Designing an Ethnic Minority Social Integration Policy

Yujung Hwang *

October 7, 2025

Abstract

Residential segregation and slow cultural assimilation among ethnic minorities pose pressing challenges. To design an effective social integration policy, I estimate a dynamic general equilibrium model of neighborhood choice and minorities' cultural assimilation using an 18-year micro panel dataset tracking neighborhood choices and measures of cultural assimilation. Both minorities and majorities select neighborhoods based on cultural similarity with neighbors, while minorities also decide on assimilation, with associated costs depending on their exposure to the majority. 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. During the preparation of this manuscript, the author used Grammarly version 14.1256.0, Chat GPT version 4o and 5 in order to improve clarity in writing. After using these tools, the author reviewed and edited the content as needed. All errors are my own.

1 Introduction

Residential segregation of minorities in Western societies—especially among Muslim communities—has drawn increasing attention. In the United Kingdom—the focus of this study—nearly half of Muslims (about 5% of the total population in 2011) lived in neighborhoods where at least 20% of residents were Muslim. Projections suggest the Muslim share could rise to 16.7% by 2050 (Pew Research Center, 2017), heightening concerns that persistent segregation may deepen social divisions.

This paper examines how ethnic minorities culturally assimilate and how this process interacts with neighborhood choice to inform non-coercive integration policy. 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. I develop a general equilibrium model of neighborhood choice and cultural assimilation using dynamic programming, estimating it with microdata from the United Kingdom, focusing on the period between 2001 and 2018. 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 that only incentivize moves to high-amenity neighborhoods 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. The proposed 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. Most prior studies treat unobserved heterogeneity in location preferences as time invariant (Kennan and Walker (2011); Bayer et al. (2016); Lessem (2018); Llull and Miller (2018); Piyapromdee (2021); Davis et al. (2021); Colas (2025)). A notable exception is Adda et al. (2022), which examines return migration to the country of origin and allows for stochastic, persistent location preferences alongside endogenous investments in country-specific human capital (e.g., language proficiency). In contrast, I study internal migration, in which neighborhood choice coevolves with cultural assimilation. 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 from 2001 to 2018 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 as-

similation, 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 a recent econometric approach for modeling unobserved heterogeneity using proxy variables (Hu (2017)). This method treats proxy variables as noisy measurements of a latent cultural assimilation type, estimating their relationships with the latent variable while accounting for measurement error. Then, the model is estimated using the Conditional Choice Probability (CCP) estimator Hwang (2025b), which extends Arcidiacono and Miller (2011) by incorporating proxy likelihoods. This approach extends the structural applied microeconomics literature on cultural evolution (Bisin et al. (2004), Bisin and Tura (2019)), which has traditionally focused on a single cultural dimension using proxy variables assumed to be free of measurement error, such as religion or interracial marriage. This paper examines cultural assimilation over the life cycle, while recognizing that intergenerational cultural transmission is also crucial for understanding ethnic minority integration. The latter is addressed in a separate study, Hwang (2025a).

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.

As emphasized by Fouka (2023), effective integration policy must account for equilibrium responses. I therefore compare general-equilibrium outcomes of two counterfactual desegregation policies: (i) an ethnicity-conscious subsidy whose amount depends on recipients' ethnicity and neighborhood composition—akin to Chicago's Gautreaux program (Chyn et al., 2025); and (ii) an ethnicity-neutral subsidy independent of ethnicity—akin to Moving to Opportunity (MTO) (Chetty and Hendren, 2018).

Simulating comparable UK policies, I find that ethnicity-conscious subsidies substantially

increase both desegregation and cultural assimilation, qualitatively consistent with evidence on Gautreaux and MTO (Chyn et al., 2025; Chetty and Hendren, 2018). 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% decrease in the segregation index (0.01). Importantly, means-tested ethnicity-conscious policies improve the welfare of both White British and non-White British residents on average. 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.

A decomposition of the counterfactuals underscores the importance of general-equilibrium feedbacks—through house prices, skill prices, and neighborhood composition. Treating these objects as exogenous overstates policy effects on welfare, assimilation, and desegregation. Endogenizing composition alone does not recover the general equilibrium benchmark; price feedbacks via housing and skills are pivotal as well. Transition dynamics further show that White British households experience modest short-run welfare losses under the proposed policies, but average gains accrue over the subsequent 30 years, yielding positive long-run effects. This pattern suggests policymakers should anticipate and mitigate short-run losses—for example, via temporary compensation or phased implementation.

The paper is structured as follows: Section 2 discusses the context and data. Section 3 presents the structural model. Section 4 explains the identification and estimation strategy. Section 5 discusses the estimation results. Section 6 analyzes counterfactual policies. Section 7 concludes.

2 UK Context and Data

This study focuses on England and Wales, where the White British population—the largest ethnic group—has experienced a steady decline, with its share falling from 87.5% in 2001 to 74.4% in 2021.[1] Over the same period, the share of non-White British residents—treated as ethnic minorities in this study—has increased, and residential segregation has drawn heightened public and policy attention. Muslims—predominantly of South Asian origin and comprising roughly 5% of the total population (UK Census 2011)—have been a particular focus of this debate. Their presence within the White British population is very small (0.2%; Office for National Statistics (2011)), and census data indicate substantial spatial concentration, with nearly half of Muslims living in neighborhoods where at least one in five residents is also Muslim. These demographic and spatial patterns have been interpreted in both academic and media discourse as potential sources of social division.

