The long-run effects of streamed primary education on socio-political participation

PRELIMINARY DRAFT

PLEASE, DO NOT QUOTE!

Francesco Bartolucci
Università di Perugia
francesco.bartolucci@unipg.it

Donata Favaro
Università di Padova
donata.favaro@unipd.it

Dario Sciulli
Università di Chieti-Pescara
dario.sciulli@unich.it

Marcello Signorelli
Università di Perugia
marcello.signorelli@unipg.it

Abstract

Until '60s British primary school widely applied streaming, i.e. the practice of assigning pupils to classes on the basis of an overall assessment of their general ability. This paper investigates the lasting impact of attending streamed primary school on socio-political participation during adulthood. By using NCDS data and 2SLS models to account for endogeneity, first, we find that cohort-members enrolled in primary school adopting streaming, would be less likely involved in socio-political activities during adulthood. Second, by focusing on sub-sample of cohort-members enrolled in streamed primary school, we find that those assigned to high-ability classes would be more likely involved in adult socio-political activities than those assigned to low-ability classes. In sum, our findings suggest that school organization would be relevant when shaping social behavior and social capital formation.

INTRODUCTION

It is a matter of fact that "social capital" may have several socioeconomic effects, including growth effects. Trust, social norms and civic engagement have been recognized as relevant factors for economic success and growth. Zak and Knack (2001) show that low trust environments reduce the rate of investment and growth. On the other hand, the role of associational activity in expanding and enhancing economic performance is more controversial. Associational life can "instill in their members habits of cooperation, solidarity, and public-spiritedness", as underlined by Putnam (1993). However, the goals of a group may conflict with those of other groups, and the overall effect of group memberships on economic activities could be negative (Knack and Keefer, 1997). Whatever the sign of the effect, most of the literature finds that social capital influences economic performance (see Sobel, 2002 for a critical analysis of the literature). Therefore, it is important to identify the determinants of social capital and how it is formed.

The literature on the determinants of social capital shows that social capital is the result of the interaction of the degree of socio-economic heterogeneity and the degree of people involvement in formal and informal groups and associations. When societies are polarized by ethnic, political, religious, or income differences, associations in such societies often form along (e.g.) ethnic lines (relatively homogeneous associations in heterogeneous societies). This may strengthen trust and cooperative norms within an ethnic group, but weaken trust and cooperation between those groups. Polarization does not help individuals sharing mutual expectations and common backgrounds, so it is more difficult to make self-enforcing agreements (Knack and Keefer, 1997). Moreover, participation in social activities is significantly lower in more unequal and in more racially or ethnically fragmented localities (Alesina and La Ferrara, 2000).

Following this literature, in this article we try to investigate the role played by school in building individual social attitude in heterogeneous societies. Specifically, we analyse how school organization can shape social behavior of individuals when adult, and affect the building of social capital. Why should school organization affect the construction of social capital and influence socio-political participation and the associational life of people when adult? What we mean by school organization is related to the way in which pupils are grouped in classes according to their abilities. Specifically, we focus on the method of school organization that became the dominant form of ability grouping in Britain after the Second World War: the method of streaming. Streaming is a type of homogeneous grouping that assigns pupils to classes on the basis of an overall assessment of their general ability. Pupils remain in their streamed class for the majority of subjects. The use of streaming was meant to reduce the heterogeneity of the learning group and the diversity of pupils' needs in the class.

The idea of this article is that the method of grouping within schools may create patterns of interactions among pupils that may persist during their lives and influence their civic behavior. Then, attendance of streamed schools, with ability grouping, may reinforce polarization, and strengthen trust and cooperative norms within homogeneous groups, but weaken trust and cooperation between groups.

We investigate this issue by analyzing the impact of attending streamed primary schools on adult socio-political participation. The National Child Development Study (NCDS)[1] is an excellent source of data for this analysis, as it provides information on the type of primary school attended by cohort-members (streamed or not streamed) and, in case of streaming, on the ability class (high, average and low) to which the individual has been assigned. In addition, NCDS provides information on socio-political participation of cohort-members during adulthood. In this, respect, we use information from the 5^{th}, 6^{th} and 8^{th} sweeps of NCDS, when individuals were aged 33, 42 and 51, respectively. Socio-political participation is defined according to the participation/registration to political parties, environmental, charity and voluntary associations, women organizations, school and parents associations, resident associations and vote to the most recent general elections. Binary responses provided by every individual at each time occasion for each participation category are collapsed in standardized quantitative scores by using the Rasch model (Rasch, 1960, 1961).

Our empirical analysis is based on a two-step approach. In the first step, we measure the impact of attending primary streamed schools (versus not streamed) on the adult socio-political participation propensity, as measured by the standardized scores mentioned above. In the second step, we restrict our analysis to the sub-sample of cohort-members attending primary streamed schools, and estimate the impact on adult socio-political participation of being assigned to a high-ability class versus an average-low ability class, and vice-versa. In both cases, because enrolment in streamed schools and allocation to ability-classes would be endogenous, the adult socio-political participation equations are estimated by using 2SLS models.

