The University Of Sheffield. | School Of Economics

The role of ethnic minority peers in fostering university aspirations among White pupils

Lucy Ward

Sheffield Economic Research Paper Series

SERPS no. 2024014

ISSN 1749-8368

28 November 2024

The role of ethnic minority peers in fostering university aspirations among White pupils*

Lucy Ward[1]

[1]School of Economics, University of Sheffield, Sheffield, S1 4DT, United Kingdom.

Contributing authors: lucy.ward@sheffield.ac.uk;

Abstract

Some theories suggest that encountering ethnic diversity broadens knowledge and perspectives, equipping individuals with cultural capital. This research examines whether high aspiring ethnic minority pupils influence the university aspirations of White pupils they interact with in the same school. We link White pupils' reported university likelihood to administrative data on all pupils in England. As an instrument for exposure to ethnic minorities in 2014, we use the proportion of employed individuals in a local area that are nurses in 1951—an indicator of post-World War II job shortages filled by immigrant workers. The findings show that increasing ethnic minorities in schools positively impacts White pupils' university aspirations. We provide some evidence of heterogeneous impacts across individual and family characteristics, suggesting that improving school diversity could aid in improving higher education participation for under-represented groups.

Keywords: Ethnic diversity; Peer effects; University; Instrumental variables

JEL Classification: J15 , I21 , C26

*Acknowledgements: I am grateful to Steven McIntosh, Bert Van Landeghem, Gurleen Popli, Nicolas Van de Sijpe, Jo Blanden and Panos Nanos for helpful comments. I also thank participants at the International Workshop on Applied Education Economics (IWAEE) (Italy, June 2024)

1 Introduction

This paper examines whether ethnic minority pupils influence the university aspirations of White pupils they interact with in the same school. In the UK, all ethnic minority groups are, on average, significantly more likely to attend university than White British pupils. Similarly in the United States, around 60% of young Black, Hispanic, and White Americans enrol in college, whereas the college enrolment rate for Asian Americans is notably higher, exceeding 80% (Barshay, 2023). These trends may be surprising, considering ethnic minorities tend to have more disadvantaged backgrounds than White individuals[1] (a characteristic associated with lower higher education participation). However, after adjusting for background characteristics and prior attainment, the findings remain.[2] Burgess (2014) was one of the first to attribute these differences to aspirations.

Modern societies are becoming increasingly ethnically diverse. Putnam (2007) found that neighbourhood diversity reduces social capital, supporting 'conflict' theory. In contrast, 'contact' theory suggests diversity reduces prejudice through positive interactions. However, since most of this research focuses on adults, questions remain about the effects of diversity on children, particularly within schools. Ethnic diversity in education can alter the learning environment and potentially affect pupils' skill development and long-term economic outcomes. Anelli et al (2023) suggest homogeneous groups may lower learning quality, while Iranzo et al (2008) suggest that diverse skill sets across ethnicities can foster cooperation. Additionally, ethnic diversity in classrooms may reshape pupils' self-assessment by shifting their reference groups (Elsner and Isphording (2017), Murphy and Weinhardt (2020)). Given these potential impacts, this research aims to understand how ethnic diversity in schools influences educational

[1] In the UK, over 60% of Bangladeshi individuals are in the lowest SES band (bottom 20%), as are around 45% of Black African and Black Caribbean individuals (Dearden and Sibieta, 2010)

[2] Black African pupils are almost 35 percentage points more likely to go to university than otherwise-identical White British pupils; most other ethnic minority groups are around 15-25 percentage points more likely to go than similar White British pupils (Crawford and Greaves, 2015).

aspirations, which could have broader implications for addressing inequalities in higher education participation.

The first contribution of this research lies in its focus on individuals' further education aspirations. Beliefs about future goals play a crucial role in shaping both educational and economic decision-making. Gorard et al (2012) present evidence of a positive relationship between aspirations and university participation, highlighting aspirations as a key pathway to higher education. However, the role that peers' ethnicity plays in shaping these aspirations remains largely unexplored, highlighting an important gap in the literature.

Most research on educational peer effects has focused primarily on academic attainment. The few studies that examine the impact on other student outcomes (De Giorgi et al (2009), Mora and Oreopoulos (2011), Mendolia et al (2018), Dickerson et al (2018), Gagete-Miranda (2022)) do not explore the effects of migrant, refugee status, or ethnicity. Findings also vary across countries, time periods, and outcome measures, making generalisability challenging. This underscores the need for targeted research on how ethnic diversity influences educational goals.

A second contribution of this research is the focus on the specific impact of ethnic minorities. Much of the existing literature has examined the effects of refugees or immigrants facing native language challenges (Figlio et al, 2024; Figlio and Özek, 2019; Morales, 2022; Chareyron et al, 2021). Despite using similar methodologies, these studies yield mixed findings. In contrast, language challenges are less relevant for many ethnic minority pupils, particularly those born and raised in the host country, allowing for a clearer examination of how ethnic diversity itself,rather than language barriers, affects educational goals. Causal evidence on the impact of peers' ethnicity remains limited, with most studies focused on the US, again showing mixed results. As countries become increasingly ethnically diverse, understanding how ethnic diversity influences the economy is more crucial than ever.

As ethnic minorities are not equally spread across the UK, there is a high degree of self-selection into areas. Parents with similar characteristics tend to select into the same areas. This self-selection into peer groups creates selection bias from the fact that an outcome we attribute to a peer effect is just a consequence of the fact that people who share similar characteristics make themselves into groups. To address this endogeneity, we create a unique instrument by leveraging historical job shortages post-World War II. Specifically, we use the distribution of nurses in 1951—a period marked by government-backed recruitment from Commonwealth countries to address labour shortages—as a predictor of ethnic minority presence in schools in 2014. This approach allows us to identify areas with historically higher proportions of ethnic minorities, providing an instrument for ethnic diversity exposure.

Our analysis uses linked data from the Millennium Cohort Study (MCS), a birth cohort dataset, and the National Pupil Database (NPD), which includes administrative records for all pupils in England, as well as using historical occupational data. We find that a higher proportion of ethnic minority peers has a positive and significant effect on White pupils' university aspirations at age 14. These effects are significant for both male and female pupils and larger for the most disadvantaged pupils.

The findings are robust across various checks, including changes to the sample definition for White and ethnic minority pupils, as well as controls for potential omitted variables such as parental aspirations and attitudes toward ethnic diversity. We also assess the impact of homophily in friendship formation, accounting for the closeness of connections with ethnic minority peers. Further robustness checks include weak instrument tests and sensitivity analyses with alternative instruments based on additional occupational data.

Recent geopolitical events, such as Brexit and the election of Donald Trump, have fueled pessimistic views toward the impacts of ethnic diversity. However, these results

suggest that diverse populations can have positive social impacts, contributing to a broader societal benefit through improved educational aspirations.

This paper is organised as follows. Sections 2 and 3 outline the data and methodology respectively. Section 4 presents the estimated impact and tests the robustness, before concluding in Section 5.

2 Data

The data used in this research combines the Millennium Cohort Study (MCS) with the National Pupil Database (NPD).[3]

The MCS is a multidisciplinary cohort survey run by the Centre for Longitudinal Studies at the University College London. The study is a valuable data source as it tracks the lives of a sample of about 19,000 babies born in the UK between 1st September 2000 and 11th January 2002.[4] The sample was constructed to be representative of the total UK population. The data collectors selected electoral wards with the aim to recruit 100 per cent of the children born in the eligible period within them.[5] They also wanted to adequately represent disadvantaged and ethnic minority children. The population of wards was therefore stratified by ethnicity and the Child Poverty Index.

The survey is conducted in several waves, the first occurred when children were aged nine months, gathering information from the parents of 18,818 children. Since then, families have been interviewed again six times at ages 3, 5, 7, 11, 14 and 17.[6] The survey was originally answered by just the parents. From wave 2 onwards, the child was also surveyed along with older siblings. Class teachers responded to a survey in waves 3-5. Early topics included parental and child health, parenting activities

[3] University College London and UCL Institute of Education and Centre for Longitudinal Studies and Department for Education (2023)

[4] 1 September 2000 and 31 August 2001 (for England and Wales), and between 24 November 2000 and 11 January 2002 (for Scotland and Northern Ireland).

[5] They achieved a response rate of 72 per cent of all the families with eligible children living at nine months in the sampled wards.

[6] Wave 8 at age 22, has been undertaken but the data has not been released.

and attitudes, physical, social and cognitive development of the child, preschool experiences, and leisure activities etc. As the cohort member aged, there was a larger focus on schooling, relationships, mental and physical health, wellbeing, aspirations and the future, identity and attitudes, alongside age-appropriate questions on more sensitive topics such as risky behaviours, including alcohol, smoking and drugs; antisocial and criminal activities and contact with the police. Many topics were covered in each wave including family composition, housing and local area, parental education, employment and income.

