Diversity and Neighbourhood Satisfaction

Monica Langella and Alan Manning

Centre for Economic Performance

London School of Economics

Houghton St.

London WC2A 2AE

March 2017

Abstract

This paper uses UK data to investigate the impact of ethnic diversity on individuals' attitudes about their neighbourhood. It considers a wider range of outcomes than most of the existing literature - overall neighbourhood satisfaction, social capital, residential mobility, fear of crime, and satisfaction with local services and social life. It uses longitudinal data, instrumental variables and sample selection methods to control for potential endogeneity of diversity and of the location choices in a more thorough way than most of the existing literature. We find that a higher white share in the neighbourhood raises overall satisfaction with the neighbourhood in our (overwhelming white) sample, but has no significant impact on generalised trust or other commonly-used measures of social capital. We suggest that part of the impact of diversity on overall neighbourhood satisfaction may be through an effect on a fear of crime.

Keywords: Neighbourhood satisfaction, social capital, diversity, deprivation.

JEL Classification: Z1

We would like to thank the Understanding Society team at the University of Essex for allowing us to use the geo-coded versions of the data sets used in this paper. And seminar participants at Stanford University, Carlos III Madrid, and Royal Economic Society conference for their comments.

Introduction

Sizeable parts of the population of most Western countries seem troubled by increased ethnic diversity in their societies. For many people, the most direct impact of these changes is felt in their communities. This paper is about how ethnic diversity within neighbourhoods in the UK affects people's satisfaction with their local areas.

We consider a wider range of outcomes than most of the existing literature; in addition to trust in others, activity in organizations (these two being commonly used measures of 'social capital' e.g. Putnam, 2000), we consider overall satisfaction with the neighbourhood, the intention to move, actual residential mobility, perceptions of crime, the quality of social life and local services. In doing so, we contribute to a more complete description of the impact of diversity. The main findings of our paper are that a lower white share leads to a lower neighbourhood satisfaction among our (overwhelmingly white) sample and increased perceptions of crime though is not strongly associated with the level of generalised trust, activity in organizations, the quality of social life or the quality of local services.

This paper makes substantial contribution to the two main strands of research that might be used to support the idea that diversity is important within local social interactions. Studies on the link between social capital and diversity, and studies on the impact of diversity on neighbourhood choices and residential mobility.

The conclusion that people care about the ethnic composition of their neighbourhood may seem a statement of the obvious. Existing studies have argued that greater diversity reduces trust (Putnam, 2007; Dinesen and Sonderskov, 2012), lowers involvement in organizations (Alesina and La Ferrara, 2000, 2002; Costa and Kahn, 2003), lowers the level of social cohesion (see the survey by Van der Meer and Tolsma, 2014), lowers the level of public good provision (see Alesina, Baqir and Easterly, 1999; the review by Alesina and La Ferrara, 2005), lowers the quality of government (Alesina and Zhuravskhaya, 2011) or changes attitudes to redistribution (Dahlberg, Edmark and Lundquvist, 2012).

The studies cited above have been controversial and subject to a number of criticisms (see Portes and Vickstrom, 2011, for an overview). First, confounding factors - both at individual or neighbourhood level – may play a role in explaining the correlation between diversity and trust. One example of this can be poverty, as shown by the Moving To Opportunity (MTO) evaluations (Kling, Liebmann and Katz, 2007, Ludwig et al., 2012, and Chetty, Hendren and

Katz, 2016), which have found important effects of neighbourhood poverty on subjective well-being, mental health and long-term child outcomes. In addition, Uslaner, (2012) argues that it is segregation rather than diversity that is important, while Tesei (2015) points to the importance of racial income inequality.

There are also concerns that most studies in this area rely on correlations in cross-sectional data in which causal evidence is limited (Portes and Vickstrom, 2011)[1]. Our contribution is to pay greater attention to causality. We do so exploiting and combining three different approaches. First, this work uses individual level longitudinal data to document the effect of diversity on neighbourhood perceptions. This allows us to control for individual and neighbourhood fixed effects. Second, we instrument for diversity using an instrument popularized by Altonji and Card (1991) and Card (2001) which uses predicted ethnic mix based on neighbourhood-specific initial ethnic mix and national growth rates in the population of different ethnic minorities. Third, we develop an empirical approach to control for the possible selection bias caused by the endogeneity of residential choice. We hope that in addressing these empirical issues, we make some progress in providing causal estimates of the impact of diversity. Throughout, we report a wide variety of specifications (including some very demanding ones) in order to convey the robustness (or lack of it) of our empirical findings.

One other relevant literature is that on how ethnic composition affects neighbourhood choice. In the US the most striking evidence for this is 'white flight', the process by which some US neighbourhoods and cities rapidly became majority black (Card, Mas and Rothstein, 2008; Boustan, 2010, 2012, for the US and Kaufmann and Harris, 2015, for the UK) but other studies have also estimated preferences for racial composition (e.g. Bayer, Ferreira and McMillan, 2007). Studies of residential mobility reveal the preferences of the marginal resident of a neighbourhood but find it harder, without further assumptions, to make more general statements about the impact of diversity on preferences. In contrast, our variable on neighbourhood satisfaction is informative about the preferences of infra-marginal residents which may be the most important impact especially in countries like the UK where residential mobility is quite low. To give an extreme example, if residential mobility costs are so high nobody ever changes neighbourhood studies of residential mobility would yield no information

[1] There are some exceptions that focus on the impact on diversity on slightly different outcomes. Most of these study the MTO experiment in the US (Kling, Liebmann and Katz, 2007, Ludwig et al, 2012, and Chetty, Hendren and Katz, 2016) and exploit the random provision of housing vouchers to encourage low-income families to more to lower-poverty neighbourhoods. An example applied to a different context is Algan, Hemet and Laitin (2016). They exploit the random assignment of tenants to apartment blocks in France to investigate the impact of ethnic diversity on social relationships and housing quality.

about preferences over neighbourhood composition but our variable would be able to identify the effect.

A further area where this paper makes a contribution is in its study of the UK[2]. Most of the literature on diversity studies the US, and there is a risk that conclusions do not generalise to other countries e.g. because of the different history and nature of inter-ethnic relations[3]. The influential study of Putnam (2007) concludes by noting that the rise in diversity is probably here to stay and societies need to work out how to manage its consequences – consideration of other countries can then be helpful in deciding whether some impacts are inevitable or can be mediated. The UK is an interesting country to study the impact of diversity as the last 25 years have seen substantial change in the structure of the population, which means that many neighbourhoods have undergone quite sizeable changes. The fraction of the population that is non-white has risen from 5% in the 1991 census to 15.5% in the 2011 census. However the pattern of residential segregation is very different in the UK and US which might indicate differences in attitudes to diversity. Figure 1 presents the density of the white share of the population in US census tracts in 2010 and UK wards in 2011[4]. The US shows a marked bimodality which might be taken as prima facie evidence that people care about the ethnic composition of their neighbourhood. But the UK does not show any evidence of bimodality possibly suggesting ethnic diversity is less salient in the UK. We do find some differences between the UK and the US, notably that we fail to find any significant link between diversity and generalized trust. However we do find that diversity does have a significant impact on overall neighbourhood satisfaction.

To summarize we think our paper makes a contribution in the following areas: by investigating the impact of diversity on a wider range of outcomes than considered by most of the literature, by paying more attention than most studies to confounding factors and causality, by providing information about the impact of diversity on infra-marginal neighbourhood residents and by providing estimates of the impact of diversity for a country other than the US.

The plan of the paper is as follows. In the next section we describe the individual data that we use, and introduce the outcome measures we study. The third section presents the

[2] The existing studies of the UK (Laurence and Heath, 2008; Letki, 2008; Andrews, 2009; Fieldhouse and Cutts, 2010; Twigg, Taylor and Mohan, 2010; Laurence, 2011; Becares et al, 2011; Sturgis et al, 2011; Demireva and Heath, 2014) have focused on the link between diversity and trust with the main focus being whether neighbourhood deprivation or diversity is the most important factor.

[3] For example black-white interactions might be thought to be different in the two countries as evidence by the different levels of inter-racial marriage. In the US, Fryer (2007) reports that in 2000 about 5% of blacks in the US are in an inter-racial marriage while for the UK in 2001, 35% of blacks are in a an inter-racial couple.

[4] Both tracts and wards have average population of about 4000 so are broadly comparable in size

neighbourhood data and discusses the measure of diversity we use. The fourth section describes our empirical methodology and the instrumental variables and selection controls we use to address endogeneity of our variable of interest and of residential choices. The fifth section reports our basic results for satisfaction with the neighbourhood – we find robust evidence that a lower white share is associated with lower neighbourhood satisfaction. The sixth section presents evidence that neighbourhood satisfaction does predict intention to move and actual residential moves. The seventh section considers outcomes that have often been studied in the social capital literature (trust, and activity in organizations) – we find no significant, robust, relationship with diversity. The eighth section considers satisfaction with particular aspects of the neighbourhood - perceptions of crime, and quality of local services and social life. We find negative effects of the white share on perceptions of crime, while results on the quality of services and of social life are more mixed and generally not significantly different form zero. The ninth section provides an account of the transmission mechanism from neighbourhood characteristics to overall satisfaction with the neighbourhood. These estimates cannot be thought of as causal but we argue that they are informative on which channels may play a role in explaining how diversity affects neighbourhood satisfaction.

2. Individual Data

This study uses data for the period 1991-2014 from the British Household Panel Study (BHPS)[5] and its successor Understanding Society (UKHLS)[6], two longitudinal British surveys with a similar sample structure to the PSID in the US, though asking a wider range of questions on social attitudes. They follow a representative sample of households over time, interviewing all individuals aged 16 or above[7]. BHPS started in 1991 and lasted for 18 waves, finishing in 2008. The first wave included around 10,300 individuals from 5,500 households in Great Britain[8]. UKHLS started in 2009, and is still ongoing. The first wave surveyed individuals from approximately 40,000 households. Since 2010, UKHLS also includes the whole BHPS panel[9] that, at the time, surveyed individuals from about 8,000 households. Together, they

[5] University of Essex. Institute for Social and Economic Research. (2014). *British Household Panel Survey, Waves 1-18, 1991-2009: Special Licence Access, Lower Layer Super Output Areas and Scottish Data Zones*. [data collection]. *3rd Edition.* UK Data Service. SN: 6136, http://dx.doi.org/10.5255/UKDA-SN-6136-2

[6] University of Essex. Institute for Social and Economic Research, NatCen Social Research, Kantar Public. (2016). *Understanding Society: Waves 1-6, 2009-2015: Special Licence Access, Census 2001 Lower Layer Super Output Areas*. [data collection]. *7th Edition.* UK Data Service. SN: 6670, http://dx.doi.org/10.5255/UKDA-SN-6670-7

[7] Since 1994, BHPS includes a short module for individuals aged 11-15.

[8] Following the first wave sampling, new entrants in the sample are mainly represented by people reaching the minimum age for the interview and people joining the original households. Additional samples of households from Scotland and Wales were included in 1999, and for Northern Ireland in 2001.

[9] The attrition rate for the BHPS panel between 2008 and 2010 was of 20%.

allow for the construction of a panel that covers more than 20 years. They include a wide variety of detailed questions on perceptions and attitudes towards the neighbourhood where people live. Unfortunately, not all questions appear in each year and there is no year in which all questions we consider appear[10] – this has implications for our empirical enquiry that we discuss below. Table 1 presents descriptive statistics on the variables, grouped into broad categories. First, there is our main question of interest – overall, whether you like your present neighbourhood – 92.5% of people do. We also consider the fraction who plan to stay in their current neighbourhood (67.9%) and the actual mobility from one year to the next (6.7%). The next panel of Table 1 considers some measures of social capital – generalized trust, whether active in or member of at least one organization and whether one is willing to improve one's neighbourhood. The third panel of Table 1 presents measures relating to the perception of crime, both an overall worry about being a victim of crime (47.8% of people are) as well as fears about specific types of crime. As a summary index of fear of crime we use the overall "worry" question as this has the largest sample size having been a question in more waves of the survey. The fourth panel of Table 1 summarizes responses to questions about the quality of local services – schools, medical, transport, shopping and leisure. The final panel of Table 1 summarizes measures relating to the quality of social life – interactions and friendships with neighbours etc. We also combine measures of social capital, quality of local services, and quality of life in separate indices that are constructed using the first principal component of each group of variables. Principal component analysis results are reported in Table B2 of the Appendix.

Although they are not the focus of interest we also include individual-level controls as regressors in most of our specifications – these are summarized in Table 2. The individual-level controls comprise information on age, working status, housing status, education level, gender, marital status, number of children, and a dummy variable for being non-white.

3. Neighbourhood level data

The geo-coded versions of the BHPS and UKHLS also contain detailed information on the residence of the respondents in each wave, specifically the Lower Super Output Area (LSOA)

[10] Table B1 in the Appendix lists the relevant questions and the waves in which they appear.

level[11] that will be our main geographical reference[12]. There are 40,880 LSOAs in Britain[13], containing on average 1,416 people[14]. The sample size of BHPS/UKHLS is too small to be able to compute reliable neighbourhood characteristics at this spatial scale so we use other data sources to measure them, mostly the decennial censuses, 1991-2011 inclusive.

Measuring Diversity

The main variable of interest in our study is ethnic diversity. The number of ethnic groups categorized varies across censuses and in our analysis we use nine groups that can be defined on a consistent basis - White, Indian, Pakistani, Bangladeshi, Chinese, Black Caribbean, Black African, other Asian, other Black, and a residual category grouping together all other ethnicities. We impute values for the inter-censual years using linear interpolation for each area. We use information from 1991 and 1971 censuses to construct an instrumental variable for the ethnic mix, as explained in the following section.