Our results suggest that attending primary streamed schools negatively affects adult socio-political participation. In addition, when focusing on streamed cohort-members, we find that adult socio-political participation is positively affected by being assigned to high-ability classes, and negatively affected by being assigned to low-ability classes.

Section 2 presents the data and the sample of analysis, Section 3 describes the econometric approach, Section 4 reports the results, and Section 5 concludes.

[1] NCDS collects information on individuals born from March 3-9, 1958, in England, Wales and Scotland, selecting data from the Perinatal Mortality Survey. In addition to information from 1958, NCDS sweeps were performed in 1965, 1969, 1974, 1981, 1991, 1999-2000, 2004-2005 and 2008-2009.

2. THE NCDS DATA

2.1 The data and definition of variables

The study is carried out employing information from the National Child Development Studies (NCDS). The NCDS is a cohort study following all UK births during the week of 3-9 March 1958. The main aim of the study is to improve the understanding of the factors affecting human development over the entire lifespan. The NCDS has its origin in the Perinatal Mortality Survey (PMS) that collected information on a cohort of approximately 17,000 children. Successively, the PMS became the NCDS that has gathered information on the same individuals at different times in their lives (1965, 1969, 1974, 1981, 1991, 1999-2000, 2004-2005 and 2008-2009). The available data have been reduced considerably since 1991, consisting of approximately 9800 observations in the latest sweeps[2].

We use five sweeps of the NCDS database. The 1969 sweep, when cohort-members were 11 years old, includes retrospective information allowing us to identify if an individual attended or not a primary school applying streaming[3] (variable n861). Another variable (n862), allows us to distinguish cohort-members, attending a streamed school, accordingly to the allocation by ability class (higher, average or lower ability). The 1965 sweep contains information when cohort-members were aged 7. Some of these variables are possibly useful to identify mechanism leading individuals to attend schools applying streaming and, in case, to be allocated in different ability classes. Particularly, because both decisions are possibly endogenous, some variables drawn from the 1965 NCDS sweep are used as instruments in the econometric approach.

The adoption of streamed classes is based on principal's decisions. However, familiar background and parents' views potentially guide the enrollment of children in schools adopting or not streamed classes. This represents a possible source of endogeneity as familiar background and parents' views may affect also cohort-members' views and, then, social participation decisions in the long-run. In this respect, we identify two variables potentially being used as excluded instruments, i.e. a dummy variable indicating if educational meetings are arranged for parents/teachers associations at school and a dummy variable indicating if social functions are arranged for parents at school. In particular, the hypothesis is that schools allowing for educational meetings and social functions are less likely to provide for streamed classes.

---

[2] The selection and the attrition bias problems in the NCDS data have been investigated in several papers. Among others, Dearden et al. (1997) show that attrition in the NCDS has tended to weed out individuals with lower ability and lower educational qualifications, while Hawkes and Plewis (2006) have found that the attrition and non-response issues can be associated with only a few significant predictors.

[3] Streaming indicates the practice of placing students in classes with other with comparable skills or needs (see the 1967 Plowden Report, chapter 20.3, for more details).

Once cohort-members have been enrolled in a school applying streaming by abilities, they are allocated (according to their abilities), to higher, average or lower ability classes. The allocation to different ability-classes is possibly guided by individual unobserved factors, including cognitive and non-cognitive abilities and familiar background, which may contribute to explain later social participation decisions. This results in endogeneity problems. In this respect, we identify four potential instruments to be used in the empirical approach dealing with endogeneity issues. Particularly, we use a math and a reading test measured at age 7, the British Social Adjustment Guide score at age 7, approximating problematic children behaviors, and a dummy variable taking value one if the father of cohort-members, when he/she is aged 7 years old, belonged to the highest social class.

We investigate the impact of streaming during primary school on adult social participation propensity. By drawing information from the 5^{th}, 6^{th} and 8^{th} sweeps, we analyze the social participation propensity when cohort-members were aged, respectively, 33, 42 and 51. Social participation is measured according to a score based on a number of indicators. These indicators identify the participation of cohort-members to different type of organizations, including political parties, charity, environmental and voluntary associations, women organizations, school associations, resident associations and trade unions. In addition, the social participation score is also based on a binary variable indicating if the cohort-member voted in the last general election.

With the aim of accounting for observable heterogeneity in the social participation equation, we include a number of control variables. These include a gender dummy variable and another indicating if the individual is married or not. We also control for labour market status, since time devoted to social participation would be smaller in case of full-time employment (being inactive is the base-category), and health status, by including a dummy variable taking value one if the cohort-member reports a poor health status. In addition, we control the role of educational level, by including five dummy variables, for which no education is the base-category and, then, we distinguish among NVQ1, NVQ2, NVQ3, NVQ4 and (the highest) NVQ5-6 levels. Finally, we control for regional specific effects (Wales is the base-category) and time dummy variables (where 1991 is the base-category).