The main advantage of the MCS is the rich range of information regarding the experiences and outcomes of the MCS children and their families. This allows many individual and family characteristics to be controlled for. They also ask the same questions at multiple ages which allows for changes over time to be analysed. The main limitation of the survey is that the longitudinal pattern of response is complex, with attrition and re-entry. By age 17, 10,757 cohort members responded to the survey, a reduction of 8,061 from the original sample.

Plewis (2007) analyses factors influencing non-response between the first and second wave of the Millennium Cohort Study. They provide evidence to show that young mothers as well as breast-feeding mothers, respondents from minority groups, notably Black and 'other' minority ethnic groups, respondents with fewer educational qualifications, poorer families, living in rented accommodation and not in a house are more likely to leave the sample. This research controls for many of these factors and Plewis (2007) argues that although the cases lost from the sample were different from those that remained, they were not substantially different.

The second dataset used in this research is the National Pupil Database (NPD). The NPD is an individual-level administrative database controlled by the Department for Education in England.[7] The database combines information held by schools, exam

[7]The first version of the NPD was produced in 2002 and is mainly used for funding purposes, school performance tables, policy making, and research.

awarding bodies, and local authorities on all pupils, aged 2-18, in English state schools. The NPD is a valuable data resource as it provides a near-complete picture of school outcomes for the majority of children in England. The data is only limited by the fact that both children who attend private school and children who are home schooled are not included.

The NPD consists of a range of data sources that provide information on pupils' education attainment at different Key Stages. The main data source is the School Census which is carried out three times a year (January, May, and October). The School Census includes information on gender, ethnicity, first language, eligibility for free school meals, special educational needs, and any absences and exclusions. It also includes information at the school level such as the number of pupils within a school in each ethnic group, the proportion of pupils on free school meals and the average achievement of pupils. The NPD also reports information about pupils' test results at each Key Stage. This includes English, Mathematics, and Science scores in Key Stage 1, 2 and 3, aged 7 and 11 and 14. Achievement data in Key Stage 4, aged 16, includes GCSEs and equivalent qualifications.[8] The grades achieved in A levels and equivalent qualifications at age 16 in KS5 are also reported.

The MCS can be linked to the NPD using a pupil-level identifier. Up to age 16, the linkage is based on parental consent collected at various time points. At age 17, cohort members are asked for their own consent. At age 11, out of a possible 7,942 cohort members who participated in the survey in England, 7,508 gave consent to match the survey to their education data. 7,252 were successfully matched (Rihal, 2021). This linking is crucial for this study as it enables information on cohort members' school and achievement to be matched with detailed individual and parental responses to survey questions.

[8]Equivalent qualifications include vocational qualifications such as BTECs, which are taken by 16 year olds within schools.

2.1 Sample

In this research we use data from wave 6 of the MCS, which took place when the cohort members were age 14.[9] We restrict our sample to children living in England, as the NPD is only available for children who attend school in England. For our baseline model we also restrict the sample to White pupils as the research question specifically looks at the impact of ethnic minority peers on White pupils' perceived likelihood of attending university. Due to the limited sample size, we are unable to examine the impact of White pupils on ethnic minority pupils.

The sample covers 3,606 White pupils who report their likelihood of attending university at age 14 as well as having other information including parental education and income.[10] We also require that their survey data has been successfully matched to school level information on peers' ethnicity.[11]

2.2 Measure of university ambition

The outcome of interest in this research is the reported likelihood of attending university. We concentrate on this measure because previous literature has focused on the impact of ethnic minorities on the academic achievement of White pupils. It is well established that, on average, ethnic minority pupils have higher educational aspirations than their White counterparts. Consequently, it is worthwhile to explore whether these heightened aspirations can be transmitted to White pupils through their interactions with ethnic minority peers. As our outcome measure, we use a question asked to cohort members in wave 6 when they are 14. The cohort members were asked how likely it was, on a scale of 0-100%, that they would go to university.

[9] Wave 6 was carried out between January 2015 and April 2016 when the cohort members were on average aged 14.

[10] The Millennium Cohort Study is nationally representative. Whilst 5,309 White individuals were surveyed at age 14, only 5,117 English White pupils reported their university likelihood. The further reduction in the sample is due to missing data in the control variables. Despite being a relatively small sample, it is representative.

[11] We define White pupils as any pupil who self-identifies as belonging to the racial category commonly referred to as "White". The results are robust to restricting the sample to White British pupils only.

While the reported likelihood of attending university likely captures elements of both aspirations and expectations, we argue that any change in this variable driven by changes in exposure to ethnic minority peers predominantly reflects a shift in aspirations. Expectations tend to be more stable and are influenced by factors like socioeconomic status, academic performance, and family background, which help create a realistic picture of what outcomes are likely. On the other hand, aspirations are more flexible; they can shift based on new experiences, different influences, and the role models we encounter.

Ethnic minority peers, who are statistically more likely to aspire to and attend university, may serve as aspirational influences for White pupils. By observing these peers' educational goals and commitment to higher education, White pupils may develop heightened aspirations for themselves, independent of any immediate change in their objective circumstances or academic outcomes. Therefore, while the reported likelihood of attending university is a composite measure, any shift linked to changes in ethnic diversity is more likely to represent an aspirational change. This study thus uses perceived university likelihood as an aspiration measure, with the expectation that exposure to diverse peers primarily enhances pupils' aspirations and sense of possibility regarding higher education.

Whilst the question on university likelihood is also asked in wave 7 of the MCS when the members are 17, we do not examine the impact of peers' ethnicity on this measure. Pupils in England decide whether to take an academic or vocational route at age 16, this suggests that by age 17, the desire to attend university has already been determined.

2.3 Measure of peers' ethnicity

The NPD provides school level information on the number of pupils from different ethnic groups. We use this information to calculate the percentage of pupils within

a school who have a non-White ethnicity. We define a pupil as having a non-White ethnicity if they identify as Asian (Pakistani, Indian, Bangladeshi, other Asian and Chinese), Black (Black African and Black Caribbean), mixed raced (White and Black African and White and Asian) or other.[12]

The pupils in English secondary schools are grouped with different peers for different subjects and consequently they interact with a much larger group of students than in primary school. Despite this, we acknowledge that peer effects might be stronger from interactions with smaller groups such as peers in their year or close friends than from the overall school group, unfortunately the data does not provide information on peer ethnicity at the year group level or information on friendship groups.

To address this concern, we make two points. Firstly, the NPD reports the ethnic makeup of a school every year. The variation of non-White pupils within a school across years is small, indicating that the ethnic make-up of each year group within a school is similar. Secondly, as a robustness check, we make use of a question in the MCS asked to the cohort members about their friends. In wave 6, cohort members are asked "How many of your friends are from the same ethnic group as you?" We drop cohort members who report that all of their friends are from the same ethnic group. This means that every pupil left in the sample says they have some close connection to a peer who is of a different ethnicity. The results are robust to the change in the sample.

2.4 Conditioning variables

The MCS linked with the NPD provides us with rich background information on the cohort member, household and school attended. We include these in the model to control for factors that may impact upon an individual's reported likelihood of attending university and ethnicity of peers in order to identify the impact of peers'

[12] The results are robust to the removal the mixed ethnicity category from the calculations of the non-White peer group.

ethnicity on the individual's perceived likelihood of attending university. See Table 2 in Section 4.1 for the full summary statistics.

We consider the individual characteristics (gender, month of birth, KS2 score (a proxy for ability.); household characteristics (number of siblings, real weekly equivalent income[13] , and local deprivation index); main parent's characteristics (age and level of education and economic activity status)[14] ; and school characteristics (number of students enrolled in the school, average KS2 performance of the current KS4 pupils (a proxy for school quality), number of pupils on free school meals).

3 Method

3.1 OLS estimation

The data allows this research to observe university likelihood at age 14 as well as the ethnic makeup of the school attended. This analysis aims to identify a causal effect of non-White peers on White pupils' university likelihood. As previously discussed, identifying causality between peers' ethnicity and pupils' university likelihood faces methodological challenges. There are two main threats to causality: (i) omitted variable bias, and (ii) self-selection in to peer groups. Section 2 discussed the controls we include to minimise omitted variable bias. This section will outline how this research will try to correct for self-selection.

The analysis starts by estimating an ordinary least squares regression (OLS) in which we regress university likelihood of White pupils on the proportion of peers who have a non-White ethnicity and the covariates. The model takes the following form:

[13]The equivalent income is the income of the household taking into account the number of people in the family and assigning weights. The one provided in the MCS follows the OECD equivalence scale, which assigns a value of 1 to the first household member, of 0.7 to each additional adult, and of 0.5 to each child.

[14]The main parent is identified as the parent who responds to the survey about the cohort member. The estimated coefficients are robust to the inclusion of both parents characteristics.

$$UniversityLikelihood_i = \beta_0 + \beta_1 Non-WhitePeers_i + \beta_2 Child_i + \beta_3 Family_i + \beta_4 School_i + \epsilon_i \quad (1)$$

Where $UniversityLikelihood_i$ is the self-reported likelihood of attending university for individual i. $Non-WhitePeers_i$ is the proportion of pupils within the individual's school who have a non-White ethnicity.

