Gender heterogeneity and the role of ethnic capital in the intergenerational transmission of educational attainment

Agnieszka Postepska*

Abstract

This paper investigates how gender and ethnic capital interact in the intergenerational transmission of human capital. Relying on heteroskedasticity to identify parameters in the presence of endogenous regressors, I find evidence that while the transfer of ethnic capital benefits girls, its effect can be negative for boys. Moreover, introducing heterogeneity with respect to parental educational attainment reveals that while the effect of ethnicity is relatively constant along the distribution of parental schooling among girls, it varies significantly for boys. It could reflect differential investments in children schooling driven by substitution patterns between parental and ethnic capital applied to sons and daughters.

Keywords: intergenerational transmission of education, ethnic capital, gender heterogeneity, identification through conditional second moments

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

1 Introduction

Intergenerational transmission of human capital has attracted much attention in the literature with the primary focus on the link between parents and children schooling. An advantageous ethnic environment has been found to have a moderate in magnitude but positive effect on the human capital accumulation. However, not much attention has been paid to the intermediary role of gender in this transfer despite the evidence that ethnic socialization differs among boys and girls.

A major problem in this literature is identification due to unobserved ability. Some studies have focused on twins (Behrman & Rosenzweig, 2002), while others considered adoptees (Björklund, Lindahl, & Plug, 2006) or changes in educational systems (Black, Devereux, & Salvanes, 2005). While all these studies are important contributions, these identification strategies are difficult to apply in many settings and identification comes from selected samples of the population, such as twins, adoptees or individuals who respond to reforms. This paper follows identification strategy developed by Klein and Vella (2010) which relies on the heteroskedasticity of the error term and exploits the fact that the correlations of unobservables across generations is constant and invariant to the individuals' socioeconomic environment. Among others, this method has been applied to estimate the intergenerational transmission of education in the US (Farré, Klein, & Vella, 2013) and more recently to estimate the intergenerational transmission of ethnic capital (Postepska, 2019). This paper relies on the methodology laid out in Postepska (2019) to investigate whether gender plays an intermediary role in the transfer of ethnic capital and its effect on transmission of educational attainment between generations.

The ethnic hierarchy in educational attainment has been well documented in the literature (see Kao and Thompson (2003) for an excellent review). Ethnic socialization, a concept that describes maturing to ethnic identity, has also been recognized to vary significantly between boys and girls implying that ethnicity can have different impact on human capital accumulation among men and women. Since, in general, girls are more susceptible to social influences, they might be more likely to be isolated due to parental fear of the "bad" influence of the majority. Therefore, girls are prone to a much stricter control over their brothers (Sung, 1987; Olsen, 1997). This finding is consistent over time and across almost all ethnic groups (Dasgupta, 1998; Gupta, 1997; Williams, Alvarez, & Hauck, 2002; Yung, 1999; Sung, 1987). As a result, girls might be more likely to have contacts with peers of the same ethnicity than than boys. Moreover, such an increased supervision has proved to have a positive effect on schooling among Vietnamese girls (Zhou & Bankston III, 2001). Also, this could lead to a stronger importance of gender roles within ones ethnic group. High correlation between parental education implies that

high average education within ethnic group is directly related to a high average education among women within this ethnic group. Even though girls are also more likely to rebel against the traditions and values, they have been found to be more flexible in choosing ethnic identity and to built more complex ethnic identity by bridging home and host country identities (Rumbaut, 1997; Olsen, 1997). Girls have also higher educational and career aspirations, while boys tend to express more concern about social mobility (Suárez-Orozco & Qin, 2006). Also, boys are more pressured to take on their ethnic identity and are more likely to see the host country as hostile and unwelcoming (Suárez-Orozco & Qin, 2006). This might result in low self esteem and low aspirations and, therefore, boys might perceive a more limited set of opportunities in comparison to girls (Qin-Hilliard, 2003) regardless of the socioeconomic position of their ethnic group. This paper aims to investigate to what extent the transfer of ethnic capital varies by gender.

This study contributes to the literature by directly assessing the intermediary role of gender in the transfer of ethnic human capital within the framework developed in Borjas (1992) which to my knowledge has not been studied before. It is a first step toward gaining insight into the channels through which the transfer of ethnic capital affects schooling outcomes. The results show that the transfer of ethnic capital is more likely to affect women and that for men the effect varies with parental schooling. Among others, this might reflect different socialization patterns and imply that the environment play a much bigger role for educational outcomes of the disadvantaged youth than for their more advantaged peers contributing to the widening of the income gap.