The existing literature uses a variety of measures to summarize the ethnic mix of an area. One popular measure is the fractionalization index (see e.g. Alesina and La Ferrara, 2000), that is defined as:

$$FRAC_n = 1 - \sum_g s_{gn}^2 \quad (1)$$

where s_{gn} is the share of ethnic group g in neighbourhood n. This can be interpreted as the chance that two randomly chosen people in the area belong to different groups[15]. The fractionalization index is simply one of many possible ways in which the ethnic mix of a neighbourhood might affect outcomes. It might be that it is only the share of one's own ethnic groups that is important or it might be the shares of particular ethnic groups. Or it might also be that it is the immigrant share rather than the ethnic group that is important. In the UK historically most minorities were migrants and most migrants were minorities. But this correlation has weakened over time – many minorities are now UK–born and there has been

[11] LSOAs were created from 2001 census of population to improve the reporting of small area statistics, they were then revised according to 2011 census. Both BHPS and UKHLS contain information at the 2001 LSOA level. In 2001, LSOAs in England and Wales were constructed to have a minimum of 1,000 inhabitants, and 400 households, and a maximum of 3,000 inhabitants, and 1,200 households. Scotland designed statistical areas following the same criteria. Where other area codes were available, information was harmonised using Postcodes Directories (EDINA, University of Edinburgh) and Postcode Headcounts (Office for National Statistics).

[12] In all text neighbourhood means Lower Super Output Area.

[13] 32,476 of which are in England, 6,502 in Scotland, and 1,896 in Wales. Northern Ireland is excluded from this analysis.

[14] This datum refers to the 1991 *Census of Population*.

[15] An alternative interpretation relates to individuals putting a positive weight on their own-group share so that the 'treatment' effect varies across ethnic groups within their neighbourhood. The fractionalization index is then the average treatment effect across neighbourhoods.

substantial white immigration into the UK following the accession of the Eastern European A8 countries into the EU in 2004.

In principle one can distinguish between these different hypotheses by conducting a 'horse race', testing one measure against another and seeing which has the greatest explanatory power. In practice we do not have enough power in the data to resolve this question beyond reasonable doubt as there is a high degree of collinearity between different possible diversity measures as shown in Table 4[16]. For example, the correlation between the white share and the ethnic fractionalization index is -0.96 because there are only a small number of neighbourhoods where the minority share is very high[17]. The practical implication is that one cannot distinguish clearly between the hypothesis that it is the white share that is the relevant neighbourhood characteristic or the fractionalization index. This is in spite of the fact that these have different implications e.g. a linear effect of the white share implies a monotonic relationship between the white share and outcomes while the fractionalization index does not – however the values of the white share where they are different is sparse in our data.

After some experimentation we decided to use the white share as the diversity measure in our main specification as this can be considered a parsimonious model for diversity that seems to work best for most specifications (see Becares et al, 2011, for another study that uses ethnic group shares as the diversity measure). But we recognize that others might prefer other variables and we also present results for a variety of alternative specifications – including the fractionalization index, the shares of people in the other ethnic groups as additional variables, as well as the share of migrants and the share of Muslims. While we do think that our paper presents robust evidence that there is an impact of some measure of ethnic mix on neighbourhood satisfaction, we think it remains a more open question as to the exact measure of the ethnic mix that is important.

One other issue that we do not explore is the difference between fractionalization and segregation (e.g. see Uslaner, 2012; Alesina and Zhuravskhaya, 2011). Many measures of segregation (see Massey and Denton, 1988, for a classic review) depend not just on the ethnic composition of the own neighbourhood but on its comparison with the ethnic composition of a wider area (e.g. a city) and the segregation index applies to this larger area. A neighbourhood

[16] Information on the country of birth is also available for 1991-2011 censuses but the country of birth classification changes quite extensively across censuses. To estimate the migrant mix we use 4 groups that are consistently available throughout all censuses – United Kingdom, Ireland, Europe, and Other countries.

[17] Less than 1% of the LSOAs has a proportion of white people that is lower than 50% of whites in 1991, and, even though the white share fell over time, only 5% of areas had less than 50% of white residents in 2011.

with a low white share would not be classed as a segregated neighbourhood if surrounding areas also had the same white share. It is ultimately an empirical matter which measures affect people's welfare[18] but it is perhaps plausible that it is the ethnic composition of the local neighbourhood itself that is the most important. However, Echenique and Fryer (2007) develop a "spectral segregation index" that can be used to compute measures of segregation for individuals using information on social networks. We do not have direct measures of social networks but if all connections are within the neighbourhood and those connections are drawn at random from people in the neighbourhood then the spectral segregation index corresponds to the white share (Ballester and Vorsatz, 2014). So our preferred measure of diversity can – under some assumptions – be given a theoretical justification.

Other Neighbourhood Characteristics

Even though our main interest is in the measures of diversity described above, our specifications include time varying controls for deprivation, which is likely to be another factor influencing satisfaction and which has received a lot of attention in the UK literature (e.g. Demireva and Heath, 2014 inter alia). We use two measures. The first is the 'claimant count' (an administrative measure of the numbers claiming Job Seekers Allowance, the UK's unemployment-related benefit) normalized by the working age population so this can be interpreted as a measure of the unemployment rate and we refer to it as such. The claimant count is available at the LSOA level on an annual basis through NOMIS[19]. We include this as a control variable in all our specifications. The second is the UK government's Index of Multiple Deprivation (IMD) that combines a range of indicators of disadvantage including the claimant count. This is the measure of deprivation most commonly used in the existing literature for the UK (e.g. Demireva and Heath, 2014). The IMD is available at five-yearly intervals so has to be interpolated for intervening years. There is the concern that some of the indicators in the IMD might be considered as potential outcomes, and the IMD varies in the way it is constructed across UK countries. However, our results are robust to whether we include or exclude the IMD. As an additional time varying area level control, we include the fraction of households in the area who are homeowners: we experimented with other time-varying controls but this was the only one we found to be significant.

[18] We provide some comparison with the effect of diversity calculated at different spatial scales in the Appendix, finding that our results are almost entirely explained by the local area level variation rather than the one measured at higher layers.
[19] This corresponded to the count of the number of people claiming Jobseeker's Allowance (JSA) from 1996 until 2012. With the introduction of the Universal Credit system in 2013, means tested elements of JSA have been replaced by this new system (Nomis, Official Labour Market Statistics).

In specifications that do not include neighbourhood fixed effects we control for some other time-fixed area characteristics, namely the baseline ethnic and country of birth mix[20], the 1991 industrial composition, the size of the neighbourhood and a dummy for the urbanisation of the area in 1991[21]. This is important as some studies (e.g. Sturgis et al, 2011) have argued that the estimated impacts of diversity are sensitive to the other neighbourhood controls that are included.

We now turn to the empirical specification we use.

4. Empirical Specification

We are interested in how the white share (or some other neighbourhood characteristics) affects an outcome variable, y . Suppose we can model the outcome variable for individual i in neighbourhood n in period t, y_{int}, as:

$$y_{int} = \beta^n W_{n,t} + \beta^c x^c_{int} + u_{int} \tag{2}$$

where x^c_{int} are individual and neighbourhood characteristics, while $W_{n,t}$ is the white share in neighbourhood n in period t, and u_{int} residuals that might have both a neighbourhood and an individual component.

There is a number of issues associated with the estimation of (2). First it might be that even if people were randomly assigned to their neighbourhoods (which they are not), the white share is correlated with unobserved individual or neighbourhood characteristics so that the errors in (2), u_{int}, are not independent of the white share W_{nt} and OLS estimation of (2) would lead to bias. One strategy for dealing with this issue is to control for a wide range of individual and neighbourhood characteristics and the longitudinal data we use is helpful in that regard as we can include individual and/or area fixed effects. This may not eliminate all bias (e.g. there could be other time-varying covariates) so our main strategy is to instrument for the white share (the precise instrument is described later) i.e. to assume there is a set of instrumental variables z_{nt} independent of u_{int} but correlated with W_{nt}. The first-stage of this instrumental variable approach will then be:

$$W_{nt} = \pi^n z_{nt} + \pi^c x_{int} + \eta_{int} \equiv \widehat{W}_{nt} + \eta_{int} \tag{3}$$

[20] Constructed upon 5 country of birth groups: United Kingdom and Europe, Africa, India, Pakistan, and Other countries, data from 1971 *Census of population*.
[21] Table 3 shows descriptive statistics for those control variables.

Sample Selection

Even if one assumes that the errors in (2) - u_{int} - are independent of W_{nt}, each individual is only observed in one neighbourhood in each period. The observed neighbourhood (and hence the observed white share) is therefore likely to be correlated with u_{int} as individuals are more likely to be found in neighbourhoods with a white share that they prefer. In other words, the neighbourhood in which we observe people is the result of a choice. We do have evidence that people do respond in this way – e.g. 'white flight', the process by which some US neighbourhoods and cities rapidly became majority black (Card, Mas and Rothstein, 2008; Boustan, 2010, 2012, for the US and Kaufmann and Harris, 2015, for the UK). A less dramatic example would be the literature on how immigration into an area affects the migration decisions of natives (Borjas, 1987, 1994; Borjas, Freeman and Katz, 1996; Card and di Nardo, 2000; Card, 1990, 2001, 2005; Saiz and Wachter, 2011; Amior, 2015).

The first step of the approach we use to deal with the endogeneity of the choice is to substitute (3) into (2) to have:

$$y_{int} = \beta^n \widehat{W}_{nt} + \beta^c x_{int}^c + \beta^n \eta_{int} + u_{int} \quad (4)$$

We then assume that neighbourhood choice is based on the maximization of some objective function, V_{int} which is given by:

$$V_{int} = \gamma^n W_{nt} + \gamma^c x_{it}^c + \nu_{int} \quad (5)$$

Where the residuals ν_{int} may be correlated with (η_{int}, u_{int}) but are assumed independent of the instruments z_{nt}. Substituting (3) into (5) leads to:

$$V_{int} = \gamma^n \widehat{W}_{nt} + \gamma^c x_{it}^c + \gamma^n \eta_{int} + \nu_{int} \quad (6)$$

Following Das, Newey and Vella (2003), we also assume that the expectation of the error in (4) can be written as a function of the propensity scores, p_{int}, the probability of individual i choosing neighbourhood n, conditional on the covariates, and the neighbourhood being chosen. In the non-binary case, this will generally be the probabilities of choosing all neighbourhoods, not just the chosen one[22]. That is one can write (4) as:

[22] This might seem an arbitrary assumption but it is satisfied by all the most commonly used discrete choice models. See Das, Newey and Vella (2003) for details. Dahl (2002) building on Lee (1983) presents assumptions under which the sample selection term depends only on the probability of the chosen option – as this probability depends on the payoffs from all options this leads to an empirical formulation similar to the more general case.

$$E[y_{int}|X_{int}, D_{int} = 1] = \beta^n \widehat{W}_{nt} + \beta^c x_{int}^c + \lambda(p_{i1t}, \dots, p_{iNt}) \quad (7)$$

where are X_{int} all the individual and area level characteristics for which we control, and D_{int} is a binary variable taking the value 1 if individual i is observed in neighbourhood n at time t. The final term $\lambda(p_{i1t}, \dots, p_{iNt})$ can be thought of as a more complicated version of the familiar sample selection correction term popularized by Heckman. Using (6), the propensity scores can be written in the following form:

$$p_{ijt} = p_j[\gamma^n(\widehat{W}_{1t} - \widehat{W}_{nt}), \dots, \gamma^n(\widehat{W}_{Nt} - \widehat{W}_{nt})] \quad (8)$$

The propensity scores have this form because only the differences in the white shares affect choices, while the individual characteristics cancel out. In our context, where the number of neighbourhoods that an individual might choose is very large, it is not computationally straightforward to estimate (7) and (8) in its general form. Our approach is to approximate the terms using a linear form. That is, we write (7) as:

$$E[y_{int}|X_{int}, D_{int} = 1] = \beta^n \widehat{W}_{nt} + \beta^c x_{int}^c + \sum_{j \neq n} \omega_{nj}[\gamma^n(\widehat{W}_{jt} - \widehat{W}_{nt})] \quad (9)$$

where ω_{nj} is the weight put on the white share of neighbourhood j in influencing the sample selection term for neighbourhood n. It is natural to assume that more distant neighbourhoods have less influence and we assume that the weights have the form:

$$\omega_{nj} = e^{-\alpha d_{nj}} \quad (10)$$

Where d_{nj} is the distance between the neighbourhoods and α is a measure of the cost of distance. Using (10) in (9) leads to:

$$E[y_{int}|X_{int}, D_{int} = 1] = \beta^n \widehat{W}_{nt} + \beta^c x_{int}^c + \beta^a(\widehat{W}_{(\therefore n)t} - \widehat{W}_{nt}) \quad (11)$$

Where:

$$\widehat{W}_{(\therefore n)t} = \sum_{j \neq n} e^{-\alpha d_{nj}} \widehat{W}_{jt} \quad (12)$$

i.e. the sample selection correction term is a function of the difference between the chosen neighbourhood white share and a weighted average of other neighbourhood white shares[23]. In implementing (12) we use a value α=0.03 derived from census data on residential moves

[23] Weights ω_{nj} are rescaled to sum up to 1.

though our results are not sensitive to this choice over plausible values[24]. The sample selection term can be interpreted as a first-order Taylor series approximation to the difference in utility between living in this neighbourhood and others in a discrete choice model where one nest is the current neighbourhood and the other nest is all other neighbourhoods.

The intuition for how one can distinguish between the impact of the absolute value of the white share on neighbourhood satisfaction and sample selection is that the fact that people have to live somewhere means that the choice of neighbourhood will be influenced by the white share in this neighbourhood relative to that in other possible residential choices. So sample selection will be determined by the relative not the absolute white share as shown in (11).

Equation (11) is not quite estimable because it includes the predicted values for the white share. To estimate the model, we replace with the actual white share and instrument it. For the own neighbourhood we use an own-neighbourhood instrument and for the sample selection term we use the equivalent relative instrument.