The main interest of this research is the estimation of β_1, which is the effect of non-White peers on university likelihood. To interpret β_1 as the causal effect, we require independent variation in the proportion of non-White peers, meaning the zero conditional mean assumption must hold,$E(\epsilon_i|Child_i, Family_i, School_i) = 0$. Due to the endogeneity of the ethnic makeup of a school, it can be argued that this assumption may not hold.

The school the pupil attends is largely driven by parental location choice. Ethnic minorities are not spread equally across the country and tend to be very geographically concentrated. Individuals who attend a school with a high percentage of ethnic minorities may differ in both observable and unobservable ways to those who attend schools with low levels of ethnic minorities. These differences could be driven by child and parental factors. Parents with similar characteristics tend to select into the same areas. Dustmann and Preston (2001), using data from England, find that individuals who are more hostile towards ethnic minorities will not settle in neighbourhoods with a high ethnic concentration. Betts and Fairlie (2001) and more recently, Cascio and Lewis (2012) provide evidence in support of 'White flight', where White pupils move to private schools as a response to increasing levels of ethnic minorities. Ethnic minorities, specifically immigrants, are more likely to settle in areas with a lower-than-average level of education (Dustmann and Preston, 2001). However, ethnic minority parents tend to have high expectations for their children and therefore may select

higher-quality schools (d'Este and Einiö, 2021). Many of these factors may not be observable and may influence the child's reported likelihood of attending university, biasing the estimates. The direction of the bias is not straightforward. Take school quality for example. Whilst this is likely to increase the reported likelihood of attending university it is not clear whether this would increase or decrease the number of ethnic minority peers. Much of the literature to date has suggested a downward bias mainly driven by high-ability individuals choosing not to live in areas with many ethnic minorities (Dustmann and Preston, 2001). In this case we can interpret the OLS results as a lower bound of the association between ethnic minority peers and the perceived likelihood of attending university.

3.2 Instrumental variable

To check whether the estimated effect is subject to bias, we exploit exogenous variation in peers' ethnicity by using an instrumental variable. We argue that the current geographical distribution of ethnic minorities in England is driven by labour shortages post World War 2 and government backed requirements specifically in the health care industry. Appendix Figure A1 provides a map of the geographical distribution of ethnic minorities in England.

We choose the percentage of individuals employed as nurses in 1951 in the Local Government District (LGD) as our instrument for two main reasons. Firstly, it was a government backed recruitment drive where immigrants were placed in the areas with the shortages. Secondly, nurses were needed across the country whereas some of the other professions were geographically concentrated. Once individuals arrived in Britain, they were dispersed to their appointed hospitals all over the United Kingdom where they lived in the Nurses' Homes attached to the hospitals.

The 1948 British Nationality Act said that all Commonwealth citizens could have British passports and work in the UK.[15] Britain encouraged mass immigration

[15] See Home Office (2020) for an overview of the history of nationality law.

from the Commonwealth countries after the Second World War due to severe labour shortages, especially in the newly created National Health Service (NHS).[16]

Hospitals in Britain were dealing with labour shortages well before the establishment of the NHS and nurse shortages had been discussed in several government inquiries.[17] The national post-war labour shortage had only made the problem worse. By 1948, there were 54,000 nursing vacancies (Snow and Jones, 2011). In 1949 the Ministries of Health and Labour, along with other healthcare industry representatives, including the Colonial Office, the General Nursing Council and the Royal College of Nursing launched campaigns to recruit hospital staff directly from the West Indies (NHS, 2023). Senior NHS staff from Britain travelled to the West Indies to recruit, and vacancies were published in local papers. By 1955 there were official nursing recruitment programmes across 16 British colonies and former colonies (Snow and Jones, 2011). The NHS, which became Britain's biggest employer in 1961, hired thousands of Commonwealth workers: by 1968, around a third of student nurse and midwife roles were filled by Commonwealth migrants (Babikian, 2021).

Until 1986, there were two nurse training programmes: State Registered Nurse (SRN) qualification and State Enrolled Nurse (SEN) qualification. The State Registered Nurse (SRN) and State Enrolled Nurse (SEN) qualifications represented different levels of nursing training and responsibilities. The SRN qualification required three years of training, focused on medical and surgical nursing skills as well as taking on leadership roles. In contrast, the SEN qualification was a two year course covering the fundamentals of nursing care with limited clinical responsibilities. Most individuals arriving from the Commonwealth, were placed on the SEN course. It has been suggested that few were accepted on the SRN course despite possessing the required qualifications, due to racial discrimination.

[16] The NHS was established in 1948.
[17] See Parliament. House of Commons (1942) where Mr. Ernest Brown the Minster of Health discusses the nurse shortages.

In 1951, England was still faced with large labour shortages. We therefore assume that areas that had relatively low levels of nurses in 1951 were the areas with the highest shortages. In the following years, these shortages were filled by individuals from the commonwealth. We argue that it is these areas that have higher numbers of ethnic minorities today. This is due to that fact that in the following years, there was growing public and political unease regarding the impact of migration, leading to numerous changes to the 1948 British Nationality Act. These changes focused on the entry of dependents and family members of those already in the United Kingdom meaning that individuals entering the country in later years normally located in areas where their family members already were. Additionally, Britton et al (2021) provides evidence to show that ethnic minorities are less likely to move areas, and the effect of higher education on mobility is much weaker. After their nurse training, many individuals stayed where they had been placed. Most of them could not get onto the SRN course, and therefore could not get promoted. Many felt unable to return to their home country as the SEN qualification was not recognised in the Caribbean. In addition, it was very difficult for them to move to other places in the United Kingdom due to difficulty finding accommodation. The infamous 'No Irish, no Blacks and no dogs' signs have become symbolic of the wave of xenophobic sentiment that arose in response to the influx of people who came to answer Britain's call for workers. The response also resulted in the 'colour bar', an informal discriminatory practice of the time whereby people of colour were denied jobs, housing and services or spaces, such as pubs, had segregated access.

Appendix Figure A2 shows a map of the geographical distribution of nurses in 1951.

It was not just nurses that were required from overseas, the labour shortages also drove the first mass wave of junior doctor recruitment from India, Pakistan,

[17] While the initial recruitment to the nursing sector was focused on the Caribbean, it created significant pathways for individuals from other Commonwealth countries, such as South Asia (e.g., India and Pakistan).

Bangladesh and Sri Lanka. The government was also involved in the recruitment of transport workers.

Additionally, labour shortages were not the only reason many individuals came to Britain. The economy of the Caribbean islands, underdeveloped by Britain, had high levels of unemployment. The partition of India and Pakistan and the civil war in Cyprus caused many to escape and seek a better life in the UK. These individuals arrived in the country and undertook employment as carpenters, typists, tailors, machinists, domestic servants, etc.[18]

3.3 Data on nurses

The occupation data comes from the 1951 census which has been computerised by the Great Britain Historical GIS Project.

The 1951 census data is recorded at the local government district level (LGD). Since then, the local government structure in England has undergone significant changes meaning that LGDs are not directly comparable to current geographical areas. The MCS records the middle layer super output area (MSOA) the cohort member lives in at each wave. Using centroid mapping, which is where if the centre of the MSOA falls within the LGD it is assigned to that LGD, we match LGDs to MSOAs.

3.4 Two-stage least-squares (2SLS) regression approach

We use a two-stage least-squares (2SLS) regression approach to first estimate the proportion of non-White peers as a function of nurses in 1951, net of child, family, and school characteristics. The predicted proportion of non-White peers is then forwarded to a second-stage regression to estimate the unbiased LATE of non-White peers on university likelihood. The first stage equation takes the following form:

[18] We use these additional occupations as instruments for robustness checks where we find very similar estimates to our baseline model.

$$Non-WhitePeers_i = \beta_0 + \beta_1 Nurses_i + \beta_2 Child_i + \beta_2 Family_i + \beta_4 School_i + \gamma_i \quad (2)$$

Where $Nurses_i$ is the percentage of employed individuals in individual i's local government region who were nurses in 1951.

The second-stage equation takes the following form:

$$UniversityLikelihood_i = \beta_0 + \beta_1 Non - \hat{WhitePeers}_i + \beta_2 Child_i + \beta_3 Family_i + \beta_4 School_i + \epsilon_i \quad (3)$$

where $Non - \hat{WhitePeers}_i$ is the predicted proportion of peers who are non-White based on the first stage.

The IV strategy requires that three assumptions be met. First, the instrument must be relevant, meaning that the proportion of nurses in the local area in 1951 is highly predictive of the exposure to non-White peers. Second, it must be exogenous, meaning it is not correlated with the error term in the explanatory (second-stage) equation. Finally, the instrument must affect university likelihood only through its effect on the proportion of non-White individuals within a school and not through any other pathway.

The first stage estimates (see Appendix Table A3) suggest that a 1 percentage point increase in the percentage of nurses in the local government district in 1951, reduces the proportion of non-White peers in 2014 by 1.67 percentage points. This is what we expected as in 1951, England was still faced with large labour shortages. In the following years, these shortages were filled by individuals from the commonwealth. The Montiel Olea and Pflueger F statistics on the strength of the excluded instrument is 58.80. This suggests that the proportion of nurses in the local area in 1951 is a strong

predictor of the ethnic composition of schools in 2014, i.e., the reliability assumption holds.

