THE
ECONOMIC
JOURNAL

The Economic Journal, **129** (*August*), 2459–2521 DOI: 10.1093/ej/uez003
Advance Access Publication Date: 23 April 2019

NEIGHBOURHOOD ETHNIC COMPOSITION AND SOCIAL PARTICIPATION OF YOUNG PEOPLE IN ENGLAND*

Elena Fumagalli and Laura Fumagalli

We analyse how neighbourhood ethnic diversity and segregation affect adolescents' social participation in England. We distinguish between participation in 'purposeful activities'—such as sports and volunteering—and hanging around with friends. We suggest a novel identification strategy to address the problem of endogeneity of ethnic diversity and segregation. We find that ethnic diversity decreases hanging around, while ethnic segregation increases it. No effects on participation in purposeful activities are found.

Britain has become increasingly concerned about whether the co-existence of different ethnicities may affect the quality of social interaction. Data from the Market & Opinion Research International (MORI) Issues Index show that, since the early 2000s, the quality of the relations among different ethnic groups has been perceived as a major issue faced by Britain, which often scores higher than other salient issues such as the state of the economy or the National Health Service.[1] This article investigates the roots of these concerns for England. It studies how social participation of young people aged 14 or 15 years old is affected by the ethnic composition of the neighbourhood where adolescents live.

We pay particular attention to the role of geography in shaping social participation. Our measures of neighbourhood ethnic composition are ethnic *diversity* and *segregation* at the local authority district (henceforth 'district') level.[2] District ethnic diversity does not consider the spatial distribution of ethnicities within the district. This spatial distribution is captured by district ethnic segregation. In our case, district ethnic segregation captures how ethnicities are distributed

* Corresponding author: Laura Fumagalli, University of Essex, Institute for Social and Economic Research, Wivenhoe Park, CO43SQ, Colchester, UK. Email: lfumag@essex.ac.uk

This paper was received on 11 October 2016 and accepted on 31 May 2018. The Editor was Kjell Salvanes.

We are grateful to the editor Kjell Salvanes and two anonymous referees for comments that have significantly improved the article. We are also grateful to Wiji Arulampalam, Agar Brugiavini, Paul Clarke, Sergio Currarini, Emilia Del Bono, John Ermisch, Adriaan Kalwij, Thomas Martin, John Micklewright, Mario Padula, Fabrizio Panebianco, Steve Pudney, Anna Salomons, Jordi Vidal-Robert and seminar participants at University of Venice, University of Essex, Lancaster University, UCL Institute of Education, University of Groningen, St. Gallen University, Utrecht University, JRC European Commission, IZA summer school, QED meeting, ESPE conference, APET conference, Popfest conference, EEA/ESEM conference, EALE conference, SIEP conference, EURODIV conference. Laura Fumagalli acknowledges support from the ESRC through the Research Centre on Micro-Social Change (MiSoC) at the University of Essex, reference (ES/L009153/1). This article is based on work carried out during Elena Fumagalli's visit to the European Centre for Analysis in the Social Sciences (ECASS) at the Institute for Social and Economic Research, University of Essex, supported by the Access to Research Infrastructures action under the EU Improving Human Potential Programme.

[1] Ipsos MORI interviews a representative sample of around 1,000 18+ people in the UK about the most important issues faced by the country. No pre-defined answers are provided.

[2] Districts are identified by the first four digits of the Office for National Statistics (ONS) coding system. At the time of our data there were 354 districts in England: metropolitan districts (36), non-metropolitan districts (239), London boroughs (33, i.e., City of London, Barking and Dagenham, Barnet, Bexley, Brent, Bromley, Camden, Croydon, Ealing, Enfield, Greenwich, Hackney, Hammersmith and Fulham, Haringey, Harrow, Havering, Hillingdon, Hounslow, Islington, Kensington and Chelsea, Kingston upon Thames, Lambeth, Lewisham, Merton, Newham, Redbridge, Richmond upon Thames, Southwark, Sutton, Tower Hamlets, Waltham Forest, Wandsworth and Westminster), and unitary authorities (46). Districts were composed on average by 23 electoral wards: the spatial units used for the election of local government councillors, identified by the first six digits of the ONS coding system.

into wards within the district. To define social participation we distinguish 'purposeful activities' (e.g., playing sports, taking part in political activities or community work, and joining youth clubs) from 'hanging around with friends'.[3] We describe the effect of district ethnic composition on the demand for social activities by proposing a simple explanatory model where purposeful activities take place in the district and hanging around with friends takes place in the ward. We test its predictions using the Longitudinal Study of Young People in England (LSYPE).

Studying the determinants of social participation is important. Social participation is an objective measure of social capital (Alesina and La Ferrara, 2000; Glaeser *et al.*, 2000; Durlauf and Fafchamps, 2005).[4] Not all forms of social participation are equal. Participation in purposeful activities at a young age predicts many positive outcomes in adolescence and adulthood. Purposeful activities improve cognitive and non-cognitive skills (Broh, 2002; Hansen *et al.*, 2003; Del Boca *et al.*, 2017), educational attainment (Long and Caudill, 1991; Barron *et al.*, 2000; Eide and Ronan, 2001; Lipscomb, 2007; Lechner, 2009; Pfeifer and Cornelißen, 2010; Rees and Sabia, 2010; Stevenson, 2010; Felfe *et al.*, 2016) and labour market outcomes (Long and Caudill, 1991; Barron *et al.*, 2000; Kuhn and Weinberger, 2005; Henderson *et al.*, 2006; Lechner, 2009; Stevenson, 2010; Weinberger, 2014; Lechner and Sari, 2015; Deming, 2017; Lechner and Downward, 2017). Hanging around with friends is often claimed to be associated with undesirable outcomes, such as drunkenness and disorders (DfES, 2006).[5] However, we are not aware of any systematic analysis of hanging around with friends, or other social activities with no specific purpose.

The article fills this gap in the literature by drawing attention to the phenomenon of hanging around. We believe that this is important. In Appendix A we discuss the characteristics and the potential consequences of hanging around. Most young people hang around with friends at least once a week. Young people's hanging around decreases adults' well-being. For young people, hanging around with friends is associated with lower human capital accumulation, higher involvement in risky behaviours and more close friends in adulthood. We use propensity score techniques to claim these associations—particularly those on risky behaviours and number of close friends—are likely to be at least partially causal. For comparison, Appendix A also shows that purposeful activities are associated with—and possibly cause—higher human capital accumulation and more close friends in adulthood.

A growing literature in economics claims that ethnic diversity discourages social participation (Alesina and La Ferrara, 2000; Costa and Kahn, 2003; Charles and Kline, 2006).[6] This literature typically assumes that individuals derive greater utility from forming ethnically homogeneous groups. As in ethnically diverse areas, the probability of making a friend of the same ethnicity is lower than in ethnically homogeneous ones, ethnic diversity results in a decrease of social interaction. Studies on the effect of segregation are limited in number and lead to mixed conclu-

[3] Purposeful activities are often referred to as 'positive activities' (e.g., DfES, 2003). We prefer the definition of 'purposeful activities', as it only implies that these activities have a specific goal, without making any judgment on their desirability.

[4] Social capital has been shown to predict a number of key outcomes (see Durlauf and Fafchamps, 2005), including economic performance (Knack and Keefer, 1997; Zak and Knack, 2001), financial development (Glaeser *et al.*, 2002), and health (Folland, 2007; Rocco *et al.*, 2014).

[5] Hanging around on street corners was antisocial behaviour under the 1998 Crime and Disorder Act.

[6] A negative effect of ethnic diversity has been found for other indicators of social capital like: trust (Alesina and La Ferrara, 2002; Delhey and Newton, 2005; Pennant, 2005; Anderson and Paskeviciute, 2006; Leigh, 2006; Putnam, 2007), number of friends (Putnam, 2007). Moreover, ethnic diversity has been found to be correlated with: economic and institutional performance of countries (Easterly and Levine, 1997; Alesina *et al.*, 2003; Alesina and La Ferrara, 2005), provision of public goods (Alesina *et al.*, 1999; Luttmer, 2001; Vigdor, 2004; Alesina *et al.*, 2015), firm productivity and innovation (Böheim *et al.*, 2014; Parrotta *et al.*, 2014), number of political jurisdictions (Alesina *et al.*, 2004).

sions. Ethnic segregation has been found to increase the willingness to contribute to local public goods (La Ferrara and Mele, 2007), but also to erode generalised trust (Uslaner, 2011, 2012). To our knowledge, the effect of ethnic segregation on social participation has never been explicitly studied.[7]

The article makes a number of contributions to the existing literature on the effect of neighbourhood ethnic composition on social participation. First, it focuses on young people, as evidence suggests that adolescence is when relational skills are formed (Cunha and Heckman, 2008; Cunha *et al*., 2010) and social participation in adolescence predicts human capital accumulation, risky behaviours, future sociability and labour market outcomes (see Appendix A). Second, it distinguishes between different forms of social participation, thus accounting for the quality of young people's interaction. Third, it provides the first attempt of modelling the effect of segregation (in addition to ethnic diversity) on social participation, by emphasising the geographical level where hanging around and purposeful activities take place. This is particularly meaningful for young people, who are more restricted in how far they can travel from home, compared with adults. Finally, the article addresses the problem of the potential endogeneity of neighbourhood ethnic diversity and segregation with respect to social participation in a number of ways, including by proposing new instruments based on historical events and legislative acts which shaped the sorting of migrants within England.

We find that district ethnic diversity decreases hanging around, while district ethnic segregation increases it. The effects of district ethnic diversity and segregation on purposeful activities are weak or null. The results are robust to alternative empirical specifications. As hanging around is widely perceived as an undesirable behaviour, we conclude that ethnic diversity does not have detrimental effects on teenagers' social participation. Therefore, policies limiting migration in an attempt to preserve ethnic homogeneity are not desirable. Desegregation policies and policies improving the supply of purposeful activities are likely to be more beneficial.

1. Data and Descriptive Statistics

1.1. *Data*

The main data that we use are from the the Longitudinal Study of Young People in England (LSYPE), a cohort study containing detailed information for approximately 15,000 pupils living in England and born between 1 September 1989 and 31 August 1990. The LSYPE is a two-stage probability sample with schools as primary sampling units and students within schools as secondary sampling units.[8] This article is based on the first wave of the data, collected in 2004 when the respondents were 14/15 years old (year 9 at school).

Additional data are added to the LSYPE. Data from the 2001 Population Census (the closest census data to wave one of LSYPE) are used to compute indices of district ethnic diversity and

[7] Alesina and La Ferrara (2000) discuss a potential effect of ethnic segregation (on top of ethnic diversity) on social participation, but acknowledge that their analysis 'does not deal directly with segregation'.

[8] Maintained schools are stratified by deprivation level (measured as number of pupils receiving free school meal). Deprived schools are over-sampled and so are ethnic minority students. Independent schools are stratified by GCSE results in 2003, boarding status and gender of the students. Pupil Referral Units (PRUs) are a stratum of their own. The selection probability of both independent schools and PRUs is proportional to the number of pupils aged 13 enrolled in the institution. Children educated at home are excluded from the sample and so are children attending very small schools (with fewer than 10 pupils for the maintained sector and fewer than 6 for the independent sector), boarders and children who are in England just for education purposes. Response rate at wave one is 74% (66% full interviews and 9% partial interviews).

segregation and to construct controls at the district level used in the robustness checks.[9] Districts have been chosen to indicate neighbourhoods as they are responsible for cultural and recreational functions, and thus are the relevant geographical level for looking at social participation. Moreover, with a population ranging from approximately 300,000 to 1,000,000 people (with even large cities like Birmingham or Leeds being formed by a single district), districts are likely to cover the area where most of the social participation takes place. Finally, weather data (monthly millimetres of rainfall and hours of sun) from the UK Met Office are added to account for weather conditions.[10]

1.2. *Measuring Ethnic Diversity and Segregation*

District ethnic diversity is measured by the 'fractionalisation' (or 'fragmentation') index, as in Easterly and Levine (1997); Alesina and La Ferrara (2000); Alesina *et al.* (2003); Ashraf and Galor (2013); Böheim *et al.* (2014); and Parrotta *et al.* (2014).

DEFINITION 1 (F INDEX). *The fractionalisation index for district j (F_j) is defined as:* $F_j = 1 - \sum_k s_{kj}^2$, *where s_{kj} is the share of ethnic group k over the population of district j.*

F_j is constructed as 1 − the Herfindahl-Hirschman (1964) index and can be interpreted as the probability that two individuals randomly drawn from the population of district j belong to two different ethnic groups. F_j ranges from 0 to 1, with higher values indicating high diversity.

District ethnic segregation is measured by the 'Duncan and Duncan segregation index', as in Duncan and Duncan (1955); Cutler and Gleaser (1997); and Cutler *et al.* (1999).

DEFINITION 2 (D INDEX). *The Duncan and Duncan segregation index is defined as:* $D_j = \frac{1}{2}\sum_{w=1}^{W_j}\left|\frac{m_{wj}}{M_j} - \frac{b_{wj}}{B_j}\right|$, *where m_{wj} and b_{wj} are, respectively, the number of people in the ethnic minority group and in the White British group living in ward $w = 1,.., W_j$. M_j and B_j are, respectively, the number of people in the ethnic minority group and in the White British group in district j, and W_j is the number of wards in district j.*

D_j can be interpreted as a share of people belonging to one of the ethnic groups of district j that should move to another ward within district j (without being replaced) in order to make ward ethnic diversity equal to district ethnic diversity. D_j ranges from 0 to 1, with higher values indicating high segregation.

[9] Census data report the raw number of people living in each ward classified by ethnic group, which we aggregate into coarser categories, as in Berthoud *et al.* (2009). Using the LSYPE to extract information on the area where respondents live (including the indices of ethnic diversity and segregation) is not a viable strategy, because of potential non-coverage and non-response bias, and because any geographical aggregation small enough to be defined as 'neighbourhood' has a very small sample size. For the use of survey data based indices, see: Carrington and Troske (1997); Jenkins *et al.* (2008); Rathelot (2012); Allen *et al.* (2015).

[10] These data are collected monthly from 37 UK meteo stations (Aberporth, Armagh, Ballypatrick Forest, Bradford, Braemar, Camborne, Cambridge, Cardiff Bute Park, Chivenor, Cwmystwyth, Dunstaffnage, Durham, Eastbourne, Eskdalemuir, Heathrow, Hurn, Lerwick, Leuchars, Lowesoft, Manston, Nairn, Newton Rigg, Oxford, Paisley, Ringway, Ross-on-Wye, Shawbury, Sheffield, Southampton, Stornoway Airport, Sutton Bonington, Tiree, Valley, Waddington, Whitby, Wick Airport, Yeovilton). Each LSYPE respondent is assigned the weather information collected from the closest meteo station in the interview month. Missing weather data are imputed by chained equation imputation (Raghunathan et al., 2001) using as predictors maximum and minimum temperature, days of frost, millimetres of rainfall and hours of sun for each area and month, plus their values in the previous and in the subsequent month.

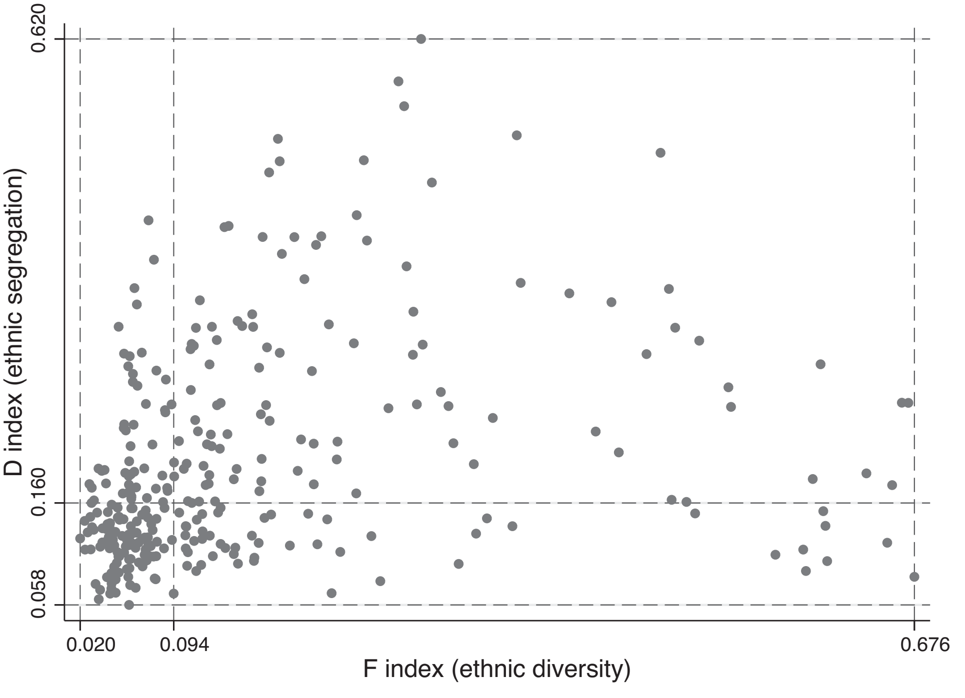

Fig. 1. *District Ethnic Diversity and Segregation.*
Notes: Dashed lines indicate the minimum, maximum and the median values for each index. N = 313 (districts included in the analysis). Census data. Pairwise correlation: 0.275.

The F index and the D index capture two different aspects of district ethnic composition. For a given positive level of the F index, the D index can take a wide range of values in the interval [0,1]. Figure 1 plots district ethnic segregation against district ethnic diversity for the districts included in our main analysis. Given levels of district ethnic diversity are generally associated with a wide range of levels of district ethnic segregation.

1.3. *Measuring Social Participation*

To measure purposeful activities and hanging around we use two LSYPE questions asking young people about the activities they have carried out when not at school in the four weeks before the interview. The first question lists a set of activities including: sport, playing an instrument, going to the cinema/theatre/concert, or partying. The second question lists activities implying an engagement in the community (e.g., going to youth clubs, doing community work) and two forms of 'hanging around' (near home and in the city centre).

We measure participation in purposeful activities through an indicator (*PA*) equal to one if the respondent reports carrying out at least one activity among: playing sport, taking part in political activities, going to youth clubs, doing community work. *PA* is zero otherwise. Two types of activities are excluded from *PA*: (*i*) those not necessarily requiring interaction among people, such as playing a musical instrument, going to the cinema/theatre and going to a concert and (*ii*) those not necessarily 'positive' in the definition of DfES (2006), e.g., partying, dancing, watching

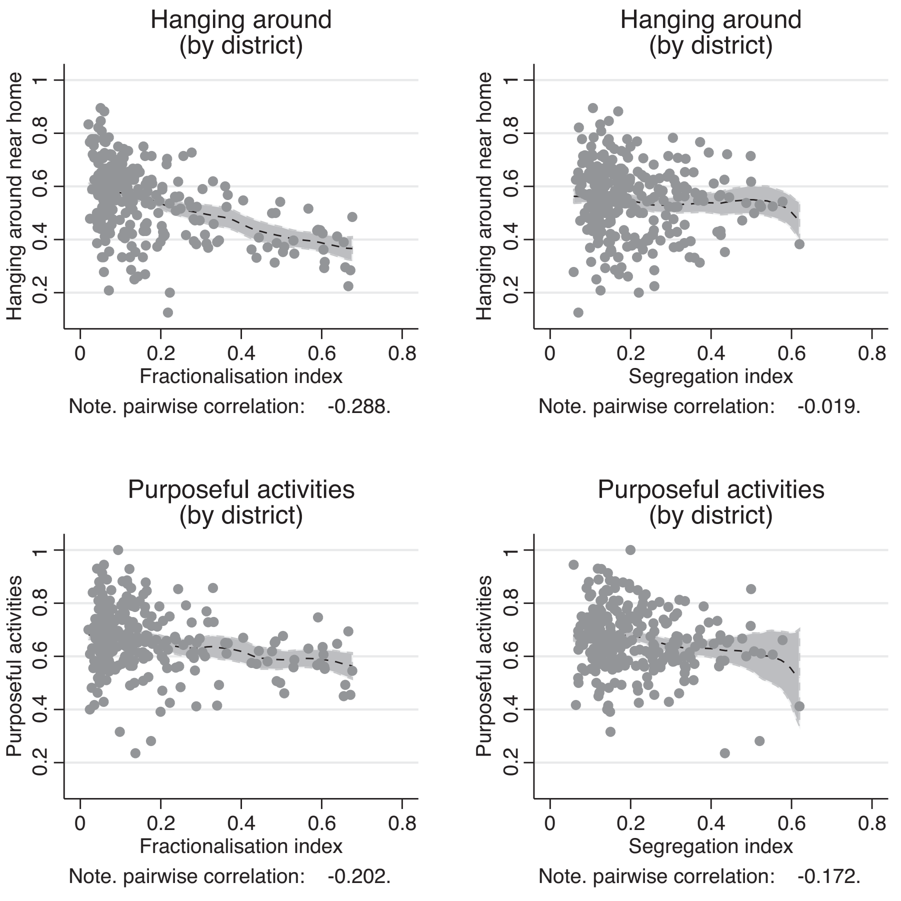

Fig. 2. *District Ethnic Composition and Social Participation.*
Notes: Fitted line: local polynomial regression (triangular kernel); shaded area: 95% confidence intervals. *HA* = 1 if the respondent reports hanging around with friends near home. *PA* = 1 if the respondent reports carrying out at least one activity among: playing sport, taking part in political activities, going to youth clubs, doing community work. Census and LSYPE data. LSYPE data are weighted. N = 274 (excludes districts with fewer than five respondents). Just districts included in the analysis have been plotted.

a football match, going to a nightclub, going to the pub, going to the amusement arcade and playing snooker, as they may be associated with risky behaviours, such as fights, drunkenness or illegal substance use. Our measure of 'hanging around' is 'hanging around near home'. It is measured through an indicator (*HA*) equal to one of the respondent reports 'hanging around near home', and zero otherwise. 'Hanging around in the city centre' is excluded because it might mean something completely different for young people from different areas of the city.[11]

Figure 2 plots district level prevalence of *PA* and *HA* against district ethnic diversity and segregation. *HA* decreases with district ethnic diversity (see top left panel). The share of young people hanging around drops from 60% in the least ethnically diverse districts to less than 40%

[11] Hanging around in the city centre is analysed in Section 6.

Table 1. *Participation in Social Activities (by Income Quintiles).*

Income quintile	*PA* Mean	*HA* Mean	Sport Mean	Political activities Mean	Community work Mean	Youth clubs Mean
All	0.567	0.657	0.572	0.017	0.045	0.045
	(0.004)	(0.004)	(0.004)	(0.001)	(0.002)	(0.002)
Bottom quintile	0.560	0.586	0.492	0.021	0.045	0.045
	(0.009)	(0.009)	(0.009)	(0.003)	(0.004)	(0.004)
Second quintile	0.583	0.601	0.507	0.015	0.040	0.040
	(0.009)	(0.009)	(0.009)	(0.002)	(0.004)	(0.004)
Third quintile	0.582	0.655	0.568	0.018	0.044	0.044
	(0.009)	(0.008)	(0.009)	(0.002)	(0.004)	(0.004)
Fourth quintile	0.583	0.696	0.615	0.018	0.049	0.049
	(0.010)	(0.009)	(0.010)	(0.003)	(0.004)	(0.004)
Top quintile	0.526	0.730	0.655	0.014	0.047	0.047
	(0.010)	(0.009)	(0.009)	(0.002)	(0.004)	(0.004)
N	14,244	14,244	14,244	14,244	14,244	14,244

Notes: LSYPE wave 1, weighted. The sample is the same used in the empirical analysis discussed in Section 3. Income is total gross household income. Standard errors in parentheses.

in the most diverse ones. *HA* does not seem to vary with district ethnic segregation (see top right panel). No relationship between *PA* and either measure of district ethnic composition seems to exist, with the average share of those who participate in *PA* being just above 60% (see bottom panels).

The bivariate relationships in Figure 2 must be interpreted with caution, as they can be partially due to respondents' characteristics differing by districts and also affecting social participation. For example, the effects of household economic deprivation might be mistakenly interpreted as effects of neighbourhood ethnic composition (see Letki, 2008; Fieldhouse and Cutts, 2010; Sturgis *et al.*, 2011). This issue is addressed in Table 1, where social participation is broken down by quintiles of total gross household income.[12] Participation in *PA* increases with household income. The share of those participating in *PA* is 73% among respondents in the top quintile of household income and only 59% among those in the bottom quintile. These results are driven by participation in sports. In contrast, participation in *HA* is pretty constant within quintiles of household income. A full set of descriptive statistics for the sample that we use can be found in Table B1 in Appendix B.

2. Explanatory Model

Our formalisation of the effect of district ethnic diversity and segregation on *HA* and *PA* is based on two assumptions. The first assumption is that *HA* and *PA* have different geographical scopes. *HA* is likely to take place near home, as made explicit by the wording of the question we use to measure it. We refer to the ward where young people live as the space where *HA* takes place. Appendix A shows that *HA* generally takes place near home and suggests that this assumption is plausible in our case. *PA* is likely to take place in an area broader than the ward. We refer to the district where young people live as the space where *PA* takes place. The reasons for this choice are discussed in Subsection 1.1.

[12] In LSYPE wave one, total household income is available as a banded variable, with bandwidth increasing with income. Missing values have been inputed by chain equation (Raghunathan *et al.*, 2001).

The second assumption is that young people care about the ethnicity of those participating in *HA* and *PA* and have a preference for people of their own ethnic group. This preference for one's own ethnic group, at the core of many explanations of social interaction (e.g., Alesina and La Ferrara, 2000; Charles and Kline, 2006), has been documented for adolescents by Currarini *et al.* (2009). In Appendix A we use information about the ethnicity of the friends of LSYPE respondents to claim this assumption is plausible in our case.

Our explanatory model focuses on the demand side of social participation. At least in the short run, the supply of social activities is given. The role of supply side is discussed in Section 7. Consider a district, divided into two wards, and assume young people come in two types: Greens (g) and Purples (p). There are two social activities: *HA* (taking place in the ward) and *PA* (taking place in the district). The young people decide whether to participate in *HA* and *PA* based on an expected utility that depends on their expectations about the ethnicity of the participants in the activity. We assume that young people approximate the ethnicity of the participants in *HA*(*PA*) with the ethnicity of the young people in the ward (district).[13]

Consider Greens and Purples in ward 1. Their expected utility for *HA* can be written as:

$$EU_{g1i}(HA) = \frac{n_{g1} + \tau_{HA} n_{p_1}}{n_{g1} + n_{p1}} + \alpha_{HAi}, \tag{1}$$

$$EU_{p1i}(HA) = \frac{n_{p1} + \tau_{HA} n_{g_1}}{n_{g1} + n_{p1}} + \alpha_{HAi}, \tag{2}$$

n_{g1} and n_{p1} are the numbers of, respectively, Greens and Purples in ward 1. τ_{HA} is a parameter $\in [0, 1]$ indicating the level of tolerance for the other ethnicity when young people do *HA*. If τ_{HA} is 0, young people only derive utility from participants of their own ethnicity; if τ_{HA} is 1, young people's utility does not depend on the ethnicity of the participants. Finally, α_{HAi} is an idiosyncratic parameter, capturing individual preferences and costs related to *HA*. $\alpha_{HAi} \in [-\bar{\alpha}_{HAi}, \bar{\alpha}_{HAi}]$, and $\bar{\alpha}_{HAi}$ is such that at least one person is always willing to do *HA*. Defining $n_{g2} = N_g - n_{g1}$ and $n_{p2} = N_p - n_{p1}$, the expected utility of *HA* for Greens and Purples in ward 2 can be written in terms of the number of Greens and Purples in the district (N_g and N_p, respectively) and the number of Greens and Purples in ward 1 (n_{g1} and n_{p1}, respectively):

$$EU_{g2i}(HA) = \frac{N_g - n_{g1} + \tau_{HA}(N_p - n_{p1})}{N_g - n_{g1} + N_p - n_{p1}} + \alpha_{HAi}, \tag{3}$$

$$EU_{p2i}(HA) = \frac{N_p - n_{p1} + \tau_{HA}(N_g - n_{g1})}{N_g - n_{g1} + N_p - n_{p1}} + \alpha_{HAi}, \tag{4}$$

The expected utility of *PA* depends on the ethnicity of the young people participating in *PA*, approximated by the ethnicity of the young people in the district. Therefore, it is the same for young people from the same ethnic group, but living in different wards. The expected utility of *PA*, for Greens and Purples, can be written:

$$EU_{g1i}(PA) = EU_{g2i}(PA) = \frac{N_g + \tau_{PA} N_p}{N_g + N_p} + \alpha_{PAi}, \tag{5}$$

[13] We believe this assumption is not unreasonable in the case of young people, as strategic behavior fully develops in adulthood. In this respect, Czermak *et al.* (2016) find that only 40% of adolescents aged 10–17 years are strategic thinkers.

$$EU_{p1i}(PA) = EU_{p2i}(PA) = \frac{N_p + \tau_{PA} N_g}{N_g + N_p} + \alpha_{PAi}, \tag{6}$$

where τ_{PA} is a parameter $\in [0, 1]$ indicating the level of tolerance for the other ethnicity when young people do *PA*. τ_{PA} can be greater, equal or smaller than τ_{HA}. $\alpha_{PAi} \in [-\bar{\alpha}_{PAi}, \bar{\alpha}_{PAi}]$ is an idiosyncratic parameter, capturing individual preferences and costs related to *PA*, and such that at least one person is always willing to do *PA*.

