Designing an Immigrant Social Integration Policy

Yujung Hwang *

September 27, 2024

Abstract

I estimate a dynamic general equilibrium model of neighborhood choice and immigrants' cultural assimilation using an 18-year-long micro panel dataset that includes neighborhood choices and proxy variables for immigrants' cultural traits. Both immigrants and natives choose their neighborhoods, considering cultural similarities with neighbors. Immigrants also make decisions about their level of cultural assimilation, with the associated costs varying based on their exposure to native residents. Using the model estimates, I show how a hypothetical rent subsidy, granted conditional on living in a diverse neighborhood, can induce neighborhood desegregation and immigrants' cultural assimilation under a budget-balancing tax.

JEL Classification: J15, J61, R23, Z12, Z13, Z18

Keywords: Social Engineering, Social Integration, Neighborhood Choice, Residential Sorting, Assimilation

*yujungghwang@gmail.com. This paper is a revised version of my Ph.D dissertation chapter titled "An Estimable General-Equilibrium Structural Model of Immigrants' Neighborhood Sorting and Social Integration". I thank Yale University and Johns Hopkins University for their generous grant support for this project. I am also grateful to University College London and the Institute for Fiscal Studies for hosting me to use their computer facilities. Finally, I thank my Ph.D advisors, Joseph Altonji, Mark Rosenzweig, and Aureo de Paula, for their guidance and seminar participants for their comments. All errors are my own.

1 Introduction

Integrating minorities is central to achieving functioning political institutions, social harmony, and fast economic growth as shown in the previous literature (Easterly and Levine (1997), Alesina et al. (1999), Alesina and La Ferrara (2000), Alesina et al. (2003), Alesina and La Ferrara (2005), Miguel and Gugerty (2005), Dahlberg et al. (2012), Algan et al. (2016), Ray and Esteban (2017)). Yet, many immigrants live in segregated neighborhoods with limited contact with natives, which has triggered concern about immigrant social integration. For example, in Great Britain, nearly half of Muslim immigrants were living in Muslim neighborhoods in which at least one out of every five residents was Muslim, according to Census 2011. Little work has been done about how the government may integrate culturally different immigrant populations, which is the subject of this paper.

This paper quantitatively assesses a residential mixing policy as a social integration tool. First, I build a *non-stationary* dynamic model of neighborhood choice and immigrants' cultural assimilation that links how the residential choice is intertwined with immigrants' assimilation efforts. Specifically, moving to culturally dissimilar neighborhoods is costly, but exposure to natives lowers the assimilation cost for immigrants. The model considers various factors that affect neighborhood choices, including rents in the neighborhoods, skill prices in the local labor markets[1], exogenous neighborhood amenities, and proximity to the current neighborhood. Conditional on the current neighborhood location, immigrants choose the level of cultural assimilation for the next period, the high level of which is financially beneficial because it encourages female spouses to work. This is a general equilibrium model that explains skill prices in local labor markets, rents for neighborhoods, and the social profile of neighborhood residents in equilibrium. So the model estimates can credibly predict the outcomes of large-scale intervention in the housing market, taking into account changing equilibrium conditions in labor and housing markets and natives' outmigration, so-called "White-flight"(Schelling (1971)).

This paper provides a unified framework combining the separate literature on neighborhood sorting (Bayer et al. (2016), Davis et al. (2021), Davis et al. (2023)) and cultural assimilation (Fouka et al. (2022),Bisin et al. (2004), Bisin and Tura (2019), Adda et al. (2020)). Previous papers on cultural assimilation focused on other mechanisms such as inter-religious or inter-ethnic marriages (Bisin et al. (2004), Bisin and Tura (2019), Adda et al. (2020)). In contrast, I emphasize endogenous neighborhood sorting as a key mechanism for explaining the immigrants' cultural assimilation, consistent with the empirical evidence in Fouka et al. (2022). My paper is distinguished from other structural papers on immigration (Albert et al. (2021), Adda et al. (2022)), which focus on economic assimilation instead of cultural assimilation.

[1]Local labor market is a larger geographic unit than neighborhoods. Each neighborhood belongs to at most one local labor market. I assumed there are five local labor markets in Great Britain.

This paper also extends the literature on spatial general equilibrium model with externality (Aliprantis and Carroll (2018), Fogli and Guerrieri (2019), Piyapromdee (2021), Eckert and Kleineberg (2021), Anstreicher (2023)) in multiple dimensions to preserve realistic features in the environment, thereby allowing credible counterfactual policy analysis. The former papers assumed a stationary environment, whereas my paper assumed a non-stationary environment (e.g., due to population growth from immigration/emigration, generation transition, and Total Factor Productivity (TFP) change in each labor market), so it accounts for changes in the environment over the sample period (18 years) used in estimation and generates predictions for the transition periods in the near future[2]. In addition, I define much smaller geographic areas (Lower Super Output Area in Great Britain, or LSOA) as neighborhoods so the model can explain segregation at the granular level, LSOAs, which are much smaller than US Census tracts on average[3]. In contrast, previous papers considered migration between two types of neighborhoods, US states, MSAs, and counties. Finally, previous papers studied how neighborhood sorting affects income mobility across generations. The challenges for studying cultural assimilation instead of income lie in the imperfect measurement. These non-trivial extensions are possible because of the new Conditional Choice Probability (CCP) estimator based on the factor model approach (Hwang (2024))[4] and a micro panel data (Millennium Cohort Study) that includes imperfect measurements on the cultural traits and spans sufficiently many years (18 years) to observe cultural assimilation among immigrants.

The structural estimates show that living close to White British lowers the assimilation cost for non-White British, who have been historically immigrants in Great Britain, so desegregating the housing market changes their assimilation speed. Another notable finding from the estimates is that it is more costly for White British to move into a diverse neighborhood than for non-White British. Often, immigrants are blamed for self-isolation, but the estimates imply that segregation might have occurred mostly due to White British avoiding neighbor-

[2] Aliprantis and Carroll (2018) assumed a stationary environment for estimation but showed a transition path to a new steady state in their numerical experiment.

[3] There are 34,378 LSOAs in England and Wales and 5,462 LSOAs in London. Each LSOA includes 1000 to 3000 residents, and the US Census tract includes 2500 to 8000 residents on average.
https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeographies/2001censusgeographies
https://www2.census.gov/geo/pdfs/reference/GARM/Ch10GARM.pdf

[4] Estimating this model using a full-solution simulation-based approach, such as an Indirect Inference, would be challenging because of the non-stationarity assumption, the possibility of multiple equilibria, many endogenous aggregate state variables, and imperfect measurement of cultural assimilation efforts. The CCP estimator can reduce the computational time for a general-equilibrium model with non-stationarity and potential multiple equilibria compared to most simulation-based estimators because it does not require solving for equilibrium for estimation. Compared to the previous CCP estimator, à la Arcidiacono and Miller (2011), which pools information from longitudinal choice information only, the Hwang (2024) estimator makes use of noisy proxies of cultural traits, taking into account potential measurement errors in the survey responses (Hu (2008), Hu (2017)), and allows the unobservable state variable, here the level of assimilation, to evolve *endogenously* conditional on *all* state variables.

hoods with many non-White British. Indeed, the data shows that many non-White British gradually moved to White British neighborhoods over time, whereas many White British moved to neighborhoods with even more White British than the previous neighborhood.

Next, using the model estimates, I discuss how to design a rent subsidy conditional on the cultural dissimilarity from the neighbors and a budget-balancing tax. Although this policy is unprecedented, it shares the same spirit with the existing Singaporean policy, which imposes an ethnicity quota for each neighborhood to prevent segregation of ethnic minorities (Wong (2013)). I simulate the outcomes when the UK government offers a substantial monetary incentive in the form of a rent subsidy if people reside in a culturally different neighborhood. Specifically, the White British rent subsidy is set proportionally to the share of non-White British in the neighborhood, and the non-White British rent subsidy is proportional to the share of White British in the neighborhood. To subsidize such a large-scale intervention, I assume the UK government must adopt an additional linear income tax imposed on everyone. Using the estimated model, I simulate the speed of neighborhood desegregation, immigrants' cultural assimilation, and welfare gain and loss[5] from introducing the social engineering policy.

I find that under such a hypothetical rent subsidy policy, we can induce immigrants' cultural assimilation and neighborhood desegregation, the effect size of which depends on the rent subsidy rate. The rent subsidy rate is determined by the maximum subsidy rate chosen by the government and the resident profile of the residing neighborhood. Under most policy parameters considered, White British would experience a welfare loss, and non-White British would gain welfare, but there also exists a set of policy parameters that yield welfare improvement on average for both White British and non-White British. When non-White British perfectly culturally assimilate under such a policy, migration barriers to move into any neighborhood are lowered, so it can increase the option value from migration opportunities, thereby increasing White British's welfare.[6] For example, adopting a 40% maximal subsidy rate and a budget-balancing tax rate of 2.2% for White British and non-White British would not hurt White British welfare on average but would improve non-White British welfare substantially. Under a 40% maximal subsidy policy, the unassimilated population share would drop from 9.94% to 9.6% (3.4% change) in the year 2040, and the segregation index[7] would drop from 0.55 to 0.51 (7.2% change). The average subsidy amount under a 40% maximal subsidy policy is modest compared to the current average social housing subsidy rate 36% in London and 30% in England (Adam et al. (2015)): White British would receive 7.3% of

[5]Welfare is defined as Utilitarian welfare with equal weight on everyone.

[6]The presence of externalities in the model prevents applying the First Welfare Theorem, which guarantees the Pareto efficiency of competitive equilibrium.

[7]This is often called a dissimilarity index, varying between 0 (no segregation) and 1 (perfect segregation). When a dissimilarity index is $x \in [0,1]$, it means x% of the population needs to be moved to achieve an even distribution of subgroups.

rent on average as a subsidy, and non-White British would be subsidized for 20.4% of rent on average. Other policies can induce much bigger reductions in the unassimilated population share and segregation index at the cost of White British's welfare loss. For example, if White British are willing to accept welfare loss up to 1.7%, we can expect to reduce 16% unassimilated population share (from 9.94% to 8.30%) and 39% of the segregation index (from 0.55 to 0.34) by the year 2040 by adopting a 100% maximal subsidy rate.

This paper is organized as follows: Section 2 explains context and data used for analysis and reduced-form evidence regarding neighborhood sorting and cultural assimilation. Section 3 presents a structural model consistent with the reduced-form evidence. Section 4 discusses the estimation strategy and the results. For brevity, I relegate technical details about estimation to Supplemental Appendix Section C. Section 5 provides the main counterfactual analysis of rent subsidy policy under a budget-balancing tax rate. Section 6 concludes.

2 Descriptive Evidence

2.1 Context and Data

I study the research question in the context of England and Wales, where the share of the largest ethnic group, White British, is constantly falling due to immigration. The White British share fell from 87.5% to 74.4% between the years 2001 and 2021, according to national Census[8]. The largest non-White British ethnic group is South Asian (Indian, Pakistani, Bangladeshi), many of whom follow non-Western religions, including Islam. Muslims, who comprise about 5% of the population (Census 2011), are often identified as socially isolated. Nearly half of Muslims live in a neighborhood where at least one out of every five residents is Muslim, according to the Census 2011.

The main dataset used for estimation is the Millennium Cohort Study (MCS, hereafter; University of London. Institute of Education. Centre for Longitudinal Studies. (2022)). As the name indicates, the survey tracks cohort members born between 2000 and 2002 and their natural guardians (mostly natural parents). I mostly use information from guardian surveys. This data is useful for this study for many reasons: first, it oversampled ethnic minority households, including 500 Muslim households. Second, it repeatedly surveyed cultural preferences, which are identified as the most important cultural dimensions to understand Muslim immigrants' assimilation in Western Europe (Adida et al. (2016)); See Table 1 for these variables[9].

[8]Source: https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity

[9]The MCS data includes many gender attitude-related variables. However, to improve computational speed, I selected a subset of gender attitude variables to use for estimation. To choose the variables, I first chose a

Table 1: Subjective Measures Included in Estimation

Questionnaire	*Wave*
Variables on Religiosity	
• How often, if at all, do you attend services or meetings connected with your religion. Do you attend...	1, 2†, 3,4
Gender Attitude	
• A child is likely to suffer if his or her mother works before he/she starts school	1, 2
• Couples who have children should not separate	1,2
• All in all, family life suffers when the woman has a full-time job	1
Language Ability	
• Is English the language usually spoken at home?	1,2,3‡, 4,5,6
• Whether interviewed in English	1,2,3,4,5,6,7

Note: This table shows the MCS questionnaires for the proxies used in the estimation. †: Religious attendance in wave 2 was asked only for new respondents. ‡: Response scale changed in wave 3.

Third, the Secure Access version of the data provides retrospective migration history at a very fine neighborhood level (called Lower Super Output Area, or LSOA hereafter), having only 1500 residents on average. There are 34,378 LSOAs[10] in England and Wales.

I complement the MCS with several other datasets that are summarized in Supplemental Appendix A. I use the administrative payroll tax records, called the Annual Survey of Hours and Earnings (ASHE; Office for National Statistics (2022)), to estimate the local labor market skill price and Total Factor Productivity (TFP). And I use the annual British Household Panel Survey (BHPS) and the UK Household Longitudinal Study (UKHLS) (University of Essex, Institute for Social and Economic Research (2022)) to estimate a yearly income process because the MCS does not provide income information every year, so it is inadequate for estimating yearly income process. I use Census and Census Migration Flow data to obtain local area statistics and migration distance distribution. The (median) house price data is from the HM Land Registry data, and the neighborhood deprivation index data is from Abel et al. (2016).

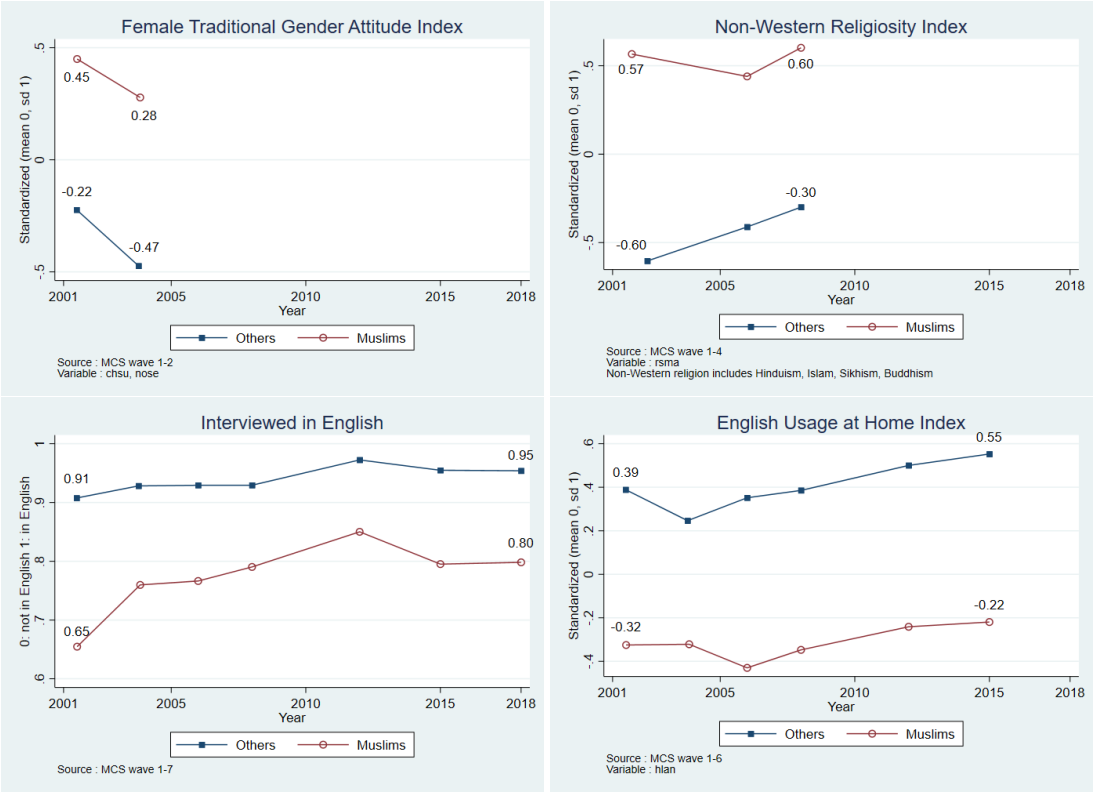
2.2 Reduced-Form Evidence

In this section, I show a few stylized facts from data, which motivate a structural model presented in the next section.

proxy variable, which seems most relevant based on the questionnaire wording, "A child is likely to suffer if his or her mother works.", and decided on other proxies that are the best predictor of the first chosen variable by conducting a LASSO estimation.

[10] I used the geographic boundaries in 2001.

Figure 1: Change in Cultural Traits



Note: This figure shows the raw means of proxy variables for cultural traits from MCS data. 'Others' is defined as non-White British, non-Muslims.

2.2.1 Fact 1: Evidence of Time-Varying Cultural Traits

The cultural traits of the immigrant sample have gradually changed over time, as shown by the raw mean of the proxy variables of these traits (Figure 1). The cultural proxy variables are summarized into a traditional gender attitude index, non-Western religiosity index, English usage at home index, and an indicator of whether a respondent was interviewed in English. Every index was normalized to have a mean of 0 and a standard deviation of 1 for interpreting the scale. I use both female and male guardian responses unless otherwise specified. Only the English interview variable was measured in all waves, so the figure shows the raw means for available years (see Table 1). The gender attitude questions were asked only in the first two waves, and for male respondents, the responses to these questions in wave two are mostly missing. So, I only report the female traditional gender attitude index changes over time. For the non-Western religiosity index, I count Hinduism, Islam, Sikhism, and Buddhism as non-Western religions. I construct this index by using a variable on the frequency of religious meeting attendance.

Table 2: Neighborhood Sorting by Cultural Preferences

	Dependent Var: NBHD White British density (std.)		
	Sample: Non White British		
Traditional Gender Attitude Index	-0.212***		
	(0.0325)		
Foreign Language Usage Index		-0.241***	
		(0.0149)	
Non-Western Religiosity Index			-0.194***
			(0.0129)
Controls	Yes	Yes	Yes
Sample	NWHB	NWHB	NWHB
N	1067	7443	4477
R-square	0.179	0.196	0.211
Partial R-square	0.038	0.077	0.088

Note: Standard errors are clustered at the household level and are presented in parentheses. All index variables and the dependent variable were standardized. Controls include male and female age quadratic polynomials, education dummies, and log labor income. For details about how each index was constructed, see Supplemental Appendix B. 'NWHB' stands for non-White British.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 1 shows that non-White British's gender attitude has become more liberal, more similar to that of White British, and they exhibited better English proficiency and more frequent use of English over time. However, interestingly, there is no evidence that their non-Western religiosity has weakened over time. Instead, it shows a slight increase in non-Western religiosity in later surveys. Muslims appear to be the most culturally distinct from White British compared to the "Others", defined as non-White British, non-Muslims. Still, both groups appear to have assimilated in gender attitudes and language dimensions over time. To sum up, there is evidence of change in non-White British cultural preferences over time, but not all cultural dimensions changed in the same way: in terms of gender attitude and language, non-White British appear to have assimilated into White British, most notably in language dimension, but their religiosity has not changed very much.

2.2.2 Fact 2: Neighborhood Sorting by Cultural Preferences

In this section, I show that cultural preferences are closely associated with neighborhood choice. This evidence motivates to build a model of a neighborhood choice conditional on (ethnic minorities') cultural preferences.

Table 2 shows how a neighbor's ethnicity – measured as (standardized) White British density in a neighborhood – varies by a respondent's cultural preferences. I use similar cultural preference indices used in the previous section. For details about how each index was constructed, see Supplemental Appendix B. Every regression controlled for economic variables (education, income, age), and standard errors were clustered at the household level. Partial R-squares of cultural preference indices are presented in the last row of the table. The results show that non-White British exhibiting more distinct cultural traits from White British tend to live in neighborhoods with fewer White British neighbors: one standard deviation increase in cultural proxy indices is associated with 0.19 to 0.24 standard deviation decreases in the neighborhood White British density. That is, there is a strong sorting into different neighborhoods by cultural preferences, even after controlling for key economic variables.

2.2.3 Fact 3: Migration Patterns in the Estimation Sample

Next, I show the migration patterns observed in the estimation sample[11] in Table 3. In the selected sample, the attrition is not severe: I observed the same individuals for 15 years on average out of 18 total years. In particular, the number of observed years is similar between White British and non-White British. This assuages concerns about selective attrition for non-White British samples.[12] The migration rate is slightly higher for non-White British: 50% of White British and 55% of non-White British moved at least once during the sample period, and the number of moves conditional on having moved at least once was 1.5 for White British and 1.3 for non-White British. Both White British and non-White British samples exhibited the migration pattern of moving to neighborhoods with more White British (56% and 68% of moves) and fewer Muslims (55% and 68% of moves) over time, but non-White British showed a stronger tendency to prefer such neighborhoods: Compared to White British, the non-White British sample moved to a very different neighborhood from the current one. Conditional on moving, non-White British moved to a new neighborhood with 13% higher White British

[11]For details on sample selection, see Section 4.1. The MCS survey is not an annual survey, but I construct a yearly migration history based on retrospective questions. Specifically, I use the variable, "year when Cohort Member moved to the current address." I use a natural parent sample only, excluding the step-parent sample, so I assume parents lived with the cohort member all the time. To minimize the sample loss, I impute the exact year to move to the current address when the information is missing by drawing a random year between the last and current survey years. The average year gap between surveys is not large, only 2 - 4 years on average, so this imputation would create only a 1-3 year difference from the actual moving year. For further information on the missing rate of the year moved to the current address, see Table E.1 in Supplemental Appendix.

[12]The non-White British sample includes more immigrants, which means they might be more likely to have emigrated during the sample period and have disappeared from the sample. The lack of detailed attrition in data precludes identifying emigration events and making a direct comparison. However, we observe a similar number of years for White British and non-White British samples, so selective emigration would not be a serious concern for this sample.

Table 3: Summary Statistics (Mean, SD) of Migration Patterns

	White British	Non-White British
Num of Observed Years	15.387	14.981
	(2.401)	(2.381)
% Ever Moved 2001-2018	0.5	0.545
	(0.5)	(0.498)
Num of Moves (Cond. on Ever Moved)	1.474	1.306
	(0.798)	(0.622)
Δ NBHD White British Density (%, Cond. on Moving)	0.024	0.130
	(0.111)	(0.256)
Δ NBHD White British Density (%, Cond. on Staying)	-0.004	-0.011
	(0.005)	(0.008)
% Moved to Less White British NBHD	0.438	0.317
	(0.496)	(0.466)
% Moved to More White British NBHD	0.562	0.683
	(0.496)	(0.466)
Change in NBHD WHB Density (Cond. on Moving to Less WHB NBHD)	-0.039	-0.115
	(0.066)	(0.134)
Change in NBHD WHB Density (Cond. on Moving to More WHB NBHD)	0.074	0.244
	(0.113)	(0.217)
Δ NBHD Muslim Density (%, Cond. on Moving)	-0.008	-0.088
	(0.063)	(0.219)
Δ NBHD Muslim Density (%, Cond. on Staying)	0.001	0.006
	(0.003)	(0.007)
% Moved to Less Muslim NBHD	0.551	0.682
	(0.498)	(0.466)
% Moved to More Muslim NBHD	0.449	0.318
	(0.498)	(0.466)
Change in NBHD Muslim Density (Cond. on Moving to Less Muslim NBHD)	-0.026	-0.171
	(0.073)	(0.204)
Change in NBHD Muslim Density (Cond. on Moving to More Muslim NBHD)	0.013	0.09
	(0.037)	(0.121)
Num of Individuals	1234	1234
Num of Individual-Year	18987	18487

Source: Millennium Cohort Study, wave 1-7.

Note: This table shows the mean and standard deviation of migration pattern-related variables in the estimation sample. Mean is shown without parentheses, and the standard deviations are presented in parentheses. For sample selection, see Section 4.1. 'WHB'=White British, 'NBHD'=Neighborhood.

density and 8.8% lower Muslim density. On the other hand, White British moved to a new neighborhood with only 2.4% higher White British density and 0.8% lower Muslim density.