Instrument validity requires that the instrument should be as good as randomly assigned. The proportion of nurses in 1951 should affect university likelihood only through its effect on the ethnic composition of the area, and not through any other direct channels.

The nurse shortages in 1951 were sector-specific, driven largely by the high demand for healthcare workers after World War II. These shortages were not a reflection of broader economic conditions but were instead related to the unique staffing needs of the healthcare sector. The formation of the NHS and the demographic shifts caused by the war created a need for nurses, a role that required specialised training and skills. Healthcare needs, unlike certain industries such as coal mining or shipbuilding, were universal across the UK, which means that nurse shortages occurred nationwide.

Babikian (2021) provides examples of shortages including in Derbyshire, where hospitals had to close wards that could have accommodated 403 patients due to insufficient staff. Similarly, Nottinghamshire lost 344 beds, while Lincoln hospitals needed an additional 241 nurses to meet demand. The severity of these shortages is illustrated by the case of a Northamptonshire ward that had only one nurse for sixty patients. Mental health facilities across the country were particularly affected, with nearly two thousand unused beds reported due to a lack of staffing. In Lincolnshire, they had to close a maternity ward due to a shortage of trained midwives, and a new mother in London reported that the sole midwife in her maternity ward was also responsible for serving breakfast. Babikian (2021) describes key shortage hospitals including Scunthorpe, Wrightington, Bristol and Manchester. This broad geographical distribution of healthcare needs makes it less likely that the instrument is confounded by region-specific economic conditions.

Babikian (2021) describes the variation in nurse demand across regions stemming from two main factors, hospital expansion plans and healthcare specialisations. We argue that this supports instrument validity for two main reasons.

Firstly, these determinants were rooted in the immediate post-war recovery efforts. In the 1940s, most large towns had two main types of general hospitals: voluntary hospitals funded by charitable donations and municipal hospitals supported by government and local authorities. In addition to these, local 'cottage' hospitals provided simpler forms of treatment, while specialist facilities addressed needs in mental health, infectious diseases, orthopaedics, and children's care. For wartime purposes, hospitals were classified into two categories: Class 1 hospitals, equipped to perform surgeries and treat casualties, and Class 2 hospitals, which focused on convalescence and general medical care. The war underscored the need for additional capacity, leading to the creation of an extra 100,000 hospital beds through various means, such as 'crowding' existing wards, repurposing private homes, and constructing hutted accommodation. This reorganisation of healthcare, along with the subsequent establishment of the NHS, highlighted an urgent need for modernisation across hospitals. In the aftermath of the war, significant upgrades were made to less well-endowed hospitals, resulting in the opening of outpatient departments and enhanced collaboration with teaching hospitals. This historical context demonstrates that the distribution of nurses was primarily influenced by immediate healthcare priorities due to restructuring during the war rather than other local factors, thereby reinforcing the exogeneity of the instrument.

Secondly, while there might be long-run impacts of historical healthcare developments, such as the establishment of community norms that value education and the potential economic benefits of a robust healthcare system, we should not be overly concerned about these influences. This is because the specific pathways linking hospital expansion and healthcare specialisations to contemporary educational aspirations are

likely weak and indirect. Furthermore, unobserved historical factors, including endogenous policy responses related to nurse shortages and immigration patterns, have likely diminished in influence over time. This decreases the risk of these factors having a direct impact on educational outcomes decades later. Additionally, controlling for the index of multiple deprivation in 2014 helps to account for persistent socioeconomic factors.

The persistence of ethnic minority communities in areas where Commonwealth nurses were initially placed was driven by social and political factors, such as residential discrimination and limited mobility, rather than by ongoing industrial or economic conditions. Ethnic clustering in certain areas was often due to factors like housing discrimination and the need for proximity to family members and community networks. Furthermore, family reunification policies in the 1960s and 1970s ensured that ethnic composition in these areas continued to reflect the initial placement of nurses and their families. These factors were largely independent of broader industrial or economic changes, further distancing the instrument from the risk of being influenced by local economic shifts.

4 Results & Discussion

4.1 Descriptive Statistics

This research focuses on the impact of non-White peers on White pupils' self-reported likelihood of attending university. To examine the difference in reported likelihood of attending university across ethnic groups, we focus on an initial sample of 6,794 individuals who report university likelihood at age 14. 74.2% are White, 16.2% are Asian, 4.6% are Black, 3.8% are mixed raced, and 1.2% classify themselves as another ethnicity.[19]

[19]This sample is representative of the UK population. See Office of National Statisitcs (2023B) for data on the population of England.

The motivation for this research, as set out in Section 1, is that ethnic minorities have higher university aspirations than White individuals. Table 1 presents the average reported likelihood of attending university at age 14 for each broad ethnic group.[20] Ethnic minority pupils are significantly more likely to report higher likelihoods of attending university than White pupils at age 14.

Table 1 Average reported likelihood of attending university at age 14.

Ethnicity	Probability of attending university (age 14)
White	66.75
Asian	80.13
Black	83.80
Mixed ethnicity	74.31
Other ethnic group	80.94

To examine the descriptive relationship between university likelihood and peers' ethnicity, we restrict the sample to White pupils only. The sample is now comprised of 3,606 White individuals who report university likelihood at age 14.

Descriptive statistics for this sample are shown in Table 2[21]. The average percentage of peers who have a non-White ethnicity ranges from 0 to 96.8%. The average is 8.93%.

[20] See Table A1 in the appendix for reported university likelihood across a wider range of ethnic groups. All groups report a higher likelihood of attending university than White pupils.
[21] Parental education is measured in levels. See Table A2 in appendix.

Table 2 Summary statistics for covariates

Variable	Mean	Standard deviation
University likelihood age 14	65.72%	29.03
Percentage of peers who have a non-White ethnicity at age 14	8.93%	14.89
Female	0.51	0.50
KS2 score (Standardised)	0	1.00
Main parent employed	0.69	0.46
Main parent self-employed	0.10	0.29
Main parent out of the labour market	0.15	0.36
Main parent unemployed	0.01	0.11
Main parent employment missing	0.05	0.23
Main parent education L1	0.13	0.33
Main parent education L2	0.39	0.49
Main parent education L3	0.10	0.30
Main parent education L4	0.26	0.44
Main parent education L5	0.03	0.18
Main parent other qualification	0.02	0.12
Main parent no qualification	0.07	0.26
Equivalised household income	458.43	170.09
Number of siblings	1.35	0.98
Main parent age at birth of cohort member	29.49	5.58
Most deprived decile	0.07	0.26
10-20%	0.08	0.27
20-30%	0.09	0.29
30-40%	0.08	0.28
40-50%	0.11	0.31
50-60%	0.10	0.30
60-70%	0.10	0.30
70-80%	0.11	0.31
80-90%	0.11	0.32
Least deprived decile	0.12	0.33
Total number of pupils enrolled full time	1,015	39.62
KS2 average point score of the cohort at the end of KS4	28.68	1.84

To conclude the descriptive analysis, Figure 1 documents the association between the reported likelihood of attending university at age 14 and the percentage of peers who have a non-White ethnicity. The correlation is positive. As the percentage of non-White peers increases, the average reported university likelihood for White pupils also increases. For White pupils who have less than 50% exposure to non-White peers, the average reported likelihood of attending university at age 14 is 65.5%. For White pupils who have over 50% exposure, the average reported likelihood of attending university is 71.9%.

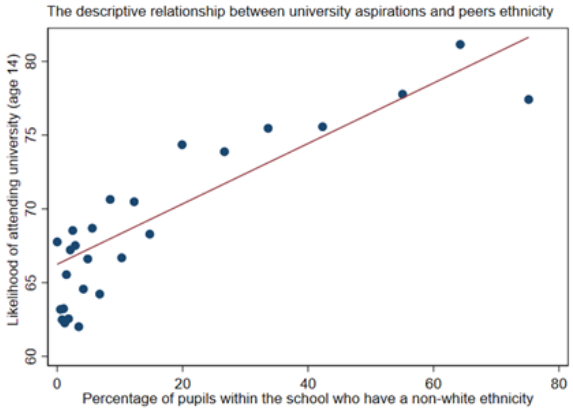

Fig. 1 Descriptive relationship between the percentage of peers who have a non-White ethnicity and perceived university likelihood for White pupils (age 14). The bin scatter plot shows the correlation between the percentage of non-White pupils and university aspirations for White pupils at age 14. The percentage of non-White peers is divided into 36 equally sized groups, and the average university aspirations are calculated for each group. The fitted line is taken from a simple linear regression of university aspirations on the percentage of non-White peers

4.2 Regression Results

Table 3 presents OLS estimates of the effect of peers' ethnicity on reported university likelihood at age 14. The dependent variable is university likelihood. Controls include gender, ability proxy, main parent's employment and education, household income, number of siblings, main parent's age at birth, and local deprivation index. Non-White is defined as the percentage of peers who are non-White.