To guide the interpretation of our empirical results, we study how the incentives of doing *HA* and *PA* change when: (*i*) district ethnic segregation changes with district ethnic diversity constant; (*ii*) district ethnic diversity changes with district ethnic segregation constant.[14] We distinguish the case where *HA* and *PA* are substitutes (i.e., young people are willing and are allowed to switch from activities) from the case where *HA* and *PA* are not substitutes. If *HA* and *PA* are not substitutes, changes in the utility of *HA* (*PA*) can only affect the incentives of doing *HA* (*PA*). If *HA* and *PA* are substitutes, changes in the utility of *HA* (*PA*) may change the relative incentives of doing *HA* and *PA*, and thus induce young people to switch between activities.

Let us study the effect of changes in district ethnic segregation. We introduce the *perfect integration* scenario where: (*i*) district ethnic segregation is zero: the district and both its wards have the same share of Greens and thus the same ethnic diversity ($F_d = F_{w1} = F_{w2} = \bar{F}$); (*ii*) Greens and Purples are equal in number in the district ($N_g = N_p$). (*i*) and (*ii*) imply (*iii*) Purples in ward 1 and Purples in ward 2 are equal in number ($n_{p1} = n_{p2}$), and Greens in ward 1 and Greens in ward 2 are equal in number ($n_{g1} = n_{g2}$). Condition (*iii*) implies (3) and (4) become, respectively:

$$EU_{g2i}(HA) = \frac{(N_g - n_{g1}) + \tau_{HA} n_{p1}}{N_g - n_{g1} + n_{p1}} + \alpha_{HAi}, \tag{7}$$

$$EU_{p2i}(HA) = \frac{n_{p1} + \tau_{HA}(N_g - n_{g1})}{N_g - n_{g1} + n_{p1}} + \alpha_{HAi}, \tag{8}$$

PROPOSITION 1. *a) Starting from the perfect integration scenario, district ethnic segregation can be increased by moving the Greens from ward 2 to ward 1 (without moving the Purples), keeping the number of Greens in the district (N_g) constant.*

b) If HA and PA are not substitutes, an increase in district ethnic segregation keeping district ethnic diversity constant increases the incentives of doing HA and leaves the incentives of doing PA unaffected.

c) if HA and PA are substitutes, an increase in district ethnic segregation keeping district ethnic diversity constant translates into a further increase in the incentives of doing HA at the expense of PA.

Proofs of propositions 1.a, 1.b and 1.c are shown in Appendix C. Intuitively, an increase in district ethnic segregation affects the ethnic mix of the wards and thus the utility of doing *HA*: the incentives of doing *HA* increase. An increase in district ethnic segregation does not change the ethnic mix of the district and thus it does not affect the utility of doing *PA*: the incentives of doing *PA* remain unaffected. If *HA* and *PA* are substitutes, the majority of young people may have incentives to switch from *PA* to *HA* (the incentives of doing *HA* increase more than the incentives

[14] To capture incentives we use marginal utilities. This permits to abstract from the distributions of α_{HAi} and α_{PAi}.

of doing *PA*). This can make *HA* increase at the expense of *PA*. Therefore, we expect an increase in district ethnic segregation increases *HA*. The effect on *PA* is null or negative, depending on the degree of substitutability between *HA* and *PA*.

To study the effects of changes in district ethnic diversity, we introduce the *maximum diversity* scenario. In the *maximum diversity* scenario (*i*) district ethnic diversity is at the maximum achievable level in our case, i.e., $F_d = \frac{1}{2}$ (Greens and Purples are equal in numbers in the district); (*ii*) district ethnic segregation is at a generic positive level $\bar{D}$; (*iii*) Purples in ward 1 and Purples in ward 2 are equal in number ($n_{p1} = n_{p2}$).[15]

PROPOSITION 2. *a) Starting from the maximum diversity scenario, district ethnic diversity can be decreased by increasing the number of Greens in the district (N_g) keeping Purples (N_p) constant, and allocating Greens into wards such that district ethnic segregation remains constant.*

b) If HA and PA are not substitutes, a decrease in district ethnic diversity keeping district ethnic segregation constant increases the incentives of doing both HA and PA. The incentives of doing PA increase more than the incentives of doing HA if $\Delta\tau = \tau_{PA} - \tau_{HA} \leq 0$. The incentives of doing HA increase more than the incentives of doing PA if $\Delta\tau = \tau_{PA} - \tau_{HA} > \bar{\Delta}\tau(\tau_{HA}, \tau_{PA}) > 0$.

c) If HA and PA are substitutes, a decrease in district ethnic diversity keeping district ethnic segregation constant translates into a further increase of PA at the expense of HA if $\Delta\tau = \tau_{PA} - \tau_{HA} \leq 0$; it translates into a further increase of HA at the expense of PA if $\Delta\tau = \tau_{PA} - \tau_{HA} > \bar{\Delta}\tau(\tau_{HA}, \tau_{PA}) > 0$.

Proofs of propositions 2.a, 2.b and 2.c are shown in Appendix C. Intuitively, a decrease in ethnic diversity increases both the incentives of doing *HA* and the incentives of doing *PA*. If τ_{PA} is ‘large enough’ compared with τ_{HA}, the incentives of doing *HA* increases more than the incentives of doing *PA*.[16] This can happen if, when participating in *HA*, young people care about the ethnicity of those they hang around with, when participating in *PA*, they care mainly about the purpose of the activity (helping the community, keeping fit). If *HA* and *PA* are substitutes, the majority of the young people may want to switch from the activity whose incentives increase less to the activity whose incentives increase more.

The threshold $\bar{\Delta}\tau$ depends on the relative magnitude of τ_{PA} and τ_{HA}. We present some simulations for Greens in ward 1, the largest group. We consider a baseline maximum diversity scenario with 20 Greens and 20 Purples, for different levels of district ethnic segregation (0.1, 0.2, 0.3, 0.4, 0.5). Figure 3 shows the difference in the marginal expected utility (with respect to N_g) of *HA* and *PA* for Greens in ward 1. In the top panel, $\Delta\tau = 0$. The incentives of doing *PA* increase more than the incentives of doing *HA* for any district ethnic segregation and N_g. In the middle panel, $\bar{\Delta}_{g1}\tau > \Delta\tau > 0$ ($\tau_{PA} > \tau_{HA}$, and τ_{PA} and τ_{HA} are similar in magnitude). Whether the incentives of doing *HA* increase more or less than the incentives of doing *PA* depends on the level of district ethnic segregation and N_g. In the bottom panel, $\Delta\tau > \bar{\Delta}_{g1}\tau > 0$ ($\tau_{PA} > \tau_{HA}$, and τ_{PA} is substantially larger than τ_{HA}). The incentives of doing *HA* increase more than the incentives of doing *PA* for any district ethnic segregation and N_g.[17]

[15] The new benchmark has been chosen because it is compatible with many levels of segregation. In contrast, a benchmark with zero district ethnic diversity has, by definition, also zero district ethnic segregation. Like in the previous case, when $n_{p1} = n_{p2}$, the maximum achievable level of segregation is $\frac{1}{2}$.

[16] In what follows, we refer to this situation as the case where τ_{PA} is ‘substantially larger’ than τ_{HA}.

[17] Note that $\frac{\partial EU_{g1}(HA)}{\partial N_g} - \frac{\partial EU_{g1}(PA)}{\partial N_g} > 0$ ($\forall D$, N_g) if $\frac{1}{2}(\tau_{HA} + 1) < \tau_{PA}$. Consider Figure 3. In the top panel: $\tau_{HA} = \tau_{PA} = 0.5$ and thus $\Delta\tau = 0$. In the middle panel: $\tau_{HA} = 0.4$ and $\tau_{PA} = 0.6$, thus $\Delta\tau = 0.2$. When $\tau_{HA} = 0.4$,

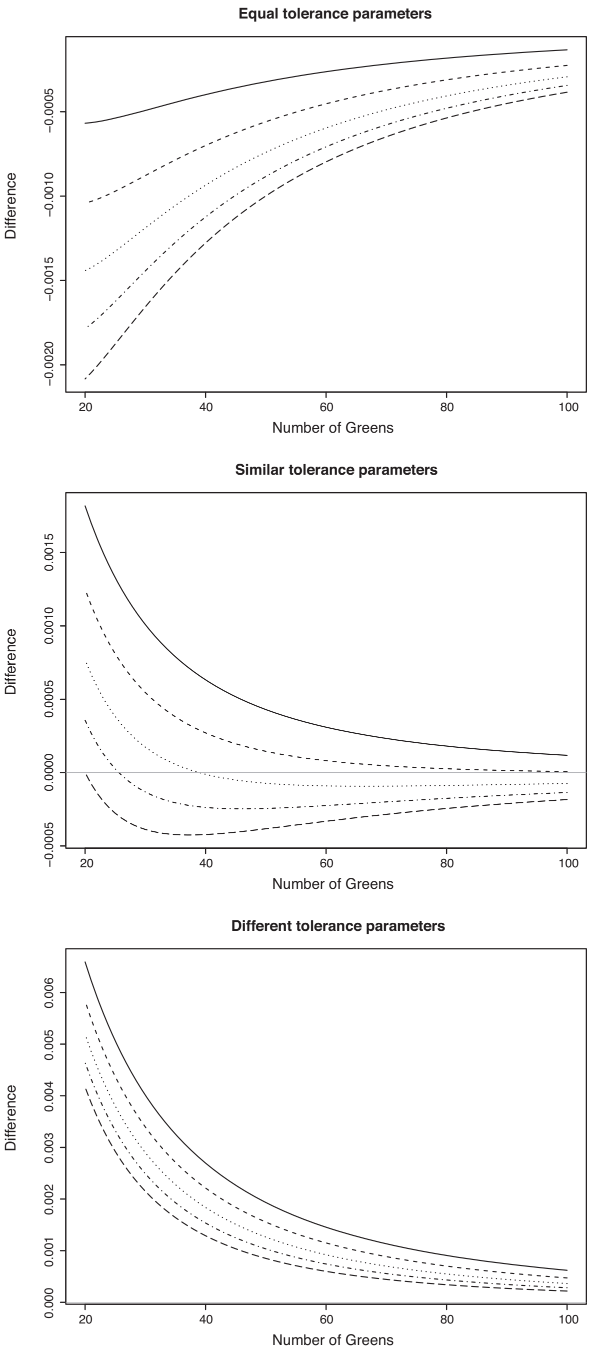

Fig. 3. *Simulation of* $\frac{\partial EU_{g1}(HA)}{\partial N_g} - \frac{\partial EU_{g1}(PA)}{\partial N_g}$

Notes: Each panel plots the value of the difference $\frac{\partial EU_{g1}(HA)}{\partial N_g} - \frac{\partial EU_{g1}(PA)}{\partial N_g}$ when $\bar{D} = 0.1$ (solid), $\bar{D} = 0.2$ (dashed), $\bar{D} = 0.3$ (dotted), $\bar{D} = 0.4$ (dot-dash), and $\bar{D} = 0.1$ (long dash). Parameters: $n_{p1} = 10$, $\tau_{HA} = \tau_{PA} = 0.5$ (top panel), $\tau_{HA} = 0.4$ and $\tau_{PA} = 0.6$ (middle panel), $\tau_{HA} = 0.2$ and $\tau_{PA} = 0.8$ (bottom panel).

3. Empirical Analysis

3.1. *Empirical Model*

We model *HA* and *PA* as linear functions of district ethnic diversity and segregation as follows:

$$PA_{ij} = \alpha_1 + \beta_1 F_j + \beta_2 D_j + \mathbf{x}'_{ij}\gamma_1 + \xi_{1ij}, \tag{9}$$

$$HA_{ij} = \alpha_2 + \beta_3 F_j + \beta_4 D_j + \mathbf{x}'_{ij}\gamma_2 + \xi_{2ij}, \tag{10}$$

F_j and D_j are, respectively, the fractionalisation and the segregation index for the district $j =$ *1,..., m* where person $i =$ *1,...,n* lives. Thus, β_1, β_2, β_3 and β_4 are the parameters of interest. Finally, $\mathbf{x}_{ij}$ is a vector of controls (individual and family characteristics, plus weather indicators, presented in Subsection 3.2) and ξ_{1ij} and ξ_{2ij} are composite errors, decomposed as:

$$\xi_{1ij} = \zeta_{1j} + \eta_{1i} + \epsilon_{1ij},$$

$$\xi_{2ij} = \zeta_{2j} + \eta_{2i} + \epsilon_{2ij},$$

ζ_{1j} and ζ_{2j} are unobserved effects at the district level, η_{1i} and η_{2i} are unobserved effects at the individual level, and ϵ_{1ij} and ϵ_{2ij} are idiosyncratic components.

Equations (9) and (10) are first estimated via weighted linear probability models (LPM), with standard errors clustered at the district level.[18] In a second approach (IV), ethnic diversity and segregation are instrumented, and equations (9) and (10) are estimated via weighted two stages least squares (2SLS), general method of moments (GMM), and limited information maximum likelihood (LIML), with errors clustered at the district level.[19]

3.2. *Endogeneity of District Ethnic Diversity and Segregation*

Identification of causal effects of district characteristics can be problematic. In our case, there are two types of endogeneity. Endogeneity of type I can arise if people select into districts on the basis of some unobserved individual characteristics η_{1i} and η_{2i} correlated with F_j and D_j (e.g., preferences for local public goods, as in Tiebout, 1956). This makes it difficult to distinguish between the effect of the district as such and the sum of the individual effects of its inhabitants.

$\frac{\partial EU_{g1}(HA)}{\partial N_g} - \frac{\partial EU_{g1}(PA)}{\partial N_g} > 0$ $(\forall D, N_g)$ if $\tau_{PA} > 0.7$, which implies $\bar{\Delta}_{g1}\tau = 0.3$, and thus $\Delta\tau < \bar{\Delta}_{g1}\tau$. In the bottom panel: $\tau_{HA} = 0.2$ and $\tau_{PA} = 0.8$, thus $\Delta\tau = 0.6$. When $\tau_{HA} = 0.2$, $\frac{\partial EU_{g1}(HA)}{\partial N_g} - \frac{\partial EU_{g1}(PA)}{\partial N_g} > 0$ $(\forall D, N_g)$ if $\tau_{PA} > 0.6$ which implies $\bar{\Delta}_{g1}\tau = 0.4$, and thus $\Delta\tau > \bar{\Delta}_{g1}\tau$.

[18] Weights are sampling weights. We use LPM, as linear specifications are very tractable and permit a direct comparison with the IV results. The LPM was also suitable for our case as all the predicted probabilities estimated for *HA* through the LPM fell in the [0,1] range, while only one probability for *PA* fell outside that range. Note that equations (9) and (10) are estimated individually, although participating in *PA* and *HA* is likely to be a joint decision, and thus ϵ_{1ij} and ϵ_{2ij} are likely to be correlated. Failing to account for this correlation may lead to a loss of efficiency. We chose not to estimate the equations jointly for comparability with the IV approach. Joint and non-linear estimation of (9) and (10) is presented as a robustness check in Section 5.

[19] The 2SLS, GMM and LIML estimators have different strengths and weaknesses. The GMM estimator is more efficient than 2SLS in large samples with heteroscedasticity of the error term caused by clustering. However, GMM can be undesirable in small samples where precise estimates of higher moments are difficult (Baum and Schaffer, 2003). LIML are generally preferred in presence of many (potentially week) instruments and in a small sample, where 2SLS can be seriously biased, while the LIML is median-unbiased.

Endogeneity of type II can arise when unobserved district characteristics are correlated with both district ethnic diversity/segregation and social participation, that is when ζ_{1j} and ζ_{2j} are correlated with F_j and D_j.

Endogeneity of type I can arise when people select into districts for reasons correlated with both districts' ethnic diversity/segregation and districts' opportunities of socialisation. This requires, for example, that antisocial people disproportionally select into ethnically diverse districts (Putnam, 2007). Endogeneity of type I would also occur if poorer and lower educated people can only afford districts with worse provision of local public goods (e.g., sport facilities, parks).[20] If those districts are also the most diverse/segregated, the correlation between district ethnic diversity/segregation and social participation could be spurious.

Endogeneity of type II can arise when the central government disproportionally invests in purposeful activities in ethnically diverse or segregated districts to engender community cohesiveness.[21] This problem is unlikely to exist in our case, as recreational activities are provided by districts. Indeed, for the case of sport, evidence has claimed that investment in sport undertaken by the central government in the years we consider 'was not primarily aimed at getting the local community involved in sport' (Gratton *et al.*, 2005).

Our empirical strategy limits both types of endogeneity. Endogeneity of type I is limited as young people are less mobile than adults and are not directly involved in parental location decisions. Therefore, households' residential choices are not directly correlated with children's taste for social participation and depend mainly on parental characteristics we can control for. This reduces potential correlations between η_{1i} and η_{2i} and the indices of ethnic diversity and segregation. Endogeneity of type II is limited as district ethnic diversity and segregation are measured using 2001 census data (i.e., three years before 2004, when our data on social participation were collected). This reduces potential correlations between F_j and D_j and the error term due to districts-specific shocks that occurred in 2004.

To address the problem of endogeneity of type I, we include in $\mathbf{x}_{ij}$ a set of proxies for preferences and constraints affecting social participation, potentially correlated with F_j and D_j through endogenous residential choices. At the individual level we control for gender and ethnicity: two important predictors of preferences for and constraints to social participation (Antunes and Gaitz, 1975; Platt, 2009). We also control for whether the respondent was born in 1989 and whether English is not her mother-tongue language. The former captures differences in age and anything else affecting people born in different calendar years, the latter captures ability and willingness to interact with natives.[22] At the household level we control for mother's characteristics (education, employment status and age) and whether the main parent is the mother.[23] These variables capture heterogeneity in parental decisions about young people's leisure time allocation. We also control

[20] Appendix A shows parks and green areas are where young people hang around the most.

[21] The economic literature generally does not see differences in investment in local public goods by the local government as an endogeneity problem. It claims ethnic diversity and segregation determine both social capital and provision of local public goods (Alesina *et al.*, 1999; La Ferrara and Mele, 2007). In our case, this implies β_1, β_2, β_3 and β_4 pick up the total effect of district ethnic diversity and segregation on social participation, including supply side effects through investment in local public goods. For a discussion of the supply side effects, see Section 7.

[22] Most LSYPE respondents were born in 1989 or 1990. Nineteen people were born before or after these dates. They are excluded from the sample.

[23] Mothers' characteristics are captured through a set of dummy variables. Education variables are: no education, GCSE and below, above GCSE, degree, missing information on education. Employment status variables are: mother working full-time, part-time, not working, missing information on employment status. Age variables are: below 35, 35/50, above 50, missing age. In LSYPE the main parent is the parent most involved in the young person's education, who provides household level information. Fathers' characteristics are strongly correlated with mothers' and their measurement is poor due to high non-response. Therefore, they are not included in the analysis.

for household income (in quintiles) to capture availability of resources. Finally, we control for the millimetres of rain and the hours of sun in the month of interview. Our controls may not be enough to avoid any potential source of endogeneity, especially endogeneity of type II. Type IV tackles this.

3.3. *Instrumenting Ethnic Diversity and Segregation*

3.3.1. *The instruments*

We propose two new sets of instruments motivated by historical research. The first set uses measures of proximity to the ports listed in the Aliens Act: a 1905 piece of legislation restricting the number of ports through which individuals could legally enter Britain (see Pellew, 1989). The second set of instruments is based on measures of proximity to the so-called mill towns.[24] At the end of the Second World War, mill-town jobs were becoming unattractive for native people (Simpson, 2004). At the same time, constraints on migration were loosened, creating an exogenous supply of workers for jobs in the textile industry (Kalra, 2000).These new workers, who worked night hours and weekends and spoke very little English, never mixed with the English majority (Cantle, 2001; Simpson, 2004; Swanton, 2010).

These two historical events had different effects on district ethnic composition. Migrants located by the ports were likely to be seamen, inclined to temporary migration and used to travelling and interacting with other ethnicities (Tabili, 1994). Migrants to mill towns were a fairly homogeneous group, spoke poor English, and had relative long-term perspectives of staying in England. This led to high diversity and low segregation in ports and high segregation in mill towns. Subsections 3.3.2 and 3.3.3 discuss instruments' exogeneity and relevance.

3.3.2. *Exogeneity*

Distance from the main ports of entry has been used to instrument cultural diversity in the USA (Ottaviano and Peri, 2006), following the argument that a port's location is exogenously determined by land morphology. Our strategy rests on a similar assumption and has the further advantage that ports are chosen based on a historical legislative act. This avoids arbitrary ports selection by the researcher and simultaneous determination of port status (determined in 1905) and social participation (determined in 2004). To be exogenous, Aliens Act port status should be uncorrelated with district level unobserved effects potentially correlated with social participation. In other words, cities chosen as ports of entry in 1905 should not be substantially different from other areas in the provision of social activities, or, if such difference existed, it should not persist until 2004. This is difficult to test, as, to the authors' knowledge, there are no data on social activities for the early twentieth century.

To proxy for the provision of social activity we use data on population growth from the Urban Population database 1801–1911 (Bennett, 2012).[25] We want to explore if before the introduction of the Alien Act future ports of entry were already growing faster (in terms of population, and thus development) than other areas. The idea is that high development could be associated with

[24] These ports are: Cardiff, Dover, Folkestone, Grangemouth, Grimsby, Harwich, Hull, Leith, Liverpool, London, Newhaven, Southampton, the Tyne Ports and Plymouth. These mill towns are Manchester, Bolton, Leigh, Ashton under Lyne, Warrington, Wigan, St Helens, Blackburn, Chorley, Preston, Lancaster, Ramsbottom, Rochdale, Burnley, Accrington, Colne, Bury, Oldham.

[25] The rationale is that population growth is correlated with economic development (Easterlin, 1967; Ashraf and Galor, 2011, 2013) and thus potentially to social activities. We are aware this is a coarse proxy determined by the scarcity of disaggregated historical data.

high provision of social activities. Figure B1 in Appendix B plots the logarithm of population against time. It shows that, although the Aliens Act ports of entry were on average bigger than other towns or cities, the population growth rate in the nineteenth century (the slope of the line) for the two groups is almost identical. This suggests ports listed in the Alien Act were not very different from other areas of Britain in terms of population growth, and possibly also in terms of unobservables.[26]

Proximity to mill towns could not be considered an exogenous instrument if the area where the first mill towns were established was more advanced than other areas in terms of a number of aspects potentially endogenous to social participation, such as technical or scientific knowledge. This is unlikely for three reasons. First, the spin-off of the industrial revolution was the presence, in particular areas across Britain, of a strong watch industry (Allen, 2009) that had only a mediated effect on the British industrial development. The watch industry increased the availability of cheap, good quality gears (and skilled workers capable of assembling them), which facilitated the introduction of crucial 'macro inventions' (Mokyr, 1990; Allen, 2009).[27] Second, the key inventors of the cotton revolution were generally artisans, rather than leading scientists.[28] Third, similar technologies to those incorporated in the 'macro inventions' were already in use in other areas of Britain (Allen, 2009).[29]

3.3.3. *Relevance*

Relevant instruments must predict district ethnic diversity (segregation) in 2004. This condition is satisfied if: (*i*) the Aliens Act affected the ethnic composition of the areas by the ports (the decline of the textile industry and the contemporary loosening of migration constraints affected the ethnic composition of the areas by the mill towns); (*ii*) the ethnic composition of Britain resulting from (*i*) persisted until 2004.

We claim that condition (*i*) is satisfied for the case of ports of entry, as the Aliens Act was successfully enforced and thus changed the ethnic composition of the areas by the ports (Pellew, 1989; Tabili, 1994; Little, 2013). Condition (*i*) is satisfied for the case of mill towns, as the spread of the cotton industry outside its original location was promptly inhibited by licensing policies requiring would-be cotton spinning firms to bear the burden of assembling water frames. This favoured producers located in proximity to the watch industry (Allen, 2009). Condition (*ii*) is satisfied as the migration legislation introduced in the late twentieth century hindered temporary migration and forced new migrants to move where pioneers had already settled down (Holmes, 2001).[30] This created a strong correlation between location decisions of new and old migrants,

[26] Note that differences in population growth between Aliens Act ports and other cities would not invalidate our identification strategy, however, similarities in population growth could suggest similarity in other aspects, including development, and thus provision of social activities.

[27] The 'Spinning Jenny' (1764), the water frame (1769), the carding machine (1775) and the mule (1779).

[28] Other areas in Britain had stronger links with the scientific community, like London, home of the the Royal Society, Birmingham, hosting the 'lunar society', and Edinburgh, with a world leading school of chemistry. Allen (2009) shows that, among all the key sectors of the British industrial revolution, the textile industry was the one with the weakest link with the scientific community.

[29] Spinning machines have been documented in the early 1700s in Norwich, spinning and carding machine in Birmingham between late 1730 and late 1750, and carding machines in 1748 in Herefordshire (Allen, 2009).

[30] For example, in 1962, the 'Commonwealth Immigrants Act' introduced an entry system for Commonwealth workers based on vouchers. This system was followed in 1968 by a quota system and by further restrictions based on evidence of partiality in 1971.

Table 2. *IV1, Instruments Relevance and Exogeneity.*

Test	Test statistics	*p*-value
Underidentification		
Kleibergen-Paap Ch-sq	41.396	0.001
Sanderson-Windmeijer Ch-sq (Diversity)	377.005	0.000
Sanderson-Windmeijer Ch-sq (Segregation)	2,492.583	0.000
F tests		
Kleibergen-Paap	28.507	
Sanderson-Windmeijer (Diversity)	22.034	
Sanderson-Windmeijer (Segregation)	145.682	
Overidentification (2SLS)		
Sargan-Hansen (PA)	20.658	0.192
Sargan-Hansen (HA)	13.607	0.628

Notes: Tests and are computed using the ivreg2 stata command by Baum *et al.* (2002).

with the effect that England's ethnic composition has remained stable throughout the twentieth century (Pellew, 1989; Holmes, 2001).[31]

To create our instruments, we compute the inverse of the geodesic distance between each district in England and the closest Aliens Act port/mill town.[32] These measures, together with the variables in $\mathbf{x}_{ij}$, are used to predict district ethnic diversity and segregation. Results are shown in Table B2 in Appendix B. Proximity to ports of entry is positively (negatively) and significantly correlated with the diversity (segregation) index. Proximity to mill towns is positively and significantly correlated with the segregation index only.