The paper is organized as follows. The next section briefly explains the estimation method and identification. Section 3 describes the data and section 4 follows with empirical results and discussion. Section 5 concludes.

2 Model and Identification

The identification and estimation strategy follows directly Postepska (2019) and I direct the reader to this manuscript for details while below I present the key features of the model and identification. Let *edu* denote the individual's education, *edup* parental education[1] and *eduav* the average education of the ethnic group measured as the average education in the parents generation[2]. To account for heterogeneous effects among boys and girls in the base model ethnic capital is interacted with a female dummy (*fem*) and the basic specification of the main equation

[1]Parental education is measured with father's schooling

[2]Average education in parents generation is computed as average education among the fathers within a cohort.

of interest is as follows:

$$edu_i = \gamma_1 edup_i + \gamma_2 eduav_i + \gamma_3 eduav_i \times fem_i + \delta_0 X_i + u_i \quad (1)$$

where X denotes a vector of socioeconomic characteristics. As parents make their investment decision about their children education within their ethnic environment, ethnic capital might affect children differently along the distribution of parental schooling. Equation 2 extends equation 1 to allow for this interaction:

$$edu_i = \gamma_1 edup_i + \gamma_2 eduav_i + \gamma_3 eduav_i \times fem_i + \gamma_4 edup_i \times eduav_i + \delta_0 X_i + u_i \quad (2)$$

To further allow for differential effect of the intermediary role of parental capital in transmission of ethnic capital by gender, parental education is interacted with ethnic capital and gender so that the model becomes:

$$\begin{aligned} edu_i = {} & \gamma_1 edup_i + \gamma_2 eduav_i + \gamma_3 edup_i \times fem_i + \gamma_4 eduav_i \times fem_i + \gamma_5 edup_i \times eduav_i \\ & + \gamma_6 edup_i \times eduav_i \times fem_i + \delta_0 X_i + u_i \end{aligned} \quad (3)$$

The specification in equation 3 aims to capture the possibility that the role of ethnicity in the transfer of educational attainment can affect girls and boys differently depending on their socioeconomic standings.

We assume that all variables in X are exogenous and there are no instruments available for the two endogenous regressors: parental and ethnic capital. In absence of exclusion restrictions, identification of the parameters relies on assumptions about the structure of the error term and heteroskedasticity in the model. The two endogenous equations are of the following form:

$$\begin{aligned} edu f_i &= \delta_2 X_i + v_i^p \\ eduav_i &= \delta_3 X_i + v_i^{av} \end{aligned} \quad (4)$$

Two conditions are necessary for identification:

(1) Presence of heteroskedasticity in the error term of the main equation and/or in the error terms of the auxiliary equations for parental and ethnic capital. That means that we can write the error terms of the main equation and of the two endogenous equations as:

$$u_i = H_u(X_i) u_i^*, \quad v_i^p = H_v^p(X_i) v_i^{p*} \text{ and } v_i^{av} = H_v^{av}(X_i) v_i^{av*} \quad (5)$$

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.

(2) Constant correlation:

$$E[u_i^* v_i^{p*}|X_i] = E[u_i^* v_i^{p*}] = \rho^p \text{ and } E[u_i^* v_i^{av*}|X_i] = E[u_i^* v_i^{av*}] = \rho^{av} \tag{6}$$

This assumption implies that there exists a constant transfer that is not affected by any socioeconomic characteristics or behavior of the individual or the parents.

As argued in details in Postepska (2019), both assumptions appear plausible 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. The heteroskedasticity implies that the contribution of the unobservables to the formation of educational attainment differ depending on characteristics. If unobserved ability is transferred genetically, than this assumption is clearly satisfied. This approach was successfully applied in (Farré, Klein, & Vella, 2012; Farré et al., 2013). Since systematic differences in the "quality" of cohorts of immigrants have been found (Borjas, 2006, 1987), 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). In addition to peer effects, Heteroskedacticty in the child's schooling equation is granted by the presence of peer effects as well as by the fact that parents investments in child's education depend on the ethnic environment (Bisin & Verdier, 2001; Patacchini & Zenou, 2011). Moreover, 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 father 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.