Table 3 Impact of ethnic minority peers on reported university likelihood (OLS estimation)

	University likelihood
Non-White	0.0883**
	(0.0351)
N	3606
R^2	0.13

Notes: Standard errors in parentheses. The dependent variable is the reported likelihood at age 14 of attending university (as a percentage). Non-White is the percentage of peers who have a non-White ethnicity. Controls include gender, ability proxy, month of birth, main parent's age at birth, number of siblings, main parent's education and labour market status, household income, local deprivation index, and school characteristics. $*p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Turning to the IV results, the instrument performs well. The first-stage F statistic (Montiel Olea and Pflueger F statistics) is 58.80.[22]

The first stage estimates (see Appendix Table A3) suggests that a 1 percentage point increase in the percentage of nurses in the local government district in 1951, reduces the proportion of non-White peers in 2014 by 1.67 percentage points. This is what we expected as in 1951, England was still faced with large labour shortages. In the following years, these shortages were filled by individuals from the commonwealth.

As discussed in Section 3.2, in 1951 there was a shortage of nurses. Whilst government-backed recruitment of nurses from the commonwealth had begun in 1949 there were still significant shortages in 1951. As nursing is a profession required across the country, the 1951 census data shows where the shortages were. If the local government district had high numbers of nurses in 1951, they had less labour shortage and were therefore assigned few of the government requirements from overseas. Britton et al (2021) provided evidence to show that ethnic minorities are less likely to move areas. We therefore expect to see fewer ethnic minorities in these areas today.

[22] To check the reliability of our t-ratio, we use Lee et al (2022) method to calculate valid t-ratios for the model 3. With an Montiel Olea and Pfluger F statistic of 58.80, we calculate our standard error adjustment factor of 1.07. Multiplying the standard error by the adjustment factors increases the standard error slightly but the estimated coefficient remains significant at a 1% significance level.

The second-stage IV results, shown in Table 4, are larger in magnitude than the OLS estimates.[23][24] This suggests a downwards bias in the OLS estimates.[25] This is consistent with the literature, specifically with studies that use instrumental variables (De Giorgi et al, 2009; Mendolia et al, 2018). The IV results indicate that on average, increasing the percentage of non-White peers by 1 percentage point, increases White individuals' reported likelihood at age 14 of attending university by 0.50 percentage points. This finding is in line with Dickerson et al (2018), who also use data from England and provide evidence to show that a 10 percentage point increase in the proportion of an individual's peers who aspire to an academic route is associated with a 5 percentage point increase in the likelihood that the individual has similar aspirations themselves.[26] The simplifying assumption made in the baseline model is that there are neither diminishing nor increasing benefits to the proportion of non-White peers. We found no evidence of non-linearity.

[23] Full model second stage results are shown in Table A4 in the appendix

[24] The coefficient of proportionality, as demonstrated by Ciacci (2021) using Oster (2019) bounds, quantifies how much stronger the selection on unobservables must be compared to observables in order to estimate the IV coefficient with the OLS model. In our analysis, we find low values for this coefficient, indicating that selection on unobservables is relatively small compared to observables to estimate the IV coefficient with the OLS model. This provides supportive evidence that the IV estimates are not excessively large in relation to the OLS estimates.

[25] OLS estimates an Average Treatment Effect (ATE), and IV estimates a Local Average Treatment Effect (LATE) so we do have to be aware that they are estimating different things.

[26] Table **??** presents the IV estimates from a just identified model when controls for both parents are included. The estimated impact is consistent across the different model.

Table 4 Impact of ethnic minority peers on reported university likelihood (Just-identified IV model)

	University likelihood
Non-White	0.500***
	(0.154)
Montiel Olea and Pfluger F Stat	58.80
N	3606

Notes: The dependent variable is the reported likelihood at age 14 of attending university (provided as a percentage). Non-White is the percentage of peers who have a non-White ethnicity. Montiel Olea and Pflueger F statistic is shown. Controls include gender, a proxy for ability, month of birth, main parent's age at birth, number of siblings, main parent's education and labour market status, household equivalised income, the local deprivation index, the number of pupils enrolled in the school, and a proxy for school quality. Standard errors are clustered at the school level and are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.3 Heterogeneity

On average, increasing the proportion of ethnic minority peers has a positive impact on the perceived likelihood at age 14 of attending university for White pupils.

The next stage of the analysis explores the extent to which the impact of peers' ethnicity differs across observable characteristics. To do this we conduct a series of subgroup analyses based on individual and family characteristics.[27] Due to small sample sizes, the sub-sample analysis provides suggestive evidence of how impact varies across observable characteristics.

We start by examining the impact across gender. Females are much more likely to go to university than males and have been for many years. The higher education participation level for females is 56.6%, compared to 44.1% for males (Hewitt, 2020). Assuming that the reported likelihood of attending university is a good predictor of actual university attendance, if we find a more positive impact on male pupils

[27] We estimate all subgroup analysis using the just-identified IV model. Montiel Olea and Pfluger First stage F statistics remain above 10. The average Montiel Olea and Pfluger F stat across the sub-group analysis is 48.1.

then increasing the exposure to non-White pupils could aid in reducing the gender differences in university participation.

Table 5 The impact of ethnic minority peers on the perceived likelihood of attending university for White pupils, by gender. (Estimated by a just-identified IV model)

	Female	Male
Non-White	0.519**	0.477**
	(0.222)	(0.217)
Montiel Olea and Pfluger F Stat	46.69	40.12
N	1832	1774

Notes: The dependent variable is the reported likelihood at age 14 of attending university (provided as a percentage). Non-White is the percentage of peers who have a non-White ethnicity. Controls include gender, a proxy for ability, month of birth, main parent's age at birth, number of siblings, main parent's education and labour market status, household equivalised income, the local deprivation index, the number of pupils enrolled in the school, and a proxy for school quality. Standard errors are clustered at the school level and are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5 shows a positive impact of ethnic minority peers for both males and females, of very similar magnitudes.[28] Diette and Oyelere (2014) and Legewie and DiPrete (2012), who both examine the impact of peers on educational achievement, find males to be most impacted. Dickerson et al (2018), who focus on aspirations, shows that female and males' aspirations at age 14 are both impacted by their peers.

We also consider the effects for White pupils who are more exposed to ethnic minority peers, for those who are economically disadvantaged, and those who performed poorly in tests at age 11. More generally, Table 6 shows sub-group analysis across the income distribution whilst Table 7 presents the impact for individuals whose parents, at best, have school level qualifications or lower compared to parents with higher levels of qualifications, Table 8 estimates the impact across the ability distribution.[29]

Table 6 shows that the estimated impact of non-White peers on White pupils' university likelihood decreases across the income distribution. The estimated impact

[28] The difference is insignificant.

[29] To examine the impact across the income and ability distribution we generate three dummy variables relating to the bottom 20%, top 20% and middle 60% of the distribution.

for pupils at the bottom of the income distribution is three times larger than the estimated impact for those in the top of the income distribution. We interpret these coefficients with caution as we have small sample sizes and large standard errors.

Similarly, Table 7 shows that the estimated coefficient is larger for individuals who have parents with lower level of qualifications.

The literature provides varying results when analysing the heterogenous impacts of peer effects. Hoxby (2000) and Gould et al (2009) argue that disadvantaged pupils are more responsive to changes in school resources. On the other hand, Geay et al (2013) provides evidence to show that disadvantaged pupils are less affected. The most recent study in the literature that focuses on aspirations, Gagete-Miranda (2022), shows that peer impacts are homogenous across individual and family characteristics.

Table 6 The impact of ethnic minority peers on the perceived likelihood of attending university of White pupils, by household income (Estimated by a just-identified IV model)

	Top	Middle	Bottom
Non-White	0.229	0.626**	0.634*
	(0.186)	(0.271)	(0.381)
N	722	2160	724

Notes: See Table 5 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7 The impact of ethnic minority peers on university aspirations of White pupils, by parental education. (Estimated by a just-identified IV model.)

	School level qualification	Higher than school qualification
Non-White	0.531**	0.429*
	(0.209)	(0.224)
N	1400	2206

Notes: See Table 5 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Similarly, we can consider how the impact of non-White peers on White pupils' reported university likelihood differs across the ability distribution. Wiseman et al

(2017) shows that prior educational attainment is the main predictor of university participation and accounts for much of the variation in participation. The higher the number of GCSEs attained, the more likely pupils are to attend university. We therefore assess the impact of peers' ethnicity across the ability distribution. Table 8 shows that as KS2 test scores increase, the impact of peers on university likelihood decreases. The estimated impact for those in the bottom of the ability distribution is double the impact for those at the top of the ability distribution. We find a significant impact in the middle of the distribution.

Table 8 The impact of ethnic minority peers on the perceived likelihood of attending university for White pupils, by ability. (Estimated by a just-identified IV model)

	Top	Middle	Bottom
Non-White	0.314	0.425*	0.762
	(0.260)	(0.219)	(0.489)
N	810	1930	847

Notes: See Table 5 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.4 Aspiration distribution

We additionally examine the impact of ethnic minority peers across the distribution of reported university likelihood. If exposure to ethnic minority peers is only having an impact on individuals who already have high aspirations then the policy response would be different to if the impact was focused at the bottom of the distribution.