Table 4: Correlation Between Change in NBHD Demographics and Non-White British Cultural Assimilation

	Depvar: Last - First Cultural Preference Indices					
	All Sample			Movers Only		
	Traditional Gender Attitude	Non-Western Religiosity	Foreign Language Usage	Traditional Gender Attitude	Non-Western Religiosity	Foreign Language Usage
Δ NBHD WHB %	0.392	-0.772**	-0.467*	-0.230	-0.918**	-0.522*
	(0.303)	(0.287)	(0.181)	(0.377)	(0.348)	(0.207)
Δ NBHD Muslim %	0.157	-0.167	0.0832	-0.168	-0.445	0.0285
	(0.413)	(0.357)	(0.205)	(0.460)	(0.425)	(0.230)
Δ Years	2.318	4.974	9.320	2.412	5.711	10.495
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample	NWHB	NWHB	NWHB	NWHB Movers	NWHB Movers	NWHB Movers
N	705	1415	1873	177	491	794
R-square	0.168	0.235	0.248	0.207	0.216	0.276

Source: MCS Wave 1-7

Note: Standard errors are in parentheses. Controls include the number of years between the first and last survey, first-year cultural preference indices, and first-year neighborhood White British and Muslim densities. All indices are standardized, so the dependent variable changes are interpreted as changes in the index's standard deviation. 'WHB' means White British, 'NWHB' is non-White British. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.2.4 Fact 4: Correlation Between Migration and Cultural Preferences

In this section, I show the evidence that cultural assimilation is closely intertwined with migration. I use the same non-White British sample and similar cultural preference indices[13] used in Table 2 and examine the correlation between *changes* in the cultural preference indices and *changes* in their neighborhood environment, measured as White British and Muslim densities. The first three columns used all samples, and the last three columns used movers only. Those who were living in ethnic enclaves and who were least assimilated in the first year are the ones who have the potential to show the biggest change in cultural preference. Therefore, I control for their characteristics in the first survey year in all regressions, includ-

[13] I made one minor change in the cultural preference index construction to minimize missing values. For the traditional gender attitude index, I used only female responses and 'chsu' and 'nose' variables only. First, another gender attitude variable 'fasu' was asked only once, so using this variable would prevent constructing a longitudinal gender norm index. Second, too few male respondents answered any gender attitude question repeatedly, so I used female responses only to minimize missing values.

ing the first-year cultural preference indices and first-year neighborhood White British and Muslim densities. Also, I control for the number of years between the first and last survey years.

Table 4 shows that those who lived with more White British over time are the ones who showed more cultural assimilation in terms of lower non-Western religiosity and less foreign language use. In all samples, a 100% increase in the neighborhood White British density is associated with a 0.772 standard deviation decrease in non-Western religiosity and a 0.467 standard deviation decrease in foreign language usage. The point estimates for the mover sample imply a much stronger association, but they are statistically not different from the ones for the entire sample. There is no evidence of a significant correlation between changes in a traditional gender attitude and changes in the neighborhood environment. However, gender attitude questions were asked in only waves 1 and 2, so the time difference between the first and last survey years is too small to observe substantial cultural assimilation. To sum up, there is strong evidence that non-White British cultural assimilation is associated with changes in the neighborhood demographics, and this provides a direct motivation to study a dynamic model of neighborhood choice and ethnic minorities' cultural assimilation.

2.2.5 Fact 5: Association Between Economic and Cultural Assimilation

In this section, I investigate if there are any economic returns in the labor market to cultural assimilation. Table 5 shows individual fixed-effect regressions, regressing the labor market outcomes – female labor earnings and female employment – onto cultural indices, which were used in the previous sections. For the male sample, I ran a similar regression for male labor earnings and the results are similar to those of the female sample (Table E.2 in Supplemental Appendix).

The results show that cultural assimilation is not associated with changes in labor market earnings: this is plausible if ethnic minorities with limited cultural and language capital work in occupations that do not require much of those skills but other manual skills. And in those occupations, a modest improvement in language skills must not result in a wage increase. However, a reduction in traditional gender attitudes is associated with a higher female employment rate. There is no evidence that either changes in non-Western religiosity or English proficiency are associated with female employment. To sum up, there is no evidence of a direct monetary return to cultural assimilation in the labor market other than through changing female work preferences and additional female spouse labor income.

Table 5: Association Between Economic and Cultural Assimilation

	Dependent Variable:					
	Female Log Labor Income			Whether Female Works		
Traditional Gender Attitude Index	0.0633			-0.0430*		
	(0.0672)			(0.0194)		
Non-Western Religiosity Index		-0.00843			-0.00216	
		(0.0288)			(0.00590)	
Foreign Language Usage Index			0.00556			0.000697
			(0.0284)			(0.00569)
Sample	NWHB	NWHB	NWHB	NWHB	NWHB	NWHB
Age Pol	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Labor Market FE	Yes	Yes	Yes	Yes	Yes	Yes
N	1037	2107	3383	2538	6155	10384
R-square	0.060	0.036	0.039	0.054	0.043	0.062

Source: MCS Wave 1-7

Note: Standard errors are clustered at the household level and are in parentheses. In every regression, a quadratic age polynomial, individual fixed effect, year fixed effect, and labor market fixed effect are controlled. The unit for labor income is in 2015 £. 'NWHB' stands for non-White British. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3 Model

This section presents a Schelling (1971)-type segregation model where every household chooses its own position in the physical space with a preference to be close to culturally similar neighbors. However, there is a key difference from a Schelling (1971) model: ethnic minority households can make costly efforts to assimilate culturally. Their cultural assimilation decision is conditional on their plan about where to move in the physical space. If it is better to move close to the majority, for example, for access to higher paying jobs, then ethnic minorities are better off investing in cultural assimilation ahead to reduce the cost of living close to a culturally different group. On the other hand, if there is an option to move to a nice ethnic enclave that is culturally similar, then ethnic minorities are less likely to assimilate culturally.

I take this idea to build a *realistic* version of a neighborhood choice model that considers various market equilibrium factors for the neighborhood choice, as well as ethnic minorities' cultural assimilation. To represent cultural assimilation, I assume everyone has a position in a

one-dimensional social space, where all natives are assumed to be positioned in the left-most corner, and ethnic minorities can choose a position relative to natives.[14]

Below, I formally introduce the model environment, timing assumptions, and notations.

I assume a *non-stationary* infinite time-horizon, $t \in \mathcal{T} = \{1, \cdots, \infty\}$. The non-stationarity assumption is useful to capture the fast-changing social environment in the UK context: immigration has been rapidly increasing during the time period captured in the data, so a stationary environment is unlikely. Inflows and outflows of the population are taken exogenously. I study an equilibrium under perfect foresight (for the next two periods) assumption, which I will elaborate on later. The decision-maker i is a married household whose female spouse is aged between $\mathscr{A}_0 > 0$ and $\mathscr{A}_1 < \infty$. For tractability, I abstract away from marriage and divorce problems.

I assume each household makes multiple discrete choices sequentially within a period. State variables are updated after each discrete choice is made, so I introduce a notation of three subperiods within a period, $t(s), s \in \{1, 2, 3\}$. I denote the state variables at the beginning of each subperiod at time t as $\Omega_{it(s)}$. The state variables at time t in a subperiod s, $\Omega_{it(s)}$, includes (1) neighborhood index $j_{it(s)}$, (2) social position $u_{it(s)}$, (3) female human capital $h_{it(s)}$, (4) female age $a_{it(s)}$, (5) a dummy whether the female respondent is White British D_i, (6) time index $t(s)$. Due to the discrete nature of survey measurements, I assume social position $u_{it(s)}$ is a discrete variable, so $u_{it(s)} \in \{1, \cdots, \overline{u}\}, \overline{u} < \infty$.[15] I include only female human capital stock into state variables to limit the number of state variables for computation, and I assume the male human capital stock is a parametric function of female human capital stock. To sum up, the state variables are

$$\Omega_{it(s)} = \{j_{it(s)}, u_{it(s)}, h_{it(s)}, a_{it(s)}, D_i, t(s)\}. \tag{1}$$

At the beginning of each period t, the household first chooses a new neighborhood $j_{it(2)}$ to live in, and next female spouse decides whether to work, $P^F_{it(2)}$, conditional on the skill price at the new location $j_{it(2)}$. Finally, the (non-White British) household can choose the new social position for the next period $u_{it+1(1)}$[16]. Before making each discrete choice, relevant stochastic shocks are realized. There are four types of shocks in the model. There are choice-specific Gumbel shocks for neighborhood and social position choice denoted as

[14] Although natives can also assimilate culturally into ethnic minorities in principle, I assume away cultural assimilation of natives mainly due to data limitation: the survey data does not have questions to measure how natives assimilate into ethnic minorities' culture. On the other hand, there are available survey questions for ethnic minorities' cultural assimilation. And this assumption is likely to be reasonable in the specific context I study, the United Kingdom: for example, most white British people do not learn Hindu or Urdu when more South Asians immigrate to the UK. On the other hand, South Asian immigrants learn English after arrival.

[15] I use both female and male spouse survey responses to estimate the social position.

[16] The timing assumption that the household chooses the social position one period in advance is motivated by the "limited feedback" assumption used in Hu and Shum (2012).

$\{\epsilon_{1it(1)}, \cdots, \epsilon_{Jit(1)}\}$, $\epsilon_{jit(1)} \overset{i.i.d.}{\sim} Gumbel(0,1)$[17] and $\{\xi_{1it(3)}, \cdots, \xi_{\overline{u}it(3)}\}$, $\xi_{uit(3)} \overset{i.i.d.}{\sim} Gumbel(0,\sigma_u)$. And there is a normally distributed human capital shock to explain wage variation, $\omega_{it(2)} \sim N(0,\sigma_\omega^2)$, and a logistically distributed shock for the household utility from female leisure to match female employment variation, $\eta_{it(2)} \sim logistic(0,\sigma_p)$.

I assume every structural parameter can vary by White British ethnicity D_i, but I omit the superscript D_i for notational brevity. The flow payoff associated with the location choice j' conditional on current location j is modeled as a flexible function of origin and destination characteristics and household characteristics. Specifically, γ_0^j captures the base housing price elasticity, and γ_1^j is a vector of coefficients for the exogenous neighborhood amenity vector[18]. The neighborhood amenity vector includes the (minus log of) neighborhood deprivation index, the log of college graduate share, the housing quality index, and the control function, which is crucial to address the endogeneity of the housing price. See Appendix D for details on the construction of the control function. To explain briefly, the control function is a measure of the unobserved neighborhood amenities, which can explain why people prefer to move to an expensive neighborhood. The instrumental variable used for the control function is similar to the one introduced in Bayer et al. (2016), and it is the predicted housing price in the non-adjacent but nearby neighborhoods, defined as an area between 5 miles and 10 miles away from the neighborhood. These non-adjacent but close neighborhoods are substitutes in the local housing market and can affect housing prices in equilibrium, but they are non-adjacent, so their unobserved amenities are likely to be not correlated.

I assume moving to a culturally dissimilar neighborhood incurs a moving cost, which depends on the distance between one's own social position and the average social position in the destination neighborhood $C_2(|u_{it(1)} - \overline{u}_{j'}|)$.[19] This term can explain residential segregation between White British and non-White British. Other terms in the moving cost function include the physical distance between origin and destination and the female spouse's (rescaled) age.[20]

$$\Pi^1(\Omega_{it(1)}; j') = \gamma_0^j log(\text{house price}_{j't(1)}) + \gamma_1^j f_t^{Amenity}(j')' + MC(j_{it(1)}, j') \tag{2}$$

[17] The Gumbel shock dispersion parameter for a location choice is normalized to 1 for identification.

[18] I investigated an extended model where the utility for amenity vector depends on the social position type. I find the estimates to be very imprecise, and they are not statistically different by social position type. So I use a simple model assuming the same γ_1^j for everyone.

[19] I find that the social distance term in the moving cost function can fit the data better than the social distance term in the flow utility. Intuitively speaking, there is a strong tipping point behavior observed in the data (Card et al. (2008)). The inflow of culturally dissimilar residents does not immediately lead to the outmigration. However, conditional on moving, people tend to move to a much more culturally similar neighborhood. I also investigated an extended version including an additional social distance term in the flow utility, but the model likelihood remains very similar, and the coefficient for non-White British is statistically not different from zero. The coefficient for White British is small and significantly positive.

[20] The rescaled age, $\widetilde{Age} = (Age - 20)/10$, was used for a computational advantage.

$$MC(j_{it(1)}, j') = (C_0 + C_1 \widetilde{Age}_{it} + C_2(|u_{it(1)} - \overline{u}_{j't}|) \times 1(j_{it(1)} \neq j') + C_3 log(distance(j_{it(1)}, j') + 1), \quad C_0, C_1, C_2, C_3 > 0 \tag{3}$$

The flow payoff in the second subperiod following the female employment decision comprises of utility from the log of household income ($log(R_{lt(2)} \times (h_{it(2)} + h^{SP}_{it(2)}))$ if the female spouse works and $log(R_{lt(2)} \times h^{SP}_{it(2)})$ if the female spouse does not work and disutility from working conditional on social position, $\sum_{k=1}^{\overline{u}} \gamma^P_{1k} 1(u_{it(s)} = k)$. $R_{lt(2)}$ is the equilibrium skill price in the labor market l that the neighborhood j belongs to.

$$\Pi^2(\Omega_{it(2)}; P^F_{it(2)}) = \left\{ \gamma^P_0 log(R_{lt(2)} \times (h_{it(2)} + h^{SP}_{it(2)})) + \sum_{k=1}^{\overline{u}} \gamma^P_{1k} 1(u_{it(s)} = k) \right\} P^F_{it(2)} + \gamma^P_0 log(R_{lt(2)} \times h^{SP}_{it(2)})(1 - P^F_{it(2)}) \tag{4}$$

I assume people accrue human capital stocks by having more work experience, and they also face a stochastic random shock every period. Non-White British social position does not affect wage growth, which is consistent with the reduced-form evidence. My specification for human capital dynamics is similar to a popular specification in the literature (Meghir and Pistaferri (2004) and Meghir and Pistaferri (2011)), and it is a mixture of a stochastic random walk and a deterministic wage gain term from work experience.[21] As for other parameters, I assume both stochastic human capital shock distributions and the deterministic wage gain parameter differ by White British ethnicity to capture potential discrimination in the labor market and skill level differences. In regards to the timing assumption, the stochastic human capital shock is realized before making a female employment decision $P^F_{it(2)}$ in the second subperiod, and the experience gain is accrued right after the second subperiod, following the female employment decision $P^F_{it(2)}$.

$$logh_{it(2)} = logh_{it(1)} + \omega_{it(2)} \tag{5}$$

$$logh_{it(3)} = logh_{it(2)} + \lambda_0 P^F_{it(2)} \tag{6}$$

$$logh_{it+1(1)} = logh_{it(3)} \tag{7}$$

I allow an i.i.d. measurement error in the *observed* wage as follows.

$$\tilde{w}_{lt(2)}(h_{it(2)}) = w_{lt(2)}(h_{it(2)}) m_{it(2)}, \quad logm_{it(2)} \underset{i.i.d.}{\sim} (0, \sigma^2_m), \tag{8}$$

[21] I do not allow for a transitory human capital shock because they are not separately identifiable from measurement errors in the absence of an external validation study providing the variance of the measurement errors. For further discussion on this, see Meghir and Pistaferri (2004) and Meghir and Pistaferri (2011). Note also that I do not assume human capital depreciation during unemployment spells to keep a finite dependence property, which provides much computational gain. The other papers which made the same assumption to have finite dependency include Altuğ and Miller (1998), Bishop (2012), Humphries (2017).

The third subperiod flow payoff includes only the psychic cost for changing social position, which depends on the fraction of White British in the neighborhood, $\overline{D}_{jt(3)}$. That is, the exposure to more White British in the same neighborhood may change the cultural assimilation cost.

$$\Pi^3(\Omega_{it(3)}; u') = PC_{u_{it(3)}u'} + PC^D_{u_{it(3)}u'}\overline{D}_{jt(3)} \tag{9}$$

The model is a general equilibrium model which includes three equilibrium objects: (i) average characteristics of neighbors ($\overline{u}_{jt(2)}$), (ii) skill price $R_{lt(2)}$, and (iii) housing (rental) price r_{jt}. To close the model, next, I specify the aggregate production function in the local labor market and housing market equilibrium equation.

For each local labor market, I assume a one-sector (tradable good) and small open economy. Each local labor market has a time-varying total factor productivity A_{lt}. The only input in the production function is the aggregate human capital stock, $H_{lt(2)}$, in the local labor market, including both working female's and male's. I assume perfect substitutability for human capital stock as an input[22]. The production function is written as

$$Y_{lt(2)} = A_{lt}H^{\zeta}_{lt(2)} \tag{10}$$

$$H_{lt(2)} = \sum_{l_{it(2)}=l} \left\{ h_{it(2)}P^F_{it} + E\left[h^{SP}_{it(2)} | h_{it(2)} \right] \right\}. \tag{11}$$

The production function implies the following skill price and wage equation.

$$R_{lt(2)} = A_{lt}\zeta H^{\zeta-1}_{lt(2)} \tag{12}$$

$$w_{lt(2)}(h_{it(2)}) = R_{lt(2)}h_{it(2)}. \tag{13}$$

Next, for the housing market, I assume the housing supply in each neighborhood, $\overline{H}_j$, is perfectly inelastic, which is a reasonable approximation in the UK context, where the housing supply is highly inelastic (Ball et al. (2010), Kim et al. (2012)). Equilibrium rent r_{jt} in each neighborhood is determined to clear the housing market.

$$r_{jt} \times \overline{HS}_j = sh \times TW_{jt} \tag{14}$$

$\overline{HS}_j$: total housing supply in neighborhood j, sh: expenditure share on rents,

[22]To allow for imperfect substitutabilities between demographic groups, such as CES production function, a large dataset of income, demographics, and location, including many labor market switchers, is required. However, the Annual Survey of Hours and Earnings (ASHE), used for estimation in this paper, does not include much demographic information, as it comes from administrative payroll tax records. Other panel datasets are unsuitable for estimation because the datasets are usually small and do not include many labor market switchers, which are key for identification.

TW_{jt}: total wage of residents in neighborhood j at time t.

Total wage is simply the sum of a couple's income in the neighborhood. $f(\Omega_{it})$ is the distribution of state variables, and Pop_t is the population stock from Census.[23]

$$TW_{jt} = \int_{j_{it}=j} \left(w_{it}P_{it} + E\left[w_{it}^{SP}|h_{it}\right]\right) f_t(\Omega_{it}) di * Pop_t \tag{15}$$

To impute male spouse's earnings using female spouse human capital, I adopt the following linear function in equation 16.

$$E\left[log(h_{it(2)}^{SP})\middle|h_{it(2)}\right] = b_1 + b_2\ log h_{it(2)} \tag{16}$$

Now I define the ex-ante value function for each subperiod, $V_{t(1)}(\Omega_{it(1)}), V_{t(2)}(\Omega_{it(2)}), V_{t(3)}(\Omega_{it(3)})$. Note that there is no discounting within each period, but only between subperiod 3 and the next period subperiod 1.

$$V_{t(1)}(\Omega_{it(1)}) = E\left[\max_{j'\in\mathcal{N}} \left\{\Pi^1(\Omega_{it(1)}; j') + \epsilon_{j'it(1)} + E\left[V_{t(2)}(\Omega_{it(2)})|\Omega_{it(1)}, j'\right]\right\}\right], \tag{17}$$

$$V_{t(2)}(\Omega_{it(2)}) = E\left[\max_{p\in\{0,1\}} \left\{\Pi_j^2(\Omega_{it(2)}; p) + \eta_{it(2)}(1-p) + E\left[V_{t(3)}(\Omega_{it(3)})|\Omega_{it(2)}, p\right]\right\}\right], \tag{18}$$

$$V_{t(3)}(\Omega_{it(3)}) = E\left[\max_{u'\in\{1,\cdots,\overline{u}\}} \left\{\Pi^3(\Omega_{it(3)}; u') + \xi_{u'it(3)} + \beta E\left[V_{t+1(1)}(\Omega_{it+1(1)})|\Omega_{it(3)}, u'\right]\right\}\right]. \tag{19}$$

It is worth mentioning that the above value function has two-period finite dependence, which speeds up estimation later.[24] Therefore, the value function can be written as a function of the next two-period choice probabilities only, and the perfect foresight for the next two periods is sufficient. Appendix C.4 provides the formal statement of this result and the derivation.

I define an equilibrium concept in this model.

Definition 1 (Markov Perfect Equilibrium). *An equilibrium consists of policy functions*

[23]Because the estimation sample is small, I utilize both Census aggregate statistics as well as panel dataset to estimate $f(\Omega_{it})$. First, using the estimation survey sample, I estimate the conditional distribution of other state variables based on age and neighborhood characteristics (White British, Muslim density, median house price, skill price, neighborhood amenity vector). Next, combining with the age distribution for each neighborhood from Census 2011, I compute the full state variable distribution $f(\Omega_{it})$.

[24]In page 1836, Arcidiacono and Miller (2011) provides a more precise mathematical definition of ρ-period finite dependence. "A pair of choices exhibits ρ-period dependence if there exists a sequence from initial choice j and a corresponding sequence from initial choice j' such that for all (states) $z_{t+\rho+1}$, (the transition probabilities to state in period $t+\rho+1$ are equalized) $\kappa^*_{t+\rho}(z_{t+\rho+1}|z_t, j) = \kappa^*_{t+\rho}(z_{t+\rho+1}|z_t, j')$". For example, any life-cycle model that has a terminal age and terminal value function normalized to zero automatically has a finite dependence property. However, the finite dependence property is useful only when the choice paths require a small number of periods.

$\{j_{it(2)}, P^F_{it(2)}, u_{it+1(1)}\}_{t\in\mathcal{T}}$ *and endogenous aggregate state variables* $\{R_{lt(s)}, r_{jt(s)}, \overline{u}_{jt(s)}\}_{l\in\mathcal{L}, j\in\mathcal{N}, t\in\mathcal{T}}$ *such that*

(i) Policy function $\{j_{it(2)}, P^F_{it(2)}, u_{it+1(1)}\}_{t\in\mathcal{T}}$ *is a solution to the household maximization problem defined in equation (17, 18, 19) and under the information set assumption, given the payoff-relevant endogenous aggregate state variables* $\{R_{lt(s)}, r_{jt(s)}, \overline{u}_{jt(s)}\}_{l\in\mathcal{L}, j\in\mathcal{N}, t\in\mathcal{T}}$ *and exogenous state variables in a current period t.*

(ii) The skill price vector $\{R_{lt(s)}\}$ *is determined to clear each competitive local labor market. That is, equation 12 holds.*

(iii) Rent vector $\{r_{jt(s)}\}$ *is determined to clear each neighborhood. That is, equation 14 holds.*

(iv) Aggregate state transition is consistent with household policy functions.

Due to strategic complementarity in the moving cost, this model might have multiple equilibria. Deriving an analytical condition that guarantees a single equilibrium is difficult for a high-dimensional model like this. However, potential multiple equilibria are not a concern when carrying out estimation because the CCP estimation does not require solving for an equilibrium. For counterfactual simulations, I choose an equilibrium that is close to a status quo.[25]

3.1 Model Discussion

The model highlights the relationship between geographic sorting and immigrant assimilation.

First, the economic factors changing geographic dispersion will inevitably change immigrants' assimilation. Such factors include monetary incentives, like local skill prices and rents, as well as the economic disparity (in human capital) between immigrants and natives. These factors make immigrants choose a different neighborhood, meet different neighbors, and alter their incentive to invest in cultural assimilation. For example, even if the main reason for immigrants to live in an ethnic enclave is economic reasons, once they start living in an ethnic enclave, they may stop learning the natives' language because they can use their ethnic language in their enclave. In the model, this is captured by higher assimilation costs when living in an ethnic enclave.