This assumed error structure allows construction of control functions which inclusion in equation 1, equation 2 and equation 3 makes estimation of the unknown parameters γ =

$\{\gamma_1, \gamma_2\, \gamma_3, \gamma_4, \gamma_5, \gamma_6\}$ feasible:

$$edu_i = \gamma_1 edup_i + \gamma_2 eduav_i + \gamma_3 eduav_i \times fem_i + \delta_0 X_i + \rho^p \frac{H_u(X_i)}{H_v^p(X_i)} v_i^p + \rho^{av} \frac{H_u(X_i)}{H_v^{av}(X_i)} v_i^{av} + \epsilon_i \quad (7)$$

$$edu_i = \gamma_1 edup_i + \gamma_2 eduav_i + \gamma_3 eduav_i \times fem_i + \gamma_4 eduav_i \times edup_i + \delta_0 X_i + \rho^p \frac{H_u(X_i)}{H_v^p(X_i)} v_i^p \quad (8)$$
$$+\rho^{av} \frac{H_u(X_i)}{H_v^{av}(X_i)} v_i^{av} + \epsilon_i$$

$$edu_i = \gamma_1 edup_i + \gamma_2 eduav_i + \gamma_3 edup_i \times fem_i + \gamma_4 eduav_i \times fem_i + \gamma_5 edup_i \times eduav_i \quad (9)$$
$$+\gamma_6 edup_i \times eduav_i \times fem_i + \delta_0 X_i + \rho^p \frac{H_u(X_i)}{H_v^p(X_i)} v_i^p + \rho^{av} \frac{H_u(X_i)}{H_v^{av}(X_i)} v_i^{av} + \epsilon_i$$

3 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. Also, only individuals who grew up with both parents are included. Individuals for whom information about their own or their parents education attainment is not available are omitted from the sample. Individuals in the sample were born between 1913 and 1992 and they are divided into 5 cohorts[3]. Also, only individuals for whom there is at least 30 other individuals in the same cohort of the same ethnic origin are included[4]. Parental human capital is measured with father's education and ethnic capital as the average education in the father's generation[5].

The final sample contains individuals coming from 26 different origins. Table 1 presents the breakdown by country or region of origin in the whole sample. Descendants of German, English, Welsh and Irish immigrants are most represented in the sample, while other origins individually constitute a small share of the total sample.

Table 2 and Table 3 present the distribution of gender across cohorts and regions pointing to some important differences. Women are over represented in the oldest cohorts (70% in the oldest cohort), which is consistent with longer life expectancy among women. There is also

[3]Data on parents age is not available so I use year of birth to categorize into cohorts. Finer division is not possible due to small cell sizes.

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

[5]Due to the high correlation between father's and mother's education and since data on father's education was available for more observations, I only include father's education in the estimation. Moreover, (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.

Country of ancestry	Men	Women	Country of ancestry	Men	Women
Africa	0.08	0.15	Netherlands	1.35	1.06
Austria	0.09	0.09	Norway	2.29	2.03
French Canada	1.04	1.05	Philippines	0.33	0.64
Other Canada	0.03	0.04	Poland	3.51	3.30
China	0.62	0.62	Puerto Rico	0.64	0.71
Czechoslovakia	0.69	0.94	Russia	0.79	0.89
Denmark	0.05	0.12	Scotland	4.47	3.54
England & Wales	20.03	18.93	Mexico	5.74	5.58
Finland	0.18	0.10	Spain	0.46	0.82
France	1.38	1.77	Sweden	1.45	1.62
Germany	26.04	25.73	India	1.06	0.49
Hungary	0.24	0.45	Portugal	0.04	0.13
Ireland	16.77	18.95	Arabic	0.21	0.09
Italy	7.36	7.21	Other Spanish	1.23	1.16
Japan	0.13	0.11	Non-span Windies	0.03	0.03
American only	0.79	0.80	Other Asian	0.36	0.41

Table 1: Sample composition

some gender imbalance across regions, with New England and Mountain having slightly higher shares of female than other regions noting that this is not a direct effect of age distribution.