Instrumental variable

As an instrument for the white share we use using an instrument popularized by Altonji and Card (1991) and Card (2001). This type of instrument builds on the idea that, for historical reasons, area varies in terms of ethnic composition. Denote by s_{gnt} the share of minority group g in employment in neighbourhood n in year t. The white share (and other ethnic mix measures we use in the Appendix) can all be written as a function of the current ethnic shares $I(s_{1nt,}, \dots, s_{Gnt})$. Our instrument for the ethnic shares is constructed in the following way. Denote by s_{gn0} the share of minority group g in neighbourhood n in some base year. Denote by $(logP_{gt} - logP_{gt-1})$ the change in log population of minority group g at time t. Then we define the predicted ethnic mix based on national population movements as:

$$\hat{s}_{gnt} = \frac{s_{gn0}(logP_{gt} - logP_{gt-1})}{\sum_{g'} s_{g'n0}(logP_{g't} - logP_{g't-1})} \quad (13)$$

And we use as the instrument for the ethnic mix measure $I(\hat{s}_{1nt,}, \dots, \hat{s}_{Gnt})$. In levels equations when we do not have neighbourhood fixed effects we control for the initial ethnic mix in the area: this ensures that variation in the instrument comes from the interaction of initial ethnic

[24] Specifically, to determine the value of α we exploit LSOA level flow data from 2001 and 2011 censuses of population. We estimate the probability of moving from one area to other group areas - based on the distance from the area of origin – as a function of the distance between areas. The resulting α is approximately 0.03. As a robustness check we also set α to be equal to 1. Results are quite similar in magnitude to the ones displayed in the tables.

mix and the changing population shares of different ethnic groups in the UK as a whole. Our first stages, reported in Table X, are always very strong.

5. Results for Overall Satisfaction with Neighbourhood

This section reports results for regressions where the dependent variable is a dummy taking value 1 for people answering *Yes* to the question *Do you like your neighbourhood?*. On average 92.5% of people do. The variable we are interested in is the white share (our measure of diversity). To convey the robustness of the results we report a wide range of specifications:

- With or without individual fixed effects, area fixed effects, (individual*area) fixed effects, and (individual+area) fixed effects
- OLS and IV
- With and without corrections for sample selection
- In levels and differences

These vary in the type of variation in the white share that is being used to estimate the effect on neighbourhood satisfaction. These specifications are much more wide-ranging and demanding than those found in almost all of the existing literature. For example, Putnam (2007), only presents cross-sectional evidence without any fixed effects, controlling for endogeneity or sample selection (though there is a discussion of many of these issues). The results for the levels specifications are contained in Table 5 and the differences specifications in Table 6.

Levels

The first column of the top panel (which we will refer to as specification 1A) shows the results for a model estimated by OLS and without any fixed effects or sample selection effects although it does contain individual characteristics and a variety of baseline neighbourhood characteristics, as well as time varying measures of local deprivation. This specification is the closest to those estimated in most of the literature on the impact of diversity on outcomes such as the ones we consider[25]. It uses variation in the white share across the neighbourhoods chosen by different individuals and variation over time in the white share of the same neighbourhood. There is a significant positive effect of the white share on neighbourhood satisfaction – a fall of 10 percentage points (approximately the change over the period 1991-2011) in the white

[25] This is not to say there are no studies that attempts to deal with endogeneity issues – for example, see Leigh (2006) and Bjørnskov (2007) for papers using an IV approach in trust equations.

share is estimated to reduce neighbourhood satisfaction by 1.2 percentage points. So column 1A suggests that diversity may be important. This is broadly consistent with the findings in Letki (2008) who includes neighbourhood satisfaction as one component of her 'neighbourhood attitude' index.

However, it is possible that these effects cannot be interpreted as causal, as they may be biased for a number of reasons. It may be that the types of individuals who live in more diverse areas are different in some unobserved way that also affects neighbourhood satisfaction. A natural way to explore this hypothesis is to exploit the longitudinal nature of our data and include individual fixed effects – results when we do so are reported in column 2A of Table 5. One is now using variation in the white share for the same individual, both from changes within neighbourhoods and changes that result from residential mobility. The diversity variable remains significantly different from zero and the estimated effect is almost double that found in column 1A. This might be what one would expect e.g. individuals who are more tolerant of diversity might be found in more diverse areas.

Column 3A reports results when neighbourhood fixed effects are included. In this specification the impact of the white share is somewhat smaller than that found in column 1A. In this specification one is exploiting variation in the white share within neighbourhoods over time but there might be correlations with individual characteristics. For this reason, we also compare results that include different fixed effects for each individual-area cell (Column 4A), i.e. a different fixed effect if an individual changes area. In this specification one is only using variation in the white share within neighbourhoods over time. The estimated impact of the white share is now similar to the individual fixed effect specification. Finally, column 5A includes both area and individual fixed effects – results are very close to the individual fixed effects estimates.

The use of fixed effects does not entirely rule out the possibility that our results are driven by some endogeneity bias e.g. because there may be some omitted time-varying covariates, other than our measures of deprivation of the area and the time varying individual level characteristics as employment status and house tenure we are controlling for[26]. For this reason, we introduce an instrumental variable strategy i.e. we instrument the diversity with variables that one can argue are uncorrelated with unobserved neighbourhood characteristics. The IV estimates are in specifications 1B-5B that mirror the specifications 1A-5A. The first stages are

[26] Tables 2 and 3 illustrate the full set of characteristics that we include in our specifications.

reported in Table B3 in the Appendix. The Kleibergen-Papp test statistic is reported and the values suggest that the instruments are generally very strong. The estimated coefficients are very similar to the OLS estimates though the standard errors are larger and the coefficient estimate with area fixed effects only is not significantly different from zero at conventional significance levels (the instrument becomes weaker in this case as well though still very strong). However, including individual as well as area fixed effects or an individual*area fixed effect leads to larger and more significant estimated effects. Overall, the results seem quite robust.

Even the IV estimates do not control for possible sample selection of individuals into areas and the right-hand panel of Table 5 presents estimates when a sample selection term is included along the lines of that suggested in the previous section – essentially a weighted average of the diversity variable in the surrounding neighbourhood. The coefficients on these sample selection terms are reported in Table B4 in the Appendix. When the sample selection term is included, the estimated coefficients are, for the most part, similar to those in the equivalent specification without sample selection terms. This is true for both OLS and IV estimates. However, the Kleibergen-Papp test statistics do suggest that the instruments become weak once area fixed effects are included, but one should also take into account that these are very demanding specifications – there are 2 endogenous variables once the sample selection term is included as well as a great number of fixed effects.

The sample selection terms themselves are not generally significant from zero. This is not too surprising given that the rates of residential mobility are low so that most individuals are in the same neighbourhood from one year to the next.

Overall, Table 5 suggests a significant positive effect of the white share on neighbourhood satisfaction. The estimates suggest that a fall in the white share of 10 percentage points is likely to reduce neighbourhood satisfaction by 1-3 percentage points (the sample average is 92.5%) with perhaps a central estimate of 2 percentage points.

First Differences

Equation (2) can also be written in first-differenced form as:

$$\Delta y_{int} = \beta^n \Delta W_{it}^{FD} + \beta^c \Delta x_{int}^c + \Delta u_{int} \qquad (14)$$

where we define $\Delta W_{it}^{FD} = W_{n(i,t)t} - W_{n(i,t-1)t-1}$, using $n(i,t)$ to represent the area in which individual i lives at time t. The first column Table 6 reports estimates for the model in this form, both OLS (Panel A) and IV (Panel B) without and with sample selection. First-differencing is an alternative way to eliminate individual fixed effects so we do not report specifications with them included. Column (1A) estimates the model by OLS and column (1B) by IV. The OLS results are in line with the results for 'levels' that an increase in the white share increases neighbourhood satisfaction (though the magnitude of the effect is larger in the differences specification) – in the IV specification the results are quite similar.

One should notice, in fact, that while for people who do not move one year from the other the variation in ΔW_{it}^{FD} comes only from changes over time in the white share, for people who move the variation comes from changes in the white share across areas and over time. The latter source of variation might be thought to be problematic because the residential mobility decision is clearly endogenous. To deal with this, one might consider estimating a model in which the change in diversity is measured for the original area. This could be interpreted as an 'intention to treat' (ITT) estimator as some individuals can avoid to face the decrease in the white share – the "treatment" in this case - by moving to a different area. Hence, define $\Delta W_{it}^{ITT} = W_{n(i,t-1)t} - W_{n(i,t-1)t-1}$ i.e. the change in the white share experienced by the area in which the individual was living last year.

The ITT estimate does not directly measure how much people care about the white share in their current neighbourhood that is the main aim of our enquiry. Suppose people do care about diversity but residential mobility is very high and the range of neighbourhoods on offer so great that any change in the current neighbourhood that one disliked could be avoided by moving to another area. In this case the ITT estimate would be zero but it would be wrong to conclude people do not care about their neighbourhood – it would be more accurate to say that residential mobility insures them against any changes they do not like. Similarly, if residential mobility itself is costly, one should not conclude that individuals do not care about diversity on the basis that the ITT estimate is zero.

Using ΔW_{it}^{ITT} as the regressor leads to the results reported in column (2A) for the OLS estimator and column (2B) for the IV estimator (we call this the ITT estimate). The estimated coefficients are very different from the equivalent specifications in columns (1A) and (1B). Most strikingly, the coefficient on the ITT white share in the OLS specification becomes negative, though not significantly different from zero.

To investigate the differences between the FD and ITT results, we include both white share variables in the same regression i.e. estimate:

$$\Delta y_{int} = \beta_1^n \Delta W_{it}^{FD} + \beta_2^n \Delta W_{it}^{ITT} + \beta^c \Delta x_{it}^c + \Delta u_{in(i,t)t} \quad (15)$$

Columns 3A and 3B present estimates of (15). Column 3A shows that ΔW_{it}^{FD} always has a positive sign, significantly different from zero. However, conditional on ΔW_{it}^{FD} , ΔW_{it}^{ITT} has a negative coefficient in most specifications but has a positive insignificant coefficient in the instrumental variable estimate without sample selection. A significant effect of ΔW_{it}^{ITT} could be interpreted as suggestive that how the white share evolves in an area that individuals have left affects satisfaction with the current neighbourhood. A negative coefficient could perhaps interpret it as individuals experiencing relief if they have moved away from an area that was becoming less white because they have avoided changes they would have been uncomfortable with. This explanation is speculative, but the results do suggest that neighbourhood satisfaction may not simply be driven by characteristics of the current neighbourhood. Finally, we also consider first-difference specifications within individual*neighbourhood pairs i.e. using only within area changes for each individual – these are reported columns 4A (for OLS) and 4B (for IV)[27]. Also in this case we find that a higher white share is associated with higher neighbourhood satisfaction. The right-hand panel of Table 6 presents estimates of fixed difference specifications with sample selection corrections. The first-difference specifications are similar though the instruments become weaker once sample selection is accounted for.

Heterogeneity in Coefficients

One obvious concern with the estimates presented so far is that they assume that all individuals are affected by neighbourhood characteristics in the same way i.e. the effects are homogeneous.

First, it is quite possible that individuals prefer to be surrounded by their own ethnic group so that the coefficient on the white share would be different, possibly differently signed for whites and ethnic minorities. 89.9% of our sample is white so the estimates reported above will largely reflect their preferences but the preferences of minorities might well be different[28]. This has been explored by Becares et al (2011) who use data from the British Citizenship Survey that

[27] This approach is similar to the one used by Dustmann and Fasani (2016) to estimate the impact of crime on mental health at the local level.

[28] BHPS does not over-sample minorities though UKHLS does. However, UKHLS is a small part of our sample. See Knies, Nandi and Platt (2016) for an analysis of UKHLS data showing that people have higher life satisfaction when surrounded by co-ethnics. The short sample period does not allow for as rich a set of covariates as we consider.

over-samples ethnic minorities to investigate the impact of diversity and deprivation on social cohesion.

Other differences often discussed are that the old and the less educated may be less comfortable with diversity than the young, and the more highly educated, or that home ownership is important because it affects the ability to move areas and any impact on house prices might also be a consideration.

Table 7 investigates possible heterogeneity introducing in the baseline specifications interactions of the diversity variable with individuals' ethnicity, level of education, age, and home ownership status[29]. We present four sets of estimates, OLS and IV, with and without individual fixed effects. Across all four specifications graduates' neighbourhood satisfaction is found to be less affected by the white share, while people aged more than 50 appear more strongly affected, as expected. For the non-whites, the estimates with and without fixed effects are different – without fixed effects, non-whites are less affected by the white share but with fixed effects they are more affected. People who own a house seem to be more affected by diversity than private tenants but results are only significant in the IV specification. Overall, the main evidence for heterogeneity is that graduates seem less concerned about the white share.

Other Robustness Checks

In an earlier section we discussed how it is hard to separately identify whether the right variable to measure diversity is the white share or the fractionalization index. So it is also possible that our results could really be picking up the impact of some other correlated measure of diversity. For this reason, we explore results using alternative measures for diversity. Table 8 shows the results obtained with IV strategy, while in the Appendix (Table B5) we show the OLS counterpart. We use the fractionalization index (Panel A), the white share and the black share (Panel B), the white share and the Asian share (Panel C), the white share and the Pakistani/Bangladeshi share (to pick up the possible hostility towards Muslims) (Panel D) and the white share and the immigrant share (Panel E). An overall theme is that there is a robust significant impact of diversity on neighbourhood satisfaction but that the high degree of collinearity between different diversity measures means that one can be less sure about exactly which aspect of diversity is important. When we compare the white share coefficient to other

[29] We also considered gender but this was never significant.

group shares (Panels B-E) the white share tends to prevail, although there is loss of precision in some specifications, in particular when area fixed effects are introduced.

In the existing UK literature (Laurence and Heath, 2008; Letki, 2008; Andrews, 2009; Fieldhouse and Cutts, 2010; Twigg, Taylor and Mohan, 2010; Laurence, 2011; Becares et al, 2011; Sturgis et al, 2011; Demireva and Heath, 2014) on the impact of diversity on social capital, there is considerable discussion of the impact of deprivation. All our estimates control for the time-varying unemployment rate as a measure of deprivation. There is perhaps some independent interest in the impact of the unemployment rate on neighbourhood satisfaction (see also, Kling, Liebmann and Katz, 2007, and Ludwig et al, 2012 for the impact of neighbourhood poverty on various measures of well-being) and these results are reported in Appendix C. There is a significant impact of unemployment on neighbourhood satisfaction in some specifications but not all. Because one might also be concerned about the endogeneity of the unemployment rate, Appendix C also reports results when it is treated as endogenous.

We have assumed that it is the LSOA that is the appropriate level of geographical aggregation for affecting neighbourhood satisfaction. We check this using two alternative methods. Firstly, we estimate the same model using Travel to Work Areas (TTWA) fixed effects instead of neighbourhood fixed effects (Table B6 in the Appendix), finding similar results. Secondly, we control for measures of diversity calculated both at the neighbourhood and at the TTWA level and we find that the effect of neighbourhood level diversity is essentially unaltered and is the most important factor (Table B7 in the Appendix).