2.2. Descriptive statistics

The outcome of interest in our analysis is adult social participation, as measured by a standardized score (see Section 3) summarizing the individual binary responses to a number of social activities during adulthood. Participation in social activities here analyzed includes political party, environmental, charity and voluntary associations, women organizations, school and parents

associations, resident associations and trade union memberships, and participation in most recent general elections. The average participation in social activities is reported in Table 1.

Looking at the full sample, political party memberships is slightly greater for cohort-members having attended schools adopting streaming (3.7%) rather than not adopting streaming (3.4%). When looking at other activities, with the only exception of trade union memberships, the observed participation is greater for those having attended schools not adopting streaming. In particular, environmental, charity and voluntary associations' memberships regard 14.3% of not streamed cohort-members against 13% of streamed ones. A smaller gap is found when focusing on participation of women organization activities, which concern 2.3% of not streamed individuals and 2.1% of streamed ones. Participation in parents and/or school associations engage 14.1% of not streamed individuals and 12.6% of those enrolled in schools adopting streaming classes. Resident associations involve 5.9% of individuals enrolled in schools not applying streaming and 5.8% of those enrolled in schools applying streaming. Finally, 78.5% of individuals enrolled in schools not applying streaming and 76.6% of those applying streaming voted at the last general elections. In sum, preliminary evidence suggests a slight prevalence of participation in social activities for individuals enrolled in school not applying streaming when compared with their counterpart.

Table 1. Average standardized social participation score and social activity indicators

	Full sample			Streamed school sample		
		Not streamed	Streamed	High	Average	Low
Political parties	0.035	0.034	0.037	0.046	0.035	0.022
Environmental, charity, voluntary assoc.	0.138	0.143	0.130	0.167	0.106	0.085
Women organizations	0.022	0.023	0.021	0.024	0.018	0.020
Parents/School associations	0.136	0.141	0.126	0.160	0.114	0.074
Resident associations	0.059	0.059	0.058	0.078	0.044	0.035
Voted in last general elections	0.778	0.785	0.766	0.792	0.783	0.694
Observations	16598	10745	5853	2779	1679	1395

Source: our elaboration of NCDS data

When focusing on the sub-sample of individuals enrolled in schools applying streaming, we observe a noticeable heterogeneity in adult social participation according to the ability class to which individuals where allocated. Those allocated in high ability classes show the highest participation rate in social activities among streamed individuals. Conversely, those allocated in low ability classes show the lowest participation rate in social activities, with the only exception of women organization memberships. Individuals allocated in average-ability classes show an intermediate level of social participation. In addition, when comparing cohort members enrolled in schools not adopting streaming and those allocated in high-ability classes of schools adopting streaming, it emerges that social participation is higher among the latter individuals.

Table 2. Streaming, control variables and instruments

	Full sample (obs. 16598)		Not streamed (obs. 10745)		Streamed (obs. 5853)	
	Mean	Std dev.	Mean	Std dev.	Mean	Std dev.
Streamed schools	0.352	0.478	-	-	-	-
Low ability class	0.084	0.278	-	-	0.239	0.427
Average ability class	0.101	0.301	-	-	0.286	0.452
High ability class	0.167	0.373	-	-	0.475	0.499
Female	0.527	0.499	0.532	0.499	0.519	0.500
Married	0.786	0.410	0.788	0.409	0.783	0.412
Inactive	0.042	0.201	0.042	0.200	0.044	0.205
Domestic work	0.086	0.281	0.085	0.279	0.089	0.285
Unemployed	0.022	0.146	0.021	0.143	0.023	0.151
Self-employed	0.126	0.332	0.130	0.337	0.119	0.324
Part-time employed	0.169	0.374	0.168	0.374	0.170	0.376
Full-time employed	0.554	0.497	0.555	0.497	0.554	0.497
Poor health status	0.031	0.174	0.031	0.173	0.032	0.177
Less than education NVQ1	0.086	0.281	0.087	0.281	0.086	0.280
Education NVQ1	0.130	0.336	0.125	0.331	0.139	0.346
Education NVQ2	0.347	0.476	0.344	0.475	0.350	0.477
Education NVQ3	0.145	0.352	0.151	0.358	0.134	0.340
Education NVQ4	0.152	0.359	0.148	0.356	0.157	0.364
Education NVQ5-6	0.141	0.348	0.144	0.351	0.134	0.341
North	0.072	0.259	0.063	0.242	0.090	0.286
North-West	0.108	0.310	0.105	0.307	0.111	0.315
Yorkshire	0.100	0.300	0.103	0.304	0.095	0.293
West-Midlands	0.103	0.305	0.099	0.299	0.111	0.314
East-Midlands	0.085	0.279	0.088	0.283	0.079	0.270
Wales	0.062	0.240	0.056	0.230	0.072	0.258
East-Anglia	0.049	0.216	0.049	0.216	0.049	0.215
South-West	0.102	0.303	0.114	0.317	0.081	0.272
South-East	0.319	0.466	0.323	0.468	0.313	0.464
Educational meetings arranged	0.617	0.486	-	-	-	-
Social functions arranged for parents	0.530	0.499	-	-	-	-
Math test at age 7	-	-	-	-	5.435	2.447
Reading test at age 7	-	-	-	-	24.368	6.478
BSAG score at age 7	-	-	-	-	7.419	8.085
High father social class at age 7	-	-	-	-	0.203	0.403