For our type IV analysis we allow for a number of non-linearities between proximity to ports/mill towns and district ethnic composition, adding higher orders of the measures of proximity up to cubic. Further, to capture the effect of ports/mills other than the closest, we compute the number of ports/mills within a radius of 20, 50 and 100 km, plus the interactions with the measures of proximity (without higher terms). We call this type IV approach 'IV1'.[33]

Tests of instruments' relevance are presented in Table 2 (top two panels). They all suggest that the instruments are relevant. The first panel presents underidentification tests, leading to the conclusion that the model is not underidentified.[34] The second panel presents the Kleibergen-Paap (KP) Kleibergen and Paap (2006); Kleibergen and Schaffer (2007) and the Sanderson-Windmeijer (SW) (Sanderson and Windmeijer, 2016) F-tests on the instruments in the first stages.[35] All F-

[31] Holmes (2001) documents that the main ethnic groups living in Britain at the beginning of the twentieth century were the main ethnic groups we observe in Table A10 (African, Caribbean, Indian, Bangladeshi, Pakistani and Chinese, with the Chinese community being the smallest in size).

[32] The shortest path between two points a, b of coordinates (lat(a), long(a)) and (lat(b), long(b)) on the earth's surface can be computed using the great circle distance, corresponding to the length of a straight line in an Euclidean space. The great circle distance is obtained by multiplying the central angle between the two points (in radians) by the radius of the sphere, i.e., $dist_{ab} = \arccos(\sin(lat(a)) * \sin(lat(b)) + \cos(lat(a)) * \cos(lat(b)) * \cos(|\ (long(b) - long(a))\ |)) * 6371$, where 6371 is the radius of the earth in kilometres.

[33] Slightly different IV approaches (IV2, IV3 and IV4) are presented as robustness checks.

[34] The Kleibergen-Paap (KP) underidentification test tests if the matrix of the reduced form coefficients is full rank; that is if the correlations between the instruments and the endogenous variables are different from zero (null hypothesis: the model is underidentified). The Sanderson-Windmeijer (SW) underidentification tests (Sanderson and Windmeijer, 2016) test underidentification regressor by regressor.

[35] The SW F-test is considered the most appropriate in the case of multiple endogenous variables. We also computed the Angrist and Pischke (AP) F-statistics (Angrist and Pischke, 2009). With two endogenous variables they converge to the SW statistics, and thus results are not reported. In addition, the comparison between partial R^2 of excluded

Table 3. *Linear Probability Model, Coefficients of Interest.*

	(1)		(2)		(3)		(4)	
	PA	*HA*	*PA*	*HA*	*PA*	*HA*	*PA*	*HA*
Diversity	− 0.146***	− 0.316***	− 0.123***	− 0.380***	− 0.066*	− 0.204***	− 0.065*	− 0.185***
	(0.033)	(0.049)	(0.037)	(0.040)	(0.037)	(0.046)	(0.036)	(0.046)
Segregation			− 0.079	0.216***	− 0.070	0.242***	− 0.001	0.199***
			(0.051)	(0.054)	(0.049)	(0.050)	(0.048)	(0.049)
Individual controls	No	No	No	No	Yes	Yes	Yes	Yes
Household controls	No	No	No	No	No	No	Yes	Yes
Weather controls	No	No	No	No	No	No	Yes	Yes
N	14,244	14,244	14,244	14,244	14,244	14,244	14,244	14,244

Notes: Standard errors in parentheses, *p < 0.1, **p < 0.05, ***p < 0.01. Individual controls: male, born in 1989, ethnicity dummies. Family controls: income quintiles, English notmother tongue, main parent is female, mother's education, age and employment status. Weather controls: total sunshine duration, total rainfall. Each LSYPE respondent is assigned the weather information collected from the closest meteo station in the interview month. Weighted data.

tests are greater than 10, providing additional evidence in favour of instruments' relevance. Instruments' exogeneity is difficult to test. In case of over-identified models, most scholars refer to the Sargan-Hansen test to derive an indication of instruments' exogeneity. Sargan-Hansen tests for both *PA* and *HA* (see bottom panel of Table 2) fail to reject the null of orthogonality of the instruments.

4. Results

Table 3 shows the coefficients for district ethnic diversity and segregation estimated via LPM. Four specifications are presented: with district ethnic diversity only (specification 1), with both district ethnic diversity and segregation (specification 2), with individual controls (specification 3), with individual, household and weather controls (specification 4). Specification 4 is our main specification.[36]

The estimated coefficient for district ethnic diversity on *HA* is negative and strongly statistically significant in all specifications. The estimated coefficient for district ethnic diversity on *PA* is negative in specifications (1) and (2), but goes towards zero when controls are added in specifications (3) and (4). This lack of effect of district ethnic diversity on *PA* is compatible with our explanatory model in a case where τ_{PA} is substantially larger than τ_{HA}. If *HA* is motivated by a pure desire of socialisation, while *PA* is also motivated by other goals (keeping fit, helping the community), young people may care more (less) about other people's ethnicity when doing *HA* (*PA*). The estimated coefficient for district ethnic segregation on *HA* is positive and statistically significant in all specifications. The estimated coefficient for district ethnic segregation on *PA* is negative but statistically insignificant in all specifications. These results are compatible with our explanatory model.

To have an idea of the magnitude of the effects, consider specification (4). Going from the median (0.094) to the highest observed ethnic diversity (0.676) leads to a decrease of around 11 percentage points (p.p.) in the probability of participating in *HA*. This negative effect is

instruments and Shea's R^2 suggests that both instruments are relevant (R^2(diversity) = 0.371, Shea's R^2(diversity) = 0.432, R^2(segregation) = 0.248, Shea's R^2(segregation) = 0.289).

[36] Estimated coefficients for the control variables are shown in Table B3. They generally suggest young people from richer and more educated households are more likely to choose *PA*, while young people from more deprived background, but also young people with working mothers, are more likely to choose *HA*.

Table 4. *IV1: Second Stages, Coefficients of Interest.*

	PA				HA			
	LPM	*2SLS*	*GMM*	*LIML*	*LPM*	*2SLS*	*GMM*	*LIML*
Diversity	– 0.065*	– 0.042	– 0.034	– 0.041	– 0.185***	– 0.228***	– 0.226***	– 0.228***
	(0.036)	(0.064)	(0.060)	(0.064)	(0.046)	(0.077)	(0.061)	(0.077)
Segregation	– 0.001	0.182*	0.132	0.184*	0.199***	0.508***	0.516***	0.511***
	(0.048)	(0.094)	(0.086)	(0.095)	(0.049)	(0.098)	(0.032)	(0.099)
N	14,244	14,244	14,244	14,244	14,244	14,244	14,244	14,244

Notes: Standard errors in parentheses, *p < 0.1, **p < 0.05, ***p < 0.01. IV results are computed using the ivreg2 stata command by Baum *et al.* (2002). Individual controls: male, born in 1989, ethnicity dummies. Family controls: income quintiles, English not mother tongue, main parent is female, mother's education, age and employment status. Weather controls: total sunshine duration, total rainfall. Each LSYPE respondent is assigned the weather information collected from the closest meteo station in the interview month. Instruments: proximity to the closest port of entry and its higher terms (up to the third), number of ports in a radius of 20, 50, 100 km, and interactions between the proximity to the closest port and the number of ports within 20, 50, 100 km. Weighted data.

comparable to the one associated with the dummy variable indicating respondents whose mother tongue is not English (see Table B3 in Appendix B). Similarly, going from the lowest observed ethnic diversity (0.02) to the median leads to a decrease in *HA* of just above 1 p.p. Finally, going from the lowest observed ethnic segregation (0.058) to the median (0.160) leads to an increase of around 2 p.p. in the probability of *HA*, while going from the median segregation to the highest observed value (0.620) increases this probability by 9 p.p. This latter effect is larger than the effect of having a mother with no education, as opposed to having a mother with a university degree (see again Table B3).

Table 4 shows the estimated coefficients for ethnic diversity and segregation on *PA* and *HA* when endogeneity is accounted for. They are similar to those obtained via LPM.[37] The results are robust across estimation methods (2SLS, GMM, LIML). The 2SLS and the LIML estimates are very similar, suggesting our instruments are not weak (Angrist and Pischke, 2009). The estimated type IV coefficients for district ethnic diversity and segregation are generally statistically insignificant for *PA* and strongly statistically significant for *HA*. For *HA*, the coefficients for district ethnic diversity become more negative than the one estimated via LPM, with increases in absolute value of up to 23%. The coefficients for district ethnic segregation become more positive and more than double in magnitude. For example, 2SLS and LIML suggest that passing from the median to the highest observed district ethnic diversity (segregation) leads to a 13 p.p. decrease (23–24 p.p. increase) in *HA*.

The lack of correlation between *PA* and district ethnic diversity is coherent with our explanatory model when τ_{PA} is substantially larger than τ_{HA}. Moreover, it is also in line with the critique that part of the negative effects of neighbourhood ethnic diversity on social participation is due to omitted neighbourhood deprivation (see Letki, 2008). This critique is valid in our case if: (*i*) ethnically diverse districts are the most deprived; and (*ii*) district deprivation is associated with poor access to *PA*. Indeed, in the specifications without individual, household and weather controls (Table 3, specifications (1) and (2)), ethnic diversity is found to be negatively correlated with *PA*, but such correlation disappears when the endogeneity problem is addressed first by adding controls and then through the type IV approach.

For *HA*, a similar argument explains why controlling for endogeneity makes the coefficient of ethnic diversity even more negative. Suppose that district ethnic diversity leads to *less HA*.

[37] The estimated coefficients for the first stages are shown in Tables B4 and B5 in Appendix B.

Table 5. *CMP: Marginal Effects at the Mean.*

	PA				HA			
	Probit	*CMP1*	CMP2	CMP3	*Probit*	*CMP1*	*CMP2*	*CMP3*
Diversity	– 0.069*	– 0.047	– 0.065	– 0.060	– 0.189***	– 0.277***	– 0.231***	– 0.232***
	(0.037)	(0.070)	(0.057)	(0.060)	(0.047)	(0.084)	(0.071)	(0.072)
Segregation	– 0.004	0.195*	0.183*	0.113	0.205***	0.524***	0.512***	0.461***
	(0.050)	(0.114)	(0.110)	(0.108)	(0.052)	(0.111)	(0.107)	(0.110)
N	14,244	14,244	14,244	14,244	14,244	14,244	14,244	14,244

Notes: Standard errors in parentheses, *p < 0.1, **p < 0.05, ***p < 0.01. Results are computed using the cmp stata command by Roodman (2011). Individual controls: male, born in 1989, ethnicity dummies. Family controls: income quintiles, English not mother tongue, main parent is female, mother's education, age and employment status. Weather controls: total sunshine duration, total rainfall. Each LSYPE respondent is assigned the weather information collected from the closest meteo station in the interview month. Instruments: proximity to the closest port of entry and its higher terms (up to the third), number of ports in a radius of 20, 50, 100 km, and interactions between the proximity to the closest port and the number of ports within 20, 50, 100 km. Weighted data.

Suppose also that ethnic diversity is positively correlated with deprivation, and deprivation leads to more hanging around. Thus, deprivation implies young people in more ethnically diverse districts do *more HA*. This positive bias would mask the 'real' negative effect of district ethnic diversity on *HA*. Thus, the type IV coefficient of district ethnic diversity on *HA* is likely to be more negative than the LPM one.

There is no literature on the effects of ethnic segregation on the demand of social activities. On the supply side, La Ferrara and Mele (2007) suggest that district ethnic segregation leads to higher provision of local public goods. If local public goods include parks, security, sport facilities and clubs, this may facilitate participation in both *PA* and *HA*. If less social families self-select into more segregated districts, this causal effect may show up only when endogeneity is controlled for. Thus, the type IV coefficient of district ethnic segregation on *HA* and *PA* is likely to be more positive than the LPM one.

5. Robustness Checks

We present a number of robustness checks. First, we test the robustness of the results to the functional form used in the estimation, by estimating the model in (9) and (10) non-linearly. We first estimate the model by probit and then we use a conditional mixed process (CMP), together with IV1 instruments, to account for endogeneity of district ethnic composition.[38] We estimate three different specifications of the CMP. Each specification consists in four jointly estimated equations. Two equations, equivalent to the first stages in the fully linear case, estimate district ethnic diversity and segregation. Two equations, estimated by probit, estimate the probability of taking part in *PA* and *HA*. The exogenous variables included in the first stages are: CMP1, IV1 instruments and variables $\mathbf{x}_{ij}$ in (9) and (10); CMP2, IV1 instruments without variables $\mathbf{x}_{ij}$ in (9) and (10); CMP3, IV1 instruments on ports of entry to predict ethnic diversity and IV1 instruments on mill towns to predict ethnic segregation.

Marginal effects at the mean (in Table 5) are consistent across specifications and similar to those obtained through linear models. The probit estimation suggests district ethnic diversity has a small (marginally significant) negative effect on *PA*, which disappears once endogeneity is accounted for. The effect of district ethnic segregation on *PA* is not significant when estimated

[38] Estimation is via limited-information maximum likelihood (stata command cmp, Roodman, 2011).

Table 6. *RS with IV1 Instruments: LPM and Second Stages (Selected Coefficients).*

	PA				HA			
	LPM	*2SLS*	*GMM*	*LIML*	*LPM*	*2SLS*	*GMM*	*LIML*
Diversity	− 0.001	0.019	0.009	0.019	− 0.287***	− 0.460***	− 0.406***	− 0.462***
	(0.081)	(0.135)	(0.104)	(0.136)	(0.075)	(0.130)	(0.106)	(0.131)
Segregation	0.029	0.270*	0.178	0.273*	0.193***	0.447**	0.456***	0.452**
	(0.085)	(0.156)	(0.142)	(0.157)	(0.074)	(0.181)	(0.114)	(0.185)
Population	− 0.000	− 0.001*	− 0.001	− 0.001*	0.000	− 0.001	− 0.001	− 0.001**
	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Number of	0.025***	− 0.025***	0.025***	0.025***	− 0.013*	− 0.012*	− 0.014**	− 0.012*
vehicles	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)	(0.007)
Unemployed	− 1.855	− 0.894	− 1.141	− 0.883	0.364	1.946	0.613	1.976
	(1.997)	(2.265)	(1.793)	(2.271)	(1.699)	(2.025)	(1.586)	(2.036)
No	0.167	0.156	0.028	0.155	− 0.438	− 0.729**	− 0.374	− 0.733**
qualification	(0.242)	(0.324)	(0.288)	(0.325)	(0.281)	(0.335)	(0.279)	(0.337)
Household	− 0.013	− 0.100	− 0.040	− 0.101	0.183*	0.164	0.106	0.163
size	(0.073)	(0.098)	(0.089)	(0.098)	(0.098)	(0.122)	(0.110)	(0.123)
Good	− 0.103	− 0.068	− 0.224	− 0.067	− 0.730	− 0.544	− 0.069	− 0.541
health	(0.420)	(0.462)	(0.418)	(0.462)	(0.520)	(0.557)	(0.516)	(0.558)
Conception	− 0.001	− 0.001*	− 0.002**	− 0.001*	0.001	0.001	0.000	0.001
rate (U18)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
All road	51.490*	50.552*	43.716*	50.549*	32.710	39.386	24.715	39.471
casualties	(27.361)	(28.127)	(26.129)	(28.144)	(33.749)	(34.124)	(30.920)	(34.158)
IDACI	0.250	0.151	0.225	0.150	0.032	0.093	0.189	0.093
	(0.188)	(0.211)	(0.189)	(0.212)	(0.198)	(0.230)	(0.175)	(0.231)
N	12,042	12,042	12,042	12,042	12,042	12,042	12,042	12,042

Notes: Standard errors in parentheses, $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. IV results are computed using the ivreg2 stata command by Baum *et al.* (2002). Individual controls: male, born in 1989, ethnicity dummies. Family controls: income quintiles, English not mother tongue, main parent is female, mother's education, age and employment status. Weather controls: total sunshine duration, total rainfall. District level controls: population of the district (divided by 10,000), number of vehicles in household, percentage of unemployed people (16–74 years old), percentage of people without any educational qualification (16–74 years old), average household size, percentage of people in good health, under-18 conception rate (multiplied by 1,000), number of road casualties (divided by population of the district), income deprivation affecting children index (IDACI). Each LSYPE respondent is assigned the weather information collected from the closest meteo station in the interview month. Instruments IV1: proximity to the closest port of entry and its higher terms (up to the third), number of ports in a radius of 20, 50, 100 km, and interactions between the proximity to the closest port and the number of ports within 20, 50, 100 km. Weighted data.

via probit and becomes slightly significant and positive in some specifications accounting for endogeneity of neighbourhood ethnic composition. Results for *HA* confirm a strong negative effect of district ethnic diversity and a strong positive effect of district ethnic segregation.

Second, we test the robustness of the results to small changes in the construction of the instruments. We present three different specifications of the instruments (IV2, IV3 and IV4) in addition to the IV1 specification used in the body of the article. The functional form and the set of controls used remain the same as those used in IV1. IV2 instruments district ethnic diversity and segregation with measures of proximity to both the first and second closest port/mill town, their squares, the number of towns within 20, 50 and 100 km, and the interactions between these variables and the proximity from the first and second port/mill town. IV3 instruments district ethnic diversity and segregation with measures of proximity to all the ports and mill towns. IV4 instruments ethnic diversity and segregation with a series of binary variables capturing the number of ports (mill towns) within a 20, 50 or 100 km radius. Table B6 shows the tests on the instruments are fully satisfactory for all strategies. Second stages, shown in Table B7, are in line with those from our main specification. In conclusion, these robustness checks show the results are robust to modest changes in how instruments are constructed.

Table 7. *RS with IV1 Instruments and County Fixed Effects: LPM and Second Stages (Selected Coefficients).*

	PA				HA			
	LPM	*2SLS*	*GMM*	*LIML*	*LPM*	*2SLS*	*GMM*	*LIML*
Diversity	− 0.020	− 0.159	− 0.179*	− 0.160	− 0.204***	− 0.412***	− 0.398***	− 0.414***
	(0.084)	(0.135)	(0.107)	(0.135)	(0.071)	(0.127)	(0.110)	(0.127)
Segregation	0.023	0.308*	0.255*	0.311*	0.193***	0.450***	0.535***	0.454***
	(0.096)	(0.163)	(0.146)	(0.165)	(0.074)	(0.161)	(0.131)	(0.163)
Population	− 0.000	− 0.001	− 0.001	− 0.001	0.001	− 0.000	− 0.000	− 0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Number of	0.024***	0.024***	0.024***	0.025***	− 0.013*	− 0.012*	− 0.013**	− 0.012*
vehicles	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)	(0.007)
Unemployed	− 1.906	0.059	0.268	0.078	0.290	2.319	1.592	2.345
	(2.380)	(2.899)	(2.018)	(2.908)	(1.570)	(2.133)	(1.745)	(2.143)
No	− 0.197	− 0.522	− 0.644*	− 0.524	− 0.351	− 0.770*	− 0.560*	− 0.774*
qualification	(0.309)	(0.408)	(0.330)	(0.409)	(0.322)	(0.401)	(0.337)	(0.402)
Household	0.059	0.046	0.104	0.046	0.099	0.125	0.077	0.125
size	(0.088)	(0.108)	(0.093)	(0.109)	(0.097)	(0.114)	(0.095)	(0.114)
Good	− 0.856	− 0.777	− 0.949*	− 0.776	− 1.031*	− 0.887	− 0.443	− 0.886
health	(0.557)	(0.612)	(0.548)	(0.613)	(0.618)	(0.655)	(0.568)	(0.656)
Conception	− 0.000	− 0.000	− 0.000	− 0.000	− 0.000	− 0.000	− 0.000	− 0.000
rate (U18)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
All road	63.470*	68.255**	62.865*	68.278**	1.313	9.908	7.484	9.964
casualties	(32.929)	(34.190)	(32.212)	(34.216)	(37.005)	(38.338)	(32.866)	(38.371)
IDACI	0.275	0.285	0.311	0.285	0.087	0.163	0.227	0.163
	(0.212)	(0.260)	(0.199)	(0.260)	(0.188)	(0.214)	(0.168)	(0.215)
N	12,042	12,042	12,042	12,042	12,042	12,042	12,042	12,042

Notes: Standard errors in parentheses, *p < 0.1, **p < 0.05, ***p < 0.01. IV results are computed using the ivreg2 stata command by Baum *et al.* (2002). Individual controls: male, born in 1989, ethnicity dummies. Family controls: income quintiles, English not mother tongue, main parent is female, mother's education, age and employment status. Weather controls: total sunshine duration, total rainfall. District level controls: population of the district (divided by 10000), number of vehicles in household, percentage of unemployed people (16–74 years old), percentage of people without any educational qualification (16–74 years old), average household size, percentage of people in good health, under-18 conception rate (multiplied by 1,000), number of road casualties (divided by population of the district), income deprivation affecting children index (IDACI). Each LSYPE respondent is assigned the weather information collected from the closest meteo station in the interview month. Instruments IV1: proximity to the closest port of entry and its higher terms (up to the third), number of ports in a radius of 20, 50, 100 km, and interactions between the proximity to the closest port and the number of ports within 20, 50, 100 km. Weighted data.

Third, we test the robustness of our results to the inclusion of additional district level characteristics that are likely to affect *HA* and *PA*. Such characteristics are: population of the district (divided by 10,000), average number of vehicles in households, unemployment rate (for people 16–74 years old), percentage of people with no educational qualification (16–74 years old), average household size, percentage of people in good health, under 18 conception rate (multiplied by 1,000), number of all road casualties (divided by population of the district), income deprivation affecting children index (IDACI).[39]

We estimate this richer specification (RS) first via linear probability models and then by instrumenting district ethnic diversity and segregation with IV1 instruments. In a further sets of specifications, we also include county fixed effects.[40] RS rests on different assumptions than the main specification of the article. Controlling for a larger number of district level characteristics

[39] Data on district population, average number of cars, unemployment, health, education, and household size are from the 2001 Population Census data, while data on road casualties and under 18 conception rate are from the 'neighbourhood statistics' collected by the ONS. IDACI is computed at the Super Output Area and can be obtained with the LSYPE. Data are aggregated by district using sample weights.

[40] County is the administrative level higher than the district. We included one dummy for each county in England and a dummy indicating 'unitary authorities': a special type of districts (generally bigger cities) which do not belong to any county.

Table 8. *Heteroscedasticity Constructed Instruments (2SLS): Second Stages (Coefficients of Interest).*

	PA	HA	PA	HA	PA	HA
Diversity	0.042	– 0.301***	0.108	– 0.269***	– 0.029	– 0.229***
	(0.085)	(0.094)	(0.110)	(0.097)	(0.107)	(0.082)
Segregation	– 0.054	0.401***	– 0.075	0.275***	0.038	0.266***
	(0.102)	(0.103)	(0.119)	(0.100)	(0.120)	(0.085)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes
District level controls	No	No	Yes	Yes	Yes	Yes
County FE	No	No	No	No	Yes	Yes
N	14,244	14,244	12,042	12,042	12,042	12,042

Notes: Standard errors in parentheses, *p < 0.1, **p < 0.05, ***p < 0.01. IV results are computed using the ivreg2h stata command by Baum and Schaffer (2012). Individual controls: male, born in 1989, ethnicity dummies. Family controls: income quintiles, English not mother tongue, main parent is female, mother's education, age and employment status. Weather controls: total sunshine duration, total rainfall. Each LSYPE respondent is assigned the weather information collected from the closest meteo station in the interview month. District level controls: population of the district (divided by 10,000), number of vehicles in household, percentage of unemployed people (16–74 years old), percentage of people without any educational qualification (16–74 years old), average household size, percentage of people in good health, under-18 conception rate (multiplied by 1,000), number of road casualties (divided by population of the district), income deprivation affecting children index (IDACI). Weighted data.

may reduce the district level endogeneity left in the error term. However, while district ethnic diversity and segregation are allowed to be endogenous (and are instrumented), the other district level characteristics are assumed exogenous.

Instrument diagnostics for RS models with IV1 are shown in Tables B8 and B9. The instruments perform slightly worse than in our main specification. However, the tests are overall satisfactory. Results from the RS specifications (both LPM and IV) are shown in Tables 6 and 7. They confirm there is no effect of district ethnic composition on *PA* and there is a negative (positive) effect of district ethnic diversity (segregation) on *HA*. These results are in line with those from our main specification, in spite of the differences in the identifying assumptions.

Fourth, we test the robustness of our results to the use of a completely different type IV strategy. We use the instrumental variable approach introduced by Lewbel (2012) and we apply it both to the specification without district level controls and to the richer specification with district level controls, with and without county fixed effects. Lewbel's approach derives identification from the standard assumption of exogeneity of the controls and the additional assumption of heteroscedasticity of the error term. It can be used when valid exclusion restrictions are not available. Results are shown in Table 8: they are in line with those obtained with the other identification strategies.[41]

The results in this section and in Section 4 come from models relying on a very diversified set of identification assumptions.[42] In spite of these differences, they all tell the same story. All models suggest that district ethnic diversity (segregation) has a negative (positive) impact on *HA*. These effects become even larger when endogeneity is accounted for, although most robustness checks suggest the impact of district ethnic segregation on *HA* is smaller than the one implied by the type IV results in Table 4. Where endogeneity is controlled for, we are generally unable

[41] Tests of relevance and exogeneity of the instruments as well as tests on the heteroscedasticity of the error term are shown in Tables B10 and B11.

[42] For another paper comparing results from a geographically inspired IV strategy and result from Lewbel's IV strategy, see Emran and Hou (2013).

to find negative effects of district ethnic diversity on *PA*. This is coherent with our explanatory model where τ_{PA} is substantially larger than τ_{HA}. We also find some—rather limited—evidence that there may be a positive effect of district ethnic segregation on *PA*. In Section 7 we provide evidence suggesting that this result is likely to be driven by supply side effects.

6. Demand Side: More on Mechanisms

Our explanatory model focuses on the demand side of social participation. It emphasises the role of geography: *HA* takes place in the ward and *PA* takes place in the district. Moreover, the model acknowledges that *HA* and *PA* may be different in nature. In particular, they may differ in the importance young people attach to the ethnicity of those participating in the activity (captured by the tolerance parameters τ_{HA} and τ_{PA}). In this section we argue that differences in the tolerance parameters τ_{HA} and τ_{PA} capture a meaningful difference in the nature of *PA* and *HA*, and that geography is important in explaining our results.

In Section 4 we claim that the relative magnitude of τ_{HA} and τ_{PA} may explain why district ethnic diversity affects participation in *HA* more than participation in *PA*. Moreover, we claim that geography explains why district ethnic segregation affects participation in *HA* more than participation in *PA*. Is it possible that the different effect of district ethnic segregation on *HA* and *PA* is not driven by geography and, instead, is driven by the nature of *PA* and *HA*, in a way that is not captured by the tolerance parameters τ_{HA} and τ_{PA}? In what follows, we compare two activities similar in nature to claim that it is unlikely that our results on the effect of district ethnic segregation are driven by a difference in the nature of *HA* and *PA* not captured by our model, and that geography is likely to play an important role.

