Does ethnic diversity affect well-being and allostatic load among people across neighbourhood in England?

Supplementary information

British Household Panel Study (BHPS) and UK Household Longitudinal Study (UKHLS)

The BHPS was conducted by the Institute for Social and Economic Research at the University of Essex. It was designed as an annual survey tracking a nationally representative probability sample of 10,000 individuals nested in 5,000 households over time. Information such as individual and household characteristics, demographics, as well as socio-economic status were collected. It started in 1991, and was terminated in 2009 when the UKHLS was launched, resulting in 18 waves of BHPS data in total. The UKHLS followed the same design features, instruments, and survey questions as the BHPS, but expanded the sample size to include approximately 40,000 households in the UK. Data were collected face-to-face via computer-aided personal interviews. To date, UKHLS has 9 survey waves. From wave 2 of UKHLS, part of the BHPS respondents were added to UKHLS and continued as part of the UKHLS sample. This offers the opportunity of combining two studies into one long panel data series. In both studies, geographic information on households and individuals, such as their lower layer super output area (LSOA) and local authority district (LAD) codes, were collected and made available subject to approval of special license access. Using geographic identifiers, data users can link both studies to UK Censuses in order to investigate how individuals respond to contextual-level circumstances.

Both BHPS and UKHLS provide samples for England, Scotland, Wales and Northern Ireland; however, we only analyse the English sample due to lack of contextual-level information for Scotland, Wales and Northern Ireland. For the same reason, we only use data from BHPS and UKHLS between 2004 and 2011. Our final sample consists of 121,736 individual-wave observations.

In wave 2 (for the general sample) and wave 3 (for the remaining BHPS sample) of UKHLS, 20,070 eligible adult participants received a follow-up health assessment from a trained nurse. A range of bio-medical measures were collected, including, for example, blood pressure, weight, and height. Among the eligible participants, blood samples of 13,107 of them were taken during these visits, from which a

set of biomarkers were extracted. We calculate allostatic load (AL), a measure of individuals' stress level, using data from the UKHLS Nurse Health Assessment. Our main analysis includes 7441 respondents with a complete set of biomarkers and bio-medical measures listed in SI Appendix 2.

UK Census geography

The UK Census is conducted every ten years with the latest one being conducted in 2011. In England and Wales, the Office for National Statistics (ONS) is the responsible body. Output Areas (OA) comprise the lowest geographical level at which census estimates are produced. The minimum OA size is 100 residents to ensure data confidentiality. Super Output Areas (SOA) are larger than OAs and designed to improve the reporting of small area statistics. Two layers of SOAs are available – the lower layer super output areas (LSOA) and middle layer super output areas (MSOA). LSOAs in England have a minimum population of 1000 and maximum of 3000. MSOAs have a minimum of 5000 and maximum of 15 000 individuals. There are 32 844 LSOAs and 6791 MSOAs in England in 2011. Local authority districts are a level of subnational division in England, and local governments operate at this level. Both OAs and SOAs align to LAD boundaries. BHPS and UKHLS provide authorised data users with respondents' LSOA, MSOA and LAD information.

In the literature, LSOA, MSOA and LAD have all been used to define a community or neighbourhood. However, studies have illustrated that estimates are more robust and the effects are stronger when neighbourhood is defined at a smaller geographic level (Dinesen and Sønderskov, 2015). In addition, if ethnic diversity changes individuals' well-being or stress levels, this process is more likely to happen within their local areas because individuals perceive threats around them more strongly. We thus use LSOA as the contextual-level of our analysis.

Calculating ethnic diversity of English LSOAs

1. Calculate the total population of each ethnic group for each LAD and year, using data from Rees and colleagues (Rees et al., 2017).

2. Calculate the growth rate of each ethnic group for each LAD and year, obtaining:

$$Growth_{i,t,j} = (Population_{i,t,j} - Population_{i,t-1,j})/Population_{i,t-1,j} \quad (1)$$

for LAD *i*, year *t* (*t*=2002, …, 2011) and ethnicity *j* (*j*=White, Caribbean, African, Other Black, Indian, Pakistani, Bangladesh, Chinese, Other Asian, and other).

3. Assuming growth of each ethnic population is the same for all LSOAs in same LAD, merge growth rates calculated in Step 2 into the LSOA-level 2001 UK census ethnic population dataset.

4. Calculate ethnic population for each LSOA and year, starting from the LSOA-level ethnic population provided in the 2001 UK census ($Population_{k,2001,j}$) and rolling forward for subsequent years, so that:

$$Population_{k,2002,j} = Population_{k,2001,j} * (1 + Growth_{i,t,j}) \text{ and,} \ldots$$

$$Population_{k,t,j} = Population_{k,t-1,j} * (1 + Growth_{i,t,j}), \quad (2)$$

Where *k* indexes a LSOA in LAD *i*, *t* indexes year (*t*=2003, …, 2011) and the j^{th} ethnicity.

5. Calculate 10-group Herfindahl Index based on ethnic populations calculated in Step 4, using the formula:

$$Herfindahl_{k,t} = 1 - \sum_j S_{k,t,j}^2 \quad (3)$$

where *k* stands for the LSOA area, and *t* for year, and *j* for the j^{th} ethnic group.

6. Compare the Herfindahl index calculated using Steps 1-5 to that calculated using 2011 UK census, and find out only a marginal difference in the indexes at LSOA-level. This validates our assumption made in Step 3.