Share of females	
Born 1889-1909	0.70
Born 1910-1929	0.58
Born 1930-1949	0.54
Born 1950-1969	0.52
Born 1970-1999	0.52

Table 2: Gender distribution by cohort

Share of females	
New England	0.57
Middle Atlantic	0.55
East North Central	0.53
West North Central	0.55
South Atlantic	0.54
East South Central	0.55
West South Central	0.53
Mountain	0.56
Pacific	0.52

Table 3: Gender distribution by region

Table 4 presents the summary statistics for all variables used in this analysis separately for men and women, noting that 54 percent of the sample are women. The only difference between men and women appears along the age dimension with women being on average two years older than men. The average individual has about 3 siblings and has completed almost 14 years of

schooling, which is surprisingly high. The average parental and ethnic capital are approximately the same at 11 years of schooling. 41 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.

	Men	Women
Age	44.69	46.27
	(16.00)	(16.74)
Number of siblings	3.24	3.32
	(2.39)	(2.43)
Years of schooling	13.88	13.64
	(3.04)	(2.82)
Parental capital	11.20	11.07
	(4.25)	(4.20)
Ethnic capital	11.11	10.95
	(2.19)	(2.15)
Living in a city at 16	0.42	0.42
	(0.49)	(0.49)
Living in a Southern state at 16	0.23	0.23
	(0.42)	(0.42)
At least one parent born abroad	0.17	0.18
	(0.38)	(0.38)
Number of observations	7804	9127

Notes: Standard deviations in brackets.

Table 4: Summary statistics of the main variables

Table 5 presents the key variables, self, parental and average ethnic educational attainment by region of origin separately for men and women. There is significant variation in father's education and ethnic capital across different origins. Individuals of Chinese, Indian, Russian and Arabic decent and their fathers have the highest educational attainment throughout the years. Individuals of Mexican origin have the lowest educational attainment and the least advantageous parental and ethnic environment with respect to schooling. It is worth noticing though that this gap has decreased over years. Nevertheless, their fathers are still ranked last and so is the overall ethnic capital across all cohorts. Overall, the gap in the average individual educational attainment decreased from about 8 years for the cohort born between 1910-1929 to less than 3 years among individuals born between 1970 and 1999. This decrease is partially driven by the large increase (about 6 years) in average schooling among individuals of Mexican origin. Also the gap in the average schooling among ethnic groups has shrunk. The variation between men and women within ethnic origin is much smaller. While on average men have more years of schooling than women, this difference usually does not exceed one year of schooling. As expected, even smaller differences are observed along parental capital and almost no difference

is present in ethnic environments. However, some notable exception are present. While women of Danish decent have on average 2.5 years of schooling less than men, women of Portuguese and Spanish decent tend to obtain more schooling than men of the same origin. Surprisingly, Danish women, similarly to Portuguese women, appear to grow up in a more advantageous ethnic environment than men do.

	Men			Women		
Country of ancestry	Self	Father	Ethnic capital	Self	Father	Ethnic capital
Africa	13	10.8	12.2	13.1	11.5	11.5
Austria	13.4	6	6.5	13.3	7.25	6.5
French Canada	13.2	9.35	10	12.9	10.6	9.96
Other Canada	16.5	10	9.07	11.5	7	9.07
China	16.6	13	13.1	15.5	13.1	13.1
Czechoslovakia	14.1	10.6	9.94	13.2	9.36	9.76
Denmark	15.8	11.5	11.8	13.3	12.3	11.8
England & Wales	14.5	11.9	11.8	14.1	11.8	11.6
Finland	12.8	8	7.77	11.3	7.22	7.69
France	13.8	11.8	11.6	13.8	11.7	11.6
Germany	13.8	11.5	11.3	13.5	11.2	11.2
Hungary	13.9	11.4	11.3	14.4	12	11.2
Ireland	13.9	11.6	11.5	13.7	11.5	11.3
Italy	14.1	10.8	10.8	13.7	10.8	10.6
Japan	14.2	13.1	12.7	15.5	12.8	12.7
Mexico	11.3	6.46	6.57	11.3	6.93	6.89
Netherlands	13.3	10.8	10.6	13	10.4	10.4
Norway	13.4	11.3	11.4	13.7	10.8	10.7
Philippines	15.8	11.9	11.8	14.9	11.5	12
Poland	14	10.8	10.6	13.6	10.8	10.6
Puerto Rico	12	8.8	8.94	12.8	9.55	8.96
Russia	15.7	12.3	12	15.4	12.2	12.3
Scotland	14.5	12.1	12	14.4	12	11.9
Spain	12.4	9.97	9.99	13.7	9.91	10.1
Sweden	14.1	12.2	11.7	14.3	11.5	11.6
Other	15	11	11.4	14.2	11.9	11.4
India	17	14.2	14.1	16.2	14.2	14.4
Portugal	12	10	11.9	15	12.8	11.9
Arabic	16.1	13.9	13.8	15.9	13.5	13.8
Other Spanish	12	8.92	9.03	11.6	8.75	8.85
Non-span Windies	14	14	11.5	15	12	11.1
Other Asian	15.1	12	12.1	15.4	12	12.1
American only	12.1	9.5	9.06	11.4	8.89	8.59