Second, immigrants' assimilation effort is undersupplied in a competitive equilibrium because immigrants do not internalize the welfare gain of natives when they assimilate into natives' culture. For example, non-White British in the model do not consider that their cul-

[25] Although I do not prove my model has a unique equilibrium, there is a good reason to believe a single unique equilibrium is likely: Bayer and Timmins (2005)'s extensive simulation shows that a sorting model with a large choice set, higher heterogeneity in location characteristics, and agent's heterogeneous preference makes it more likely to have a unique sorting equilibrium. These are the conditions that my model exhibits.

tural assimilation would lower the moving cost for White British to move into their ethnic enclaves. The implication is that competitive equilibrium may not be optimal, and social engineering-style government intervention may improve social welfare by inducing immigrants' assimilation efforts.

Third, segregation can occur due to either White British or non-White British's high reluctance to move into a culturally dissimilar neighborhood. Often, immigrants are blamed alone for their segregation and lack of assimilation efforts. However, according to the model, immigrants might become segregated due to natives' high prejudice (represented as high moving cost C_2) and their consequent "White flight". Once immigrants become isolated, it is more costly for them to assimilate culturally.

4 Estimation

4.1 Sample Selection

The main estimation sample comes from parent surveys in MCS data between years 2001 and 2018[26], and I select samples by several criteria: I select married couples[27] living in England and Wales, whose female spouse is aged between 25 and 64. I use consecutive spells, including estimation-relevant information[28], and I dropped short spells (less than ten years)[29]. I use all eligible non-White British samples, but I randomly select the same number of White British samples as non-White British samples because of computational capacity constraints in the data server[30]. In the end, I obtained 2468 households (1234 White British, 1234 non-White British households).

Table 6 summarizes the basic demographics of the female sample members in the first interview. I show a similar table for male sample members in the Supplemental Appendix (Table E.3). The average age for both White British and non-White British samples is in their early 30s. By construction, the White British sample does not include any Muslims[31]. Nearly half of the non-White British sample are Muslims, and most (98%) are first-generation or second-generation immigrants. First-generation immigrants spent 12 years on average at the time of

[26]Most interviews in England and Wales for wave 7 were completed in the year 2018, so I do not use a small number of interviews done in year 2019.

[27]To minimize the sample loss due to missing marital status information, I assume the couple is married if both mother and father of the cohort member filled out a survey.

[28]This includes prejudice proxy variables, neighborhood geocode, mother's detailed ethnicity (whether White British), religion, mother's and father's education, and mother's employment.

[29]It is difficult to identify unobservable cultural type with a shorter spell of data.

[30]I could access the Millennium Cohort Survey Secure Access license data only inside a secure data server.

[31]I excluded a small number of White British Muslims from an estimation sample.

Table 6: Main Parent (Female) Characteristics in the First Interview

	White British	Non-White British
Survey Year	2001.968	2002.021
	(1.267)	(1.386)
Age	32.56	30.987
	(4.312)	(4.56)
Muslim		0.532
		(0.499)
Not an Immigrant		0.022
		(0.147)
First Gen Immigrant		0.686
		(0.464)
Second Gen Immigrant		0.292
		(0.455)
Years in the UK (1st Gen)		12.009
		(8.426)
College Graduation	0.456	0.335
	(0.498)	(0.472)
Employed	0.686	0.347
	(0.464)	(0.476)
% NBHD WHB Density	0.916	0.458
	(0.109)	(0.29)
% NBHD Muslim Density	0.017	0.292
	(0.051)	(0.28)
Sample Size	1234	1234

Source: MCS wave 1-7

Note: This table shows the average and standard deviations of female sample characteristics in the first interview. The standard deviation of each variable is in parentheses. By construction, the White British sample does not include any Muslims. I do not report the immigration-related statistics for the White British sample, as they are small in number.

the first survey. Compared to White British females, non-White British females had a much lower rate of college graduation and female employment. The White British sample mostly lived in White British towns (where the average neighborhood White British density is 92%), compared to the non-White British sample (their average neighborhood White British density is only 46%). Similarly, the White British sample rarely had any Muslims in their neighborhood (average neighborhood Muslim density is 2%), but the non-White British sample had more Muslim neighbors (average neighborhood Muslim density is 29%).

Table 7: Preset Parameters

Parameter	Description	Value	Source
β	Time Discount	0.95	Kennan and Walker (2011), Bishop (2012), Bayer et al. (2016)
$\frac{r_{it}}{\text{house price}_{it}}$	Rent-to-house price ratio	0.037	Table A1.1 in Marsden (2015)
sh^{WHB}	White British Expenditure Share on Housing	0.199	Calibration Using UKHLS 2009-2018
sh^{NWHB}	Non-White British Expenditure Share on Housing	0.226	Calibration Using UKHLS 2009-2018

Note: Marsden (2015) used rent and housing price data from the year 2014. When computing the expenditure share on housing, I imputed rent for homeowners using the rent-to-house price ratio 0.037 reported in Marsden (2015). The housing expenditure did not include utilities.

4.2 Choice Set Construction and Preset Parameters

First, I construct a small stratified random choice set, following the approach in Keane and Wasi (2012) because there are too many neighborhoods (34378 LSOAs) compared to our sample size, so considering migration to all possible neighborhoods poses a curse of dimensionality problem. I stratify neighborhoods into three types: White British neighborhoods, which are defined as neighborhoods whose White British density exceeds the national average, 80.5%; Muslim neighborhoods, which are defined as Muslim density being above 20%; and other neighborhoods, which are not either White British or Muslim neighborhoods. I made a random drawing of 10 neighborhoods in each category. First, I drew a migration distance using the migration distance distribution from Census 2001 Migration Flow data to match the fact that people consider moving to physically closer neighborhoods. Next, I drew a neighborhood whose distance from the origin is closest to the drawn migration distance. Including the origin neighborhood, this gives 31 potential destination neighborhoods. When an actual neighborhood is not included in this neighborhood, I replace one neighborhood in the stratum to which the actual neighborhood belongs with the actual neighborhood. The choice set of size 31 is large enough that every 34,378 neighborhood has both inflows and outflows of migration.

Several parameters were preset, and they are summarized in Table 7. The time discount parameter was set to 0.95 following the previous migration literature (Kennan and Walker (2011), Bishop (2012), Bayer et al. (2016)). The rent-to-house price ratio was taken from Marsden (2015), which used the rent and house price data from the year 2014. Expenditure share on housing by White British ethnicity was calibrated using UKHLS data between the years 2009 and 2018.

Table 8: Income Process Estimates

Parameter	Description	Estimate	
		WHB	NWHB
λ_0	Experience Gain	0.045	0.023
		(0.009)	(0.013)
σ^2_ω	Variance of Permanent Shock	0.036	0.035
		(0.002)	(0.005)
σ^2_m	Variance of Measurement Error	0.024	0.028
		(0.001)	(0.004)

Source: Harmonized British Household Panel Survey & UK Household Longitudinal Study, 2001-2018. Standard errors are in parentheses. "WHB"="White British", "NWHB"="Non-White British".

4.3 Income Process Estimates

The MCS data does not have information on annual income changes, so I use other annual panel datasets, Harmonized British Household Panel Survey (BHPS) and UK Household Longitudinal Study (UKHLS), to estimate the experience gain parameter λ_0, the variance of the permanent income shock σ^2_ω, and variance of the measurement error σ^2_ω. I estimate parameters separately by White British ethnicity.

The estimation procedure is similar to Meghir and Pistaferri (2004), Blundell et al. (2008), Voena (2015)[32]. I correct for selection in female employment, following Voena (2015). The instrumental variable to correct for selection in female employment was whether (own and spouse's) mother had worked when the respondent was age 14. The exclusion restriction requires the instrumental variable to be independent of current period income *shock* (not the *level*), and therefore, the pre-determined lifetime event must satisfy the exclusion restriction. Also, the variable strongly correlates with female employment in the current period through preference to work. Table A.2 shows the first stage estimation result, which shows that the IVs have strong predictive power for female employment.

The results show that the estimate for non-White British experience gain is smaller than that of White British. The variance of permanent income shock and the variance of measurement error are similar for both White British and non-White British.

[32]The only small difference is in that I consider endogenous skill prices which vary across labor markets and years.

Table 9: Production Function Estimates

Parameter	Description	Estimate	SE
ζ	Production Function Curvature	0.653	0.066
$\overline{A}_{lt}$	Average log skill price		
	London	0.242	0.010
	South East	0.207	0.010
	Midland	0.104	0.010
	South	0.094	0.008
	North	0.090	0.010

Source: Annual Survey of Hours and Earnings, 2003-2018
Note: Skill price was estimated as labor market - year specific. I report the average skill price and average standard errors here for brevity. For the full list of log skill prices for each local labor market every year, see Table F.1 in the Appendix.

4.4 Production Function Estimation & Human Capital Stock Calibration

First, I construct local labor market groups by clustering the Travel-to-Work areas (TTWA)[33], which are the official local labor market units used in national statistics.[34] I applied the k-means clustering algorithm to the TTWAs' average real income between the years 2003 and 2018 and the distance to three major metropolitan areas, London, Manchester, and Birmingham, and obtained five local labor market groups.[35] Figure A.1 in the Appendix shows the map of the labor market classified by this algorithm.

I follow Heckman et al. (1998) to estimate production function parameters. For details on this estimation, see Section C. The main identification idea is that I instrument the endogenous labor supply using an instrumental variable, which is plausibly independent of Total Factor Productivity (TFP) in each local labor market. I deploy a Bartik instrument (Bartik (1991)) using an industry composition share in the base year (2003). Table A.1 in the Appendix shows the first stage estimation result, which shows that the instrumental variable has strong predictive power for total human capital stock (labor supply) in the local labor market.

Table 9 shows the production function curvature parameter estimate and the average log skill price for the five local labor markets. London is the highest-paying labor market, and the South Eastern region surrounding London is the next highest-paying market. The midland area, including Manchester and Birmingham, provides the third highest pay. The South and North rural areas are the lowest-pay areas. Figure E.4 in the Supplemental Appendix shows the estimated log skill prices and log total human capital stock for every labor market each

[33] I used the 2001 geography boundary.
[34] The ASHE dataset is large but still not large enough to estimate the annual skill price for each TTWA.
[35] For the geographic continuity of each labor market, I manually switched two TTWAs into a geographically closer local labor market group. The TTWAs are Berwick, and Morpeth, Ashington & Alnwick in the North near Scotland.

year.

Given the production function and income process estimates, I calibrate the female human capital stock $h_{it(s)}$ before estimating the model, which leaves only the social position type as an unobservable state variable and simplifies the estimation. The calibration considers the selection into work by human capital stock. First, the initial human capital stock for *ever-worked* individuals is calibrated as the one maximizing the simulated likelihood of the observed wage sequence. The income process parameters and the skill prices are used to simulate the wage sequence, given the initial human capital stock. For those who never worked during the sample period, first, I predict their human capital stock using a regression model estimated from the *unemployed, ever worker* sample only[36], and calibrate the initial human capital stock to maximize the simulated likelihood of the predicted human capital sequence from regression-based prediction.

4.5 Estimation of Social Position Types

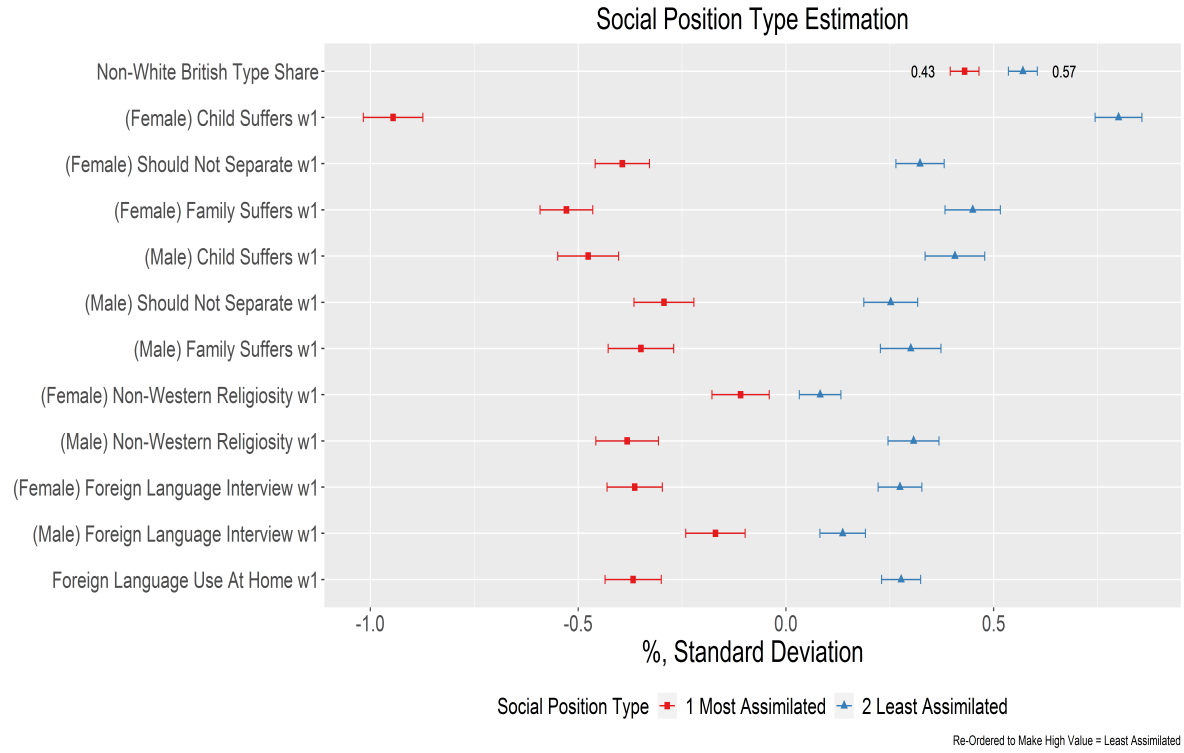
I estimate the likelihood of proxy variable responses about social position using the EM algorithm and the first wave data only before the rest of the estimation procedure, thereby reducing the computational time. This is feasible because all proxy variables about culture were collected in the first wave (Table 1). For identification, I adopt a monotonicity assumption[37] in Hu (2017) and use the variable whether a female spouse interviewed in a foreign language in wave 1 as an anchoring variable; that is, I label a social position type "Least Assimilated" if the female spouse is more likely to interview in a foreign language.

I choose the number of social positions $\overline{u}$ to two based on the proxy likelihood estimation result: Figure E.2 in Supplemental Appendix shows the social position type share estimates and the average responses to the proxy variables when assuming three social position types. The middle type (Type 2) is largely indistinguishable from the least assimilated type (Type 3). Also, the estimation results vary more when choosing a different anchoring variable, as types are harder to distinguish (Figure E.3 in Supplemental Appendix). On the other hand, when I assume two social position types, each type is clearly distinguished from the others (Figure 2), and the estimation result is invariant to the choice of anchoring variable (Figure E.1 in Supplemental Appendix). Based on this, I conclude three types are too many to identify each type's characteristics separately in this sample.

[36] The regression model includes fairly detailed demographic information. They are female and male education dummies, female quadratic age polynomial, (log) house price, (log) skill price in the labor market, (log) neighborhood Muslim and White British density, and ethnicity group dummies.

[37] The assumption is, "There exists an observable variable Y_{it} such that $E[Y_{it}|u_{it}]$ is increasing in social position u_{it}." Y_{it} is a variable whether a female spouse interviewed in a foreign language.

Figure 2: Initial Social Position Type Estimates (Assuming Two Types)



Source: MCS wave 1
Note: This figure shows the initial social position type share estimates and the average survey responses of each type, assuming two social position types. Whiskers show the 95% confidence intervals, which were computed from bootstrapping 100 times.

4.6 Two-Stage CCP Estimation

The remaining structural parameters in the flow payoff are estimated using the Hwang (2024)'s two-step Conditional Choice Probability (CCP, hereafter) estimator, which extends Arcidiacono and Miller (2011) estimator to incorporate likelihoods from proxy variables, and to allow a flexible evolution of an unobserved state variable. See Hwang (2024) for further explanation on the properties and performance of this estimator.

Algorithm 1 summarizes the first stage estimation procedure; see Appendix C.3 for more details. In the first step, the algorithm is initialized by setting the individual social position type probability $P^{(0)}(u_{i1} = s)$, initial prior $P^{(0)}(u_{i1} = s|\Omega_{i1(1)})$, and transition CCP $P^{(0)}(u_{it+1} = s'|u_{it} = s, \Omega_{it})$.[38] Next, the algorithm continues iterating between M-, E-step until the type probability $P^{(m)}(u_{it} = s)$ converges. In the M-step, the per-period likelihood, consisting of the likelihood of neighborhood choice, female work choice, and proxy variables on social

[38]For details on initialization, see footnotes in Appendix C.3.

Algorithm 1 First Step Reduced-Form CCP Estimation, More Details in Appendix C.3

- Initialization
- Iterating M-step, and E-step until type share probabilities converge

M-Step. Estimate (per-period) "likelihoods" treating Ω_{it} observable.

$$L_{it} = \underbrace{L_{it}^{j}}_{\text{NBHD choice}} \quad \underbrace{L_{it}^{P}}_{\text{Female Work}} \quad \underbrace{L_{it}^{Y}}_{\text{Proxies}} \tag{20}$$

- Transition likelihood L_{it}^{u} estimated from regressing nonparametric forward probability $Prob(u_{it+1(s)} = s' | \Omega_{it})$ on the state space Ω_{it} (Papke and Wooldridge (1996))

E-Step. Update $u_{it(s)}$

- $u_{it(s)}$: deterministic updating using Bayes rule, $\frac{L_{it}(u_{it(s)}=u')}{L_{it}}$

position, is estimated using a parametric (multinomial) logit model. To save computational time, the likelihood for proxy variables on social position is estimated beforehand using proxy variables existing in the first wave. Once the per-period likelihood is computed, the transition likelihood for the social position is estimated using the parametric fractional multinomial logit (Papke and Wooldridge (1996)). In E-step, I update the social position type probability using the Bayes rule.

The second stage estimation to estimate the structural parameters is done via quasi-MLE, where the likelihood function exploits the two-period finite dependence. Appendix C.4 shows the derivations to construct the likelihood function.

4.7 Social Position Type Share Estimates

The social position type share estimates (Figure 3) from the two-step CCP estimation shows substantial cultural assimilation for the estimation sample for 18 years. The least assimilated type among non-White British share fell from 59% in the year 2001 to 42% in the year 2018. Table 10 shows the average characteristics of each estimated social position type. The least assimilated type (Type 2) includes 2-3 years younger sample members, more Muslims, and more first-generation immigrants who spent fewer years in the UK on average. Also, they include more non-working female spouse members, less-educated female and male spouses, and lower-income households. They live in a neighborhood that has more Muslim residents and fewer White British residents. The moving rates were similar for both types, although the

Figure 3: Social Position Type Share Estimate

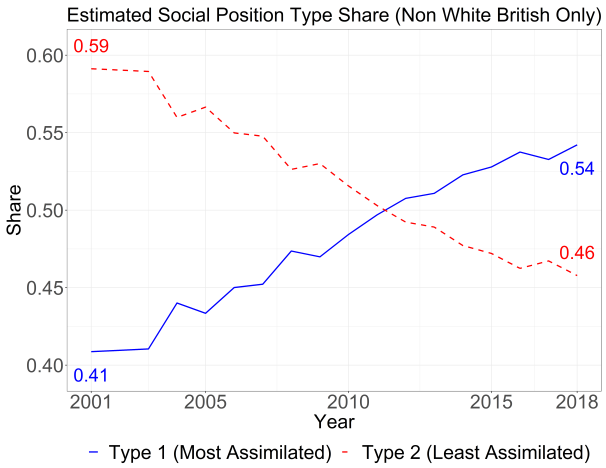


Table 10: Estimated Social Position Type Average Characteristics

	Type 1 (Most Assimilated)	Type 2 (Least Assimilated)
Female Age	39.820	36.851
Male Age	42.937	40.786
Female Muslim	0.293	0.724
Female Not an Immigrant	0.029	0.017
Female First Gen Immigrant	0.617	0.755
Female Second Gen Immigrant	0.354	0.229
Female Years in the UK (First Gen Only)	20.737	18.042
Female Employment	0.936	0.029
Female College graduate	0.513	0.197
Male College Graduate	0.539	0.313
Female Average Annual Income (2015 £)	20170.239	16026.993
Male Average Annual Income (2015 £)	39257.921	25079.443
Neighborhood Muslim Density (%)	0.182	0.390
Neighborhood White British Density (%)	0.530	0.361
Annual Move Rate	0.045	0.050

Source: Millennium Cohort Study, Wave 1-7
Note: This table shows the average characteristics of each social position type.

moving rate for type 2 was estimated to be slightly higher than that of the type 1 households.

4.8 Structural Parameter Estimates

Table 11 shows the structural parameter estimates. To highlight a few estimates, first, the social distance moving cost, C_2, was estimated to be much bigger for White British than non-White British, which implies that segregation is mostly driven by White British's reluctance

Table 11: Structural Parameter Estimate

Parameter	Description	Estimate/SE	
		White British	Non-White British
	Spouse Income Parameter Estimates		
b_1	Spouse Income Parameter (level)	9.742	7.799
		(0.162)	(0.275)
b_2	Spouse Income Parameter (slope)	0.061	0.229
		(0.017)	(0.030)
	Flow Utility Parameters For Location Choice		
γ_0^j	House Price Elasticity	-1.057	-0.657
		(0.397)	(0.321)
γ_{11}^j	-Log(Neighborhood Deprivation Index)	0.140	0.207
		(0.066)	(0.070)
γ_{12}^j	Neighborhood College Grad Share	0.730	0.275
		(0.347)	(0.303)
γ_{13}^j	Neighborhood Housing Quality Index	1.241	0.581
		(0.387)	(0.397)
γ_{14}^j	Neighborhood Control Function	1.139	1.052
		(0.399)	(0.320)
C_0	Base Fixed Moving Cost	-2.679	-3.537
		(0.068)	(0.079)
C_1	Age-dependent Fixed Moving Cost	-0.934	-0.886
		(0.033)	(0.033)
C_2	Social Distance Moving Cost	-6.960	-0.374
		(0.287)	(0.097)
C_3	Variable Fixed Cost by Log Distance	-0.406	-0.767
		(0.015)	(0.021)
	Flow Utility Parameters for Female Employment		
γ_0^P	Skill Price/Income Elasticity	2.610	1.719
		(1.272)	(1.049)
γ_{11}^P	Gender Norm Cost, WHB or NWHB Type 1	-1.500	-0.677
		(0.598)	(0.420)
γ_{12}^P	Gender Norm Cost, NWHB Type 2		-1.109
			(0.678)
	Adjustment Cost Parameters For Social Position Choice		
PC_{12}	Cultural Assimilation Cost, From Type 1 to 2, Intercept		-7.949
			(6.624)
PC_{21}	Cultural Assimilation Cost, From Type 2 to 1, Intercept		-7.761
			(6.955)
PC_{12}^D	Cultural Assimilation Cost, From Type 1 to 2, White British Share		-2.384
			(0.626)
PC_{21}^D	Cultural Assimilation Cost, From Type 2 to 1, White British Share		2.438
			(0.590)
	Gumbel Shock Dispersion Parameters		
σ_p	Logistic Dispersion Parameter for Female Work	0.451	0.114
		(0.164)	(0.027)
σ_u	Logistic Dispersion Parameter for Social Position		2.836
			(0.414)

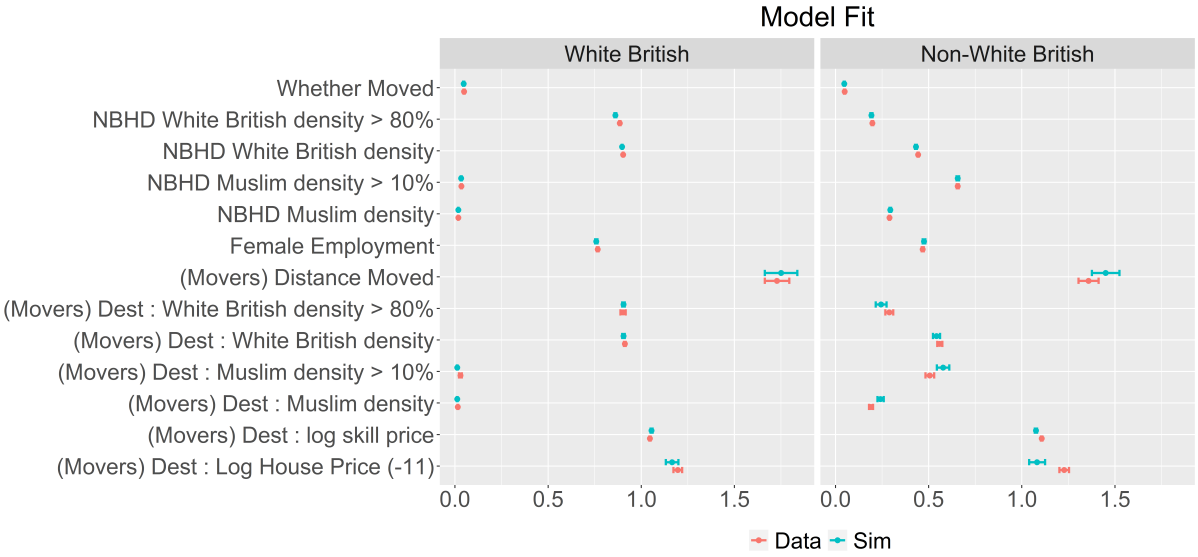
Source: Millennium Cohort Study, wave 1-7 (2001-2018)
Note: This table shows the structural parameter estimates. Standard errors are in parenthesis.

to live close to non-White British rather than the non-White British avoiding White British neighborhoods. Second, the cultural assimilation cost estimates, PC_{12}^D, PC_{21}^D, imply that living close to White British substantially lowers the cost of cultural assimilation while increasing

the cost of cultural dissimilation (Change from Type 1 to 2).