Source: our elaboration of NCDS data

Table 2 informs about descriptive statistics. First two columns refer to the full sample, while the remaining informs about statistics distinguishing by streaming status. Cohort-members enrolled in schools adopting streaming represent 35.2% of the analyzed sample and, among them, those allocated in high ability classes represent 47.5%. The remaining streamed students were allocated to average ability classes (28.6%) and low ability classes (23.9%). Looking at the sample composition, differences between streamed and not streamed cohort members are relatively small. Therefore, we

briefly describe statistics referred to the full sample. In this regard, females represent 52.7% of sampled individuals, while 78.6% of cohort-members are married. Labor market status and, therefore, time allocation is a potential relevant predictor of adult social participation. Because we are considering individuals aged between 33 and 51 years old, just 4.2% of cohort members are inactive, while 8.6% carry out domestic work. The remaining cohort-members participate to the labor market. In particular, unemployed cohort members represent 2.2% of sampled individuals, 12.6% are self-employed, 16.9% are part-timers and 55.4% are full-timers. Individuals self-reporting a poor health status represent 3.1% of the sample. Educational level is resumed by six dummy variables, identifying the NVQ level achieved by cohort-members. 8.6% of sampled individuals achieved an education al level lower than the NVQ1, while 13% just achieved that mentioned level. The modal value in the educational level distribution is represented by the NVQ2 level, which has been achieved by 34.1% of sampled individuals. 14.5% of cohort-members achieved the NVQ3 educational level, while those achieving NVQ4 and NVQ5-6, the highest educational levels in the British system, represent 15.2% and 14.1%, respectively. Finally, we control heterogeneity in adult social participation, by including nine regional dummy variables. In this context, Scotland has been excluded from our analysis, because of a different application of streaming. When looking at the analyzed sample, 31.9% of sampled individuals live in the South-East region (including London). About 40% of individuals are quite equally distributed among the following regions: North-West, Yorkshire, West-Midlands and South-West. Finally, 8.5% of sampled cohort-members live in East-Midlands, 6.2% in Wales and 4.9% in East-Anglia.

Table 2 also reports descriptive statistics of instruments adopted in our econometric analysis. When focusing on the whole analysis, we use two instruments, indicating the arrangement of educational meetings and social functions for parents at school. The observed frequency in the data is 61.7% and 53%, respectively. When analyzing the sub-sample of individuals enrolled in schools applying streaming and, therefore, the allocation among ability classes, we instrument the suspected endogenous variable by using four instruments, namely a math and a reading test scores at age 7, a non-cognitive score (the British Social Adjustment Guide at age 7, and a dummy variable indicating if the father's social class when cohort-members were aged 7 was high.

3. ECONOMETRIC ANALYSIS

3.1 Obtain scores on the latent dimension

The first step of the analysis is assigning quantitative scores corresponding to the binary responses provided by every individual at each time occasion. For this aim, we use the Rasch model (Rasch, 1960, 1961), which is the most popular Item Response Model (Hambleton & Swaminathanan,

1984; Bartolucci et al., 2015). In summary, let Y_{ijt} the binary response provided by individual i at occasion t to item of type j, with $i = 1,\ldots,n$, $j = 1,\ldots,J$, and $t = 1,\ldots,T$. This model assumes that

$$\log \frac{p(Y_{ijt}=1|q_{it})}{p(Y_{ijt}=0|q_{it})} = q_{it} - b_j \qquad (1)$$

where q_{it} is interpretable as the "ability level" and b_j is interpretable as the "difficulty level" of item j. These definitions are typical of the educational context, where the Rasch model finds one of its main applications. In our specific context, there are $J = 6$ items (org_polit, org_voluntary, org_women, org_school, org_resident, voted) and then q_{it} is interpretable as the tendency of towards social participation of individual i at occasion t. On the other hand, b_j is an intercept that accounts for the general tendency of a certain behavior.

Regarding the distribution of q_{it}, we assume a discrete distribution if with an arbitrary number k of support points denoted by $\xi_1,\ldots,\xi_k$ and corresponding probabilities $\pi_1,\ldots,\pi_k$ (Lindsay *et al*., 1991). This avoids parametric assumptions as in Heckman and Singer (1984), while allowing to cluster individual in homogenous classes, named latent classes, and predicting the latent trait level. This prediction is the score that we use for the analysis. In order to estimate the resulting latent class Rasch model to the observed data, we use an Expectation Maximization algorithm Dempster *et al*. (1977), implemented in the R package MultiLCIRT; Bartolucci et al. (2014). Moreover, to select the number of latent classes we rely on the Bayesian Information Criterion (Schwarz, 1978) that is commonly used for this aim. This criterions leads to selecting $k = 3$ classes; the corresponding parameter estimates are reported in the following table.