We consider a third activity: 'hanging around near the city centre' (*HC*). *HC* is important for a number of reasons. First, it is popular among young people: around 30% of our sample reports having done *HC* in the four weeks before the interview. Second, *HC* is similar in nature to *HA*, as it is likely to be motivated by the desire to spend time with friends with no specific goal. Third, due to this similarity, *HC* is likely to display a high degree of substitutability with *HA*. Finally, like *PA*, *HC* does not necessarily take place in the ward. We estimate the effect of district ethnic diversity and segregation on *HC*. We are particularly interested in the coefficient for district ethnic segregation. If the results for district ethnic segregation are only driven by the nature of the activity, the effects of district ethnic segregation on *HC* should resemble the one found for *HA*. If the geographical scope of the activity also matters, the effect should get closer to the one found for *PA*.

Results (in Table 9) show district ethnic segregation has a strongly significant positive effect on *HA* and a negative effect on *HC* (significant at the 10% level). As previously discussed, district ethnic segregation increases the incentives to participate in *HA* (which takes place in the ward), but not the incentives to participate in activities such as *PA* or *HC*, which may take place elsewhere in the district. If *HA* and *HC* display high substitutability, an increase in *HA* can lead to a decrease in *HC*. In segregated districts, where wards are on average more homogeneous than the district, people may prefer to hang around near home rather than in the city centre. This suggests that the nature of the activity is not the only driver of the results, and that geography plays a role in explaining the impact of district ethnic composition on social participation.[43]

[43] The estimated coefficients of interests obtained instrumenting district ethnic diversity and segregation using IV1 are shown in Table B12. They confirm a statistically significant negative effect of district ethnic segregation on *HC*

Table 9. *Hanging Around Near the City Centre: LPM (Coefficients of Interest).*

	PA	HA	*HC*
Diversity	– 0.065*	– 0.185***	– 0.004
	(0.036)	(0.046)	(0.046)
Segregation	– 0.001	0.199***	– 0.118*
	(0.048)	(0.061)	(0.049)
Individual controls	Yes	Yes	Yes
Household controls	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes
N	14,244	14,244	14,244

Notes: Standard errors in parentheses, *p < 0.1, **p < 0.05, ***p < 0.01. Individual controls: male, born in 1989, ethnicity dummies. Family controls: income quintiles, English not mother tongue, main parent is female, mother's education, age and employment status. Weather controls: total sunshine duration, total rainfall. Each LSYPE respondent is assigned the weather information collected from the closest meteo station in the interview month. Weighted data.

To investigate the role of geography further, we estimate an augmented model by adding to our main LPM specification the fractionalisation index at the ward level. The coefficient for ward ethnic diversity should be negative for *HA* (*HA* is discouraged by diverse wards). Ward ethnic diversity should be irrelevant for *PA*: its coefficient should be close to zero.[44] This augmented model also helps decompose the effect of district ethnic segregation. The coefficient for ward ethnic diversity captures the effect of district ethnic segregation due to ward ethnic diversity. Due to the inclusion of ward ethnic diversity, the coefficient for district ethnic segregation in the augmented model measures the effect of district ethnic segregation due to the set of available ethnic mix of wards within the district (see La Ferrara and Mele, 2007).

Results are presented in the top panel of Table 10 (last two columns). *HA* is negatively correlated with ward ethnic diversity, but not with district ethnic diversity. This suggests *HA* is likely to take place in the ward, and thus ward ethnic diversity is what primarily affects the choice of engaging in *HA*. The inclusion of ward ethnic diversity leads to a reduction in the coefficient for district ethnic segregation if compared with the one in the second column. This is what we expect, as the coefficient for district ethnic segregation in the second column captures both the effect of ward ethnic diversity and the effect of the available ethnic mix of wards within the district. The estimated coefficient for district ethnic segregation in the augmented model in Table 10 remains positive and statistically significant at the 1% level. In sum, results in Table 10 suggest geography plays an important role in explaining participation in *HA* and *PA*.

Ward ethnic diversity may be endogenous to social participation due to endogenous sorting of households into wards. To address this potential problem, we restrict the sample to respondents living in social housing (see Algan *et al.*, 2016; Bonomi Bezzo, 2017).[45] In 2004 social houses were allocated by districts through a waiting list system, and households had very limited

and a negative—but in most cases not statistically significant—effect of district ethnic diversity. Tests on instruments relevance, in table B13, show the instruments are valid. Note that the substitutability between *HA* and *HC* may also explain the coefficient for district ethnic diversity on *HC*. According to our explanatory model, social activities (including *HC*) decrease when district ethnic diversity increases. *HC* and *HA* are similar in nature, and thus it is also likely that $\tau_{HC} \approx \tau_{HA}$. However, for some $\tau_{HC} > \tau_{HA}$, an increase in ethnic diversity decreases the incentives of doing *HA* more than the incentives of doing *HC*. For this reason, young people may switch from *HA* to *HC*. This may mitigate the negative effect of district ethnic diversity on *HC*.

44 Reality may partially depart from our simplified framework. For example, when *PA* are provided in multiple wards within the districts, young people may choose the activity taking place closer to their house. Therefore, the coefficient for ward ethnic diversity may become slightly negative also in the case of *PA*.

45 We selected respondents who rented either from the council or from a housing association.

Table 10. *Augmented Model with Ward Ethnic Diversity: LPM (Coefficients of Interest).*

	PA	HA	PA	HA
Full estimation sample				
District ethnic diversity	– 0.065*	– 0.185***	– 0.025	0.053
	(0.036)	(0.046)	(0.058)	(0.067)
District ethnic segregation	– 0.001	0.199***	– 0.040	0.176***
	(0.048)	(0.049)	(0.048)	(0.117)
Ward ethnic diversity			– 0.039	– 0.212*
			(0.044)	(0.117)
Individual controls	Yes	Yes	Yes	Yes
Household controls	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
N	14,244	14,244	14,244	14,244
Social housing sample				
District ethnic diversity	0.076	– 0.248***	0.135	– 0.072
	(0.065)	(0.057)	(0.102)	(0.099)
District ethnic segregation	– 0.018	0.123*	– 0.017	0.127*
	(0.099)	(0.071)	(0.098)	(0.070)
Ward ethnic diversity			– 0.061	– 0.181**
			(0.083)	(0.089)
Individual controls	Yes	Yes	Yes	Yes
Household controls	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
N	3,543	3,543	3,543	3,543

Notes: Standard errors in parentheses, *p < 0.1, **p < 0.05, ***p < 0.01. Individual controls: male, born in 1989, ethnicity dummies. Family controls: income quintiles, English not mother tongue, main parent is female, mother's education, age and employment status. Weather controls: total sunshine duration, total rainfall. Each LSYPE respondent is assigned the weather information collected from the closest meteo station in the interview month. Weighted data.

discretionality in the choice of where to live.[46] This strongly limits the scope of endogenous sorting into wards. Results for the 'social housing' subsample are presented in the bottom panel of Table 10 (last two columns) and are in line with those for the full estimation sample. This suggests that the results are not driven by endogeneity.[47]

[46] The allocation of social housing was regulated by the Housing Act 1996. Houses were allocated on a first-come-first-served basis. Priority could be given to subjects who were homeless, lived in unhealthy or overcrowded accommodations, were ill or disabled. Preferences for location were taken into account only when a failure to relocate would have caused hardship for the subject or for others. Houses could also be allocated to avoid the concentration of residents with similar socio-economic characteristics, such that the greatest possible social mix could be achieved. With the introduction of Choice Based Letting (CBL) schemes, the Homelessness Act 2002 made it possible for the households to bid for specific properties, although the success of the bid was not guaranteed. In 2002, only 15% of the districts participated in CBL. This share was just above 20% in 2004 (DCLG, 2010). In our case, CBL schemes could only affect those who applied for social housing between late 2002 and early 2004 in the restricted group of districts participating in the scheme. Therefore, CBL schemes are unlikely to have substantially affected the randomness of the residential allocation of our 'social housing' subsample.

[47] The results in the first and the second column of the bottom panel of Table 10 are derived from estimating our main specification on the social housing subsample. District ethnic diversity and segregation are likely to be less endogenous for this subsample than for the full estimation sample. This is because, in 2004, households who wanted to relocate were pushed to the back of the housing waiting list in the new district. This discouraged endogenous sorting into districts. Therefore, these results, in line with those obtained in Sections 4 and 5, can be seen as a further robustness check for our main specification.

7. Supply Side: the Provision of Social Activities

Our explanatory model abstracts from the supply side effects of district ethnic composition, i.e., how district ethnic composition can affect social participation through the provision of local public goods. In this section we discuss these supply side effects and we provide evidence that district ethnic segregation may incentivise *PA* through increased or better targeted provision of local sport facilities.

The literature suggests that district ethnic diversity discourages the provision of local public goods, while district ethnic segregation encourages it (Alesina *et al.*, 1999; La Ferrara and Mele, 2007). If local public goods include parks, sport facilities and clubs, district ethnic diversity (segregation) may decrease (increase) *HA* and *PA*. Our empirical model captures the total effect of district ethnic composition, i.e., the combination of the effects stemming from the demand and the supply side. We do not have an identification strategy to separate the demand and the supply effect. However, it is still interesting to discuss the correlation between district ethnic diversity and segregation and the provision of local public goods.

We are not aware of the existence of comprehensive data on provision of local public goods in England. However, we have information on the satisfaction with local sport facilities. As sport is an important component of *PA*, this could shed light on how district ethnic diversity and segregation correlate with the supply of *PA*. We use the 2005/2006 wave of the Active People Survey (APS): the closest in time to the LSYPE data used in our main analysis. APS contains information on around 360,000 respondents, including an indication of the district respondents live in and a question on satisfaction with local sport facilities.

We measure satisfaction with local sport facilities through an indicator equal to one if the respondent is fairly/very satisfied with local sport provision.[48] We collapse this variable at the district level and we regress it on the indices of district ethnic diversity and segregation plus the district level controls used to estimate the RS model in Table 6. Table 11 shows satisfaction with sport facilities is negatively correlated with district ethnic diversity and positively correlated with district ethnic segregation.

We are more confident about making causal statements about the results on district ethnic segregation in Table 11 than about the results on district ethnic diversity in the same table. Omitted deprivation positively correlated with district ethnic diversity and segregation and negatively correlated with the provision of sport facilities may lead to overestimating the negative effect of district ethnic diversity and underestimating the positive effect of district ethnic segregation. Therefore, we believe the results in Table 11 provide some evidence that district ethnic segregation may incentivise participation in *PA* through an increased or better targeted provision of local public goods.

8. Conclusions

This article studies the effect of district ethnic diversity and segregation on social participation. We focus on social participation for young people, as evidence suggests relational skills are formed during adolescence. We distinguish 'purposeful activities', which have been found to be beneficial for young people's development, from 'hanging around with friends', which we show to be associated with risky behaviours. We study the case of England, where improving the provision of purposeful activities has become central in the policy agenda (DfES, 2006).

[48] The other options are 'neither' and 'fairly/very dissatisfied' with local sport provision.

Table 11. *The Role of Supply: Satisfaction with the Provision of Local Sport Facilities (LPM).*

	Satisfaction
District ethnic diversity	– 0.109***
	(0.041)
District ethnic segregation	0.113***
	(0.034)
Population	– 0.000
	(0.000)
Number of vehicles	– 0.006
	(0.004)
Unemployed	– 1.141
	(0.765)
No qualification	0.185
	(0.141)
Household size	0.001
	(0.043)
Good health	0.220
	(0.248)
Under-18 conception rate	– 0.001*
	(0.000)
All road casualties	– 20.629
	(10.137)**
IDACI	0.136
	(0.107)
Constant	0.651***
	(0.165)
N	241

Notes: Standard errors in parentheses, *p $<$ 0.1, **p $<$ 0.05, ***p $<$ 0.01. Weighted data.

We account for the potential endogeneity of district ethnic diversity and segregation in a number of different ways, including by proposing a novel set of instruments. All the methods that we use lead to the same conclusions. District ethnic diversity does not discourage participation in purposeful activities, but it does discourage hanging around with friends. The former is in line with Letki (2008); the latter is in line with the literature finding negative effects of ethnic diversity on social interaction (e.g., Alesina and La Ferrara, 2000, 2002; Costa and Kahn, 2003; Pennant, 2005; Anderson and Paskeviciute, 2006; Charles and Kline, 2006; Putnam, 2007). District ethnic segregation increases hanging around with friends. Positive effects of district ethnic segregation on participation in purposeful activities are weak and are likely to be driven by supply side effects. To the authors' knowledge, these findings on the effect of district ethnic segregation on social participation are new.

Our findings have important policy implications. At least for the case of England, we show that ethnic diversity does not seem detrimental for young people's social participation. In fact, greater ethnic diversity could limit hanging around with friends and the potential risks associated with it. In contrast, ethnic segregation may be detrimental, as it is found to encourage hanging around. Thus, desegregation policies, in combination with policies improving the supply of purposeful activities, are more likely to have beneficial effects on young people's well-being than policies limiting migration to preserve ethnic homogeneity.

Our analysis focuses on the impact of ethnic diversity and segregation on social participation. Other forms of diversity and segregation, for example regarding income, social status, education, are likely to play a role in determining social participation. Our explanatory model can be adapted to analyse other forms of diversity and segregation. This is a fascinating exercise, which is left for future research.

Utrecht University
University of Essex

Additional Supporting Information may be found in the online version of this article:

Replication Package

Appendix A: Social Participation for Young People

A1. Introduction

This appendix describes young people's social participation. The first part describes the phenomenon of hanging around with friends in England. It answers four questions: (*i*) is hanging around near home pervasive among teenagers?; (*ii*) when do teenagers hang around?; (*iii*) is hanging around a problem?; (*iv*) what are the consequences of teenagers hanging around on adults' well-being? This analysis gives an idea of why hanging around has received so much attention in the policy debate and why a study of the phenomenon is needed. The second part investigates the short- and long-term consequences of hanging around and purposeful activities on teenagers' human capital accumulation, involvement in risky behaviours and social ties. We conclude social participation in teenage years has potentially long-lasting effects on crucial domains of people's lives. The third part presents evidence that young people prefer to interact with others from the same ethnic group. This is a crucial assumption in our explanatory model and thus it is important to show that it has some empirical support in our case.

A2. Hanging Around with Friends: A Description

This section describes what it means for teenagers to hang around with friends. We use the the British Crime Survey (BCS), which collects data on major crimes, as well as (perceived) antisocial behaviours, such as teenagers hanging around, noise and loud parties, vandalism and graffiti, drug dealing and using, drunkenness and alcohol-related violence. We use data from the 2004–2005 BCS adult main sample (over 45,000 respondents) and data on young people aged 14 to 15 from the 2009 under-16 sample (over 3,600 respondents). We restrict the analysis to young people aged 14 to 15: the age group we study in the body of the article.[49]

Hanging around near home is a pervasive phenomenon among young people. Figure A1.a shows that almost 60% of the 14/15-year-old 2009 BCS respondents hang around at least once a week, and more than one in four respondents (27.41%) hangs around more than three times a week. In contrast, only 1 in 10 respondents (10.52%) never hangs around. Young people tend to hang around locally. When asked to report where they hang around (see Figure A1.b), respondents mention local shops in almost 6% of the cases and streets outside home in almost 13%. Some of the green areas, mentioned in around 38% of the cases, are also likely to be local. In sum, Figure A1 suggests that hanging around near home is a common activity among British 14/15-year-olds.

Data on when teenagers hang around—from the 2004 BCS main sample of adult respondents—help to shed light on the type of activities displaced by hanging around. Hanging around does not

[49] The BCS has been collected biannually from 1982 and 2001 and annually afterward. For adults, we use 2004-2005 data as this is the time span covered by our main analysis. Data on under 16 are not available for 2004. Therefore, we use data from 2009, the closest available year. Data are weighted to account for sample design and non-response.

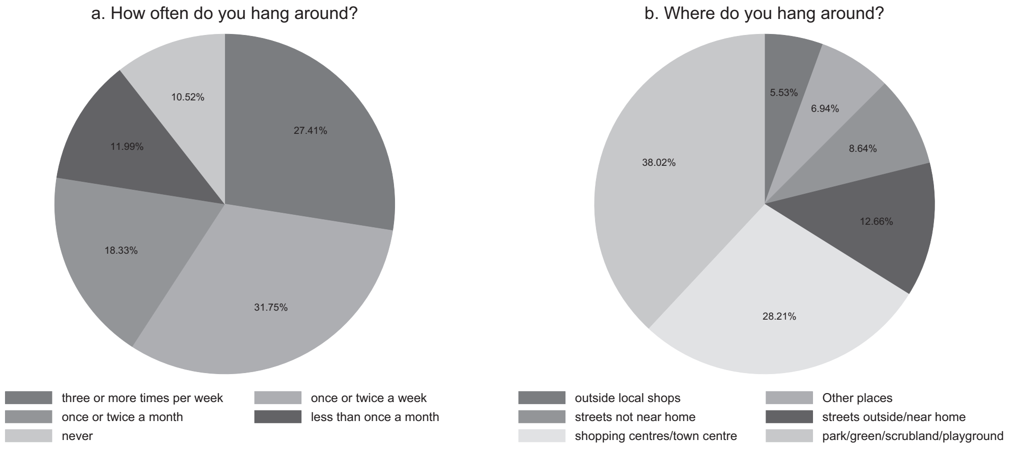

Fig. A1. *Is Hanging Around with Friends Near Home Pervasive?*
Notes: British Crime Survey, under-16 sample, 2009. Weighted data. Graph a: N = 478. Graph b: The question has not been asked to those who never hang around. N = 422. The category Other Places includes: river/lake: 0.13%; community/youth centre: 0.19%; subway and underpasses: 0.25%; outside school: 0.32%; at/outside cinemas: 0.32%; beach: 0.44%; outside pubs, clubs, bars: 0.51%; leisure centre/swimming pool/other sports: 0.95%; railway, underground, bus stations: 1.10%; home or friend's home: 1.08%; car parks 1.14%; somewhere else 0.82%.

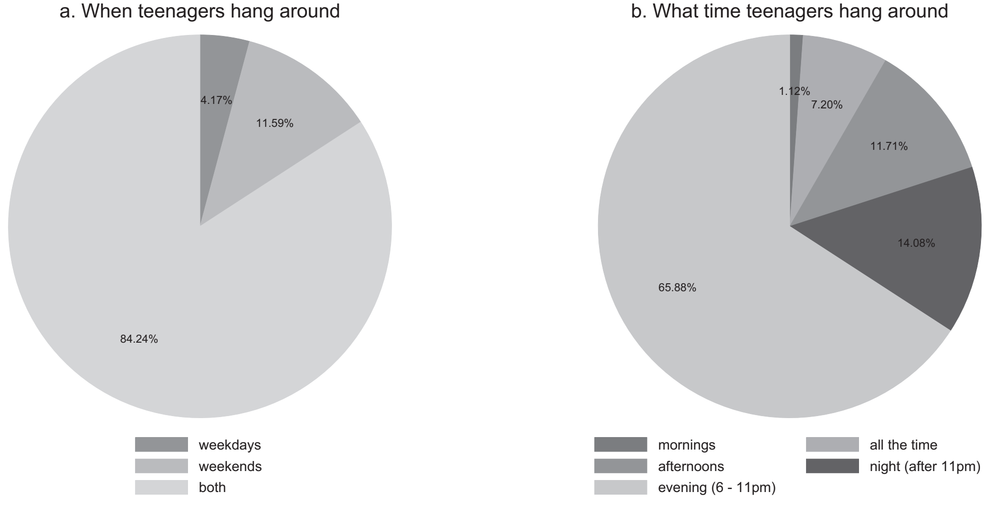

Fig. A2. *When Teenagers Hang Around.*
Notes: British Crime Survey, main adults sample, 2004/2005. Weighted data. Graph a: N = 3,526. Graph b: N = 3,529.

take place only at weekends, when teenagers have more free time. An overwhelming majority of BCS adult respondents (84.24%) report seeing teenagers hanging around during both weekends and weekdays (see Figure A2.a). Hanging around is mainly concentrated in the evening from 6 p.m. to 11 p.m. (see Figure A2.b). Hanging around in the morning is recorded in 1.12% of the

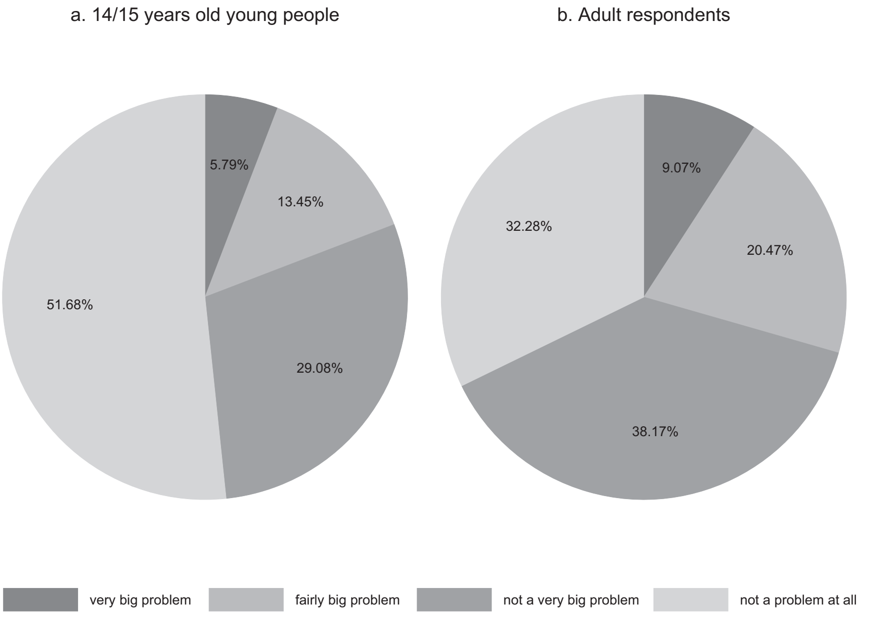

Fig. A3. *Problems in the Neighbourhood.*
Notes: British Crime Survey. Weighted data. Graph a: under-16 sample, 2009; weighted data; N = 475. Graph b: main adults sample, 2004/2005; weighted data; N = 42,442.

cases. This share reaches 8.32% if 'hanging around all the time' is added. Hanging around in the afternoon before 6 p.m. and in the night after 11 p.m. accounts for around 8% and 14.08% of the cases, respectively. In sum, the data suggest hanging around does not generally happen during school time (mornings and early afternoon), although the share of teenagers observed to hang around during school time is not negligible. Hanging around also takes place at night. This is potentially harmful for teenagers, as it interferes with their sleep.

Figure A3 shows teenagers hanging around is perceived as a pervasive problem. Figure A3.a shows that around 20% of the young people consider teenagers hanging around a 'very big' or a 'fairly big' problem. Among adults, this figure reaches 29.56% (see Figure A3.b). Problems of comparable magnitude to teenagers hanging around, as reported by adults, are: rubbish lying around (very or fairly big problem for 29.29% of the respondents), vandalism and graffiti (very or fairly big problem for 27.26% of the respondents), people using or dealing drugs (very or fairly big problem for 24.98% of the respondents). Less than 22% of the adults consider drunk people to be a very or fairly big problem. Abandoned and burnt-out cars and racially motivated attacks are very or fairly big problems for only 11.25% and 6.23% of the adults, respectively.[50] Similar conclusions can be derived from Figure A4. Adults consider teenagers' hanging around the main problem of their neighbourhood both when they report the three main problems (Figure A4.a) and when they select the main one (Figure A4.b).

[50] To ensure comparability, the figures on the perception of different antisocial behaviour are derived on the same subset of respondents with valid data for all antisocial behaviours. Note that data in figures A3 and A4 are likely to measure a combination of how worrisome and how prevalent a behaviour is perceived.

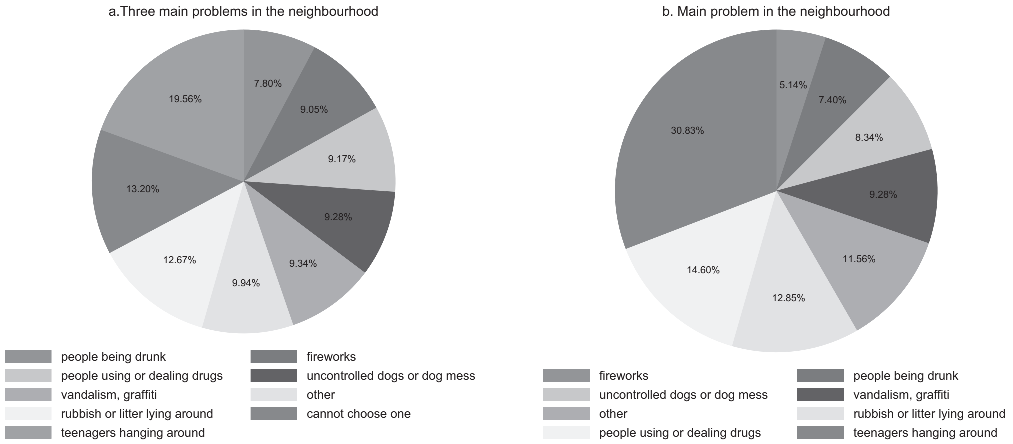

Fig. A4. *Problems in the Neighbourhood.*

Notes: British Crime Survey, main adults sample, 2004/2005. Weighted data. Graph a: N = 21,702. Graph b: N = 11,914. Graph a: other is constructed as follows: people being attacked due to ethnicity 0.78%; people begging: 1.06%; people being insulted: 2.13%; abandoned or burnt-out cars: 2.75%; noisy neighbours or loud parties 4.00%. Graph b: other is constructed as follows: people being attacked due to ethnicity 0.80%; people begging: 0.84%; people being insulted: 1.42%; abandoned or burnt out cars: 1.84%; noisy neighbours or loud parties 4.63%; cannot choose one 2.03%.

Figure A5 sheds light on why teenagers hanging around is perceived so negatively. Almost 65% of the adults consider teenagers hanging around on the streets to be deliberately antisocial. The antisocial behaviours most frequently reported are generally minor and include: swearing

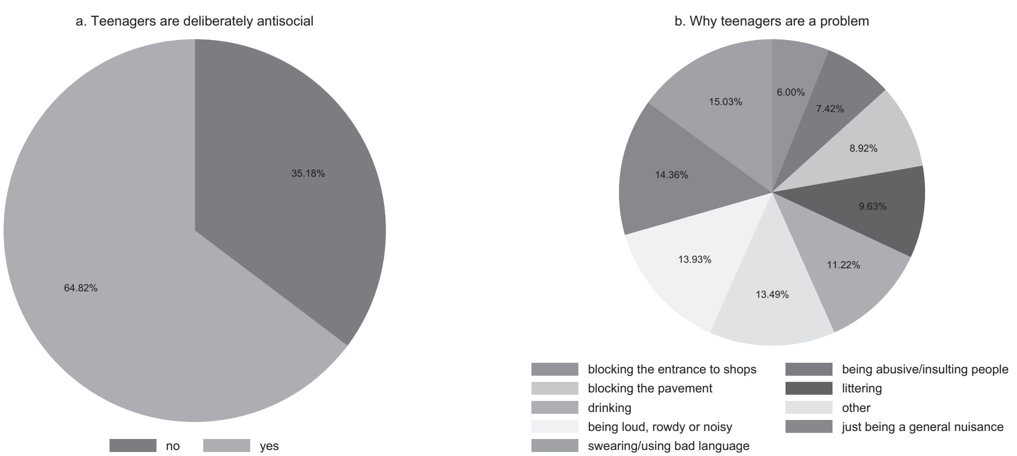

Fig. A5. *Why Teenagers are a Problem.*

Notes: British Crime Survey, main adults sample, 2004/2005. Weighted data. Graph a: N = 3,729; Graph b: N = 3,603. Other is constructed as follows: other: 0.15%; mugging or robbing people: 0.59%; physically assaulting people: 0.85%; not doing anything in particular: 0.90%; damaging property or cars: 3.73%; fighting with each other 3.94%; doing graffiti: 4.05%; taking drugs: 4.05%; intimidating or threatening people: 4.85%.