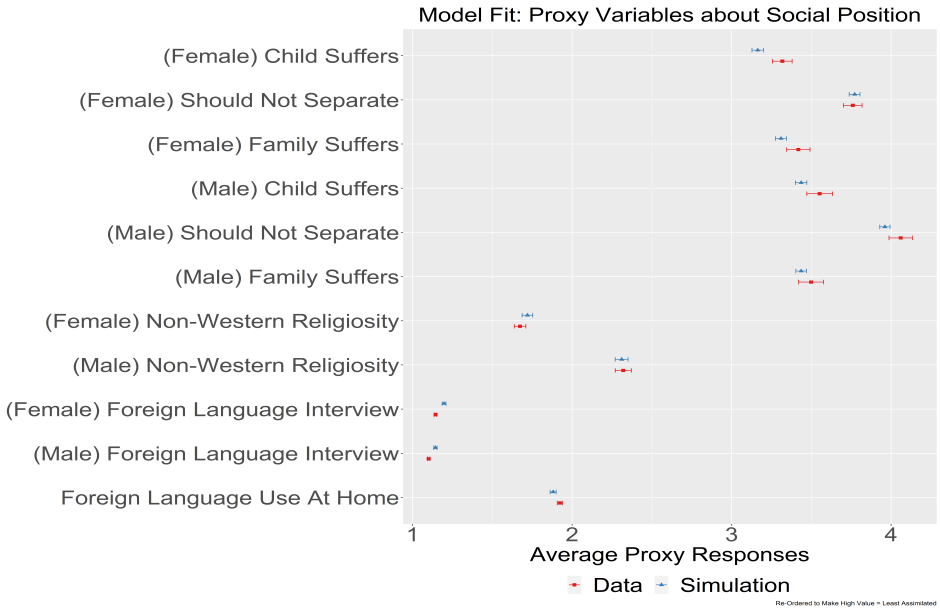
4.9 Model Fit

Figure 4: Model Fit: Migration Patterns, Female Employment



Source: Millennium Cohort Study, wave 1-7
Note: To represent on the same scale, I subtracted 11 from the log median house price.

Figure 5: Model Fit: Proxy Variables about Social Position



Source: Millennium Cohort Study, wave 1-7

Overall, model fit for most key features of the model, including migration patterns, female employment, and proxy responses about social position, is good, as shown by Figure 4 and Figure 5. First, Figure 4 shows the model fit for migration patterns and female employment. I show both neighborhood statistics for all samples and for the mover's sample only (labeled as "(Movers)"). I also display the share of respondents living in or moving to White British towns (defined as neighborhoods whose White British density exceeds 80.5%) and Muslim towns (defined as neighborhoods whose Muslim density exceeds 20%). To display model fit statistics in one graph, I adjust the scale of the log median house price of the mover's new destination neighborhood by subtracting 11. Figure 5 shows the model fit for all social position proxy variable responses.

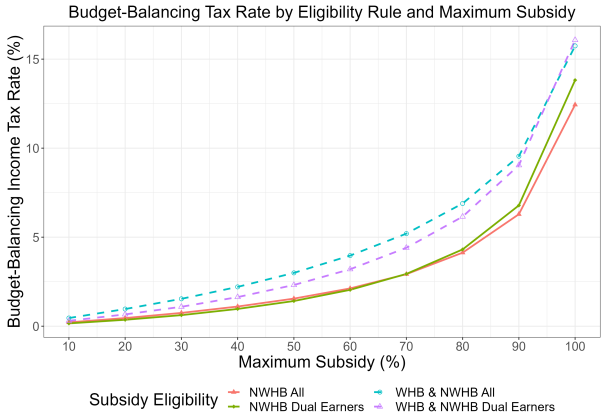
5 Counterfactual Analysis of a Rent Subsidy Policy Under a Budget-Balancing Tax

Using the structural model estimates, this section assesses a hypothetical rent subsidy policy, which is granted conditional on living in a diverse neighborhood. Specifically, non-White British receive a rent subsidy proportional to the White British share in the neighborhood, and White British receive a rent subsidy proportional to the non-White British share in the neighborhood. This rent subsidy is financed by a linear income tax levied on everyone. This rent subsidy policy provides direct monetary incentives to increase inter-ethnic mixing and can indirectly induce non-White British assimilation.[39] This policy shares a similar spirit with existing desegregation policies, such as the neighborhood ethnicity quota policy in Singapore (Wong (2013)) or race-blind social housing allocation policy in France (Algan et al. (2016)). I also investigated a variant of this policy that restricted this rent subsidy recipient to only dual-earner households. This policy expands the tax base and may incentivize non-White British females to participate in the labor market, update their traditional gender norms, and improve language skills.

Structural approach is particularly useful for this policy evaluation not only because the policy is unprecedented but also because such a large-scale intervention in the housing market would inevitably change market prices and prompt some "White flights"(Card et al. (2008)). The general-equilibrium structural model can predict how quickly the housing market will

[39]For the sake of maximizing cultural assimilation only, one may consider an alternative policy compensating for cultural assimilation with monetary rewards. For example, the government may provide tax benefits for immigrants who prove their language proficiency or dual-earner immigrant households who confront traditional gender norm. Compared to this alternative policy, the residential mixing policy can induce cultural assimilation holistically without costly screening of immigrants' cultural assimilation.

Figure 6: Budget-Balancing Linear Income Tax Rate for Different Subsidy Policies



Note: 'WHB' = White British, 'NWHB' = Non-White British.

be desegregated and how the housing market rent, the labor market skill prices, and neighborhood ethnic compositions will change. The structural model can also predict how the non-White British population would adjust their level of cultural assimilation, so it can assess the residential desegregation policy as a tool for inducing immigrant assimilation.

A few assumptions underpin this counterfactual analysis. First, I assume people expect the current price to prevail forever when forming their dynamic expectation, but they update this expectation every year after observing new prices. This alternative expectation assumption was adopted instead of the perfect foresight assumption because of computational challenge[40]. Second, I assume deterministic demographic change. Every year, a new youngest cohort (aged 25) enters the economy, and the oldest cohort (aged 65) dies. The new cohort is assumed to have the same demographic profile as the youngest cohort (aged 25) in the year 2011. I assume the population growth rate of 0.7%, which is the average population growth rate between the years 2011 and 2023[41]. Third, Total Factor Productivity (TFP) in the labor market and the housing stock is assumed to remain at the same level of year 2011.[42] The interest rate for government borrowing and saving is set at 0.053, which is consistent with the time discount β assumed ($r = (1/\beta) - 1 \approx 0.053$). The budget-balancing tax rate is the one that will make the government budget meet for the 30 years from year 2011 to year 2040, with borrowing and saving allowed.

I consider various maximum rent subsidy rates $\kappa \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$,

[40]To adopt the perfect foresight assumption, one needs to solve for a new equilibrium vector of *all* prices, including labor and housing market prices and *all* neighborhood ethnic compositions for *all* future periods, and this computation is practically infeasible.

[41]https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/populationestimatesforenglandandwales/mid2023

[42]The Total Factor Productivity estimates from the ASHE data show that they are similar between the years 2011 and 2018, and there is no evidence of TFP growth.

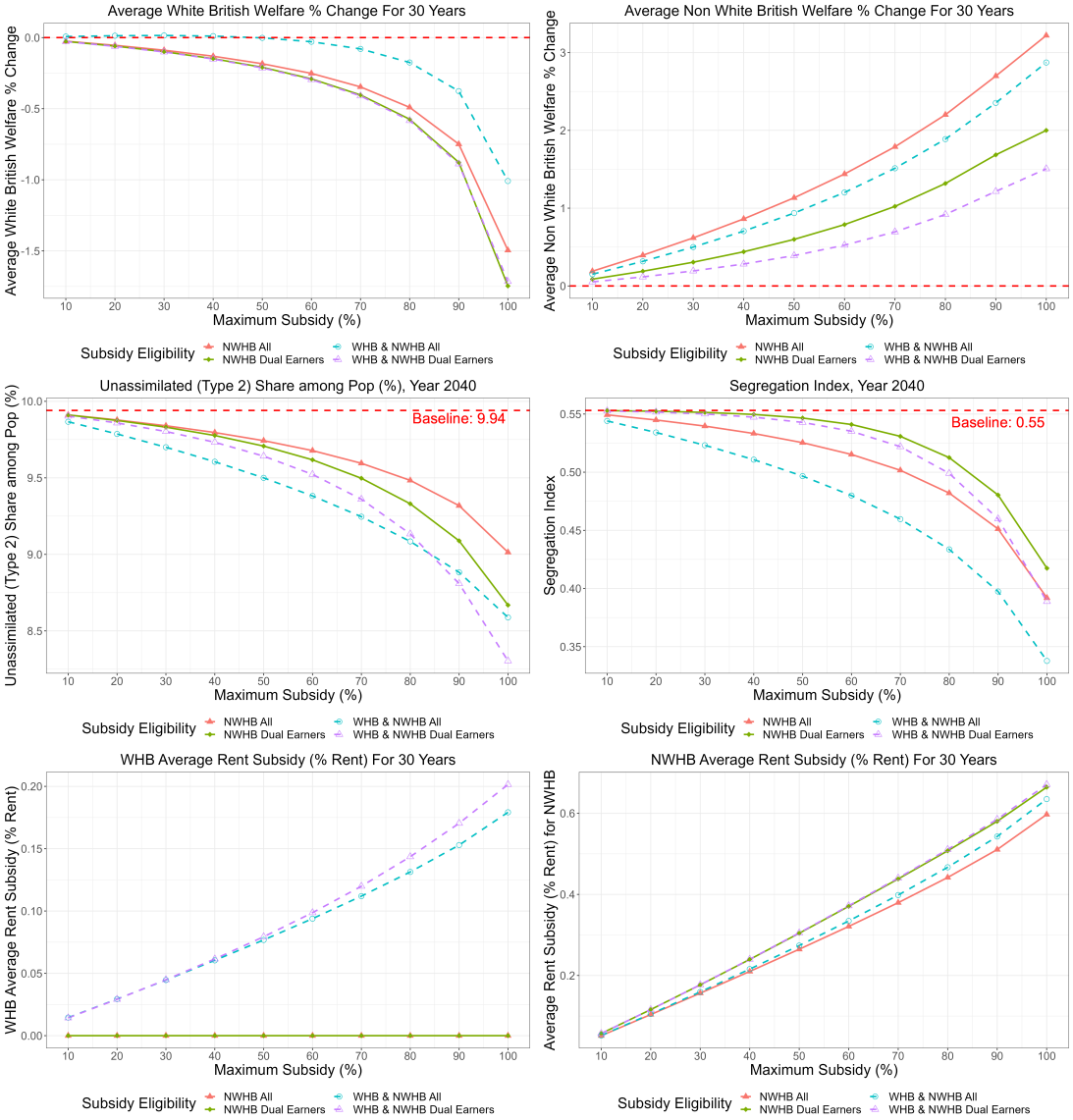
where the total rent subsidy is $\kappa \times$ (Neighborhood White British Share)% of market equilibrium rent for non-White British and $\kappa \times (1 -$ Neighborhood White British Share)% of market equilibrium rent for White British. I consider four types of subsidy policies: 1) only Non-White British dual-earner households are eligible, 2) only non-White British households (including single-earner households) are eligible, 3) non-White British and White British dual-earner households are eligible, 4) all households are eligible. I discuss which of these four policies would be efficient and achieve desegregation and cultural assimilation while curving welfare loss.

Figure 6 shows the budget-balancing linear income tax rate for different subsidy recipient policies. The budget-balancing tax rate was computed to account for endogenous state variables, including market price changes, female labor supply, and housing demand. The budget-balancing tax rate increases at a much steeper rate when the maximum subsidy rate exceeds 80%. Not surprisingly, the tax rate must be higher when White British and Non-White British are eligible for subsidy. The difference in tax rate when only dual earners are eligible and when everyone is eligible is small because households rationally increase female employment to take up subsidies, and they increase the housing demand as well, which increases the government subsidy expenditure.

Figure 7 shows the simulated outcomes under various subsidy policies. The top two panels show the average welfare change compared to baseline prediction for White British and non-White British for thirty years from the year 2011 to 2040 under various policies. Under most policies, White British experience welfare loss, but non-White British gain in welfare. However, there are a few policies that are Pareto-improving for both White British and non-White British, that is, when the maximum subsidy rate given to both White British and non-White British households is below 40%. The general equilibrium model includes externality, so the competitive equilibrium may not be optimal, and this hypothetical subsidy policy may improve both White British and non-White British welfare.

The middle two panels show the outcome of cultural assimilation and desegregation in year 2040. Compared to the baseline prediction indicated by the red dashed line, we can expect a sizable change in immigrants' cultural assimilation and segregation, whose change increases by the maximum subsidy rate. Under the 40% maximal subsidy rate for everyone, the unassimilated share would drop from 9.94% to 9.6% (-3.4%), and the segregation index would decrease from 0.55 to 0.51 (-7.2%). There are other policies maximizing cultural assimilation and desegregation, which would result in White British welfare loss on average. The policy maximizing immigrants' assimilation is when both White British and Non-White British dual-earner households can receive up to 100% of rent as a subsidy, whereas the one maximizing desegregation is when all White British and non-White British households can receive up to 100% of rent as a subsidy. The policy restricting the subsidy recipients to only dual-earner

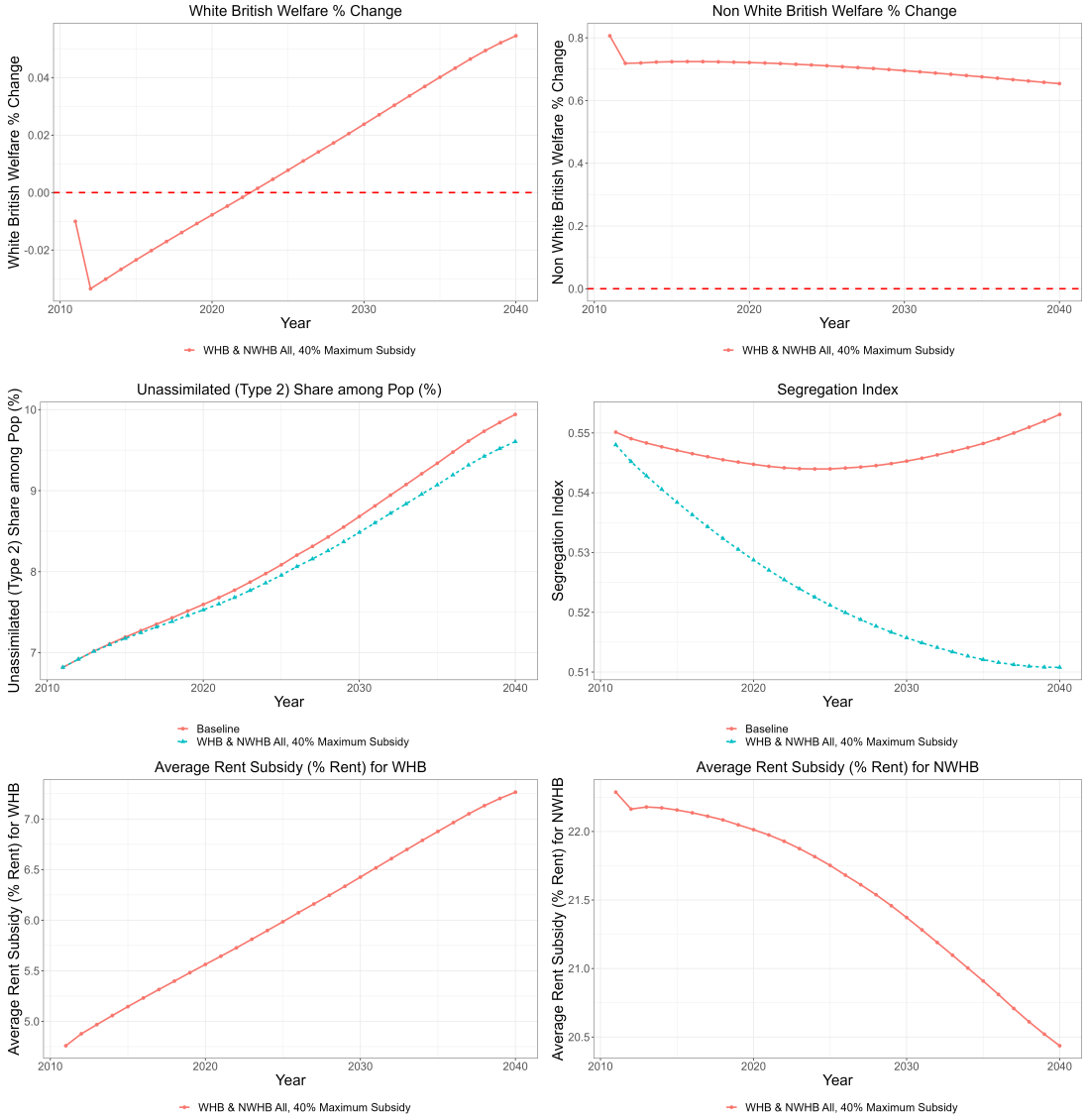
Figure 7: Simulated Outcomes Under Different Subsidy Policies



Note: 'WHB' = White British, 'NWHB' = Non-White British. The Segregation index is an index of dissimilarity, ranging from 0 to 1, and the formula is $0.5 \times \sum_{j=1}^{J}(|W_j/W - N_j/N|)$, where W_j is the number of WHB living in neighborhood j, W is the total number of WHB in the country, N_j is the number of NWHB living in neighborhood j, and N is the total number of NWHB in the country. If the segregation index is s, it means that s% of the population needs to be relocated to achieve an even distribution by ethnicities.

households encourages unassimilated female immigrants to participate in the labor market and promotes their assimilation. These two policies would be politically difficult to adopt, as White British would experience welfare loss on average, as confirmed by the top left graph in Figure 7. However, if we are willing to accept White British's welfare loss up to 1.7%, these policies would be an effective assimilation and desegregation policy (Table A.3 in Appendix): Under the 100% maximal subsidy rate for everyone policy, the unassimilated population share would drop by 1.352 percentage points (-13.605%), and the segregation index would decrease

Figure 8: Prediction for 30 Years Under 40% Maximal Subsidy Rate for All



Note: 'WHB' = White British, 'NWHB' = Non-White British. The Segregation index is an index of dissimilarity, ranging from 0 to 1, and the formula is $0.5 \times \sum_{j=1}^{J}(|W_j/W - N_j/N|)$, where W_j is the number of WHB living in neighborhood j, W is the total number of WHB in the country, N_j is the number of NWHB living in neighborhood j, and N is the total number of NWHB in the country. If the segregation index is s, it means that s% of the population needs to be relocated to achieve an even distribution by ethnicities.

by 0.215 (-38.938%). Under the 100% maximal subsidy rate for dual earners only policy, the unassimilated population share would decrease by 1.637 percentage points (-16.464%) and the segregation index would be reduced by 0.164 (-29.699%).

The bottom two panels show the average rent subsidy taken by White British and non-White British under various policies. The average subsidy taken is much lower than the maximum subsidy level because the subsidy is determined by both the maximum subsidy rate and the neighborhood's ethnic composition. To compare the generosity of this subsidy policy to a

benchmark, the average subsidy rate for social housing in the UK is 36% in London and 30% in England (Adam et al. (2015)). So, setting the maximum subsidy to 50% ($\kappa = 0.5$) would result in a similar level of rent subsidy for non-White British in our counterfactual analysis.

Figure 8 shows 30-year prediction under a 40% maximal subsidy rate for everyone, which is the maximal subsidy rate that is welfare-improving for both White British and non-White British on average. The top two panels show that White British would experience welfare loss on average for the first eleven years[43], but non-White British would not experience such welfare loss throughout 30 years. The middle two panels show that immigrants' cultural assimilation and neighborhood desegregation would monotonically improve over time compared to baseline prediction in the absence of any policy intervention. The bottom two panels show the average subsidy rate taken by White British and non-White British over time. The average subsidy rate taken by White British would gradually increase as neighborhoods become desegregated and White British move to a diverse neighborhood, but the average subsidy rate taken by non-White British would gradually decline.

To sum up, a hypothetical rent subsidy granted conditional on living in a diverse neighborhood has the potential to induce sizable immigrants' cultural assimilation and neighborhood desegregation. And there exists a welfare-improving subsidy policy for both White British and non-White British.

6 Conclusion

This paper estimated a dynamic general equilibrium model of neighborhood sorting and immigrants' cultural assimilation using 18-year-long panel data, including noisy measurements of immigrants' cultural traits. Using the model estimates, this paper discussed a hypothetical rent subsidy policy, granting rent subsidies conditional on living in a diverse neighborhood. To meet the government budget, these policies would require raising an extra linear income tax, leading to potential welfare loss of White British.

The counterfactual analysis shows that such policy can induce substantial immigrants' cultural assimilation as well as neighborhood desegregation, where the policy effect size depends on the maximum rent subsidy rate. Under most policies, White British would experience welfare loss on average, unlike non-White British, who always experience welfare gain on average. However, there also exists a set of welfare-improving rent subsidy policies for both White British and non-White British that can promote immigrants' cultural assimilation and neighborhood desegregation. The presence of externality in the model can explain this

[43]Potentially, we can mitigate their welfare loss for the first eleven years by reallocating resources across years, but computing this time-dependent policy would be computationally challenging.

welfare improvement compared to baseline competitive equilibrium. Under the 40% maximal subsidy rate and 2.2% linear income tax for everyone, White British would not lose welfare on average for the first 30 years, the least assimilated population share would drop from 9.94% to 9.6% (-3.4%), and the segregation index would decrease from 0.55 to 0.51 (-7.2%). For the first eleven years, however, White British would experience welfare loss, which is compensated by welfare gain later. If White British are willing to accept welfare loss up to 1.7%, the 100% maximal subsidy rate policy for either everyone or dual-earners only would reduce the unassimilated population share by 1.637 percentage points (-16.464%), and the segregation index by 0.215 (-39%).