Table 3. Parameter estimates under the Rasch model with $k = 3$ latent classes

Parameter	Estimate	Parameter	Estimate	Parameter	Estimate
b_1	0.000	x_1	-4.328	p_1	0.791
b_2	-1.663	x_2	-2.089	p_2	0.203
b_3	-0.175	x_3	0.083	p_3	0.006
b_4	-1.592				
b_5	-0.564				
b_6	-5.338				

Source: our elaboration of NCDS data.

On the basis of these estimates, we obtain the prediction of q_{it}, denoted by $\widehat{q}_{it}$, by an a-posteriori expected value for every individual and time occasion. Descriptive statistics for the distribution of

these predictions for each time occasions are reported in Table 4, denoting a decrease of the tendency towards social participation in time.

Table 4. Distribution of the scores for each period

Period	Mean	St.dev.
1	-3.658	0.777
2	-3.938	0.501
3	-3.946	0.494

Source: our elaboration of NCDS data.

3.2 The impact of streaming on social participation

In the second step of the analysis, we investigate the impact of streaming on adult social participation, as measured by the score predicted by the Rasch model.

In this regard, our first purpose is to assess the effect of attending a school adopting streaming or not on social participation propensity. This would be potentially estimated by applying a (pooled) ordinary least square (OLS) model including a binary indicator taking value one if the cohort-member has attended a school adopting streaming and zero otherwise. Although adopting or not streaming is a school principal's choice and, therefore, it would be potentially exogenous with respect to the adult social participation of cohort members, the parents' choice of enrolling cohort members in schools adopting streaming is, instead, potentially endogenous. In that case, the orthogonality assumption of OLS models would be violated and related estimates would be not consistent. With the aim of handling the issue of endogeneity we estimate a two-stage least square regression models (2SLS), for which the explanatory variables suspected of endogeneity is instrumented by using instruments correlated with the instrumented variable and not correlated with the disturbance[4].

The 2SLS model would be represented by the following two equations:

$$s_i = X_i\alpha_1 + Z_i\gamma + u_i \tag{2}$$

$$y_i = X_i\alpha_2 + \beta\hat{s}_i + \varepsilon_i \tag{3}$$

where s indicates if the individual i has been enrolled in a school adopting streaming or not, y represents the standardized score of social participation during adulthood, X is a vector of

[4] Even though the endogenous variables are dichotomous and, then, one would apply a binary first stage model, only the application of an ordinary least square estimation of the first stage guarantees that residuals are uncorrelated with fitted values and covariates (Angrist and Pischke, 2008).

exogenous determinants of social participation, Z is a set of instruments, α_1, α_2, β and γ are unknown parameters to be estimated and, finally, u and ε are normally distributed disturbances, with $E[\varepsilon \mid X, Z = 0]$. In particular, Z includes a couple of school-related variables, i.e. one indicating f educational meeting are arranged at school, and another one indicating if social functions are arranged for parents at school.

Once enrolled in schools adopting streaming, cohort-members were allocated to different classes according to their abilities. Therefore, a second purpose of our analysis is to investigate the effect of being allocated to a specific streamed class, by focusing on the sub-sample of cohort-members enrolled in schools adopting streaming. In particular, we investigate, in turn, the impact on adult social participation of being allocated to a high ability class (versus being allocated to average and low ability classes) and the impact on adult social participation of being assigned to a low ability class (versus being allocated to average and high ability classes). In both cases, the streamed-class effect is controlled for by including a binary indicator taking, in turn, value one if the individual has been allocated to a high ability class and a low ability class, respectively.

The allocation to different ability classes is possibly guided by cognitive and non-cognitive skills, as well as, parental background. These variables potentially affect the child development and, therefore, social participation in the long-run, resulting, in endogeneity problems. Similarly to above, we handle this issue by estimating two-stage least square regression models (2SLS). In particular, we run two specifications; the first one focuses on the impact of being enrolled in high ability classes (versus average and low ability classes), while the second one focuses on the impact of being enrolled in low ability classes (versus average and high ability classes). The related representation for both specifications is the following one:

$$c_i = X_i\omega_1 + W_i\eta + \nu_i \tag{4}$$

$$y_i = X_i\omega_2 + \delta\hat{c}_i + \varsigma_i \tag{5}$$

where c indicates if the individual attending a streamed school has been assigned, in turn, to a high ability class or not, or to a low ability class or not. W is a set of instruments, ω_1, ω_2, η and δ are unknown parameters to be estimated and, finally, υ and ς are normally distributed disturbances, with $E[\varsigma \mid X, W = 0]$. In particular, W includes four instruments, namely results in math and reading tests at age 7, the level of the British Social Maladjustment Guide (BSAG) score at age 7, to approximate non-cognitive skills, and a dummy variable indicating the father belonged to the high social class when the cohort-member was aged 7.

With the aim of testing the robustness of our analysis we run some specific tests. First, we run an endogeneity test of (suspected) endogenous regressor. A rejection of the test indicates the explanatory variable is, actually, endogenous. Second, for each specification, we run a LM test of whether the equation is identified, i.e., that the excluded instruments are relevant, meaning they are correlated with the endogenous regressors. A rejection of the null hypothesis indicates that the matrix is full column rank, i.e., the model is identified. Third, we run a weak identification test. Weak identification arises when the excluded instruments are correlated with the endogenous regressors, but only weakly. Weaker are the instruments, greater is the bias of the 2SLS estimator. Finally, we run an over-identification test for all instruments, for which the joint null hypothesis that the excluded instruments are valid instruments, i.e., uncorrelated with the error term and correctly excluded from the estimated equation is tested. In this case, a rejection would cast doubt on the validity of the instruments.