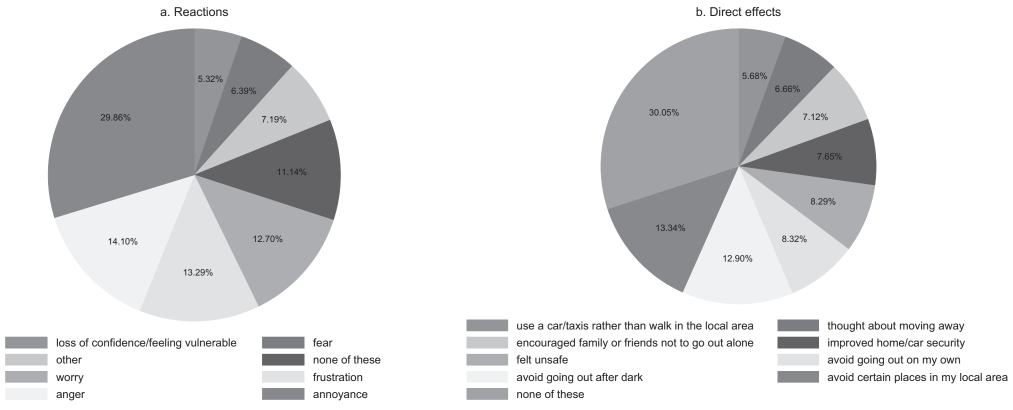

Fig. A6. *Reaction to HA and Action Taken.*
Notes: British Crime Survey, main adults sample, 2004/2005. Weighted data. Graph a: N = 3,792. Other is constructed as follows: other: 0.47%; crying/tears: 0.49%; depression: 1.08%; anxiety/panic attacks: 1.37%; shock: 1.51%; stress: 4.58%. Graph b: N = 3,792. Other is constructed as follows: avoid staying home: 0.40%; being assaulted: 0.48%; moved out of an area: 0.50%; carry a personal security device: 0.89%; not been able to sleep at times: 2.97%; not very trusting of people in the local area 4.85%; other: 0.47%.

(15.3%), generally being a nuisance (14.36%), and being loud and noisy (13.93%). Drinking is the most frequently mentioned risky behaviour (11.22%). Other risky or illegal behaviours such as mugging, robbing, intimidating, threatening or physically assaulting people, damaging property or cars, fighting, doing graffiti, or taking drugs are infrequent individually, but all together account for 13% of the cases (see Figure A5.b). These data suggest that hanging around, while not criminal *per se*, is associated with illegal and antisocial behaviours.[51]

Teenagers hanging around affect adults' well-being. Adults evaluated the effect on teenagers hanging around on the quality of their life using a 10-point scale, where 10 means 'very affected'. In 2004, the mean was 3.64 (for comparison, the means for vandalism and for people using or dealing drugs were respectively 3.34 and 3.10). The most common emotional response triggered by teenagers hanging around was annoyance (almost 30% of the cases, as shown in Figure A6). Many respondents also reported anger (14.10%), frustration (13.29%), worry (12.70%). More extreme expressions of distress are less frequent (crying: 0.49%; depression: 1.08%; anxiety/panic attacks: 1.37%; shock: 1.51%; stress: 4.58%), but all together account for more than 9% of the cases. Around 70% of the respondents reported direct consequences of the emotional distress caused by the teenagers hanging around. These include: avoiding certain places in the local area (13.34%), avoiding going out after dark (12.90%) or alone (8.32%), improving home or car security (7.65%), encouraging family or friends not to go out alone (7.12%), thinking about moving away (6.66%), using the car instead of walking (5.68%). Only 34% of the respondents made a complaint, but about one-third of them (11.13% of the total) complained directly to the police (see Figure A7.a). Interestingly, 16.7% of those who did not complain, failed to do it for

[51] The under 16 module does not contain self-reported data on illegal behaviours taking place while young people hang around. However, it contains data on the reasons why adults (mainly the police or members of the public) have asked young people to move from the place where they were hanging around. These data are in line with those from the adult module, and show the young people are generally asked to move due to minor antisocial behaviours (e.g., being noisy), but also that behaviours such as smoking, drinking, intimidating others and damaging properties or areas are associated with hanging around.

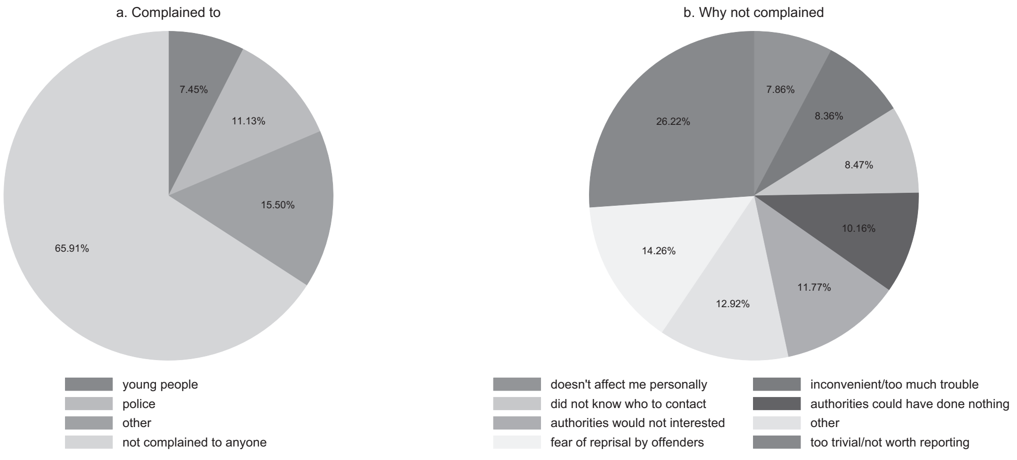

Fig. A7. *HA Reported to Somebody.*

Notes: British Crime Survey, main adults sample, 2004/2005. Weighted data. Graph a: N = 3,530. Other is constructed as follows: pub landlord/bar manager: 0.60%; landlord of the house or flat 0.72%; neighbourhood warden 0.92%; teachers or local school: 1.64%; tenants/residents association 1.70%; local councillor/MP: 2.20%; neighbourhood watch: 2.29%; local council department 3.85%; parents/family/friends of young people 4.17%; other: 1:47%. Graph b: The question is only asked to those who did not complain to anyone. N = 2,301. Other is constructed as follows: dislike/fear of police/authorities: 0.31%; previous bad experience of the police 0.80%; just accept it 0.84%; private/personal/family matter 1.84%; waste of time 1.99%; problem is already known 2.37%; dealt with matter myself/ourselves 72%; other 2.14%.

fear of reprisal by offenders. This suggests the relative low incidence of complaints is unlikely to be due to a low perceived salience of the problem.

A3. The Impact of Hanging Around with Friends and Purposeful Activities on Young People's Well-Being

To assess the effect of hanging around on teenagers' contemporaneous and future well-being, we use LSYPE data (and its longitudinal dimension), together with propensity score techniques. Hanging around with friends is measured at wave one. As outcome variables, we use contemporaneous (at wave one) and future (at waves five, six and seven) indicators of human capital investment and risky behaviours, and future indicators (at wave seven) of strength of social ties (see Table A1).[52] The analysis investigates whether *HA* is as negative as described in the BCS data

[52] The wording of the questions used to derive the variables in Table A1 is the following: 1. 'During an average week in term time, on how many evenings do you do any homework?' (respondents are prompted to consider only weekdays from Monday to Friday inclusive); 2. 'How often do you read books, magazines or newspapers for pleasure?' (available answers: most days, more than once a week, once a week, less than once a week, never); 3. 'Have you applied for a place on a university course which will start either this year that is in September/October 2008 or next year, that is in September/October 2009?'; 4. 'Have you received any offers of places yet, either conditional offers or unconditional offers?'; 7. 'Do you ever smoke cigarettes at all?'; 8. 'Have you ever had a proper alcoholic drink? That is a whole drink, not just a sip. Please do not count drinks labelled low alcohol'; 9. 'Have you ever tried cannabis even if only once?'; 10. 'Have you ever written on walls with spray cans?'; 11. 'Have you ever smashed, slashed or damaged public property or something in a public place?'; 12. 'Have you ever taken something from a shop, supermarket, or department store without paying?' 13. 'Have you ever taken part in fighting or some sort of disturbance in public for example, at a football ground, a railway station, music festival, riot, demonstration or just in the street?'; 14. 'Have you ever had sex without using precautions or contraception?' (respondents are prompted not to include any times when trying for a baby); 15.

Table A1. *Outcomes Description.*

Measure	Variable creation	Wave	Age
Human capital			
1. Homework	Number of days spent doing homework (for respondents who are given homework)	1	14/15
2. Reading	Binary variable equal to one if the respondent reads for pleasure at least once a week, and zero otherwise	1	14/15
3. Applied for higher education	Binary variable equal to one if the respondent has applied for a place in HE, and zero otherwise	5	18/19
4. Offers from higher education	Binary variable equal to one if the respondent has received an offer from higher education institutions, and zero otherwise	5	18/19
5. In higher education	Binary variable equal to one if the respondent attends HE, and zero otherwise	7	20/21
6. Russell group	Binary variable equal to one if the respondent attends a HE institution in the Russel group, and zero otherwise	7	20/21
Risky behaviours			
7. Tried cigarette	Binary variable equal to one if the respondent has ever smoked cigarettes, and zero otherwise.	1	14/15
8. Tried alcohol	Binary variable equal to one if the respondent has ever had a proper alcoholic drink, and zero otherwise.	1	14/15
9. Tried cannabis	Binary variable equal to one if the respondent has ever tried cannabis, and zero otherwise.	1	14/15
10. Done gaffiti	Binary variable equal to one if the respondent has ever written on walls with spray cans, and zero otherwise.	1	14/15
11. Vandalised	Binary variable equal to one if the respondent has ever vandalised public property, and zero otherwise.	1	14/15
12. Shoplifted	Binary variable equal to one if the respondent has ever shoplifted, and zero otherwise.	1	14/15
13. Involved in fights	Binary variable equal to one if the respondent has ever shoplifted, and zero otherwise.	1	14/15
14. Had unsafe sex	Binary variable equal to one if the respondent has ever had unsafe sex, and zero otherwise.	6	19/20
15. Tried cannabis	Binary variable equal to one if the respondent has ever taken cannabis, and zero otherwise.	6	19/20
16. Tried other drugs	Binary variable equal to one if the respondent has ever taken drugs other than cannabis, and zero otherwise.	6	19/20
Social ties			
17. Number of close friends	Numerical variable obtained by taking the mid-points of the bands used to record number of friends.	6	19/20

presented above. We investigate whether *HA* displaces investment in human capital (reading or study time) and encourages risky behaviours, and whether these effects persist over time. We also investigate whether *HA* engenders the creation of long-lasting friendship ties. This potentially positive effect of *HA* is overlooked in the political debate.

Table A2 compares the means of the outcome variables for respondents who hang around at wave one (*HA* respondents) and respondents who do not hang around at wave one (non-*HA* respondents). All differences are statistically significant. *HA* respondents experience lower human capital accumulation than non-*HA* respondents. For example, at age 14/15, *HA* respondents spend 14% less time (almost half the evening) doing homework. By age 18/19, *HA* respondents are over 25% (12 p.p.) less likely to apply to and receive offers from higher education institutions.

'Have you ever taken cannabis?'; 16. 'Have you ever taken other drugs?'; 17 'How many close friends do you have, that is friends you could talk to if you were in some sort of trouble?' (available answers are: none, 1, 2-3, 4-5, 6-9, 10 or more). Variables in rows 5 and 6 are derived from a number of different questions.

Table A2. *Tests for Equality of Means for HA.*

	Mean			t-tests		N. Observations		
	HA	*Non-HA*	*Difference*	*SE*	*p-value*	*HA*	*Non-HA*	Age
Human capital								
1. Homework	2.546	2.950	– 0.404	0.027	0.000	6794	6101	14/15
2. Reading	0.725	0.783	– 0.058	0.008	0.000	6943	6239	14/15
3. Applying to higher education	0.346	0.466	– 0.119	0.010	0.000	4627	4249	18/19
4. Offers from higher education	0.327	0.434	– 0.108	0.010	0.000	4623	4246	18/19
5. In higher education	0.299	0.409	– 0.110	0.010	0.000	4291	4001	20/21
6. In Russell group	0.064	0.092	– 0.029	0.006	0.000	4282	3996	20/21
Risky behaviours								
7. Tried cigarettes	0.132	0.058	0.074	0.005	0.000	6575	5987	14/15
8. Tried alcohol	0.579	0.343	0.236	0.009	0.000	6552	5935	14/15
9. Tried cannabis	0.117	0.052	0.065	0.005	0.000	6758	6068	14/15
10. Done graffiti	0.088	0.039	0.049	0.004	0.000	6823	6098	14/15
11. Vandalised	0.136	0.060	0.076	0.005	0.000	6698	6044	14/15
12. Shoplifted	0.147	0.081	0.066	0.006	0.000	6720	6041	14/15
13. Involved in fights	0.247	0.122	0.125	0.007	0.000	6698	6018	14/15
14. Had unsafe sex	0.628	0.390	0.238	0.016	0.000	4289	3996	19/20
15. Tried cannabis	0.381	0.228	0.152	0.010	0.000	4239	3966	19/20
16. Tried other drugs	0.134	0.072	0.062	0.007	0.000	4246	3968	19/20
Social ties								
17. Number of close friends	5.660	5.339	0.321	0.075	0.000	4237	3947	19/20

Notes: LSYPE waves 1, 5, 6, 7.

Two years later, they are 27% (11 p.p.) less likely to be in higher education and 31% (3 p.p.) less likely to be in a Russell group university.[53] *HA* respondents are also more likely than non-*HA* respondents to be involved in risky behaviours. At the age of 14/15, *HA* respondents are 69% (23 p.p.) more likely to have tried alcohol and 81% (7 p.p.) more likely to have shoplifted. They are more than twice as likely to have tried cigarettes or cannabis, and to have been involved in graffiti, vandalising, and fights. These differences in risky behaviours reduce only slightly with time. By age 19/20, *HA* respondents are 61% (24 p.p.) more likely to have had unsafe sex, 67% (15 p.p.) more likely to have tried cannabis, and 86% (6 p.p.) more likely to have tried other drugs. Finally, Table A2 shows that *HA* respondents have 6% more close friends in adulthood.

The differences in means in Table A2 cannot be interpreted as causal. They are likely to be due to a combination of selection and causal effect. Young people who are less interested in studying and reading and more attracted by risky behaviours are likely to self select into *HA*. This is the selection effect. The potential causal effect is three-fold. First, the time spent hanging around is subtracted to studying and reading. Thus, *HA* may harm human capital accumulation through a decrease in study/reading time. Second, hanging around takes place without the supervision of adults. This reduces the likelihood of being sanctioned and thus the cost of risky behaviours.

[53] The Russell Group is a self selected group of 24 public funded universities. Affiliation to the Russell group is often perceived as a signal of high quality education.

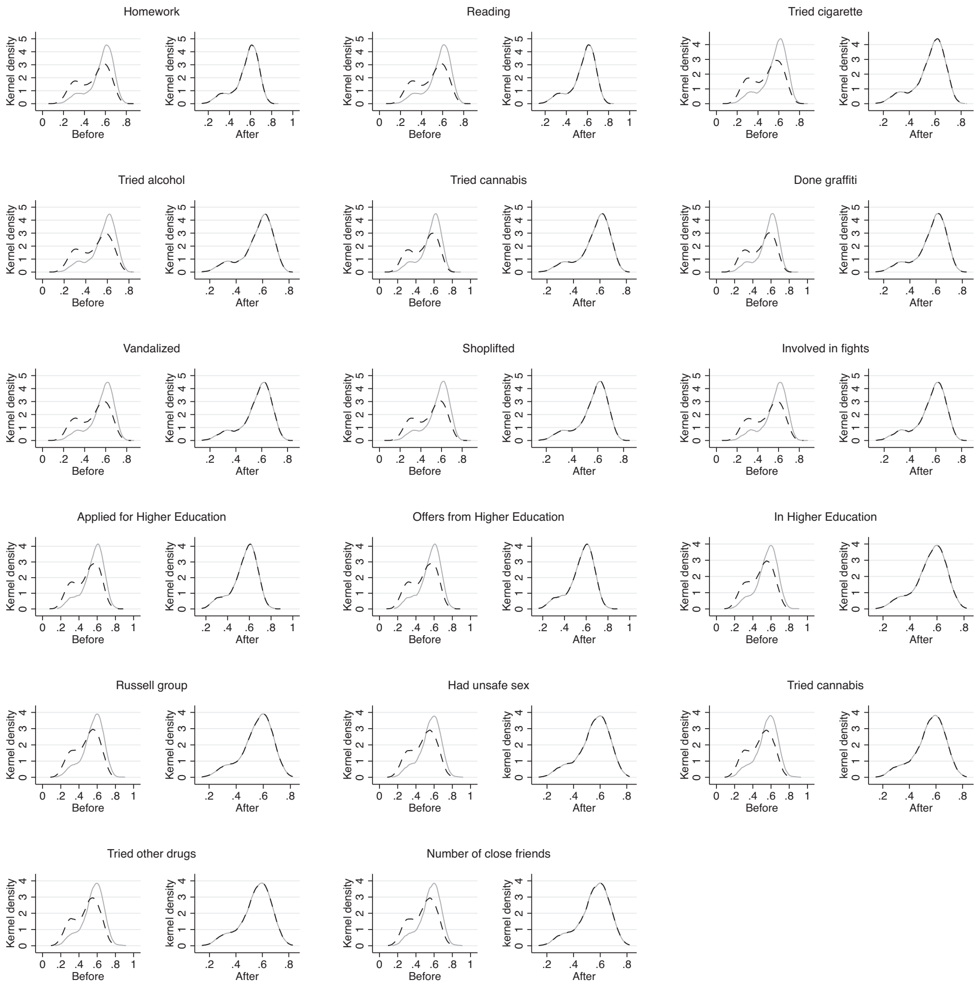

Fig. A8. *Propensity Scores for Treated and Controls Before and After Matching (HA).*
Notes: The 'Before' graphs show the distribution of the propensity scores for the treated (black dashed line) and the control observations (grey solid line) before the matching. The 'After' graphs show the distribution of the propensity scores for the treated (black dashed line) and the control observations (grey solid line) where the control observations within a certain radius from the treated have been weighted proportionally to the inverse of their distance from the treated. Weights for the treated observations are equal to 1.

Third, if 'risky' young people self-select into *HA,* peer effects may encourage less study and more risky behaviours. The fact that *HA* respondents form stronger friendship ties, while positive *per se*, might exacerbate the role of negative peer effects.

To shed light on the causal effect of hanging around (Average Treatment Effect on the Treated (ATET)) on young people's human capital accumulation, risky behaviours and social ties, we use propensity score techniques. We use the distance-weighted radius matching with bias adjustment by Lechner *et al.* (2011), which weights the control observations using the inverse of their

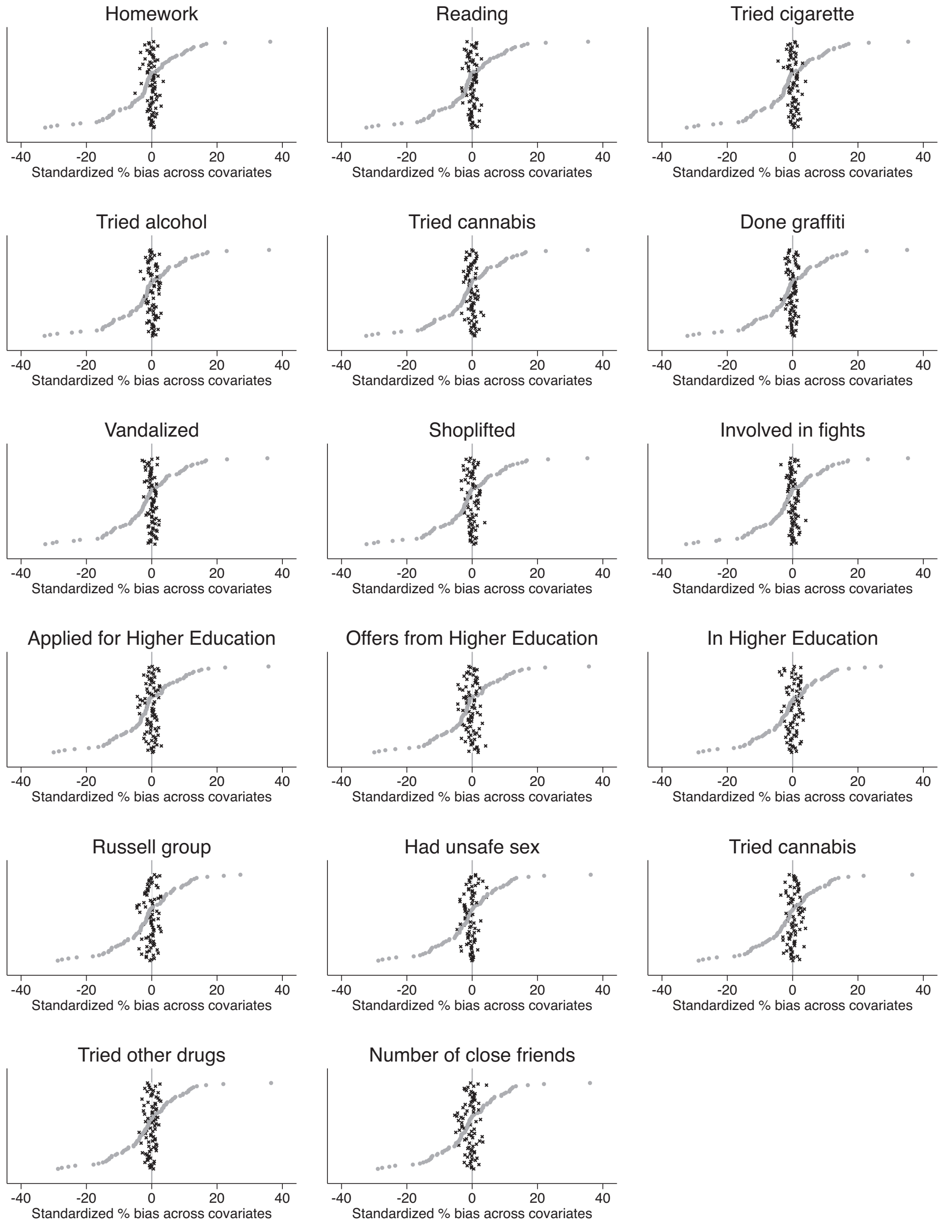

Fig. A9. *Standardised Bias for Covariates Before and After Matching (HA).*
Notes: The grey dots indicate the bias in the covariates before the matching, the black 'x' signs indicate the bias in the covariates after the matching. The bias after the matching is computed after weighting the control observations using the inverse of their distance from the corresponding treated observation.

distance from the corresponding treated observation.[54] Distance-weighted radius matching limits the likelihood of a bad match, and thus is more precise and less biased than nearest neighbour matching. This is particularly desirable in our case, where the treatment group is larger than the control group. Moreover, the estimator has been proven to be robust to misspecification of the propensity score (Huber *et al.*, 2013, 2015).

With propensity scores, identification of ATET relies on having observable covariates that permit to predict the treatment such that, conditional on the propensity score, the potential outcomes in the absence of treatment are independent of the treatment assignment (Conditional Independence Assumption CIA). To try to satisfy this assumption, we estimate the propensity score using a large set of individual, household and neighbourhood characteristics from wave one, including a set of characteristics measured at birth and in early age.[55] In addition, identification of ATET relies on each treated observation having a non-treated counterpart with similar characteristics and thus propensity score (common support assumption).

For each outcome analysed, Figure A8 shows the distribution of the propensity score for the treated and the control observations before and after the matching. The post-matching distributions are derived by re-weighting the pre-matching propensity scores using weights obtained in the matching procedure. The post-matching distributions for the treated and the control observations overlap for all outcomes.[56] This shows the common support assumption holds. While the pre-matching distributions of the propensity score for the treated and the control observations are clearly different, the post-matching distributions are almost identical. This provides a first glance into the ability of the propensity scores of making treated (*HA* respondents) and control (non-*HA* respondents) observations comparable.

Figure A9 shows the standardised percentage bias (i.e., covariate imbalance) for each covariate before and after the matching. While the pre-matching bias is substantial (between −40 and +40 standardised percentage bias), the post-matching bias is close to zero for all covariates. Diagnostics on overall post-matching covariates imbalance are reported in Tables A3 and A4.[57] All tests suggest our propensity scores are satisfactory. The *R* statistics are close to one, the B statistics are generally below 16 (see Table A3). The mean and the median bias, that before the propensity score exceeded 5% and 7%, respectively, now drop at or below 1% (see Table A4). The pseudo R^2 are very close to zero (and much smaller than those before the matching), and

[54] The weights are then used in a regression to correct the bias due to the mismatch. The estimator is implemented in stata through the command 'radiusmatch' (Huber *et al.*, 2012).

[55] These variables are: respondent's characteristics, i.e., gender, ethnicity, birth weight, whether the respondent was born early, whether the respondent has attended a nursery in her early childhood, month of birth (in months of distance from the previous August, when school starts); household characteristics, i.e., household income (in quintiles), whether the household has a computer at home, whether the household has internet access, whether the household has family dinners every or most days, whether the main parent is female, mother's and father's education (degree, above GCSE, GCSE and below, no education), mother's and father's employment status (full time, part time, not working), religion of main parent, health of main and second parent, number of cars (none, one, two, more than two), whether mother/father was NEET when the respondent was five, total months of unemployment of mother/father since the respondent was born, whether the respondent has ever lived in a single-parent household; neighbourhood characteristics, i.e., indices of ethnic fractionalisation and segregation, as used in the paper. We also add survey weight as in Rubin (2001). Dummy variables are used with categorical variables, with a separate dummy indicating missing data.

[56] Weights for the treated observations are equal to 1. The number of off-support observations ranges from a minimum of zero to a maximum of 14. These observations have not been used in the analysis.

[57] R is the ratio between the variance of the propensity score index for *HA* and non-*HA* respondents; B is the standardised difference in the means of the propensity scores computed for the *HA* and the matched non-*HA* respondents. To be satisfactory, R should be close to one, and surely between 0.5 and 2, while B must be below 25 (Rubin, 2001). The pseudo R^2 is derived from regressing the propensity score on the covariates used to estimate the propensity score. Reported p-values are from the likelihood-ratio test of the joint insignificance of those covariates.

Table A3. *Propensity Score: Diagnostics, Part One (HA).*

	R		B		
	Before	*After*	*Before*	*After*	Age
Human capital					
1. Homework	0.671	1.018	56.441	13.569	14/15
2. Reading	0.671	1.150	56.129	12.111	14/15
3. Applied to higher education	0.712	1.150	55.183	14.138	18/19
4. Offers from higher education	0.710	1.043	55.196	16.073	18/19
5. In higher education	0.747	1.228	54.649	15.896	20/21
6. In Russell group	0.748	1.183	54.655	13.925	20/21
Risky behaviours					
7. Tried cigarettes	0.677	1.207	56.970	13.123	14/15
8. Tried alcohol	0.678	1.176	57.145	11.848	14/15
9. Tried cannabis	0.675	1.015	56.516	12.693	14/15
10. Done graffiti	0.670	1.117	56.332	11.331	14/15
11. Vandalised	0.669	1.118	56.616	11.653	14/15
12. Shoplifted	0.669	0.930	56.334	13.151	14/15
13. Involved in fights	0.665	1.106	56.661	10.809	14/15
14. Had unsafe sex	0.760	1.395	55.407	15.564	19/20
15. Tried cannabis	0.758	1.565	55.405	14.559	19/20
16. Tried other drugs	0.749	0.836	55.156	14.986	19/20
Social ties					
17. Number of close friends	0.749	0.998	55.076	15.626	19/20

Notes: LSYPE waves 1, 5, 6, 7. ATET are estimated using the radiusmatch stata command (Huber *et al.*, 2012).

the likelihood test is not able to reject the hypothesis of joint insignificance of the observables in explaining the propensity score (see Table A4).