It is worth mentioning that this counterfactual policy analysis does not consider other potential benefits of promoting social integration, such as better functioning political institutions and fewer social conflicts, which facilitate fast economic development. This paper mainly considered people's reluctance to live close to culturally different people and how it prevents moving to different neighborhoods, which might provide better amenities otherwise. Assessing an integration policy that considers these other potential benefits discussed in the literature would be an ambitious work in the future.

References

Abel, G., R. Payne, and B. Matt (2016). UK Deprivation Indices. https://doi.org/10.5523/bris.1ef3q32gybk001v77c1ifmty7x.

Adam, S., D. Chandler, A. Hood, and R. Joyce (2015). Social rent policy: choices and trade-offs. Technical report, IFS report.

Adda, J., C. Dustmann, and J.-S. Görlach (2022). The dynamics of return migration, human capital accumulation, and wage assimilation. *The Review of Economic Studies 89*(6), 2841–2871.

Adda, J., P. Pinotti, and G. Tura (2020). There's more to marriage than love: the effect of legal status and cultural distance on intermarriages and separations.

Adida, C. L., D. D. Laitin, and M.-A. Valfort (2016). *Why Muslim integration fails in Christian-heritage societies*. Harvard University Press.

Albert, C., A. Glitz, and J. Llull (2021). Labor market competition and the assimilation of immigrants.

Alesina, A., R. Baqir, and W. Easterly (1999). Public goods and ethnic divisions. *The Quarterly journal of economics 114*(4), 1243–1284.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003). Fractionalization. *Journal of Economic growth 8*, 155–194.

Alesina, A. and E. La Ferrara (2000). Participation in heterogeneous communities. *The quarterly journal of economics 115*(3), 847–904.

Alesina, A. and E. La Ferrara (2005). Ethnic diversity and economic performance. *Journal of economic literature 43*(3), 762–800.

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

Aliprantis, D. and D. R. Carroll (2018). Neighborhood dynamics and the distribution of opportunity. *Quantitative Economics 9*(1), 247–303.

Altuğ, S. and R. A. Miller (1998). The effect of work experience on female wages and labour supply. *The Review of Economic Studies 65*(1), 45–85.

Anstreicher, G. (2023). Spatial influences in upward mobility. *Forthcoming at Journal of Political Economy*.

Arcidiacono, P. and R. A. Miller (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica 79*(6), 1823–1867.

Arellano, M., R. Blundell, and S. Bonhomme (2017). Earnings and consumption dynamics: a nonlinear panel data framework. *Econometrica 85*(3), 693–734.

Ball, M., G. Meen, and C. Nygaard (2010). Housing supply price elasticities revisited: Evidence from international, national, local and company data. *Journal of Housing Economics 19*(4), 255–268.

Bartik, T. J. (1991). Who benefits from state and local economic development policies?

Bayer, P., R. McMillan, A. Murphy, and C. Timmins (2016). A dynamic model of demand for houses and neighborhoods. *Econometrica 84*(3), 893–942.

Bayer, P. and C. Timmins (2005). On the equilibrium properties of locational sorting models. *Journal of Urban Economics 57*(3), 462–477.

Bishop, K. C. (2012). A dynamic model of location choice and hedonic valuation. *Mimeo, Washington University in St. Louis*, 1–38.

Bisin, A., G. Topa, and T. Verdier (2004). Religious intermarriage and socialization in the united states. *Journal of political Economy 112*(3), 615–664.

Bisin, A. and G. Tura (2019). Marriage, fertility, and cultural integration in italy. Technical report, National Bureau of Economic Research.

Blundell, R., L. Pistaferri, and I. Preston (2008). Consumption inequality and partial insurance. *American Economic Review 98*(5), 1887–1921.

Card, D., A. Mas, and J. Rothstein (2008). Tipping and the dynamics of segregation. *The Quarterly Journal of Economics 123*(1), 177–218.

Dahlberg, M., K. Edmark, and H. Lundqvist (2012). Ethnic diversity and preferences for redistribution. *Journal of Political Economy 120*(1), 41–76.

Davis, M. A., J. Gregory, and D. A. Hartley (2023). Preferences over the racial composition of neighborhoods: Estimates and implications. *Available at SSRN 4495735*.

Davis, M. A., J. Gregory, D. A. Hartley, and K. T. Tan (2021). Neighborhood effects and housing vouchers. *Quantitative Economics 12*(4), 1307–1346.

Easterly, W. and R. Levine (1997). Africa's growth tragedy: policies and ethnic divisions. *The quarterly journal of economics*, 1203–1250.

Eckert, F. and T. Kleineberg (2021). Saving the american dream? education policies in spatial general equilibrium.

Fernández, R., A. Fogli, and C. Olivetti (2004). Mothers and sons: Preference formation and female labor force dynamics. *The Quarterly Journal of Economics 119*(4), 1249–1299.

Fogli, A. and V. Guerrieri (2019). The end of the american dream? inequality and segregation in us cities. Technical report, National Bureau of Economic Research.

Fouka, V., S. Mazumder, and M. Tabellini (2022). From immigrants to americans: Race and assimilation during the great migration. *The Review of Economic Studies 89*(2), 811–842.

Hasan, A., Z. Wang, and A. S. Mahani (2016). Fast estimation of multinomial logit models: R package mnlogit. *Journal of Statistical Software 75*(3), 1–24.

Heckman, J. J., L. Lochner, and C. Taber (1998). Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of economic dynamics 1*(1), 1–58.

Hu, Y. (2008). Identification and estimation of nonlinear models with misclassification error using instrumental variables: A general solution. *Journal of Econometrics 144*(1), 27–61.

Hu, Y. (2017). The econometrics of unobservables: Applications of measurement error models in empirical industrial organization and labor economics. *Journal of econometrics 200*(2), 154–168.

Hu, Y. and M. Shum (2012). Nonparametric identification of dynamic models with unobserved state variables. *Journal of Econometrics 171*(1), 32–44.

Humphries, J. E. (2017). *The causes and consequences of self-employment over the life cycle*. The University of Chicago.

Hwang, Y. (2024). Identification and estimation of a dynamic discrete choice model with an endogenous time-varying unobservable state using proxies. *Available at SSRN 3535098*.

Keane, M. P. and N. Wasi (2012). *Estimation of discrete choice models with many alternatives using random subsets of the full choice set: With an application to demand for frozen pizza*. Nuffield College.

Kennan, J. and J. R. Walker (2011). The effect of expected income on individual migration decisions. *Econometrica 79*(1), 211–251.

Kim, K., S.-Y. Phang, and S. Wachter (2012). Supply elasticity of housing.

Low, H., C. Meghir, and L. Pistaferri (2010). Wage risk and employment risk over the life cycle. *American Economic Review 100*(4), 1432–1467.

Marschner, I. C. (2011). glm2: Fitting generalized linear models with convergence problems. *The R Journal 3*, 12–15.

Marsden, J. (2015). House prices in london–an economic analysis of london's housing market. *Greater London Authority Economics 72*, 1–62.

McFadden, D. et al. (1973). Conditional logit analysis of qualitative choice behavior.

Meghir, C. and L. Pistaferri (2004). Income variance dynamics and heterogeneity. *Econometrica 72*(1), 1–32.

Meghir, C. and L. Pistaferri (2011). Earnings, consumption and life cycle choices. In *Handbook of labor economics*, Volume 4, pp. 773–854. Elsevier.

Miguel, E. and M. K. Gugerty (2005). Ethnic diversity, social sanctions, and public goods in kenya. *Journal of public Economics 89*(11-12), 2325–2368.

Office for National Statistics (2022). Annual Survey of Hours and Earnings, 1997-2022: Secure Access. [data collection]. 21st Edition. http://doi.org/10.5255/UKDA-SN-6689-20.

Papke, L. E. and J. M. Wooldridge (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of applied econometrics 11*(6), 619–632.

Piyapromdee, S. (2021). The impact of immigration on wages, internal migration, and welfare. *The Review of Economic Studies 88*(1), 406–453.

Ray, D. and J. Esteban (2017). Conflict and development. *Annual Review of Economics 9*, 263–293.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of mathematical sociology 1*(2), 143–186.

University of Essex, Institute for Social and Economic Research (2022). Understanding Society: Waves 1-11, 2009-2020 and Harmonized BHPS: Waves 1-18, 1991-2009. End User License and Special License Access, Census 2001 Lower Layer Super Output Areas [data collection]. http://doi.org/10.5255/UKDA-SN-6614-16, http://doi.org/10.5255/UKDA-SN-6670-14.

University of London. Institute of Education. Centre for Longitudinal Studies. (2022). Millennium Cohort Study End User License And Secure Access (Geographical Identifiers), from First Survey to Seventh Survey [data collection]. http://doi.org/10.5255/UKDA-SN-7758-2, http://doi.org/10.5255/UKDA-SN-7759-2, http://doi.org/10.5255/UKDA-SN-7760-2, http://doi.org/10.5255/UKDA-SN-7761-2, http://doi.org/10.5255/UKDA-SN-7762-2, http://doi.org/10.5255/UKDA-SN-8231-1, http://doi.org/10.5255/UKDA-SN-8758-1, http://doi.org/10.5255/UKDA-SN-4683-6, http://doi.org/10.5255/UKDA-SN-5350-6, http://doi.org/10.5255/UKDA-SN-5795-6, http://doi.org/10.5255/UKDA-SN-6411-9, http://doi.org/10.5255/UKDA-SN-7464-6, http://doi.org/10.

5255/UKDA-SN-8156-7, http://doi.org/10.5255/UKDA-SN-8682-2, http://doi.org/10.5255/UKDA-SN-8172-4.

Voena, A. (2015). Yours, mine, and ours: Do divorce laws affect the intertemporal behavior of married couples? *American Economic Review 105*(8), 2295–2332.

Wong, M. (2013). Estimating ethnic preferences using ethnic housing quotas in singapore. *Review of Economic Studies 80*(3), 1178–1214.

A Additional Tables and Figures

Table A.1: First Stage of IV regression Using Bartik IV

	(1)
	Log Total Human Capital Stock
Bartik IV (Predicted Employment in 10000 persons)	0.0938*
	(0.0256)
Year FE	Yes
Labor Market FE	Yes
N	80

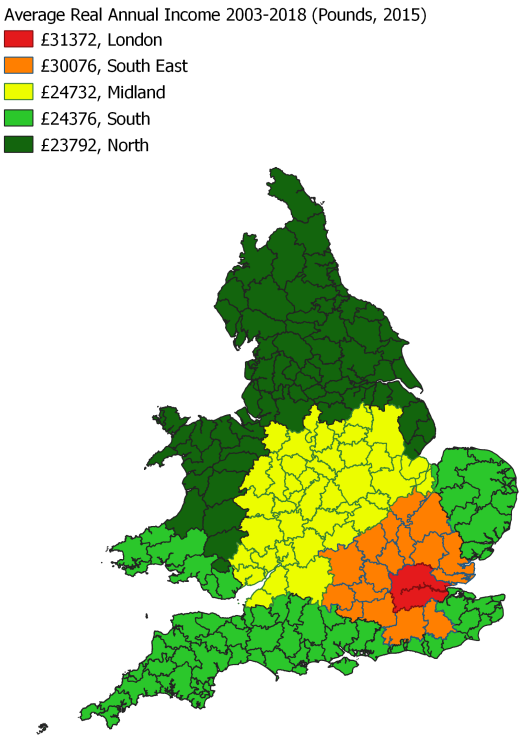
Source: Annual Survey of Hours and Earnings, 2003-2018
Note: Standard errors are clustered at the labor market level and are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.2: Income Process First Stage Regression

	Dependent Variable: Whether Female Works	
	White British	Non-White British
Whether Mother has Worked at Age 14	0.332***	0.942***
	(0.0319)	(0.0425)
Whether Spouse's Mother has Worked at Age 14	0.281***	0.874***
	(0.0311)	(0.0429)
Baseline Control	Yes	Yes
Labor Market X Time FE	Yes	Yes
N	35621	13787

Source: Harmonized British Household Panel Survey & UK Household Longitudinal Study, 2001-2018
Note: Standard errors are in parentheses. Baseline control includes a quadratic age polynomial.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure A.1: Labor Market Classification



Source: Annual Survey of Hours and Earnings, 2003-2018
Note: This figure shows labor market classification from k-means clustering of Travel-to-Work Areas (TTWA) in England and Wales using average real income (in 2015 pounds) in each TTWAs and normalized distance from London, Manchester, and Birmingham, which are three largest metropolitan areas in England and Wales.

Table A.3: Year 2040 Predictions Under Politically Less Feasible Policies

	100% Maximum Subsidy Rate	
	WHB & NWHB All	WHB & NWHB Dual Earners
Avg. WHB Welfare Change for 30 Yrs (%)	-1.009	-1.715
Avg. NWHB Welfare Change for 30 Yrs (%)	2.872	1.506
Unassimilated Pop Share Change (p.p, %) in 2040	-1.352 p.p.	-1.637 p.p.
	[-13.605%]	[-16.464 %]
Desegregation Index Change (%) in 2040	-0.215	-0.164
	[-38.938%]	[-29.699%]
Avg. WHB Subsidy Rate (% Rent) for 30 Yrs	22.773	24.953
Avg. NWHB Subsidy Rate (% Rent) for 30 Yrs	63.883	65.712

Note: This table shows predictions for the year 2040 under 100% maximum subsidy rate for everyone and for dual-earners only.

SUPPLEMENTAL APPENDIX For
“Designing an Immigrant Social Integration Policy" (Yujung Hwang)

A Data Appendix

In this section, I summarize the datasets used in this paper.

- **Annual Survey of Hours and Earnings (ASHE):** This is a 1% random sample of longitudinal Pay As You Earn records. This dataset was used to estimate skill price, local Total Factor Productivity (TFP), and production function and to classify different Travel-to-Work Areas (TTWAs) into labor markets.
- **British Household Panel Survey (BHPS, Special License):** This is a longitudinal annual household panel survey, which started in the year 1991. I use the data from the year 2001 onwards. This dataset was used to estimate the annual income process.
- **UK Household Longitudinal Study (UKHLS, Special License):** This is a longitudinal annual household panel survey. This survey started in the year 2009. The former BHPS sample was incorporated into wave 2. This dataset was used to estimate the annual income process along with BHPS.
- **Millennium Cohort Study (MCS, Secure Access):** This is a longitudinal household panel survey started in the year 2001. This dataset includes a sizable minority sample, including over 500 Muslim households, and has repeatedly been surveyed about cultural preferences. This is the main dataset used for estimation.
- **UK Census 2001, 2011, 2021:** I use local aggregate statistics computed from three UK Censuses.
- **UK Census Flow Data 2001:** I use the UK Census Flow Data 2001 for every origin-destination LSOAs pair to construct a stratified random choice set in migration decision.
- **HM Land Registry Data 2001-2018 (ONS Publication for small area statistics, LSOA 11):** This dataset publishes the quarterly median house prices paid for each small area (LSOA 2011 geography) in England and Wales. I compute the yearly average median house price and convert its geography to LSOA 2001 geography[44] to be consistent with the MCS data. The price is deflated to use 2015 British pounds as a unit. The rent-to-house price ratio, 3.7%, is from the last row of Table A1.1 in Marsden (2015). Their

[44] I compute a weighted average of yearly median house prices where the weight is how much LSOA 2001 geography is overlapped with different LSOA 2011 geography neighborhoods.

calculation is originally based on Land Registry price paid data 2015, VOA private rents 2013/14 for England.

- **Neighborhood Deprivation Index:** I use data published by Abel et al. (2016), specifically 'uk_imd_england_score', which is the recalculated IMD score for the UK, based on English IMD score. The index represents how deprived each neighborhood is, and I multiply minus to it so that a higher number implies a better amenity.

B Construction of Unobservable Attribute Indices

In this section, I describe the details about how the cultural preference indices in Table 2 were constructed. The variables used to construct the indices are listed in Table 1.

To construct the traditional gender attitude index, the variables on gender attitude in Table 1 were re-labeled so that a higher number implies a more traditional gender attitude. The overall gender attitude index was computed by summing every male and female responses to gender attitude questionnaires, and the final number was standardized again to have a mean 0 and a standard deviation 1. Only the observations with complete responses were used.

To construct the foreign language usage index, I re-labeled variables on English usage and English interview in Table 1 so that a higher number implies more foreign language usage. I summed male and female responses and normalized the sum. Only the observations with complete responses for all foreign language questions were used. There is a discontinuity in the response scale of the English usage at home variable between waves 2 and 3. I recoded the variable between waves 3 and 6 to make it comparable to a variable between wave 1 and 2: "mostly English and sometimes other language" and "about half English and half other language" in wave 3 and afterwards were recoded to "English and other language" in wave 1 and 2 variable. "Mostly other language and sometimes English" and "Other languages only" in wave 3 and afterwards were recoded to "Other languages only" in wave 1 and 2 variable. "English only" response does not need a re-coding.

To construct a non-Western religiosity index, I used variables on the male and female spouse's attendance at non-Western religious services. The non-Western religions include Hinduism, Islam, Sikh, and Buddhism. I summed up the variables for each spouse and normalized the sum to construct the non-Western religiosity index.

C Details on the Estimation

This section explains estimation details omitted for brevity in the main text.

C.1 Step 1. Income Process

In this section, I explain how to estimate the income process, taking into account selective female labor force participation and the general equilibrium effect in the local labor market. I estimate the income process parameters separately by White Britishness, and for brevity, I omit the White Britishness index $D_i \in \{\text{WHB}, \text{NWHB}\}$ in notations for parameters.

First, following the usual convention in the income process literature (for instance, Blundell et al. (2008)), I trim the income outliers and income growth outliers. I have set any monthly income below 100 pounds as equivalent to not working. I dropped the top 1% income in the data. I dropped any income growth greater than 150% and income drop greater than 150%, which is roughly the bottom and top 1% of annual income growth observed in the data. I restrict the analysis to England and Wales and use only married people aged between 25 and 64 to be consistent with the MCS sample. I drop the sample that appears only once in the dataset as I can not observe income growth. I use the sample who *did not move across labor markets* [45].

Estimation proceeds sequentially as below.

- Step 1. Estimate the income profile slope $\{\lambda_0\}$.
  - Step 1.1. Compute the individual log wage residuals. That is, subtract the average log wage by labor market and year, $\overline{logw_{lit(2)}}$. I allow additional i.i.d. component $logm_{it(2)} \sim (0, \sigma_m^2)$ in the observed log wage variable, which can be interpreted as the mixture of transitory income shock and measurement error.[46]

$$logw_{lit(2)} = logR_{lt(2)} + logh_{it(2)} + logm_{it(2)}$$

[45] There are only a small number of labor market switchers in the panel dataset because the labor market definition used in this paper covers a large geographic area. So, this leads to only a small selection in the sample. However, using labor market stayers only simplifies the estimation procedure. In addition, the structural model assumed in this paper implies that there is no selection in the income shock between labor market switchers and stayers because the wage shock is assumed to be realized after the migration decision. So this treatment is fully consistent with the structural model.

[46] As explained in Meghir and Pistaferri (2004), Low et al. (2010), Arellano et al. (2017), a transitory component in income shock is not separately identifiable from measurement error. If there is a validation study, which provides the estimate for the variance of measurement error in the income variable, then one may be able to estimate the variance of a transitory income shock. Unfortunately, I am not aware of any validation study for UK income data. Another partial identification available is to use the bounding method, as discussed in Meghir and Pistaferri (2004).

$$= \underbrace{\left(logR_{lt(2)} + \overline{logh_{lt(2)}}\right)}_{\approx \overline{logw_{lit(2)}}} + \underbrace{\left(logh_{it(2)} - \overline{logh_{lt(2)}} + logm_{it(2)}\right)}_{\equiv s_{it(2)}}$$

$$\Leftrightarrow s_{it(2)} \equiv logw_{lit(2)} - \overline{logw_{lit(2)}} = logh_{it(2)} - \overline{logh_{lt(2)}} + logm_{it(2)}$$

– Step 1.2. Estimate the control function $\lambda_{it(2)}(\alpha_{it})$ for female. To obtain a closed-form control function, I assume the dependence between preference shock to working η_{it} and permanent shock to human capital stock ω_{it} can be modeled with Gaussian copula with correlation parameter $\rho_{\omega\eta}$. The instrumental variable (Z_{it}), which is excluded from the *change*[47] in individual wage residual, is whether (own and spouse's) mother worked when the respondent (and the spouse) was age 14. The idea is that these variables are pre-determined so they are not correlated with *human capital shock* in the current period. However, they are strongly correlated with an individual's preference for working.[48] Other exogenous regressors X_{it} include labor market X time dummy and quadratic age polynomials.

$$P(P_{it}^F = 1) = P(X_{it}\beta_1 + Z_{it}\beta_2 + \eta_{it} > 0), \quad \eta_{it} \sim \text{logistic}(0,1)$$

$$(\Phi^{-1}(F_\omega(\omega_{it(2)})), \Phi^{-1}\left(F_\eta\left(\eta_{it(2)}\right)\right)) \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{\omega\eta} \\ \rho_{\omega\eta} & 1 \end{pmatrix}\right)$$

$$\lambda_{it(2)}(\alpha_{it}) = \frac{\phi(\Phi^{-1}(F_\eta(\alpha_{it})))}{1 - F_\eta(\alpha_{it})}$$

$$E\left[\omega_{it} | \eta_{it} > \alpha_{it}\right] = \sigma_\omega \rho_{\omega\eta} \lambda_{it(2)}(\alpha_{it})$$

ϕ: PDF for standard normal distribution, Φ: CDF for standard normal distribution,
F_η: CDF for logistic distribution, F_ω: CDF for normal distribution

– Step 1.3. Using labor market *stayers only*, estimate the income profile slope $\{\lambda_0\}$. For notation, I add a tilde to human capital stock permanent shock $\omega_{it(2)}$ to make it clear this is a *selected* random variable because the residual can be computed only when the sample member worked in year t and $t-1$.
The first difference of residual is

$$s_{it(2)} - s_{it-1(2)} = \underbrace{-\left(\overline{logh_{lt(2)}} - \overline{logh_{lt-1(2)}}\right)}_{\text{labor market - year fixed effect}} + \lambda_0$$

[47] Not the level equation. This IV may be an invalid IV for the level equation if it is correlated with the human capital stock, for instance.

[48] Fernández et al. (2004) finds that whether the mother has worked in the past has a strong effect on whether her daughter-in-law works.

$$+\tilde{\omega}_{it(2)} + logm_{it(2)} - logm_{it-1(2)}$$

I control for the current and lagged control functions $\lambda_{it(2)}(\alpha_{it}), \lambda_{it-1(2)}(\alpha_{it-1})$ to account for selection in $\tilde{\omega}_{it(2)}$.

– Step 2. Estimate the variance of permanent income shock σ^2_ω, the Gaussian copula parameter $\rho_{\omega\eta}$, and the variance of the measurement error σ^2_m.