4. RESULTS

We estimated a 2-stage least square (2SLS) regression. At the first stage, we instrumented the covariate "school with streamed classes" using alternative instruments, depending on the model specification (Table A1 in the Appendix). In the estimates using the whole sample, we employed instruments that were related to "social" activities arranged by schools. Specifically, the database supplies information on "educational meetings arranged" and "social functions arranged for parents". When the analysis was restricted to the sub-sample of individuals that attended streamed schools, we employed the following instruments: the results for the Math and Reading tests at age 7, the BSAG score and a dummy for those with "high father social class". The estimated coefficients for all instruments have the expected signs. In particular, in the whole sample, the two instruments "educational meetings arranged" and "social functions arranged for parents" have negative and significant coefficients. This suggests that schools that organize these activities are less likely to be schools that adopt streaming. Unsurprisingly, in the sub-sample regression, the Math and Reading tests instruments are significant and positive (negative) in high (low) ability classes. On the opposite, the coefficients of the instrument "BSAG score", related to aggressive behaviors, are negative (positive) in high (low) quality classes. Finally, the instrument "high father social class" has a significant and negative sign only for low ability classes.

Table 5 reports the results for the second-stage estimation. The first two columns show the results for the regression based on the full sample. This analysis focuses on the impact of attending schools that adopt streaming on adult social participation. The subsequent columns show the results of the model based on the sub-sample of the only individuals that enrolled in schools that stream, and

investigates the impact of having attended a high ability class or a low ability class on adult social participation.

Table 5. 2SLS estimation results.

	FULL SAMPLE			STREAMED STUDENTS SUB-SAMPLE					
				High vs Low-Middle			Low vs Middle-High		
	Coeff.	R.s.e.		Coeff.	R.s.e.		Coeff.	R.s.e.	
School with streamed classes	-0.614	0.244	**	-	-		-	-	
High ability class	-	-		0.219	0.059	***	-	-	
Low ability class	-	-		-	-		-0.233	0.061	***
Female	0.191	0.020	***	0.094	0.034	***	0.099	0.033	***
Married	0.078	0.019	***	0.064	0.032	**	0.056	0.032	*
Inactive				base-category					
Domestic work	0.008	0.051		0.051	0.081		0.066	0.081	
Unemployed	-0.080	0.062		0.033	0.102		0.036	0.102	
Self-employed	-0.075	0.048		-0.022	0.078		-0.020	0.078	
Part-time employed	-0.023	0.047		0.098	0.077		0.104	0.077	
Full-time employed	-0.143	0.044	***	-0.062	0.071		-0.059	0.071	
Poor health status	-0.095	0.048	**	-0.100	0.076		-0.094	0.076	
No education				base-category					
Education NVQ1	0.128	0.026	***	-0.021	0.043		-0.057	0.045	
Education NVQ2	0.297	0.023	***	0.154	0.046	***	0.141	0.047	***
Education NVQ3	0.440	0.029	***	0.244	0.058	***	0.242	0.058	***
Education NVQ4	0.586	0.029	***	0.399	0.056	***	0.386	0.058	***
Education NVQ5-6	0.981	0.033	***	0.708	0.077	***	0.733	0.072	***
North	-0.092	0.036	**	-0.223	0.042	***	-0.224	0.043	***
North-West	-0.104	0.027	***	-0.118	0.047	**	-0.117	0.046	**
Yorkshire	-0.133	0.027	***	-0.065	0.046		-0.069	0.046	
West-Midlands	-0.055	0.027	**	-0.130	0.044	***	-0.139	0.044	***
East-Midlands	-0.090	0.029	***	-0.171	0.051	***	-0.145	0.051	***
Wales	0.039	0.037		0.014	0.053		-0.001	0.052	
East-Anglia	-0.102	0.036	***	-0.159	0.060	***	-0.152	0.060	**
South-West	-0.100	0.032	***	-0.091	0.050	*	-0.080	0.050	
South-East				base-category					
Year 1991				base-category					
Year 2000	0.003	0.016		-0.016	0.026		-0.015	0.026	
Year 2009	0.478	0.020	***	0.428	0.034	***	0.428	0.034	***
Constant	-0.359	0.103	***	-0.489	0.092	***	-0.339	0.104	***

Note. R.s.e. indicates robust standard error

The general result of the estimation supports the main idea of this article: attending streamed schools reduces social participation when adult. This is shown by the estimated coefficient for the covariate “School with streamed classes” in the full sample regression. The coefficient is significant and negative, meaning that individuals that attended a streamed school when young have a lower attitude to participate in social activities when adult. However, when the analysis is restricted to the

only individuals that attended streamed schools and the estimated coefficient detects the effect of being in a high-ability class versus a medium-low ability class, then the grouping of students with high abilities has a positive impact on social participation. Being in a streamed school and in a high-ability class increases social participation when adult. The result is robust to the change in the base category: when the regression is carried out with the dummy that identifies individuals that were in low-ability classes, the coefficient is significantly negative, meaning that being in low-ability classes negatively affects social participation.