Because data on ethnic mix is only available at Census years we use interpolation for the intervening years. One might be concerned that this interpolation influences the results in some ways. But Table B8 in the Appendix shows that the results are very similar if we restrict attention to Census years.

Summary

Most of the specifications that we report suggest that an increase in the white share increases neighbourhood satisfaction among our (overwhelmingly white) sample. We find this in levels and differences, with and without individual fixed effects, allowing for endogeneity and sample selection. However, one should recognize that there are limits to how robust these conclusions are and we have tried to be open about that – when instrumental variables, area fixed effects

and sample selection corrections are included the instruments become weaker, the standard errors larger and the estimated coefficients not significantly different from zero.

Most of the estimated coefficients are in the region 0.1-0.3 with perhaps a central estimate around 0.2. As the white share has fallen by about 10 percentage points in the period 1991-2011 these estimates would imply that neighbourhood satisfaction has fallen by between 1 and 3 percentage points over this period because of rising diversity. This effect is not enormous but the baseline probability is 92.5% so this is perhaps a sizeable rise in the fraction who are not satisfied with their neighbourhood.

6. Residential Mobility

One potential criticism of the analysis so far is that response to the neighbourhood satisfaction question simply reflects people's subjective response to which no significance can be attached. One way of addressing this is to consider whether responses to the neighbourhood satisfaction question are correlated with intentions to move neighbourhood (itself subjective) and actual residential mobility.

In Table 9 we provide evidence that satisfaction with the neighbourhood has predictive power for the decision to move. Panel A shows results for the actual moving. The dependent variable is a binary variable taking the value 1 if the person is observed in a different LSOA in time t than in time t-1. This is regressed on the lagged values of the neighbourhood satisfaction in columns (1) to (3) of the first row. In all specifications, a higher satisfaction is associated with a lower probability of moving. The bottom part of panel A shows a reduced form counterpart of the previous models, as it shows the results of a regression of the probability on moving on the white share. Results show that the white share is negatively associated to the probability of moving once fixed effects are taken into account. Panel B of Table 9 shows that current neighbourhood satisfaction is also strongly correlated with the expression of an intention to move. Our conclusion is that responses to the neighbourhood satisfaction question are informative.

7. Social Capital

Our analysis so far has focused on the impact of diversity on the level of satisfaction with the neighbourhood. While we would argue this is an outcome of interest as a summary measure of how well the neighbourhood 'works' for individuals, much of the literature on the impact of diversity on community focuses on two commonly-used measures of 'social capital' –

generalized trust and activity in organizations. This section considers these two outcomes. These questions are asked in fewer years than neighbourhood satisfaction so the sample sizes are much smaller particularly. A number of prominent authors have suggested that diversity erodes generalized trust (e.g. Putnam, 2007), a view that has been the subject of considerable controversy (e.g. Nannestad, 2008, Uslaner, 2012; Gerritsen and Lubbers, 2010; Gesthuize, Van der Meer and Scheepers, 2009).

Our method for controlling for sample selection into neighbourhoods has rarely found evidence that this is important in practice, perhaps unsurprising given the low rate of residential mobility. In the interests of brevity and clarity, the estimates that follow only report specifications without controls for sample selection – though our results are very similar if they are included.

The upper panel of Table 10 presents results using generalised trust as the outcome variable. We find no evidence that diversity affects the level of trust even in the OLS model without fixed effects that is closest to the specification commonly estimated in the existing literature. This result is very different from the conventional wisdom on the topic mostly derived from US evidence. It could be that there is no inevitable close link between generalised trust and diversity and that the US results are driven by the particular interactions between ethnic groups within that society that do not translate to other societies.

As an alternative measure of social capital we use a dummy variable taking value one for people who are active in at least one organisation - results are reported in the lower panel of Table 10. We find no significant impact of diversity on activity in organizations in any specification.

8. Other Neighbourhood Outcomes

So far we have investigated the impact of diversity on neighbourhood satisfaction and residential mobility. While there is some impact on overall satisfaction we have not provided any evidence on the aspects of the neighbourhood that changes that influences overall satisfaction. This section investigates this. We consider possible impacts on the perception of crime, the quality of local services and the quality of one's social life.

Fear of Crime

The questions asked about perceptions and fear of crime are listed in Table 1. For this section we only use the answers to whether the respondent worries about being a victim of crime because the sample size is largest for this question. The results are shown in Table 11 where

the top panel reports OLS estimates with a variety of individual and area fixed effects and the bottom panel the equivalent specifications but using IV. The estimates suggest that a higher white share is generally associated with a significantly lower level of concern about crime.

One should note that this is fear of crime and not actual crime and these may not be the same. For example, the literature on the link between migration and crime sometimes finds an impact on fear of crime but little impact on actual crime once one controls for labour market status (that would be expected to affect crime incentives as argued by Becker, 1968)[30].

Quality of Local Services

Table 12 does a similar exercise for the quality of local services index, which is the first principal component of the quality of local services variables in Table 1. One should note that the sample sizes for these outcomes are even smaller than for the fear of crime variable and many specifications have very large standard errors. This together with the considerably variability in the estimated coefficient on the white share across specifications mean that the results for this outcome are not very robust.

Quality of Social Life

Table 13 does a similar exercise for the quality of social life index, which is the first principal component of the quality of social life variables in Table 1. All the estimated coefficients are positive suggesting that a high white share is associated with a higher quality of social life for our respondents. However, the standard errors are very large so that many of the estimated coefficients are not significantly different from zero.

9. A Production Function for Neighbourhood Satisfaction

So far we have documented what we have argued are the causal effects of neighbourhood characteristics on various measures of feelings about neighbourhoods, from the high level overall satisfaction to different domains such as social capital, fear of crime, quality of local services and social life. One hypothesis is that feelings about specific domains go into

[30] Bell and Machin, (2013), provide a broader literature review on the topic, here we report some of the most recent works on the argument. In the US, Chalfin (2014) finds no causal effect of Mexican migration on crime. Spenkuch (2014) finds that there is some small effect, concentrated on property and financial crimes, and for migrants with low labour market prospects. Moheling and Piehl (2009, 2014) find that prison commitment rates for new migrants is in general lower or equal to the natives' one. Evidence from Europe has been targeted mostly in studying the EU enlargements during the 2000s (Bell, Fasani and Machin, 2013, Bianchi, Buonanno and Pinotti, 2012, Mastrobuoni and Pinotti, 2015) find modest effect of migration on property crimes only. Nunziata (2015) finds no effect on victimisation, but a significant impact on fear of crime. Sà (2015) finds that the negative effect of migration on house prices is not explained by any migration related increase in crime.

producing an overall satisfaction with the neighbourhood. A simple linear production function for individual i's overall satisfaction SAT_i, would be:

$$SAT_i = \sum_j \beta_j f_{ji} + \beta^c x_i + \epsilon_i \quad (16)$$

Where f_{ji} is the level of feeling about domain j for individual i. What we have estimated are the impact of neighbourhood characteristics on these feelings i.e.:

$$f_{ji} = \gamma_i W_i + \gamma_j^c x_i + u_{ji} \quad (17)$$

Substituting (17) into (16) implies that overall satisfaction can be written as:

$$SAT_i = \left[\sum_j \beta_j \gamma_j\right] W_i + \left[\beta^c + \sum_j \beta_j \gamma_j^c\right] x_i + \epsilon_i + \sum_j \beta_j u_{ji} \quad (18)$$

which is what we have also estimated. What is not identified in this estimates are the factor loadings – the β_j - on different domains in (16). These are however of some interest e.g. whether it is crime or social interactions that is the main transmission channel from neighbourhood characteristics to overall satisfaction. Ideally one would estimate (16) instrumenting the different domains using the first-stages implied by (17). However, this approach only works if we have at least as many instruments as domains and there is independent variation in the domains.

In the absence of our ability to do that, we report estimates of (16) by OLS. These estimates cannot be given a causal interpretation but we do think they are of some interest. Results that include factors[31] for the different groups of variables are reported in Table 14[32]. Columns (1), (3) and (5) report estimates of neighbourhood satisfaction on the various domain i.e. (16). Neighbourhood satisfaction is, as one might expect, positively related to social capital, quality of local services and social life and negatively related to the fear of crime though the importance of social capital does not survive the introduction of individual fixed effects. The relation with deprivation is negative, but significant only when individual fixed effects are considered. Our earlier results would suggest that the impact of diversity is not through the social capital variables that has been the focus of much of the literature but through fear of crime, the quality of services and social life. Columns (2), (4) and (6) include both the diversity variables and the domain satisfaction measures showing that the white share still has explanatory power: this

[31] Factors are obtained grouping single variables with Principal Component Analysis techniques.
[32] Table B9 in the Appendix reports results obtained including all variables. In both cases, as there is no year for which all questions are asked, information is pooled for the two closest years in which information is available.

suggests that we have not identified all the channels through which diversity affects satisfaction with the neighbourhood.

10. Conclusion

This paper has investigated the impact of the white share of the population on a variety of measures of neighbourhood satisfaction. Our estimates suggest that diversity does affect overall neighbourhood satisfaction and the fear of crime, confirming the general conclusion of the existing literature. But, there is one area where our findings are at odds with the existing literature – we do not find any significant link between diversity, generalized trust and activity in organizations, commonly used measures of social capital.

The value-added of the paper is to consider a wider variety of neighbourhood outcomes than the existing literature, to pay closer attention to issues of causality and endogeneity through the use of fixed effects, instrumental variables and sample selection, and to provide estimates on the impact of diversity on infra-marginal residents which may be the most important effects especially when residential mobility rates are low.

While people may care about the nature of their neighbours, they cannot control who they are. My presence in an area may have some externalities on my neighbours, and my decision to move is not something they can control[33]. This combination of caring about something but being unable to control it is the classic recipe for stress so it is not surprising that changing communities stir up strong emotions and reactions. As Putnam (2007) noted, the rise in diversity is probably here to stay and societies need to work out how to manage its consequences to make communities thrive.

[33] If there is, as seems likely, residential sorting, we have known since the work of Schelling (1971, 1972) that there is no presumption that the resulting equilibrium is efficient. There is no strong prediction on whether there is too little or too much segregation in equilibrium but a number of studies have documented the impact of segregation on wages, rental prices, and in general on economic performance (e.g. Cutler and Glaeser, 1997, Peri and Ottaviano, 2006; Ananat, 2011; Chetty et al, 2014).

References

Alesina, A., Baqir, R., and Easterly, W. (1999). "Public goods and ethnic divisions". *The Quarterly Journal of Economics*, vol. 114, no. 4, pp. 1243–1284.

Alesina, A. and La Ferrara, E. (2000). "Participation in Heterogeneous Communities". The *Quarterly Journal of Economics*, vol. 115, no. 3, pp. 847-904.

Alesina, A. and La Ferrara, E. (2002). "Who trusts others?". *Journal of Public Economics*, vol. 85, no. 2, pp. 207–234.

Alesina, A. and La Ferrara, E. (2005). "Ethnic Diversity and Economic Performance". *Journal of Economic Literature*, vol. 63, no. 3, pp. 762–800.

Alesina, A. and Zhuravskaya, E. (2011). "Segregation and the Quality of Government in a Cross Section of Countries.", *The American Economic Review,* vol. 101, no. 5, pp. 1872-1911.

Algan, Y., Hémet, C., and Laitin, 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, no. 3, pp. 696-733.

Altonji, J. and Card, D. (1991) "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives," in Immigration, Trade and the Labor Market, NBER Chapters, pp. 201–234. National Bureau of Economic Research, Inc.

Amior, M. (2015). "Why are Higher Skilled Workers More Mobile Geographically? The Role of the Job Surplus", CEP Discussion Paper 1338.

Ananat, E.O. (2011). "The Wrong Side(s) of the Tracks: The Causal Effects of Racial Segregation on Urban Poverty and Inequality", *The American Economic Journal: Applied Economics*, vol. 3, no. 2, pp. 34-66.

Andrews, R. (2009). "Civic Engagement, Ethnic Heterogeneity and Social Capital in Urban Areas: Evidence from England", *Urban Affairs Review*, vol. 44, no. 3, pp. 428-440.

Ballester, Coralio Ballester and Marc Vorsatz (2014) "Random Walk-Based Segregation Measures", *Review of Economics and Statistics*, 96, 383-401.

Bartik, T. J. (1991). *Who Benefits from State and Local Economic Development Policies?,* W.E. Upjohn Institute for Employment Research: Kalamazoo, MI.

Bayer, P, Ferreira, F and McMillan, R. (2007) "A Unified Framework for Measuring Preferences for Schools and Neighborhoods", *Journal of Political Economy*, Vol. 115, pp. 588-638

Becares L., Stafford M., Laurence J., and Nazroo, J. (2011). "Composition, concentration and deprivation: Exploring their association with social cohesion among different ethnic groups in the UK", *Urban Studies*, vol. 48, no. 13, pp. 2771–2787.

Becker, G. (1968). "Crime and punishment: An economic approach", *Journal of Political Economy*, vol. 76, no. 2, pp. 169-217.

Bell, B., Fasani, F. and Machin, S. (2013). "Crime and immigration: Evidence from large immigrant waves", *The Review of Economics and Statistics*, vol. 95, no. 4, pp. 1278-1290.

Bell, B. and Machin, S. (2013). "Immigration and crime", International Handbook on the Economics of Migration, ch. 19, pp. 353-372.

Bianchi, M., Buonanno, P. and Pinotti, P. (2012). "Do immigrants cause crime?", *Journal of the European Economic Association*, vol. 10, no. 6, pp. 1318-1347.

Bjørnskov, Christian (2007). "Determinants of generalized trust: A cross-country comparison," Public Choice, Springer, vol. 130, no. 1, pp. 1-21.

Borjas, G.J. (1987). "Self-Selection and the Earnings of Immigrants", *The American Economic Review*, vol. 77, no. 4, pp. 531-553.

Borjas, G.J. (1994). "The Economics of Immigration", *Journal of Economic Literature*, vol. 32, no. 4, pp. 1667-1717.