Figure 2 presents results from an instrumental variable quantile regression. Figure 2 shows us that the impact of ethnic minority peers is close to the 2SLS estimate across the aspiration distribution but only significant for the bottom half of the aspiration distribution. However, we should be cautious in our interpretation as the confidence intervals increase as we move up the distribution.

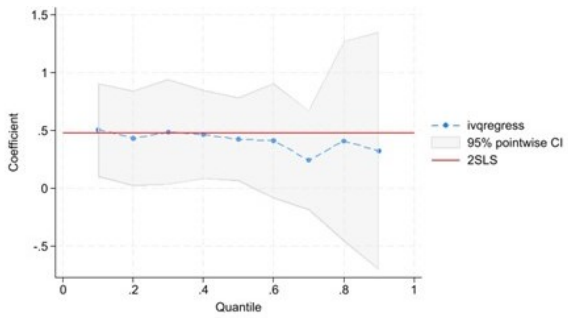

Fig. 2 The impact of ethnic minority peers on the distribution of university aspirations of White pupils. (Estimated by a just-identified quantile IV model.) Notes: Reported university likelihood is separated in to 9 quantiles. The estimated coefficient provides the impact on ethnic minority peers on individual reported likelihood of attending university from a IV quantile version of equation 3

4.5 Ethnic groups

Non-White pupils represent a heterogeneous group. To determine whether the peer effect differs across ethnic groups we disaggregate them into the 19 ethnic groups used by the census. Whilst non-White pupils on average have higher perceived likelihood of attending university than White pupils, there is variation within non-White pupils. We may expect larger peer effects from ethnic groups with the highest reported university likelihood. We focus our analysis on the share of Indian, Pakistani and African pupils within schools as they are the largest minority groups in England.[30] We groups all other ethnic groups in to one category.[31] We estimate separate equations for each group.

Whilst our baseline IV results suggest an impact of 0.50, disaggregating by ethnicity produces much larger estimates. Being exposed to Indian, Pakistani and African peers increases university likelihood by between 2 and 1.5 percentage points. These findings fit with the differences in university aspirations of these ethnic groups. The

[30] 16.1% of the population is non-White. 3.1% of the population is Indian, 2.7% are Pakistani and 2.5% are Black African (Office of National Statisitcs, 2023B).
[31] The other group include Bangladeshi, other Asian ethnicities, Black Caribbean, other Black ethnicities, White and Black African, White and Asian, other mixed ethnicities and other ethnicities.

average effect observed earlier therefore seems to be pushed down by the impact of all other ethnicities. As Indian, Pakistani and African pupils make up the majority of non-White peers, we suggest that the impact for some White pupils could be larger than estimated in our baseline.

Table 9 The impact of different groups of ethnic minority peers on the perceived likelihood of attending university for White pupils. (Estimated by a just-identified IV model)

	Indian	Pakistani	Black African	Other
Ethnic groups	2.086***	1.503***	1.965***	0.603
	(0.802)	(0.547)	(0.677)	(0.539)
N	3581	3581	3581	3581

Notes: See Table 5 'Ethnic groups' is the percentage of peers in the ethnic group given by the respective column headings. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.6 Robustness checks

4.6.1 Tobit model

Figure 3 illustrates the distribution of university likelihood at age 14. As with many subjective probability questions, responses are heaped to different degrees at multiples of 5 and 10 and the fraction of '50%' answers is particularly large. This suggests that the question suffers from various types of reporting behaviour (rounding and focal answers). Kleinjans and Soest (2014) who investigate how individuals respond to survey questions about their personal beliefs or expectations, particularly when asked to assign probabilities to uncertain events, find that these behaviors can lead to biased data and missing responses, posing challenges for accurately interpreting results. They discuss the use of Tobit models as one of the potential methods to handle issues with subjective probability data. In their analysis, they indicate that using Tobit models allows for the estimation of the underlying distribution of subjective probabilities while accounting for the censoring at the bounds. Whilst this might not fully address the rounding behavior, they provide evidence to show that estimating

a Tobit model provides signs and significance levels that are very similar to models that fully account for the reporting behaviour. They conclude that ignoring reporting behaviour only leads to modest biases in the estimated means and standard deviations of the true probabilities. Based on these finding we also estimate a Tobit model to test the sensitivity of our findings.

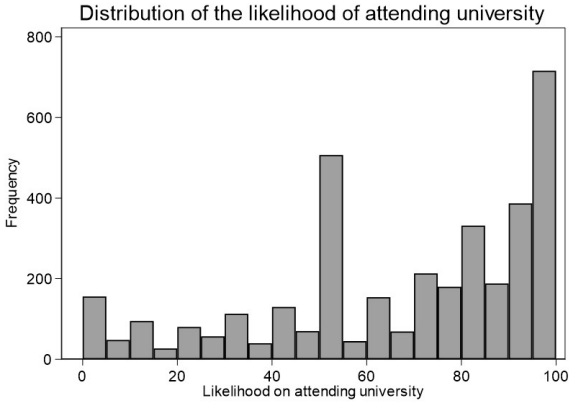

Fig. 3 Distribution of university likelihood reported at age 14

Table 10 Just-identified tobit IV model

	University likelihood
Non-White	0.578***
	(0.175)
N	3606

Notes: The dependent variable is the reported likelihood of attending university at age 14 (provided as a percentage). Non-White is the percentage of peers who have a non-White ethnicity. Controls include gender, a proxy for ability, month of birth, main parent's age at birth, number of siblings, main parent's education and labour market status, household equalised income, the local deprivation index, the number of pupils enrolled in the school, and a proxy for school quality. Standard errors are clustered at the school level and are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tabel 10 shows that the estimated coefficient from our 2SLS estimation are consistent with those from the Tobit model.

4.6.2 Omitted variables

All specifications control for individual, family, school and local characteristics. Despite controlling for a good amount of selection into schools, there is however still potential for selection on unobservables. As school attended is normally a decision made by the parents of a child, the ethnic composition of that school may have impacted the parents' decision due to beliefs held on ethnic minorities. Dustmann and Preston (2001), using data from England, find that individuals who are more hostile towards ethnic minorities will not settle in neighbourhoods with a high ethnic concentration.

The concept of White flight demonstrates how parents' attitudes and preferences regarding ethnic diversity can affect the demographics of schools. When families with negative views toward ethnic minorities choose to move to neighbourhoods with fewer ethnic residents, this behaviour has a direct effect on the ethnic makeup of local schools. By including controls for parental attitudes toward mixed-race schools in our analysis, we can account for the impact of White flight on school demographics. The validity of the instrumental variable could be compromised if White flight is linked to the ethnic composition of schools and affects educational outcomes. By including controls for parental preferences and aspirations, our model minimises the potential confounding effects of White flight. This enhances the reliability of the instrument, reducing the likelihood that unobserved elements associated with White flight are affecting the connection between nurse shortages and educational results.

We use two measures of parental attitudes. Firstly, when the cohort member is 9 months old, years prior to them starting school, the parent is asked how they would feel about their child attending a 50/50 mixed raced school. The responses are recorded in a categorical variable ranging from strongly agree to strongly disagree. We use this categorical variable to create three dummy variables, agree, disagree and neither

agree nor disagree. We include agree and disagree in the model as a proxy for parental beliefs. Secondly, when the child is 14, parents are asked how likely they think it is that the child will attend university. This is reported as a categorical variable ranging from very unlikely to very likely.

Table 11 presents the IV estimates with parental preferences on mixed raced schools whilst Table 12 presents the IV estimates including parental aspirations. The results are robust to the control of beliefs and parental aspirations with very little change in the estimated coefficient.[32]

Table 11 Just-identified IV model including parental preferences on mixed raced schools

	University likelihood
Non-White	0.512***
	(0.157)
N	3606

Notes:See 5 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 12 Just-identified IV model including parental aspirations.

	University likelihood
Non-White	0.289**
	(0.129)
N	3379

Notes: See 5 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Another potential omitted variable is whether the ethnic minorities are first language English. The language skills of peers is likely to influence the impact they have. Additionally, peers' language skills can also influence a pupil's exposure to ethnic minorities by shaping educational experiences and school policies. We therefore additionally control for the percentage of pupils within the school for whom English is their second language.

[32]The Montiel Olea and Pfluger F statistics is 57.25.

Table 13 presents the IV estimates with this additional control. The coefficient is slightly larger. We might expect this as an increase in English second language peers is likely to be positively correlated with exposure to ethnic minorities and negatively correlated with individual aspirations if language differences leads to less interaction.

Table 13 Just-identified IV model including percentage of pupils within the school who have English as a second language.