Our analysis controls for a number of factors that can affect social participation during adulthood. These include gender, marital status, labor market status, health status, education and region of living.

Females have a higher attitude for social participation. All models confirm this result. Also being married increases social participation. The importance of controlling for gender status is supported by the results of Clark and Del Bono (2016) in their study about the long-run effects of attending an elite school.

The results on the occupational status sharply show how the time constraint is a key factor for social participation. In fact, being employed full-time significantly and negatively influences participation in social activities. In addition to that, health is important for social participation: a poor health has a negative impact on civic behavior.

Education is determinant for social participation: the higher the level of education, the more individuals participate in social activities.

The controls for the macro regions (South-West is the base category and it includes London) highlight how important is urbanization and socio-economic density for social involvement.

The significant and positive sign on the dummy for year 2009 seems to suggest that social participation is higher in this year (relative to year 1991)) because of the economic and social impact of the "Great Recession".

Finally, we briefly comment results concerning the instruments we used in the 2SLS first stage estimations (see Table A1) which display expected signs. In particular, in the full sample, the two instruments "educational meetings arranged" and "social functions arranged for parents" have negative and significant coefficients, also suggesting that these activities are less probable in streamed schools. Not surprising, the "Math" and "Reading" tests instruments both are significant and positive (negative) in high (low) ability classes. On the opposite, the coefficients of the instrument "BSAG score", related to aggressive behaviors, are negative (positive) in high (low) quality classes. Finally, the instrument "high father social class" has a significant and negative sign only for low ability classes.

CONCLUSIONS

School organization would be an important aspect for the formation of individual social behavior and it may affect the building of social capital that has been recognized as a relevant factor for economic success and growth.

In this context, we analyze how streaming, i.e. the practice of homogenous grouping that assigns pupils to classes on the basis of an overall assessment of their general ability, affects socio-political participation during adulthood. The underlying idea is that applying streaming in primary schools may create patterns of interactions among pupils that may persist during their lives and influence their civic behavior. In particular, the attendance of primary schools applying streaming may reinforce polarization, and strengthen trust and cooperative norms within homogeneous groups, but weaken trust and cooperation between groups.

We empirically test our hypothesis, by using information from the 1958 National Child Development Study that collects information on the type of school attended during childhood and provides data on socio-political activities during adulthood. The outcome variables are obtained by collapsing individual binary responses for each participation category in standardized quantitative scores, as determined by using the Rasch model. Because the choice of being enrolled in a streaming school and the assignment to a specific ability-class would be endogenous, our estimation strategy consists in adopting 2SLS methods.

The obtained results are suggestive of a significant role of streaming during primary school on adult socio-political participation. Specifically, our first analysis, clarifies that cohort-members enrolled in primary schools applying streaming, are negatively affected in terms of socio-political participation during adulthood, when compared with cohort-members enrolled in primary schools do not adopting streaming and, therefore, assigning students to mixed-ability classes. Nevertheless, the situation of streamed students is heterogeneous. Focusing just on streamed students, our second-step analysis reveals that cohort-members assigned to high-ability classes are more likely to participate in socio-political activities during adulthood, when compared to other streamed students. In other terms, our results are suggestive of a relevant marginalization in terms of social and civic behavior especially for those cohort-members banished in low-ability classes, while mixed classes are more likely to reinforce cooperation between groups.

REFERENCES

Alesina A. and E. La Ferrara (2000) Participation in heterogeneous communities, *The Quarterly Journal of Economics* CXV, 847-904.

Angrist J., Pischke J.S. (2008). Mostly Harmless Econometrics. Princeton University Press, Princeton, New Jersey.

Bartolucci, F., Bacci, S., and Gnaldi, M. (2014). MultiLCIRT: An R package for multidimensional latent class item response models, *Computational Statistics & Data Analysis*, 71, pp. 971-985.

Bartolucci, F., Bacci, S., and Gnaldi, M. (2015), *Statistical Analysis of Questionnaires: A Unified Approach Based on Stata and R*, Chapman and Hall/CRC press.

Clark and Del Bono (2016), The long-run effects of attending an elite school: Evidence from the United Kingdom, *American Economic Journal: Applied Economics*, vol. 8(1): 150-176.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society*, *Series B*, 39, pp. 1-38.

Hambleton, R. K. and Swaminathan, H. (1985). Item Response Theory: Principles and Applications. Kluwer Nijhoff, Boston.

Heckman, J. and Singer, B. (1984), A method for minimizing the impact of distributional assumptions in econometric models for duration data, 52, *Econometrica*, pp. 271-320.

Knack S. and P. Keefer (1997) Does social capital have and economic payoff? A cross-country investigation, *The Quarterly Journal of Economics* CXII, 1252-1288.