The UK provides unusually rich microdata for studying how spatial segregation interacts with cultural assimilation. Several nationally representative longitudinal surveys have oversampled non-White British households, enabling the study of cultural assimilation dynamics over extended horizons. The primary dataset used in this paper is the Millennium Cohort Study (MCS; University of London. Institute of Education. Centre for Longitudinal Studies. (2022)), which oversampled ethnic minority households, including approximately 500 Muslim households, and repeatedly collected information on cultural preferences and English language proficiency—that are key dimensions for measuring cultural assimilation. The MCS follows individuals born between 2000 and 2002 and their guardians (primarily biological parents). For this study, I primarily exploit the guardian survey. Importantly, the Secure Access version of the MCS includes detailed retrospective migration histories and precise geographic identifiers. I conduct my analysis at the level of Lower Super Output Areas (LSOAs)—small statistical units with an average population of about 1,500 residents, with 34,378 LSOAs in total in England and Wales[2]—allowing fine-grained measurement of residential segregation.

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

[2] Geographic boundaries are defined as of 2001.

Table A.1 in Online Appendix summarizes the selected measures of religiosity, gender attitudes[3], and language ability that serve as inputs to the estimator for the cultural assimilation index. Following Adida et al. (2016), these dimensions are widely recognized as salient indicators of assimilation in Muslim immigrant populations in Western Europe. I exclude name-based measures of assimilation (e.g., foreignness index; Fouka (2020)) because names are unavailable in the MCS and, in any case, change very little over an individual's life cycle, making them unsuitable for modeling cultural assimilation dynamics over the life-cycle.

The structural model developed in Section 3 focuses on married households, for which I construct a household-level cultural assimilation index by pooling spouses' survey responses. Single-person households are excluded to avoid modeling marriage entry and exit, which would introduce additional state variables and complicate the life-cycle problem. Roughly 13.5% of non-White British households in the sample are interethnic couples (Table A.9 in the Online Appendix). Rather than excluding these households—an approach that would reduce representativeness—I treat them as potentially highly assimilated. This assumption is operationalized by incorporating the survey responses of the White British spouse, which shifts the household's assimilation index toward a fully assimilated state. Among non-White British respondents, 69% are first-generation immigrants who, at the time of their first interview, had lived in the UK for a substantially long period, an average of 12 years; 53% identify as Muslim (Table 1). Given the modest size of the non-White British sample ($N = 1{,}234$), it is infeasible to study the heterogeneous groups within the non-White British respondents separately. Therefore, the model assumes separate structural parameters for White British/non-White British ethnicity, and assumes a common set of parameters for non-White British households, capturing heterogeneity through household-specific cultural assimilation index.

The analysis sample from the MCS consists of 2,468 households, which is insufficient to estimate general equilibrium effects in local housing or labor markets. To recover aggregate statistics required for counterfactual simulations, I supplement the MCS with additional administrative and survey data sources. In several parts of the analysis, I estimate household

[3]Among numerous gender attitude questions, I selected a subset of questions 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.

behavior conditional on state variables using the MCS data and subsequently reweight these estimates with auxiliary large-scale datasets to obtain representative aggregate statistics. Local labor market skill prices and total factor productivity are derived from administrative payroll records from the Annual Survey of Hours and Earnings (ASHE; Office for National Statistics (2022)). Annual income processes are estimated using the British Household Panel Survey (BHPS) and the UK Household Longitudinal Study (UKHLS; University of Essex, Institute for Social and Economic Research (2022)), as the MCS does not report annual income changes. I also use Census and Census Migration Flow data to measure neighborhood ethnic composition and migration distance distributions. Median house price data are taken from HM Land Registry (Office for National Statistics (2023)), and neighborhood-level deprivation indices from Abel et al. (2016).

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[4]. 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

[4]The survey does not include measures of majority adoption of minority cultural traits, but it does track assimilation into mainstream British culture.

assimilated) and Type 2 (least assimilated)[5].

The model incorporates several assumptions for tractability and identification. First, I assume that Type 1 individuals—the most assimilated non-White British—share the same social position as White British due to the identification challenge in the given data. 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. 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 A.4 and A.5). 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 – for example, for better neighborhood amenities – may find assimilation beneficial. The structure accommodates bi-directional causality between neighborhood sorting and cultural assimilation by treating each as a state variable in the other decision.

The model operates in general equilibrium, with three key endogenous aggregate vari-

[5]Rather than imposing Type 1 as an absorbing state, the extent of reverse assimilation is left to be determined by the data.

ables: 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.

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

Environment: Choices, Timing, Shocks

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 a *non-stationary* infinite time-horizon, $t \in \mathscr{T} = \{1, \cdots, \infty\}$, where inflows and outflows of the population are taken exogenously. I study an equilibrium under perfect foresight (for all future periods, including all subperiods) assumption, which I will elaborate on later.

I assume each household i 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. 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 $\Omega_{it(s)} = \{j_{it(s)}, u_{it(s)}, h_{it(s)}, a_{it(s)}, D_i, t(s)\}$.

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 $u_{it+1(1)}$[6]. 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 $\{\epsilon_{1it(1)}, \cdots, \epsilon_{Jit(1)}\}$, $\epsilon_{jit(1)} \overset{i.i.d.}{\sim} Gumbel(0,1)$[7] 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[8] 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.

First Subperiod: Location Choice

Households consider housing costs, neighborhood amenities, and moving costs when deciding on moving. Formally, the flow payoff from choosing location j' conditional on current location j is specified as

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

where γ_0^j is the housing price elasticity, γ_1^j is a vector of coefficients on the amenity vector, and $MC(j_{it(1)}, j')$ denotes the moving cost. Neighborhood amenities include (i) the negative log of the neighborhood deprivation index, which is a comprehensive index summarizing many aspects of the neighborhood, including school quality, (ii) the share of college graduates, (iii)

[6]Choosing the unobserved state one period ahead aligns the timing with the "Limited Feedback" assumption in Hu and Shum (2012), enabling use of their identification results.