Alternative measures of ethnic diversity

To test the robustness of our results, we measure ethnic diversity with two additional diversity indices. We calculate Shannon's H Index (Shannon, 1948) and the Evenness Index (Mulder et al., 2004, Pielou, 1966) which measure biodiversity and are widely used in the environmental and ecological literature (Ramos et al., 2019). Similar to the Herfindahl Index, a higher Shannon or Evenness Index indicates higher diversity. The Shannon Index is defined by the following formula:

$$H(Shannon's\ H) = -\sum_{i=1}^{S} p_i * \ln(p_i) \quad (4)$$

where S equals 10 as our index is calculated based on 10 ethnic groups (white, black Caribbean, black African, other black, Indian, Pakistani, Bangladeshi, Chinese, other Asian, and others), p_i is the proportion of the i^{th} ethnic group. Compared to the Herfindahl index, it allocates more weight to groups with smaller proportions. Hence, the Shannon Index may be more consistent with the fact that people from both majority and minority groups often have misperceptions about the group size of minority groups(Alba et al., 2005).

The Evenness Index is defined by the following formula:

$$J' = \frac{H'}{H'_{max}} \quad (5)$$

where H' is Shannon's H calculated using formula (4), and H'_{max} is the maximum possible value of H' which also equals lnS. J' is constrained between 0 and 1, and the less evenness in communities between the groups, the lower J' is.

We re-estimate equation [1] (in the main article) based on alternative ethnic diversity measures, and the results are presented in Table S1. The short-term ethnic diversity measured by deviation from the mean is negative and significantly associated with subjective well-being and health. However, in the long term, ethnic diversity affects neither individuals' subjective well-being and health nor their AL. Once again, ethnic diversity is not a statistically significant predictor of AL, even if it potentially could be in

the short term. Therefore, our main findings are robust to measuring ethnic diversity with other measures such as the Shannon Index and Evenness Index.

Measure of subjective well-being and health

Individual-level well-being and health outcome variables are derived from the BHPS and UKHLS. The subjective well-being and health indicators are respondents' overall life satisfaction, self-perceived health status, score for the General Health Questionnaire (0-36 scale; GHQ36), drawn from the main survey. We combine three subjective measures of health and well-being into one score representing individuals' overall life quality using factor analysis (see SI Appendix 1). The subjective well-being and health score is available in every wave of BHPS/UKHLS. Hence, it enables us to test how individuals' subjective well-being and health change in response to changes in ethnic diversity of local areas over time.

Among the three subjective well-being and health indicators, life satisfaction is measured by the answer to the question, "Satisfaction with life overall". Answers range from 1 (completely dissatisfied) to 7 (completely satisfied). Self-perceived health is derived from answers to the question, "In general, would you say your health is …", and answers range from 1 (poor) to 5 (excellent), representing individuals' physical health status. GHQ36 is derived from answers to 12 questions of the General Health Questionnaire, reflecting respondents' mental health status. We reverse the original scale of GHQ36 to have a higher score indicating a better mental health status: 0 (least distressed) – 36 (most distressed).

Measure of allostatic load

We adopt the Allostatic Load (AL) model proposed by McEwen (McEwen, 1998) to measure objective well-being and health. AL measures the 'wear-and-tear' on one's body when exposed to stressful situations (Juster et al., 2010). It is a widely used measure in the literature when studying humans' physiologic responses to stressors (Juster et al., 2010, Chandola and Zhang, 2017, Prior et al., 2018, Duong et al., 2017). We calculate a 12-item AL score which represents individuals' objective stress levels based on biomarkers available in UKHLS. The calculation of AL is similar to that in previous studies using the same dataset (Chandola and Zhang, 2017, Prior et al., 2018), with cardiovascular,

metabolic, neuroendocrine and immune biomarkers all being included. There are considerable variations in calculating AL (Juster et al., 2010, Duong et al., 2017), thus we also conduct sensitivity analyses to verify our main results, using alternative measures and cut-off points when calculating AL. Main results are reported in Table 1 (in the main article) and additional results using an alternative measure of AL are reported in Table S2.

AL is derived from a smaller representative sample in the UKHLS as part of the Nurse Health Assessment dataset in which adult respondents' blood samples are taken and analysed. Biomarkers are only available in wave 2 (BHPS sample) or wave 3 (UKHLS sample) of UKHLS, accounting for 9000 valid observations. Twelve biomarkers, namely glycosylated haemoglobin (HbA1c), triglycerides, C - reactive protein (CRP), clauss fibrinogen, dehydroepiandrosterone sulphate (DHEA-S), insulin-like growth factor 1 (IGF-1), total cholesterol, high-density lipoprotein cholesterol (HDL), pulse, systolic blood pressure (SBP), diastolic blood pressure (DBP) and body mass index (BMI) are used to calculate AL. Around 3000 individuals with one or more missing biomarkers are excluded from the analysis. We use clinical cut-off points advised in the UKHLS Biomarker User Guide and Glossary (Benzeval et al., 2014). For each biomarker, individuals with valid values above the clinical cut-off point of the corresponding biomarker are coded as "1", and those within the normal ranges are coded "0". The only exceptions are HDL for which values below the cut-off point are coded as "1", and DHEA-S and IGF for which values below the lower cut-off point or above the higher cut-off point are coded as "1". The AL is a count-based summary of all 12 biomarkers. A higher AL indicates higher health risk and vice versa. Table SA3 summarises the biomarkers used in the present study, and their cut-off points and functions. Although it is evident that high AL is linked to incident cardiovascular disease (CVD), and decline in physical functioning decline in cognitive functioning and mortality (Seeman et al., 1997), in the literature, there is no consensus on which biomarkers to include in AL and which cut-off values to use when grouping observations into high/low risk groups. In a review of both theoretical and empirical work using the AL framework (Juster et al., 2010), the 58 reviewed studies use 4 - 17 biomarkers and adopt various cut-off points and algorithmic formulations to calculate AL. To avoid potential biases resulting from different algorithms being used, we conduct a robustness check by testing the effect of

ethnic diversity on an additional measure of AL that is based on different biomarkers and cut-off points. Specifically, we follow strictly Prior and colleagues (Prior et al., 2018) who selected 13 biomarkers from the Understanding Society survey, representing five aspects: cardiovascular (SBP, DBP, pulse), inflammatory (CRP, fibrinogen, albumin), lipid and glucose metabolism systems (HDL, total: DHL ratio, triglycerides, BMI, waist circumference, HbA1c), and the hypothalamic-pituitary axis (DHEAs). The high-risk cut-off values are given in their study. A system risk score of AL is then created by summing the proportions of biomarkers within each of the subsystems that fall into the high-risk ranges. Again, higher scores represent worse outcomes. This algorithm adopts the maximum bias approach where respondents with one or more biomarkers missing are treated as 'not at risk' unless all biomarkers are missing (Solís et al., 2015).