Table 5: Self, parental and ethnic capital by ancestry

4 Empirical strategy and results

I follow closely Farré et al. (2012); Postepska (2019) in the estimation strategy[6]. 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. I start with a guess of coefficients for the main equation (OLS estimates). Then, given these coefficients I compute the residuals and estimate the heteroskedastic index of the main equation. Given these estimates, I improve the guess of coefficients by including this correction term into the equation and estimating it by OLS to get new set of coefficients. This process continues until the coefficients values converge.

Before proceeding to the discussion of results of the main equation, let us consider the results of the estimation of the heteroskedasticity indexes in the three equations. Note that only the heteroskedastcity index of the main equation differ between the base and extended model. Moreover, the sets of variables included in the parental and ethnic capital equations are almost the same, so I discuss them together. Since I do not have information about the age of the parents, I include the age of the children (and age squared) in both equations. This, together with the dummy variable indicating the cross section, controls for the age of the fathers. Dummy variables for regions control for geographic differences in educational attainment that might result from labor market specific needs of given region or different access to educational institutions. I also include a dummy variable indicating whether the child was living in the South or in the city at the age of 16. Unfortunately, this information is not available for the parents so I use the information for the children as proxies. In the ethnic capital equation I also include a dummy indicating whether at least one of the parents is foreign born and in the father's schooling equation a dummy variable indicating whether the father is foreign born.

The OLS estimates of the conditional means of the parental and ethnic capital equations presented in Table 6 are in line with the literature. All the year dummies (with the exception of 1978) are significant and indicate an increasing trend in educational attainment among parents. Younger individuals and those with fewer siblings have not only better educated father's but also more favorable ethnic environment. Individuals living in the city at 16 have, on average, better educated parents than their counterparts residing outside of the cities. The correlation with the ethnic capital is also positive but it's of much smaller magnitude and appear to be

[6]Details of estimation are explained in Appendix A.

insignificant. Residence in southern states lowers the average educational attainment of the fathers as well as the average ethnic capital individuals are exposed to by almost one year. Fathers born outside of the US have on average 2.7 years less of schooling than US born fathers. Also, average ethnic capital decreases for individuals with at least one parent born abroad by 1.4 years on average. This might reflect the fact that more recent immigrant groups are on average less educated than members of established groups within the US. There is also some evidence of regional differences for both parental and ethnic capital.

Lower panel of Table 6 presents the test statistics for White and Breush-Pagan tests for heteroskedasticity in both equations. The null hypothesis of homoskedastic errors is strongly rejected in both equations confirming presence of heteroskedasticity in parental and ethnic capital equations necessary for identification.

	Parental capital	Ethnic capital
Age	-0.089	-0.061
	(0.009)	(0.004)
Female	0.011	-0.031
	(0.057)	(0.021)
Living in a city at the age of 16	1.171	0.009
	(0.059)	(0.022)
Living in the South at the age of 16	-0.828	-0.703
	(0.099)	(0.037)
Number of siblings	-0.303	-0.084
	(0.0125)	(0.005)
Father born abroad	-2.706	
	(0.108)	
At least one parent born abroad		-1.406
		(0.036)
Constant	12.199	13.979
	(0.316)	(0.112)
Breush-Pagan test	544.42	5030.34
White test	1095.95	3126.06
Number of observations	15390	

Notes: Robust standard errors in brackets. All regressions also include age squared, dummy variables for region of residence and year dummies for cross section.

