Designing an Ethnic Minorities' Social Integration Policy

Yujung Hwang *

March 10, 2025

Abstract

The persistence of residential segregation and slow cultural assimilation among ethnic minorities has become an increasingly pressing policy concern in Western societies. Understanding how neighborhood segregation influences cultural assimilation is key to designing effective integration policies. To address this, I estimate a dynamic general equilibrium model of neighborhood choice and ethnic minorities' cultural assimilation using an 18-year micro panel dataset that tracks neighborhood choices and proxy measures of cultural assimilation. Both ethnic minorities and majorities select neighborhoods based on cultural similarity with neighbors, while ethnic minorities also decide on assimilation, with associated costs depending on their exposure to the majority population. Counterfactual policy experiments show that a rent subsidy policy encouraging ethnic minorities to relocate from enclaves to majority neighborhoods effectively promotes cultural assimilation and desegregation.

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, Peter Arcidiacono, Patrick Bayer, Jonathan Elliott, Matthew Wiswall, and seminar participants for their comments. All errors are my own.

1 Introduction

The residential segregation of minorities in Western societies, particularly among Muslim populations, has become a growing concern in recent years. In the United Kingdom—the focus of this study—nearly half of the Muslim population, which comprises only 5% of the total population, lived in neighborhoods where at least 20% of residents were Muslim, according to the 2011 Census. Projections by the Pew Research Center estimate that the Muslim population in the UK could rise to 16.7% by 2050 (Pew Research Center (2017)), raising concerns that persistent segregation may deepen social divisions. Similar patterns exist across other European countries, prompting discussions on integration policies. For example, Denmark implemented the *Ghetto Law* in 2018, which aims to disperse ethnic minorities living in so-called "ghettos" or *parallel societies*—defined as areas where at least 50% of residents are from non-Western backgrounds—into other parts of the country to promote integration (Lundsteen (2023)). Assessing whether such desegregation policies successfully foster social integration requires a deeper understanding of how residential segregation influences cultural assimilation, which is the focus of this paper.

This paper examines how ethnic minorities culturally assimilate and how this process interacts with neighborhood selection. Following the approach of Bisin et al. (2004), one of the first empirical studies to apply rigorous economic modeling to cultural evolution, I conceptualize both cultural assimilation and residential sorting as optimal decisions made by rational, forward-looking individuals. To study this, I develop a general equilibrium model of neighborhood choice and cultural assimilation using dynamic programming, estimating it with microdata from the United Kingdom. Using structural estimates from the model, I simulate counterfactual housing integration policies to assess their effectiveness in reducing segregation and fostering cultural assimilation. Structural modeling is particularly valuable for such counterfactual analysis, as it accounts for endogenous market price adjustments and potential "White flight" dynamics (Schelling (1971)) that could arise under large-scale desegregation policies. The results indicate that targeted rent subsidies encouraging ethnic minorities to relocate to majority neighborhoods are particularly effective in promoting cultural assimilation and reducing segregation, while ethnicity-neutral rent subsidy policies have a more limited impact.

The key economic intuition underlying the model is that cultural assimilation functions as a valuable human capital asset, facilitating social interactions and labor market participation. However, assimilation often entails substantial effort costs, making it an attractive investment primarily for those who expect long-term benefits—such as younger individuals with longer time horizons. A crucial factor influencing this decision is individuals' expectations about future social interactions. Individuals who expect to remain in ethnic enclaves, often due to

financial constraints on housing, may have little incentive to culturally assimilate, as their interactions will primarily be with culturally similar neighbors. In contrast, those planning to move to majority neighborhoods—whether for better amenities or labor market opportunities—are more likely to view cultural assimilation as beneficial, anticipating more frequent interaction with the majority population. At the same time, living in majority neighborhoods further facilitates cultural assimilation by increasing exposure to the dominant culture, reinforcing the interdependence between neighborhood choice and cultural integration. This dynamic life-cycle general equilibrium framework captures these interactions.

The model advances the structural migration literature by endogenizing cultural assimilation, a key factor influencing location choice. Previous studies have typically treated individual unobserved heterogeneity in location preferences as fixed (Kennan and Walker (2011), Bishop (2012), Galiani et al. (2015), Piyapromdee (2021), Bayer et al. (2016), Davis et al. (2021), Davis et al. (2023)). Each period, households make sequential decisions about their neighborhood, female employment, and future cultural assimilation status, taking into account equilibrium prices—housing prices, local skill prices, and neighborhood ethnic compositions—along with other exogenous location characteristics. A key implication of the model is that cultural assimilation by ethnic minorities is undersupplied in a competitive equilibrium, as individuals do not internalize the broader societal benefits of their assimilation. This externality in neighborhood choice, driven by homophily, prevents applying the First Welfare Theorem that guarantees Pareto-efficiency of competitive equilibrium and provides a rationale for government intervention to desegregate housing markets and incentivize assimilation.

The primary dataset used for estimation is the Millennium Cohort Study (MCS), which spans 18 years and includes repeated observations on detailed neighborhood locations and ethnic minorities' assimilation status. This dataset allows for a life-cycle analysis of neighborhood sorting and cultural assimilation. In selecting proxy variables for cultural assimilation, I follow Adida et al. (2016), who identify religiosity, gender attitudes, and language proficiency as key dimensions of Muslim immigrants' integration in Western Europe. To manage the high dimensionality of this information, I employ recent econometric approaches for modeling unobserved heterogeneity using proxy variables (Hu (2008), Hu and Shum (2012), Hu (2017), Cunha et al. (2010), Agostinelli and Wiswall (2025)). These methods treat proxy variables as noisy measurements of a latent cultural assimilation type, estimating their relationships with the latent variable while accounting for measurement error. This approach extends the structural applied microeconomics literature on cultural evolution (Bisin et al. (2004), Fogli and Veldkamp (2011), Fernández (2013), Bisin and Tura (2019), Adda et al. (2020)), which has traditionally focused on single cultural dimensions using proxy variables assumed to be free of measurement error, such as female labor force participation, fertility, religion, and interracial marriage.

The model is estimated using the Conditional Choice Probability (CCP) estimator Hwang (2024), which extends Arcidiacono and Miller (2011) by incorporating proxy likelihoods. The CCP estimator is particularly well-suited for this setting, as the presence of multiple equilibria and the model's general equilibrium structure make full-solution-based estimators, such as the Simulated Method of Moments, challenging to apply.[1]

Structural estimates confirm that homophily—the tendency to migrate to culturally similar neighborhoods—is strong for both groups but significantly stronger among White British than non-White British. Additionally, the neighborhood environment plays a crucial role in cultural assimilation for non-White British individuals. Those living in closer proximity to White British residents experience significantly lower assimilation costs. This finding implies that housing market desegregation policies will influence the rate of cultural assimilation among non-White British, with the speed of assimilation varying based on the degree of exposure to majority-group neighbors.

The counterfactual experiment compares two desegregation policies—an ethnicity-conscious policy (similar to the Gautreaux Program in Chicago, Chyn et al. (2025)) and an ethnicity-neutral policy (similar to the Moving to Opportunity (MTO) program in various U.S. cities, Chetty et al. (2016)). The Gautreaux Program, which relocated African American families to majority White neighborhoods, was found to be highly effective in fostering racial desegregation and integration, as evidenced by increased interracial marriage rates and upward mobility for treated families. In contrast, the MTO program focused on relocating participants to low-poverty areas, which were not necessarily predominantly White. As Chyn et al. (2025) notes, *"Among MTO participants who relocated to low-poverty neighborhoods, just 18 percent moved to neighborhoods that were majority white."* While MTO successfully promoted upward mobility for young children (Chetty and Hendren (2018)), it was less effective at reducing racial segregation. For discussions on other social integration policies, see Fouka (2023).

Simulating similar policies in the UK, I find that ethnicity-conscious subsidies significantly enhance both desegregation and cultural assimilation, while ethnicity-neutral subsidies have minimal impact. Under a welfare-maximizing means-tested ethnicity-conscious subsidy (capped at 90%), the unassimilated non-White British share decreases by 5% (1.91 percentage points), and the segregation index declines by 22% (0.12) by the year 2040. A more modest means-tested ethnicity-conscious policy that provides a similar average subsidy to the current UK housing subsidy program (capped at 60%) still reduces the unassimilated non White British by 2% (0.67 percentage points) and segregation by 9% (0.05). In contrast, an ethnicity-neutral subsidy with a similar average subsidy to the current UK program (capped at 70%) results in a 1% increase in the unassimilated non White British (0.33 percentage points) and only a 1% de-

[1]However, counterfactual simulations require solving for equilibrium, in which case I adopt an explicit equilibrium selection rule to identify the outcome closest to the baseline.

crease in the segregation index (0.01). Importantly, means-tested ethnicity-conscious policies improve the welfare of both White British and non-White British residents. These findings highlight the potential of well-designed, targeted integration policies to simultaneously reduce segregation, promote cultural assimilation, and enhance overall social welfare—offering a viable path for policymakers seeking to foster more integrated societies.

The paper is structured as follows: Section 2 discusses the data and reduced-form evidence. Section 3 presents the structural model. Section 4 outlines the estimation strategy and explains estimation results. Section 5 analyzes counterfactual policies. Section 6 concludes.

2 Descriptive Evidence

2.1 Context and Data

My study focuses on England and Wales, where Muslims—who make up about 5% of the total population (Census 2011)—are among the most segregated minority groups. Nearly half of all Muslims reside in neighborhoods where at least one in five residents is also Muslim. Meanwhile, the White British population, the largest ethnic group, has been steadily declining due to immigration, with its share falling from 87.5% in 2001 to 74.4% in 2021[2].

The primary dataset used for estimation is the Millennium Cohort Study (MCS; University of London. Institute of Education. Centre for Longitudinal Studies. (2022)), which tracks individuals born between 2000 and 2002 along with their guardians, primarily biological parents. This dataset is particularly valuable for this study for several reasons. First, it oversampled ethnic minority households, including 500 Muslim households. Second, it repeatedly collected data on cultural preferences and English proficiency, which are critical for identification. Third, the Secure Access version provides detailed retrospective migration histories at a fine geographic scale—Lower Super Output Areas (LSOAs)—which have an average population of only 1,500 residents. The dataset covers all 34,378 LSOAs[3] in England and Wales.

Following Adida et al. (2016), I select proxy variables to identify cultural assimilation; Table 1 lists the selected measures[4]. Adida et al. (2016) highlight religiosity, gender attitudes, and language proficiency as key cultural dimensions shaping the assimilation of Muslim immigrants, one of the largest non-White British subgroups in the dataset. To reduce dimension-

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

[3]Geographic boundaries as of 2001 were used.

[4]The MCS includes numerous gender attitude-related questions. To improve computational efficiency, I selected a subset based on relevance. The primary variable chosen was: "A child is likely to suffer if his or her mother works." Additional proxies were identified using LASSO regression to select the best predictors of this main variable.