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

[8]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.

a housing quality index, and (iv) an unobserved amenity component. School quality—often emphasized as a key driver of residential mobility—is incorporated via the deprivation index (Abel et al. (2016)). An explicit school quality index is excluded because available measures are not harmonized across England and Wales and therefore lack comparability.

Moving costs depend on the physical distance between origin and destination (e.g., people tend to move to a closer location), $log(distance(j_{it(1)}, j') + 1)$, the female spouse's (rescaled) age $\widetilde{Age}_{it} = (Age_{it} - 20)/10$[9] (e.g., young couples move more frequently) and the cultural distance between the household's social position $u_{it(1)}$ and the average social position of the destination neighborhood $\overline{u}_{j't}$. This cultural distance term, $|u_{it(1)} - \overline{u}_{j't}|$, captures households' preference for culturally similar neighbors and helps explain observed patterns of residential segregation.

$$MC(j_{it(1)}, j') = (C_0 + C_1 \widetilde{Age}_{it} + C_2(|u_{it(1)} - \overline{u}_{j't}|) \times 1(j_{it(1)} \neq j') + C_3 \; log(distance(j_{it(1)}, j') + 1). \quad (2)$$

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[10] 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. There-

[9] Age is rescaled so that the associated coefficient magnitudes are sufficiently large to be reliably estimated by the numerical optimizer.

[10] 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.

fore, 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

In the second subperiod, the household chooses female employment by weighing the gain in household income against the disutility of work, which varies with social position that reflects household gender norms. Household income is determined by the total human capital stock of working spouses multiplied by the local skill price per unit of human capital, $R_{lt(2)}$. Here, l denotes the labor market that neighborhood j belongs to. The flow payoff is given by the indirect utility from log household income plus the work disutility term, $\sum_{k=1}^{\overline{u}} \gamma_{1k}^P \mathbb{1}(u_{it(s)} = k)$, if the female spouse works. Social position-dependent disutility of work is consistent with evidence that more assimilated non-White British women have higher labor market participation rates (Table A.2).

$$\begin{aligned} \Pi^2(\Omega_{it(2)}; P^F_{it(2)}) \quad = \quad & \left\{ \gamma_0^P log(R_{lt(2)} \times (h_{it(2)} + h^{SP}_{it(2)})) + \sum_{k=1}^{\overline{u}} \gamma_{1k}^P 1(u_{it(2)} = k) \right\} P^F_{it(2)} + \\ & \gamma_0^P log(R_{lt(2)} \times h^{SP}_{it(2)})(1 - P^F_{it(2)}). \end{aligned} \tag{3}$$

I assume that individuals accumulate human capital through work experience while also facing stochastic shocks each period. The wage growth is assumed not to depend on social position, since I did not find the evidence that the social position of non-White British individuals influences the wage growth once individual fixed effect is accounted for (Table A.2). 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 5), 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. 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 (2011)), and it is a mixture of a stochastic random walk and a deterministic wage gain term from work experience. 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 (Equation 4), and the experience gain is accrued right after the second subperiod, following the female employment decision $P^F_{it(2)}$ (Equation 5). There is no change in human capital stock between the third subperiod in time t and the first subperiod in time $t+1$ (Equation 6).

$$logh_{it(2)} = logh_{it(1)} + \omega_{it(2)}, \quad \omega_{it(2)} \sim N(0, \sigma^2_\omega). \tag{4}$$

$$logh_{it(3)} = logh_{it(2)} + \lambda_0 P^F_{it(2)}. \tag{5}$$

$$logh_{it+1(1)} = logh_{it(3)}. \tag{6}$$

Third Subperiod: Next Period Social Position (Assimilation Status)

Cultural assimilation status may persist due to a fixed cost of changing social position, which can depend on the surrounding environment. Greater exposure to White British neighbors is assumed to influence this cost through horizontal cultural transmission, consistent with the mechanism proposed by Bisin and Verdier (2001). This channel is captured in the third-subperiod flow payoff, which consists solely of the psychic cost of changing social position:

$$\Pi^3(\Omega_{it(3)}; u') = PC_{u_{it(3)}u'} + PC^D_{u_{it(3)}u'} \overline{D}_{jt(3)}, \tag{7}$$

where $\overline{D}_{jt(3)}$ denotes the White British share in neighborhood j at time $t(3)$. By construction, this payoff is normalized to zero whenever the household's social position remains unchanged, i.e., $u_{it(3)} = u'$.

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.

$$V_{t(1)}(\Omega_{it(1)}) = E\left[\max_{j' \in \mathcal{N}} \left\{\Pi^1(\Omega_{it(1)}; j') + \epsilon_{j'it(1)} + E\left[V_{t(2)}(\Omega_{it(2)})|\Omega_{it(1)}, j'\right]\right\}\right], \tag{8}$$

$$V_{t(2)}(\Omega_{it(2)}) = E\left[\max_{p \in \{0,1\}} \left\{\Pi^2_j(\Omega_{it(2)}; p) + \eta_{it(2)}(1-p) + E\left[V_{t(3)}(\Omega_{it(3)})|\Omega_{it(2)}, p\right]\right\}\right], \tag{9}$$

$$V_{t(3)}(\Omega_{it(3)}) = E\left[\max_{u' \in \{1,\cdots,\overline{u}\}} \left\{\Pi^3(\Omega_{it(3)}; u') + \xi_{u'it(3)} + \beta E\left[V_{t+1(1)}(\Omega_{it+1(1)})|\Omega_{it(3)}, u'\right]\right\}\right]. \tag{10}$$

It is worth mentioning that the above value function has two-period finite dependence, which speeds up estimation later; see Online Appendix B.5 for further details. Therefore, the value function can be written as a function of the next two-period choice probabilities only. This implies that the perfect foresight assumption for all future periods (including all subperiods) reduces to the perfect foresight over only the next two periods.