	University likelihood
Non-White	1.27**
	(0.507)
N	3596

Notes: See Table 5 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5 Conclusion

While there is evidence on the impact of peers' ethnicity on pupil's educational outcomes, very little exists on the impact on aspirations. With aspirations seen as a pathway to higher education, this is key gap in the literature. This research fills that gap by providing the first analysis of the relationship between ethnic minority peers and White pupils' perceived likelihood of attending university. Additionally, we identify the causal relationship by creating a unique instrumental variable, the proportion of nurses in 1951—an indicator of post-World War II job shortages filled by immigrant workers.

We find that increasing the proportion of ethnic minorities in the school by 1 percentage point leads to an 0.50 percentage point increase in White pupils' reported university likelihood at age 14. We believe age 14 is an important time to examine the impact on reported university likelihood as at this age, it is the first time that pupils get a choice over what subjects to study when picking their GCSEs. Pupils tend to see this choice as a first step in determining their future educational pathway.

The findings are robust across various checks, including changes to the sample definition for White and ethnic minority pupils, as well as controls for potential omitted variables such as parental aspirations and attitudes toward ethnic diversity. We also assess the impact of homophily in friendship formation, accounting for the closeness of connections with ethnic minority peers. Further robustness checks include weak instrument tests and sensitivity analyses with alternative instruments based on additional occupational data. The robustness checks consistently suggest that the existence of weak instruments or omitted variable bias is rather unlikely.

The findings are consistent with Dickerson et al (2018) and Gagete-Miranda (2022) who both find that exposure to peers with high aspirations leads to large, positive, and significant peer effects on individuals' aspirations. Building on this research, we find that high aspiring ethnic minority pupils influence the university aspirations of White pupils they interact with in the same school.

Appendix A

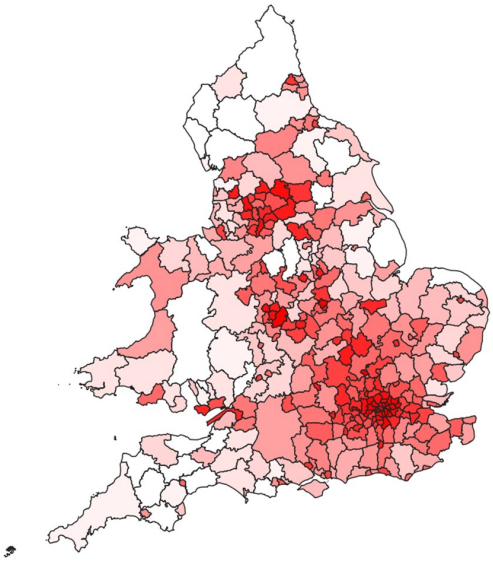

Fig. A1 Ethnic minority density in 2011 by local authority.Notes: The map above shows the proportion of ethnic minorities in each local authority. The darker the colour the more ethnic minorities in that area.

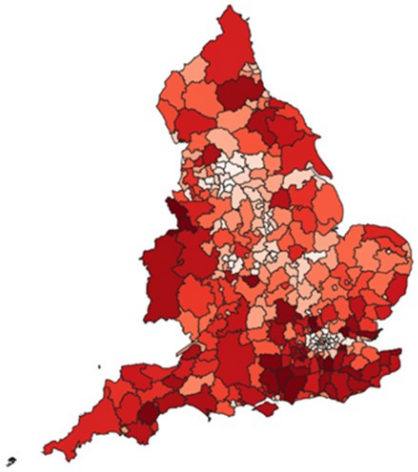

Fig. A2 Proportion of nurses in 1951 by local authority.Notes: The map above shows the proportion of nurses in each local authority in 1951. The darker the colour the more nurses in that area. We assume that areas with fewer nurses in 1951 (lighter colour) have more ethnic minorities today (darker colour).

Table A1 Average reported likelihood of attending university at age 14.

Ethnicity	Probability of attending university (age 14)
White and Black Caribbean	67.26
	(30.38)
White and Black African	77.24
	(24.50)
White and Asian	74.79
	(25.93)
Any other mixed background	84.21
	(18.29)
Indian	84.68
	(20.68)
Pakistani	77.54
	(24.02)
Bangladeshi	77.66
	(22.77)
Any other Asian background	87.94
	(17.86)
Caribbean	74.96
	(23.97)
African	88.61
	(15.68)
Any other Black background	82.89
	(18.30)
Any other background	80.94
	(22.53)

Table A2 Levels of Parents' Education

Level	Qualifications
L1	Foundation Diploma
	GCSE (grades D–G)
	Scottish National level below and equal to level 4
	NVQ Level 1
	City and Guilds foundation part 1
	GNVQ foundation level
	BTEC first certification
	RSA level 1
	Entry level qualification
	Level 1 Award, Basic Skill qualification
	Key Skill qualification
	YT/YTP Certificate
L2	Higher diploma
	O-level - GCSE (grades A*–C)
	Scottish National level 5
	Intermediate Welsh Baccalaureate
	NVQ Level 2
	City and Guilds Craft part 2
	GNVQ intermediate
	BTEC level 2
	RSA level 2
	Level 2 Diploma/Certificate Advanced diploma
	A-level
	International Baccalaureate
	SCE higher
	Scottish Baccalaureate
	Advanced Welsh Baccalaureate
	NVQ Level 3
	City and Guilds Advanced Craft
	GNVQ Advanced
	BTEC National
	RSA level 3
	Level 3 Award/Certificate

Level	Qualifications
L3	Certificate of Higher Education NVQ Level 4 Higher National Certificate (HNC) BTEC Professional award certificate and diploma level 4 RSA level 4 Level 4 Award/ Certificate/Diploma Nursing Diploma of Higher Education Foundation degree Teaching foundation stage/ primary education/secondary education/further education Higher National Diploma (HND) BTEC Professional Award Level 5 Certificate/Diploma/ Award.
L4	First Degree Foundation degree BTEC Advanced Professional award Level 6 Certificate/Diploma/Award
L5	Master's degree Integrated master's degree BTEC Advanced Professional award Postgraduate certificate/ diploma level 7 NVQ Level 5 Doctorates

Table A3 The impact of ethnic minority peers on the reported university likelihood of White pupils (First Stage for 2SLS estimation using the proportion of nurses in 1951 as the instrumental variable)

	Non-White
Nurse (IV)	-1.6707***
	(0.2168)
Female	0.2196
	(0.5065)
February	1.3249
	(1.3228)
March	1.2780
	(1.1391)
April	0.3801
	(1.1322)
May	1.4180
	(1.1873)
June	1.5885
	(1.1448)
July	1.8032
	(1.1839)
August	1.2096
	(1.1287)
September	0.7974
	(1.0963)
October	0.7974
	(1.1298)
November	0.3645
	(1.0646)
December	0.0896
	(1.0378)
Ability	0.3378
	(0.2158)
Household income	0.0153***
	(0.0028)
Mother L1 education	-2.4056***
	(2.7452)
Mother L2 education	-1.6771***
	(0.6088)
Mother L3 education	-1.0888
	(0.6914)
Mother L4 education	-2.0873***
	(0.8020)
Mother L5 education	1.7255
	(1.7849)
Mother other education	8.4825***
	(2.7452)

Main employed	-3.481
	(2.9601)
Main self-employed	-3.6957
	(3.0203)
Main out of labour market	-2.8002
	(2.9564)
Main employment missing	-3.5610
	(3.0618)
Siblings	1.3721***
	(0.2793)
Main birth age	0.1511***
	(0.0468)
Deprivation 1	4.0206**
	(1.8859)
Deprivation 2	-0.5057
	(1.5367)
Deprivation 3	-0.5402
	(1.6203)
Deprivation 4	-4.3401***
	(1.4795)
Deprivation 5	-1.4262
	(1.7758)
Deprivation 6	-4.5857***
	(1.5583)
Deprivation 7	-51730***
	(1.5200)
Deprivation 8	-5.5637***
	(1.7016)
Deprivation 9	-5.1902***
	(1.7016)
School ability	0.8035***
	(0.2906)
N	3606
R^2	0.14

Notes: First stage regression. We also include the number of pupils enrolled in the school as a control. This is a categorical variables split in to many groups. Coefficient for these groups can be requested from the author. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4 The impact of ethnic minority peers on the reported university likelihood of White pupils(Second Stage for 2SLS estimation)

	University likelihood
Non-White	0.500***
	(0.1541)
Female	6.8841***
	(0.9104)
February	-2.0176
	(2.5887)
March	0.1068
	(2.4751)
April	0.0554
	(2.4829)
May	-1.2269
	(2.4342)
June	-4.3079*
	(2.3078)
July	-1.4775
	(2.5057)
August	0.4056
	(2.3328)
September	0.2357
	(2.4319)
October	2.1623
	(2.4319)
November	2.4375
	(2.2420)
December	3.4567
	(2.1880)
Ability	1.2767***
	(0.4340)
Household income	0.0080*
	(0.0043)
Mother L1 education	-1.4621
	(1.7046)
Mother L2 education	-3.9315***
	(1.4983)
Mother L3 education	6.1086***
	(1.4708)
Mother L4 education	11.6819***
	(1.5404)
Mother L5 education	17.5521***
	(2.3674)
Mother other education	-1.4980
	(3.6096)

	University likelihood
Main employed	5.2528
	(4.7456)
Main self-employed	6.1765
	(4.9180)
Main out of labour market	2.1614
	(4.8228)
Main employment missing	4.2843
	(5.1304)
Siblings	-1.1899*
	(0.5542)
Main birth age	-0.0929
	(0.0918)
Deprivation 1	-3.4852
	(2.6539)
Deprivation 2	-3.3051
	(2.6539)
Deprivation 3	-4.9518*
	(2.5422)
Deprivation 4	0.1118
	(2.7928)
Deprivation 5	-4.5569*
	(2.5652)
Deprivation 6	-1.6772
	(2.7183)
Deprivation 7	-1.3772
	(2.5407)
Deprivation 8	-0.2457
	(2.6527)
Deprivation 9	-0.4663
	(2.7323)
School ability	1.90079***
	(0.3991)
N	3606
R^2	0.09

Notes: Second stage regression. We also include the number of pupils enrolled in the school as a control This is a categorical variables split in to many groups. Coefficient for these groups can be requested from the author. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Declarations

- No funding was received for conducting this study.
- The authors have no relevant financial or non-financial interests to disclose.