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

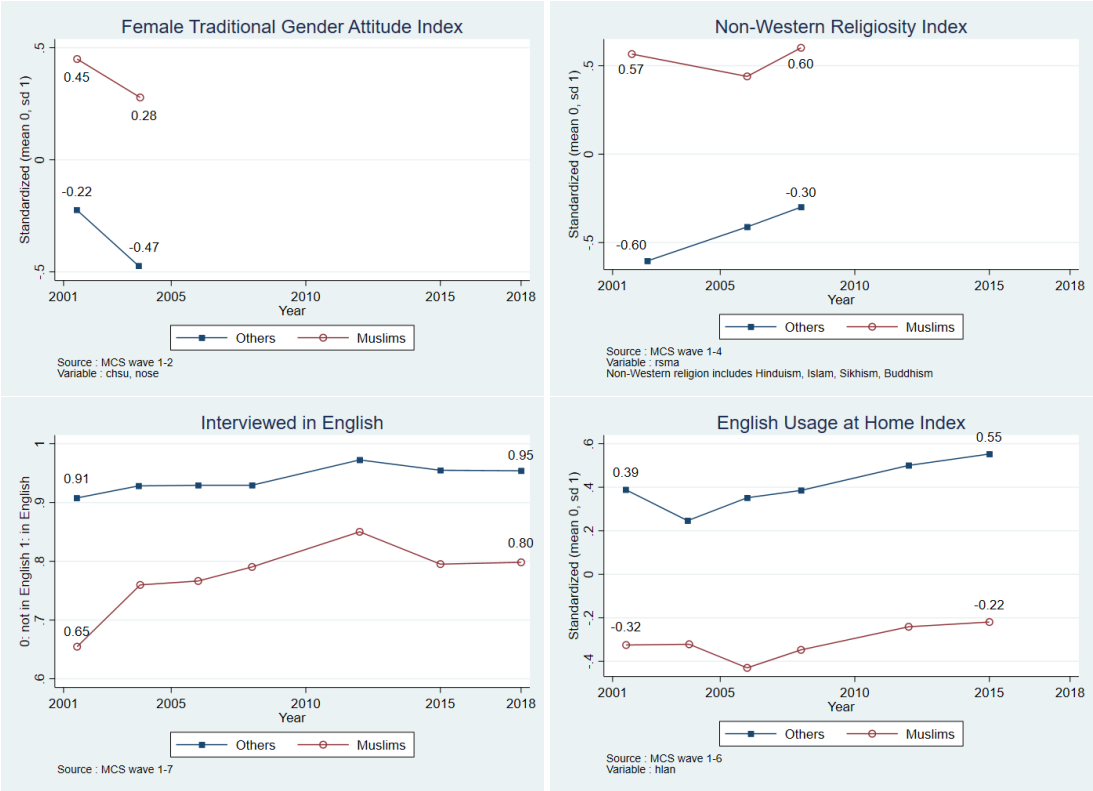
Note: This table shows the MCS questionnaires for the proxies used in the estimation. †: Religious attendance in wave 2 was asked only for new respondents. ‡: Response scale changed in wave 3.

ality, I estimate a single latent cultural assimilation variable, which is indirectly measured using proxy variables for both male and female spouses (Table 1). Additionally, I account for measurement error in these proxies by estimating its extent and incorporating it into the estimation procedure (Online Appendix Section F).

Another commonly used measure of cultural assimilation—the foreignness index of names (Fouka (2020))—is not included in this analysis for several reasons. First, respondent names are not available in this dataset. Second, because names are typically assigned at birth and rarely change, they do not effectively capture the gradual process of cultural assimilation over an individual's lifetime.

I supplement the MCS with several additional datasets, summarized in Supplemental Appendix A. To estimate local labor market skill prices and Total Factor Productivity (TFP), I use administrative payroll tax records from the Annual Survey of Hours and Earnings (ASHE; Office for National Statistics (2022)). To estimate the yearly income process—since MCS does not provide annual income data—I use the British Household Panel Survey (BHPS) and the UK Household Longitudinal Study (UKHLS) (University of Essex, Institute for Social and Economic Research (2022)). Additionally, I incorporate Census and Census Migration Flow data for local area statistics and migration distance distributions. Median house price data is sourced from HM Land Registry, and neighborhood deprivation index data is obtained from Abel et al. (2016).

Figure 1: Evidence of Time-Varying 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 Reduced-Form Evidence

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

2.2.1 Fact 1: Evidence of Time-Varying Cultural Traits

The cultural traits of ethnic minorities 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

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

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.

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

This section demonstrates that cultural preferences are strongly linked to neighborhood choice, even after accounting for economic factors. This finding motivates the development of a model in which neighborhood selection is influenced by the cultural preferences of ethnic minorities.

Table 2 examines how the ethnic composition of a neighborhood—measured as the standardized White British density—varies with respondents' cultural preferences. The cultural preference indices used here are similar to those in the previous section[5]. Details on index construction are provided in Supplemental Appendix B. All regressions control for key economic variables, including education, income, and age, and standard errors are clustered at the household level. The last row of the table presents the partial R-squares for the cultural preference indices.

The results indicate a strong relationship between cultural preferences and neighborhood sorting. Non-White British individuals with cultural traits more distinct from White British tend to live in neighborhoods with fewer White British residents. Specifically, a one-standard-deviation increase in cultural preference indices corresponds to a 0.19 to 0.24 standard deviation decrease in neighborhood White British density. This suggests that cultural preferences play a significant role in residential sorting, even after controlling for economic factors.

2.2.3 Fact 3: Migration Patterns in the Estimation Sample

This section examines the migration patterns observed in the estimation sample[6]. Table 3 summarizes these migration patterns.

Attrition in the selected sample is not severe: on average, individuals are observed for 15 out of the 18 total years. Notably, the number of observed years is similar between White British and non-White British individuals, alleviating concerns about selective attrition[8].

Migration rates are slightly higher for non-White British individuals: 50% of White British

[5] A minor difference is that these indices aggregate responses from both male and female spouses.

[6] For details on sample selection, see Section 4.1. The MCS survey is not conducted annually, but I construct a yearly migration history using retrospective questions.[7] The analysis focuses on the natural parent sample, excluding step-parents, under the assumption that parents lived with the cohort member throughout. To minimize sample loss, when the exact year of relocation is missing, I impute it by randomly selecting a year between the last and current survey waves. Since the average gap between survey waves is relatively short (2–4 years), this imputation likely introduces only a 1–3 year discrepancy from the actual moving year. Further details on missing data for the year of relocation are provided in Table E.1 in the Supplemental Appendix.

[8] Since the non-White British sample includes more immigrants, they might be more likely to have emigrated during the sample period, leading to sample dropout. Although the data lacks detailed attrition information to directly identify emigration events, the comparable number of observed years between White British and non-White British individuals suggests that selective emigration is not a significant concern.

and 55% of non-White British moved at least once during the sample period. Among those who moved, White British relocated an average of 1.5 times, while non-White British moved 1.3 times. Both groups exhibited a tendency to move to neighborhoods with a higher share of White British residents (56% and 68% of moves, respectively) and a lower share of Muslim residents (55% and 68%, respectively). However, non-White British individuals showed a stronger preference for such neighborhoods. On average, when moving, non-White British relocated to neighborhoods with 13 percentage points higher White British density and 8.8 percentage points lower Muslim density. In contrast, White British individuals moved to neighborhoods with only a 2.4 percentage point increase in White British density and a 0.8 percentage point decrease in Muslim density.

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.

2.2.4 Fact 4: Correlation Between Migration and Cultural Preferences

Table 4: Correlation Between Change in Neighborhood (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. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

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[9] 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, including 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

[9]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.

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

This section examines whether cultural assimilation leads to economic gains in the labor market. Table 5 presents individual fixed-effects regressions of female labor earnings and employment on cultural preference indices used in previous sections. A similar analysis for male labor earnings yields comparable results (Table E.2 in the Supplemental Appendix).

The results indicate that *within-individual changes* in cultural assimilation do not influence *within-individual changes* in labor earnings. While this may seem inconsistent with prior research linking cultural assimilation levels to wage levels (Abramitzky et al. (2020)), the findings suggest that although initial cultural assimilation may be correlated with wage levels, subsequent assimilation does not drive wage growth. For example, ethnic minorities with limited cultural and language skills may work in occupations where these skills are less essential, such as manual labor. In these jobs, incremental improvements in language proficiency may have little effect on wages. However, a reduction in traditional gender attitudes is associated with a higher likelihood of female employment, while changes in non-Western religiosity or English proficiency show no significant relationship with employment outcomes.

These findings have direct implications for the structural model introduced in the next section. The *within-individual* wage growth specification should exclude cultural assimilation changes as explanatory variables, provided that initial wage levels remain unrestricted in their relationship to initial cultural assimilation. However, the structural female labor supply equation must incorporate *within-individual* cultural assimilation changes, as the evidence

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

suggests that assimilation influences female labor force participation.

3 Model

Before presenting the formal model, I provide an overview of its key components. The model conceptualizes cultural assimilation as a latent discrete type, referred to as "social position," which represents a household's position in a one-dimensional "social space." This social position serves as a state variable in a dynamic life-cycle model of neighborhood choice.

The majority population, White British, are assumed to have a fixed social position at the leftmost end of the social space ("Type 1") throughout their lifetime. While, in principle, the majority population could assimilate into minority cultures, I do not model cultural assimilation for White British households due to data limitations. The survey lacks measures for how

majorities adopt minority cultures, whereas it does track cultural assimilation among ethnic minorities. This assumption is also reasonable in the UK context, where White British people typically do not learn minority languages such as Hindi or Urdu as immigration increases, whereas South Asian immigrants commonly learn English upon arrival. In contrast, ethnic minorities can actively change their social position relative to White British by exerting costly effort. They can transition between Type 1 (most assimilated) and Type 2 (least assimilated). I assume that Type 1 individuals—the most assimilated non-White British—are indistinguishable from White British.

The model incorporates several assumptions for tractability and identification. First, I explored an extended version where the most assimilated non-White British are distinguished from White British, but this specification was not identified given the available data. As a result, I impose the assumption that the most assimilated non-White British share the same social position as White British. Second, I originally considered allowing for three social position types instead of two but found that the available proxy data does not cleanly distinguish among them, and estimation results were sensitive to different identification assumptions (see Online Appendix Figures E.2 and E.3). Therefore, I assume only two social position types, though the framework can be extended to accommodate additional types if future data allows.

The model follows a dynamic discrete choice framework in which households make three sequential decisions in each period: 1. Neighborhood choice – Households choose where to live, considering economic factors (e.g., housing prices, local skill prices) and social factors (e.g., cultural proximity to neighbors), 2. Female employment decision – This decision is explicitly modeled because female labor market participation often signals cultural assimilation, providing important identifying variation. 3. Social position choice for the next period – The cost of cultural assimilation depends on neighborhood composition, particularly the share of White British neighbors.

These decisions are forward-looking: when choosing their social position type, individuals account for the impact on their future migration choices, employment opportunities, and assimilation incentives. For instance, ethnic minorities who expect to remain in ethnic enclaves have little incentive to assimilate, while those planning to move into majority neighborhoods may find assimilation beneficial.

The model operates in general equilibrium, with three key endogenous aggregate variables: 1. Housing prices – Determined by the local demand for housing, therefore migration decision, 2. Skill prices in the labor market – Influenced by the human capital distribution of workers in each region. 3. Neighborhood composition – Evolving based on individual migration decisions.

To close the model, I specify equilibrium equations for housing prices and skill prices, en-

suring that market conditions adjust in response to individual choices. Neighborhood composition is also endogenous, shaped by the collective migration decisions of all households.

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

Environment: Choices, Timing, Shocks

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, including all subperiods) 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 $\Omega_{it(s)}$. Under the assumption of perfect foresight, households are assumed to have complete knowledge of future aggregate conditions, including prices and neighborhood compositions, in all subsequent subperiods and periods. 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 $t(s)$. The social position $u_{it(s)}$ is a representative position of the household that is identified from the responses of both male and female spouses. So when a non-White British female has a White British spouse, which is 13.5% of our non-White British sample, such a household is likely to be identified as an assimilated ethnic minority due to the male spouse's responses in our estimation. Due to the discrete nature of survey measurements, social position $u_{it(s)}$ is assumed to be discrete, so $u_{it(s)} \in \{1, \cdots, \overline{u}\}, \overline{u} < \infty$. 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 $u_{it+1(1)}$[10]. 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 $\{\epsilon_{1it(1)}, \cdots, \epsilon_{Jit(1)}\}$, $\epsilon_{jit(1)} \overset{i.i.d.}{\sim} Gumbel(0,1)$[11] 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 allow every structural parameter to vary by White British ethnicity (D_i) to capture potential differences in preferences[12] or potential market discrimination. However, for notational simplicity, I omit the superscript D_i. This specification improves model fit and ensures that homophily preferences are not confounded with broader differences between White British and non-White British individuals, as neighborhood composition is often correlated with various other neighborhood characteristics.

First Subperiod: Location Choice

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. 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 j' from current neighborhood

[10]While the choice of social position could, in principle, be made at a different time, adopting a timing assumption where the unobserved state is determined one period in advance ensures consistency with the "Limited Feedback" assumption in Hu and Shum (2012). This, in turn, allows me to apply the identification results established in Hu and Shum (2012).

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

[12]For instance, non-White British individuals may have distinct preferences for housing quality due to cultural factors or exhibit different price elasticity, possibly because many send remittances to family members abroad.