Borjas, G. J., Freeman, R. B., and Katz, L. F. (1996). "Searching for the Effect of Immigration on the Labor Market", *The American Economic Review*, vol. 86, no. 2, pp. 246–251.

Boustan, L. P. (2010). "Was Postwar Suburbanization 'White Flight'? Evidence from the Black Migration", *Quarterly Journal of Economics*, vol. 125, no. 1, pp. 417-443.

Boustan, L. P. (2012). "School Desegregation and Urban Change: Evidence from City Boundaries", *The American Economic Journal: Applied Economics*, vol. 4, no. 1, pp. 85-108.

Card, D. (1990). " The Impact of the Mariel Boatlift on the Miami Labor Market", *Industrial and Labor Relations Review*, vol. 43, no. 2, pp. 245–257.

Card, D. (2001). "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration", *Journal of Labor Economics*, vol. 19, no. 1, pp. 22-64.

Card, D. (2005). "Is the New Immigration Really so Bad?", *The Economic Journal*, vol. 115, no. 507, pp. F300–F323.

Card, D. and Di Nardo, J. (2000). " Do immigrant inflows lead to native outflows?", *The American Economic Review*, vol. 90, no. 2, pp. 360–367.

Card, D., Mas, A., and Rothstein, J. (2008), "Tipping and the Dynamics of Segregation", *Quarterly Journal of Economics*, vol. 123, no. 1, pp. 177-218.

Chalfin, A. (2014). "What is the contribution of Mexican immigration to US crime rates? Evidence from rainfall shocks in Mexico". *American Law and Economics Review*, vol. 16, no. 1, pp. 220-268.

Chetty, Raj, Nathaniel Hendren and Lawrence F. Katz. (2016) "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review,* 106, 855-902.

Chetty, R., Hendren, N., Kline, P., and Saez E. (2014). "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States", *The Quarterly Journal of Economics*, vol. 129, no. 4, pp. 1553-1623.

Costa, D. L. and Kahn, M. E. (2003). "Civic Engagement and Community Heterogeneity: An Economist's Perspective". *Perspective on Politics*, vol. 1, no. 1, pp. 103-111.

Cutler, D. M. and Glaeser, Edward L. (1997). "Are Ghettos Good or Bad?", *Quarterly Journal of Economics*, vol. 112, no.3, pp. 827-872.

Dahl, G. (2002) "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets", *Econometrica* Vol. 70, pp. 2367-2420

Dahlberg, M., Edmark, K. and Lundqvist, H. (2012). "Ethnic Diversity and Preferences for Redistribution", *Journal of Political Economy*, vol. 120, no. 1, pp. 41-76.

Das, M., Newey, W.K. and Vella, F. (2003). "Nonparametric Estimation of Sample Selection Models", *The Review of Economic Studies*, vol. 70, no. 1, pp. 33-58.

Demireva, Neli and Anthony Heath (2014). "Diversity and the Civic Spirit in British Neighbourhoods: An Investigation with MCDS and EMBES 2010 Data". *Sociology*, vol. 48, pp. 643-662.

Dinesen, P.T. and Sonderskov, K.M. (2012). "Trust in a Time of Increasing Diversity: On the Relationship between Ethnic Heterogeneity and Social Trust in Denmark from 1979 until Today", *Scandinavian Political Studies*, vol. 35, no. 4, pp. 273–294.

Dustmann, C. and Fasani, F. (2016). "The Effect to Local Area Crime on Mental Health", The Economic Journal, vol. 126, no. 593, pp. 978-1017.

Echenique, F. and Fryer, R. (2007). "A Measure of Segregation Based on Social Interactions", *Quarterly Journal of Economics*, vol. 122, no. 2, pp. 441-485.

Fieldhouse, E. and Cutts, D. (2010). "Does Diversity Damage Social Capital? A Comparative Study of Neighbourhood Diversity and Social Capital in the US and Britain." *Canadian Journal of Political Science-Revue Canadienne de Science Politique*, vol. 43, no. 2, pp. 289-318.

Fryer, R. (2007) "Guess Who's Been Coming to Dinner? Trends in Interracial Marriage over the 20th Century." *Journal of Economic Perspectives*, vol. 21, pp. 71-90.

Gerritsen D. and Lubbers M. (2010). "Unknown is unloved? Diversity and inter-population trust in Europe", *European Union Politics* vol. 11, no. 2, pp. 267–87.

Gesthuizen M., Van der Meer T. and Scheepers P. (2009). "Ethnic diversity and social capital in Europe: Tests of Putnam's thesis in European countries". *Scandinavian Political Studies* vol. 32, no. 2, pp. 121–142.

Kaufmann, E. and Harris, G. (2015). " 'White Flight' or Positive Contact?: Local Diversity and Attitudes to Immigration in Britain, 2009-11", *Comparative Political Studies*, pp. 1–28.

Kling JR, Liebman JB, Katz LF. (2007) "Experimental Analysis of Neighborhood Effects". *Econometrica*, 75, 83-119.

Knies, Gundi Alita Nandi and Lucinda Platt. (2016) "Life satisfaction, ethnicity and neighbourhoods: Is there an effect of neighbourhood ethnic composition on life satisfaction?", *Social Science Research*, in press.

Lee, L-F. "Generalized Econometric Models with Selectivity", *Econometrica*, vol. 51, pp. 507-512.

Laurence, J. (2011). "The Effect of Ethnic Diversity and Community Disadvantage on Social Cohesion: A Multi-Level Analysis of Social Capital and Interethnic Relations in UK Communities." *European Sociological Review*, vol. 27.

Laurence, J. and A. Heath (2008). "Predictors of Community Cohesion: Multi-Level Modelling of the 2005 Citizenship Survey." Department for Communities and Local Government, London.

Leigh, Andrew 2006. "Trust, Inequality and Ethnic Heterogeneity," *The Economic Record, The Economic Society of Australia*, vol. 82, no. 258, pp. 268-280.

Letki, N. (2008). "Does Diversity Erode Social Cohesion? Social Capital and Race in British Neighbourhoods.", *Political Studies*, vol. 56, no. 1, pp. 99-126.

Ludwig J, Duncan GJ, Gennetian LA, Katz LF, Kessler R, Kling JR, Sanbomatsu L. (2012) "Neighborhood Effects on the Long-Term Well-Being of Low-Income Adults", *Science* [Internet]. 2012, 337, 1505-10.

Massey, Douglas S.; Denton, Nancy A. (1988) "The Dimensions of Residential Segregation", *Social Forces*, 1988; 67:281–315

Mastrobuoni, G. and Pinotti, P. (2015). "Legal status and the criminal activity of migrants", *American Economic Journal: Applied Economics*, vol. 7, no. 2, pp. 175-206.

Moehling, C. and Piehl, A. Morrison (2009). "Immigration, crime, and incarceration in early twentieth-century America", *Demography*, vol. 46, no. 4, pp. 739-763.

Moehling, C. and Piehl, A. Morrison (2014). "Immigrant assimilation into US prisons, 1900-1930", *Journal of Population Economics*, vol. 27, pp. 173-200.

Nannestad P. (2008). "What have we learned about generalized trust, if anything?", *Annual Review of Political Science*, vol. 11, pp. 413-436.

Nunziata, L. (2015). "Immigration and crime: Evidence from victimisation data", *Journal of Population Economics*, vol. 28, pp. 697-736.

Peri, G. and Ottaviano, G. (2006). "The Economic Value of Cultural Diversity: Evidence from US Cities". *Journal of Economic Geography*, vol. 6, no. 1, pp. 9–44.

Portes, A. and Vickstrom, E. (2011). "Diversity, Social Capital, and Cohesion", *Annual Review of Sociology*, vol. 37, pp. 461-479.

Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community*, New York: Simon and Schuster.

Putnam, R. D. (2007). "Et Pluribus Unum: Diversity and Community in the Twenty-first Century" *The 2006 Johan Skytte Prize Lecture*, vol. 30, no. 2, pp. 137–174.

Sà, F. (2015). "Immigration and house prices in the UK", *The Economic Journal*, vol. 125, pp. 1393-1424.

Saiz, A. and Wachter, S. (2011). "Immigration and the Neighborhood", *The American Economic Journal: Economic Policy*, vol. 3, no. 2, pp. 169-188.

Schelling, T. (1971). "Dynamic Models of Segregation", *Journal of Mathematical Sociology*, vol. 1, pp. 143-186.

Schelling, T. (1972). "A Process of Residential Segregation: Neighborhood Tipping", in Pascal, A. (ed.), *Racial Discrimination in Economic Life*, D.C. Heath, Lexington, MA.

Spenkuch, JL (2014). "Understanding the impact of immigration on crime". *American Law and Economics Review,* vol. 16, no. 1, pp. 177–219.

Sturgis P, Brunton-Smith I, Read S, Allum N, (2011). "Does ethnic diversity erode trust? Putnam's `hunkering-down' thesis reconsidered". *British Journal of Political Science,* vol. 41, no. 1, pp. 57-82.

Tesei, A. (2015). "Trust, Racial Fragmentation and Income Inequality: Evidence from the U. S.", CEP Discussion Paper n. 1331

Twigg, L., Taylor, J., & Mohan, J. (2010). "Diversity or disadvantage? Putnam, Goodhart, ethnic heterogeneity, and collective efficacy", *Environment and Planning A*, vol. 42, pp. 1421-1438.

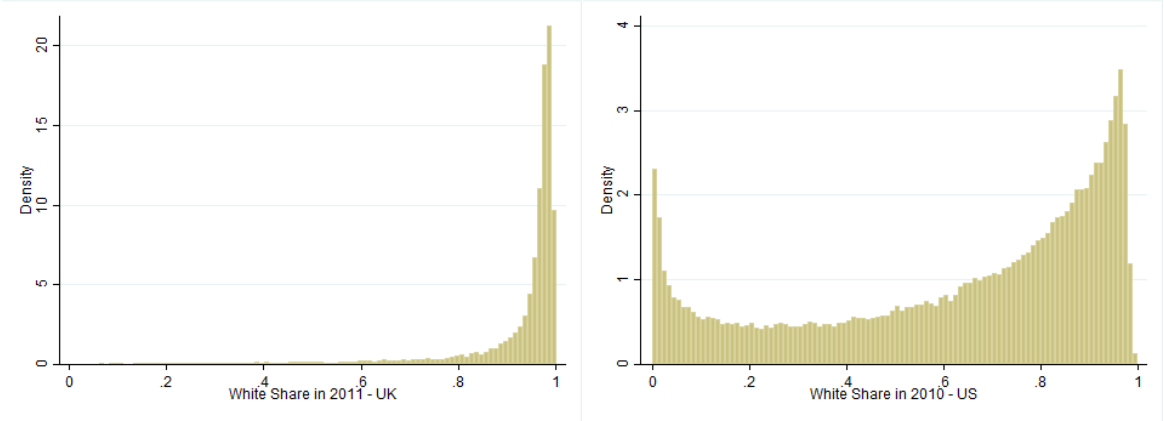
Uslaner, Eric M. (2012). *Segregation and Mistrust: Diversity, Isolation, and Social Cohesion*, Cambridge, Cambridge University Press.

Van der Meer, T. and Tolsma, J. (2014) "Ethnic Diversity and its Effects on Social Cohesion", *Annual Review of Sociology*, vol. 40, pp. 459-478.

Tables / Figures

Figure 1. Distribution of the share of white people across small areas.

Panel A. UK, 2011, Census Wards **Panel B.** US, 2010, Census Tracts



Notes: Authors' elaboration of UK Census Data (2011), and US Census Data (2010).

Table 1. Descriptive statistics for values and attitude variables in the British Household Panel and in Understanding Society.

Variable	Mean	Standard Deviation	N
Satisfaction:			
Like your present neighbourhood	.925	.263	249,084
Plan to stay in your neighbourhood	.679	.467	109,247
Actual mobility	.067	.250	428,516
Social capital:			
Generally speaking, most people can be trusted	.363	.481	110,002
Active in at least one organisation	.582	.493	168,418
Member of at least one organisation	.527	.499	168,650
Willing to improve your neighbourhood	.757	.430	109,322
Crime:			
Worry you're being victim of crime	.478	.500	77,287
Feel unsafe walking alone at night	.184	.387	77,250
Likely home broken into	.238	.426	33,959
Likely car stolen/broken into	.079	.270	33,970
Likely drunks/tramps on the street	.151	.358	34,597
Likely graffiti on the walls	.234	.423	34,665
Likely people being attacked on the street	.084	.278	34,302
Likely racial insults/attacks	.057	.232	33,896
Likely teenagers hanging about	.569	.495	34,632
Likely vandalism	.264	.441	34,638
Quality of local services:			
Good schools	.698	.459	62,005
Good medical services	.705	.456	74,787
Good transportation	.501	.500	71,677
Good shopping facilities	.578	.494	76,234
Good leisure facilities	.442	.497	72,234
Suitable for children	.653	.476	33,337
Social life:			
Meet your neighbours often	.769	.412	149,840
Friends in the local neighbourhood	.609	.488	109,647
Can obtain advice locally	.558	.497	109,406
Can you borrow from people in the neighbourhood	.437	.496	109,236
Feel similar to people in the neighbourhood	.601	.490	109,378
Talk to people in your neighbourhood	.688	.463	109,709
Satisfied with social life	.657	.475	144,185

Notes: All variables are dummies constructed upon affirmative replies to the corresponding questions. Table B1 reports the waves where each variable appears.

Table 2. Descriptive statistics for individual-level control variables for the British Household Panel and Understanding Society sample.

Variable	Mean	Standard Deviation	Min	Max
Age	46.42	18.46	14	104
Female	0.539	0.498	0	1
Unemployed	0.045	0.208	0	1
Retired	0.215	0.411	0	1
Full-time student	0.007	0.085	0	1
Other non working	0.049	0.216	0	1
Married	0.523	0.499	0	1
Number of children	0.511	0.929	0	10
Non-white	0.098	0.296	0	1
Higher education	0.293	0.455	0	1
Low education	0.242	0.428	0	1
No education	0.175	0.380	0	1
Social house tenant	0.180	0.384	0	1
Private tenant	0.099	0.299	0	1

Notes: Excluded dummies for each set – working, high school education, and home owners, for labour force status, education and house status, respectively – are not reported.

Table 3. Descriptive statistics for main area-level variables.