Table A5 shows the ATETs estimated via propensity score matching. The effect of *HA* suggested by the ATETs are all statistically significant (generally at the 1% level), and maintain the same sign as the t-tests in Table A2. However, the ATETs are smaller in magnitude. They suggest *HA* respondents spend 7% less time doing homework (one-fifth of a day) than non-*HA* respondents and are 4% (3 p.p.) less likely to read for pleasure at least once a week. *HA* respondents are 10% and 12% (6 and 5 p.p.) less likely to apply to and receive offers from higher education institutions, respectively. They are also 13% (5 p.p.) less likely to be in higher education at age 20/21. The ATET for attending a Russell group university is −12% (1 p.p.), but it is at the verge of statistical significance.

The estimated ATETs for risky behaviours at wave one are between five (done graffiti) and 13 p.p. (tried alcohol). These effects translate into 112% and 39%, respectively. The ATETs for having had unsafe sex (17 p.p.) and having tried cannabis by age 19/20 (12 p.p.) are sizeable. In percentage terms, these ATETs suggest that *HA* increases the probability of both risky behaviours by around a half. Equally, the ATET for trying other drugs, only 4 p.p., signals an increase of more than half, when seen in percentage. Finally, the ATET for the number of close friends at age 19/20 is small in magnitude (around 5%), and very close to the estimate obtained using simple t-tests and reported in Table A2.

In summary, our results suggest the differences observed in the mean outcomes of *HA* and non-*HA* respondents are partially due to selection. However, there might be a causal effect of

Table A4. *Propensity Score: Diagnostics, Part Two (HA).*

	Median bias		Mean bias		Chi prob		R^2		
	Before	*After*	*Before*	*After*	*Before*	*After*	*Before*	*After*	Age
Human capital									
1. Homework	5.664	0.658	8.200	0.993	0.000	0.937	0.055	0.003	14/15
2. Reading	5.544	0.791	8.177	1.007	0.000	0.996	0.055	0.003	14/15
3. Applied to higher education	5.108	0.970	7.986	1.221	0.000	0.999	0.053	0.004	18/19
4. Offers from higher education	5.165	1.206	7.985	1.456	0.000	0.964	0.053	0.005	18/19
5. In higher education	5.313	1.256	7.688	1.435	0.000	0.990	0.052	0.005	20/21
6. In Russell group	5.420	0.938	7.705	1.340	0.000	1.000	0.052	0.004	20/21
Risky behaviours									
7. Tried cigarettes	5.726	0.860	8.323	1.067	0.000	0.982	0.056	0.003	14/15
8. Tried alcohol	5.328	0.974	8.302	1.122	0.000	0.999	0.057	0.003	14/15
9. Tried cannabis	5.529	0.828	8.202	0.997	0.000	0.990	0.055	0.003	14/15
10. Done graffiti	5.378	0.812	8.191	0.899	0.000	1.000	0.055	0.002	14/15
11. Vandalised	5.371	0.776	8.270	0.859	0.000	1.000	0.056	0.002	14/15
12. Shoplifted	5.519	0.912	8.259	1.186	0.000	0.974	0.055	0.003	14/15
13. Involved in fights	5.607	0.720	8.306	0.905	0.000	1.000	0.056	0.002	14/15
14. Had unsafe sex	5.231	0.930	7.789	1.310	0.000	0.991	0.053	0.005	19/20
15. Tried cannabis	5.539	1.111	7.835	1.343	0.000	0.999	0.053	0.004	19/20
16. Tried other drugs	5.546	1.080	7.816	1.169	0.000	0.999	0.053	0.004	19/20
Social ties									
17. Number of close friends	5.466	1.010	7.776	1.558	0.000	0.993	0.053	0.005	19/20

Notes: LSYPE waves 1, 5, 6, 7. ATET are estimated using the radiusmatch stata command (Huber *et al.*, 2012).

HA. In the case of the human capital outcomes, the ATETs estimated through propensity score matching are about half the size of those suggested by t-tests. Such a drop in the estimated effect following the use of propensity score matching is not observed in the case of risky behaviours and social ties (compare, for example, the results in Tables A2 and A5 for outcomes 10 and 17). This suggests that the negative effect of *HA* on risky behaviours and the (small) positive effect on social ties is more likely to be causal than the one on education. The results on risky behaviours are in line with our analysis of BCS data: they suggest risky behaviours do take place when young people hang around. Moreover, *HA* is likely to affect risky behaviours through peer effects, as vandalism, fights and drug initiation are more likely to be group activities than studying. The results on social ties are also likely to be at least partially causal, as spending time hanging around with friends may strengthen friendship ties.

We now look at the effects of purposeful activities (*PA*). Table A6 compares the means of the outcomes in Table A1 for those who take part in *PA* at wave one (*PA* respondents) and those who do not take part in *PA* at wave one (non-*PA* respondents). *PA* is associated with higher human capital accumulation. For example, compared with non-*PA* respondents, *PA* respondents spend around one-fifth of an evening more doing homework at age 14/15 and are 8% more likely to be in higher education at age 20/21. The picture on risky behaviour is more nuanced. At age 14/15, *PA* respondents are less likely to have tried cannabis, but are more likely to have tried alcohol or to have been involved in graffiti, vandalising, shoplifting, and fights. At age 19/20, *PA* respondents are more likely to have had unsafe sex, and tried drugs (both cannabis and other

Table A5. *Propensity Score: Estimated ATET (HA).*

	ATET			
	Estimate	*SE*	*p-value*	Age
Human capital				
1. Homework	– 0.202	0.035	0.000	14/15
2. Reading	– 0.030	0.010	0.002	14/15
3. Applied to higher education	– 0.056	0.013	0.000	18/19
4. Offers from higher education	– 0.045	0.013	0.000	18/19
5. In higher education	– 0.052	0.013	0.000	20/21
6. In Russell group	– 0.011	0.007	0.091	20/21
Risky behaviours				
7. Tried cigarettes	0.061	0.006	0.000	14/15
8. Tried alcohol	0.133	0.011	0.000	14/15
9. Tried cannabis	0.050	0.006	0.000	14/15
10. Done graffiti	0.045	0.005	0.000	14/15
11. Vandalised	0.061	0.007	0.000	14/15
12. Shoplifted	0.053	0.007	0.000	14/15
13. Involved in fights	0.108	0.009	0.000	14/15
14. Had unsafe sex	0.170	0.020	0.000	19/20
15. Tried cannabis	0.117	0.013	0.000	19/20
16. Tried other drugs	0.044	0.009	0.000	19/20
Social ties				
17. Number of close friends	0.289	0.097	0.003	19/20

Notes: LSYPE waves 1, 5, 6, 7. ATET are estimated using the radiusmatch stata command (Huber *et al.*, 2012).

drugs). Finally, at age 19/20 *PA* respondents have on average almost one close friend more than non-*PA* respondents.

Again, the differences in Table A6 cannot be considered causal. The potential causal effect of *PA* is unclear *a priori*. On the one hand, *PA* foster skills like grit, self-confidence, trust and cooperation. These skills are likely to boost human capital accumulation and discourage risky behaviours. On the other hand, *PA* subtract time to homework and reading and put young people in contact with other people of the same age. Therefore, *PA* may interfere with human capital accumulation and expose young people to peer effects encouraging risky behaviours.

As with *HA*, we use propensity score matching to shed light on the causal effect of *PA*.[58] Figures A10 and A11, and Tables A7 and A8 suggest the matching has successfully achieved comparability between the treatment and the control group. Figure A10 shows the propensity scores for treated and controls have a large common support. Moreover, it shows the distributions of the propensity score computed for treated and control observations are initially quite different, but then become indistinguishable when reweighed. Figure A11 shows the standardised bias for the covariates is initially sizeable, but it dramatically reduces after the matching. Finally, all measures for overall covariate imbalance are satisfactory (see Tables A7 and A8).

The ATETs for *PA* obtained via propensity score matching are shown in Table A9. The results on human capital accumulation and social ties maintain the same sign—and roughly the same size—as the t-tests in Table A6. For example, the ATETs suggest *PA* respondents are likely to

[58] The technique, and the variables used to estimate the propensity score are the same used for *HA*.

Table A6. *Tests for Equality of Means for PA.*

	Mean			t-tests		N. Observations		
	PA	*Non-PA*	*Difference*	*SE*	*p-value*	*PA*	*Non-PA*	Age
Human capital								
1. Homework	2.794	2.631	0.164	0.028	0.000	8408	4487	14/15
2. Reading	0.767	0.726	0.041	0.008	0.000	8546	4636	14/15
3. Applying to higher education	0.435	0.343	0.092	0.011	0.000	5836	3040	18/19
4. Offers from higher education	0.410	0.316	0.094	0.011	0.000	5831	3038	18/19
5. In higher education	0.378	0.303	0.076	0.011	0.000	5454	2838	20/21
6. In Russell group	0.093	0.048	0.044	0.006	0.000	5446	2832	20/21
Risky behaviours								
7. Tried cigarettes	0.085	0.120	− 0.035	0.006	0.000	8190	4372	14/15
8. Tried alcohol	0.488	0.427	0.061	0.009	0.000	8117	4370	14/15
9. Tried cannabis	0.087	0.087	0.000	0.005	0.995	8330	4496	14/15
10. Done graffiti	0.068	0.059	0.009	0.005	0.058	8393	4528	14/15
11. Vandalised	0.109	0.084	0.026	0.006	0.000	8269	4473	14/15
12. Shoplifted	0.121	0.107	0.015	0.006	0.014	8293	4468	14/15
13. Involved in fights	0.205	0.156	0.049	0.007	0.000	8263	4453	14/15
14. Had unsafe sex	0.530	0.483	0.047	0.017	0.005	5456	2829	19/20
15. Tried cannabis	0.328	0.267	0.061	0.011	0.000	5405	2800	19/20
16. Tried other Drugs	0.112	0.090	0.023	0.007	0.002	5412	2802	19/20
Social ties								
17. Number of close friends	5.817	4.903	0.914	0.078	0.000	5391	2793	19/20

Notes: LSYPE waves 1, 5, 6, 7.

spend over one-fifth of an evening doing homework more than non-*PA* respondents. Moreover, *PA* respondents are 19% (almost 7 p.p.) more likely to apply to university than non-*PA* respondents at age 18/19, and 21% more likely (over 6 p.p.) to be in higher education two years later. The ATETs for risky behaviours suggest at age 14/15 *PA* respondents are 20% less likely (around 2 p.p.) to have tried cigarettes and cannabis, but also 30% more likely (5 p.p.) to be involved in fights.[59] For risky behaviours the ATETs suggest a more desirable effect of *PA* than the t-tests, as most of the positive signs estimated through t-tests become zeros or turn negative when propensity scores are used.

In summary, we find that *PA* are associated with—and possibly determine—higher human capital accumulation and greater number of close friends in adulthood. The association between *PA* and risky behaviours is less clear. Our analysis suggests *PA* participants are more likely to select into risky behaviours, but the causal effect of *PA* may go in the opposite direction, and *PA* may prevent young people from being involved in (at least some) risky behaviours.

Our findings on *HA* are entirely new, as no literature investigates this pervasive and policy-relevant behaviour. Our findings on *PA* are in line with the existing literature that generally finds positive effects of *PA* on educational attainment (Long and Caudill, 1991; Barron *et al.*, 2000; Eide and Ronan, 2001; Lipscomb, 2007; Lechner, 2009; Pfeifer and Cornelißen, 2010; Rees and

[59] The result on fights may be partially due to the fact that LSYPE reports as fights also those occurred in political rallies, and participating in political rallies is a component of *PA*.

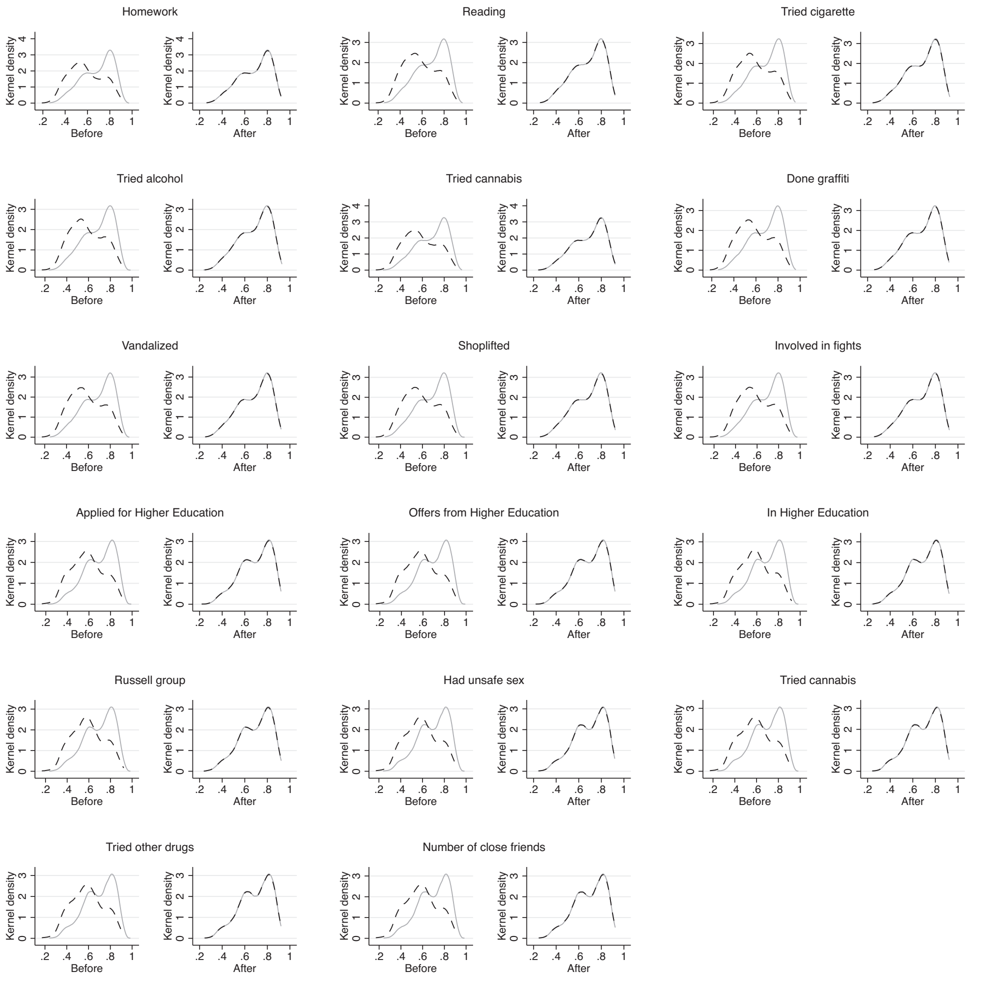

Fig. A10. *Propensity Scores for Treated and Controls Before and After Matching (PA).*
Note. The 'Before' graphs show the distribution of the propensity scores for the treated (black dashed line) and the control observations (grey solid line) before the matching. The 'After' graphs show the distribution of the propensity scores for the treated (black dashed line) and the control observations (grey solid line) where the control observations within a certain radius from the treated have been weighted proportionally to the inverse of their distance from the treated. Weights for the treated observations are equal to 1.

Sabia, 2010; Stevenson, 2010; Felfe *et al.*, 2016) but no effects on antisocial behaviours (Felfe *et al.*, 2016).[60]

The results in this appendix rest on the strong assumption of selection on observables, and thus one needs to be cautious in claiming causality.[61] However, our propensity score analysis is

[60] Negative effects of sport participation on crime are found by Caruso (2011), but the identification of the effect only relies on regional variation for Italian regions.

[61] We are fairly confident about the robustness of our propensity score analysis for two reasons. First, our propensity scores are based on a large set of variables. This reduces any unobservable effect potentially left in the error term. Second,

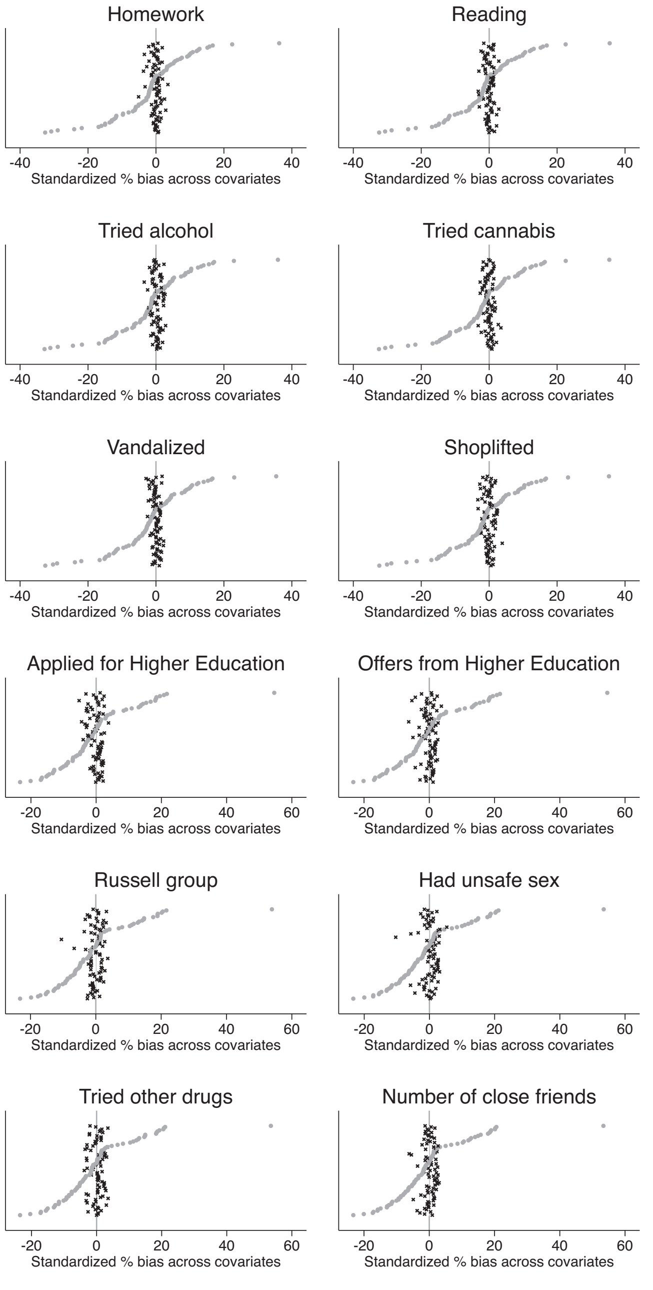

Fig. A11. *Standardised Bias for Covariates Before and After Matching (PA).*

Notes: The grey dots indicate the bias in the covariates before the matching, the black 'x' signs indicate the bias in the covariates after the matching. The bias after the matching is computed after weighting the control observations using the inverse of their distance from the corresponding treated observation.

Table A7. *Propensity Score: Diagnostics, Part One (PA).*

	R		B		
	Before	*After*	*Before*	*After*	Age
Human capital					
1. Homework	0.986	0.999	70.712	13.920	14/15
2. Reading	0.990	0.963	70.354	13.601	14/15
3. Applying to higher education	0.996	1.077	71.671	15.544	18/19
4. Offers from higher education	0.996	0.955	71.620	17.656	18/19
5. In higher education	0.995	1.114	70.210	17.223	20/21
6. In Russell group	0.996	0.816	70.439	17.447	20/21
Risky behaviours					
7. Tried cigarettes	0.964	0.846	71.008	14.453	14/15
8. Tried alcohol	0.984	1.025	70.756	13.213	14/15
9. Tried cannabis	0.980	1.047	70.424	12.904	14/15
10. Done graffiti	0.984	0.926	70.398	12.879	14/15
11. Vandalised	0.983	1.065	70.670	14.605	14/15
12. Shoplifted	0.981	1.206	70.271	13.749	14/15
13. Involved in fights	0.978	1.102	70.568	14.528	14/15
14. Had unsafe sex	0.992	0.936	70.517	18.301	19/20
15. Tried cannabis	0.994	1.085	70.751	17.448	19/20
16. Tried other Drugs	0.990	3.136	70.579	15.277	19/20
Social ties					
17. Number of close friends	0.992	0.856	70.491	17.228	19/20

Notes: LSYPE waves 1, 5, 6, 7. ATET are estimated using the radiusmatch stata command (Huber *et al.*, 2012).

not able to revert the initial conclusions based on raw data. We still conclude *HA* is negatively associated with human capital accumulation and positively associated with risky behaviours. In contrast, *PA* are positively associated with accumulation of human capital. This suggests that there is scope for economists to work on the identification, under weaker assumptions, of the causal mechanisms linking *HA* and *PA* to these crucial outcomes. Also, this suggests the importance of studying the determinants of *HA* and *PA*, as the body of the article does.

A4. Social Interaction and Preference for Own Ethnic Group.

In this section, we use LSYPE data to provide empirical support to our hypothesis that young people prefer to interact with other young people from their own ethnic group. For each of the main ethnic groups (White, Indian, Pakistani, Bangladeshi, Black Caribbean, Black African), Table A10 reports: (*i*) the share of that ethnic group in the sample (as the data are weighted, this is an estimate of the share in the population); (*ii*) information on the ethnicity of the friends

the selection on observables, which can be inferred by the difference between the t-tests and the ATETs are tiny, and so is likely to be the selection on unobservables. This argument is based on the assumption that the bias due to observables has the same sign as the bias due to unobservables, and the latter is at most as important as the former (Altonji *et al.*, 2005).

Table A8. *Propensity Score: Diagnostics, Part Two (PA).*

	Median bias		Mean Bias		Chi Prob		R^2		
	Before	*After*	*Before*	*After*	*Before*	*After*	*Before*	*After*	Age
Human capital									
1. Homework	5.417	1.193	8.035	1.373	0.000	0.466	0.082	0.004	14/15
2. Reading	5.176	1.152	8.036	1.187	0.000	0.547	0.081	0.003	14/15
3. Applying to higher education	6.279	1.221	8.649	1.482	0.000	0.788	0.084	0.004	18/19
4. Offers from higher education	6.279	1.113	8.640	1.485	0.000	0.211	0.084	0.006	18/19
5. In higher education	6.429	1.390	8.533	1.480	0.000	0.482	0.081	0.005	20/21
6. In Russell group	6.479	1.408	8.564	1.606	0.000	0.409	0.081	0.006	20/21
Risky behaviours									
7. Tried cigarettes	5.299	0.980	8.137	1.320	0.000	0.325	0.083	0.004	14/15
8. Tried alcohol	5.235	1.006	8.150	1.166	0.000	0.782	0.082	0.003	14/15
9. Tried cannabis	5.210	1.292	8.045	1.344	0.000	0.819	0.081	0.003	14/15
10. Done graffiti	5.033	1.201	8.034	1.254	0.000	0.812	0.081	0.003	14/15
11. Vandalised	5.186	1.291	8.097	1.433	0.000	0.279	0.082	0.004	14/15
12. Shoplifted	5.190	1.068	8.106	1.196	0.000	0.552	0.081	0.003	14/15
13. Involved in fights	5.409	1.067	8.033	1.314	0.000	0.304	0.082	0.004	14/15
14. Had unsafe sex	6.529	1.435	8.549	1.691	0.000	0.187	0.081	0.006	19/20
15. Tried cannabis	6.393	1.603	8.579	1.697	0.000	0.437	0.082	0.006	19/20
16. Tried other Drugs	6.450	1.361	8.548	1.513	0.000	0.278	0.082	0.006	19/20
Social ties									
17. Number of close friends	6.597	1.272	8.502	1.465	0.000	0.499	0.081	0.005	19/20

Notes: LSYPE waves 1, 5, 6, 7. ATET are estimated using the radiusmatch stata command (Huber *et al.*, 2012).

respondents from that ethnic group have at school; (*iii*) information on the ethnicity of the friends respondents from that ethnic group have outside school.[62]

Table A10 suggests that young people form ethnically homogeneous groups of friends and do it by choice. Consider the case of Pakistanis (third panel from the top), who are around 2% of the total. If friends' ethnicity was random, Pakistanis should have around 2% of Pakistani friends and thus report very few friends who are Pakistani themselves. However, 68.6% of the Pakistanis say at least half of their friends at school are Pakistani. This percentage reaches 79.12% in the case of friends outside school. This suggests that groups of friends are more homogeneous than if they were formed randomly. Table A10 also shows that friendship ties outside school are more ethnically homogeneous than those formed at school. While the ethnic mix of friends from school is driven by the ethnic composition of the school catchment areas, friends outside schools are more likely to be freely chosen. This suggests that the ethnic homogeneity of friendship groups is at least partially driven by preferences. If young people prefer to form ethnically homogeneous groups, *HA* and *PA* are likely to be affected by district ethnic diversity and segregation.

[62] The exact questions are: *'How many of your friends at your school are (e.g., White)?'* and *'How many of your friends outside school are (e.g White)?'*. The questions are asked at wave two. White British and other Whites are not distinguished in wave two.

Table A9. *Propensity Score: Estimated ATET (PA).*

	ATET			
	Estimate	*se*	*p-value*	Age
Human capital				
1. Homework	0.232	0.039	0.000	14/15
2. Reading	0.077	0.011	0.000	14/15
3. Applying to higher education	0.065	0.014	0.000	18/19
4. Offers from higher education	0.071	0.014	0.000	18/19
5. In higher education	0.064	0.015	0.000	20/21
6. In Russell group	0.026	0.009	0.003	20/21
Risky behaviours				
7. Tried cigarettes	− 0.024	0.008	0.002	14/15
8. Tried alcohol	0.021	0.013	0.105	14/15
9. Tried cannabis	− 0.018	0.008	0.024	14/15
10. Done graffiti	0.006	0.006	0.289	14/15
11. Vandalised	0.010	0.008	0.190	14/15
12. Shoplifted	0.012	0.008	0.139	14/15
13. Involved in fights	0.046	0.009	0.000	14/15
14. Had unsafe sex	0.007	0.023	0.759	19/20
15. Tried cannabis	0.024	0.015	0.102	19/20
16. Tried other Drugs	0.017	0.010	0.081	19/20
Social ties				
17. Number of close friends	0.629	0.108	0.000	19/20

Notes: LSYPE waves 1, 5, 6, 7. ATET are estimated using the radiusmatch stata command (Huber *et al.*, 2012).