$j_{it(1)}$ incurs a moving cost $MC(j_{it(1)}, j')$, which depends on the distance between one's own social position and the average social position in the destination neighborhood $C_2(|u_{it(1)} - \overline{u}_{j'}|)$. 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.[13]

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

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

Several alternative specifications were explored but ultimately not chosen due to identification challenges or poor model fit. First, the current specification models homophily as a moving cost factor rather than a flow utility amenity. An alternative specification[14] that includes homophily in the flow utility function failed to capture the strong tipping point behavior observed in the data, even with fixed moving costs (Card et al. (2008)). Specifically, the inflow of culturally dissimilar residents does not trigger immediate outmigration, but movers tend to relocate to culturally similar neighborhoods, reflecting strong homophily in location choice. The current specification better aligns with this pattern. Second, the model assumes that White British individuals like the most assimilated non-White British ($u_{it(1)} = 1$) the same as other White British. An extended specification introducing a parameter to capture potential differences in this preference was explored but proved difficult to identify in the data. Therefore, this restriction was adopted for identification. Third, cultural assimilation is assumed not to influence preferences for exogenous neighborhood amenities such as housing quality. This assumption was maintained because an extended model allowing utility for amenities to vary by social position type yielded imprecise estimates that were not statistically different across groups. Given this, I adopt a simpler specification, assuming a common γ_1^j for all individuals.

Second Subperiod: Female Employment

The flow payoff in the second subperiod following the female employment decision comprises of utility from the log of household income ($log(R_{lt(2)} \times (h_{it(2)} + h_{it(2)}^{SP})$) if the female spouse works and $log(R_{lt(2)} \times h_{it(2)}^{SP})$ if the female spouse does not work and disutility from working conditional on social position, $\sum_{k=1}^{\overline{u}} \gamma_{1k}^P 1(u_{it(s)} = k)$. $R_{lt(2)}$ is the equilibrium skill

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

[14] Under this specification, the model likelihood remains similar, with the social distance amenity coefficient for non-White British statistically indistinguishable from zero, while for White British, it is small but significantly positive.

price in the labor market l that the neighborhood j belongs to.

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

I assume that individuals accumulate human capital through work experience while also facing stochastic shocks each period. The social position of non-White British individuals does not influence *wage growth*. Importantly, this specification does not contradict prior research on labor market discrimination against immigrants, which finds a relationship between *wage levels* and cultural assimilation status (Abramitzky et al. (2020)). This is because, when calibrating initial human capital stock (detailed in Section 4), I do not impose any assumptions about its dependence on initial cultural assimilation status. Instead, this specification implies that, conditional on initial human capital—potentially shaped by early cultural assimilation—subsequent cultural assimilation does not affect wage growth during the sample period, as confirmed in Table 5. For instance, while non-native English speakers in the UK may initially earn lower wages, incremental improvements in their English proficiency alone do not predict wage growth after controlling for accumulated labor market experience post-arrival.

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.[15] 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 $P^F_{it(2)}$.

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

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

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

[15] 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).

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

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

Third Subperiod: Next Period Social Position (Assimilation Status)

The third subperiod flow payoff includes only the psychic cost, $\{PC_{u_{it(3)}u'}, PC^D_{u_{it(3)}u'}\}$, 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.

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

Labor and Housing Market Clearing

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[16]. 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

[16]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.

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,
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.[17]

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

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

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

Household Value Functions

Now I define the ex-ante value function for each subperiod, $V_{t(1)}(\Omega_{it(1)}), V_{t(2)}(\Omega_{it(2)}), V_{t(3)}(\Omega_{it(3)})$. Within a given period, there is no discounting across subperiods; discounting only occurs between subperiod 3 of the current period and subperiod 1 of the next period. The state variables $j_{it(s)}$ (neighborhood choice) and $u_{it(s)}$ (social position type) remain constant until the next decision point. This implies that $j_{it(2)} = j_{it(3)} = j_{it+1(1)}$ and $u_{it(1)} = u_{it(2)} = u_{it(3)}$ always hold. After each decision point, the state variables are updated, so $j_{it(1)} \neq j_{it(2)}$ and $u_{it(3)} \neq u_{it+1(1)}$ may hold. The age of the female spouse is constant within a period and increases by one after the third subperiod, so $a_{it(1)} = a_{it(2)} = a_{it(3)}$ and $a_{it+1(1)} = a_{it(3)} + 1$ hold.

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

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

[17] 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 $f(\Omega_{it})$.

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

It is worth mentioning that the above value function has two-period finite dependence, which speeds up estimation later.[18] 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* $\{j_{it(2)}, P^F_{it(2)}, u_{it+1(1)}\}_{t\in\mathcal{T}}$ *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.[19]

[18] 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.

[19] 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.

3.1 Model Discussion

The model highlights the relationship between geographic sorting and ethnic minorities' assimilation.

First, the economic factors changing geographic dispersion will inevitably change ethnic minorities' assimilation. Such factors include monetary incentives, like local skill prices and rents, as well as the economic disparity (in human capital) between ethnic minorities and majorities. These factors make ethnic minorities choose a different neighborhood, meet different neighbors, and alter their incentive to invest in cultural assimilation. For example, even if the main reason for ethnic minorities to live in an ethnic enclave is economic reasons, once they start living in an ethnic enclave, they may stop learning the majority 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, ethnic minorities' assimilation effort is undersupplied in a competitive equilibrium because ethnic minorities do not internalize the welfare gain of majorities when they assimilate into the majority culture. For example, non-White British in the model do not consider that their cultural 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 ethnic minorities' 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, ethnic minorities are blamed alone for their segregation and lack of assimilation efforts. However, according to the model, ethnic minorities might become segregated due to majorities' high prejudice (represented as high moving cost C_2) and their consequent "White flight". Once ethnic minorities become isolated, it is more costly for them to assimilate culturally.

The current model does not incorporate all potentially relevant mechanisms. First, it assumes that interactions between non-White British and White British do not alter prejudice or homophily, as reflected in the constant C_2 parameter, which captures social barrier-related moving costs. However, the "Contact Hypothesis" from social psychology (Paluck et al. (2019)) suggests that increased interaction between racial groups may reduce intergroup prejudice. If such contact effects exist, C_2 would not be constant but instead evolve with changes in the social environment. This mechanism is excluded from the model due to the lack of high-quality data on prejudice. However, if such data were available, the model could be extended to account for endogenous changes in prejudice.

Another omitted mechanism is the impact of segregation on public goods provision and intergroup conflict. A rich political economy literature (Easterly and Levine (1997), Alesina et al. (1999), Alesina et al. (2003), Alesina and La Ferrara (2005), Dahlberg et al. (2012)) explores

how diversity influences these outcomes. The primary challenge in incorporating this mechanism is the lack of reliable measurements. Even if such data were available, identification would require sufficient variation across locations and a measurable influence on migration decisions. Without these conditions, estimating how much individuals value high-quality public goods or lower intergroup conflict would be difficult. Given these challenges, I leave this extension for future research.

4 Estimation

4.1 Sample Selection

The main estimation sample is drawn from parent surveys in the MCS dataset, covering the years 2001 to 2018.[20] The sample selection follows several criteria. I include married couples[21] living in England and Wales, where the female spouse is aged between 25 and 64. I use only consecutive spells that contain estimation-relevant information,[22] and I exclude short spells (fewer than ten years),[23] as longer observations are necessary for identification.

Household ethnicity is determined by the female spouse's ethnicity. Among non-White British females, 13.5% had a White British spouse (Table E.3 in Appendix). Since the household's assimilation status is identified using survey responses from both spouses, interracial non-White British couples are likely to be classified as assimilated ethnic minorities due to the male spouse's responses.

I include all eligible non-White British households but randomly select an equal number of White British households to match the sample size, as computational capacity constraints on the secure data server limit the dataset size.[24] The final estimation sample consists of 2,468 households (1,234 White British and 1,234 non-White British).

Table 6 presents the demographic characteristics of female sample members at the time of their first interview. A corresponding table for male sample members is provided in the Supplemental Appendix (Table E.3). The average age for both White British and non-White British females is in the early 30s.

[20] Most interviews in England and Wales for wave 7 were completed in 2018, so I exclude a small number conducted in 2019.

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

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

[23] Shorter spells provide insufficient information to reliably identify unobserved cultural types.

[24] The Millennium Cohort Survey Secure Access data can only be accessed within a secure data server.

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 to protect confidentiality, as they are small in number.

By construction, the White British sample does not include any Muslims,[25] whereas nearly half of the non-White British sample identifies as Muslim. Additionally, 98% of non-White British females are either first-generation or second-generation immigrants. Among first-generation immigrants, the average time spent in the UK at the time of the first survey was 12 years.

Non-White British females had significantly lower college graduation and employment rates compared to their White British counterparts. The residential patterns also differ markedly: the White British sample predominantly lived in White British neighborhoods, with an average White British density of 92%, whereas for non-White British females, the average White

[25] A small number of White British Muslims were excluded from the estimation sample.

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.

British density in their neighborhood was only 46%. Similarly, the presence of Muslim neighbors was minimal for the White British sample (average Muslim density of 2%), whereas non-White British females lived in neighborhoods with a substantially higher average Muslim density of 29%.

4.2 Choice Set Construction and Preset Parameters

First, I construct a stratified random choice set of size 31, following the approach in Keane and Wasi (2012), to address the curse of dimensionality that arises from the large number of potential migration destinations (34,378 LSOAs) relative to the sample size. A choice set of this size ensures that all neighborhoods receive both in-migration and out-migration while keeping computation feasible. To ensure a diverse set of potential destinations, I classify neighborhoods into three categories and randomly select 10 neighborhoods from each: (1) White British neighborhoods – where the White British population density exceeds the national average (80.5%), (2) Muslim neighborhoods – where the Muslim population density is above 20%, (3) other neighborhoods – neighborhoods that do not fall into either of the first two categories. Including the individual's current neighborhood, this results in a total choice set of 31 neighborhood LSOAs.

To account for the fact that people tend to migrate to nearby areas, I incorporate a distance-based selection mechanism when drawing the 10 neighborhoods in each category. First, I randomly draw a migration distance from the 2001 Census Migration Flow data to match observed migration distances. Then, within each category, I select the neighborhood whose

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".

actual geographic distance from the household's current location is closest to this drawn migration distance. Finally, if the household's actual migration destination is not included in the constructed choice set, I replace one neighborhood in the corresponding category with the actual destination.

Several parameters were preset, as 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.

4.3 Income Process Estimates

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

The estimation procedure follows the approach used in Meghir and Pistaferri (2004), Blundell et al. (2008), and Voena (2015), with one modification: I account for endogenous skill prices, which vary across labor markets and years. To correct for selection into female employment, I apply the "Heckit" correction method, following Voena (2015). The instrumental variable (IV) for female employment selection is whether the respondent's mother (and spouse's mother) worked when the respondent was 14 years old. The exclusion restriction requires that the IV be independent of current-period income *shocks* (though not necessarily income *levels*), making a predetermined lifetime event a valid instrument. Additionally, this

IV is strongly correlated with female employment through preferences for work. Table A.2 presents the first-stage estimation results, confirming the IV's strong predictive power for female employment.

The results indicate that the estimated experience gain for non-White British individuals is only half that of White British individuals, suggesting potential differences in occupational sorting or differential treatment in the labor market. However, the variance of permanent income shocks and measurement errors appears similar across both groups.

4.4 Production Function Estimation & Human Capital Stock Calibration

First, I construct local labor market groups by clustering Travel-to-Work Areas (TTWAs; 2001 geography), which are the official labor market units used in national statistics and are larger than neighborhood LSOAs. Clustering is necessary because the ASHE dataset, used to estimate skill prices, does not have sufficient observations for every TTWA. To define these clusters, I apply a k-means clustering algorithm to TTWAs based on their average real income (2003–2018) and their distance to three major metropolitan areas: London, Manchester, and Birmingham. This procedure yields five distinct local labor market groups.[26] Figure A.1 in the Appendix presents a map of these labor market classifications.