Variable	Mean	Standard Deviation	Min	Max
Britain overall (census data)				
White share	0.918	0.143	0.007	1
Ethnic fractionalization index	0.128	0.181	0	0.870
Immigrant fractionalization index	0.132	0.135	0	0.681
Black share	0.020	0.051	0	0.637
Immigrant share	0.080	0.102	0	0.939
Asian share	0.048	0.101	0	0.997
Muslim share	0.029	0.075	0	0.953
Unemployment rate	0.039	0.035	0	0.986
Area (Ha)	568	2,740	0.73	115,963
Country of birth mix in 1971 (area level)	0.057	0.073	0	0.629
House owners	0.642	0.229	0	1
Urban areas in 1991	0.111	0.314	0	1
BHPS and UKHLS sample				
White share	0.896	0.179	0.004	1
Ethnic fractionalization index	0.152	0.207	0	0.872
Immigrant fractionalization index	0.152	0.148	0	0.683
Black share	0.023	0.057	0	0.617
Immigrant share	0.094	0.118	0	0.757
Asian share	0.059	0.128	0	0.962
Muslim share	0.026	0.086	0	0.952
Unemployment rate	0.037	0.029	0	0.708
Area (Ha)	599	2,152	1.24	77,870
Country of birth mix in 1971 (area level)	0.055	0.073	0	0.629
House owners	0.639	0.229	0	1
Urban areas in 1991	0.106	0.308	0	1

Notes: Area-level information refers to the Lower Super Output Area codification related to the 2001 census. *Britain overall* panel displays descriptives for all English, Welsh, and Scottish LSOAs in the non-interpolated years only. *BHPS and UKHLS sample* panel displays results for the subset of LSOAs that appears in the British Household Panel and in Understanding Society. Se Appendix A for a description of the variables. *Muslim share* is proxied by the share of Pakistani and Bangladeshi people in the area.

Table 4: Correlations.

	White share	Ethnic F- index	Immigrant F- index	Black share	Immigrant share	Asian share	Muslim share	Unempl. rate
White share	1							
Ethnic F-index	-0.958*	1						
Immigrant F-index	-0.809*	0.853*	1					
Black share	-0.727*	0.771*	0.653*	1				
Immigrant share	-0.839*	0.847*	0.977*	0.658*	1			
Asian share	-0.900*	0.800*	0.641*	0.382*	0.683*	1		
Muslim share	-0.832*	0.728*	0.597*	0.420*	0.633*	0.870*	1	
Unemployment rate	-0.147*	0.147*	0.018*	0.193*	0.027*	0.107*	0.139*	1

Notes: * $p < 0.01$. Correlations are estimated for the universe of the 40,880 UK LSOAs, for the years for which census data are available, namely 1991, 2001, and 2011. Se Appendix A for a description of the variables. F-index denotes fractionalization index.

Table 5: The impact of ethnic mix on how you like your neighbourhood.

	No sample selection					Sample selection				
	(1) No FE	(2) Individual FE	(3) Area FE	(4) Individual* Area FE	(5) Individual+ Area FE	(6) No FE	(7) Individual FE	(8) Area FE	(9) Individual* Area FE	(10) Individual+ Area FE
	A. OLS									
White share	0.120***	0.268***	0.090*	0.208***	0.212***	0.170***	0.364***	0.056	0.217***	0.235***
	(0.015)	(0.034)	(0.050)	(0.053)	(0.053)	(0.030)	(0.061)	(0.073)	(0.070)	(0.070)
N	233,548	200,344	229,637	192,296	198,698	233,548	200,344	229,637	192,296	198,698
	B. IV									
White share	0.097***	0.287***	0.075	0.340***	0.329***	0.080**	0.234***	0.045	0.196**	0.229**
	(0.020)	(0.047)	(0.072)	(0.081)	(0.080)	(0.040)	(0.088)	(0.093)	(0.090)	(0.090)
N	231,649	198,590	227,761	190,669	196,957	231,649	198,590	227,761	190,669	196,957
KP	9034.987	2472.030	475.184	1012.597	1034.063	405.313	1162.229	107.076	235.197	245.486

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. Standard errors in parentheses. *OLS* is the baseline specification not including any individual or area Fixed Effects. Standard errors account for clusters at the individual level in OLS, *Individual FE,* and *Area+Individual FE* specifications. Standard errors account for clusters at the area level for *Area FE,* and for clusters ate the area-individual level in the *Area*Individual FE* specifications. *Area FE* specifications include Lower Super Output area Fixed Effects, *Area*Individual FE* include LSOA-individual level fixed effects, and *Area+Individual FE* include both LSOA and individual fixed effects, taken as separate FE sets. Panel B shows Instrumental Variable estimates. Right-hand panels include controls for sample selection, as illustrated by equation (13), with α equal to 0.03. IV-Sample Selection specifications use instrumental variables for both the share of white people and for the corresponding sample selection variable. Sample selection coefficients are reported in Table B3 of the Appendix. All regressions include individual, area-level controls, and year dummy variables. Se Appendix A for a description of the control variables and of the variables of interest. KP is the Kleibergen-Paap weak instrument statistic.

Table 6: The impact of ethnic mix on how you like your neighbourhood. ***First difference results.***

	No Sample Selection				Sample Selection			
	(1)	(2)	(3)	(4) within area*individual	(5)	(6)	(7)	(8) within area*individual
				A. OLS				
White share (first difference)	0.391***		0.370**	0.506***	0.743***		0.742**	0.554***
	(0.046)		(0.047)	(0.095)	(0.105)		(0.113)	(0.134)
White share (ITT)		-0.150	-0.181*			-0.045	-0.466***	
		(0.098)	(0.099)			(0.130)	(0.164)	
N	169,086	169,141	169,074	154,902	169,086	169,141	169,074	154,902
				B. IV				
White share (difference)	0.358***		0.327***	1.117***	0.721***		0.726**	0.535***
	(0.056)		(0.056)	(0.177)	(0.118)		(0.120)	(0.170)
White share (ITT)		-0.222	0.235			-0.055	-0.406***	
		(0.164)	(0.157)			(0.158)	(0.183)	
N	167,672	167,770	167,661	153,692	167,672	167,770	167,661	153,692
KP	3495.885	1959.263	970.911	1723.450	1091.676	388.580	207.817	355.483

Notes: * $p < 0.1$ ** $p < 0.05$ *** $p < 0.001$. Robust standard errors clustered at the individual level in parentheses. In each regression the dependent variable is a dummy that takes value 1 if the respondent states to like the neighbourhood in which he/she lives. See Appendix A for a description of the control variables and of the variables of interest. See Section 4 for a discussion of the various First Difference Specifications, Instrumental Variables, and controls for sample selection. KP is the Kleibergen-Paap weak instrument statistic.

Table 7: Heterogeneity of results. **Dependent variable:** 1 if affirmative answer to *Do you like your neighbourhood?*

	OLS		***IV***	
	(1) No FE	(2) Individual FE	(3) No FE	(4) Individual FE
White share	0.174***	0.304***	0.128***	0.277***
	(0.013)	(0.038)	(0.020)	(0.054)
White share *NonWhite	-0.101***	0.054	-0.077***	0.239***
	(0.014)	(0.059)	(0.015)	(0.072)
White share *Higher Education	-0.050***	-0.188***	-0.080***	-0.275***
	(0.012)	(0.045)	(0.012)	(0.053)
White share*Low Education	-0.000	0.071	0.000	0.098
	(0.002)	(0.062)	(0.002)	(0.080)
White share*No Education	-0.002	0.030	-0.003	0.122
	(0.002)	(0.065)	(0.002)	(0.082)
White share *Age50	0.011***	0.008***	0.010***	0.007***
	(0.002)	(0.003)	(0.002)	(0.003)
White share *Home Owner	0.003	0.002	0.029*	0.093***
	(0.007)	(0.007)	(0.017)	(0.035)
White share*Social Tenant	-0.038**	0.002	-0.023	0.021
	(0.015)	(0.043)	(0.020)	(0.052)
N	237,921	200,385	235,980	198,631

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. Standard errors in parentheses, clustered at the individual level. Please refer to Notes of Table 5 for references about different specifications, and to Appendix A for full control variables description.

Table 8. Comparison between different specifications for diversity. Instrumental Variable results. **Dependent variable:** 1 if affirmative answer to *Do you like your neighbourhood?*

		IV				
		(1)	(2)	(3)	(4)	(5)
		No FE	Individual FE	Area FE	Individual* Area FE	Individual+ Area FE
A	Fractionalisation	-0.016***	-0.038***	-0.016	-0.053***	-0.052***
		(0.003)	(0.007)	(0.012)	(0.013)	(0.013)
	N	231,649	198,590	227,761	190,669	196,957
	KP	1.6e+04	5090.232	569.488	1282.873	1323.010
B	White share	0.110***	0.224***	0.247	0.504*	0.378
		(0.023)	(0.058)	(0.235)	(0.263)	(0.253)
	Black share	0.074	-0.289*	0.917	0.859	0.254
		(0.056)	(0.149)	(1.201)	(1.287)	(1.230)
	N	231,649	198,590	227,761	190,669	196,957
	KP	1950.056	1184.676	6.536	10.013	10.923
C	White share	-0.021	0.272***	0.022	0.199	0.314
		(0.040)	(0.095)	(0.299)	(0.350)	(0.343)
	Asian share	-0.157***	-0.022	-0.080	-0.259	-0.028
		(0.048)	(0.133)	(0.518)	(0.626)	(0.615)
	N	231,649	198,590	227,761	190,669	196,957
	KP	1694.259	769.909	22.320	43.410	44.881
D	White share	0.087***	0.209***	0.030	0.263*	0.314**
		(0.025)	(0.057)	(0.119)	(0.135)	(0.133)
	Muslim share	-0.024	-0.224*	-0.175	-0.315	-0.064
		(0.037)	(0.123)	(0.393)	(0.539)	(0.530)
	N	231,649	198,590	227,761	190,669	196,957
	KP	3999.067	560.241	14.920	25.432	26.493
E	White share	0.272***	0.555***	0.155	1.394	0.825
		(0.031)	(0.078)	(0.617)	(1.000)	(0.963)
	Foreigners share	0.347***	0.493***	0.072	0.925	0.434
		(0.049)	(0.102)	(0.550)	(0.872)	(0.838)
	N	231,556	198,503	227,668	190,589	196,870
	KP	4260.506	738.035	10.546	13.208	13.676

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. See Notes for Table 5. Each panel corresponds to a different set of regressions.

Table 9: Actual moving and propensity to stay in the area. Linear probability models.

A	**(1) No FE**	**(2) Individual FE**	**(3) Area FE**	**(4) FD (lagged)**
Dependent variable: **Actual moving**				
Like your neighbourhood (lagged)	-0.083***	-0.106***	-0.083***	-0.037***
	(0.003)	(0.005)	(0.004)	(0.004)
N	206,240	175,162	201,902	146,461
White share (lagged)	0.026***	-0.182***	-0.103***	-0.032
	(0.006)	(0.029)	(0.051)	(0.031)
N	284,508	272,395	282,262	213,676
IV				
White share (lagged)	0.020***	-0.239***	-0.222***	-0.064*
	(0.007)	(0.043)	(0.089)	(0.039)
N	282,615	270,557	280,359	212,227

B	**(1) No FE**	**(2) Individual FE**	**(3) Area FE**
Dependent variable: **Propensity to stay**			
Like your neighbourhood	0.483***	0.361***	0.408***
	(0.007)	(0.014)	(0.009)
N	68,091	33,268	63,591
White share	0.060***	0.318***	-0.070
	(0.017)	(0.069)	(0.118)
N	98,870	77,985	96,018
IV			
White share	0.010	0.338***	-0.080
	(0.023)	(0.109)	(0.165)
N	98,160	77,400	95,336

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. Standard errors in parentheses. *Actual moving* is a dummy that takes value 1 if the respondent is observed in 2 different LSOAs from one year to the other. *Propensity to stay* is a dummy variable that takes value 1 if the respondent states that he/she is willing to stay in the area. See Notes of Table 5.

organisation, results in levels.

OLS			**IV**				
(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Area FE	Individual*Area FE	Individual+Area FE	No FE	Individual FE	Area FE	Individual*Area FE	Individual+Area FE
Dependent variable: *Generalised trust*							
-0.097	-0.050	-0.079	-0.027	0.159**	0.079	0.141	0.117
(0.112)	(0.124)	(0.130)	(0.028)	(0.074)	(0.182)	(0.209)	(0.220)
95,103	58,986	63,191	97,977	63,897	94,221	58,331	62,449
			7810.244	937.152	507.381	1053.068	957.938
Dependent variable: *Active in any organisation*							
-0.081	-0.045	-0.046	-0.012	-0.036	0.076	0.125	0.096
(0.065)	(0.078)	(0.079)	(0.024)	(0.062)	(0.099)	(0.131)	(0.132)
153,974	112,208	118,870	156,638	119,941	152,921	111,456	117,992
			1.3e+04	2620.269	503.462	933.914	922.968

rrors in parentheses. See notes of Table 5.

Table 11: Crime, results in levels.
Dependent variable: 1 if affirmative answer to *Are you worried of being victim of a crime?*

	(1) No FE	(2) Individual FE	(3) Area FE	(4) Individual*Area FE	(5) Individual+Area FE
			OLS		
White share	-0.094***	-0.141*	-0.459***	-0.357***	-0.350**
	(0.021)	(0.075)	(0.114)	(0.138)	(0.149)
N	73,511	34,921	69,165	28,115	32,850
	(1) No FE	(2) Individual FE	(3) Area FE	(4) Individual*Area FE	(5) Individual+Area FE
			IV		
White share	-0.071**	-0.214*	-1.040***	-1.403***	-1.372***
	(0.028)	(0.111)	(0.183)	(0.236)	(0.255)
N	72,876	34,475	68,564	27,819	32,461
KP	8328.835	1143.694	779.469	1117.506	957.120

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. Standard errors in parentheses. See notes of Table 5.