Lindsay, B., Clogg, C., and Greco, J. (1991). Semiparametric estimation in the Rasch model and related exponential response models, including a simple latent class model for item analysis. *Journal of the American Statistical Association*, 86, pp. 96-107.

Putnam R. (1993) Making democracy work. Princeton, NJ: Princeton University Press.

Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. Danish Institute for Educational Research, Copenhagen.

Rasch, G. (1961). On general laws and the meaning of measurement in psychology. In Neyman, J. editor, Proceedings of the IV Berkeley Symposium on Mathematical Statistics and Probability, volume 4, pp. 321-333.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, pp. 461-464.

Sobel J. (2002) Can we trust social capital? *Journal of Economic Literature*, 40, 139-154.

The Plowden Report (1967). Children and their primary school. A Report of the Central Advisory Council for Education, London, England.

Zak Paul J. and S. Knack (2001) Trust and growth, *The Economic Journal*, 111, 295-321.

APPENDIX

Table A1. 2SLS first stage estimations

	FULL SAMPLE			STREAMED STUDENTS SUB-SAMPLE					
Endogenous variable:	Streamed school			High ability class			Low ability class		
	Coeff.	R.s.e.		Coeff.	R.s.e.		Coeff.	R.s.e.	
Female	-0.028	0.009	***	0.085	0.013	***	-0.044	0.012	***
Married	-0.008	0.009		-0.010	0.014		-0.024	0.013	*
Inactive				base-category					
Domestic work	0.020	0.023		-0.024	0.035		0.091	0.032	***
Unemployed	0.008	0.032		-0.063	0.049		0.069	0.046	
Self-employed	-0.024	0.022		-0.074	0.034	**	0.080	0.032	**
Part-time employed	0.011	0.022		-0.032	0.033		0.055	0.030	*
Full-time employed	-0.009	0.020		-0.016	0.031		0.036	0.029	
Poor health status	0.002	0.023		-0.037	0.030		0.060	0.034	*
No education				base-category					
Education NVQ1	0.031	0.016	*	-0.007	0.021		-0.140	0.026	***
Education NVQ2	0.017	0.014		0.146	0.020	***	-0.182	0.024	***
Education NVQ3	-0.012	0.016		0.203	0.024	***	-0.188	0.026	***
Education NVQ4	0.027	0.016	*	0.173	0.024	***	-0.210	0.026	***
Education NVQ5-6	0.001	0.016		0.368	0.025	***	-0.239	0.025	***
North	0.077	0.016	***	-0.013	0.022		0.004	0.018	
North-West	0.014	0.013		0.026	0.019		-0.026	0.016	
Yorkshire	-0.023	0.013	*	-0.003	0.021		-0.014	0.017	
West-Midlands	0.027	0.013	**	0.000	0.019		-0.039	0.016	**
East-Midlands	-0.021	0.014		0.128	0.021	***	-0.015	0.019	
Wales	0.047	0.017	***	-0.062	0.023	***	-0.018	0.019	
East-Anglia	-0.001	0.018		-0.060	0.028	**	0.090	0.025	***
South-West	-0.064	0.013	***	-0.038	0.022	*	0.076	0.021	***
South-East				base-category					
Year 1991				base-category					
Year 2000	0.002	0.009		0.002	0.014		0.003	0.012	
Year 2009	0.003	0.009		0.002	0.014		0.000	0.012	
Constant	0.453	0.028	***	-0.439	0.045	***	1.010	0.046	***
Instruments:									
Educational meetings arranged	-0.033	0.008	***						
Social functions arranged for parents	-0.047	0.008	***						
Math test				0.036	0.003	***	-0.014	0.002	***
Reading test				0.022	0.001	***	-0.024	0.001	***
BSAG score				-0.005	0.001	***	0.007	0.001	***
High father social class				0.001	0.015		-0.022	0.012	*

Source: our elaboration of NCDS data.
Note. R.s.e. indicates robust standard error.

Table A2. Tests

	Full sample	Sub-sample	
	Streamed school	High ability class	Low ability class
Endogeneity test of endogenous regressor	6.38	10.26	10.05
Chi-square(1) P-value	0.01	0.00	0.00
Underidentification test (Kleibergen-Paap rk LM statistic):	69.62	943.08	688.04
Chi-square(2) P-value	0.00	0.00	0.00
Weak identification test (Kleibergen-Paap rk Wald F statistic):	35.00	405.00	259.56
Stock-Yogo weak ID test critical values:			
5% maximal IV relative bias	-	16.85	16.85
10% maximal IV relative bias	-	10.27	10.27
20% maximal IV relative bias	-	6.71	6.71
30% maximal IV relative bias	-	5.34	5.34
10% maximal IV size	19.93	24.58	24.58
15% maximal IV size	11.59	13.96	13.96
20% maximal IV size	8.75	10.26	10.26
25% maximal IV size	7.25	8.31	8.31
Hansen J statistic (overidentification test of all instruments):	0.11	4.63	5.03
Chi-square(1) P-value	0.74	0.20	0.17

Source: our elaboration of NCDS data.