To estimate the production function parameters, I follow the methodology of Heckman et al. (1998). Further details on this estimation are provided in Section C. The main identification strategy relies on the Cobb-Douglas production function, which can be transformed into a log-linear equation. I address the endogeneity of labor supply by using an instrumental variable approach, employing a Bartik instrument (Bartik (1991)) based on the industry composition share in the base year (2003). Table A.1 in the Appendix presents the first-stage estimation results, confirming that the instrumental variable strongly predicts total human capital stock (labor supply) in the local labor market.

Table 9 presents the estimated production function curvature parameter and the average log skill price across the five local labor markets. London is the highest-paying labor market, followed by the South East region surrounding London. The Midlands, which includes Manchester and Birmingham, ranks third in pay, while the South and North rural areas offer the lowest wages. Figure E.4 in the Supplemental Appendix illustrates the estimated log skill prices and total human capital stock for each labor market over time.

The production function and income process parameter estimates serve as the foundation

[26]For geographic continuity, I manually reassigned two TTWAs—"Berwick" and "Morpeth, Ashington & Alnwick"—to a geographically closer local labor market group in the North near Scotland.

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 G.1 in the Appendix.

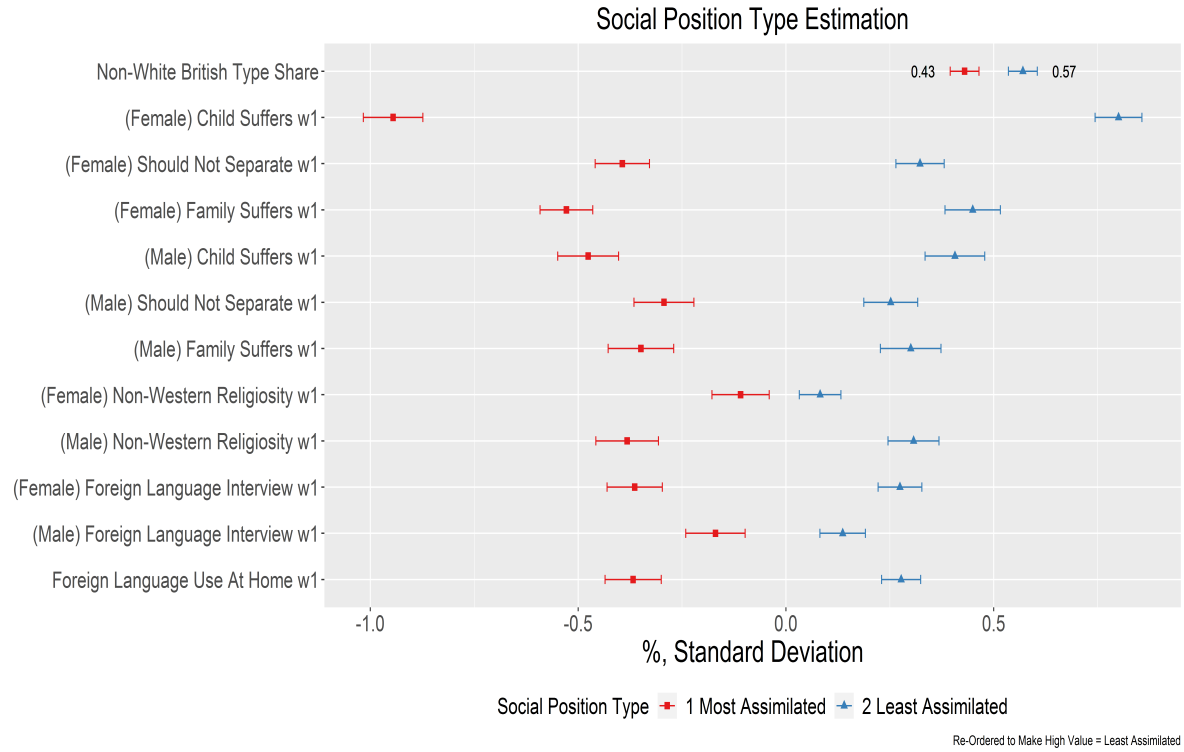
for calibrating individuals' initial human capital stocks. Since individual wages are the product of local skill prices and human capital stock, I use the estimated skill prices to recover individual human capital levels from observed wages, accounting for potential measurement errors. To impute missing wages for unemployed individuals, I apply the following procedure:

For *ever-employed* individuals (those who worked at least one year during the sample period), I calibrate their initial human capital stock by maximizing the simulated likelihood of their observed wage sequence. However, for those who were *never employed* during the sample period, wages are entirely unobserved. To address this, I first estimate a regression model using only the *unemployed-but-ever-employed* sample. The model includes detailed demographic characteristics such as female and male education dummies, a quadratic age polynomial for females, log house prices, log skill prices in the local labor market, log neighborhood Muslim and White British densities, and ethnicity group dummies. Using the estimated coefficients, I predict human capital stocks for those who were never employed. Finally, I calibrate their initial human capital stocks by maximizing the simulated likelihood of the predicted human capital sequence from this regression-based approach.

4.5 Estimation of Social Position (Assimilation Status)

The data includes repeatedly measured proxy variables related to social position, allowing for the identification of social position dynamics at the household level. Following Cunha et al. (2010) and Hu (2017), I account for measurement errors in proxy variables and their impact on likelihood estimation. Specifically, I estimate the conditional probability that a proxy variable Y_{it} takes response k, given that the household's true social position type is l. This estimation is performed for each proxy variable and all possible values of k and l.

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



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

This estimation requires an identification assumption due to the "label-swapping problem." (Hu (2017))[27] To address this, I adopt the monotonicity assumption from Hu (2017): *"There exists an observable variable Y_{it} such that $E[Y_{it}|u_{it}]$ is increasing in social position u_{it}."* I use whether a female spouse was interviewed in a foreign language in wave 1 as the anchoring variable Y_{it} for this assumption. I set the number of social position types, $\overline{u}$, to two for the reasons I explain below.

To improve computational efficiency, I estimate these probabilities before the main estimation procedure using the EM algorithm, leveraging first-wave data, which contains all proxy variables (Table 1). The results, reported in the Online Appendix (Section F), indicate that most proxy variables are informative about the latent social position type, as their conditional likelihoods vary across social position types.

Figure 2 presents the estimated initial shares of social position types and their corresponding average proxy responses. Among non-White British women in Wave 1, 43% are identified

[27]For example, in identifying a function $y = f(x)$, where x is a discrete latent variable, an identification assumption is needed for x. Suppose we define another latent variable $\tilde{x}$, which reverses the order of x. We could then construct another function $\tilde{f}(\tilde{x})$ such that $y = \tilde{f}(\tilde{x}) = f(x)$, meaning $\tilde{f}(.) \neq f(.)$. This ambiguity prevents identification unless we impose a restriction on how x is labeled.

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, $\frac{L_{it}(u_{it(s)}=u')}{L_{it}}$

as the most assimilated, while the remaining 57% are categorized as the least assimilated. The two types are clearly distinguished based on their proxy responses, as seen in the second-to-last set of average proxy responses conditional on each type. Figure E.1 confirms the robustness of these estimates across different anchoring variables.

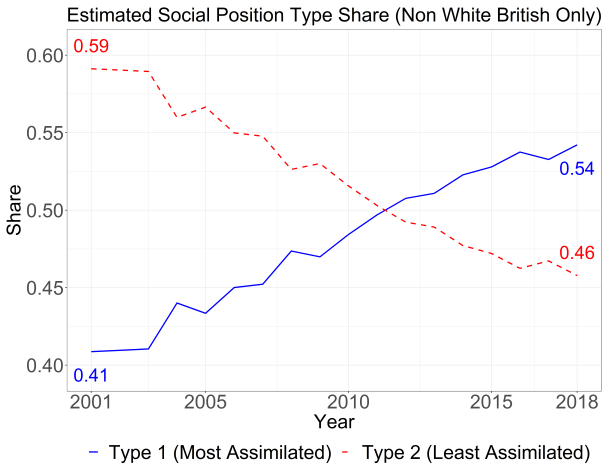
In contrast, assuming three types results in weaker differentiation between them and less robust estimates. Figure E.2 (Supplemental Appendix) shows that the middle type (Type 2) is largely indistinguishable from the least assimilated type (Type 3). Additionally, the estimation results become more sensitive to the choice of anchoring variable, indicating weaker separation between types (Figure E.3). These findings suggest that a three-type classification does not sufficiently capture distinct characteristics in this sample, supporting the decision to use two types.

4.6 Two-Stage CCP Estimation

The remaining structural parameters in the flow payoff are estimated using the two-step Conditional Choice Probability (CCP) estimator developed by Hwang (2024), which extends the Arcidiacono and Miller (2011) estimator. This approach incorporates likelihoods from proxy variables and allows for a flexible evolution of an unobserved state variable. For further details on the properties and performance of this estimator, see Hwang (2024).

Algorithm 1 outlines the first-stage estimation procedure; additional details are provided in Appendix C.3. The algorithm begins by initializing the individual social position type prob-

Figure 3: Social Position Type Share Estimate



ability $P^{(0)}(u_{i1} = s)$, the initial prior $P^{(0)}(u_{i1} = s|\Omega_{i1(1)})$, and the transition CCP $P^{(0)}(u_{it+1} = s'|u_{it} = s, \Omega_{it})$.[28] The estimation then iterates between the M-step and the E-step until the type probability $P^{(m)}(u_{it} = s)$ converges.

In the M-step, the per-period likelihood is estimated using a parametric (multinomial) logit model. This likelihood consists of three components: (1) neighborhood choice, (2) female employment decision, and (3) proxy variables measuring social position (assimilation status). To reduce computational time, the likelihood for proxy variables is pre-estimated using data from the first survey wave. Once the per-period likelihood is computed, the transition likelihood for assimilation status is estimated using a parametric fractional multinomial logit model (Papke and Wooldridge (1996)).

In the E-step, the social position type probability is updated using Bayes' rule. This iterative process continues until convergence.

The second-stage estimation of the structural parameters is conducted via quasi-maximum likelihood estimation (quasi-MLE), where the likelihood function leverages the two-period finite dependence property of the model. Appendix C.4 provides detailed derivations for constructing the likelihood function.

4.7 Social Position Type Share Estimates

The estimated social position type shares from the two-step CCP estimation (Figure 3) indicate significant cultural assimilation over the 18-year sample period. The share of the least assimilated type (Type 2) among non-White British households declined from 59% in 2001 to 46% in 2018.

[28] For details on initialization, see Appendix C.3.

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.

Table 10 summarizes the average characteristics of each estimated social position type, revealing intuitive patterns. Households classified as the least assimilated (Type 2) tend to have younger spouses (by 2–3 years on average), a higher proportion of Muslims, and a greater share of first-generation immigrants who have spent fewer years in the UK. This group also has lower female employment rates, lower educational attainment for both spouses, and lower household income. Additionally, they predominantly reside in neighborhoods with a higher concentration of Muslim residents and a lower share of White British residents.

While both social position types exhibit similar migration rates, Type 2 households have a slightly higher estimated moving rate than Type 1 households. This suggests that even the least assimilated individuals exhibit mobility, though their relocation choices may reinforce ethnic residential patterns.

4.8 Structural Parameter Estimates

Table 11 presents the structural parameter estimates from the CCP estimation. Many parameters are estimated with sufficient precision to be distinguishable from zero.

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.

The estimated spouse income level parameter for non-White British (b_1) is significantly lower than that of White British, suggesting systematic income differences across groups. Additionally, the correlation between spouses' incomes (b_2) is much stronger for non-White British households compared to White British households, indicating greater financial inter-

dependence. The estimated house price elasticity parameters (γ_0^j) are negative, confirming that higher housing prices deter migration into those neighborhoods. Moreover, the neighborhood control function parameter (γ_{14}^j) appears to play a crucial role in migration decisions.

The estimated moving cost associated with social position type (C_2) is substantially larger for White British than for non-White British, implying that segregation is primarily driven by White British households' reluctance to live near non-White British, rather than non-White British actively avoiding White British neighborhoods. This finding challenges the common perception that ethnic minorities self-segregate and instead suggests that segregation patterns are largely shaped by majority-group preferences.

The estimated gender norm cost parameters ($\gamma_{11}^P, \gamma_{12}^P$) suggest that cultural assimilation among non-White British households promotes female labor force participation. However, these estimates exhibit some degree of statistical noise and are not significantly different from each other.