Labor and Housing Market Clearing

It is important to account for endogenous market prices and neighborhood compositions to make credible counterfactual predictions, so this model endogenizes three equilibrium ob-

jects: (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 in the local labor market, $H_{lt(2)}$, that includes both working female's and male's. I assume perfect substitutability for human capital stock as an input; the perfect substitutability assumption is difficult to relax due to the data limitation[11]. The production function is written as

$$Y_{lt(2)} = A_{lt} H_{lt(2)}^{\zeta}. \tag{11}$$

$$H_{lt(2)} = \sum_{l_{it(2)}=l} \left\{ h_{it(2)} P_{it}^{F} + E\left[h_{it(2)}^{SP} | h_{it(2)} \right] \right\}. \tag{12}$$

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

$$R_{lt(2)} = A_{lt} \zeta H_{lt(2)}^{\zeta - 1}. \tag{13}$$

$$w_{lt(2)}(h_{it(2)}) = R_{lt(2)} h_{it(2)}. \tag{14}$$

Next, for the housing market, I assume the housing supply in each neighborhood, $\overline{HS}_j$, is perfectly inelastic. This assumption is a reasonable approximation in the UK context, where housing supply is well-documented to be highly inelastic relative to many other countries (Ball et al. (2010), Kim et al. (2012), Drayton et al. (2025)). In each neighborhood, the equilibrium rent r_{jt} is pinned down by the housing market-clearing condition (Equation 15), that the rents multiplied by the housing stocks equals the total household expenditure on rents. Here, TW_{jt} denotes total wages of residents in neighborhood j at time t, and for tractibility, I assume all

[11] 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.

households devote a fixed share sh of income to rent.

$$r_{jt} \times \overline{HS}_j = sh \times TW_{jt}. \tag{15}$$

The total wage in neighborhood j at time t is

$$TW_{jt} = \int_{j_{it}=j} \left(w_{it} P^F_{it} + E\left[w^{SP}_{it} | h_{it} \right] \right) f_t(\Omega_{it}) di * Pop_t. \tag{16}$$

$f(\Omega_{it})$ denotes the distribution of state variables, and Pop_t is the total population at time t. To impute the male spouse's earnings using the female spouse's human capital, I adopt the following linear function in equation 17[12].

$$E\left[log(h^{SP}_{it(2)}) \middle| h_{it(2)} \right] = b_1 + b_2 \ log h_{it(2)}. \tag{17}$$

Next, 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 (8, 9, 10) 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 13 holds.*

(iii) Rent vector $\{r_{jt(s)}\}$ *is determined to clear each neighborhood. That is, equation 15 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. That is, there might be more than one vector of endogenous aggregate state variables,

[12]Table 5 indicates that b_1 and b_2 are estimated with a reasonable precision.

$\{R_{lt(s)}, r_{jt(s)}, \overline{u}_{jt(s)}\}_{l\in\mathscr{L}, j\in\mathscr{N}, t\in\mathscr{T}}$, that satisfy the equilibrium conditions. Deriving an analytical uniqueness condition is difficult in a high-dimensional model like this[13]. 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 make an explicit equilibrium selection rule.

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, both segregation and the cultural assimilation of ethnic minorities are joint problems of White British and non-White British populations. Segregation can arise due to either

[13]While I do not prove uniqueness, Bayer and Timmins (2005) show via simulations that models with large choice sets, rich location heterogeneity, and heterogeneous preferences—features present here—typically have a unique sorting equilibrium.

group's reluctance to move into culturally dissimilar neighborhoods. Ethnic minorities are often solely blamed for segregation and a perceived lack of assimilation efforts. However, according to the model, segregation may result from the majority's high prejudice (captured by a high moving cost, C_2), leading to "White flight." Once ethnic minorities become isolated, the cost of cultural assimilation increases for them. Consequently, strong prejudice among the majority population can slow down the cultural assimilation of minorities.

The model omits some mechanisms due to data limits. First, it treats social-barrier moving costs C_2—and thus prejudice/homophily—as constant. Yet the "contact hypothesis" (Allport et al., 1954) implies that greater intergroup contact can lower prejudice, so C_2 could evolve with the social environment. Because high-quality prejudice measures are unavailable, I hold C_2 fixed here; with such data, the model could be extended to allow endogenous prejudice.

Another omitted mechanism is how segregation affects public goods provision and intergroup conflict, as documented in the political economy literature (Alesina et al., 1999; Alesina and La Ferrara, 2005). 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 Identification and Estimation Strategy

This section outlines the identification and estimation strategy; full details appear in Online Appendix B. I first reduce the neighborhood choice problem with 34,378 alternatives to a tractable location choice problem by drawing a stratified random choice set of 31 alternatives. I also consolidate local geographies into five labor markets by clustering official labor-market areas. I then describe the identification and estimation procedures. Aside from pre-set parameters, the structural parameters are estimated in three blocks—(i) income process parameters;

(ii) production function parameters; and (iii) flow-utility and other parameters—each under distinct identifying assumptions to be explained.