Table 12: Quality of services in the area, results in levels.
Dependent variable: Quality of services in the area index

	(1) No FE	(2) Individual FE	(3) Area FE	(4) Individual*Area FE	(5) Individual+Area FE
			OLS		
White share	-0.048	0.117	-0.265	-0.960**	-0.950*
	(0.122)	(0.280)	(0.482)	(0.475)	(0.529)
N	22,007	14,943	20,295	11,665	13,239
	(1) No FE	(2) Individual FE	(3) Area FE	(4) Individual*Area FE	(5) Individual+Area FE
			IV		
White share	-0.054	0.647	-0.655	-0.670	-0.724
	(0.154)	(0.413)	(0.694)	(0.697)	(0.776)
N	21,784	14,829	20,105	11,605	13,155
KP	2071.193	753.564	295.232	662.617	532.417

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. Standard errors in parentheses. See notes of Table 5.

Table 13: Quality of social life, results in levels.
Dependent variable: Quality of social life index

	(1) No FE	(2) Individual FE	(3) Area FE	(4) Individual*Area FE	(5) Individual+Area FE
	OLS				
White share	0.348***	0.349*	0.432	0.830**	0.815**
	(0.105)	(0.210)	(0.388)	(0.331)	(0.365)
N	31,356	24,614	29,828	18,767	22,459
	(1) No FE	(2) Individual FE	(3) Area FE	(4) Individual*Area FE	(5) Individual+Area FE
	IV				
White share	0.395***	0.255	0.384	0.782	0.777
	(0.137)	(0.307)	(0.561)	(0.494)	(0.544)
N	30,968	24,301	29,461	18,601	22,224
KP	2514.840	795.335	443.737	1192.307	982.586

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. Standard errors in parentheses. See notes of Table 5.

Table 14. Satisfaction production function.
Dependent variable: 1 if affirmative answer to Do you like your neighbourhood?

	No FE		**Individual FE**		**Area FE**	
	(1)	(2)	(3)	(4)	(5)	(6)
White share		0.114***		0.406***		0.262**
		(0.024)		(0.090)		(0.114)
Social capital	0.007***	0.007***	0.001	0.001	0.003*	0.003*
	(0.001)	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)
Crime	-0.037***	-0.036***	-0.035***	-0.034***	-0.025***	-0.025***
	(0.002)	(0.001)	(0.003)	(0.003)	(0.002)	(0.002)
Quality of local services	0.011***	0.011***	0.002	0.003	0.006***	0.006***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
Social life	0.023***	0.023***	0.018***	0.017***	0.021***	0.021***
	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)	(0.002)
Observations	20,143	20,118	13,273	13,256	18,468	18,440

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. Bootstrapped standard errors (200 replications) in parentheses. The full set of variables is not available for all years. The sample is therefore pooled to aggregate information for closest subsequent years for which information is available. See Appendix A for a description of the main variables and Table B2 in the Appendix for a description of the Principal Component Analysis used to construct the variables *Social Capital*, *Crime*, *Quality of local services*, and *Social life*.

Appendix – For Online Publication Only

A –Variables description

All area level variables refer to the Lower Super Output Areas (LSOA) level. Data for inter census years have been derived from LSOA level linear ipolation.

Other diversity variables

Ethnic mix index of ethnic fractionalisation calculated following Alesina and La Ferrara (2000)

$$Ethnic\ Mix = 1 - \sum_i s_{ia}^2$$

where s_{ia} is the share of people in each ethnic group i in the population area a. The ethnic group considered are White, Indian, Pakistani, Bangladeshi, Chinese, Other Asian, Black Carribean, Black African, Other Black, and Other Ethnic groups. Source: C*ensus of population* 1991-2011.

Migrant mix index of country of birth fractionalisation calculated in the same way as the ***Ethnic mix***, over the share of people born in UK, Europe, Africa, India, Pakistan, and other countries. Source: *Census of population* 1991-2011.

Immigrant share people born outside UK over the total population. Source: *Census of population* 1991-2011.

White share, ***Asian share***, ***Black share***, and ***Muslim share*** (we proxy the Muslim share with the share of Pakistani and Bangladeshi people in the area as religion is not available for all censuses considered). Source: *Census of population* 1991-2011.

Control variables

All regressions control for the following individual level characteristics from BHPS and Understanding Society

Individual level characteristics

- Age
- Gender dummy
- Level of education (Dummy variables: Higher Education – Primary – No education. Excluded category: Secondary education)
- Working status (Dummy variables: Unemployed - Retired – Full time student – Out of the labour force for other reasons, Excluded category: Working)

- Ethnicity (Dummy: Non-white vs White)
- Marital status (Dummy: Married vs Not married)
- Number of children
- House ownership status (Dummy variables: Private tenant, Social house tenant. Excluded category: Home owner)

Time invariant LSOA level characteristics. Source: *Census of population*

- Government Office Regions dummies
- 1971 migrant mix
- 1991 share of people employed in each 1-digit SIC industry
- Dummy for urban areas in 1991
- Logarithm of the size of the LSOA (hectares)

Time-variant LSOA level characteristics

- ***Rate of unemployment*** - average monthly number of people claiming for unemployment related benefits as a share of the working age population. Source: *Business Register and Employment Survey* (BRES) 1991-2014.
- ***Index of Multiple Deprivation*** – it combines information on different domains – income, employment, education, skills and training, health and disability, barriers to housing and services, living environment – at the Lower-layer Super Output Area. Source: Department for Communities and Local Government for England, Welsh Government National Statistics for Wales, and Scottish Government National Statistics for Scotland
- ***House tenure:*** proportion of house owners

B – Additional Tables

Table B1: Values and attitude questions in the British Household Panel (B) and in Understanding Society (U).

Variable	Waves
Satisfaction:	
Like your present neighbourhood	B: all - U: 3
Plan to stay in your neighbourhood	B: 8, 13, 18 - U: 3
Social capital:	
Generally speaking, most people can be trusted	B: 8, 10, 13, 15, 17 - U: 1
Active in at least one organisation	B: 1-5, 7, 9, 11, 13, 15, 17 – U: 3
Member of at least one organisation	B: 1-5, 7, 9, 11, 13, 15, 17 – U: 3
Willing to improve your neighbourhood	B: 8, 13, 18 - U: 3
Crime:	
Worry you're being victim of crime	B: 7, 12, 17 - U: 3
Feel unsafe walking alone at night	B: 7, 12, 17 - U: 3
Likely home broken into	B: 7, 12, 17
Likely car stolen/broken into	B: 7, 12, 17
Likely drunks/tramps on the street	B: 7, 12, 17
Likely graffiti on the walls	B: 7, 12, 17
Likely people being attacked on the street	B: 7, 12, 17
Likely racial insults/attacks	B: 7, 12, 17
Likely teenagers hanging about	B: 7, 12, 17
Likely vandalism	B: 7, 12, 17
Quality of local services:	
Good schools	B: 8, 13, 18 - U: 3
Good medical services	B: 8, 13, 18 - U: 3
Good transportation	B: 8, 13, 18 - U: 3
Good shopping facilities	B: 8, 13, 18 - U: 3
Good leisure facilities	B: 8, 13, 18 - U: 3
Suitable for children	B: 8, 13, 18
Social life:	
Meet your neighbours often	B: 8, 13, 18 - U: 3
Friends in the local neighbourhood	B: 8, 13, 18 - U: 3
Can obtain advice locally	B: 8, 13, 18 - U: 3
Can you borrow from people in the neighbourhood	B: 8, 13, 18 - U: 3
Feel similar to people in the neighbourhood	B: 8, 13, 18 - U: 3
Talk to people in your neighbourhood	B: 8, 13, 18 - U: 3
Satisfied with social life	B: 6-10, 12-18

Table B2. Correlations between indices and Principal Component Analysis.

Panel A. Correlation matrix

	Social capital	Crime	Quality of local services	Social life
Social capital	1			
Crime	-0.111*	1		
Quality of local services	0.116*	-0.184*	1	
Social life and neighbourhood	0.152*	-0.074*	0.206*	1

Notes: * $p < 0.01$.

Panel B. Principal Components Analysis - Eigenvalues

	Social capital	Crime	Quality of local services	Social life and Neighbourhood
1st Component	1.686	3.443	2.033	2.646
2nd Component	1.003	1.073	1.211	0.992
3rd Component	0.909	0.969	0.887	0.866
4th Component	0.402	0.838	0.702	0.818
5th Component	-	0.765	0.595	0.634
6th Component	-	0.667	0.572	0.576
7th Component	-	0.641	-	0.467
8th Component	-	0.570	-	-
9th Component	-	0.563	-	-
10th Component	-	0.472	-	-

Panel C. Principal Components Analysis – 1st Principal Components

Social Capital		**Crime**	
Generally speaking, most people can be trusted	.284	Worry you're being victim of crime	.154
Active in at least one organisation	.656	Feel unsafe walking alone at night	.224
Member of at least one organisation	.665	Likely home broken into	.351
Willing to improve your neighbourhood	.219	Likely car stolen/broken into	.331
		Likely drunks/tramps on the street	.331
		Likely graffiti on the walls	.345
		Likely people being attacked on the street	.351
		Likely racial insults/attacks	.309
		Likely teenagers hanging about	.292
		Likely vandalism	.402
Quality of local services		**Social life and Neighbourhood**	Crime
Good schools	.412	Meet your neighbours often	.319
Good medical services	.450	Friends in the local neighbourhood	.470
Good transportation	.356	Can obtain advice locally	.447
Good shopping facilities	.427	Can you borrow from people in the neighbourhood	.357
Good leisure facilities	.433	Feel similar to people in the neighbourhood	.375
Suitable for children	.362	Talk to people in your neighbourhood	.446
		Satisfied with social life	.105

Table B3. Main results, first stage regressions. **Dependent variable:** 1 if affirmative answer to *Do you like your neighbourhood?*

Panel A: No sample selection

Dependent variables	Specification	White share
White share	No FE	0.960*** (.010)
	Individual FE	0.871*** (.018)
	Area FE	1.255*** (.058)
	Area*Individual FE	1.245*** (.039)
	Area+Individual FE	1.239*** (.039)

Panel B: Sample selection

Dependent variable		Instrumental Variable First Stage Coefficients	
		White share	Sample selection: White share
White share	No FE	0.987*** (.016)	-0.032*** (.010)
	Individual FE	1.075*** (.038)	-0.225*** (.033)
	Area FE	1.529*** (.059)	-0.448*** (.085)
	Area*Individual FE	1.554*** (.039)	-0.527*** (.050)
	Area+Individual FE	1.548*** (.039)	-0.522*** (.049)
Sample selection: White share	No FE	0.562*** (.016)	0.329*** (.013)
	Individual FE	0.145*** (.031)	0.684*** (.031)
	Area FE	0.112* (.057)	0.971*** (.082)
	Area*Individual FE	0.135*** (.038)	0.910*** (.048)
	Area+Individual FE	0.129*** (.038)	0.912*** (.047)

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. Standard errors in parentheses. See Table B3 for notes on control variables, variables of interest, and clusters.

Table B4. Main results, levels. Coefficients of sample selection control variables.

Dependent variable: 1 if affirmative answer to *Do you like your neighbourhood?*

	OLS					IV				
	(1) No FE	(2) Individual FE	(3) Area FE	(4) Individual* Area FE	(5) Individual+ Area FE	(7) No FE	(8) Individual FE	(9) Area FE	(10) Individual*Area FE	(11) Individual+ Area FE
Sample selection:										
White share	-0.056*	-0.110*	0.048	-0.013	-0.032	0.019	0.060	0.053	0.270*	0.186
	(0.031)	(0.066)	(0.082)	(0.079)	(0.079)	(0.041)	(0.098)	(0.137)	(0.139)	(0.137)
N	233,548	200,344	229,637	192,296	198,698	231,649	198,590	227,761	190,669	196,957

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. See Notes for Table 5.

Table B5. Comparison between different specifications for diversity. Ordinary Least Squares results.

Dependent variable: 1 if affirmative answer to *Do you like your neighbourhood?*

		OLS				
		(1)	(2)	(3)	(4)	(5)
		No FE	Individual FE	Area FE	Individual* Area FE	Individual+ Area FE
A	Fractionalisation	-0.020***	-0.034***	-0.013*	-0.028***	-0.029***
		(0.002)	(0.005)	(0.007)	(0.007)	(0.007)
	N	233,548	200,344	229,637	192,296	198,698
B	White share	0.103***	0.185***	0.006	0.097	0.084
		(0.017)	(0.042)	(0.066)	(0.071)	(0.070)
	Black share	-0.107**	-0.416***	-0.448*	-0.606**	-0.700**
		(0.046)	(0.118)	(0.243)	(0.278)	(0.272)
	N	233,548	200,344	229,637	192,296	198,698
C	White share	0.137***	0.376***	0.260***	0.449***	0.496***
		(0.029)	(0.062)	(0.101)	(0.111)	(0.110)
	Asian share	0.022	0.164*	0.287*	0.409**	0.481***
		(0.035)	(0.089)	(0.153)	(0.168)	(0.167)
	N	233,548	200,344	229,637	192,296	198,698
D	White share	0.136***	0.271***	0.147**	0.294***	0.314***
		(0.019)	(0.039)	(0.064)	(0.064)	(0.063)
	Muslim share	0.037	0.007	0.188	0.296*	0.348**
		(0.030)	(0.082)	(0.159)	(0.164)	(0.164)
	N	233,548	200,344	229,637	192,296	198,698
E	White share	0.220***	0.406***	0.123*	0.090	0.121*
		(0.022)	(0.047)	(0.067)	(0.070)	(0.070)
	Foreigners share	0.209***	0.260***	0.045	-0.164**	-0.126
		(0.032)	(0.057)	(0.073)	(0.080)	(0.079)
	N	233,548	200,344	229,637	192,296	198,698

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. See Notes for Table 5. Each panel corresponds to a different set of regressions.

Table B6: The impact of diversity on how you like your neighbourhood. Travel to Work Area Fixed effects.

	No sample selection			Sample selection		
	(1)	(2)	(3)	(4)	(5)	(6)
	TTWA FE	Individual* TTWA FE	Individual+TTWA FE	TTWA FE	Individual*TTWA FE	Individual+TTWA FE
	A. OLS					
White share	0.117***	0.293***	0.258***	-0.055	0.382***	0.279***
	(0.025)	(0.037)	(0.034)	(0.099)	(0.069)	(0.065)
N	233,548	198,713	200,343	233,548	198,713	200,343
	B. IV					
White share	0.095***	0.365***	0.286***	-0.276***	0.220**	0.112
	(0.021)	(0.053)	(0.047)	(0.079)	(0.088)	(0.083)
N	231,649	196,980	198,589	231,649	196,980	198,589
KP	502.941	1491.946	2405.018	87.842	732.307	1199.174

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. Standard errors in parentheses.