Measure of other covariates

We control for a comprehensive set of individuals' demographic and other characteristics that are associated with health and well-being, including age, gender, marital status, highest education level, ethnicity, socio-economic group, household total income, homeownership, the number of children and the number of household members. Age is included as a continuous variable and age-squared is added to the model to control for a possible non-linear effect of age. Gender is a binary variable which equals 1 for male and 0 for female. Marital status is included as a set of binary variables (married, divorced/separated/widowed) which equal 1 if the respondent belongs to the corresponding category. Ethnicity is included as a set of binary variables (Indian, Pakistani, Bangladeshi, Black, others) which equals 1 if the respondent belongs to corresponding ethnic group and 0 otherwise. Highest education level is included as a set of binary variables (GCSE, A-level, degree, other degree, other qualification) which equals 1 if the respondent belongs to the corresponding education category and 0 otherwise. Whether or not the respondent is currently working and whether or not he/she was born in the UK, are included (both 0/1 dummy variables). Household total income is included as the natural logarithm. Homeownership is measured by a binary variable equal to 1 if the respondent's current accommodation is rented. A detailed descriptive summary is presented in Appendix 3 Table SA4 based on our full sample.

In the AL equation, as well as all covariates noted above, additional covariates are added to the model to control for factors affecting some of the biomarkers. We control for the nurse blood draw's length of time, the types of containers used for holding the blood sample, how long the blood sample was stored in the lab, whether the participant took anti-inflammatory drugs in the past 7 days, and whether statins had been taken in the past 7 days. Anti-inflammatory drugs affect the levels of CRP and fibrinogen, while statins affect the level of low-density lipoprotein cholesterol.

Alternative allostatic load measure

We calculate an alternative measure of AL based on different biomarker items, cut-off points and algorithm as a robustness check to ensure that our findings are not the result of the operationalization of our AL measure. For this purpose, we follow strictly Prior and colleagues (Prior et al., 2018). Thirteen biomarkers are selected from the Understanding Society project, representing five aspects: cardiovascular (SBP, DBP, pulse), inflammatory (CRP, fibrinogen, albumin), lipid and glucose metabolism systems (HDL, total: DHL ratio, triglycerides, BMI, waist circumference, HbA1c), and the hypothalamic-pituitary axis (DHEAs). The high-risk cut-off values are given in Table 1 in their study (11, p. 27) . A system risk score of AL is then created by summing the proportions of biomarkers within each of the subsystems that fall into the high-risk ranges. Again, higher scores represent worse outcomes. This algorithm adopts the maximum bias approach where respondents with one or more biomarkers missing are treated as 'not at risk' unless all biomarkers are missing. We estimate equation [1] using the alternative measure of AL as a robustness check, controlling for a full set of covariates. Results are reported in Table S2 below. Neither short- nor long-term ethnic diversity are associated with AL.

References

Alba, R., Rumbaut, R. G. & Marotz, K. 2005. A distorted nation: Perceptions of racial/ethnic group sizes and attitudes toward immigrants and other minorities. *Social Forces,* 84**,** 901-919.

Benzeval, M., Davillas, A., Kumari, M. & Lynn, P. 2014. Understanding society: the UK household longitudinal study biomarker user guide and glossary. *Institute for Social and Economic Research, University of Essex*.

Chandola, T. & Zhang, N. 2017. Re-employment, job quality, health and allostatic load biomarkers: prospective evidence from the UK Household Longitudinal Study. *International Journal of Epidemiology,* 47**,** 47-57.

Dinesen, P. T. & Sønderskov, K. M. 2015. Ethnic diversity and social trust: Evidence from the micro-context. *American Sociological Review,* 80**,** 550-573.

Duong, M. T., Bingham, B. A., Aldana, P. C., Chung, S. T. & Sumner, A. E. 2017. Variation in the Calculation of Allostatic Load Score: 21 Examples from NHANES. *Journal of racial and ethnic health disparities,* 4**,** 455-461.

Juster, R.-P., Mcewen, B. S. & Lupien, S. J. 2010. Allostatic load biomarkers of chronic stress and impact on health and cognition. *Neuroscience & Biobehavioral Reviews,* 35**,** 2-16.

Kaiser, H. F. 1974. An index of factorial simplicity. *Psychometrika,* 39**,** 31-36.

Mcewen, B. S. 1998. Stress, adaptation, and disease: Allostasis and allostatic load. *Annals of the New York Academy of Sciences,* 840**,** 33-44.

Mulder, C., Bazeley‐White, E., Dimitrakopoulos, P., Hector, A., Scherer‐Lorenzen, M. & Schmid, B. 2004. Species evenness and productivity in experimental plant communities. *Oikos,* 107**,** 50-63.

Pielou, E. C. 1966. The measurement of diversity in different types of biological collections. *Journal of Theoretical Biology,* 13**,** 131-144.

Prior, L., Manley, D. & Jones, K. 2018. Stressed out? An investigation of whether allostatic load mediates associations between neighbourhood deprivation and health. *Health & Place,* 52**,** 25-33.