4.1 Choice Set And Labor Market Construction

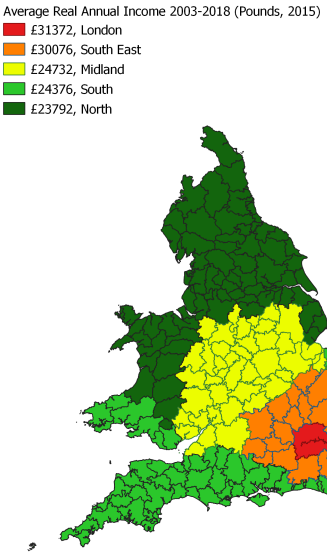
Estimating a migration model over the full choice set of 34,378 neighborhoods is computationally infeasible given the available sample size. Therefore, I follow the approach in Keane and Wasi (2016) and construct a stratified random choice set of size 31[14]. This choice set is sufficiently large to ensure that every neighborhood in the sample exhibits both in-migration and out-migration while remaining computationally tractable. To ensure diversity in potential destinations, neighborhoods are partitioned into three categories: (i) those with White British population density above the national average (80.5%), (ii) those with Muslim population density above 20%, and (iii) all remaining neighborhoods. From each category, 10 neighborhoods are randomly selected. Including the individual's current neighborhood yields a final choice set of 31 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[15]. First, I draw a migration distance at random from the empirical migration distance distribution in the 2001 Census Migration Flow data. Then, within each category, I select the neighbor-

[14] As noted in the Appendix of Keane and Wasi (2016), the use of a random choice set can introduce finite-sample bias in the presence of unobserved heterogeneity, which is the case here. However, there are several reasons why this concern is likely to be limited in the present context. First, the primary source of bias arises because the probability of drawing a particular choice set may vary across unobserved types—in this case, social position types. Since the same stratification scheme that balances the characteristics of alternative locations is applied uniformly across all households, the main channel for bias is the correlation between origin neighborhood characteristics and social position types. Under the chosen choice set size of 31, the origin neighborhood accounts for only about 3% of the available alternatives, limiting the impact of this correlation. Second, the same stratified random choice set is employed for both estimation and counterfactual simulation. Given that computational constraints prevent the use of the full choice set of 34,378 neighborhoods in simulation, the parameters corresponding to the full choice set are of limited practical relevance.

[15] The distance-based selection mechanism yields a better fit than a random, non-distance-based alternative.

Figure 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.

hood whose 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.

Next, 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. Online Appendix Figure 1 presents a map of these labor market classifications. Clustering is necessary because the ASHE dataset, which is 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 and their distance to three major metropolitan areas: London, Manchester, and Birmingham. This procedure yields five distinct local labor market groups.[16]

[16]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.

4.2 Income Process Parameters

I estimate the income process parameters $\{\sigma_\omega^2, \lambda_0\}$ – experience gain and permanent income shock variance – in addition to the variance of measurement error in observed wages, σ_m^2, separately by White British status; following standard income process literature, I account for potential measurement errors in observed wages. The identification and estimation closely follow Meghir and Pistaferri (2011), and Voena (2015), with one modification: I allow for endogenous local skill prices. Section B details this extension. 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).

The key identifying requirement is an instrument for female employment to correct selection into work via a Heckit-style procedure (following Voena (2015)). As instruments, I use whether the respondent's mother and the spouse's mother were employed when the respective child was age 14. The exclusion restriction is that these predetermined indicators are uncorrelated with contemporaneous income *shocks* (though not necessarily with income *levels*), while being strongly correlated with female employment through preferences for work. First-stage estimates in Table A.5 confirm that the instruments strongly predict female employment.

4.3 Production Function Parameters and Local TFPs

To estimate the production function parameters, I closely follow the methodology of Heckman et al. (1998). The identification strategy is based on a Cobb–Douglas production function, which can be expressed in log-linear form. To address the endogeneity of labor supply, I follow Bartik (1991) and Blanchard and Katz (1992) to construct a shift-share instrument, constructed from 2003 base-year industry composition shares using 9 industries—the first year with available geographic identifiers.[17] Because the ASHE data contain low-level geographic identifiers

[17] Recent work on shift–share instruments (Goldsmith-Pinkham et al. (2020); Borusyak et al. (2022)) clarifies that their validity relies on either the exogeneity of the base-year shares or the exogeneity of the shocks. Even if the base-year shares are not perfectly exogenous, a sufficiently large number of industries can render the instru-

only after 2002, I use the 2003–2018 sample to estimate local labor market skill prices and extrapolate values for 2001–2002 using a quartic polynomial. First-stage estimates, reported in Table A.6 in the Online Appendix, confirm that the instrument strongly predicts local human capital supply.

4.4 Flow Utility Parameters and Other Parameters

Remaining parameters, including flow utility parameters, are estimated via the two-step conditional choice probability (CCP) estimator developed in Hwang (2025b). Hwang (2025b) extends the Arcidiacono and Miller (2011) to utilize proxy variables for the unobserved state to gain identification for the flexible dynamics of an unobserved state, which corresponds to social position in this paper. The availability of repeated measurements on cultural assimilation is crucial to apply this methodology.

To improve computational efficiency, I estimate the measurement structure for proxy variables before the main estimation procedure using the EM algorithm, leveraging first-wave data, which contains all proxy variables (Table A.1). The measurement structure is expressed as the conditional probability that a proxy variable Y_{it} takes response k, given that the household's true social position type is l. I estimate these probabilities for each proxy variable and all possible values of k and l.

Identification of these conditional probabilities requires an additional assumption due to the well-known "label-swapping problem" (Hu (2017)).[18] To resolve this issue, I adopt the classic monotonicity assumption of Hu (2017): "*There exists an observable variable Y_{it} such that $E[Y_{it} \mid u_{it}]$ is increasing/decreasing in the social position u_{it}.*" In the data, English profi-

ment approximately exogenous, provided that industry-specific shocks are plausibly independent. In this paper, I use a moderately fine industry classification with nine categories. Because low-level geocodes are unavailable prior to 2003, my ability to test for pre-trends in labor supply by base-year shares is limited.

