirming the importance of unobserved ability. Moreover, coefficients on both control functions are positive which confirms the conjecture that the unobservables are positively correlated across generations and justifies the interpretation of the assumed error

---

[8]In the paper results using the preferred specification are presented. Corresponding results with all variables entering the heteroskedastcity index can be obtained on request. The results are qualitatively unaffected by the choice of the form of heterskedasticity. However, some small quantitative differences are present.

structure. The magnitude of the effect of unobserved ability is similar to the one found by

Farré et al. (2012). To summarize, I find an important effect of parental education as well as an

Table 3: Relationship between parental and ethnic capital and children education

|  | 1977-2014 | | 1977-1989 | | 1990-2014 | |
|---|---|---|---|---|---|---|
|  | OLS | CF | OLS | CF | OLS | CF |
| Parental capital | 0.237 (0.006) | 0.177 (0.006) | 0.245 (0.009) | 0.215 (0.011) | 0.234 (0.007) | 0.170 (0.007) |
| Ethnic capital | 0.138 (0.015) | 0.070 (0.017) | 0.182 (0.026) | 0.068 (0.028) | 0.122 (0.019) | 0.075 (0.019) |
| $\rho^p$ | | 0.099 (0.009) | | 0.068 (0.020) | | 0.099 (0.010) |
| $\rho^{av}$ | | 0.031 (0.006) | | 0.085 (0.014) | | 0.015 (0.004) |

**Notes:** Bootstrapped standard errors in parenthesis

evidence that ethnic capital plays a role in intergenerational transmission beyond the transfer of

unobserved ability, even though not controlling for endogeneity results in a non trivial upward

bias on both parental and ethnic capital coefficients. The effect of the unobserved ability is

stronger in case of the parental education. Moreover, even though the OLS estimates in table 3

suggests that the role of ethnicity in intergenerational transmission of education has declined

over the years, the CF estimates indicate that it remained relatively constant. The effect of

parental capital decreased from 0.22 to 0.17. At the same time the role of unobserved ability in

the transfer of parental capital increased while its role in the transfer of ethnic capital decreased

significantly.

## 5   Conclusions

This paper focuses on the role of ethnic capital in the intergenerational transmission of human

capital. Its focus is on consistent estimates of the ethnicity effects. I find evidence that the

OLS estimates of the effect of ethnic capital on intergenerational transmission of education are

biased upwards. Unobserved ability has an important effect on educational choices and not

accounting for its confounding effect biases the estimates of parental and ethnic capital.

## References

Aydemir, A., Chen, W.-H., & Corak, M. (2013). Intergenerational education mobility among the children of Canadian immigrants. *Canadian Public Policy*, *39*(Supplement 1), S107–S122. doi: 10.3138/CPP.39.Supplement1.S107

Bauer, P., & Riphahn, R. T. (2007). Heterogeneity in the intergenerational transmission of educational attainment: evidence from Switzerland on natives and second-generation

immigrants. *Journal of Population Economics*, *20*(1), 121–148. doi: 10.1007/s00148-005-0056-5

Bisin, A., & Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, *97*(2), 298–319. doi: 10.1006/jeth.2000.2678

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *The American Economic Review*, 531–553.

Borjas, G. J. (1992). Ethnic capital and intergenerational mobility. *The Quarterly Journal of Economics*, 123–150. doi: 10.2307/2118325

Borjas, G. J. (1995). Ethnicity, neighborhoods, and human-capital externalities. *American Economic Review*, *85*(3), 365–90.

Borjas, G. J. (2006). Making it in America: Social mobility in the immigrant population. *The Future of Children*, *16*(2), 55.

Bourdieu, P. (2011). The forms of capital.(1986). *Cultural theory: An anthology*, 81–93.

Dion, K. K., & Dion, K. L. (2001). Gender and cultural adaptation in immigrant families. *Journal of Social Issues*, *57*(3), 511–521. doi: 10.1111/0022-4537.00226

Emran, S., & Sun, Y. (1988). Magical transition? Intergenerational educational and occupational mobility in rural China: 1988-2002. *Intergenerational Educational and Occupational Mobility in Rural China*, *2002*. doi: 10.1596/1813-9450-7316

Farré, L., Klein, R., & Vella, F. (2012). Does increasing parents schooling raise the schooling of the next generation? Evidence based on conditional second moments. *Oxford Bulletin of Economics and Statistics*, *74*(5), 676–690. doi: 10.1111/j.1468-0084.2011.00667.x

Farré, L., Klein, R., & Vella, F. (2013). A parametric control function approach to estimating the returns to schooling in the absence of exclusion restrictions: an application to the NLSY. *Empirical Economics*, *44*(1), 111–133. doi: 10.1007/s00181-010-0376-5

Holmlund, H., Lindahl, M., & Plug, E. (2011). The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of Economic Literature*, *49*(3), 615–651. doi: 10.1257/jel.49.3.615

Kao, G., & Thompson, J. S. (2003). Racial and ethnic stratification in educational achievement and attainment. *Annual Review of Sociology*, *29*(1), 417–442. doi: 10.1146/annurev.soc.29.010202.100019

Klein, R., & Vella, F. (2010). Estimating a class of triangular simultaneous equations models without exclusion restrictions. *Journal of Econometrics*, *154*(2), 154–164. doi: 10.1016/j.jeconom.2009.05.005

Nielsen, H. S., Rosholm, M., Smith, N., & Husted, L. (2003). The school-to-work transition of $2^{nd}$ generation immigrants in Denmark. *Journal of Population Economics*, *16*(4), 755–786. doi: 10.1007/s00148-003-0164-z

Patacchini, E., & Zenou, Y. (2011). Neighborhood effects and parental involvement in the intergenerational transmission of education. *Journal of Regional Science*, *51*(5), 987–1013. doi: 10.1111/j.1467-9787.2011.00722.x

Portes, A. (2000). The two meanings of social capital. In *Sociological forum* (Vol. 15, pp. 1–12). doi: 10.1023/A:1007537902813

Rosen, B. C. (1959). Race, ethnicity, and the achievement syndrome. *American Sociological Review*, 47–60. doi: 10.2307/2089582

Saniter, N., et al. (2012). Estimating heterogeneous returns to education in Germany via conditional second moments. In *Annual conference 2012 (Goettingen): New approaches and challenges for the labor market of the 21st century*.

Suárez-Orozco, C., & Qin, D. B. (2006). Gendered perspectives in psychology: Immigrant origin youth. *International Migration Review*, *40*(1), 165–198. doi: 10.1111/j.1747-7379.2006.00007.x