  * Given the estimate for the income profile slope λ_0 and regression coefficients for the control functions, compute the new residual $r_{it(2)}$,

$$\begin{aligned} r_{it(2)} &\equiv (s_{it(2)} - s_{it-1(2)}) + (\widehat{\overline{logh_{lt(2)}} - \overline{logh_{lt-1(2)}}}) - \widehat{\lambda}_0 \\ &= \tilde{\omega}_{it(2)} + logm_{it(2)} - logm_{it-1(2)} \end{aligned} \tag{C.1}$$

, where $-(\widehat{\overline{logh_{lt(2)}} - \overline{logh_{lt-1(2)}}})$ is the average of $(s_{it(2)} - s_{it-1(2)}) - \widehat{\lambda}_0 - \widehat{E}\left[\tilde{\omega}_{it(2)}\right]$ within each labor market and year.

$$-(\widehat{\overline{logh_{lt(2)}} - \overline{logh_{lt-1(2)}}}) = \overline{(s_{it(2)} - s_{it-1(2)}) - \widehat{\lambda}_0 - \widehat{E}\left[\tilde{\omega}_{it(2)}\right]} \tag{C.2}$$

$$\tag{C.3}$$

  * Run nonlinear least squares regression to estimate $\sigma^2_\omega, \rho_{\omega\eta}, \sigma^2_m$ using the following three-moment conditions. The derivation is presented below.

$$\begin{aligned} E\left[r_{it(2)} | P_{it} = 1, P_{it-1} = 1\right] &= \sigma_\omega \rho_{\omega\eta} E\left[\lambda(\alpha_{it})\right] \\ E\left[r^2_{it(2)} | P_{it} = 1, P_{it-1} = 1\right] &= \sigma^2_\omega \left(1 + \rho^2_{\omega\eta} E\left[\alpha_{it}\lambda(\alpha_{it})\right]\right) + 2\sigma^2_m \\ E\left[r_{it(2)}(r_{it-1(2)} + r_{it(2)} + r_{it+1(2)}) | P_{it+q} = 1, q = -2, -1, 0, 1\right] & \\ = \sigma^2_\omega \left(1 + \rho^2_{\omega\eta} E\left[\alpha_{it}\lambda(\alpha_{it})\right]\right) & \end{aligned}$$

Derivation for Moment Conditions Used for Income Process Estimation

Here, I derive the moment conditions used for income process estimation. First, I show the formula for the control function. Remember the following shock distribution assumptions were made.

$$\begin{aligned} \omega_{it(2)} &\sim N(0, \sigma^2_\omega) \\ P(P_{it} = 1) &= P(X_{it}\beta + \eta_{it(2)} > 0) \\ F(\eta_{it(2)}) &= \frac{1}{1 + exp(-\eta_{it(2)})} \end{aligned}$$

$$(\Phi^{-1}(F(\omega_{it(2)})), \Phi^{-1}(F(\eta_{it(2)})) \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{\omega\eta} \\ \rho_{\omega\eta} & 1 \end{pmatrix}\right)$$

$F(.)$ is a CDF of each random variable, $\Phi^{-1}(.)$ is an inverse of a standard normal CDF.

Selection term :

$$\begin{aligned}
E\left[\omega_{it(2)}|\eta_{it(2)} > \alpha_{it}\right] &= E\left[\omega_{it(2)}|F(\eta_{it(2)}) > F(\alpha_{it})\right] \quad (\because \text{F(.) is 1-1, monotonic}) \\
&= E\left[\omega_{it(2)}|\Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it}))\right] \quad (\because \Phi^{-1}(.) \text{ is 1-1, monotonic}) \\
&= E\left[E\left[\omega_{it(2)}|\Phi^{-1}(F(\omega_{it(2)}))\right]|\Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it}))\right] \quad (\because \text{L.I.E.}) \\
&= E\left[\sigma_\omega \Phi^{-1}(F(\omega^F_{it(2)}))|\Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it}))\right] \quad (\because \text{Normality}) \\
&= \sigma_\omega E\left[E\left[\Phi^{-1}(F(\omega_{it(2)}))|\Phi^{-1}(F(\eta_{it(2)}))\right]|\Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it}))\right] \quad (\because \text{L.I.E}) \\
&= \sigma_\omega E\left[\rho_{\omega\eta}\Phi^{-1}(F(\eta_{it(2)}))|\Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it}))\right] \quad (\because \text{bivariate normality}) \\
&= \sigma_\omega \rho_{\omega\eta} \frac{\phi(\Phi^{-1}(F(\alpha_{it})))}{1 - \Phi(\Phi^{-1}(F(\alpha_{it})))} \quad (\because \text{truncated normal}) \\
&= \sigma_\omega \rho_{\omega\eta} \frac{\phi(\Phi^{-1}(F(\alpha_{it})))}{1 - F(\alpha_{it})}
\end{aligned}$$

I define the truncated human capital shock $\tilde{\omega}_{it(2)}$ conditional on that female income is observed. Remember the residual used for the moment condition is

$$r_{it(2)} = \tilde{\omega}_{it(2)} + logm_{it(2)} - logm_{it-1(2)}$$

and $logm_{it(2)} \underset{i.i.d.}{\sim} (0, \sigma^2_m)$.
I write the female work equation, $P(P^F_{it(2)} = 1) = P(\eta_{it(2)} > \alpha_{it(2)})$, and the new inverse mill's ratio $\lambda(\alpha) \equiv \frac{\phi(\Phi^{-1}(F(\alpha)))}{1-F(\alpha)}$.

$$\begin{aligned}
&E\left[r_{it(2)}|P_{it(2)} = 1, P_{it-1(2)} = 1\right] \\
&= E\left[\tilde{\omega}_{it(2)} + logm_{it(2)} - logm_{it-1(2)}\right] \\
&= E\left[E\left[\omega_{it(2)}|\eta_{it(2)} > \alpha_{it(2)}\right]\right] \\
&= \sigma_\omega \rho_{\omega\eta} E\left[\lambda(\alpha_{it(2)})\right] \\
&E\left[r^2_{it(2)}|P_{it(2)} = 1, P_{it-1(2)} = 1\right] \\
&= E\left[\tilde{\omega}^2_{it(2)} + logm^2_{it(2)} + logm^2_{it-1(2)}\right] \\
&= E\left[\omega^2_{it(2)}|\eta_{it(2)} > \alpha_{it(2)}\right] + 2\sigma^2_m \\
&E\left[r_{it(2)}(r_{it-1(2)} + r_{it(2)} + r_{it+1(2)})|P_{it+q(2)} = 1, q = -2, -1, 0, 1\right] \\
&= E\left[\tilde{\omega}_{it(2)}(\tilde{\omega}_{it-1(2)} + \tilde{\omega}_{it(2)} + \tilde{\omega}_{it+1(2)})\right]
\end{aligned}$$

$$= E\left[\omega^2_{it(2)} | \eta_{it(2)} > \alpha_{it(2)}\right]$$

$$\begin{aligned}
E\left[\omega^2_{it(2)} | \eta_{it(2)} > \alpha_{it(2)}\right] &= E\left[\omega^2_{it(2)} | \Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it(2)}))\right] \\
&= E\left[E\left[\omega^2_{it(2)} | \Phi^{-1}(F(\omega_{it(2)}))\right] | \Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it(2)}))\right] \\
&= \sigma^2_\omega E\left[\Phi^{-1}(F(\omega_{it(2)}))^2 | \Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it(2)}))\right] \\
&= \sigma^2_\omega E\left[E\left[\Phi^{-1}(F(\omega_{it(2)}))^2 | \Phi^{-1}(F(\eta_{it(2)}))\right] | \Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it(2)}))\right] \\
&= \sigma^2_\omega E\Bigg[Var\left(\Phi^{-1}(F(\omega_{it(2)})) | \Phi^{-1}(F(\eta_{it(2)}))\right) + \\
&\quad E\left[\Phi^{-1}(F(\omega_{it(2)})) | \Phi^{-1}(F(\eta_{it(2)}))\right]^2 \Big| \Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it(2)}))\Bigg] \\
&= \sigma^2_\omega \left(1 - \rho^2_{\omega\eta} + \rho^2_{\omega\eta} E\left[\Phi^{-1}(F(\eta_{it(2)}))^2 | \Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it(2)}))\right]\right) \\
&= \sigma^2_\omega \left(1 + \rho^2_{\omega\eta} E\left[\alpha_{it(2)} \frac{\phi(\Phi^{-1}(F(\alpha_{it(2)})))}{1 - F(\alpha_{it(2)})}\right]\right) \\
&= \sigma^2_\omega \left(1 + \rho^2_{\omega\eta} E\left[\alpha_{it(2)} \lambda(\alpha_{it(2)})\right]\right)
\end{aligned}$$

To sum up, the moment conditions to use for the female income process are

$$\begin{aligned}
E\left[r_{it(2)} | P_{it(2)} = 1, P_{it-1(2)} = 1\right] &= \sigma_\omega \rho_{\omega\eta} E\left[\lambda(\alpha_{it(2)})\right] \\
E\left[r^2_{it(2)} | P_{it(2)} = 1, P_{it-1(2)} = 1\right] &= \sigma^2_\omega \left(1 + \rho^2_{\omega\eta} E\left[\alpha_{it(2)} \lambda(\alpha_{it(2)})\right]\right) + 2\sigma^2_m \\
E\left[r_{it(2)} (r_{it-1(2)} + r_{it(2)} + r_{it+1(2)}) | P_{it+q(2)} = 1, q = -2, -1, 0, 1\right] & \\
= \sigma^2_\omega \left(1 + \rho^2_{\omega\eta} E\left[\alpha_{it} \lambda(\alpha_{it(2)})\right]\right). &
\end{aligned}$$

C.2 Step 2. Production Function, Skill Price, TFP in local labor market

In this section, I explain how to estimate the skill price, total factor productivity in each local labor market, and the production function curvature parameter. First, using only labor market stayers, I estimate the skill price change over time for each labor market. At this stage, I normalize the log skill price in the first year to zero, following the idea of Heckman et al. (1998). Second, using the labor market switchers, I recover the skill price level differences across the labor markets. The basic idea is that the residual change in income after accounting for the wage change following the individual income profile must come from skill price differences in different labor markets (See the wage equation 12). After computing the final skill price, I renormalize the lowest log skill price to zero.

The key identification assumption here is that there is no individual-labor market-specific match component in income. This assumption would be violated if individuals climb up a job

ladder by receiving better wage offers from different labor markets. This paper is abstracting from this alternative mechanism, but instead, to minimize the potential bias in the estimates if this alternative mechanism were true, I clean the data in the following way. First, I allow for a nonparametric individual-specific income growth rate $f(i)$.[49] Second, to remove the work hour differences right before/after the labor market move, I use the derived annual income, which is the gross weekly payment in the reference week (excluding overtime payment) multiplied by a common factor of 52. 52 was the median of the ratios dividing annual gross pay by gross weekly payment (excluding overtime payment). The income variable was deflated by the CPI index, whose base year was 2015.

I use ASHE data between year 2003 and 2018 because ASHE data includes low-level geocode only after 2002. I first estimate the skill price and TFP estimates from the years 2003 and 2018 and extrapolate the series between the years 2001 and 2002 using a quartic polynomial.

- Step 1. Estimate the (log) skill price $logR_{lt}$ in the local labor markets
  - Step 1.1. Estimate the gender-specific income growth rate $\{\lambda_0^g\}, g \in \{M, F\}$, and individual-specific income growth rate $f(i)$. First, estimate the gender-specific income growth rate, $\{\lambda_0^g\}$, following Step 1 in Section C.1. I assume there is selective participation only for females, so I control for the inverse mill's ratio only when estimating λ^F. Next, compute the labor market-year fixed effect as in equation C.2. Then compute the mean of the residual within each person, which is individual-specific income growth rate $f(i)$.

$$\widehat{f}(i) = \sum_{t=1}^{T} \frac{s_{it(2)} - s_{it-1(2)} + (\overline{logh_{lt(2)}} - \overline{logh_{lt-1(2)}}) - \lambda_0^g}{T}$$

  - Step 1.2. Estimate the (log) skill price change in each labor market over time,$\{(logR_{l(j_{it})t} - logR_{l(j_{it})t-1})\}$ given the estimate $\{\lambda_0^g, f(i)\}$ using *labor market stayers only.*

$$\Delta logw_{it(2)} = (logR_{l(j_{it})t} - logR_{l(j_{it})t-1}) + \underbrace{(logR_{l(j_{it})t-1} - logR_{l(j_{it-1})t-1})}_{\text{zero for labor market stayers}} + \lambda_0^g + f(i) + \omega_{it} + logm_{it} - logm_{it-1} \tag{C.4}$$

  - Step 1.3. Given the sequence $\{(logR_{l(j_{it})t} - logR_{l(j_{it})t-1})\}$, estimate the initial skill price differences across labor markets for the labor market switchers. Only $(L-1)$ (L = number of labor markets) parameters need to be estimated. For this, use the

[49] The ASHE does not include detailed demographic information, such as work experience, immigration status, and years in the UK, so it is impossible to estimate a growth rate as a function of these demographic variables. Instead, I estimate individual-specific growth fixed effect $f(i)$ to capture heterogeneity in income growth rate.

following relationship.

$$logR_{l(j_{it})t-1} - logR_{l(j_{it-1})t-1} = (logR_{l(j_{it})0} - logR_{l(j_{it-1})0}) \tag{C.5}$$
$$+ \underbrace{\sum_{q=2}^{t-1} \left(logR_{l(j_{it})q} - logR_{l(j_{it})q-1}\right)}_{\text{estimated in step 1.2}} - \underbrace{\sum_{q=2}^{t-1} \left(logR_{l(j_{it-1})q} - logR_{l(j_{it-1})q-1}\right)}_{\text{estimated in step 1.2}}$$

In the initial periods, with L labor markets, only L-1 level differences need to be recovered since

$$logR_{l(j_{it})0} - logR_{l(j_{it-1})0} = (logR_{l(j_{it})0} - logR_{l_00}) - (logR_{l(j_{it-1})0} - logR_{l_00})$$
$$\text{, where } l_0 \text{ is a base labor market.}$$

I run a constrained linear regression using equation C.4 including *both* labor market stayers and switchers with $(L-1)^2+2$ constraints.(L=4) [50]

- Step 1.4. Normalize the sequence $R_{l0}, \cdots, R_{lt}$ using $R_{l_00} = 1$ for one labor market $l_0 \in \mathscr{L}$. (Heckman et al. (1998)).

- Step 2. Estimate the production curvature parameter ζ

  - Step 2.1. Estimate the aggregate human capital stock for each labor market at each time t, $logH_{lt(2)}$, using the following equation, given the estimates for $logR_{lt}$ from Step 1. The last term is the summation of the measurement errors in the wage reports, which converges to zero by the law of large numbers and the assumption that the mean of the (log of) measurement errors is zero.

$$log(\text{Total Wage Bill})_{lt} = logR_{lt} + logH_{lt(2)} + \underbrace{\sum_{i \in l} logm_{it(2)}}_{\approx 0}$$
$$\Leftrightarrow logH_{lt(2)} \approx log(\text{Total Wage Bill})_{lt} - logR_{lt}$$

  - Step 2.2. Estimate the production function curvature parameter ζ by running a two-stage least square regression using a Bartik instrument (Bartik (1991)) for $logH_{lt}$. I use the industry composition in each local labor market in the base year (2003) and national industry employment growth to construct the predicted employment in each local labor market. Specifically, the Bartik instrument was

[50] I use $(L-1)^2+2$ constraints. $(L-1)^2$ to recover the $(L-1)$ level differences only and 2 more constraints to ensure that the coefficients in front of the last two terms in equation C.5 are equal to one.

constructed using the following equation.

$$Z_{lt} = \sum_{f} \lambda^{0}_{l(j)f} * M_{ft}$$

where $\lambda^{0}_{l(j)f}$ is the employment share of labor market $l(j)$ for industry f national employment in the base year, M_{ft} is the total employment in industry f at time t. The base year was 2003.[51] The regression function was derived from the below relationship. To approximate the TFP A_{lt}, I control for a year and labor market fixed effects in addition to $logH_{lt}$ in the regression.

$$logR_{lt} = logA_{lt} + (\zeta - 1)logH_{lt} + log\ \zeta \tag{C.6}$$

$logH_{l(j)t}$ is instrumented with the Bartik instrument $Z_{l(j)t}$.

- Step 3. Estimate the log TFP, $logA_{lt}$, given the estimated skill price $logR_{lt}$, total human capital stock $logH_{lt}$, and production curvature parameter ζ using the equation,

$$logA_{lt} = logR_{lt} - log\ \zeta - (\zeta - 1)logH_{lt}$$

The estimate table F.1 is in the Appendix.

C.3 Step 3-1. Estimating the reduced-form CCPs

In this section, I explain how I estimate the reduced-form CCPs, $p_{j't}, p_{p't}, p_{u't}$, which will be used to estimate the structural parameters in the flow payoff in the next step. Considering that the estimation is done in micro panel datasets with limited sample size, I approximate the conditional choice probabilities using the parametric conditional logit model (McFadden et al. (1973)), logit model, and fractional multinomial logit model (Papke and Wooldridge (1996)).

Before estimating reduced-form CCPs, I constructed a neighborhood choice set of size 31, following the suggestion of Keane and Wasi (2012) to handle the large choice set. As Keane and Wasi (2012) explain, in a finite sample, using a random subset of full-choice alternatives gives more precise estimates when there are too many choice alternatives compared to the sample size. In this paper, the full set of neighborhoods is 34378, which is too large compared to the sample size. The number of alternatives, 31, was chosen to make the estimation feasible but still large enough to represent all neighborhoods. The constructed choice set is large enough to ensure that every neighborhood is included as a potential destination for another

[51] Geocode necessary to construct local labor market variable was available only after year 2003 in the dataset.

neighborhood so that there is no neighborhood that experiences only out-migration.

To guarantee the representativeness of the choice set, the choice set was further stratified to include 10 Muslim neighborhoods, 10 White British neighborhoods, and 10 non-Muslim, non-White British neighborhoods. The last neighborhood in the choice set is an origin neighborhood, which is always included as a choice option. Muslim neighborhoods are defined as neighborhoods whose Muslim resident share exceeds 20%.[52] The White British neighborhoods are defined as neighborhoods whose White British resident share exceeds 80.5%.[53] During the estimation, if the destination different from the current neighborhood is not included in the constructed random choice set, I replace one neighborhood belonging to the same stratum as the destination neighborhood with the destination neighborhood.

I estimate all CCPs separately for White British and non-White British samples. To estimate the neighborhood CCP, I use the interaction between the neighborhood-specific characteristics and (one vector for a constant and) the household characteristics. The neighborhood-specific characteristics include the following 18 variables: log skill price in the local labor market $log(R_{jt})$, log median house price in the neighborhood $log(hp_{jt})$, log[54] of neighborhood White British density $log(WHB_{jt}+0.1)$, log of neighborhood Muslim density $log(MUS_{jt}+0.1)$, (minus of) neighborhood deprivation index (IMD, Abel et al. (2016)), $-log(IMD_j)$, log of neighborhood college graduate share $log(ColEd_{jt}+0.1)$, neighborhood housing quality index $hsqual_j$[55], control function $controlfun_{jt}$, log distance between origin and destination $log(distance(j, j_{it(1)})+1)$, whether the neighborhood is different from the origin neighborhood $1(j \neq j_{it(1)})$, and the interaction between whether the neighborhood is different from the origin and neighborhood characteristics $1(j \neq j_{it(1)}) \times \{log(R_{jt}), log(hp_{jt}), log(WHB_{jt}+0.1), log(MUS_{jt}+0.1), -log(IMD_j), log(ColEd_{jt}+0.1), hsqual_j, controlfun_{jt}\}$. For White British households, the household characteristics include the female log human capital stock $log(h_{it(1)})$, female age Age_{it}, and the time index t. For non-White British households, I add a dummy variable whether the household is a least assimilated social position type $1(u_{it(1)} = \overline{u})$ to a household characteristics vector of White British. So in total, the number of regressors is 72 for White British households and 90 for non-White British households.

The female work CCP is also estimated separately for White British and non-White British sample. For White British households, the regressors are one vector for a constant, log of human capital stock $log(h_{it(2)})$, female age Age_{it}, log of median house price $log(hp_{jt})$, log of skill price in the labor market $log(R_{jt})$, log of neighborhood Muslim density, $log(MUS_{jt}+$

[52] This is a cutoff obtained from the k-means clustering of neighborhoods using Muslim resident share in the year 2011.

[53] This is the national White British share in the year 2011.

[54] I add 0.1 to density to prevent a negative infinity when the density is close to 0.

[55] The construction of the housing quality index is explained in Section D.

0.1), log of neighborhood White British density $log(WHB_{jt}+0.1)$, (minus of) neighborhood deprivation index (IMD), $-log(IMD_j)$, log of neighborhood college graduate share $log(ColEd_{jt}+0.1)$, neighborhood housing quality index $hsqual_j$, control function $controlfun_{jt}$, and time index t. This gives 12 regressors, including a constant. For the non-White British sample, I use the same regressor for the White British sample and its interaction with a dummy variable whether the household is a least assimilated social position type $1(u_{it(1)} = \overline{u})$. So the number of regressors for the non-White British sample is 24, including a constant.

I run a social position type transition regression only for the non-White British sample, as I assume White British is the same as the most assimilated type. I use 13 regressors including a constant: they are one vector, log of female human capital stock $log(h_{it(3)})$, female age Age_{it}, log of median house price in the neighborhood $log(hp_{jt})$, log of skill price in the labor market $log(R_{jt})$, log of neighborhood Muslim density $log(MUS_{jt}+0.1)$, log of neighborhood White British density $log(WHB_{jt}+0.1)$, (minus of) neighborhood deprivation index (IMD), $-log(IMD_j)$, log of neighborhood college graduate share $log(ColEd_{jt}+0.1)$, neighborhood housing quality index $hsqual_j$, control function $controlfun_{jt}$, time index t, and a dummy whether the household is the least assimilated social position type $1(u_{it(3)} = \overline{u})$.

The estimation algorithm is the variation of the EM algorithm. By assumption the White British sample is assumed to be the most assimilated social position type. So, in each iteration, only the social position type probability for the non-White British sample is updated. I elaborate on this updating process below:

1. First, I calibrate the female's initial human capital stock and the subsequent human capital stock sequence as explained in Section 4.4.
2. Similarly, I estimate the initial social position type probability $P^{(0)}(u_{i1} = s)$ for each non-White British household using an EM algorithm and the proxy variables in wave 1. I obtain estimates for the social position proxy likelihoods $L^Y_{it} \equiv l(Y_{it}|u_{it})$ from this estimation, which will be used in the social position type updating rule below. I take the average of the initial social position type probability to set the prior $P^{(0)}(u_{i1} = s)$.
3. I set the initial transition CCP $P^{(0)}(u_{it+1} = s'|u_{it} = s, \Omega_{it})$ close to 1 if $s' = s$ and close to 0 if $s' \neq s$. This completes the initialization of the algorithm.
4. In m-th iteration, I estimate the likelihood for neighborhood choice $l(j_{it}|\Omega_{it})$ treating the unobservable states as known, but weighting by the updated individual social position type probability $P^{(m-1)}(u_{it} = s)$. This can be done by applying a weighted conditional multinomial logit regression. I used a fast multinominal logit regression R package 'mnlogit' (Hasan et al. (2016)), which can handle a large number of regressors.
5. In m-th iteration, I estimate the likelihood for female work choice $l(P^F_{it}|\Omega_{it})$ treating the unobservable states as known, but weighting regression by the updated individual

social position type probability $P^{(m-1)}(u_{it} = s)$. In practice, this can be done by applying a weighted logit regression. Then I obtain the likelihood of female work choice. I used a 'glm2' R package (Marschner (2011)) for this estimation.

6. I update the social position type probability $P^{(m)}(u_{it} = s|\Omega_{it})$ using the following Bayes rule.

$$P^{(m)}(u_{it} = s|\Omega_{it}) = \frac{L_{it}(u_{it} = s)}{\sum_{s'} L_{it}(u_{it} = s')} \tag{C.7}$$

Next, one can update $P^{(m)}(u_{it+1} = s'|u_{it} = s, \Omega_{it})$ by first finding the nonparametric forward probabilities and estimating a weighted parametric/nonparametric regression using the updated individual probability $P^{(m-1)}(u_{it} = s)$. Specifically, the forward probability for individual i of type u to have a type u' in next period is defined as :

$$\frac{l(u_{it+1} = s'|u_{it} = s, \Omega_{it}) \left[\sum_{u_{it+2}=1}^{r} \cdots \sum_{u_{i\mathcal{T}}=1}^{r} l(u_{it+2}|u_{it+1} = s', \Omega_{it+1}) \left(\Pi_{t'=t+2}^{\mathcal{T}} L_{it'}^{u} L_{it'}\right)\right]}{\sum_{u'=1}^{r} l(u_{it+1} = s'|u_{it} = s, \Omega_{it}) \left[\sum_{u_{it+2}=1}^{r} \cdots \sum_{u_{i\mathcal{T}}=1}^{r} l(u_{it+2}|u_{it+1} = s', \Omega_{it+1}) \left(\Pi_{t'=t+2}^{\mathcal{T}} L_{it'}^{u} L_{it'}\right)\right]}$$

where the per-period likelihood $L_{it'}$ and the transition likelihood is defined as follows :

$$\begin{aligned}
\text{per-period likelihood} \quad & L_{it} = L_{it}^{j} L_{it}^{P} L_{it}^{Y} && \text{(C.8)}\\
\text{likelihood from neighborhood choice} \quad & L_{it}^{j} \equiv l(j_{it}|\Omega_{it}) && \text{(C.9)}\\
\text{likelihood from female work choice} \quad & L_{it}^{P} \equiv l(P_{it}|\Omega_{it}) && \text{(C.10)}\\
\text{likelihood from proxies} \quad & L_{it}^{Y} \equiv l(Y_{it}|u_{it}) && \text{(C.11)}\\
\text{transition likelihood} \quad & L_{it}^{u} \equiv l(u_{it+1}|\Omega_{it}) && \text{(C.12)}\\
& && \text{(C.13)}
\end{aligned}$$

7. I repeat from steps 3 to 9 until the overall likelihood converges.

C.4 Step 3-2. Estimating Flow Parameters Using Finite Dependence

With reduced-form CCP estimates on hand, I conduct a maximum likelihood estimation for the structural parameters in the flow payoff. The moment conditions used for the estimation are derived using the two-period finite dependence property.