[18] For example, consider identifying a function $y = f(x)$ where x is a discrete latent variable. Without further restrictions, one could define another latent variable $\tilde{x}$ that simply reorders the values of x, and then construct a function $\tilde{f}(\tilde{x})$ such that $y = \tilde{f}(\tilde{x}) = f(x)$. Because $\tilde{f}(\cdot) \neq f(\cdot)$, the model is not identified unless we impose a restriction that anchors the labeling of x.

ciency is a strong proxy for cultural assimilation. I therefore use whether the female spouse was interviewed in a foreign language in wave 1 as the anchoring variable Y_{it}. To ensure robustness, I re-estimated the model using alternative anchoring variables and confirmed that the results are invariant to this choice (Online Appendix Figure A.3). Following the rationale explained in Section 3, the number of social position types, $\overline{u}$, is set to two.

Prior to the two-step CCP estimation, several parameters are preset as summarized in Online Appendix Table A.3. The time-discount factor is set to 0.95, following the migration literature (Kennan and Walker (2011); Bayer et al. (2016)). The rent-to-house price ratio is taken from Marsden (2015), based on 2014 rent and house price data. The expenditure share on housing for White British households is calibrated using UKHLS data from 2009–2018.

In addition to presetting parameters, I calibrate individuals' initial human capital stocks so that the only unobserved state in the CCP estimation is assimilation status. The production function and income process parameter estimates serve as the foundation for this calibration. See Online Appendix Section B.3 for details.

The CCP estimation proceeds in two steps. In the first step, the reduced-form CCPs and social position type probabilities are estimated via the EM algorithm in Hwang (2025b). After initialization, in the M-step, the per-period likelihood consists of three components: (1) neighborhood choice, (2) female employment decision, and (3) proxy variables measuring social position (assimilation status). (1) and (2) likelihoods are estimated by parametric multinomial logit estimation. The (3) likelihood is estimated beforehand when estimating the measurement structure of social position proxy variables. 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 and the per-period likelihoods. This iterative process for M- and E-step continues until convergence.

The (1) neighborhood choice likelihood estimation warrants further discussion. The neighborhood amenity vector in the model includes unobserved amenities, and it is crucial to control for this to avoid bias, especially in the estimated house price elasticity. If households systematically move toward wealthier areas and unobserved amenities are not controlled for, the

price coefficient may be biased upward, falsely implying a preference for higher prices. Also, if there is a systematic correlation between unobserved neighborhood amenities and neighborhood ethnic compositions, this would also bias the preference parameter for the cultural distance. I address this using a control function approach following Bayer et al. (2016); Online Appendix C provides details. Intuitively, the control function proxies for latent neighborhood quality that co-moves with prices. Guided by Bayer et al. (2016), I use as an instrument the predicted housing prices in nearby but non-adjacent neighborhoods—those 5–10 miles from j. These locations are close substitutes in the local housing market and shift equilibrium prices in the neighborhood j, while non-adjacency limits correlation with j's unobserved amenities, supporting the exclusion restriction. Operationally, I regress house prices on the instrument and other observable neighborhood characteristics and take the residual as the control function (Equation C.2 in Online Appendix), which is then included among the regressors in the CCP estimation to purge correlation between prices and unobserved amenities.

The second-stage estimation is conducted via quasi-maximum likelihood estimation (quasi-MLE), given the social position type probabilities from the first stage estimation. The likelihood function in this step leverages the two-period finite dependence property. See the Online Appendix B.5 for further details on the finite dependence property and detailed derivation of the likelihood function. The structural parameters estimated in this step are identified jointly based on the likelihood function; identification does not require selecting moments to match.

5 Estimation Results

5.1 Sample Selection And Descriptive Statistics

The main estimation sample is drawn from the parent surveys of the MCS, covering the years 2001–2018. Sample selection proceeds as follows. I include married couples[19] residing in England and Wales, where the female spouse is between ages 25 and 64. Only consecutive spells

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

Table 1: 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.

containing estimation-relevant information are retained, and spells shorter than ten years are excluded,[20] as longer observation windows are required for identification. All eligible non-White British households, defined based on the female spouse's ethnicity, are included, and an equal number of White British households are randomly sampled to balance the sample size. This restriction is necessitated by computational capacity limits on the secure data server.[21] The final estimation sample comprises 2,468 households (1,234 White British and 1,234 non-White British).

Table 1 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 Online Appendix (Table A.9). The average age for both White British and non-White British females is in the early 30s. By construction, the White British sample does not include any Muslims,[22] whereas nearly half of the non-White British sample identifies as Muslim. Ad-

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

[21] The Millennium Cohort Study Secure Access data are available only within a secure data environment.

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

Table 2: 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 More White British NBHD	0.562	0.683
	(0.496)	(0.466)
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)
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 5.1. 'WHB'=White British, 'NBHD'=Neighborhood.

ditionally, 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 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%.

Table 2 summarizes the migration patterns in our sample[23]. Attrition in the selected sam-

[23]The MCS survey is not conducted annually, but I construct a yearly migration history using retrospective questions. Specifically, I use the variable "year when Cohort Member moved to the current address." 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

Table 3: 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".

ple 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.

Migration rates are slightly higher among non-White British: 55% moved at least once during the sample period, compared with 50% of White British. Among movers, White British relocated more frequently on average (1.5 versus 1.3 moves). Both groups tended to move toward neighborhoods with a higher White British share (56% of White British moves; 68% of non-White British moves) and a lower Muslim share (55% and 68%, respectively). However, non-White British movers experienced larger compositional changes: their destinations had, on average, 13 percentage points more White British residents and 8.8 percentage points fewer Muslims than their origins, versus +2.4 and -0.8 percentage points for White British movers.

5.2 Structural Parameter Estimates

Table 3 shows the income process parameter estimates. 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.

actual moving year. Further details on missing data for the year of relocation are provided in Table A.7 in the Online Appendix.