- This article uses English administrative data on education outcomes from the UK Department of Education and a secure access version of the Millennium Cohort Study from the Center for Longitudinal Studies. These datasets are only available upon request to the data producers, through the UK Data Service. We also use occupation data from the 1951 census which is a available upon request to the data producers, Great Britain GIS Project. The author is willing to provide assistance with data requests.

References

Anelli M, Shih K, Williams K (2023) Foreign students in college and the supply of stem graduates. Journal of Labor Economics 41(2):511–563

Babikian C (2021) "partnership not prejudice": British nurses, colonial students, and the national health service, 1948–1962. Journal of British Studies 60(1):140–168

Barshay J (2023) New higher ed data by race and ethnicity. Accessed October 2024. https://hechingerreport.org/proof-points-new-higher-ed-data-by-race-and-ethnicity/#:~:text=Roughly%2060%20percent%20of%20young,more%20than%2080%20percent%20enroll.

Betts JR, Fairlie RW (2001) Explaining ethnic, racial, and immigrant differences in private school attendance. Journal of Urban Economics 50(1):26–51

Britton J, van der Erve L, Waltmann B, et al (2021) London calling? higher education, geographical mobility and early-career earnings. Institute for Fiscal Studies Report Accessed December 2023. https://ifs.org.uk/publications/london-calling-higher-education-geographical-mobility-and-early-career-earnings

Burgess SM (2014) Understanding the success of london's schools. Working Paper No. 14/333. Bristol: The Center for Market and Public Organisation.

Cascio EU, Lewis EG (2012) Cracks in the melting pot: immigration, school choice, and segregation. American Economic Journal: Economic Policy 4(3):91–117

Chareyron S, Chung A, Domingues P (2021) Ethnic diversity and educational success: Evidence from france. Research in Economics 75(2):133–143

Ciacci R (2021) A Matter of Size: Comparing IV and OLS estimates. Doctoral thesis: Universidad Pontificia Comillas

Crawford C, Greaves E (2015) Socio-economic, ethnic and gender differences in he participation. London: Department for Business, Innovation & Skills

De Giorgi G, Pellizzari M, Redaelli S (2009) Be as careful of the company you keep as of the books you read: peer effects in education and on the labor market. NBER Working Papers 14948. National Bureau of Economic Research

Dearden L, Sibieta L (2010) Ethnic inequalities in child outcomes. London: Institute of Education. Accessed June 2023, https://cls.ucl.ac.uk/wp-content/uploads/2017/05/12_briefing_web.pdf

d'Este R, Einiö E (2021) Beyond black and white: the impact of asian peers on scholastic achievement. Economics of Education Review 83:102129

Dickerson A, Maragkou K, McIntosh S (2018) The causal effect of secondary school peers on educational aspirations. Centre for Vocational Education Research Discussion Paper 17

Diette TM, Oyelere RU (2014) Gender and race heterogeneity: The impact of students with limited english on native students' performance. American Economic Review 104(5):412–417

Dustmann C, Preston I (2001) Attitudes to ethnic minorities, ethnic context and location decisions. The Economic Journal 111(470):353–373

Elsner B, Isphording IE (2017) A big fish in a small pond: Ability rank and human capital investment. Journal of Labor Economics 35(3):787–828

Figlio D, Özek U (2019) Unwelcome guests? the effects of refugees on the educational outcomes of incumbent students. Journal of Labor Economics 37(4):1061–1096

Figlio D, Giuliano P, Marchingiglio R, et al (2024) Diversity in schools: Immigrants and the educational performance of us-born students. Review of Economic Studies 91(2):972–1006

Gagete-Miranda J (2022) An aspiring friend is a friend indeed: On the mechanisms behind peer influences on human capital accumulation. Manuscript: Bocconi University: Milan, Italy

Geay C, McNally S, Telhaj S (2013) Non-native speakers of english in the classroom: what are the effects on pupil performance? The Economic Journal 123(570):F281–F307

Gorard S, See BH, Davies P, et al (2012) The impact of attitudes and aspirations on educational attainment and participation. York: Joseph Rowntree Foundation

Gould ED, Lavy V, Daniele Paserman M (2009) Does immigration affect the long-term educational outcomes of natives? quasi-experimental evidence. The Economic Journal 119(540):1243–1269

Hewitt R (2020) Mind the gap: gender differences in higher education. Accessed March 2024. https://www.hepi.ac.uk/2020/03/07/mind-the-gap-gender-differences-in-higher-education/,

Home Office (2020) Historical background information on nationality. Accessed March 2024. https://www.gov.uk/government/publications/historical-background-information-on-nationality/historical-background-information-on-nationality-accessible

Hoxby CM (2000) Peer effects in the classroom: Learning from gender and race variation. NBER working paper 7867. National Bureau of Economic Research Cambridge, Massachusetts, USA

Iranzo S, Schivardi F, Tosetti E (2008) Skill dispersion and firm productivity: An analysis with employer-employee matched data. Journal of Labor Economics 26(2):247–285

Kleinjans KJ, Soest AV (2014) Rounding, focal point answers and nonresponse to subjective probability questions. Journal of Applied Econometrics 29(4):567–585

Lee DS, McCrary J, Moreira MJ, et al (2022) Valid t-ratio inference for iv. American Economic Review 112(10):3260–3290

Legewie J, DiPrete TA (2012) School context and the gender gap in educational achievement. American Sociological Review 77(3):463–485

Mendolia S, Paloyo AR, Walker I (2018) Heterogeneous effects of high school peers on educational outcomes. Oxford Economic Papers 70(3):613–634

Mora T, Oreopoulos P (2011) Peer effects on high school aspirations: Evidence from a sample of close and not-so-close friends. Economics of Education Review 30(4):575–581

Morales C (2022) Do refugee students affect the academic achievement of peers? evidence from a large urban school district. Economics of Education Review 89:102283

Murphy R, Weinhardt F (2020) Top of the class: The importance of ordinal rank. The Review of Economic Studies 87(6):2777–2826

NHS (2023) Nhs history. National Health Service. Accessed March 2024. https://www.england.nhs.uk/nhsbirthday/about-the-nhs-birthday/nhs-history/

Office of National Statisitcs (2023B) Ethnic group, england and wales: Census 2021. Accessed March 2024. https:https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/bulletins/ethnicgroupenglandandwales/census2021

Oster E (2019) Unobservable selection and coefficient stability: Theory and evidence. Journal of Business & Economic Statistics 37(2):187–204

Parliament. House of Commons (1942) Ministry of health. volume 381: Debated on tuesday 30 june 1942. Accessed March 2024. https://hansard.parliament.uk/Commons/1942-06-30/debates/07c05f13-740e-4e07-8861-ddebb766080c/MinistryOfHealth?highlight=nurse%20shortages#contribution-c152d47b-5894-4d9e-9ada-cc8985e8228c

Plewis I (2007) Non-response in a birth cohort study: the case of the millennium cohort study. International Journal of Social Research Methodology 10(5):325–334

Putnam RD (2007) E pluribus unum: Diversity and community in the twenty-first century the 2006 johan skytte prize lecture. Scandinavian Political Studies 30(2):137–174

Rihal GDS. (2021) Millennium cohort study: A guide to the linked education administrative datasets. https://cls.ucl.ac.uk/wp-content/uploads/2019/07/user_guide_to_the_linked_education_datasets_2019_version3_may2021_v2.pdf

Snow S, Jones E (2011) Immigration and the national health service: putting history to the forefront. Institute of Historical Research, Senate House, University of London.

University College London and UCL Institute of Education and Centre for Longitudinal Studies and Department for Education (2023) Millennium cohort study: Linked education administrative datasets (national pupil database). UK Data Service, https://doi.org/http://doi.org/10.5255/UKDA-SN-8481-2, secure Access, UK Data Service, 2nd Edition, SN: 8481

Wiseman J, Davies E, Duggal S, et al (2017) Understanding the changing gaps in higher education participation in different regions of England. England: Department for Education