Table 6: Parental and ethnic capital - conditional means

Having established the presence of heteroskedasticity we can continue with estimation of its form and further construct the two control functions[7]. Results of the non linear least squares estimation of the conditional variances are presented in Table 7. Given the assumed exponential form of heteroskedasticity, I can directly interpret the coefficients. Older individuals are exposed

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

to a smaller variation in average education among immigrants from the same origin as well as their fathers have smaller residual variance. This could result from increasing heterogeneity of immigrants coming from the same origin as well as easier access to education. Moreover, I find higher dispersion in fathers' education for individuals who lived in the city or in the South at the age of 16. Similarly, fathers born abroad and with more children have a higher variance in educational attainment. I also find bigger dispersion in ethnic capital for individuals with at least one parent born abroad.

	Parental capital	Ethnic capital
Age	-0.001	-0.399
	(0.006)	(0.021)
Living in a city at 16	0.245	
	(0.034)	
Living in the South at 16	0.271	
	(0.030)	
Number of siblings	0.036	
	(0.006)	
One parent born abroad		0.471
		(0.111)
Father born abroad	0.414	
	(0.048)	
Year dummies	Yes	Yes
Regional dummies	No	Yes
Constant	0.600	7.466
	(0.172)	(0.350)

Notes: Standard errors bootstrapped. Results are the same for the basic and the extended model as this step does not involve interaction terms.

Table 7: Heteroskedastic indexes for parental, ethnic and children capital

To construct the correction terms we still need the estimates of the heteroskesdticity index in the primary equations (equation 1,equation 2, and equation 3). It is estimated simultaneously with the coefficients of the same equation. The results are presented in Table 8 and are very similar across the two model specifications. It appears that women, younger individuals and individuals with more siblings have a smaller residual variance.

Let us now turn to the discussion of the results of the main equations of interests (equation 1, equation 2 and equation 3).The OLS estimates presented in Table 9 show that the effect of ethnic capital is about a third of the size of the effect of parental capital on educational attainment. Increasing the average educational attainment within an ethnic group by 1 year results in 0.088 extra years of schooling. However, the CF estimates of equation 1 reveal that once the unobservables are controlled for, the transfer of ethnic capital matters only for females confirming the direction of the effects suggested above. Increasing the average years of schooling

	Base model	Extended model
Age	0.109	0.115
	(0.017)	(0.017)
Age^2	-0.001	-0.001
	(0.000)	(0.000)
Female	-0.497	-0.673
	(0.080)	(0.087)
# siblings	-0.050	-0.51
	(0.018)	(0.021)
Year dummies	No	No
Regional dummies	No	No
Constant	-3.388	-3.602
	(0.413)	(0.350)

Notes: Standard errors bootstrapped.

Table 8: Heteroskedastic indexes for children capital

in the parents generation by one year increases the educational attainment of women by 0.112 years. While this effect does not appear to be very large in magnitude it is comparable to the effect of parental educational achievement. What is more striking though is that the quality of ethnic environment does not appear to matter for educational attainment among men.

However, as parents make their investment decision within an ethnic environment ethnic capital may affect individuals differently along the distribution of parental capital. In other words, individuals with highly educated parents may be less susceptible to the environmental influence due to parental investments. This is confirmed in column three of Table 9 which corresponds to equation 2. When the effect of ethnic capital is allowed to vary with parental capital, every additional year of parental schooling decreases the influence of additional year of average schooling of the ethnic group by 0.09. This means that the effect on an individual with college educated father is 0.05 smaller than on an individual with a father who dropped out of high school. Allowing for heterogenous effects of ethnic capital with respect to parental schooling revealed a different pattern for ethnic capital transmission among girls and boys. Holding parental education and investment environment constant men also appear to be affected by ethnic capital, however the effect on women remains more than twice as large (0.08 versus 0.194).

Moreover, as it appears that girls are more affected by the quality of the ethnic environment, the differential effect of ethnic capital depending on the student background might also differ with gender. Indeed, when we take this differential effect into account, gender differences prevail as documented in the last column of Table 9. Table 10 shows marginal effects of an extra year of schooling within ethnic group on the individual educational attainment by skill level of the parent. Unskilled father corresponds to at most high school completion of the father, while

skilled father implies at least some college attendance. It appears that while the effect on girls is relatively constant and does not vary much with parental education, for boys parental education plays significant role in determining how much ethnic environment matters for educational attainment. However, the fact that boys seems to be negatively affected, especially in the presence of skilled father, is somewhat puzzling and requires further investigation. It could be the case that parents rely too much on ethnic environment and underinvest in their sons schooling in more favorable ethnic environments but as boys benefit less from advantageous environment, they end up worse off than boys that grow up in less favorable ethnic environments in which parents invest more. Since girls benefit more from their ethnic capital, this can make up for differences in parental investments across ethnic environments. A more detailed data is required to verify this conjecture.