Ramos, M. R., Bennett, M. R., Massey, D. S. & Hewstone, M. 2019. Humans adapt to social diversity over time. *Proceedings of the National Academy of Sciences,* 116**,** 12244-12249.

Rees, P., Clark, S., Wohland, P., Lomax, N. & Norman, P. 2017. Using 2001 and 2011 censuses to reconcile ethnic group estimates and components for the intervening decade for English local authority districts. *In:* STILLWELL, J. & DUKE-WILLIAMS, O. (eds.) *The Routledge Handbook of Census Resources, Methods and Applications: Unlocking the UK 2011 Census* Routledge.

Seeman, T. E., Singer, B. H., Rowe, J. W., Horwitz, R. I. & Mcewen, B. S. 1997. Price of adaptation—allostatic load and its health consequences: MacArthur studies of successful aging. *Archives of internal medicine,* 157**,** 2259-2268.

Shannon, C. E. 1948. A mathematical theory of communication. *Bell System Technical Journal,* 27**,** 379-423.

Solís, C. B., Kelly-Irving, M., Fantin, R., Darnaudéry, M., Torrisani, J., Lang, T. & Delpierre, C. 2015. Adverse childhood experiences and physiological wear-and-tear in midlife: Findings from the 1958 British birth cohort. *Proceedings of the National Academy of Sciences,* 112**,** E738-E746.

Table S1 Main analysis with alternative ethnic diversity measures

	(1) WB & H	(2) WB & H	(3) AL	(4) AL
LSOA-level Variables				
Shannon Index (mean)	0.009		-0.153	
	[-0.01,0.03]		[-0.32,0.02]	
Shannon Index (change)	-0.173**		0.382	
	[-0.30,-0.05]		[-0.77,1.54]	
Evenness Index (mean)		0.020		-0.353
		[-0.02,0.06]		[-0.75,0.04]
Evenness Index (change)		-0.399**		0.880
		[-0.69,-0.11]		[-1.78,3.54]
N	121,736	121,736	7,442	7,442

Note: Unstandardized coefficients are shown, and 95% confidence intervals are given in parentheses. WB&H indicates individual subjective well-being and health, and is based on factor analysis of 3 subjective well-being and health indicators: overall life satisfaction, self-perceived physical health status and GHQ36. AL stands for allostatic load, a measure of chronic stress. Columns (1) and (3) estimate equation [1] using Shannon index as the measure of ethnic diversity, and columns (2) and (4) using the Evenness index as the measure of ethnic diversity. The mean of diversity is the mean of ethnic diversity between 2001 and 2011. The change of diversity is the deviation of diversity from its mean. A full set of individual-level control variables and the mean and deviation of IMD are included in the analysis. Individual- and LSOA-level random intercepts are included in WB&H estimations, and individual-level random intercepts are included in AL estimations.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table S2 Robustness check of main results using an alternative measure of AL

	(1)
	AL
LSOA-level Variables	
Ethnic Diversity (mean)	-0.136
	[-0.28,0.01]
Ethnic Diversity (change)	0.216
	[-0.69,1.12]
IMD (mean)	0.004***
	[0.00,0.01]
IMD (change)	-0.000
	[-0.01,0.01]
Individual-level Variables	
Age	0.021***
	[0.01,0.03]
Age^2	-0.000
	[-0.00,0.00]
Male	-0.071***
	[-0.10,-0.04]
Employed	-0.063**
	[-0.11,-0.02]
Married	0.010
	[-0.04,0.06]
Divorce/Separated/Widowed	0.060*
	[0.00,0.12]
Indian	0.193**
	[0.06,0.33]
Pakistani	0.148

	[-0.06,0.36]
Bangladeshi	-0.062
	[-0.43,0.30]
Black	0.123
	[-0.03,0.28]
Other	0.062
	[-0.06,0.18]
Degree	-0.220***
	[-0.28,-0.16]
Other Degree	-0.136***
	[-0.20,-0.07]
A-level	-0.111***
	[-0.17,-0.05]
GCSE	-0.107***
	[-0.16,-0.05]
Other Qualification	-0.101**
	[-0.16,-0.04]
Total Income	-0.011
	[-0.02,0.00]
Home Rented	0.134***
	[0.09,0.18]
Household Size	0.009
	[-0.01,0.03]
N of Children	-0.042**
	[-0.07,-0.02]
UK Born	0.080
	[-0.13,0.29]
Wave	0.050

	0.031
Venipuncture start time	[-0.01,0.07]
	0.012***
Blood collection system	[0.01,0.02]
	-0.033
Days of blood sample taken to lab	[-0.07,0.00]
	0.001
Took inflammatory med during past 7 days	[-0.00,0.00]
	0.000
Took statins during last 7 days	[0.00,0.00]
	0.271***
Constant	-0.755
	[-1.61,0.10]
Akaike Information Criterion (AIC)	15167.8
Bayesian Information Criterion (BIC)	15402.9
N of Observations	7,441
N of Respondents	7,441
N of LSOAs	4,716

Notes: Unstandardized coefficients are reported. 95% confidence intervals are given in parentheses. All individual-level covariates and LSOA-level IMD (mean and change) are included in the estimations. Column (1) presents the results for AL using the alternative AL measure.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table S3 Robustness check of main results separating subjective well-being and health components