Table A10. *Share of Friends from own Ethnic Group at School and Outside School.*

Ethnicity	Percentage in LSYPE		At school	Outside school
White	86.57	All or most of them	66.61	72.07
		More than half of them	23.86	19.30
		About half	6.87	5.16
		Less than half	1.47	1.62
		Very few of them	0.72	0.82
		None of them	0.15	0.34
		No friends at school/outside	0.31	0.68
Indian	2.38	All or most of them	14.73	25.30
		More than half of them	17.09	20.91
		About half	20.71	17.41
		Less than half	14.18	11.25
		Very few of them	23.22	14.21
		None of them	10.03	9.44
		No friends at school/outside	0.00	1.48
Pakistani	2.37	All or most of them	28.25	40.99
		More than half of them	20.26	20.60
		About half	20.09	17.53
		Less than half	11.02	9.64
		Very few of them	14.49	6.44
		None of them	5.89	3.47
		No friends at school/outside	0.00	1.33

Table A10. *Continued*

Ethnicity	Percentage in LSYPE		At school	Outside school
Bangladeshi	0.95	All or most of them	25.38	30.49
		More than half of them	14.97	22.96
		About half	16.84	17.09
		Less than half	12.05	10.18
		Very few of them	16.55	10.21
		None of them	14.21	6.91
		No friends at school/outside	0.00	2.15
Black Caribbean	1.43	All or most of them	16.97	23.81
		More than half of them	21.20	27.17
		About half	19.27	19.23
		Less than half	19.45	13.96
		Very few of them	17.31	11.96
		None of them	5.81	3.32
		No friends at school/outside	0.00	0.58
Black African	1.53	All or most of them	12.93	18.94
		More than half of them	17.17	19.69
		About half	19.50	24.67
		Less than half	22.32	16.66
		Very few of them	21.47	12.79
		None of them	5.67	5.90
		No friends at school/outside	0.34	1.36

Notes: LSYPE wave two, weighted data.

Appendix B: Additional Tables and Figures

Fig. B1. *Alien Act Port Towns and Population Growth.*
Notes: Urban Population Database 1801–1911.

Table B1. *Descriptive Statistics.*

Variable					
Household		Individual		Location	
Household income		Male	0.510	Sun	183.840
Top quintile	0.170		(0.004)		(0.332)
	(0.003)	Born in 1989	0.323	Rain	56.709
Second quintile	0.186		(0.004)		(0.215)
	(0.003)	Ethnicity			
Third quintile	0.230	*White British*	0.844		
	(0.004)		(0.003)		
Fourth quintile	0.196	*Other Whites*	0.018		
	(0.003)		(0.001)		
Top quintile	0.217	*Indian*	0.033		
	(0.003)		(0.001)		
English not mother tongue	0.029	*Black Caribbean*	0.030		
	(0.001)		(0.001)		
Female main parent	0.857	*Other Mixed*	0.020		
	(0.003)		(0.001)		
Mother's education		*Black African*	0.020		
No education	0.197		(0.001)		
	(0.003)	*Pakistani*	0.023		
GCSE or below	0.415		(0.001)		
	(0.004)	*Bangladeshi*	0.010		
Above GCSE	0.247		(0.001)		
	(0.004)	*Chinese*	0.004		
Degree	0.092		(0.001)		
	(0.002)				
Missing	0.048				
	(0.002)				
Mother's age					
Below 35	0.002				
	(0.101)				
35 to 49	0.796				
	(0.003)				
Above 50	0.067				
	(0.002)				
Missing	0.036				
	(0.002)				
Mother's employment status					
Working full time	0.324				
	(0.004)				
Working part time	0.356				
	(0.004)				
Not working	0.283				
	(0.004)				
Missing	0.038				
	(0.002)				
N	14,244		14,244		14,244

Notes: Standard errors in parentheses, *p < 0.1, **p < 0.05, ***p < 0.01. Weighted data restricted to the sample used in the estimates.

Table B2. *First Stages (Just Identified). Coefficients of the Instruments.*

Instrument	Diversity	Segregation
Proximity to the nearest port of entry	0.631*	– 0.399*
	(0.289)	(0.196)
Proximity to the nearest mill town	– 0.007	0.284***
	(0.055)	(0.082)
N	14,244	14,244

Notes: Standard errors in parentheses, *p < 0.1, **p < 0.05, ***p < 0.01. Individual controls: Male, born in 1989, ethnicity dummies. Family controls: income quintiles, English not mother tongue, main parent is female, mother's education, age and employment status. Weather controls: total sunshine duration, total rainfall. Each LSYPE respondent is assigned the weather information collected from the closest meteo station in the interview month. Weighted data.

Table B3. *Linear Probability Model and IV1, Other Coefficients.*

	PA				HA			
	OLS	*2SLS*	*GMM*	*LIML*	*OLS*	*2SLS*	*GMM*	*LIML*
Male	0.222***	0.221***	0.223***	0.221***	0.045***	0.044***	0.044***	0.044
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Born in 1989	0.013	0.013	0.014*	0.013	0.025***	0.025***	0.024***	0.025
	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)	(0.008)	(0.009)
Other Whites	0.093***	0.089***	0.087***	0.089***	– 0.106***	– 0.100***	– 0.096***	– 0.100
	(0.032)	(0.033)	(0.032)	(0.033)	(0.037)	(0.036)	(0.035)	(0.036)
Indian	– 0.051**	– 0.070**	– 0.056**	– 0.070**	– 0.228***	– 0.247***	– 0.247***	– 0.247
	(0.023)	(0.028)	(0.025)	(0.028)	(0.022)	(0.024)	(0.021)	(0.024)
Black Caribbean	0.043**	0.028	0.024	0.028	– 0.035	– 0.039	– 0.058*	– 0.040
	(0.019)	(0.024)	(0.023)	(0.024)	(0.027)	(0.034)	(0.030)	(0.034)
Other Mixed	– 0.001	– 0.010	– 0.008	– 0.010	– 0.148***	– 0.149***	– 0.148***	– 0.149
	(0.030)	(0.031)	(0.030)	(0.031)	(0.033)	(0.035)	(0.033)	(0.035)
Black African	– 0.004	– 0.011	– 0.005	– 0.011	– 0.158***	– 0.149***	– 0.143***	– 0.149
	(0.024)	(0.026)	(0.025)	(0.026)	(0.028)	(0.029)	(0.027)	(0.029)
Pakistani	– 0.038*	– 0.068***	– 0.064***	– 0.068***	– 0.242***	– 0.278***	– 0.278***	– 0.279
	(0.019)	(0.025)	(0.022)	(0.025)	(0.026)	(0.030)	(0.024)	(0.030)
Bangladeshi	– 0.037	– 0.049*	– 0.067**	– 0.050*	– 0.190***	– 0.188***	– 0.195***	– 0.188
	(0.026)	(0.029)	(0.027)	(0.029)	(0.031)	(0.038)	(0.036)	(0.038)
Chinese	– 0.028	– 0.035	– 0.025	– 0.035	– 0.150	– 0.155*	– 0.165*	– 0.155
	(0.071)	(0.072)	(0.068)	(0.072)	(0.092)	(0.092)	(0.089)	(0.093)
Second income quintile	– 0.003	– 0.002	0.002	– 0.002	– 0.001	– 0.001	0.001	– 0.001
	(0.015)	(0.015)	(0.014)	(0.015)	(0.013)	(0.013)	(0.013)	(0.013)
Third income quintile	0.028*	0.030**	0.030*	0.030*	– 0.018	– 0.016	– 0.013	– 0.016
	(0.015)	(0.015)	(0.014)	(0.015)	(0.016)	(0.016)	(0.015)	(0.016)
Fourth income quintile	0.047***	0.051***	0.046***	0.051***	– 0.015	– 0.010	– 0.001	– 0.010
	(0.014)	(0.014)	(0.014)	(0.014)	(0.017)	(0.017)	(0.016)	(0.017)
Top income quintile	0.077***	0.081***	0.080***	0.081***	– 0.057***	– 0.051***	– 0.048***	– 0.051***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.017)	(0.017)	(0.017)	(0.017)
English not mother tongue	– 0.047**	– 0.050**	– 0.051**	– 0.050**	– 0.094***	– 0.096***	– 0.098***	– 0.096***
	(0.023)	(0.023)	(0.022)	(0.023)	(0.026)	(0.026)	(0.025)	(0.026)
Female main parent	– 0.021*	– 0.021*	– 0.013	– 0.021*	0.027**	0.028**	0.026*	0.028**
	(0.013)	(0.013)	(0.012)	(0.013)	(0.014)	(0.014)	(0.013)	(0.014)
Mother's education: GCSE or below	0.061***	0.064***	0.062***	0.064***	– 0.008	– 0.003	– 0.003	– 0.003
	(0.013)	(0.013)	(0.012)	(0.013)	(0.014)	(0.014)	(0.013)	(0.014)
Mother's education: above GCSE	0.110***	0.114***	0.112***	0.114***	– 0.047***	– 0.042***	– 0.043***	– 0.042
	(0.014)	(0.015)	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Mother's education: degree	0.162***	0.164***	0.160***	0.165***	– 0.075***	– 0.071***	– 0.073***	– 0.071
	(0.019)	(0.020)	(0.019)	(0.020)	(0.020)	(0.020)	(0.019)	(0.020)
Mother's education: missing	0.031	0.034	0.036	0.034	– 0.065	– 0.060	– 0.057	– 0.060
	(0.039)	(0.039)	(0.038)	(0.039)	(0.041)	(0.041)	(0.040)	(0.041)
Mother's age: 35 to 49	0.009	0.011	0.012	0.011	– 0.049***	– 0.045***	– 0.047***	– 0.045***
	(0.016)	(0.016)	(0.015)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)

Table B3. *Continued*

	PA				HA			
	OLS	*2SLS*	*GMM*	*LIML*	*OLS*	*2SLS*	*GMM*	*LIML*
Mother's age: 50+	– 0.012	– 0.009	– 0.015	– 0.009	– 0.066***	– 0.059**	– 0.062***	– 0.059***
	(0.022)	(0.022)	(0.021)	(0.022)	(0.023)	(0.023)	(0.022)	(0.023)
Mother's age: missing	– 0.040	– 0.037	– 0.058	– 0.037	0.042	0.045	0.009	0.045
	(0.073)	(0.073)	(0.068)	(0.073)	(0.079)	(0.078)	(0.073)	(0.078)
Mother's employment status: part time	0.027**	0.028***	0.031***	0.028***	– 0.025**	– 0.023**	– 0.021*	– 0.023**
	(0.010)	(0.010)	(0.010)	(0.010)	(0.012)	(0.012)	(0.011)	(0.012)
Mother's employment status: not working	– 0.003	0.002	0.004	– 0.002	– 0.048***	– 0.045***	– 0.042***	– 0.045***
	(0.013)	(0.012)	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)
Mother's employment status: missing	– 0.013	– 0.014	0.013	– 0.014	– 0.029	– 0.028	0.010	– 0.028
	(0.080)	(0.080)	(0.075)	(0.080)	(0.082)	(0.081)	(0.077)	(0.081)
Sun	0.000	0.000	0.000	0.000	– 0.000	– 0.000	– 0.000	– 0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Rain	– 0.000**	– 0.001**	– 0.001***	– 0.001**	– 0.000	– 0.000	– 0.000	– 0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.467***	0.406***	0.427***	0.405***	0.678***	0.583***	0.596***	0.581***
	(0.046)	(0.052)	(0.047)	(0.052)	(0.049)	(0.056)	(0.049)	(0.057)
N	14,244	14,244	14,244	14,244	14,244	14,244	14,244	14,244

Notes: Standard errors in parentheses, *p < 0.1, **p < 0.05, ***p < 0.01. Reference categories: female, born in 1990, White British, bottom income quintile, mother tongue is English, mother with no education, mother younger than 35 years old, mother works full time. Weighted data.

Table B4. *IV1 First Stages. Coefficients of the Instruments.*

Exogenous variable	Diversity	Segregation
Proximity to first port of entry	– 6.266***	– 3.334
	(2.223)	(2.131)
Proximity to first port of entry (squared)	88.342***	23.314
	(21.757)	(18.995)
Proximity to first port of entry (cubic)	– 193.354***	– 42.539
	(46.826)	(37.848)
Proximity to first mill town	– 6.282	14.889***
	(5.160)	(5.072)
Proximity to first mill town (squared)	– 0.020	6.813**
	(4.482)	(3.333)
Proximity to first mill town (cubic)	– 0.125	– 3.821**
	(2.353)	(1.752)
Number of ports of entry in 20 km	0.390***	0.123
	(0.101)	(0.091)
Number of ports of entry in 50 km	0.018	– 0.032
	(0.026)	(0.025)
Number of ports of entry in 100 km	– 0.036***	0.007
	(0.008)	(0.009)
Number of mill towns in 20 km	– 0.019	– 0.012
	(0.013)	(0.017)
Number of mill towns in 50 km	0.008	0.014*
	(0.007)	(0.008)
Number of mill towns in 100 km	– 0.004**	0.006**
	(0.002)	(0.003)
Prox. to first port of entry*Num. of ports of entry in 20 km	– 5.237***	– 0.932
	(1.117)	(1.092)
Prox. to first port of entry*Num. of ports of entry in 50 km	– 1.590***	0.228
	(0.274)	(0.422)
Prox. to first port of entry*Num. of ports of entry in 100 km	1.206***	0.097
	(0.135)	(0.148)

Table B4. *Continued*

Exogenous variable	Diversity	Segregation
Prox. to first mill town*Num. of mill towns in 20 km	0.067	0.130
	(0.057)	(0.095)
Prox. to first mill town*Num. of mill towns in 50 km	0.080	– 0.080
	(0.122)	(0.115)
Prox. to first mill town*Num. of mill towns in 100 km	0.259	– 0.900***
	(0.274)	(0.284)
N	14,244	14,244

Notes: Standard errors in parentheses, *p < 0.1, **p < 0.05, ***p < 0.01. Weighted data.

Table B5. *IV1 First Stages. Other Coefficients.*

Exogenous variable	Diversity	Segregation
Hours of sun	0.000	0.000
	(0.000)	(0.000)
Rainfalls (mm)	0.000	0.000
	(0.000)	(0.000)
Male	– 0.001	0.001
	(0.003)	(0.003)
Born in 1989	0.000	– 0.001
	(0.001)	(0.001)
Other Whites	0.078***	0.022**
	(0.016)	(0.009)
Indian	0.109***	0.077***
	(0.014)	(0.011)
Black Caribbean	0.136***	0.058***
	(0.015)	(0.012)
Other Mixed	0.088***	0.044***
	(0.015)	(0.008)
Black African	0.108***	0.031***
	(0.018)	(0.010)
Pakistani	0.121***	0.109***
	(0.019)	(0.014)
Bangladeshi	0.122***	0.046**
	(0.020)	(0.019)
Chinese	0.030*	0.033**
	(0.017)	(0.017)
Second income quintile	– 0.005*	– 0.003
	(0.003)	(0.003)
Third income quintile	– 0.005	– 0.006
	(0.004)	(0.004)
Fourth income quintile	– 0.018***	– 0.015***
	(0.004)	(0.004)
Top income quintile	– 0.009**	– 0.012***
	(0.004)	(0.004)
English not mother tongue	0.023**	0.011**
	(0.009)	(0.006)
Female main parent	– 0.001	– 0.001
	(0.003)	(0.003)
Mother's education: GCSE or below	– 0.010***	– 0.009**
	(0.003)	(0.003)
Mother's education: above GCSE	– 0.011***	– 0.012***
	(0.004)	(0.004)
Mother's education: degree	– 0.000	– 0.005
	(0.006)	(0.006)
Mother's education: missing	– 0.003	– 0.011
	(0.010)	(0.009)
Mother's age: 35 to 49	0.001	– 0.001
	(0.003)	(0.003)

Table B5. *Continued*

Exogenous variable	Diversity	Segregation
Mother's age: 50+	– 0.003	– 0.006
	(0.005)	(0.004)
Mother's age: missing	– 0.021	– 0.005
	(0.015)	(0.014)
Mother's employment status: part time	– 0.003	– 0.004
	(0.003)	(0.003)
Mother's employment status: not working	0.001	– 0.002
	(0.003)	(0.003)
Mother's employment status: missing	0.012	0.001
	(0.017)	(0.014)
Constant	0.276***	0.176***
	(0.045)	(0.051)
N	14,244	14,244

Notes: Standard errors in parentheses, *p < 0.1, **p < 0.05, ***p < 0.01. Each LSYPE respondent is assigned the weather information collected from the closest meteo station in the interview month. Reference categories: female, born in 1990, White British, bottom income quintile, mother tongue is English, mother with no education, mother younger than 35 years old, mother works full time. Weighted data.

Table B6. *IV2, IV3, IV4: Instruments Relevance and Exogeneity.*

Test	*Test statistics*	*p-value*
	IV2	
Underidentification		
Kleibergen-Paap Ch-sq	48.356	0.003
Sanderson-Windmeijer Ch-sq (Diversity)	629.842	0.000
Sanderson-Windmeijer Ch-sq (Segregation)	1,316.944	0.000
F tests		
Kleibergen-Paap	25.346	
Sanderson-Windmeijer (Diversity)	25.077	
Sanderson-Windmeijer (Segregation)	54.228	
Overidentification (2SLS)		
Sargan-Hansen (PA)	29.068	0.218
Sargan-Hansen (HA)	20.520	0.667
	IV3	
Underidentification		
Kleibergen-Paap Ch-sq	64.063	0.000
Sanderson-Windmeijer Ch-sq (Diversity)	592.760	0.000
Sanderson-Windmeijer Ch-sq (Segregation)	2,018.274	0.000
F tests		
Kleibergen-Paap	37.591	
Sanderson-Windmeijer (Diversity)	18.980	
Sanderson-Windmeijer (Segregation)	64.624	
Overidentification (2SLS)		
Sargan-Hansen (PA)	29.478	0.493
Sargan-Hansen (HA)	38.982	0.126

Table B6. *Continued*

Test	*Test statistics*	*p-value*
	IV4	
Underidentification		
Kleibergen-Paap Ch-sq	71.546	0.024
Sanderson-Windmeijer Ch-sq (Diversity)	67,974.672	0.000
Sanderson-Windmeijer Ch-sq (Segregation)	32,542.763	0.000
F tests		
Kleibergen-Paap	1,972.006	
Sanderson-Windmeijer (Diversity)	1,347.634	
Sanderson-Windmeijer (Segregation)	6,434.231	
Overidentification (2SLS)		
Sargan-Hansen (PA)	50.388	0.418
Sargan-Hansen (HA)	50.925	0.396

Notes: Tests are computed using the ivreg2 stata command by Baum *et al.* (2002). Instruments IV2: proximity to the closest and the second closest port of entry and its higher terms (up to the squares), number of ports in a radius of 20, 50, 100 km, and interactions between the proximity to the closest and the second closest port and the number of ports within 20, 50, 100 km. Instruments IV3: distances from all the ports of entry and mill towns. Instruments IV4: binary variables capturing the number of ports (mill towns) within a 20, 50,100 km radius.

Table B7. *IV2, IV3, IV4: Second Stages.*

	PA				*HA*			
	LPM	*2SLS*	*GMM*	*LIML*	*LPM*	*2SLS*	*GMM*	*LIML*
				IV2				
Diversity	– 0.065*	– 0.040	– 0.049	– 0.039	– 0.185***	– 0.278***	– 0.245***	– 0.278***
	(0.036)	(0.055)	(0.051)	(0.056)	(0.046)	(0.067)	(0.053)	(0.067)
Segregation	– 0.001	0.164*	0.113	0.166*	0.199***	0.472***	0.467***	0.472***
	(0.048)	(0.091)	(0.076)	(0.091)	(0.049)	(0.086)	(0.044)	(0.086)
N	14,244	14,244	14,244	14,244	14,244	14,244	14,244	14,244
				IV3				
Diversity	– 0.065*	– 0.014	– 0.010	– 0.013	– 0.185***	– 0.213***	– 0.192***	– 0.213***
	(0.036)	(0.053)	(0.045)	(0.053)	(0.046)	(0.066)	(0.055)	(0.066)
Segregation	– 0.001	0.116*	0.096	0.117*	0.199***	0.428***	0.414***	0.430***
	(0.048)	(0.084)	(0.059)	(0.084)	(0.049)	(0.071)	(0.047)	(0.071)
N	14,244	14,244	14,244	14,244	14,244	14,244	14,244	14,244
				IV4				
Diversity	– 0.065*	– 0.017*	– 0.100**	– 0.108*	– 0.185***	– 0.371***	– 0.373***	– 0.375***
	(0.036)	(0.062)	(0.039)	(0.063)	(0.046)	(0.085)	(0.067)	(0.087)
Segregation	– 0.001	0.027*	0.013	0.028*	0.199***	0.347***	0.362***	0.350***
	(0.048)	(0.075)	(0.020)	(0.076)	(0.049)	(0.074)	(0.032)	(0.075)
N	14,244	14,244	14,244	14,244	14,244	14,244	14,244	14,244

Notes: Standard errors in parentheses, *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Results are computed using the ivreg2 stata command by Baum *et al*. (2002). Individual controls: male, born in 1989, ethnicity dummies. Family controls: income quintiles, English not mother tongue, main parent is female, mother's education, age and employment status. Weather controls: total sunshine duration, total rainfall. Each LSYPE respondent is assigned the weather information collected from the closest meteo station in the interview month. Instruments IV2: proximity to the closest and the second closest port of entry and its higher terms (up to the squares), number of ports in a radius of 20, 50, 100 km, and interactions between the proximity to the closest and the second closed port and the number of ports within 20, 50, 100 km. Instruments IV3: distances from all the ports of entry and mill towns. Instruments IV4: binary variables capturing the number of ports (mill towns) within a 20, 50,100 km radius. Weighted data.

Table B8. *RS with IV1 Instruments: Instruments Relevance and Exogeneity.*

Test	Test statistics	p-value
Underidentification		
Kleibergen-Paap Ch-sq	31.341	0.018
Sanderson-Windmeijer Ch-sq (Diversity)	161.184	0.000
Sanderson-Windmeijer Ch-sq (Segregation)	271.696	0.000
F tests		
Kleibergen-Paap	15.528	
Sanderson-Windmeijer (Diversity)	9.452	
Sanderson-Windmeijer (Segregation)	15.798	
Overidentification (2SLS)		
Sargan-Hansen (PA)	21.639	0.155
Sargan-Hansen (HA)	24.968	0.070

Notes: Tests are computed using the ivreg2 stata command by Baum *et al.* (2002). Instruments IV1: proximity to the closest port of entry and its higher terms (up to the third), number of ports in a radius of 20, 50, 100 km, and interactions between the proximity to the closest port and the number of ports within 20, 50, 100 km.

Table B9. *RS with County FE and IV1 Instruments: Instruments Relevance and Exogeneity.*

Test	Test statistics	p-value
Underidentification		
Kleibergen-Paap Ch-sq	35.633	0.005
Sanderson-Windmeijer Ch-sq (Diversity)	188.364	0.000
Sanderson-Windmeijer Ch-sq (Segregation)	98.363	0.000
F tests		
Kleibergen-Paap	5.319	
Sanderson-Windmeijer (Diversity)	11.092	
Sanderson-Windmeijer (Segregation)	5.169	
Overidentification (2SLS)		
Sargan-Hansen (PA)	16.913	0.391
Sargan-Hansen (HA)	16.625	0.410

Notes: Tests are computed using the ivreg2 stata command by Baum *et al.* (2002). Instruments IV1: proximity to the closest port of entry and its higher terms (up to the third), number of ports in a radius of 20, 50, 100 km, and interactions between the proximity to the closest port and the number of ports within 20, 50, 100 km.

Table B10. *Heteroscedasticity Constructed Instruments (I). Instruments Relevance and Exogeneity.*

	Parsimonious		Richer	
Test	*Test statistics*	*p-value*	*Test statistics*	*p-value*
Underidentification				
Kleibergen-Paap Ch-sq	84.038	0.007	120.337	0.000
Sanderson-Windmeijer Ch-sq (Diversity)	1049.95	0.000	4920.08	0.000
Sanderson-Windmeijer Ch-sq (Segregation)	832.44	0.000	5065.84	0.000
F tests				
Kleibergen-Paap	15.71		47.89	
Sanderson-Windmeijer (Diversity)	18.92		66.53	
Sanderson-Windmeijer (Segregation)	15.00		68.50	
Overidentification				
Sargan-Hansen (PA)	54.899	0.440	74.136	0.408
Sargan-Hansen (HA)	71.907	0.052	81.190	0.215
Heteroscedasticity				
Breusch-Pagan (Diversity)	851.82	0.000	37.61	0.000
Breusch-Pagan (Segregation)	76.03	0.000	26.84	0.000

Notes: Tests are computed using the ivreg2h stata command by Baum and Schaffer (2012).

Table B11. *Heteroscedasticity Constructed Instruments (II). Instruments Relevance and Exogeneity.*

	Richer with county FE	
Test	*Test statistics*	*p-value*
Underidentification		
Kleibergen-Paap Ch-sq	178.26	0.018
Sanderson-Windmeijer Ch-sq (Diversity)	15872.68	0.000
Sanderson-Windmeijer Ch-sq (Segregation)	25930.22	0.000
F tests		
Kleibergen-Paap	129.12	
Sanderson-Windmeijer (Diversity)	110.17	
Sanderson-Windmeijer (Segregation)	179.98	
Overidentification		
Sargan-Hansen (PA)	146.225	0.342
Sargan-Hansen (HA)	150.093	0.265
Heteroscedasticity		
Breusch-Pagan (Diversity)	3.17	0.075
Breusch-Pagan (Segregation)	208.73	0.000

Notes: Tests are computed using the ivreg2h stata command by Baum and Schaffer (2012).

Table B12. *IV1: Second Stages for HC (Coefficients of Interest).*

	LPM	2SLS	GMM	LIML
Diversity	– 0.004	– 0.048	– 0.117*	– 0.050
	(0.046)	(0.078)	(0.068)	(0.079)
Segregation	– 0.118*	– 0.345***	– 0.405***	– 0.351***
	(0.061)	(0.120)	(0.111)	(0.123)
N	14,244	14,244	14,244	14,244

Notes: Standard errors in parentheses, *p < 0.1, **p < 0.05, ***p < 0.01. Results are computed using the ivreg2 stata command by Baum *et al.* (2002). Individual controls: male, born in 1989, ethnicity dummies. Family controls: income quintiles, English not mother tongue, main parent is female, mother's education, age and employment status. Weather controls: total sunshine duration, total rainfall. Each LSYPE respondent is assigned the weather information collected from the closest meteo station in the interview month. Instruments IV1: proximity to the closest port of entry and its higher terms (up to the third), number of ports in a radius of 20, 50, 100 km, and interactions between the proximity to the closest port and the number of ports within 20, 50, 100 km. Weighted data.

Table B13. *HC with IV1 Instruments: Instruments Relevance and Exogeneity.*

Test	Test statistics	p-value
Underidentification		
Kleibergen-Paap Ch-sq	41.396	0.001
Sanderson-Windmeijer Ch-sq (Diversity)	377.005	0.000
Sanderson-Windmeijer Ch-sq (Segregation)	2492.583	0.000
F tests		
Kleibergen-Paap	28.507	
Sanderson-Windmeijer (Diversity)	23.456	
Sanderson-Windmeijer (Segregation)	147.868	
Overidentification (2SLS)		
Sargan-Hansen	19.553	0.241

Notes: Tests are computed using the ivreg2 stata command by Baum et al (2002). Instruments IV1: proximity to the closest port of entry and its higher terms (up to the third), number of ports in a radius of 20, 50, 100 km, and interactions between the proximity to the closest port and the number of ports within 20, 50, 100 km.

Appendix C: Proofs

Proof of Proposition 1

PROOF. (a) Call $\bar{n}_{g1}$ the number of Greens in ward 1 in the perfect integration scenario, where $D = 0$. Call D_n the level of segregation obtained by moving $n = 0, \dots \bar{n}_{g1}$ Greens from ward 2 to ward 1, starting from the perfect integration scenario.