The results above confirm that girls are much more dependent on the ethnic environment. It can result from the different socialization styles discussed above. Realizing this differential effect is of importance to policy makers as it suggests that policies addressing girls and boys of ethnic minorities should be structured differently to alleviate the disadvantaged ethnic background among girls or help boys take advantaged of the available ethnic capital within their group.

5 Conclusions

This paper investigates how gender and ethnic capital interact in the intergenerational transmission of human capital. Relying on identification through second moments, I deliver new evidence on the gender heterogeneity with respect to the role of ethnic capital in the intergenerational transmission of educational attainment. I find that the transfer of ethnic capital benefits mostly women. I also find that the positive effect of ethnicity is relatively constant along the whole distribution of parental educational attainment among women.Among men, it is not only negative but also vary significantly along parental educational attainment distribution.

	OLS Eq 1	CF Eq 1	CF Eq 2	CF Eq 4
Parental capital	0.237	0.173	0.269	0.361
	(0.006)	(0.005)	(0.027)	(0.045)
Ethnic capital	0.088	-0.003	0.080	0.161
	(0.018)	(0.025)	(0.032)	(0.049)
Female	-1.201	-1.413	-1.449	0.120
	(0.221)	(0.252)	(0.232)	(0.647)
Female × Ethnic capital	0.094	0.112	0.114	-0.036
	(0.019)	(0.021)	(0.020)	(0.061)
Female × Parental capital				-0.165
				(0.059)
Parental capital × Ethnic capital			-0.009	-0.018
			(0.002)	(0.004)
Parental capital × Ethnic capital × Female				0.015
ρ_p		0.108	0.116	0.114
		(0.010)	(0.010)	(0.011)
ρ_{av}		0.037	0.024	0.024
		(0.004)	(0.004)	(0.003)
Year dummies	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes
Constant	7.595	7.611	7.510	6.673
	(0.313)	(0.288)	(0.401)	(0.602)

Notes: Standard errors bootstrapped. OLS results are presented only for the first model for comparison. Other controls include age, a dummy indicating whether the person lived in the city and in the South at the age of 16, number of siblings and a dummy variable indicating whether at least on parent is foreign born.

Table 9: Relationship between parental and ethnic capital and children education

	Women	Men
Skilled father	0.114	-0.096
Low skilled father	0.130	-0.02

Notes: Low skilled corresponds to 11 years of schooling while high skilled to 14 years of education.

Table 10: Marginal effects of ethnic capital by parental skill level

References

Behrman, J. R., & Rosenzweig, M. R. (2002). Does increasing women's schooling raise the schooling of the next generation? *American Economic Review*, *92*(1), 323–334.

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

Björklund, A., Lindahl, M., & Plug, E. (2006). The origins of intergenerational associations: Lessons from swedish adoption data. *The Quarterly Journal of Economics*, *121*(3), 999–1028.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *American economic review*, *95*(1), 437–449.

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. (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.

Dasgupta, S. D. (1998). Gender roles and cultural continuity in the Asian Indian immigrant community in the US. *Sex roles*, *38*(11-12), 953–974.

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

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

Gupta, M. D. (1997). What is Indian about you? A gendered, transnational approach to ethnicity. *Gender & Society*, *11*(5), 572–596.

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

Olsen, L. (1997). *Made in America: Immigrant students in our public schools.* New York: New Press.

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

Postepska, A. (2019). Ethnic capital and intergenerational transmission of educational attainment. *Forthcoming in Journal of Applied Econometrics*.

Qin-Hilliard, D. B. (2003). Gendered expectations and gendered experiences: Immigrant students' adaptation in schools. *New directions for youth development*, *2003*(100), 91–109.

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

Rumbaut, R. G. (1997). Ties that bind: Immigration and immigrant families. *Immigration and the family: Research and policy on US immigrants*, 3–46.

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

Sung, B. L. (1987). *The adjustment experience of chinese immigrant children in New York City.* Staten Island, NY: Center for Migration Studies.