The adjustment cost parameters for changes in social position are estimated to be large in absolute value, leading to persistent social position types over the life cycle. Notably, the cultural assimilation cost estimates (PC_{12}^D, PC_{21}^D) suggest that proximity to White British significantly reduces the cost of cultural assimilation while simultaneously increasing the cost of cultural dissimilation (i.e., transitioning from Type 1 to Type 2). This reinforces the idea that neighborhood composition influences the feasibility and desirability of assimilation choices.

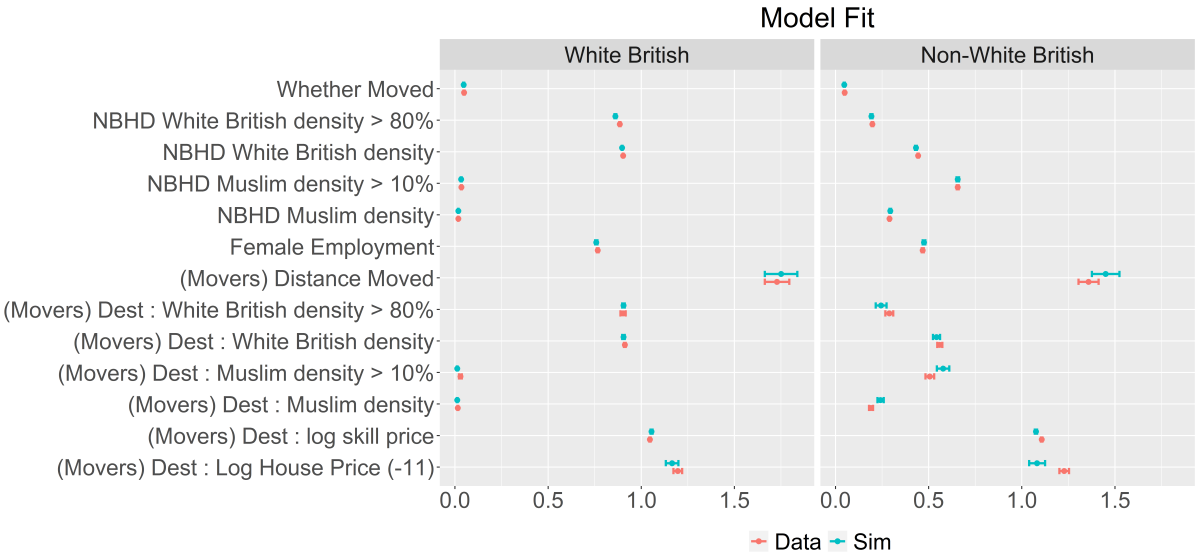
4.9 Model Fit

Overall, the model provides a good fit for key features, including migration patterns and female employment rate, as shown in Figure 4. The figure evaluates model fit separately for the full sample and movers only (labeled as "(Movers)"), demonstrating that the model estimates align well with observed data moments for both groups.

The migration statistics used to assess model fit include overall migration rates, the composition of movers' destination neighborhoods (measured by White British and Muslim densities), and the share of respondents residing in or relocating to White British towns (neighborhoods where the White British population exceeds 80.5%) and Muslim towns (neighborhoods where the Muslim population exceeds 20%). The model effectively replicates these observed migration patterns.

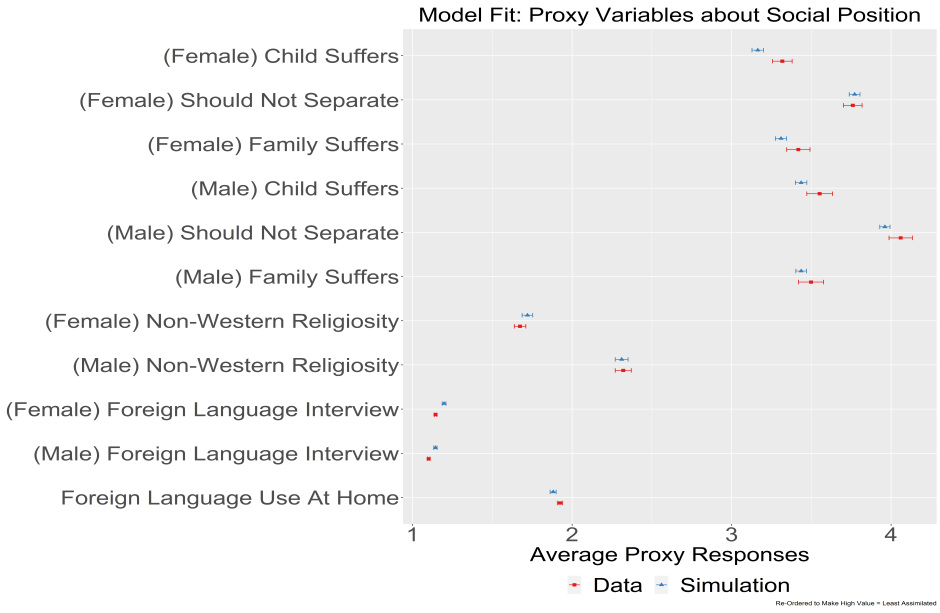
Figure 5 further demonstrates a reasonably good model fit for proxy variable responses related to social position types, disaggregated by gender.

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

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

This section evaluates a social integration policy aimed at fostering ethnic minorities' cultural assimilation and neighborhood desegregation. The policy offers rent subsidies to encourage relocation, with subsidy amounts varying by neighborhood and recipient characteristics. I

compare different subsidy conditions to determine which is most effective at increasing inter-ethnic mixing and cultural assimilation.[29] A structural model is particularly valuable for this analysis, as large-scale housing interventions inevitably affect market prices and may trigger "White flight" (Card et al. (2008)). The general-equilibrium framework allows for predictions about changes in housing rents, labor market skill prices, neighborhood compositions, and ethnic minorities' assimilation under counterfactual scenarios.

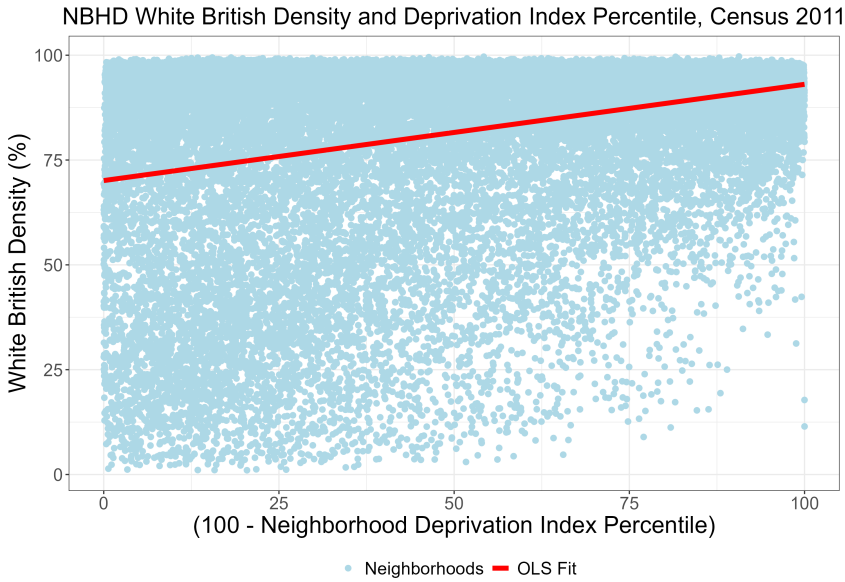
The use of housing policy as a tool for integration is well-documented. Historically, the Soviet Union and China relocated ethnic groups to promote integration (Hirsch (2005), Joniak-Lüthi (2013)). More recently, Singapore has enforced strict ethnic quotas in public housing since 1989—covering 80% of its housing market—to maintain ethnic diversity (Wong (2013)). Similarly, in 2018, Denmark adopted the 'Ghetto Law' to disperse non-Western minorities living in enclaves to different areas (Lundsteen (2023)) to promote their integration. In contrast, the U.S. government systematically promoted racial segregation through redlining, restricting African American access to White neighborhoods (Rothstein (2017)). To remedy past segregation, the Chicago Housing Authority introduced the Gautreaux Program, which provided Section 8 vouchers to help African American families relocate to affluent White neighborhoods (Chyn et al. (2025)). Research shows that program participants' children were more likely to intermarry, live in diverse neighborhoods, and experience upward mobility, demonstrating the causal effects of racial desegregation policies. However, race-neutral policies like the "Moving to Opportunity" program, which relocated low-income households to low-poverty areas, largely failed to desegregate neighborhoods, as recipients tended to move to low-poverty areas with familiar racial compositions.

This counterfactual experiment considers two kinds of rent subsidy policies: an ethnicity-conscious policy (similar to the Gautreaux Program), which tailors subsidies to encourage ethnic desegregation, and an ethnicity-neutral policy (similar to the Moving to Opportunity Program), which does not consider ethnicity. While the ethnicity-conscious policy is likely more effective at promoting integration, it may face political resistance. I quantify the efficiency trade-off when policy design is restricted to an ethnicity-neutral framework.

I evaluate two rent subsidy policies—ethnicity-conscious and ethnicity-neutral—each implemented at different subsidy levels. Let $\kappa \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$ represent the maximum subsidy rate. Under the ethnicity-conscious policy, the total rent subsidy depends on neighborhood ethnic composition. Non-White British households receive a subsidy of $\kappa \times$ (Neighborhood White British Share)% of the rent, while White British households receive $\kappa \times (1 - \text{Neighborhood White British Share})^{0.208}$% of the rent. The exponent 0.208 serves as an adjustment factor to balance average subsidy levels between White British and non-

[29] An alternative approach would be to incentivize cultural assimilation directly, such as providing tax benefits for immigrants who demonstrate language proficiency. However, a residential mixing policy fosters assimilation more holistically without requiring costly screening of cultural assimilation.

Figure 6: Correlation between White British share and Neighborhood Amenity



Note: This figure shows the scatter plot of neighborhood White British densities and the reverse ordered neighborhood deprivation index percentile, with the OLS fit.

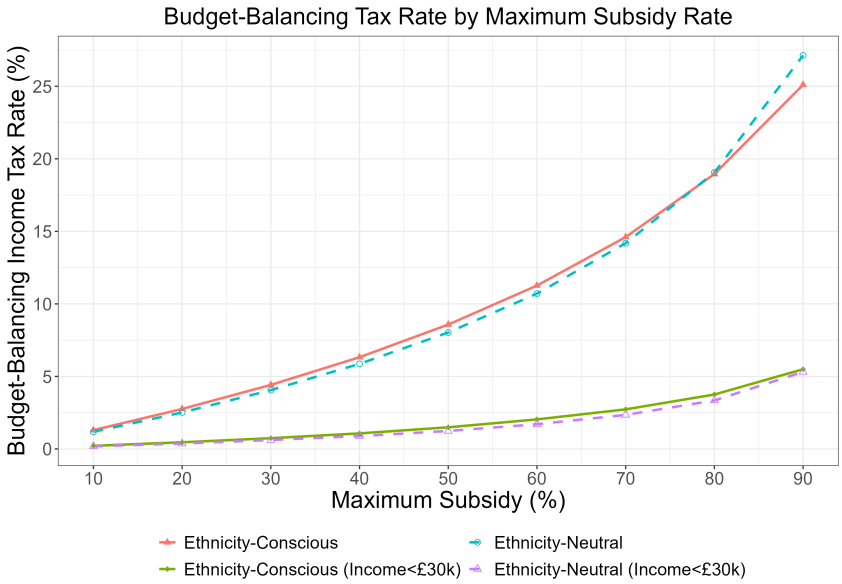
White British households. Since White British residents are the majority in most neighborhoods, this adjustment ensures that in a neighborhood where 75% of residents are White British and 25% are non-White British, both groups receive the same subsidy amount.

The ethnicity-neutral policy does not explicitly consider ethnicity but instead bases subsidies on neighborhood amenities. Subsidies are higher for individuals relocating to neighborhoods with better amenities, regardless of their ethnicity. Neighborhood amenities are measured as the inverse ranking of the neighborhood deprivation index, so the subsidy amount is given by $\kappa \times (100 - \text{NBHD Deprivation Index Percentile})/100\%$. Figure 6 shows that although neighborhood amenity levels are correlated with the White British share, there is substantial variation in White British composition across neighborhoods with similar amenity levels.