$D_1 > D_0$ given:

$$D_1 = \frac{1}{2}\left(\left|\frac{\bar{n}_{g1}+1}{N_g} - \frac{n_{p1}}{N_p}\right| + \left|\frac{\bar{n}_{g1}-1}{N_g} - \frac{n_{p1}}{N_p}\right|\right) > 0 = D_0,$$

because $\frac{\bar{n}_{g1}+1}{N_g} \neq \frac{n_{p1}}{N_p}$ and $\frac{\bar{n}_{g1}-1}{N_g} \neq \frac{n_{p1}}{N_p}$. In general $D_n > D_{n-1}\ \forall n$

$$D_n = \frac{1}{2}\left[\left|\left(\frac{\bar{n}_{g1}+n}{N_g}\right) - \frac{n_{p1}}{N_p}\right| + \left|\left(\frac{\bar{n}_{g1}-n}{N_g}\right) - \frac{n_{p1}}{N_p}\right|\right] >$$

$$\frac{1}{2}\left[\left|\left(\frac{\bar{n}_{g1}+(n-1)}{N_g}\right) - \frac{n_{p1}}{N_p}\right| + \left|\left(\frac{\bar{n}_{g1}-(n-1)}{N_g}\right) - \frac{n_{p1}}{N_p}\right|\right] = D_{n-1}$$

because $\frac{\bar{n}_{g1}+n}{N_g} > \frac{\bar{n}_{g1}+(n-1)}{N_g} > \frac{n_{p1}}{N_p}$ and $\frac{\bar{n}_{g1}-n}{N_g} < \frac{\bar{n}_{g1}-(n-1)}{N_g} < \frac{n_{p1}}{N_p}$. Notice that $\lim_{n\to\bar{n}_{g1}} D_n = \frac{1}{2}$.

Given part (a) holds, we can prove part (b) by analysing the effect of an increase in district ethnic segregation obtained when Greens are moved from ward 2 to 1. Consider the two activities in isolation. The effect is weakly positive for *HA* and null for *PA*. The partial derivatives of (1), (2), (7), (8), (5) and (6) with respect to n_{g1} are:

$$\frac{\partial EU_{g1i}(HA)}{\partial n_{g1}} = \frac{(1-\tau_{HA})n_{p1}}{(n_{g1}+n_{p1})^2} \geq 0 \qquad \text{Since } 0 \leq \tau_{HA} \leq 1,$$

$$\frac{\partial EU_{p1i}(HA)}{\partial n_{g1}} = \frac{-(1-\tau_{HA})n_{p1}}{(n_{g1}+n_{p1})^2} \leq 0 \qquad \text{Since } 0 \leq \tau_{HA} \leq 1,$$

$$\frac{\partial EU_{g2i}(HA)}{\partial n_{g1}} = \frac{-(1-\tau_{HA})n_{g1}}{(N_g-n_{g1}+n_{p1})^2} \leq 0 \qquad \text{Since } 0 \leq \tau_{HA} \leq 1,$$

$$\frac{\partial EU_{p2i}(HA)}{\partial n_{g1}} = \frac{(1-\tau_{HA})n_{g1}}{(N_g-n_{g1}+n_{p1})^2} \geq 0 \qquad \text{Since } 0 \leq \tau_{HA} \leq 1.$$

Greens in ward 1 and Purples in ward 2 have now increased incentives to do *HA*. Purples in ward 1 and Greens in ward 2 have now decreased incentives to do *HA*. As $\bar{n}_{g1}+n+n_{p1} > \bar{n}_{g1}-n+n_{p1}, \forall n = 1, \dots, \bar{n}_{g1}-1$, at the aggregate level, this translates into an increase in the total incentives of doing *HA*. Since district ethnic diversity remains unaffected:

$$\frac{\partial EU_{g1}(PA)}{\partial n_{g1}} = \frac{\partial EU_{g2}(PA)}{\partial n_{g1}} = \frac{\partial EU_{p1}(PA)}{\partial n_{g1}} = \frac{\partial EU_{p1}(PA)}{\partial n_{g1}} = 0,$$

and thus the incentives of doing *PA* remain unaffected.

(c) Assume now that young people can switch between activities. Notice that for $\bar{n}_{g1}+n+n_{p1}$ young people $\frac{\partial EU(HA)}{\partial n_{g1}} \geq \frac{\partial EU(PA)}{\partial n_{g1}}$, and for $\bar{n}_{g1}-n+n_{p1}$ young people $\frac{\partial EU(HA)}{\partial n_{g1}} \leq \frac{\partial EU(PA)}{\partial n_{g1}}$. Since $\bar{n}_{g1}+n+n_{p1} > \bar{n}_{g1}-n+n_{p1}$, $\forall n = 1, \dots, \bar{n}_{g1}-1$, more people have incentive to switch from *PA* to *HA*. □

Proof of Proposition 2

PROOF. (a) Notice that $\lim_{N_g \to \infty} F = 1 - \left(\frac{N_g}{N_g+N_p}\right)^2 - \left(\frac{N_p}{N_g+N_p}\right)^2 = 0$.

(b), (c). For a given N_g we need to find the values n_{g1} and n_{g2} that keep district ethnic segregation equal to $\bar{D}$ $\forall N_g$. Since $n_{g2} = N_g - n_{g1}$, we need to find n_{g1} that solves:

$$D = \frac{1}{2}\left(\left|\frac{n_{g1}}{N_g} - \frac{n_{p1}}{N_p}\right| + \left|\frac{N_g - n_{g1}}{N_g} - \frac{n_{p1}}{N_p}\right|\right) = \bar{D}, \tag{C1}$$

The two solutions of (C1) are:

$$n_{g1} = \left(\frac{1}{2} + \bar{D}\right)N_g \quad \text{and consequently} \quad n_{g2} = \left(\frac{1}{2} - \bar{D}\right)N_g, \tag{C2}$$

$$n_{g1} = \left(\frac{1}{2} - \bar{D}\right)N_g \quad \text{and consequently} \quad n_{g2} = \left(\frac{1}{2} + \bar{D}\right)N_g. \tag{C3}$$

The results (C2) and (C3) are symmetric and thus we focus on (C2) only, where $n_{g1} \geq n_{g2}$.

To study the effect of district ethnic diversity, we compute the derivative with respect to N_g of the utility functions (1), (2), (7), (8), (5) and (6). By substituting the value of n_{g1} obtained in (C2), we get:

$$\frac{\partial EU_{g1i}(HA)}{\partial N_g} = -\frac{2(1+2\bar{D})(-1+\tau_{HA})n_{p1}}{(N_g + 2\bar{D}N_g + 2n_{p1})^2} \geq 0, \tag{C4}$$

$$\frac{\partial EU_{g2i}(HA)}{\partial N_g} = \frac{2(-1+2\bar{D})(-1+\tau_{HA})n_{p1}}{(N_g + 2\bar{D}N_g + 2n_{p1})^2} \geq 0 \quad \forall \bar{D} \leq \frac{1}{2}, \tag{C5}$$

$$\frac{\partial EU_{p1i}(HA)}{\partial N_g} = \frac{2(1+2\bar{D})(-1+\tau_{HA})n_{p1}}{(N_g + 2\bar{D}N_g + 2n_{p1})^2} \leq 0, \tag{C6}$$

$$\frac{\partial EU_{p2i}(HA)}{\partial N_g} = -\frac{2(-1+2\bar{D})(-1+\tau_{HA})n_{p1}}{(N_g + 2\bar{D}N_g + 2n_{p1})^2} \leq 0 \quad \forall \bar{D} \leq \frac{1}{2}, \tag{C7}$$

$$\frac{\partial EU_{g1i}(PA)}{\partial N_g} = \frac{\partial U_{g2}(PA)}{\partial N_g} = -\frac{2(-1+\tau_{PA})n_{p1}}{(N_g + 2n_{p1})^2} \geq 0, \tag{C8}$$

$$\frac{\partial EU_{p1i}(PA)}{\partial N_p} = \frac{\partial U_{p2}(PA)}{\partial N_p} = \frac{2(-1+\tau_{PA})n_{p1}}{(N_g + 2n_{p1})^2} \leq 0. \tag{C9}$$

(C4), (C5), (C6), (C7), (C8) and (C9) suggest that, for a given level of district ethnic segregation $\bar{D}$, when district ethnic diversity decreases, the expected utility of both *HA* and *PA* derived by the Greens increases, while the expected utility of both *HA* and *PA* derived by the Purples decreases. As the Greens are the majority ($N_g > N_p$), at the aggregate level, a decrease in district ethnic diversity increases both the incentives of doing *HA* and the incentives of doing *PA*. The magnitude of such increases depends on the magnitude of τ_{HA} and τ_{PA}. Consider Greens in ward 1, the largest group. Call $\Delta\tau = \tau_{PA} - \tau_{HA}$. If $\Delta\tau \leq 0$, $\frac{\partial EU_{g1i}(HA)}{\partial N_g} < \frac{\partial EU_{g1i}(PA)}{\partial N_g}$ $\forall D > 0$ and $\forall N_g$. To see what happens when $\Delta\tau > 0$, we compute the limit of $\frac{\partial EU_{g1i}(PA)}{\partial N_g}$ and $\frac{\partial EU_{g1i}(HA)}{\partial N_g}$ when τ_{PA} approaches 1 and τ_{HA} approaches 0, respectively. These are $\lim_{\tau_{PA}\to 1} \frac{\partial EU_{g1i}(PA)}{\partial N_g} = 0$ and $\lim_{\tau_{HA}\to 0} \frac{\partial EU_{g1i}(HA)}{\partial N_g} = \frac{2(1+2\bar{D})n_{p1}}{(N_g+2\bar{D}N_g+2n_{p1})^2} \geq 0$. Therefore, $\exists \bar{\Delta}_{g1}\tau(\tau_{HA}, \tau_{PA}) > 0$: if $\Delta\tau > \bar{\Delta}_{g1}\tau(\tau_{HA}, \tau_{PA})$ then $\frac{\partial EU_{g1i}(HA)}{\partial N_g} > \frac{\partial EU_{g1i}(PA)}{\partial N_g}$ $\forall D, N_g$. As a consequence, at the aggregate level, $\exists \bar{\Delta}\tau(\tau_{HA}, \tau_{PA}) > 0$: if $\Delta\tau > \bar{\Delta}\tau(\tau_{HA}, \tau_{PA})$ the incentives of *HA* increase more than the incentives of *PA*. □

References

Alesina, A., Baqir, R. and Easterly, W. (1999). 'Public goods and ethnic divisions', *Quarterly Journal of Economics*, vol. 114(4), pp. 1243–84.

Alesina, A., Baqir, R. and Hoxby, C. (2004). 'Political jurisdictions in heterogeneous communities', *Journal of Political Economy*, vol. 112(2), pp. 348–96.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S. and Wacziarg, R. (2003). 'Fractionalization', *Journal of Economic Growth*, vol. 8(2), pp. 155–94.

Alesina, A., Gennaioli, C. and Lovo, S. (2015). *Public Goods and Ethnic Diversity: Evidence from Deforestation in Indonesia*, Grantham Research Institute on Climate Change and the Environment.

Alesina, A. and La Ferrara, E. (2000). 'Participation in heterogeneous communities', *Quarterly Journal of Economics*, vol. 115(3), pp. 847–904.

Alesina, A. and La Ferrara, E. (2002). 'Who trusts others?', *Journal of Public Economics*, vol. 85(2), pp. 207–34.

Alesina, A. and La Ferrara, E. (2005). 'Ethnic diversity and economic performance', *Journal of Economic Literature*, vol. 43(3), pp. 762–800.

Algan, Y., Hémet, C. and Laitin, D.D. (2016). 'The social effects of ethnic diversity at the local level: A natural experiment with exogenous residential allocation', *Journal of Political Economy*, vol. 124(3), pp. 696–733.

Allen, R. (2009). *The British Industrial Revolution in Global Perspective*, Cambridge: Cambridge University Press.

Allen, R., Burgess, S., Davidson, R. and Windmeijer, F. (2015). 'More reliable inference for the dissimilarity index of segregation', *The Econometrics Journal*, vol. 18(1), pp. 40–66.

Altonji, J., Elder, T. and Taber, C. (2005). 'Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools', *Journal of Political Economy*, vol. 113(1), pp. 151–84.

Anderson, C. and Paskeviciute, A. (2006). 'How ethnic and linguistic heterogeneity influence the prospects for civil society: A comparative study of citizenship behavior', *Journal of Politics*, vol. 68(4), pp. 783–802.

Angrist, J. and Pischke, J. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, NJ: Princeton University Press.

Antunes, G. and Gaitz, G. (1975). 'Ethicity and participation: A study of Mexican-Americans, blacks and whites', *American Journal of Sociology*, vol. 80(5), pp. 1192–1211.

Ashraf, Q. and Galor, O. (2011). 'Dynamics and Stagnation in the Malthusian Epoch', *American Economic Review*., vol. 101(5), pp. 2003–41.

Ashraf, Q. and Galor, O. (2013). 'The 'out of Africa' hypothesis, human genetic diversity, and comparative economic development', *American Economic Review*, vol. 103(1), pp. 1–46.

Barron, J.M., Ewing, B.T. and Waddell, G.R. (2000). 'The effects of high school athletic participation on education and labor market outcomes', *Review of Economics and Statistics*, vol. 82(3), pp. 409–21.

Baum, C. and Schaffer, M. (2003). 'Instrumental variables and gmm: Estimation and testing', *Stata Journal*, vol. 3(1), pp. 1–31.

Baum, C.F. and Schaffer, M.E. (2012). *IVREG2H: Stata Module to Perform Instrumental Variables Estimation Using Heteroskedasticity-based Instruments*, Boston: Statistical Software Components, Boston College Department of Economics.

Baum, C., Schaffer, M.E. and Stillman, S. (2002). *Ivreg2: Stata Module for Extended Instrumental Variables/2sls and gmm Estimation*, Boston: Statistical Software Components, Boston College Department of Economics.

Bennett, R. (2012). 'Urban population database, 1801–1911', C.M. Law, B. Robson and J. Langton, (original data producers). Colchester, Essex: UK Data Archive (distributor).

Berthoud, R., Fumagalli, L., Lynn, P. and Platt, L. (2009). *Design of the Understanding Society Ethnic Minority Boost Sample*, Understanding Society at the Institute for Social and Economic Research.

Böheim, R., Horvath, T. and Mayr, K. (2014). *Birthplace Diversity and Productivity Spill-Overs in Firms*, University of Vienna, Department of Economics.

Bonomi Bezzo, F. (2017). *The Role of the Neighbourhood on Young People's Happiness, Self-perception and Social Interactions*, Mimeo, University of Essex.

Broh, B. (2002). 'Linking extracurricular programming to academic achievement: Who benefits and why?', *Sociology of Education*, vol. 75(1), pp. 69–95.

Cantle, T. (2001). *Community Cohesion: A Report of the Independent Review Team Chaired by Ted Cantle*, Home Office.

Carrington, W. and Troske, K. (1997). 'On measuring segregation in samples with small units', *Journal of Business & Economic Statistics*, vol. 15(4), pp. 402–09.

Caruso, R. (2011). 'Crime and sport participation: Evidence from Italian regions over the period 1997–2003', *Journal of Socio-Economics*, vol. 40(5), pp. 455–63.

Charles, K.K. and Kline, P. (2006). 'Relational costs and the production of social capital: Evidence from carpooling', Economic Journal, vol. 116(511), pp. 581–604.

Costa, D.L. and Kahn, M.E. (2003). 'Cowards and heroes: Group loyalty in the American civil war', *Quarterly Journal of Economics*, vol. 118(2), pp. 519–48.

Cunha, F. and Heckman, J. (2008). 'Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation', *Journal of Human Resources*, vol. 53(4), pp. 738–82.

Cunha, F., Heckman, J. and Schennach, S. (2010). 'Estimating the technology of cognitive and noncognitive skill formation', *Econometrica*, vol. 78(3), pp. 883–931.

Currarini, S., Jackson, M.O. and Pin, P. (2009). 'An economic model of friendship: Homophily, minorities, and segregation', *Econometrica*, vol. 77(4), pp. 1003–45.

Cutler, D.M. and Gleaser, E.L. (1997). 'Are ghettos good or bad?', *Quarterly Journal of Economics*, vol. 112(3), pp. 827–72.

Cutler, D., Glaeser, E. and Vigdor, J. (1999). 'The rise and decline of the American ghetto', *Journal of Political Economy*, vol. 107(3), pp. 455–506.

Czermak, S., Feri, F., Glätzle-Rützler, D. and Sutter, M. (2016). 'How strategic are children and adolescents? Experimental evidence from normal-form games', *Journal of Economic Behavior & Organization*, vol. 128(August), pp. 265–85.

DCLG. (2010). *Local Authority Housing Statistics, England, 2009-10*, Department for Communities and Local Government.

Del Boca, D., Monfardini, C. and Nicoletti, C. (2017). 'Self investments of adolescents and their cognitive development', *Journal of Labor Economics*, vol. 35(2), pp. 565–608.

Delhey, J. and Newton, K. (2005). 'Predicting cross-national levels of social trust: Global pattern or Nordic exceptionalism?', *European Sociological Review*, vol. 21(4), pp. 311–27.

Deming, D.J. (2017). 'The growing importance of social skills in the labor market', *Quarterly Journal of Economics*, vol. 132(4), pp. 1593–1640.

DfES (2003). *Every Child Matters*, Department for Education and Skills.

DfES (2006). *Positive Activities for Young People: National Evaluation*, Department for Education and Skills.

Duncan, O. and Duncan, B. (1955). 'A methodological analysis of segregation indexes', *American Sociological Review*, vol. 20(2), pp. 210–17.

Durlauf, S. and Fafchamps, M. (2005). *Handbook of Economic Growth*, chap. Social capital, Elsevier, pp. 1639–99.

Easterlin, R. (1967). 'Effects of population growth on the economic development of developing countries', *The ANNALS of the American Academy of Political and Social Science*, vol. 369(1), pp. 98–108.

Easterly, W. and Levine, R. (1997). 'Africa's growth tragedy: Policies and ethnic division', *Quarterly Journal of Economics*, vol. 112(4), pp. 1203–50.

Eide, E.R. and Ronan, N. (2001). 'Is participation in high school athletics an investment or a consumption good? Evidence from high school and beyond', *Economics of Education Review*, vol. 20(5), pp. 431–42.

Emran, M.S. and Hou, Z. (2013). 'Access to markets and rural poverty: Evidence from household consumption in China', *Review of Economics and Statistics*, vol. 95(2), pp. 682–97.

Felfe, C., Lechner, M. and Steinmayr, A. (2016). 'Sports and child development', *PLOS ONE*, vol. 11(5), pp. 1–23.

Fieldhouse, D. and Cutts, D. (2010). 'Does diversity damage social capital? A comparative study of neighbourhood giversity and social capital in the US and Britain', *Canadian Journal of Political Science*, vol. 43(2), pp. 289–318.

Folland, S. (2007). 'Does 'community social capital' contribute to population health?', *Social Science & Medicine*, vol. 64(11), pp. 2342–54.

Glaeser, E., Laibson, D. and Sacerdote, B. (2002). 'An economic approach to social capital', Economic Journal, vol. 112(483), pp. 437–58.

Glaeser, E., Laibson, D., Scheinkman, J. and Soutter, C. (2000). 'Measuring trust', *Quarterly Journal of Economics*, vol. 115(3), pp. 811–46.

Gratton, C., Shibli, S. and Coleman, R. (2005). 'Sport and economic regeneration in cities', *Urban Studies*, vol. 42(5-6), pp. 985–99.

Hansen, D., Larson, R. and Dworkin, J. (2003). 'What adolescents learn in organized youth activities: A survey of self-reported developmental experiences', *Journal of Research in Adolescence*, vol. 13(1), pp. 25–55.

Henderson, D.J., Olbrecht, A. and Polachek, S.W. (2006). 'Do former college athletes earn more at work? A nonparametric assessment', *Journal of Human Resources*, vol. 41(3), pp. 558–77.

Hirschman, A. (1964). 'The paternity of an index', *American Economic Review*, vol. 5(54), p. 761.

Holmes, C. (2001). *A tolerant country? Immigrants refugees and minorities in Britain*, Routledge.

Huber, M., Lechner, M. and Steinmayr, A. (2012). *RADIUSMATCH: Stata Module to Perform Distance-Weighted Radius Matching with Bias Adjustment*, Statistical Software Components, Boston College Department of Economics.

Huber, M., Lechner, M. and Steinmayr, A. (2015). 'Radius matching on the propensity score with bias adjustment: Tuning parameters and finite sample behaviour', *Empirical Economics*, vol. 49(1), pp. 1–31.

Huber, M., Lechner, M. and Wunsch, C. (2013). 'The performance of estimators based on the propensity score', *Journal of Econometrics*, vol. 175(1), pp. 1–21.

Jenkins, S.P., Micklewright, J. and Schnepf, S.V. (2008). 'Social segregation in secondary schools: How does England compare with other countries?', *Oxford Review of Education*, vol. 34(1), pp. 21–37.

Kalra, V. (2000). *From Textile Mills to Taxi Ranks: Experiences of Migration, Labour, and Social Change*, Ashgate.

Kleibergen, F. and Paap, R. (2006). 'Generalized reduced rank tests using the singular value decomposition', *Journal of Econometrics*, vol. 133(1), pp. 97–126.

Kleibergen, F. and Schaffer, M.E. (2007). *Ranktest: Stata Module to Test the Rank of a Matrix Using the Kleibergen-Paap RK Statistic*, Statistical Software Components, Boston College Department of Economics.

Knack, S. and Keefer, P. (1997). 'Does social capital have an economic payoff? a cross-country investigation', *Quarterly Journal of Economics*, vol. 112(4), pp. 1251–88.

Kuhn, P. and Weinberger, C. (2005). 'Leadership skills and wages', *Journal of Labor Economics*, vol. 23(3), pp. 395–436.

La Ferrara, E. and Mele, A. (2007). 'Racial segregation and public school expenditure', CEPR Discussion Paper 5750.

Lechner, M. (2009). 'Long-run labour market and health effects of individual sports activities', *Journal of Health Economics*, vol. 28(4), pp. 839–54.

Lechner, M. and Downward, P. (2017). 'Heterogeneous sports participation and labour market outcomes in England', *Applied Economics*, vol. 49(4), pp. 335–48.

Lechner, M., Miquel, R. and Wunsch, C. (2011). 'Long-run effects of public sector sponsored training in West Germany', *Journal of the European Economic Association*, vol. 9(4), p. 742.

Lechner, M. and Sari, N. (2015). 'Labor market effects of sports and exercise: Evidence from Canadian panel data', *Labour Economics*, vol. 35, pp. 1–15.

Leigh, A. (2006). 'Trust, inequality and ethnic heterogeneity', *Economic Record*, vol. 82(258), pp. 268–80.

Letki, N. (2008). 'Does diversity erode social cohesion? Social capital and race in British neighbourhoods', *Political Studies*, vol. 56(1), pp. 99–126.

Lewbel, A. (2012). 'Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models', *Journal of Business & Economic Statistics*, vol. 30(1), pp. 67–80.

Lipscomb, S. (2007). 'Secondary school extracurricular involvement and academic achievement: A fixed effects approach', *Economics of Education Review*, vol. 26(4), pp. 463–72.

Little, K. (2013). *Negroes in Britain: A Study of Racial Relations in English Society*, Routledge.

Long, J.E. and Caudill, S.B. (1991). 'The impact of participation in intercollegiate athletics on income and graduation', *Review of Economics and Statistics*, vol. 73(3), pp. 525–31.

Luttmer, E.F.P. (2001). 'Group loyalty and the taste for redistribution', *Journal of Political Economy*, vol. 109(3), pp. 500–28.

Mokyr, J. (1990). *The Lever of Riches: Technological Creativity and Economic Progress*, New York: Oxford UP.

Ottaviano, G. and Peri, G. (2006). 'The economic value of cultural diversity: Evidence from us cities', *Journal of Economic Geography*, vol. 6(1), pp. 9–44.

Parrotta, P., Pozzoli, D. and Pytlikova, M. (2014). 'The nexus between labor diversity and firm's innovation', *Journal of Population Economics*, vol. 27(2), pp. 303–64.

Pellew, J. (1989). 'The Home Office and the Aliens Act, 1905', *Historical Journal*, vol. 32(2), pp. 369–85.

Pennant, R. (2005). *Diversity, Trust, and Community Participation in England*, Home Office.

Pfeifer, C. and Cornelißen, T. (2010). 'The impact of participation in sports on educational attainment: New evidence from Germany', *Economics of Education Review*, vol. 29(1), pp. 94–103.

Platt, L. (2009). 'Social activity, social isolation and ethnicity', *Sociological Review*, vol. 57(4), pp. 670–702.

Putnam, R. (2007). ''Et pluribus unum': Diversity and community in the twenty-first century. The 2006 Johan Skytte Prize lecture', *Scandinavian Political Studies*, vol. 30(2), pp. 137–74.

Raghunathan, T., Lepkowski, J., Van Hoewyk, J. and Solenberger, P. (2001). 'A multivariate technique for multiply imputing missing values using a sequence of regression models', *Survey Methodology*, vol. 27(1), pp. 85–95.

Rathelot, R. (2012). 'Measuring segregation when units are small: A parametric approach', *Journal of Business & Economic Statistics*, vol. 30(4), pp. 546–53.

Rees, D.I. and Sabia, J.J. (2010). 'Sports participation and academic performance: Evidence from the national longitudinal study of adolescent health', *Economics of Education Review*, vol. 29(5), pp. 751–59.

Rocco, L., Fumagalli, E. and Suhrcke, M. (2014). 'From social capital to health and back', *Health Economics*, vol. 23(5), pp. 586–605.

Roodman, D. (2011). 'Fitting fully observed recursive mixed-process models with CMP', *Stata Journal*, vol. 11(2), pp. 159–206.

Rubin, D.B. (2001). 'Using propensity scores to help design observational studies: Application to the tobacco litigation', *Health Services and Outcomes Research Methodology*, vol. 2(3), pp. 169–88.

Sanderson, E. and Windmeijer, F. (2016). 'A weak instrument F-test in linear iv models with multiple endogenous variables', *Journal of Econometrics*, vol. 190(2), pp. 212–21.

Simpson, L. (2004). 'Statistics of racial segregation: Measures, evidence and policy', *Urban Studies*, vol. 41(3), pp. 661–81.

Stevenson, B. (2010). 'Beyond the classroom: Using Title IX to measure the return to high school sports', *Review of Economics and Statistics*, vol. 92(2), pp. 284–301.

Sturgis, P., Brunton-Smith, I., Read, S. and Allum, N. (2011). 'Does ethnic diversity erode trust? Putnam 'hunkering down' thesis reconsidered', *British Journal of Political Science*, vol. 41(1), pp. 57–82.

Swanton, D. (2010). 'Sorting bodies: Race, affect, and everyday multiculture in a mill town in Northern England', *Environment and Planning A*, vol. 42(10), pp. 2332–50.

Tabili, L. (1994). 'The construction of racial difference in twentieth century Britain: The Special Restriction (Coloured Alien Seamen) Order, 1925', *Journal of British Studies*, vol. 33(1), pp. 54–98.

Tiebout, C. (1956). 'A pure theory of local expenditures', *Journal of Political Economy*, vol. 64(5), pp. 416–24.

Uslaner, E. (2011). 'Trust, diversity, and segregation in the United States and the United Kingdom', *Comparative Sociology*, vol. 10(2), pp. 221–47, .

Uslaner, E. (2012). *Segregation and Mistrust: Diversity, Isolation, and Social Cohesion*, Cambridge University Press.

Vigdor, J.L. (2004). 'Community composition and collective action: Analyzing initial mail response to the 2000 census', *Review of Economics and Statistics*, vol. 86(1), pp. 303–12.

Weinberger, C.J. (2014). 'The increasing complementarity between cognitive and social skills', *Review of Economics and Statistics*, vol. 96(4), pp. 849–61.

Zak, P. and Knack, S. (2001). 'Trust and growth', Economic Journal, vol. 111(470), pp. 295–321.