Box 1. Choice Paths Leading to Two-Periods Finite Dependence

Path 1 Choose different neighborhoods $j_{it(2)} \neq \widetilde{j}_{it(2)}$ at time t.

- $j_{it(2)} \rightarrow P_{it(2)} = 1(Working) \rightarrow u_{it+1(1)} \rightarrow j_{it+1(2)}$
- $\widetilde{j}_{it(2)} \rightarrow P_{it(2)} = 1(Working) \rightarrow u_{it+1(1)} \rightarrow j_{it+1(2)}$

Path 2 Choose different female work choices $P_{it(2)} = 1 \neq \widetilde{P}_{it(2)} = 0$ at time t.

- $P_{it(2)} = 1 \rightarrow u_{it+1(1)} \rightarrow j_{it+1(2)} \rightarrow \widetilde{P}_{it+1(2)} = 0$
- $\widetilde{P}_{it(2)} = 0 \rightarrow u_{it+1(1)} \rightarrow j_{it+1(2)} \rightarrow P_{it+1(2)} = 1$

Path 3 Choose different social position choices $u_{it+1(2)} \neq \widetilde{u}_{it+1(2)}$ at time t.

- $u_{it+1(1)} \rightarrow \tilde{j}_{it+1(2)} \rightarrow P_{it+1(2)} = 1$ (Working) $\rightarrow u_{it+2(1)}$
- $\widetilde{u}_{it+1(1)} \rightarrow \tilde{j}_{it+1(2)} \rightarrow P_{it+1(2)} = 1$ (Working) $\rightarrow u_{it+2(1)}$

Box 1 summarizes the choice paths leading to the two-period finite dependence. First, one choice path that makes an agent choose neighborhood $j_{it(2)}$ at time t and stay in the neighborhood $j_{it+1(2)}$ at time t+1 will lead to the stochastically same state variable distribution as another choice path, in which the agent chooses a different neighborhood $\widetilde{j}_{it(2)}$ at time t and moves to neighborhood $j_{it+1(2)}$ at time t+1, while making the same female work choice $P_{it(2)}$ (assumed working) and social position choice $u_{it+1(1)}$ as the first choice path. Second, one choice path making the agent work at time t, $P_{it(2)} = 1$, and not work at time t+1, $\widetilde{P}_{it+1(2)} = 0$, will lead to the stochastically same state variable distribution as another choice path making the agent not work at time t, $\widetilde{P}_{it(2)} = 0$, but work at time t+1, $P_{it+1(2)} = 1$, while making the same social position choice, $u_{it+1(1)}$, and neighborhood choice, $j_{it+1(2)}$, as the first choice path. [56] Third, a choice path to make agents choose a social position $u_{it+1(1)}$ at time t and for a female spouse to work at time t+1, $P_{it+1(2)} = 1$, and the same social position $u_{it+2(1)}$ at time t+1 will lead to the stochastically same state variable distribution as another choice path with a different social position $\widetilde{u}_{it+1(1)}$ at time t, for a female spouse to not work at time t+1, $P_{it+1(2)} = 0$, and to choose the social position $u_{it+2(1)}$ at time t+1, while making the same neighborhood choice, $\tilde{j}_{it+1(2)}$, as the first choice path. [57] I assume the household moves to a different neighborhood $\tilde{j}_{it+1(2)}$ along the two choice paths[58]. States along the choice paths

[56] No female human capital depreciation is key to having this finite dependence, which also has been assumed in the previous papers for tractability.(Altuğ and Miller (1998), Humphries (2017))

[57] These finite dependence paths are not the only ones. One may find different finite dependence paths for the estimation. Theoretically, the choice of finite dependence paths should not matter.

[58] So this condition also identifies the social distance moving cost parameter C_2. I choose $\tilde{j}_{it+1(2)}$ to be the

are stochastic due to human capital shock. I do numerical integration whenever needed using the discretized human capital shock grid.

To derive the moment conditions, I define symbols for the correction function for the first choice path as $\psi^{(1)}_{pt}, \psi^{(1)}_{ut}, \psi^{(1)}_{jt}$ (female work choice, social position choice, neighborhood choice) and for the second choice path as $\psi^{(2)}_{ut}, \psi^{(2)}_{jt}, \psi^{(2)}_{pt}$ (social position, neighborhood, female work choice) and for the third choice path as $\psi^{(3)}_{jt}, \psi^{(3)}_{pt}, \psi^{(3)}_{ut}$ (neighborhood choice, female work choice, social position choice). Due to the specific distributional assumption made (GEV shock distributions) and Lemma 3 in Arcidiacono and Miller (2011), the correction functions have the following closed-form solutions.

$$p_{j't} \equiv P(j_{it(1)} = j' | \Omega^o_{it(1)}), \quad p_{p't} \equiv P(P_{it(2)} = p' | \Omega^o_{it(2)}) \tag{C.14}$$

$$p_{u't} \equiv P(u_{it+1(1)} = u' | \Omega^o_{it(3)}) \tag{C.15}$$

$$\psi_{j't} \equiv \gamma - log p_{j't}, \quad \psi_{p't} \equiv \sigma_p \gamma - \sigma_p log p_{p't}, \quad \psi_{u't} \equiv \sigma_u \gamma - \sigma_u log p_{u't}. \tag{C.16}$$

$\gamma \approx 0.5772$ is an Euler-Mascheroni constant, σ_p is the scale parameter for logistic shock for female work, σ_u is the scale parameter for Gumbel shock for social position. The scale parameter for the location preference shock is normalized to one.

Using these correction functions, one may obtain the alternative representations for conditional choice probabilities, summarized in Proposition 1. This proposition will be used to estimate the flow payoff parameters.

Proposition 1 (Finite Dependence)**.** *The model has a two-period finite dependence described in Box 1 above. The conditional choice probabilities for each discrete choice have alternative representations, equation C.17 - C.19, and C.25 - C.27, using the finite dependence path in Box 1.*

$$p_{jt} = \frac{exp(\widetilde{v}_{jt(1)}(\Omega_{it(1)}))}{\sum exp(\widetilde{v}_{j't(1)}(\Omega_{it(1)}))} \tag{C.17}$$

$$p_{pt} = \frac{exp(\widetilde{v}_{pt(2)}(\Omega_{it(2)})/\sigma_p)}{\sum exp(\widetilde{v}_{p't(2)}(\Omega_{it(2)})/\sigma_p)} \tag{C.18}$$

$$p_{ut} = \frac{exp(\widetilde{v}_{ut(3)}(\Omega_{it(3)})/\sigma_u)}{\sum exp(\widetilde{v}_{u't(3)}(\Omega_{it(3)})/\sigma_u)} \tag{C.19}$$

where

$$\widetilde{v}_{jt(1)}(\Omega_{it(1)}) = E\Big[\Pi^1_j(\Omega_{it(1)}; j) + \gamma^P_1 log(R_{lt}) + \psi^{(1)}_{pt} \tag{C.20}$$

$$+\psi^{(1)}_{ut} + \beta\{MC(j, j_{it+1(1)}) + \psi^{(1)}_{jt+1}\}\Big] \tag{C.21}$$

$(J-1)$-th neighborhood in the choice set.

$$\widetilde{v}_{pt(2)}(\Omega_{it(2)}) = E\left[\Pi^2_j(\Omega_{it(2)};p) + \psi^{(2)}_{ut} + \beta\{\psi^{(2)}_{jt+1} + \Pi^2_j(\Omega_{it+1(2)};1-p) + \psi^{(2)}_{pt+1}\}\right] \tag{C.22}$$

$$\widetilde{v}_{ut(3)}(\Omega_{it(3)}) = E\Big[\Pi^3_u(\Omega_{it(3)};u) + \beta\{C_2(|u-\overline{u}_{jt+1(1)}|)\times 1(j_{it+1(1)} \neq j_{it(3)})$$
$$+\psi^{(3)}_{jt+1} + \gamma^P_{1u}P^F_{it+1(2)} + \psi^{(3)}_{pt+1} + \Pi^3_u(\Omega_{it+1(3)};u_{it+2(1)}) + \psi^{(3)}_{ut+1}\}\Big] \tag{C.23}$$

Proof. Using Theorem 1, Lemma 3 in Arcidiacono and Miller (2011) and the two-period finite dependence paths, the conditional value functions can be represented as follows:

$$v_{jt(1)}(\Omega_{it(1)}) = E\Big[\Pi^1_j(\Omega_{it(1)};j) + \gamma^P_1 log(R_{lt}) + \psi^{(1)}_{pt}$$
$$+\psi^{(1)}_{ut} + \beta\{MC(j,j_{it+1(1)}) + \psi^{(1)}_{jt+1} + V_{t+1(2)}(\Omega_{it+1(2)})\}\Big]$$
$$v_{pt(2)}(\Omega_{it(2)}) = E\Big[\Pi^2_j(\Omega_{it(2)};p) + \psi^{(2)}_{ut} + \beta\{\psi^{(2)}_{jt+1} + \Pi^2_j(\Omega_{it+1(2)};1-p) + \psi^{(2)}_{pt+1} + V_{t+1(3)}(\Omega_{it+1(3)})\}\Big]$$
$$v_{ut(3)}(\Omega_{it(3)}) = E\Big[\Pi^3_u(\Omega_{it(3)};u) + \beta\{C_2(|u-\overline{u}_{jt+1(1)}|)\times 1(j_{it+1(1)} \neq j_{it(3)})$$
$$+\psi^{(3)}_{jt+1} + \gamma^P_{1u}P^F_{it+1(2)} + \psi^{(3)}_{pt+1} + \Pi^3_u(\Omega_{it+1(3)};u_{it+2(1)}) + \psi^{(3)}_{ut+1} + \beta V_{t+2(1)}(\Omega_{it+2(1)})\}\Big]$$

The expected continuation value terms $E[V_{t+1(2)}(\Omega_{it+1(2)})], E[V_{t+1(3)}(\Omega_{it+1(3)})], E[V_{t+2(1)}(\Omega_{it+2(1)})]$ remain the same across different discrete choice paths in Box 1, so they can be dropped when subtracting one conditional value function from another. In an equation for $v_{ut(3)}(\Omega_{it(3)})$, the moving cost depending on the social distance in a term $MC(j_{it(3)}, j_{it+1(1)})$, and the gender norm cost for working female in a term $\Pi^2_j(\Omega_{it+1(2)}; P^F_{it+1(2)} = 1)$ do not cancel out, so they need to be taken into account when comparing flow payoff along two different choice paths. □

I use Proposition 1 to get the value function estimates.

$$\widehat{v}_{jt(1)}(\Omega_{it(1)}) = E\Big[\widehat{\Pi}^1_j(\Omega_{it(1)};j) + \widehat{\gamma}^P_1 log(\widehat{R}_{lt}) + \widehat{\psi}^{(1)}_{pt} \tag{C.24}$$
$$+\widehat{\psi}^{(1)}_{ut} + \beta\{\widehat{MC}(j,j_{it+1(1)}) + \widehat{\psi}^{(1)}_{jt+1}\}\Big] \tag{C.25}$$

$$\widehat{v}_{pt(2)}(\Omega_{it(2)}) = E\left[\widehat{\Pi}^2_j(\Omega_{it(2)};p) + \widehat{\psi}^{(2)}_{ut} + \beta\{\widehat{\psi}^{(2)}_{jt+1} + \widehat{\Pi}^2_j(\Omega_{it+1(2)};1-p) + \widehat{\psi}^{(2)}_{pt+1}\}\right] \tag{C.26}$$

$$\widehat{v}_{ut(3)}(\Omega_{it(3)}) = E\Big[\widehat{\Pi}^3_u(\Omega_{it(3)};u) + \beta\{\widehat{C}_2(|u-\widehat{\overline{u}}_{jt+1(1)}|)\times 1(j_{it+1(1)} \neq j_{it(3)})$$
$$+\widehat{\psi}^{(3)}_{jt+1} + \widehat{\gamma}^P_{1u}P^F_{it+1(2)} + \widehat{\psi}^{(3)}_{pt+1} + \widehat{\Pi}^3_u(\Omega_{it+1(3)};u_{it+2(1)}) + \widehat{\psi}^{(3)}_{ut+1}\}\Big] \tag{C.27}$$

To estimate the flow utility $\widehat{\Pi}_j^1(\Omega_{it(1)}; j)$ in the value function expression, I need the estimates for neighborhood's average social position and human capital stock. Given the previous reduced-form CCPs, I estimate these terms using the predicted values from the regression. I regress the social position, $\widehat{u}_{it(s)}$, on the neighborhood characteristics vector, $N_{it(s)} = \{1, log(R_{lt}), log(hp_{jt}), log(WHB_{jt}+0.1), log(MUS_{jt}+0.1), log(-IMD_j), log(ColEd_{jt}), hsqual_j, controlfun_{jt}\}$.

$$\begin{aligned} \text{individual characteristics regression} \quad & u_{it(s)} = N_{it(s)}\beta_1 + \epsilon_{it(s)} \\ \text{neighborhood average prediction} \quad & \widehat{u}_{jt(s)} = N_{it(s)}\widehat{\beta}_1 \end{aligned}$$

To sum up, one can estimate every structural parameter in flow utility using the ML estimation using the usual multinomial/logit formula in proposition 1.

$$p_{jt} = \frac{exp(\hat{v}_{jt(1)}(\Omega_{it(1)}))}{\sum exp(\hat{v}_{j't(1)}(\Omega_{it(1)}))} \tag{C.28}$$

$$p_{pt} = \frac{exp(\hat{v}_{pt(2)}(\Omega_{it(2)})/\sigma_p)}{\sum exp(\hat{v}_{p't(2)}(\Omega_{it(2)})/\sigma_p)} \tag{C.29}$$

$$p_{ut} = \frac{exp(\hat{v}_{ut(3)}(\Omega_{it(3)})/\sigma_u)}{\sum exp(\hat{v}_{u't(3)}(\Omega_{it(3)})/\sigma_u)} \tag{C.30}$$

D Construction of the Control Function

This section explains the details of the construction of the control function used to measure unobserved neighborhood amenities. The instrumental variable used to construct the control function is similar to Bayer et al. (2016), which used non-adjacent but nearby neighborhood housing characteristics. This neighborhood (that belongs to a donut ring drawn around the neighborhood) is likely to be a substitute in the housing market, so their house characteristics would affect the house price of the neighborhood. However, they are non-adjacent, so the unobserved neighborhood amenity would be less likely to be correlated if the donut ring is drawn sufficiently far away from the center of the neighborhood. I use a 5 miles to 10 miles donut ring to construct the instrumental variable[59].

$$\begin{aligned} log(\text{Med. House Price}_{jt}) = \beta_0 + \beta_1 \text{NBHD House Char}_{jt} + \\ \beta_2 \text{NBHD Other Char}_{jt} + \beta_3 \text{Donut House Char}_{jt} + \epsilon_{jt} \end{aligned} \tag{D.1}$$

[59] I have tried a smaller ring, a 3 miles to 6 miles donut ring, and the result was similar.

Table D.1: First-Stage Regression Result for Log of (Median) House Price Using Donut IV

	(1) log(house price)	(2) log(house price)
Donut 5-10 mile: Average Number of Rooms	-0.331***	
	(0.00627)	
Donut 5-10 mile: Share of HH with Central Heating	-0.231***	
	(0.0485)	
Donut 5-10 mile: Share of HH Owned	-13.62***	
	(0.676)	
Donut 5-10 mile: Share of Social Housing	-15.22***	
	(0.679)	
Donut 5-10 mile: Share of HH Privately Rented	-13.53***	
	(0.682)	
Donut 5-10 mile: Share of HH Rent Free	-19.98***	
	(0.671)	
Donut 5-10 mile: Share of House Detached	-0.173***	
	(0.0354)	
Donut 5-10 mile: Share of House Semi-Detached	-0.650***	
	(0.0246)	
Donut 5-10 mile: Share of House Terraced	-0.676***	
	(0.0316)	
Donut 5-10 mile: Share of Mobile Space	6.781***	
	(0.459)	
Donut 5-10 mile: log(Housing Density per Acre)	-0.0123***	
	(0.00226)	
Donut_IV		-0.551***
		(0.0179)
Partial F-stat	2166.310	953.149
Controls	Yes	Yes
N	618786	618786
R-square	0.815	0.747

Note: Standard errors are clustered at the neighborhood (lower super output area) level and are in parenthesis.
Data years 2001-2019. Annual variables were imputed from interpolation using decennial Census data.
Controls include house characteristics for the NBHD, NBHD deprivation index, college graduate share, NBHD White British and Muslim share, and log of skill price.
$^{*}p < 0.05$, $^{*}p < 0.01$, $^{*}p < 0.001$

Table D.1 shows the first stage regression results from the regression equation D.1. I regress the log of median house price for a neighborhood j in year t onto the neighborhood j house characteristics in year t, other neighborhood j characteristics in year t, and the house characteristics of the donut ring drawn around neighborhood j in year t. The house characteristics vector (for the neighborhood and for the donut ring area surrounding the neighborhood) includes the average number of rooms, the share of households with central heating, the share of homeowners, the share of social housing, the share of rental households, the share of rent-free households, the share of detached houses, the share of semi-detached houses, the share of terraced houses, the share of mobile spaces, log of housing density. The other neighborhood characteristics include (minus of) neighborhood deprivation index, college graduate share, White British share, Muslim share, and log of labor market skill price. The Donut House Char$_{jt}$ vector includes many variables, so I collapse it into a $Donut_IV$ summary index, which is computed as $\widehat{\beta}_1 \times$ Donut House Char$_{jt}$. This is a predicted house price in the donut ring based on housing characteristics. The partial F-statistics for instrumental variables are very large, far greater than 900, which shows that the instrumental variables are strong.

Next, I construct the residual from the first stage regression, which is a control function for the housing price. The interpretation is that this residual captures the unobserved amenity of the neighborhood. The structural model utilizes a summary index of the average housing quality of the neighborhood, which is constructed as the average of $\widehat{\beta}_1 \times$ NBHD House Char$_{jt}$ for 19 years, so I run a slightly modified regression equation D.2 to compute the control function residual, $controlfun_{jt} = \widehat{\epsilon}_{jt}$.

$$\begin{aligned} log(\text{Med. House Price}_{jt}) = \beta_0 + \beta_1 \text{Avg Housing Quality}_j + \\ \beta_2 \text{NBHD Other Char}_{jt} + \beta_3 \text{Donut IV}_{jt} + \epsilon_{jt} \end{aligned} \tag{D.2}$$

E Additional Tables and Figures

Table E.1: Missing Information on Year to Current Address

Wave	Survey Year	Average Year Age Gap Between Surveys	Whether Moved NBHD	Missing Year to Current Address
2	2003-2005	2.334	0.299	0.135
3	2006-2007	2.360	0.215	0.377
4	2008-2009	2.130	0.158	0.499
5	2012-2013	4.279	0.220	0.194
6	2015-2016	3.288	0.157	0.147
7	2018-2019	3.243	0.114	0.216

Source: MCS Wave 1-7

Note: This table shows the missing rates for the information on years to the current address in MCS data wave 1-7. The missing rate for wave 3 and 4 is particularly high, probably due to some changes in the survey implementation procedure for these two waves. Our estimation sample size for the ethnic minority sample is already very small, so to minimize the sample loss, I impute the year to the current address if the information is missing by randomly drawing a year between the last and the current survey year if the household has moved to a new address. The average year gap between surveys is not large, only 2-4 years on average, so this imputation would create only 1-3 year difference from the actual moving year.