	(1) Subjective well-being (Life satisfaction and GHQ36)	(2) Subjective health (Self-perceived health status)
LSOA-level Variables		
Ethnic Diversity (mean)	-0.008	0.113***
	[-0.05,0.03]	[0.06,0.17]
Ethnic Diversity (change)	-0.324**	-0.234
	[-0.57,-0.08]	[-0.58,0.11]
IMD (mean)	-0.003***	-0.007***
	[-0.00,-0.00]	[-0.01,-0.01]
IMD (change)	-0.004**	-0.008***
	[-0.01,-0.00]	[-0.01,-0.00]
Individual-level Variables		
Age	-0.032***	-0.026***
	[-0.03,-0.03]	[-0.03,-0.02]
Age^2	0.000***	0.000***
	[0.00,0.00]	[0.00,0.00]
Male	0.059***	0.031***
	[0.05,0.07]	[0.02,0.05]
Employed	0.147***	0.259***
	[0.13,0.16]	[0.24,0.28]
Married	0.103***	0.078***
	[0.09,0.12]	[0.05,0.10]
Divorce/Separated/Widowed	-0.040***	-0.020
	[-0.06,-0.02]	[-0.05,0.01]
Indian	-0.041*	-0.109***

	[-0.08,-0.01]	[-0.16,-0.06]
Pakistani	-0.107***	-0.186***
	[-0.15,-0.07]	[-0.24,-0.13]
Bangladeshi	-0.072**	-0.078*
	[-0.12,-0.02]	[-0.14,-0.01]
Black	0.044**	0.127***
	[0.01,0.07]	[0.08,0.17]
Other	-0.054***	-0.050*
	[-0.08,-0.03]	[-0.09,-0.01]
Degree	0.100***	0.493***
	[0.08,0.12]	[0.46,0.52]
Other Degree	0.053***	0.334***
	[0.03,0.08]	[0.30,0.37]
A-level	0.038***	0.263***
	[0.02,0.06]	[0.24,0.29]
GCSE	0.041***	0.222***
	[0.02,0.06]	[0.20,0.25]
Other Qualification	0.026*	0.124***
	[0.00,0.05]	[0.09,0.15]
Total Income	-0.008***	-0.023***
	[-0.01,-0.00]	[-0.03,-0.02]
Home Rented	-0.127***	-0.199***
	[-0.14,-0.11]	[-0.22,-0.18]
Household Size	0.001	-0.013***
	[-0.00,0.01]	[-0.02,-0.01]
N of Children	0.005	0.036***
	[-0.00,0.01]	[0.03,0.05]
UK Born	-0.058***	-0.111***

	[-0.08,-0.04]	[-0.13,-0.09]
Wave	-0.005***	0.047***
	[-0.01,-0.00]	[0.04,0.05]
Constant	0.728***	3.353***
	[0.67,0.79]	[3.27,3.44]
Akaike Information Criterion (AIC)	229348.5	378802.4
Bayesian Information Criterion (BIC)	229639.8	379098.2
N of Observations	121,776	141,448
N of Respondents	52,418	57,417
N of LSOAs	15,545	16,024

Notes: Unstandardized coefficients are shown, and 95% confidence intervals are given in parentheses. Life satisfaction and GHQ36 score (scale reversed) are combined into a subjective well-being indicator by factor analysis. All scales are adjusted so that a higher value indicates better well-being/health. Individual- and LSOA-level random intercepts are included in all estimations.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table S4 Testing the possible causal relationship between short-term change in ethnic diversity and well-being and health among white and non-white stayers

	(1)	(2)
Outcome: well-being and health	White stayers	Non-White stayers
Age	-0.009	-0.065**
	[-0.02,0.00]	[-0.10,-0.03]
Age²	0.000	0.001**
	[-0.00,0.00]	[0.00,0.00]
Employed	0.058***	0.094**
	[0.03,0.08]	[0.03,0.16]
Married	-0.074	0.019
	[-0.17,0.02]	[-0.30,0.33]
Divorce/Separated/Widowed	-0.085	0.223
	[-0.21,0.04]	[-0.19,0.64]
Degree	-0.181***	-0.059
	[-0.29,-0.08]	[-0.32,0.21]
Other Degree	-0.109	-0.238
	[-0.23,0.01]	[-0.51,0.03]
A-level	-0.125**	-0.083
	[-0.21,-0.04]	[-0.30,0.13]
GCSE	-0.060	-0.047
	[-0.14,0.02]	[-0.24,0.15]
Other Qualification	-0.104*	-0.026
	[-0.19,-0.02]	[-0.31,0.26]
Total Income	-0.002	-0.003
	[-0.01,0.00]	[-0.01,0.01]
Home Rented	0.023	-0.029
	[-0.06,0.11]	[-0.21,0.15]
Household Size	-0.004	0.003

	[-0.02,0.01]	[-0.03,0.03]
N of Children	0.009	0.010
	[-0.01,0.03]	[-0.03,0.05]
Ethnic Diversity (Change)	-0.630*	-0.612
	[-1.17,-0.09]	[-2.26,1.04]
IMD (Change)	-0.003	-0.006
	[-0.01,0.00]	[-0.02,0.01]
_cons	0.466**	1.574**
	[0.14,0.79]	[0.60,2.55]
N	84 939	14 029

Notes: Fixed-effects models with robust standard errors are estimated. Unstandardized coefficients are shown, and 95% confidence intervals are given in parentheses. Well-being and health is based on factor analysis of 3 subjective well-being and health indicators: overall life satisfaction, self-perceived physical health status and GHQ36. A full set of time invariant control variables are included in all estimations. Stayers are those who did not move home across all their available observational waves. Non-stayers are the rest of the sample who are not identified as stayers.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table S5. Local Authority District (LAD) level analysis