Williams, L. S., Alvarez, S. D., & Hauck, K. S. A. (2002). My name is not Maria: Young Latinas seeking home in the heartland. *Social Problems*, *49*(4), 563–584.

Yung, J. (1999). *Unbound voices: A documentary history of Chinese women in San Francisco.* Univ of California Press.

Zhou, M., & Bankston III, C. L. (2001). Family pressure and the educational experience of the daughters of Vietnamese refugees. *International Migration*, *39*(4), 133–151.

Appendix A

6 Estimation procedure

This section outlines the two step procedure to estimate the model. First, regress $edup_i$ on X_i and $eduav_i$ on X_i and obtain α^p and α^{av}. Then define the residuals from these two regressions as follows:

$$\begin{aligned} \hat{v}_i^p &= edup_i - X_i\hat{\alpha}^p \\ \hat{v}_i^{av} &= eduav_i - X_i\hat{\alpha}^{av} \end{aligned}$$

The conditional variances of the parental education and average education errors can be estimated using both parametric and non-parametric methods. In this paper I employ parametric approach and assume the following functional form of the heteroskedasticity:

$$\begin{aligned} H_{vi}^{p2} &= \exp(Z_i\theta^p) \\ H_{vi}^{av2} &= \exp(Z_i\theta^{av}) \end{aligned}$$

where Z_i is a vector of variables responsible for the heteroskedasticty of the errors. The set of variables Z_i can vary between the two equations. Note that there are no restrictions imposed over the relationship between Z_i and X_i, i.e. model is identified even if $Z_i = X_i$. If, however, there are variables that appear in Z_i but not in X_i, they do not help identify the model in a standard way. Since it is the movement in the variances that grants identification in the model, variables in Z_i aid identification only if they can explain the differences in the variance across observations.

The conditional variances are estimated using non linear least squares using $\ln(\hat{v_i}^p)$ and $\ln(\hat{v_i}^{av})$ as dependent variables. Then we can compute the standard deviation of the error terms associated with the two reduced forms: $\hat{H}_{vi}^p = \sqrt{\exp(Z_i\hat{\theta}^p)}$ and $\hat{H}_{vi}^{av} = \sqrt{\exp(Z_i\hat{\theta}^{av})}$.

Last element necessary to estimate the parameters of the main equation is the standard deviation of the child's education error (so the error term of the main equation). Since consistent residuals are nor readily available, it is estimated simultaneously with the parameters of the main equation in an iterative procedure. Let $\beta = \{\gamma_1, \gamma_2, \delta_0, \theta_u\}$. The parameters are found using a non linear least squares:

$$\min_{\beta,\rho^p,\rho^{av}} \sum_{i=1}^{n} \left(edu_i - \gamma_1 edup_i - \gamma_2 eduav_i - \delta_0 X_i - \rho^p \frac{H_{ui}}{\hat{H}_{vi}^p}\hat{v}_i^p - \rho^{av}\frac{H_{ui}}{\hat{H}_{vi}^{av}}\hat{v}_i^{av} \right)^2$$

where H_{ui}^2 denotes the conditional variance of the child's education equation. In a fully parametric specification, assume $H_{ui}^2 = exp(Z_i\theta_u)$.

To simplify the computations, Klein and Vella (2010) suggest a two step procedure. First, for a given value of $\beta = \tilde{\beta}$, define the residuals $u_i(\tilde{\beta})$ and compute the standard deviation of the child's education error in the same way as $\hat{H}_{vi}^p$ and $\hat{H}_{vi}^{av}$, so $\hat{H}_{ui} = \sqrt{\exp(Z_i\tilde{\theta}_u)}$. Second, estimate ρ^p and ρ^{av} by minimizing the sum of the squared residuals of the child's education equation:

$$\min_{\rho^p,\rho^{av}} \sum_{i=1}^{n} \left(u_i(\tilde{\beta}) - \rho^p \frac{\hat{H}_{ui}(\tilde{\beta})}{\hat{H}_{vi}^p} \hat{v}_i^p - \rho^{av} \frac{\hat{H}_{ui}(\tilde{\beta})}{\hat{H}_{vi}^{av}} \hat{v}_i^{av} \right)^2 \tag{10}$$

Repeat the last two steps until the minimum of (10) is found.