Table B7: The impact of diversity on how you like your neighbourhood. Controls for Travel to Work Area variables and fixed effects

	No sample selection					Sample selection				
	(1) No FE	(2) Individual FE	(3) TTWA FE	(4) Individual* TTWA FE	(5) Individual+ TTWA FE	(6) No FE	(7) Individual FE	(8) TTWA FE	(9) Individual* TTWA FE	(10) Individual+ TTWA FE
					A. OLS					
White share	0.111***	0.251***	0.125***	0.274***	0.253***	0.127***	0.317***	-0.064	0.319***	0.258**
	(0.016)	(0.037)	(0.030)	(0.042)	(0.037)	(0.040)	(0.085)	(0.144)	(0.112)	(0.104)
White share TTWA level	0.042*	0.077	-0.117	0.091	0.025	0.036	0.045	0.005	0.061	0.021
	(0.023)	(0.050)	(0.089)	(0.067)	(0.064)	(0.029)	(0.063)	(0.084)	(0.096)	(0.091)
N	233,548	200,344	233,548	198,713	200,343	233,548	200,344	233,548	198,713	200,343
					B. IV					
White share	0.083***	0.279***	0.105***	0.372***	0.239	-0.062	0.143	-0.391*	0.118	-0.002
	(0.022)	(0.052)	(0.024)	(0.062)	(0.151)	(0.056)	(0.128)	(0.206)	(0.159)	(0.187)
White share TTWA level	0.056**	0.024	-0.184***	-0.062	0.119	0.110***	0.080	0.117	0.089	0.219
	(0.025)	(0.058)	(0.058)	(0.087)	(0.142)	(0.032)	(0.075)	(0.130)	(0.126)	(0.210)
N	231,649	198,590	231,649	196,980	196,957	231,649	198,590	231,649	196,980	198,589
KP	5108.061	1089.258	240.745	596.793	166.474	228.332	800.170	5.228	409.724	19.073

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. Standard errors in parentheses.

Table B8. The impact of diversity on how you like your neighbourhood. Census years only

	No sample selection					Sample selection				
	(1) No FE	(2) Individual FE	(3) Area FE	(4) Individual* Area FE	(5) Individual+Area FE	(6) No FE	(7) Individual FE	(8) Area FE	(9) Individual*Area FE	(10) Individual+Area FE
						A. OLS				
White share	0.079***	0.199***	0.057	0.132*	0.117	0.115***	0.198**	0.033	0.101	0.100
	(0.012)	(0.053)	(0.067)	(0.073)	(0.081)	(0.028)	(0.089)	(0.097)	(0.098)	(0.108)
N	63,235	21,093	58,789	13,489	18,302	63,235	21,093	58,789	13,489	18,302
						B. IV				
White share	0.051***	0.260***	-0.059	0.092	0.072	0.041	-0.029	0.003	0.051	0.054
	(0.017)	(0.079)	(0.087)	(0.107)	(0.118)	(0.035)	(0.126)	(0.118)	(0.117)	(0.129)
N	62,767	20,899	58,358	13,409	18,182	62,767	20,899	58,358	13,409	18,182
KP	9461.367	1139.929	433.315	615.413	511.181	676.883	475.979	101.246	146.624	124.778

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. Standard errors in parentheses. See Appendix A and notes of Table 5 for notes on control variables, variables of interest, and clusters.

Table B9. Satisfaction production function.
Dependent variable: 1 if affirmative answer to Do you like your neighbourhood?

	No Fixed Effects		**Individual FE**		**Area FE**	
	(1)	(2)	(3)	(4)	(5)	(6)
White share		0.080***		0.384***		0.257**
		(0.026)		(0.089)		(0.121)
Social capital						
General trust	0.012***	0.012***	0.010*	0.010*	0.010***	0.010***
	(0.003)	(0.003)	(0.006)	(0.006)	(0.004)	(0.004)
Active in any organisation	-0.000	-0.000	-0.000	-0.001	-0.001	-0.001
	(0.004)	(0.004)	(0.006)	(0.006)	(0.005)	(0.005)
Member of any organisation	0.007*	0.007*	-0.003	-0.003	0.001	0.001
	(0.004)	(0.004)	(0.007)	(0.007)	(0.005)	(0.005)
Willing to improve your neighbourhood	0.016***	0.016***	0.014*	0.014*	0.012**	0.012**
	(0.005)	(0.005)	(0.008)	(0.008)	(0.006)	(0.006)
Crime						
Worry being victim of a crime	-0.003	-0.003	-0.003	-0.003	-0.009**	-0.009**
	(0.003)	(0.003)	(0.005)	(0.005)	(0.004)	(0.004)
Worry alone at night	-0.044***	-0.043***	-0.054***	-0.052***	-0.045***	-0.045***
	(0.005)	(0.005)	(0.009)	(0.009)	(0.006)	(0.006)
Likely home broken into	-0.011**	-0.011**	-0.009	-0.010	-0.007	-0.008
	(0.005)	(0.005)	(0.007)	(0.007)	(0.006)	(0.006)
Likely car stolen	-0.077***	-0.077***	-0.064***	-0.066***	-0.060***	-0.061***
	(0.011)	(0.011)	(0.016)	(0.016)	(0.012)	(0.012)
Likely drunk/trumps	-0.024***	-0.023***	-0.053***	-0.050***	-0.015*	-0.015*
	(0.007)	(0.007)	(0.011)	(0.011)	(0.009)	(0.009)
Likely graffiti	-0.015***	-0.014***	-0.024***	-0.023***	-0.013**	-0.014**
	(0.005)	(0.005)	(0.009)	(0.009)	(0.006)	(0.006)
Likely people being assaulted	-0.047***	-0.045***	-0.036**	-0.032*	-0.033***	-0.032***
	(0.012)	(0.012)	(0.017)	(0.017)	(0.012)	(0.012)
Likely racial insults	-0.055***	-0.051***	-0.049**	-0.046**	-0.034**	-0.034**
	(0.013)	(0.013)	(0.020)	(0.020)	(0.015)	(0.015)
Likely teens hanging about	-0.002	-0.003	-0.004	-0.005	0.006	0.006
	(0.003)	(0.003)	(0.006)	(0.006)	(0.004)	(0.004)
Likely vandalism	-0.034***	-0.035***	-0.028***	-0.028***	-0.023***	-0.022***
	(0.005)	(0.005)	(0.008)	(0.008)	(0.006)	(0.006)

Table B9 (cont'ed). Satisfaction production function.
Dependent variable: 1 if affirmative answer to Do you like your neighbourhood?

	No Fixed Effects		**Individual FE**		**Area FE**	
	(1)	(2)	(3)	(4)	(5)	(6)
Quality of local services						
Good schools	0.001	-0.000	-0.005	-0.005	-0.005	-0.004
	(0.004)	(0.004)	(0.007)	(0.007)	(0.005)	(0.005)
Good medical services	-0.001	-0.001	0.003	0.003	0.003	0.003
	(0.004)	(0.004)	(0.006)	(0.006)	(0.005)	(0.005)
Good transports	0.001	0.003	-0.005	-0.002	-0.002	-0.002
	(0.003)	(0.003)	(0.006)	(0.006)	(0.004)	(0.004)
Good shopping facilities	0.004	0.005	-0.004	-0.004	0.005	0.004
	(0.004)	(0.004)	(0.006)	(0.006)	(0.004)	(0.004)
Good leisure facilities	-0.002	-0.001	-0.001	-0.000	0.000	0.001
	(0.004)	(0.004)	(0.006)	(0.006)	(0.004)	(0.004)
Good children facilities	0.060***	0.058***	0.028***	0.027***	0.033***	0.032***
	(0.005)	(0.005)	(0.007)	(0.007)	(0.005)	(0.005)
Social life						
Meet Neigh. Often	0.006	0.006	0.018**	0.019**	0.011*	0.010*
	(0.005)	(0.005)	(0.009)	(0.009)	(0.006)	(0.006)
Have friends in the neighb.	0.027***	0.027***	0.015**	0.016**	0.025***	0.025***
	(0.005)	(0.005)	(0.008)	(0.008)	(0.005)	(0.005)
Can have advice in the neighb.	0.017***	0.017***	0.014**	0.013**	0.017***	0.017***
	(0.004)	(0.004)	(0.007)	(0.007)	(0.005)	(0.005)
Can borrow from neighbours	-0.008**	-0.008**	0.007	0.007	-0.004	-0.004
	(0.003)	(0.003)	(0.006)	(0.006)	(0.004)	(0.004)
Feel similar to neighbourhood	0.034***	0.033***	0.022***	0.021***	0.026***	0.026***
	(0.004)	(0.004)	(0.007)	(0.007)	(0.005)	(0.005)
Talk to neighbours	0.022***	0.021***	0.010	0.009	0.022***	0.022***
	(0.005)	(0.005)	(0.009)	(0.009)	(0.006)	(0.006)
Satisfied with social life	0.033***	0.032***	-0.003	-0.004	0.023***	0.023***
	(0.004)	(0.004)	(0.007)	(0.007)	(0.005)	(0.004)
Observations	20,143	20,118	13,273	13,256	18,468	18,440

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. Standard errors in parentheses. See Table B3 for notes on control variables, variables of interest, and clusters.

C – Endogenous deprivation

A large amount of discussion in the UK context has been focusing on how deprivation affects various aspects of social life and in particular social cohesion (Laurence and Heath, 2008; Letki, 2008; Andrews, 2009; Fieldhouse and Cutts, 2010; Twigg et al, 2010; Laurence, 2011; Becares et al, 2011; Sturgis et al, 2011; Demireva and Heath, 2014). One concern can be that deprivation can partly capture the effect that we attribute to diversity, and that our estimates could actually be partly biased. Our approach throughout the paper is to control for two measures of deprivation, the unemployment rate (source: Nomis) and the Index of Multiple Deprivation (source: ONS). Here we try to go a bit further and to consider unemployment as an additional endogenous factor.

As we do for diversity, we pursue two approaches; one relies on instrumental variables and the other attempts at controlling for the location bias. For the latter we mimic what illustrated in Section 4, *Sample Selection* paragraph, and therefore we add as a control a weighted average of the unemployment rate (u) constructed as follows

$$\hat{u}_{1(\cdot n)t} \equiv \sum_{j \neq n} e^{-\alpha d_{nj}} \hat{u}_{jt}$$

where d_{nj} is the distance between the neighbourhoods and $\alpha$is a measure of the cost of distance – that we set equal to 1 for simplicity, although trying with different cost values does not change the results.

The instrumental variable that we use is, as for diversity, based on the idea of Altonji and Card (1991) and Card (2001) though in the context of local demand shocks it is more associated with Bartik (1991). This exploits the fact that, due to historical reasons, areas differ in the industrial mix of local employment. Being ϕ_{sn} the share working age people employed in sector s i in neighbourhood n in some base year, and $(logL_{st} - logL_{st-1})$ the change in log employment in the sector at time t, the change in demand given is then:

$$\Delta C_{nt} = \sum_{s} \phi_{sn}(logL_{st} - logL_{st-1})$$

This instrument is best targeted at capturing changes in the unemployment rate as it represents the change in the demand index. In other words it is well-suited when the model is estimated in first-differences. One could write the level of the of the demand as:

$$C_{nt} = C_{n0} \sum_s \phi_{sn} (logL_{st} - logL_{st-1})$$

for some base-year measure of local demand, C_{n0}. If there are neighbourhood fixed effects then the initial level gets absorbed into that effect. If there are not then one needs to control for variables that measure the initial level of demand and we use the initial industry shares. The following table illustrates the estimated results for all specifications.

Instrumental variable results become much noisier with respect to the specifications with only diversity as endogenous variable, and while diversity remains significantly positive in some specifications, deprivation is almost never significant. It has to be stressed that the empirical models that we are estimating here are quite tight in terms of statistical requirement, in particular when both IV and selection controls are included. It somehow comes with little surprise that first stages appears to be quite weak once that also unemployment is instrumented for.

Table C1: The impact of ethnic mix and unemployment on how you like your neighbourhood.

	No sample selection					Sample selection				
	(1) No FE	(2) Individual FE	(3) Area FE	(4) Individual* Area FE	(5) Individual+ Area FE	(6) No FE	(7) Individual FE	(8) Area FE	(9) Individual* Area FE	(10) Individual+ Area FE
					A. OLS					
White share	0.120***	0.268***	0.090*	0.208***	0.212***	0.169***	0.388***	0.060	0.234***	0.254***
	(0.015)	(0.034)	(0.050)	(0.053)	(0.053)	(0.030)	(0.062)	(0.072)	(0.070)	(0.070)
Unemployment rate	-0.125***	-0.137***	0.020	0.029	-0.003	-0.123***	-0.103**	0.024	0.043	0.014
	(0.039)	(0.041)	(0.047)	(0.037)	(0.037)	(0.040)	(0.042)	(0.047)	(0.038)	(0.038)
N	233,548	200,344	229,637	192,296	198,698	233,548	200,344	229,637	192,296	198,698
					B. IV					
White share	0.098***	0.263***	0.099	0.475***	0.448***	0.107	0.296**	0.142	-0.146	0.025
	(0.021)	(0.049)	(0.119)	(0.139)	(0.138)	(0.067)	(0.126)	(0.314)	(0.389)	(0.390)
Unemployment rate	-0.078	-1.016	-0.193	-1.175	-1.062	0.322	-0.921	-0.107	-1.378	-1.165
	(0.393)	(0.657)	(0.844)	(0.949)	(0.945)	(0.458)	(0.807)	(0.970)	(1.107)	(1.085)
N	233,548	200,344	229,637	192,296	198,698	231,649	198,590	227,761	190,669	196,957
KP	129.578	45.992	5.763	19.951	20.479	55.927	15.980	2.254	7.342	7.685

Notes: * p < 0.1 ** p < 0.05 *** p < 0.001. Standard errors in parentheses. See Notes of Table 5 for further details