Table E.2: Association Between Economic and Cultural Assimilation – Additional Results

	Dependent Variable Male Log Labor Income	
(Male) Non-Western Religiosity Index	-0.0330	
	(0.0253)	
(Male) Foreign Language Usage Index		0.00341
		(0.0154)
Sample	NWHB	NWHB
Age Pol	Yes	Yes
Individual FE	Yes	Yes
Year FE	Yes	Yes
Labor Market FE	Yes	Yes
N	3149	5161
R-square	0.040	0.026

Note: Standard errors are clustered at the household level and are in parentheses. In every regression, a quadratic age polynomial, individual fixed effect, year fixed effect, and labor market fixed effect are controlled. The unit for labor income is 2015 £. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

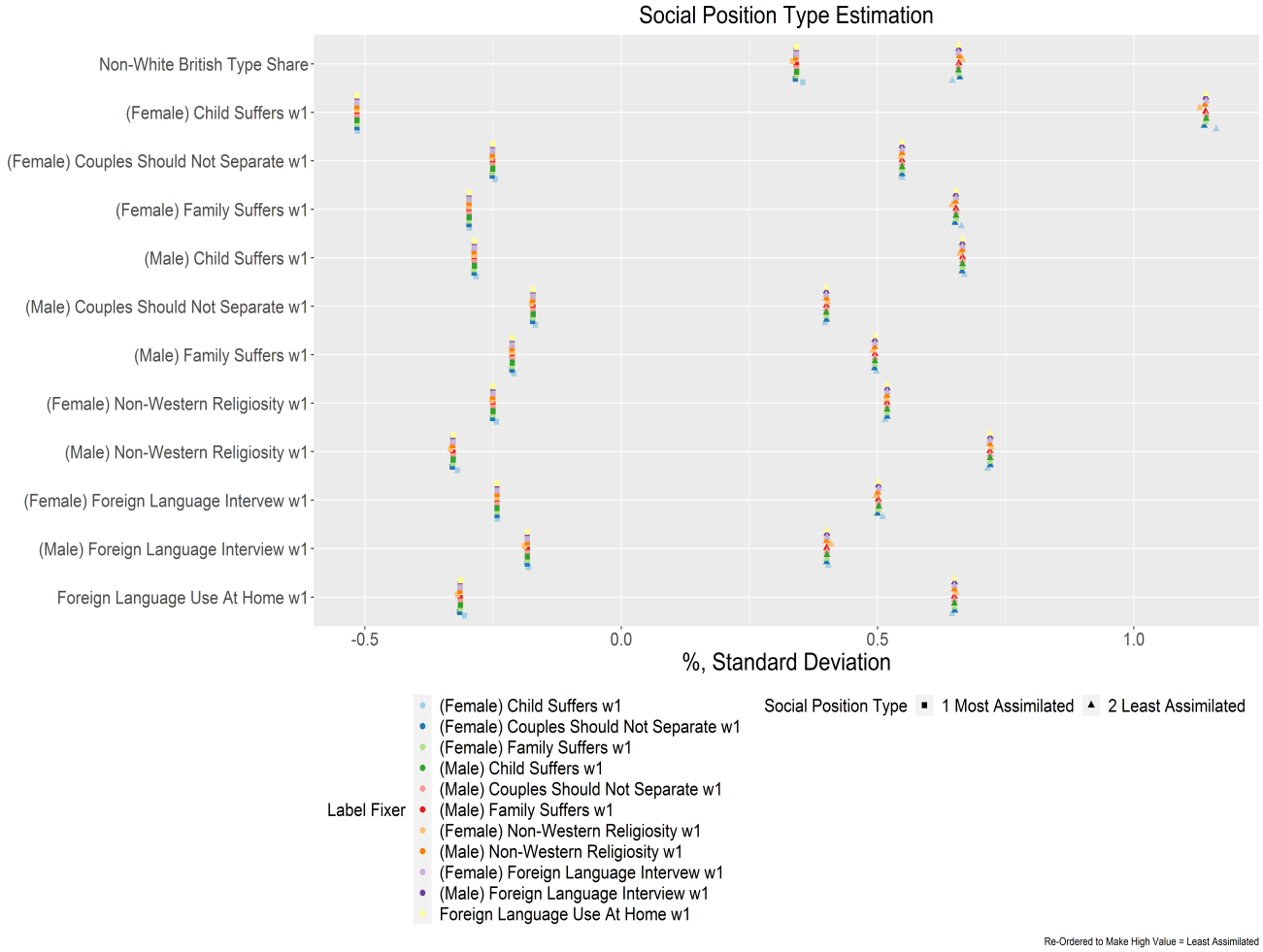
Table E.3: Main Parent (Male) Characteristics in the First Interview

Male Characteristics	(by Female Spouse Ethnicity) White British	Non-White British
White British	0.961	0.135
	(0.193)	(0.342)
Age	34.777	34.61
	(5.202)	(5.827)
Muslim		0.546
		(0.498)
Not an Immigrant		0.134
		(0.341)
First Gen Immigrant		0.644
		(0.479)
Second Gen Immigrant		0.222
		(0.416)
Years in the UK (1st Gen)		14.322
		(9.914)
College Graduation	0.429	0.412
	(0.495)	(0.492)
Employed	0.96	0.877
	(0.197)	(0.328)
Sample Size	1234	1234

Source: MCS wave 1-7

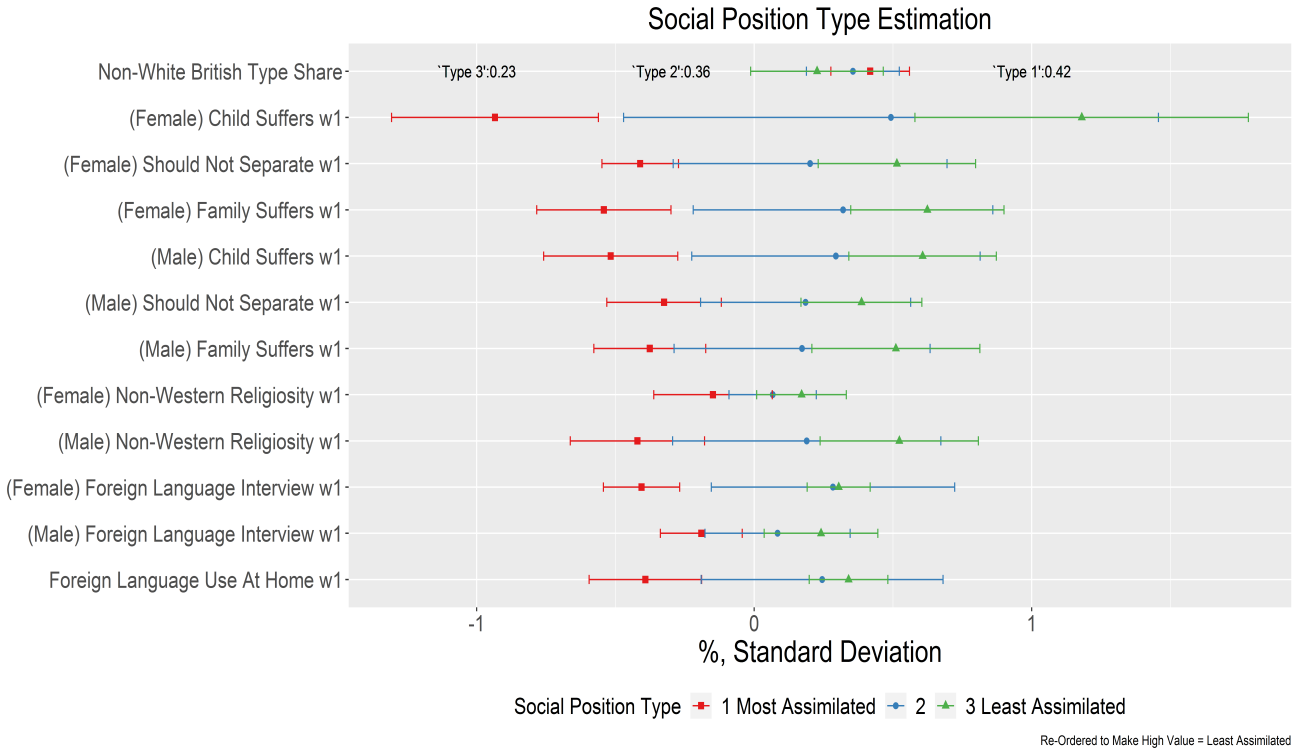
Note: This table shows the average and standard deviations of male sample characteristics in the first interview. The standard deviation of each variable is in parentheses. By construction, White British sample does not include any Muslims. I do not report the immigration-related statistics for White British sample, as they are small in number.

Figure E.1: Robustness of Initial Social Position Type Estimates by Choice of a Different Anchoring Variable (Assuming Two Types)



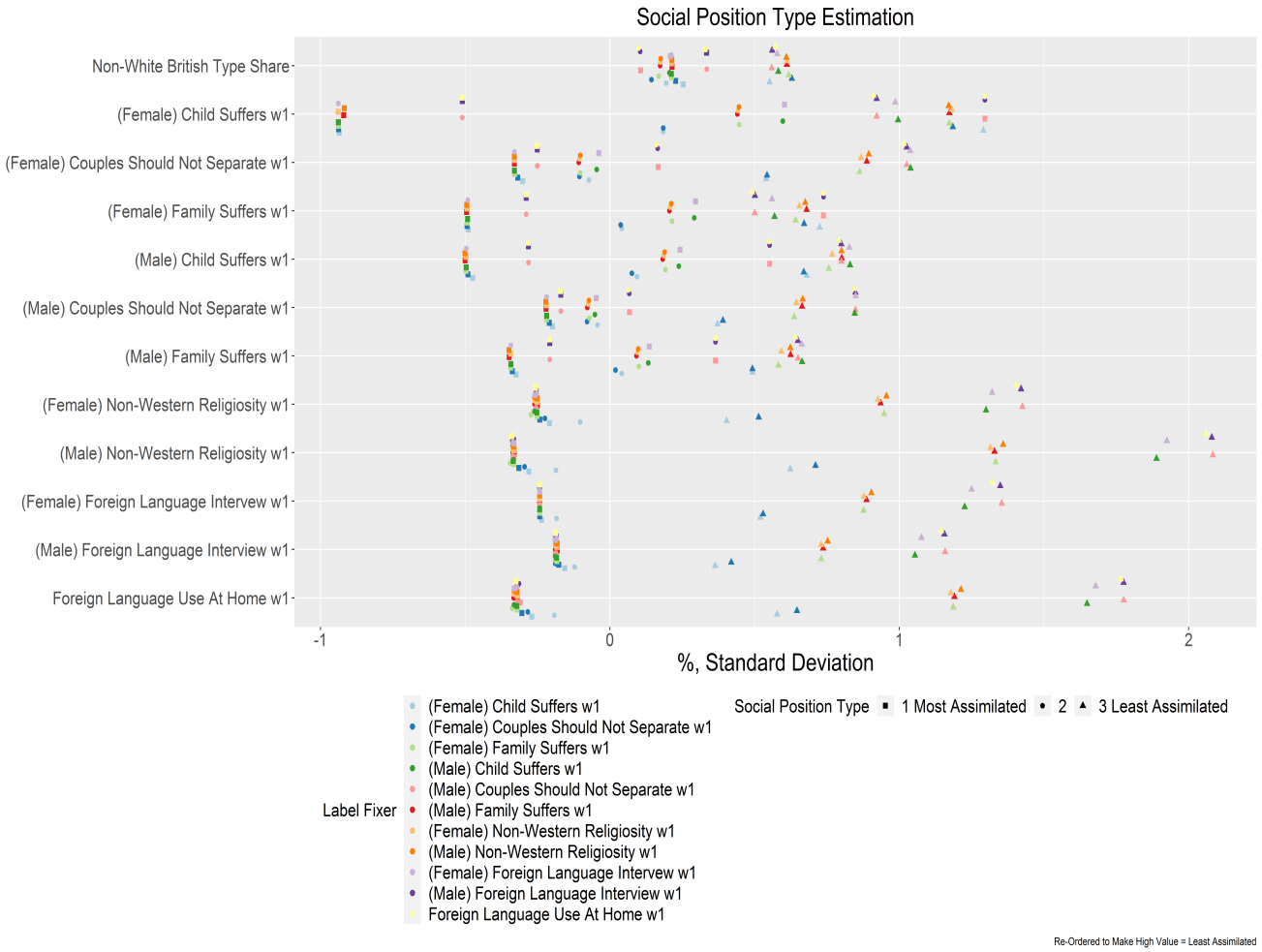
Source: MCS wave 1
Note: This figure shows the robustness of initial social position type estimates by different choices of an anchoring variable. I assumed two social position types.

Figure E.2: Initial Social Position Type Estimates (Assuming Three Types)



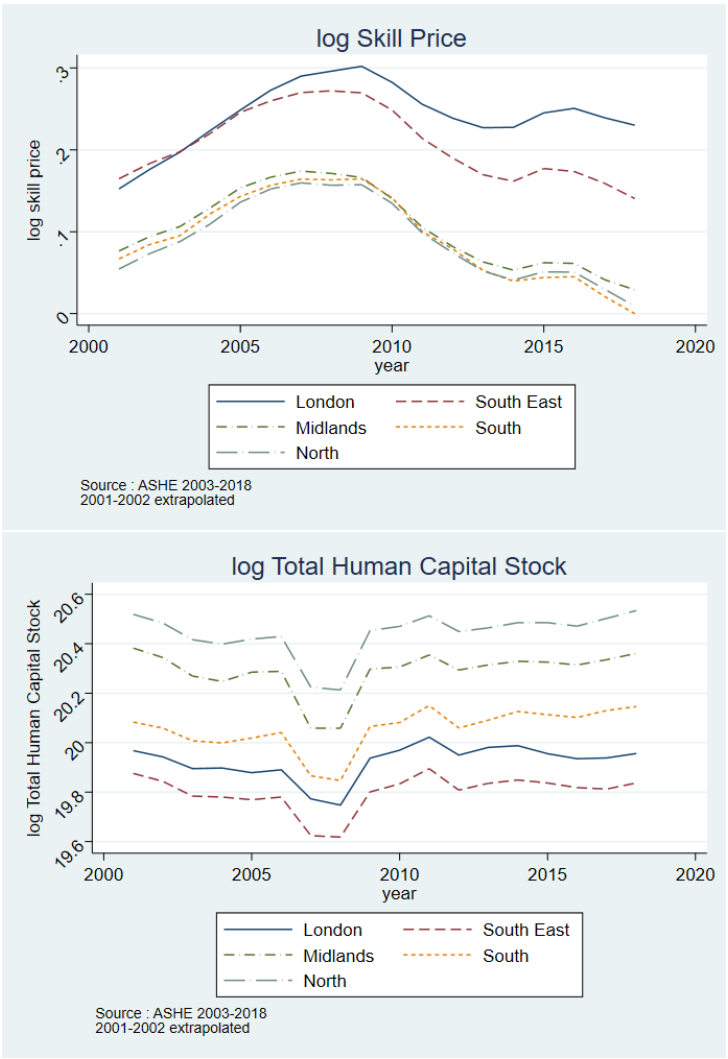
Source: MCS wave 1
Note: This figure shows the initial social position type share estimates and the average survey responses of each type, assuming three social position types. Whiskers show the 95% confidence intervals which were computed from bootstrapping 100 times.

Figure E.3: Robustness of Initial Social Position Type Estimates by Choice of a Different Anchoring Variable (Assuming Three Types)



Source: MCS wave 1
Note: This figure shows the robustness of initial social position type estimates by diffferent choice of an anchoring variable. I assumed three social position types.

Figure E.4: Log Skill Price and Log Total Human Capital Stock Estimates



Source: Annual Survey of Hours and Earnings, 2003-2018
Note: This figure shows the estimated log skill price and log total human capital stock for each local labor market for each year. The data series for the years 2001-2002 were extrapolated using a linear regression fitting the data until the year 2007 before the Great Recession.

F Proxy Measurement Structure Matrix Estimates

In this section, I report the estimates for the proxy variables measurement system. Every proxy variable is discrete, so most measurement equations can be represented in the form of a probability matrix, $M^Y_{lk} = [P(Y = k|U = l)]$, where lk-th element shows the probability of having a response k from proxy variable Y when the true latent variable has a value l. These matrices were estimated using the EM algorithm.

F.1 Social Position

I present similar measurement structure matrices, whose lk-th element is $M^Y_{lk} = [P(Y = k|U = l)]$, for proxy variables used for social position estimation. Gender norm-related variables have five response scales. Non-Western religiosity proxy variables have four response scales. Whether the interview was done in a foreign language is binary, and the Foreign Language Usage at Home proxy has three response scales. I assumed two social position types. I wrote MCS variable names inside the parentheses. The interview language variable has used different variable names throughout the survey. The probability estimates are shown without parentheses, and the standard error estimates are shown inside the parenthesis.

Female Gender Attitude Child Suffers ('chsu') Wave 1

$$M_{chsu} = \begin{bmatrix} 0.142 & 0.437 & 0.420 & 0.000 & 0.000 \\ (0.017) & (0.021) & (0.023) & (0.006) & (0.000) \\ 0.000 & 0.000 & 0.000 & 0.632 & 0.368 \\ (0.001) & (0.005) & (0.000) & (0.022) & (0.023) \end{bmatrix} \tag{F.1}$$

Female Gender Attitude Divorse ('nose') Wave 1

$$M_{nose} = \begin{bmatrix} 0.039 & 0.164 & 0.324 & 0.266 & 0.207 \\ (0.01) & (0.017) & (0.023) & (0.022) & (0.019) \\ 0.020 & 0.058 & 0.083 & 0.353 & 0.486 \\ (0.007) & (0.01) & (0.012) & (0.023) & (0.023) \end{bmatrix} \tag{F.2}$$

Female Gender Attitude Family Suffers ('fasu') Wave 1

$$M_{fasu} = \begin{bmatrix} 0.101 & 0.296 & 0.321 & 0.242 & 0.041 \\ (0.014) & (0.023) & (0.019) & (0.021) & (0.011) \\ 0.014 & 0.078 & 0.153 & 0.48 & 0.274 \\ (0.004) & (0.013) & (0.016) & (0.021) & (0.019) \end{bmatrix} \tag{F.3}$$

Male Gender Attitude Child Suffers ('chsu') Wave 1

$$M_{chsu} = \begin{bmatrix} 0.087 & 0.304 & 0.225 & 0.291 & 0.093 \\ (0.017) & (0.028) & (0.021) & (0.025) & (0.018) \\ 0.013 & 0.084 & 0.106 & 0.459 & 0.338 \\ (0.006) & (0.014) & (0.014) & (0.025) & (0.023) \end{bmatrix} \tag{F.4}$$

Male Gender Attitude Divorce ('nose') Wave 1

$$M_{nose} = \begin{bmatrix} 0.028 & 0.133 & 0.246 & 0.268 & 0.326 \\ (0.009) & (0.019) & (0.021) & (0.023) & (0.026) \\ 0.01 & 0.063 & 0.065 & 0.333 & 0.53 \\ (0.005) & (0.01) & (0.012) & (0.021) & (0.022) \end{bmatrix} \tag{F.5}$$

Male Gender Attitude Family Suffers ('fasu') Wave 1

$$M_{fasu} = \begin{bmatrix} 0.064 & 0.256 & 0.278 & 0.308 & 0.094 \\ (0.015) & (0.024) & (0.024) & (0.024) & (0.016) \\ 0.018 & 0.106 & 0.174 & 0.433 & 0.269 \\ (0.008) & (0.016) & (0.019) & (0.024) & (0.022) \end{bmatrix} \tag{F.6}$$

Female Non-Western Religious Attendance ('rsma') Wave 1

$$M_{rsma} = \begin{bmatrix} 0.701 & 0.099 & 0.087 & 0.113 \\ (0.023) & (0.014) & (0.013) & (0.016) \\ 0.609 & 0.131 & 0.083 & 0.177 \\ (0.017) & (0.015) & (0.010) & (0.016) \end{bmatrix} \tag{F.7}$$

Male Non-Western Religious Attendance ('rsma') Wave 1

$$M_{rsma} = \begin{bmatrix} 0.602 & 0.104 & 0.092 & 0.202 \\ (0.029) & (0.016) & (0.015) & (0.022) \\ 0.285 & 0.118 & 0.075 & 0.522 \\ (0.021) & (0.016) & (0.013) & (0.025) \end{bmatrix} \tag{F.8}$$

Female Interviewed in Foreign Language Wave 1

$$M_{itwfor} = \begin{bmatrix} 0.918 & 0.082 \\ (0.014) & (0.014) \\ 0.646 & 0.354 \\ (0.021) & (0.021) \end{bmatrix} \tag{F.9}$$

Male Interviewed in Foreign Language Wave 1

$$M_{itwfor} = \begin{bmatrix} 0.902 & 0.098 \\ (0.014) & (0.014) \\ 0.789 & 0.211 \\ (0.020) & (0.020) \end{bmatrix} \tag{F.10}$$

Foreign Language Usage at Home ('hlan') Wave 1

$$M_{hlan} = \begin{bmatrix} 0.370 & 0.540 & 0.090 \\ (0.022) & (0.023) & (0.012) \\ 0.115 & 0.641 & 0.244 \\ (0.014) & (0.019) & (0.017) \end{bmatrix} \tag{F.11}$$

F.2 Labor Market Skill Price

Table F.1: Log Skill Price Estimates

		Labor Market				
Year		London	South East	Midlands	South	North
2001	Estimate	0.152	0.165	0.077	0.067	0.055
	SE	0.014	0.014	0.013	0.013	0.013
2002	Estimate	0.176	0.183	0.094	0.084	0.073
	SE	0.013	0.013	0.012	0.012	0.012
2003	Estimate	0.197	0.198	0.106	0.095	0.088
	SE	0.012	0.012	0.012	0.012	0.012
2004	Estimate	0.223	0.220	0.129	0.122	0.109
	SE	0.012	0.011	0.011	0.011	0.011
2005	Estimate	0.249	0.246	0.154	0.143	0.136
	SE	0.011	0.011	0.011	0.010	0.011
2006	Estimate	0.273	0.260	0.167	0.157	0.152
	SE	0.011	0.011	0.010	0.009	0.011
2007	Estimate	0.290	0.270	0.174	0.164	0.160
	SE	0.010	0.010	0.010	0.009	0.010
2008	Estimate	0.296	0.272	0.171	0.163	0.157
	SE	0.010	0.010	0.010	0.008	0.010
2009	Estimate	0.302	0.270	0.166	0.165	0.157
	SE	0.010	0.010	0.010	0.008	0.010
2010	Estimate	0.282	0.249	0.141	0.142	0.135
	SE	0.010	0.010	0.010	0.007	0.010
2011	Estimate	0.256	0.214	0.105	0.100	0.098
	SE	0.009	0.010	0.009	0.006	0.009
2012	Estimate	0.239	0.190	0.082	0.079	0.074
	SE	0.009	0.009	0.009	0.006	0.009
2013	Estimate	0.227	0.170	0.063	0.053	0.052
	SE	0.009	0.009	0.009	0.006	0.009
2014	Estimate	0.228	0.162	0.053	0.040	0.041
	SE	0.009	0.009	0.009	0.005	0.008
2015	Estimate	0.245	0.177	0.062	0.044	0.051
	SE	0.009	0.009	0.009	0.005	0.008
2016	Estimate	0.251	0.174	0.061	0.045	0.051
	SE	0.009	0.009	0.009	0.004	0.008
2017	Estimate	0.239	0.159	0.042	0.021	0.030
	SE	0.009	0.010	0.009	0.004	0.008
2018	Estimate	0.230	0.141	0.029	0.000	0.009
	SE	0.009	0.009	0.009	0.003	0.008

Source: Annual Survey of Hours and Earnings, 2003-2018.
Note: This table shows the estimated log skill price for the five local labor markets, which were classified using k-means clustering and mapped in Figure A.1, between the years 2001 and 2018. The ASHE data before the year 2003 were not used for analysis as they do not provide low-level geographic information. The estimates for the years between 2001 and 2002 were extrapolated using a linear regression that fitted the data until 2007, before the Great Recession.

G Estimates for Counterfactual

In this section, I report auxiliary estimates to simulate population distributions in counterfactual simulations.

G.1 Average Neighborhood Characteristics

Table G.1 report multinomial logit regression coefficients that predict average neighborhood characteristics (social position type 1 share, social position type 2 share). The excluded category in the multinomial logit regression is the White British category.

Table G.1: Average Neighborhood Characteristics Estimates

	NWHB Type 1 Share	NWHB Type 2 Share
Constant	-18.152	-6.228
	(2.706)	(3.178)
Log(White British Density + 0.1)	-2.283	-2.209
	(0.126)	(0.128)
Log (Muslim Density + 0.1)	2.367	3.039
	(0.082)	(0.085)
Log (Median House Price)	1.790	0.730
	(0.208)	(0.246)
Log (Skill Price)	-1.011	1.834
	(0.612)	(0.739)
-Log(Neighborhood Deprivation Index)	0.164	-0.097
	(0.036)	(0.044)
Log (College Graduate Share + 0.1)	-1.598	-1.143
	(0.192)	(0.224)
Housing Quality	-1.298	-1.736
	(0.200)	(0.248)
Control Function	-1.256	-0.294
	(0.213)	(0.248)

Source: Millennium Cohort Study, Wave 1-7
Note: This table reports estimates to predict average neighborhood characteristics in the counterfactual simulations. The columns report the multinomial logit regression coefficients. The excluded category is White British.

G.2 State Variable Distribution Estimates

In this section, I report auxiliary estimates used to conduct counterfactual simulations. To simulate changes in the aggregate state variables, I need estimates for the distribution of various state variables. The first column reports the multinomial logit coefficients. The second and third columns report the linear regression coefficients as well as the standard deviation of the residuals.

Table G.2: Distributional Parameter Estimates for Aggregate Variables

	Least Assimilated Type	Human Capital Stock	
	$f(u\|NC,a)$	$f(h\|u,NC,a)$	
	Non-White British	*White British*	*Non-White British*
Constant	10.064	-0.746	-0.115
	(0.267)	(4.817)	(2.815)
Log (Median House Price)	-0.666	0.750	0.723
	(0.089)	(0.374)	(0.223)
Log (Skill Price)	-0.003	-2.655	-2.187
	(1.093)	(1.285)	(0.737)
Log (WHB Density+0.1)	0.108	0.400	0.102
	(0.171)	(0.364)	(0.045)
Log (Muslim Density+0.1)	0.860	0.074	-0.051
	(0.172)	(0.217)	(0.049)
-Log(Neighborhood Deprivation Index)	-0.210	-0.137	-0.025
	(0.130)	(0.068)	(0.034)
Log (College Grad Share +0.1)	0.677	-0.007	-0.343
	(0.328)	(0.321)	(0.211)
Housing Quality	-0.719	-0.449	-0.587
	(0.596)	(0.358)	(0.213)
Control Function	0.799	-0.667	-0.515
	(0.318)	(0.376)	(0.214)
Age	-0.037	0.015	0.013
	(0.013)	(0.006)	(0.003)
1(Least Assimilated Type)			-0.227
			(0.032)
SD of Human Capital shock		0.918	0.737

Source: Millennium Cohort Study, Wave 1-7

Note: This table shows the population distribution estimates for counterfactual simulation. For social position results, I run multinomial logit regressions. For human capital stock, I assume it follows a normal distribution, and I run an OLS regression.