	(1) WB & H	(2) AL
LAD-level variable		
Ethnic Diversity (mean)	0.017	-0.646**
	[-0.04,0.08]	[-1.07,-0.22]
Ethnic Diversity (change)	-0.367*	-0.140
	[-0.67,-0.06]	[-2.99,2.71]
IMD (mean)	-0.003***	0.018***
	[-0.00,-0.00]	[0.01,0.02]
IMD (change)	-0.002	0.055*
	[-0.01,0.00]	[0.01,0.10]
Individual-level variable		
Age	-0.032***	0.108***
	[-0.03,-0.03]	[0.09,0.12]
Age^2	0.000***	-0.001***
	[0.00,0.00]	[-0.00,-0.00]
Male	0.062***	0.400***
	[0.05,0.07]	[0.32,0.48]
Employed	0.183***	-0.178**
	[0.17,0.19]	[-0.29,-0.07]
Married	0.116***	-0.058
	[0.10,0.13]	[-0.18,0.07]
Divorce/Separated/Widowed	-0.038***	0.143*
	[-0.06,-0.02]	[0.00,0.28]
Indian	-0.054**	0.279
	[-0.09,-0.02]	[-0.05,0.61]
Pakistani	-0.147***	0.228
	[-0.19,-0.11]	[-0.27,0.73]
Bangladeshi	-0.104***	-0.562
	[-0.16,-0.05]	[-1.43,0.31]
Black	0.036*	-0.176
	[0.00,0.07]	[-0.55,0.19]
Other	-0.072***	0.198
	[-0.10,-0.04]	[-0.09,0.49]
Degree	0.204***	-0.561***
	[0.18,0.23]	[-0.71,-0.42]
Other Degree	0.130***	-0.316***
	[0.11,0.15]	[-0.47,-0.16]
A-level	0.103***	-0.334***
	[0.08,0.12]	[-0.48,-0.19]
GCSE	0.093***	-0.299***
	[0.07,0.11]	[-0.44,-0.16]
Other Qualification	0.048***	-0.297***
	[0.03,0.07]	[-0.45,-0.15]
Total Income	-0.011***	-0.021
	[-0.01,-0.01]	[-0.05,0.01]
Home Rented	-0.154***	0.360***
	[-0.17,-0.14]	[0.26,0.46]
Household Size	-0.003	0.062*
	[-0.01,0.00]	[0.01,0.11]
N of Children	0.010**	-0.089**
	[0.00,0.02]	[-0.15,-0.02]
UK Born	-0.068***	0.063
	[-0.09,-0.05]	[-0.44,0.57]

wave	0.005***	0.035
	[0.00,0.01]	[-0.06,0.13]
Venipuncture start time		0.037***
		[0.03,0.05]
Blood collection system		-0.020
		[-0.10,0.06]
Days of blood sample taken to lab		0.002
		[-0.00,0.01]
Took inflammatory med during past 7 days		0.642***
		[0.49,0.79]
Took statins during last 7 days		-0.213***
		[-0.32,-0.10]
Constant	0.594***	-1.539
	[0.52,0.67]	[-3.61,0.53]
N of Observations	122,717	7,487
Akaike Information Criterion (AIC)	236333.364	28612.470
Bayesian Information Criterion (BIC)	236624.893	28847.782

Notes: Unstandardized coefficients are shown, and 95% confidence intervals are given in parentheses. WB&H stands for subjective well-being and health. Well-being and health is an index combining overall life satisfaction, self-perceived physical health status and GHQ36. AL is the indicator of objective well-being and health, standing for allostatic load. A higher well-being and health score and a lower AL score indicate better health. Ethnic diversity is measured by a 10-group Herfindahl Index. The mean of diversity is the mean of ethnic diversity between 2001 and 2011. The change of diversity is the deviation of diversity from its mean. Total income is in logarithm. IMD stands for Index of Multiple Deprivation. Individual- and LAD-level random intercepts are included in subjective health and well-being estimations, and individual-level random intercepts are included in allostatic load estimations.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table S6 Taking into account respondents' length of residence