Table 4: Production Function Estimates

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

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

Figure 2: Social Position Type Share Estimate

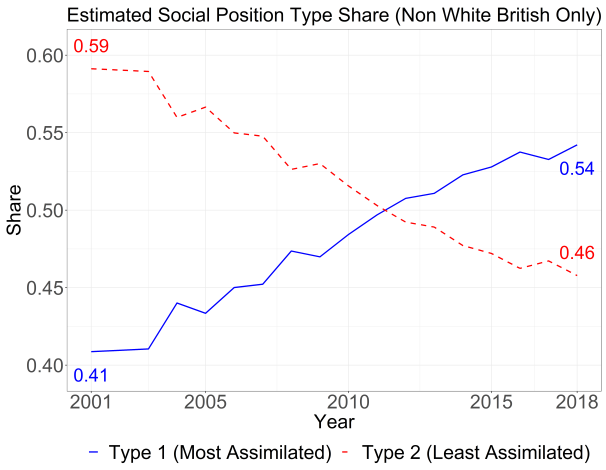


Table 4 presents the estimated production function curvature parameter and the average log skill price across the five local labor markets. London is the highest-paying labor market, followed by the South East region surrounding London. The Midlands, which includes Manchester and Birmingham, ranks third in pay, while the South and North rural areas offer the lowest wages.

Figure 2 plots the evolution of social-position type shares over the 18-year sample. The share of the least assimilated type (Type 2) among non-White British households declines from 59% in 2001 to 46% in 2018, indicating substantial assimilation. This change appears to be driven primarily by improvements in language proficiency; Figure A.1 in the Online Appendix shows marked gains in English proficiency but little change in gender attitudes or non-Western religiosity. Full estimates of the measurement structure for the proxy variables are reported in Online Appendix Section E.

Table A.4 in Online Appendix 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.

Table 5 presents the structural parameter estimates from the two-step CCP estimation. Many parameters are estimated with sufficient precision to be distinguishable from zero. 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 interdependence. 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.

Table 5: Structural Parameter Estimates

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)

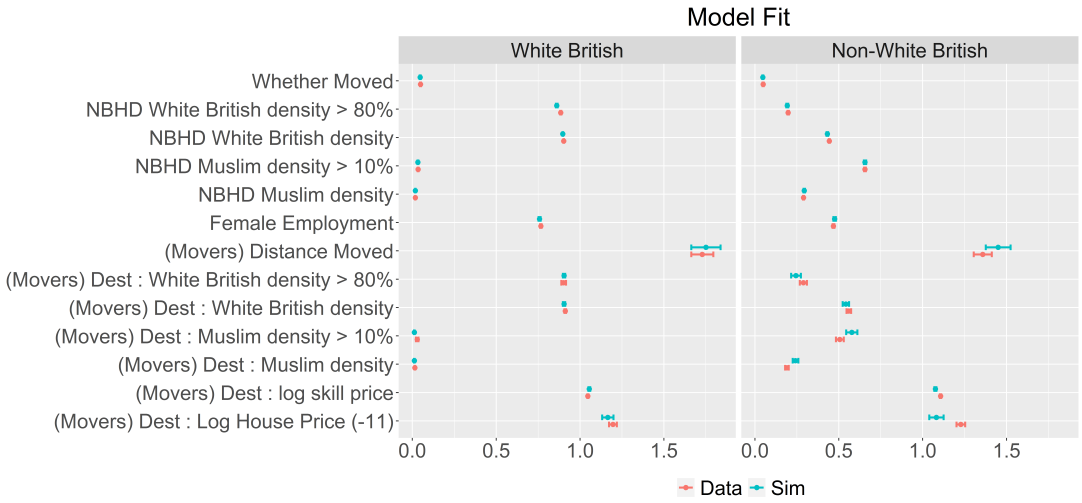
Source: Millennium Cohort Study, wave 1-7 (2001-2018)
Note: This table shows the structural parameter estimates. Standard errors are in parentheses.

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 sig-

nificantly 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.

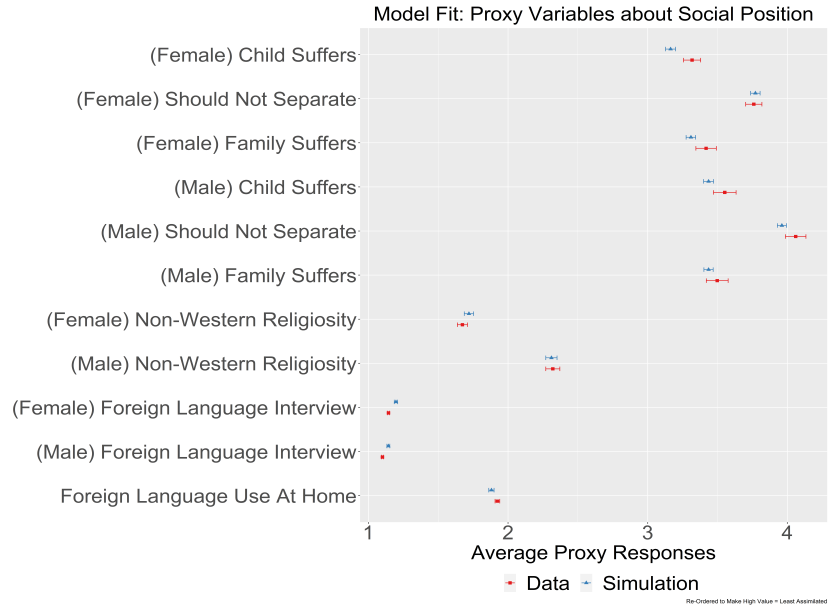
5.3 Model Fit

Figure 3: 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 4: Model Fit: Proxy Variables about Social Position



Source: Millennium Cohort Study, wave 1-7

The model matches key moments, including migration flows and female employment rates (Figure 3), as well as responses to the proxy variables that identify social position types disaggregated by gender (Figure 4). Figure 3 reports fit for the full sample and for movers (labeled "Movers"), showing close alignment between estimated and observed moments in 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.