For both ethnicity-conscious and ethnicity-neutral policies, I also consider a means-tested version, where only households earning less than 30,000 GBP per year qualify for housing subsidies. This income threshold closely aligns with the current UK housing subsidy policy (Adam et al. (2015)).

A few assumptions underpin this counterfactual analysis. First, I assume the same linear income tax rate is imposed on everyone to make the government budget meet. So the counterfactual outcomes were derived considering the welfare and behavioral changes incurred by this additional income tax. 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. The interest rate for government borrowing and saving is set at 0.053, which is consistent with the time discount β assumed ($r = (1/\beta) - 1 \approx 0.053$). Second, I assume people

Figure 7: Budget-Balancing Linear Income Tax Rate for Different Subsidy Policies



Note: This figure shows the budget-balancing linear income tax rate for different subsidy policies

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[30]. Third, 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[31]. Fourth, Total Factor Productivity (TFP) in the labor market and the housing stock is assumed to remain at the same level of year 2011.[32].

Figure 7 shows the budget-balancing linear income tax rate under different policies, which considered endogenous market prices, female labor supply, and migration. The budget-balancing tax rates for ethnicity-conscious and ethnicity-neutral policies are similar under the chosen parameters. When everyone is eligible for housing subsidy regardless of their income, the budget-balancing tax rate must be set higher, exceeding 25% when the maximum subsidy rate is 90%. When only low-income households are eligible, the required budget-balancing tax rate is much lower, close to 5% when the maximum subsidy rate is 90%. For all policies, the

[30]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.

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

[32]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.

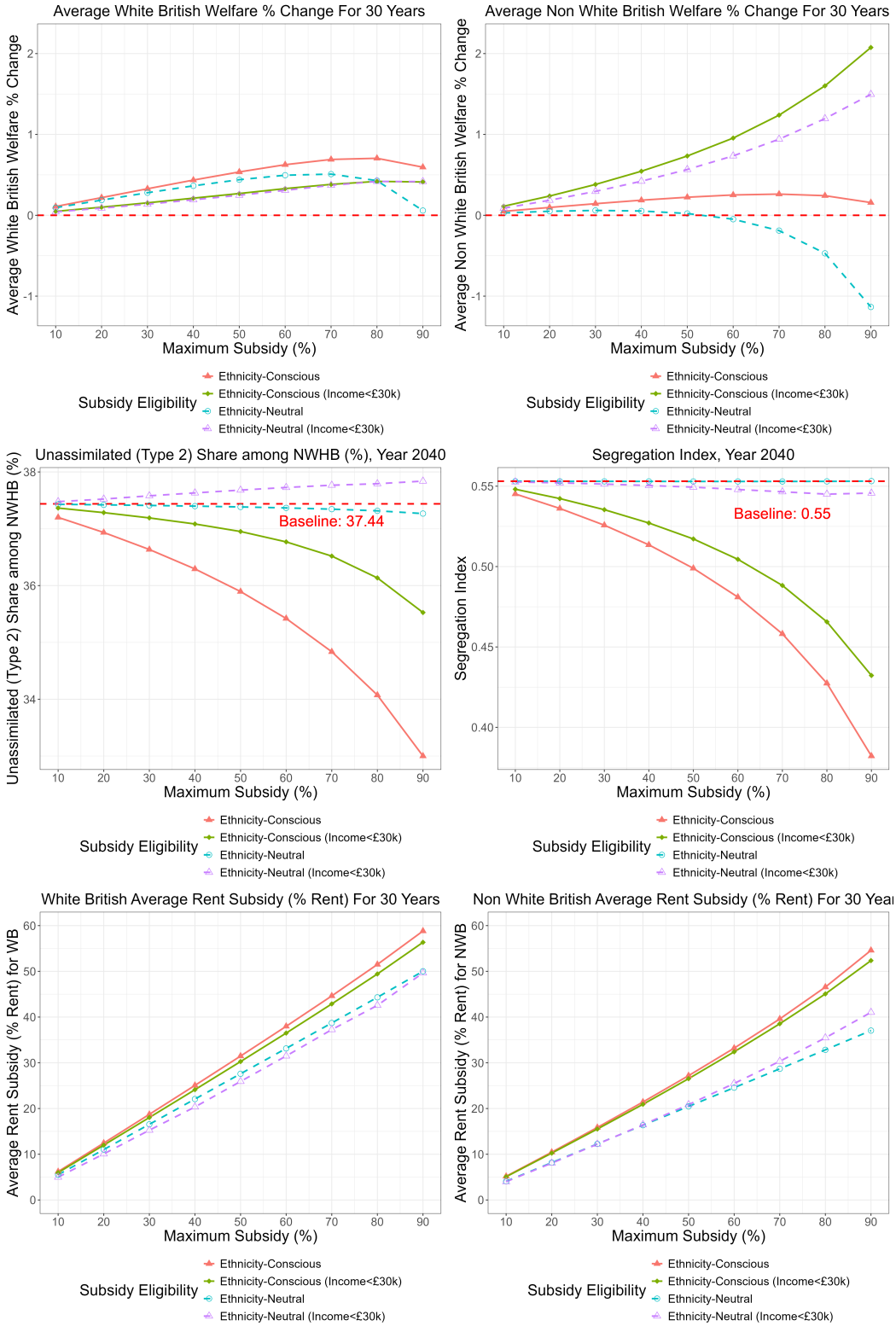
budget-balancing tax rate increases at a much steeper rate when the maximum subsidy rate exceeds 70%.

Figure 8 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. Due to the externality in location choice, competitive equilibrium may not be Pareto-optimal and the government may increase welfare through intervention. Under most policies, both White British and non-White British experience welfare gains stemming from changing market prices following desegregation and ethnic minorities' assimilation. The only case when welfare loss happens is for non-White British under the ethnicity-neutral policy with the maximum subsidy rate greater than 50%. If we restrict the focus to only means-tested one, 90% maximum subsidy rate policy maximizes both White British and non-White British welfare.

The middle two panels show the outcome of cultural assimilation and desegregation in the year 2040. Compared to the baseline prediction in the absence of any policy intervention indicated by the red dashed line, we can expect a sizable change in immigrants' cultural assimilation and segregation under ethnicity-conscious policy. The ethnicity-neutral policy, however, is not effective in inducing cultural assimilation and desegregation even when a very high maximum subsidy rate is offered. The means-tested ethnicity-neutral policy might even increase the unassimilated non-White British share. The means-tested ethnicity-conscious policy seems highly cost-effective, delivering similar assimilation and segregation outcomes as when a nearly 60 to 80% maximum subsidy rate is offered to everyone, which requires a very high budget-balancing tax rate.

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 or personal characteristics. For every policy parameter, White British people are expected to receive slightly higher subsidies than non-White British people. 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 rate to either 60% or 70% ($\kappa = 0.6$ or 0.7) for a means-tested policy would result in a similar subsidy level. For the welfare maximizing means-tested ethnicity-conscious policy setting the maximum subsidy rate at 90%, the average subsidy rate is 52% to 56% (Table 12), so it becomes slightly more generous than the current housing subsidy policy in the UK.

Figure 8: Simulated Outcomes Under Various Subsidy Policies



Note: 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.

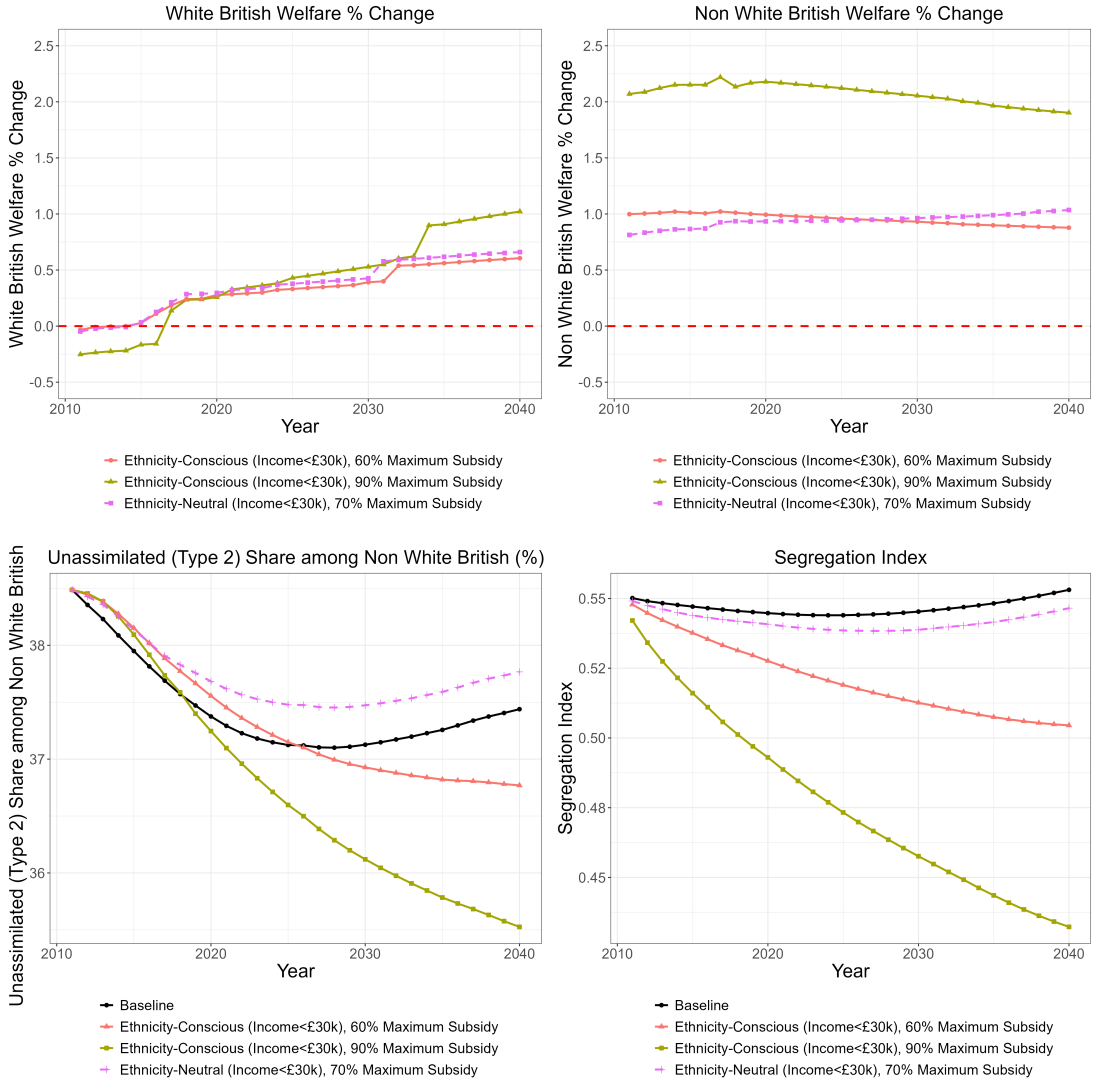
Table 12: Counterfactual Outcomes for 2011-2040 Under Various Policies

	Ethnicity-Conscious (Income<£30k) 90% Maximum Subsidy Rate	Ethnicity-Conscious (Income<£30k) 60% Maximum Subsidy Rate	Ethnicity-Neutral (Income<£30k) 70% Maximum Subsidy Rate
	Welfare-maximizing means-tested	Similar subsidy level to current policy	Similar subsidy level to current policy
Budget-Balancing Tax Rate (%)	0.05	0.02	0.02
Avg. WHB Welfare Change for 30 Years (%)	0.41	0.33	0.37
Avg. NWHB Welfare Change for 30 Years (%)	0.41	0.96	0.94
Unassimilated NWHB Share Change in 2040 (p.p., %)	-1.91	-0.67	0.33
	[-5.11%]	[-1.79%]	[0.88%]
Segregation Index Change in 2040 (%)	-0.12	-0.05	-0.01
	[-21.85%]	[-8.78%]	[-1.19%]
Avg. WHB Subsidy Rate (% Rent) for 30 Years	56.34	36.50	37.25
Avg. NWHB Subsidy Rate (% Rent) for 30 Years	52.35	32.40	30.33

Note: "WHB" means White British, "NWHB" means non-White British.