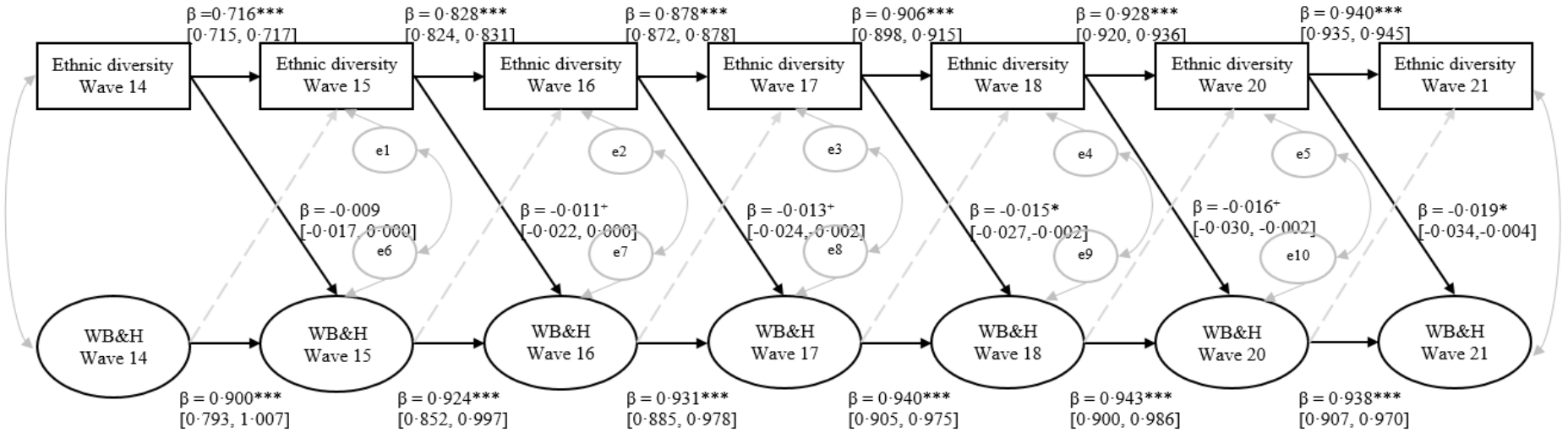
	(2) WB&H	(4) AL
LSOA-level variable		
Ethnic Diversity (mean)	0.008	-0.349
	[-0.04,0.06]	[-0.76,0.07]
Ethnic Diversity (change)	-0.442**	0.923
	[-0.76,-0.12]	[-1.73,3.58]
IMD (mean)	-0.004***	0.010***
	[-0.00,-0.00]	[0.01,0.01]
IMD (change)	-0.008***	-0.016
	[-0.01,-0.00]	[-0.04,0.01]
Individual-level variable		
Length of Residence (Year)	0.001**	0.002
	[0.00,0.00]	[-0.00,0.01]
Age	-0.035***	0.101***
	[-0.04,-0.03]	[0.08,0.12]
Age^2	0.000***	-0.001***
	[0.00,0.00]	[-0.00,-0.00]
Male	0.049***	0.421***
	[0.04,0.06]	[0.33,0.51]
Employed	0.188***	-0.184**
	[0.17,0.20]	[-0.31,-0.06]
Married	0.116***	0.018
	[0.09,0.14]	[-0.13,0.17]
Divorce/Separated/Widowed	-0.038**	0.197*
	[-0.06,-0.01]	[0.03,0.36]
Indian	-0.064**	0.201
	[-0.11,-0.02]	[-0.19,0.60]
Pakistani	-0.145***	0.025
	[-0.20,-0.09]	[-0.53,0.58]
Bangladeshi	-0.035	-0.908
	[-0.10,0.03]	[-1.94,0.13]
Black	0.055**	-0.324
	[0.02,0.09]	[-0.73,0.08]
Other	-0.043*	0.148
	[-0.08,-0.01]	[-0.18,0.48]
Degree	0.167***	-0.463***
	[0.14,0.19]	[-0.63,-0.30]
Other Degree	0.109***	-0.294***
	[0.08,0.14]	[-0.46,-0.12]
A-level	0.083***	-0.259**
	[0.06,0.11]	[-0.42,-0.09]
GCSE	0.073***	-0.239**
	[0.05,0.10]	[-0.40,-0.08]
Other Qualification	0.032*	-0.259**
	[0.01,0.06]	[-0.43,-0.09]
Total Income	-0.012***	-0.024
	[-0.02,-0.01]	[-0.06,0.01]
Home Rented	-0.157***	0.347***
	[-0.17,-0.14]	[0.22,0.47]
Household Size	-0.005	0.070*
	[-0.01,0.00]	[0.01,0.13]
N of Children	0.013**	-0.116**
	[0.00,0.02]	[-0.19,-0.04]
UK Born	-0.071***	-0.024
	[-0.09,-0.05]	[-0.55,0.50]

wave	-0.000	0.041
	[-0.00,0.00]	[-0.40,0.48]
Venipuncture start time		0.042***
		[0.03,0.05]
Blood collection system		0.036
		[-0.06,0.13]
Days of blood sample taken to lab		0.005
		[-0.00,0.01]
Took inflammatory med during past 7 days		0.634***
		[0.46,0.80]
Took statins during last 7 days		-0.286***
		[-0.41,-0.16]
Constant	0.829***	-1.529
	[0.74,0.92]	[-10.41,7.36]
N of Observations	87,101	5,592
Akaike Information Criterion (AIC)	166364.669	21309.917
Bayesian Information Criterion (BIC)	166655.289	21541.935

Notes: Unstandardized coefficients are shown, and 95% confidence intervals are given in parentheses. Samples used here only include stayers – those who never moved home across their observational waves. WB&H stands for subjective well-being and health. Well-being and health is an index combining overall life satisfaction, self-perceived physical health status and GHQ36. AL is the indicator of objective well-being and health, standing for allostatic load. A higher well-being and health score and a lower AL score indicate better health. Ethnic diversity is measured by a 10-group Herfindahl Index. The mean of diversity is the mean of ethnic diversity between 2001 and 2011. The change of diversity is the deviation of diversity from its mean. Total income is in logarithm. IMD stands for Index of Multiple Deprivation. Individual- and LSOA-level random intercepts are included in subjective health and well-being estimations, and individual-level random intercepts are included in allostatic load estimations.

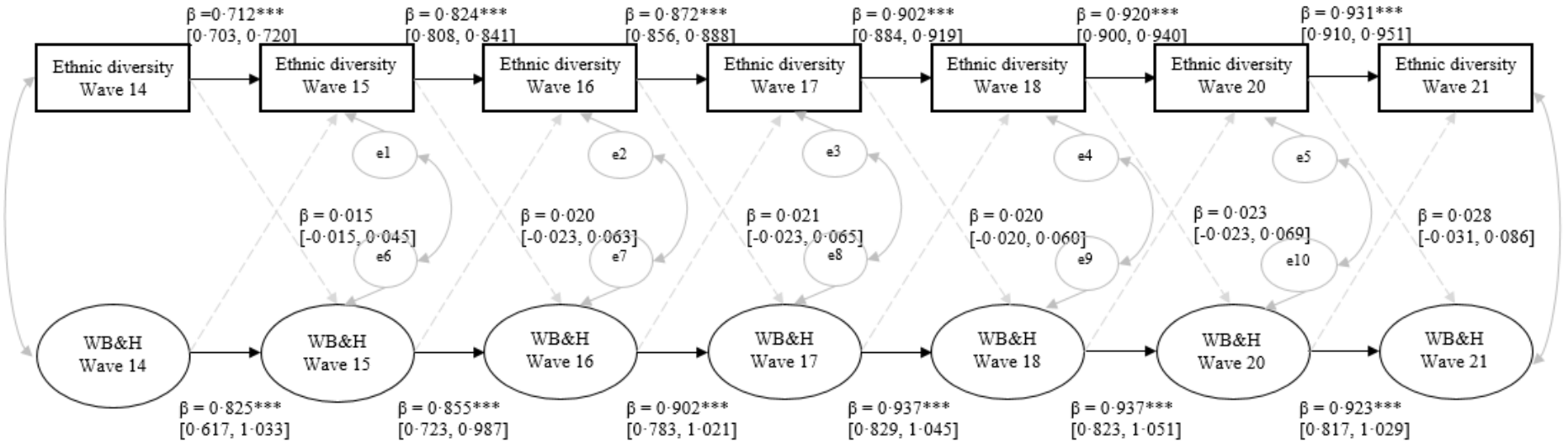
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure S1 Cross-lagged analysis among the White subsample



Notes: Dashed arrows indicate non-significance $^{+}$ p <0.10 * p < 0.05, ** p < 0.01, *** p < 0.001

Figure S2 Cross-lagged analysis among the Non-White subsample



Notes: Dashed arrows indicate non-significance $^{+}$ p <0.10 * p < 0.05, ** p < 0.01, *** p < 0.001

SI Appendix 1. Factor analysis of individuals' subjective well-being and health indicators

Table SA1. Factor analysis results

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor 1	1.03209	1.11527	1.4325	1.4325
Factor 2	-0.08318	0.14525	-0.1155	1.3171
Factor 3	-0.22844		-0.3171	1

Notes: Eigenvalue suggests the number of factors to be used in factor solution. Only factors with eigenvalue larger than 1 can be used in factor solution. Proportion reports the contribution of each factor calculated by its eigenvalue divided by the sum of eigenvalues. Cumulative is the cumulative sum of proportion.

Table SA2. Factor loadings and KMO test results

Variable	Factor 1	Uniqueness	KMO Score
	(1)	(2)	(3)
Overall life satisfaction	0.6222	0.6129	0.6019
Self-perceived health status	0.4589	0.7894	0.7237
GHQ36 (Scale reversed)	0.6590	0.5657	0.5876
Overall			0.6186

Notes: Loadings of Factor 1 presented in column (1) and uniqueness of each dimension in column (2). Column (3) presents the results of the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO test)(Kaiser, 1974). It indicates whether the variable can be factorized efficiently. The KMO scores range from 0 to 1. A higher KMO score denotes a more efficient factor analysis has been performed on this variable. However, only value greater than 0.6 could be accepted and suggests the efficiency of the performed factor analysis.

Appendix 2 The summary of biomarkers used in calculating allostatic load

Table SA 3 Summary of biomarkers used in calculating AL

Type	Biomarker	Application	Cut point
Neuroendocrine	DHEAs	Associated with cardiovascular health	Diff. between gender and across age groups[§]
Immune	CRP	Indicator of inflammatory load	>=3mg/L
	Fibrinogen	Indicator of inflammatory load	>=75th percentile of distribution
	IGF-1	Associated with heart disease and some cancers	Diff. between gender and across age groups[†]
Metabolic	HDL	Protective against CVD	<=1mmol/L
	Total cholesterol	Risk factor of CVD	>=5mmol/L
	Triglycerides	Risk factor of CVD	>=2mmol/L
	HbA1c	Indictor of diabetes risk	>= 48 mmol/mol
Cardiovascular	Pulse	Heart rate	>=75th percentile of distribution
	SBP	Indicator of hypertension and associated with CVD	>=140mmHg
	DBP	Indicator of hypertension and associated with CVD	>=90mmHg
Anthropometric	BMI	Indicator of overweight and obesity	>=25kg/m^2

Sources: *Benzeval et al. (2014)*; *Juster et al. (2010)*

Notes: Cut-off points are drawn from *Benzeval et al. (2014)*; when the cut-off points are not available, 75th percentile is calculated and observations falling in the highest quartiles are coded "1", and otherwise "0";

[§]Please refer to Table 19 Expected ranges of DHEAs in men and women by age group, *Benzeval et al. (2014)*;

[†]Please refer to Table 17 IGF-1 reference values in men and women by age, *Benzeval et al. (2014)*.

Appendix 3 Full sample descriptive statistics

Table SA4 Descriptive statistics of full sample

	Full sample (N=121,736)
Life satisfaction	
Completely dissatisfied	2,605 (2.14%)
Mostly dissatisfied	5,276 (4.33%)
Somewhat dissatisfied	9,060 (7.44%)
Neither satisfied or dissatisfied	12,925 (10.62%)
Somewhat satisfied	25,753 (21.15%)
Mostly satisfied	51,547 (42.34%)
Completely satisfied	14,570 (11.97%)
Self-perceived physical health	
Excellent	22,453(18.44%)
Very good	30,863 (25.35%)
Good	41,192 (33.84%)
Fair	19,491 (16.01%)
Poor	7,737 (6.36%)
GHQ36 (0-best; 36-worst)	11.11 (5.42)
Sex	
Male	54,306 (44.61%)
Female	67,430 (55.39%)
Marital status	
Married	63,558 (52.21%)

Divorce/Separated/Widowed	20,641 (16.96%)
Other	37,531 (30.83%)
Ethnicity	
White	105,713 (86.84%)
Indian	3,527 (2.90%)
Pakistani	2,490 (2.05%)
Bangladeshi	1,426 (1.17%)
Black	4,267 (3.51%)
Other	4,313 (3.54%)
Educational level	
Degree	25,747 (21.15%)
Other Degree	13,011 (10.69%)
A-level	25,522 (20.97%)
GCSE	27,775 (22.82%)
Other Qualification	12,826 (10.54%)
Other	16,836 (13.83%)
Other characteristics	
Age	46.33 (18.04)
Total Income (logarithm)	6.68 (1.79)
Employed	72,150 (59.27%)
Home Rented	33,671 (27.66%)
Household Size	2.87 (1.44)
N of Children	0.59 (0.98)
UK Born	115 843 (95.16%)

Note: Data are *n* (%) or mean (SD).