Table 12 summarizes a welfare-maximizing means-tested policy (column 1) and means-tested ethnicity-conscious and ethnicity-neutral policies with similar subsidy rates to the current housing subsidy policy in the UK (column 2 and 3). Under the welfare-maximizing means-tested policy, we can expect about 5% [1.91 p.p.] decline in the unassimilated non-White British share in the year 2040, and nearly 22% [0.12] decrease in the segregation index in the year 2040 compared to the baseline prediction in the absence of any intervention. When we consider a means-tested ethnicity-conscious policy with a similar subsidy level to the current amount, we can still expect about a 2% [0.67 p.p.] decrease in unassimilated non-White British share and about a 9% [0.05] drop in the segregation index in the year 2040. However, under a means-tested ethnicity-neutral policy inducing moving to neighborhoods with high amenities, the unassimilated non-White British share might increase by 1% [0.33 p.p.] compared to the baseline prediction, and the segregation index will drop only by 1% [0.01].

Figure 9: Prediction for 30 Years Under Various Policies



Note: 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.

Figure 9 shows 30-year predictions for policies considered in Table 12. From the first two top panels, it is notable that under the welfare-maximizing means-tested ethnicity-conscious policy, White British are expected to experience welfare loss on average for the first six years.[33] On the other hand, non-White British would not experience such welfare loss throughout the 30 years, and their welfare gain in percentage is much bigger than that of White British. The bottom two panels show that under ethnicity-conscious policies, ethnic minorities' cultural assimilation and neighborhood desegregation would monotonically improve over time

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

compared to baseline prediction in the absence of any policy intervention. However, under the ethnicity-neutral policy, we do not predict similar improvement in assimilation and desegregation.

To sum up, an ethnicity-conscious subsidy policy has the potential to induce sizable ethnic minorities' cultural assimilation and neighborhood desegregation, and such policy is welfare-improving on average (over 30 years) for both White British and non-White British. Replacing the current housing policy with an ethnicity-conscious one that provides a similar average subsidy would still lead to modest improvements in cultural assimilation and desegregation. However, the welfare-maximizing policy would allocate more generous subsidies, with a maximum of up to 90%

Our findings align with prior research (Chyn et al. (2025)): the Moving-to-Opportunity program in the U.S. had little effect on racial desegregation, whereas the race-based Gautreaux program significantly improved integration. This suggests that effective desegregation policies must account for ethnicity in housing policy design.

6 Conclusion

This paper estimates a dynamic general equilibrium model of neighborhood sorting and ethnic minorities' cultural assimilation using an 18-year panel, incorporating noisy cultural trait measurements. The model illustrates how ethnic minorities anticipating integration with the majority may invest in costly assimilation, while those remaining in ethnic enclaves have little incentive to do so. The estimation leverages repeated cultural measurements to track assimilation over time.

Counterfactual simulations suggest that an ethnicity-conscious, means-tested rent subsidy—offering larger subsidies for moves that promote ethnic mixing—is effective and cost-efficient. The optimal policy, with a 90% maximum subsidy rate, reduces the unassimilated non-White British share by 5% and the segregation index by 22% by 2040. A policy mirroring current UK housing subsidies (60% maximum rate) yields smaller gains: a 2% drop in the unassimilated share and a 9% reduction in segregation.

By contrast, an ethnicity-neutral policy, which incentivizes moves to high-amenity areas without considering ethnic composition, fails to promote assimilation or desegregation. This aligns with findings from the U.S., where the ethnicity-neutral Moving-to-Opportunity program had minimal impact on racial integration, whereas the ethnicity-conscious Gautreaux program significantly improved it (Chyn et al. (2025)).

This study does not capture all factors influencing integration due to data limitations. Notably, it omits the effects of intergroup contact on racial prejudice and the role of diversity in

public goods provision and social cohesion. Additionally, future research should explore how immigration policies can maximize diversity's benefits while mitigating integration costs.

References

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

Abramitzky, R., L. Boustan, K. Eriksson, and S. Hao (2020). Discrimination and the returns to cultural assimilation in the age of mass migration. In *AEA Papers and Proceedings*, Volume 110, pp. 340–346. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

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

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.

Agostinelli, F. and M. Wiswall (2025). Estimating the technology of children's skill formation. *Journal of Political Economy 133*(3), 000–000.

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 (2005). Ethnic diversity and economic performance. *Journal of economic literature 43*(3), 762–800.

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.

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.

Chetty, R. and N. Hendren (2018). The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. *The quarterly journal of economics 133*(3), 1107–1162.

Chetty, R., N. Hendren, and L. F. Katz (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review 106*(4), 855–902.

Chyn, E., R. Collinson, and D. H. Sandler (2025). The long-run effects of america's largest residential racial desegregation program: Gautreaux. *The Quarterly Journal of Economics*.

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

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.

Fernández, R. (2013). Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review 103*(1), 472–500.

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 L. Veldkamp (2011). Nature or nurture? learning and the geography of female labor force participation. *Econometrica 79*(4), 1103–1138.

Fouka, V. (2020). Backlash: The unintended effects of language prohibition in us schools after world war i. *The Review of Economic Studies 87*(1), 204–239.

Fouka, V. (2023). State policy and immigrant integration. *Annual Review of Political Science 27*.

Galiani, S., A. Murphy, and J. Pantano (2015). Estimating neighborhood choice models: Lessons from a housing assistance experiment. *American Economic Review 105*(11), 3385–3415.

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.

Hirsch, F. (2005). *Empire of nations: Ethnographic knowledge and the making of the Soviet Union*. Cornell University Press.

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*.

Joniak-Lüthi, A. (2013). Han migration to xinjiang uyghur autonomous region: between state schemes and migrants' strategies. *Zeitschrift für Ethnologie*, 155–174.

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.

Lundsteen, M. (2023). Displacing the other to unite the nation: The parallel society legislation in denmark. *European Urban and Regional Studies 30*(3), 261–281.

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.

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.

Paluck, E. L., S. A. Green, and D. P. Green (2019). The contact hypothesis re-evaluated. *Behavioural Public Policy 3*(2), 129–158.

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.

Pew Research Center (2017, November). Europe's Growing Muslim Population. Retrieved 27 May 2022.

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

Rothstein, R. (2017). *The color of law: A forgotten history of how our government segregated America*. Liveright Publishing.

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

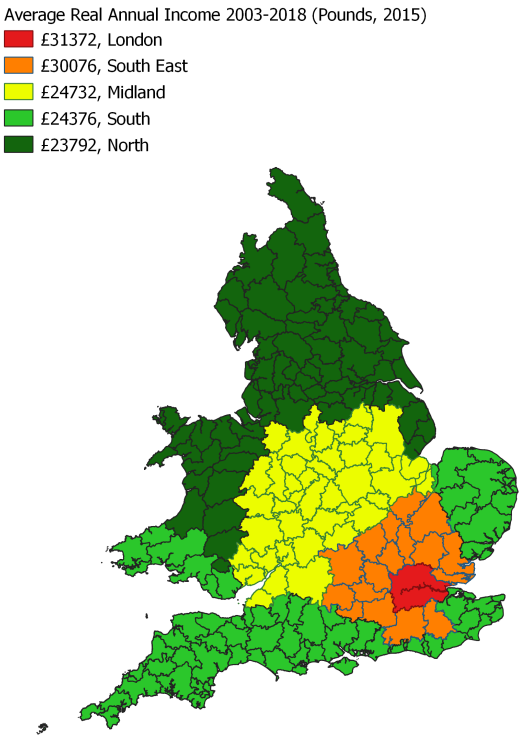
Source: Annual Survey of Hours and Earnings, 2003-2018
Note: Standard errors are clustered at the labor market level and are in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

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.

SUPPLEMENTAL APPENDIX For
"Designing an Ethnic Minorities' 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[34] to be consistent

[34] 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.

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 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* [35].

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.[36]

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

[35] 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.

[36] 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*[37] 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.[38] 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

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

[37] 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.

[38] 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 $f(i)$.[39] 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

[39] 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) [40]

- 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

[40] 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.[41] 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 G.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) explains, 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

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

enough to ensure that every neighborhood is included as a potential destination for another 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%.[42] The White British neighborhoods are defined as neighborhoods whose White British resident share exceeds 80.5%.[43] 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[44] 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 $hsqual_j$[45], 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

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

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

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

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

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 $controlfun_{jt}$, 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 the 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 :

$$\text{per-period likelihood} \quad L_{it} = L_{it}^{j} L_{it}^{P} L_{it}^{Y} \tag{C.8}$$

$$\text{likelihood from neighborhood choice} \quad L_{it}^{j} \equiv l(j_{it}|\Omega_{it}) \tag{C.9}$$

$$\text{likelihood from female work choice} \quad L_{it}^{P} \equiv l(P_{it}|\Omega_{it}) \tag{C.10}$$

$$\text{likelihood from proxies} \quad L_{it}^{Y} \equiv l(Y_{it}|u_{it}) \tag{C.11}$$

$$\text{transition likelihood} \quad L_{it}^{u} \equiv l(u_{it+1}|\Omega_{it}) \tag{C.12}$$

$$\tag{C.13}$$

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. [46] 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. [47] I assume the household moves to a different neighborhood $\tilde{j}_{it+1(2)}$ along the two choice paths[48]. States along the choice paths

[46] 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))

[47] 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.

[48] 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[49].

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

[49] 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 neighborhoods(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 £. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

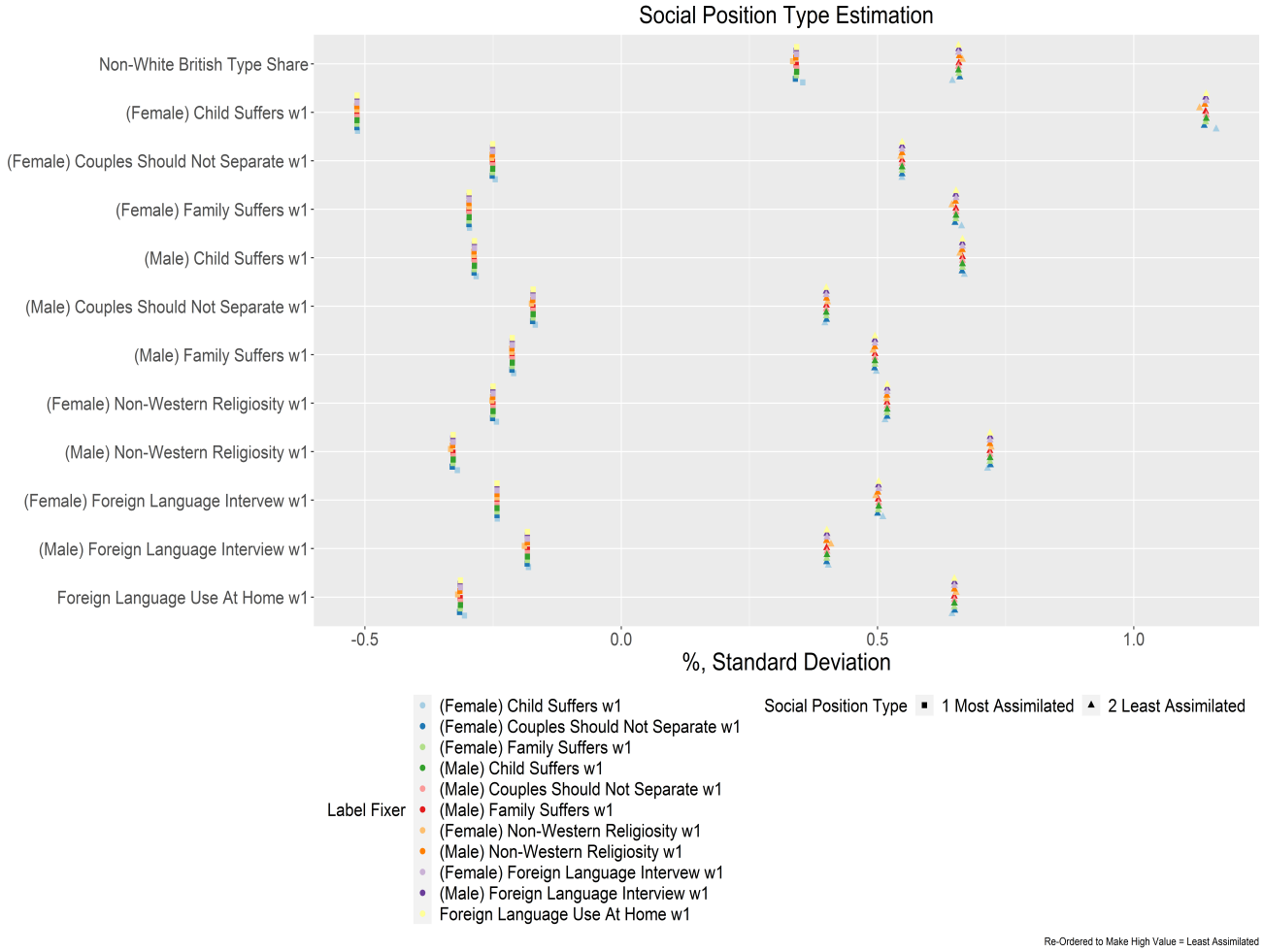
Table E.3: Main Parent (Male) Characteristics in the First Interview

	(by Female Spouse Ethnicity)	
Male Characteristics	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

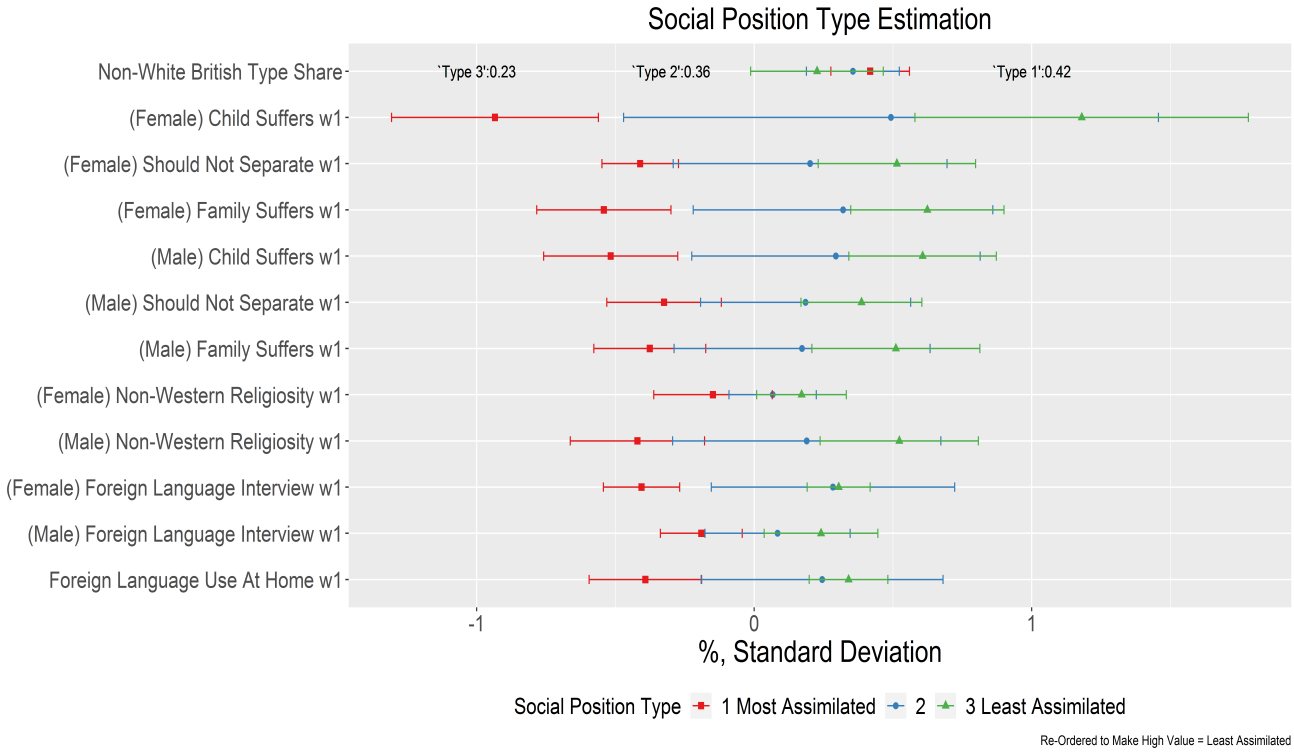
Note: This table shows the average and standard deviations of male sample characteristics in the first interview. The standard deviation of each variable is in parentheses. By construction, White British sample does not include any Muslims. I do not report the immigration-related statistics for the White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



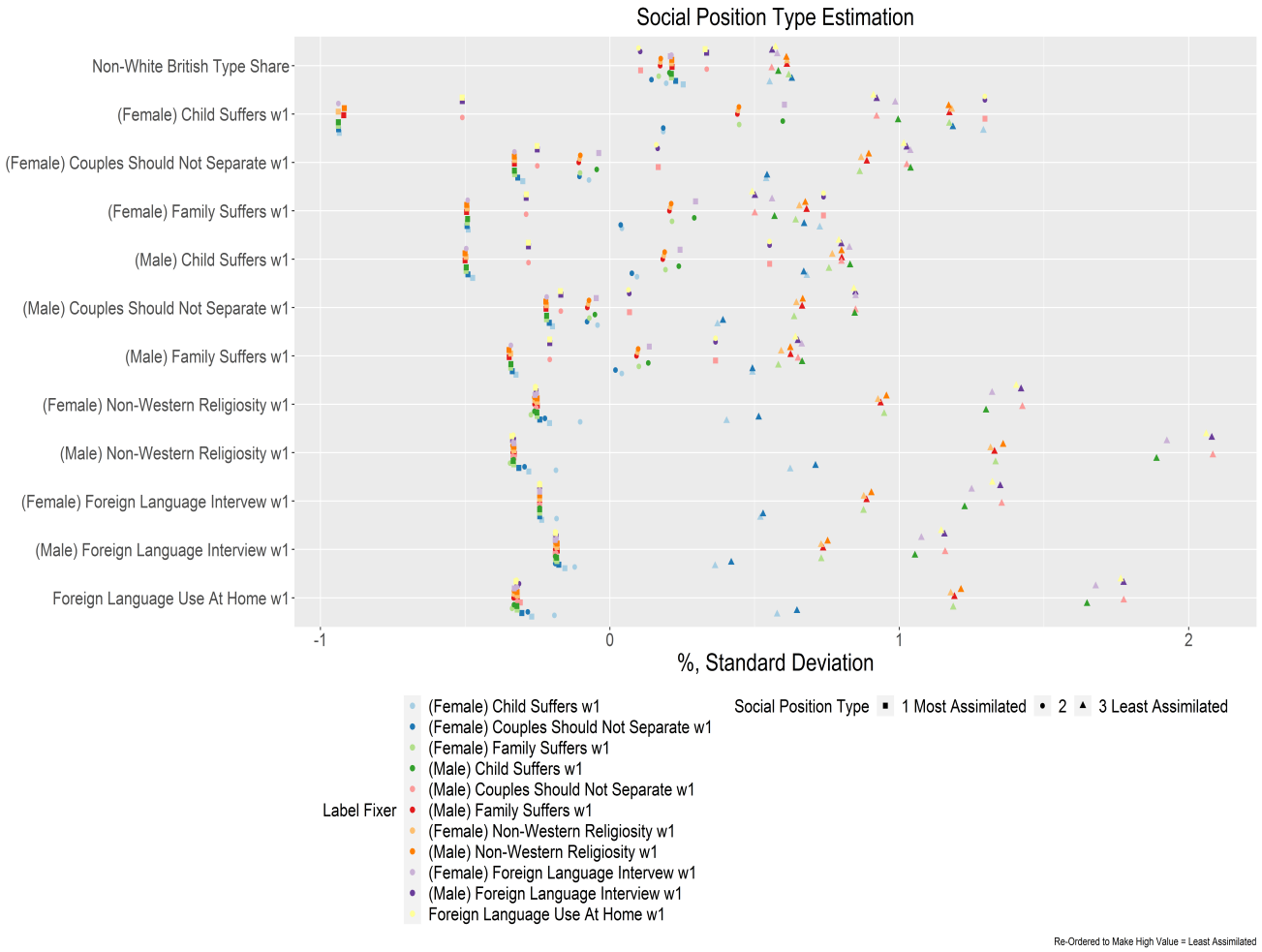
Source: MCS wave 1
Note: This figure shows the robustness of initial social position type estimates by different choices of an anchoring variable. I assumed two social position types.

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



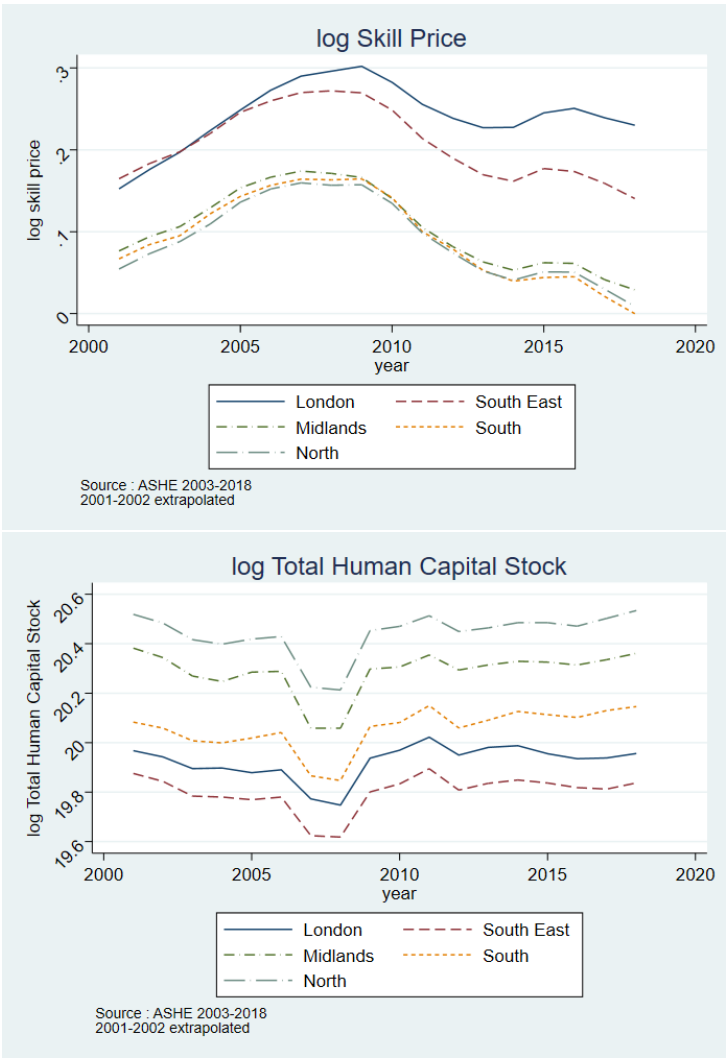
Source: MCS wave 1
Note: This figure shows the initial social position type share estimates and the average survey responses of each type, assuming three social position types. Whiskers show the 95% confidence intervals which were computed from bootstrapping 100 times.

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 Estimates for the Degree of Measurement Error in Social Position Proxies

This section presents the degree of measurement error in social position proxy variables. Specifically, I report estimates of the probability matrix $M^Y_{lk} = [P(Y = k \mid U = l)]$, where each (l, k)-th element represents the probability that proxy variable Y records a response k when the true latent variable has a value l. These estimates are obtained using the EM algorithm, as described in Section 4 Algorithm 1.

Below, I report the estimated probability matrices for different proxy variables, assuming two social position types. Since each proxy has a different response scale, the dimensions of these matrices vary. Variable names from the MCS dataset are provided in parentheses,[50] while probability estimates appear without parentheses and standard errors are shown inside parentheses.

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

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

[50] The interview language variable has changed names across survey waves, so it is not listed in parentheses.

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

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

G Labor Market Skill Price

Table G.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.

H Estimates for Counterfactual

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

H.1 Average Neighborhood Characteristics

Table H.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 H.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.

H.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 H.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.