6 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. 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. 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 (Lund-

steen (2023)) to promote their integration. In the United States, the Chicago Housing Authority introduced the Gautreaux Program, which provided Section 8 vouchers to help African American families relocate to affluent White neighborhoods to remedy past segregation (Chyn et al. (2025)).

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-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. Note that the subsidy schedule is continuous: all households are eligible regardless of whether they move, but the payment is small when the out-group (different-race) share in the current neighborhood is low.

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 A.6 in Online Appendix 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 variant in which only households with annual income below £30,000 qualify for housing subsidies. Under the subsidy schedule, eligible households receive an amount that varies with the neighborhood's out-group share. This income threshold broadly aligns with current UK housing subsidy policy (Adam et al. (2015)).

A few assumptions underpin this counterfactual analysis. First, while solving for an equilibrium, I start from the baseline aggregate states. As the model may exhibit multiple equilibria, this serves as an equilibrium selection rule. The solution process is iterative. Given aggregate states, I solve optimal individual behaviors and next, update the aggregate states to be consistent with individual behaviors. I repeat this process until convergence. Second, 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 ($r = (1/\beta) - 1 \approx 0.053$). Third, I assume people expect the current price to prevail forever when forming their dynamic expectation, but they update this expectation every year after observing new prices. This alternative expectation assumption was adopted instead of the perfect foresight assumption because of computational challenge[24]. Fourth, 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

[24]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.

Table 6: 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.

average population growth rate between the years 2011 and 2023[25]. Fifth, Total Factor Productivity (TFP) in the labor market and the housing stock is assumed to remain at the same level of year 2011.[26].

I evaluate 36 policies—ethnicity-conscious vs. ethnicity-neutral, each in non–means-tested and means-tested versions—across nine maximum subsidy caps (see Online Appendix Section D). Budget-balancing calculations show that non–means-tested designs are prohibitively costly (e.g., tax rates exceed 25% at a 90% cap; Online Appendix Figure D.1), so I restrict attention to means-tested designs. For a given cap, required tax rates are similar across ethnicity-conscious and ethnicity-neutral versions. I then (i) identify the welfare-maximizing means-tested design and (ii) select ethnicity-conscious and ethnicity-neutral variants that keep the average rent subsidy near the UK status quo, summarizing outcome changes under these rules

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

[26] 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.

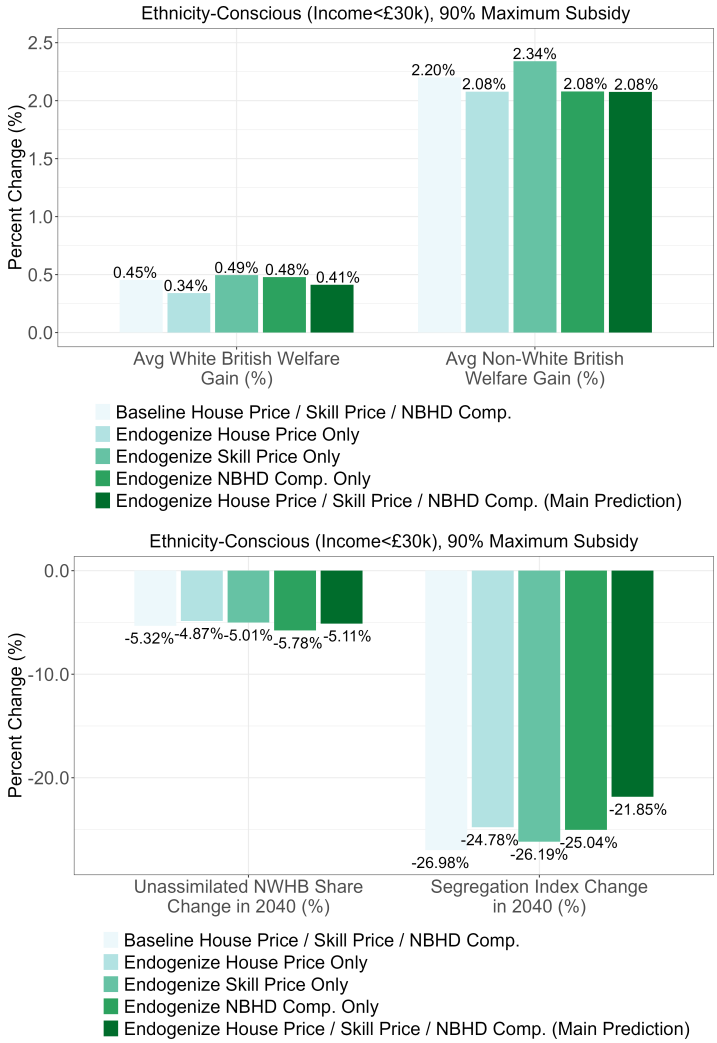
in Table 6.

The welfare-maximizing means-tested policy (Table 6, col. 1) is ethnicity-conscious with a 90% subsidy cap, implying an average subsidy rate of 52–56%. This slightly exceeds current UK averages—36% in London and 30% in England (Adam et al., 2015). Columns 2–3 report two means-tested, status-quo-level variants—ethnicity-conscious and ethnicity-neutral—obtained by setting the cap to 60% or 70%.

By 2040 (relative to the no-intervention baseline), the welfare-maximizing design reduces the unassimilated non-White British share by about 5% (1.91 p.p.) and the segregation index by roughly 22% (0.12). The status-quo-level ethnicity-conscious variant still delivers sizable gains—about a 2% (0.67 p.p.) decline in the unassimilated share and a 9% (0.05) drop in segregation. In contrast, the ethnicity-neutral variant that steers moves to high-amenity areas slightly increases the unassimilated share by 1% (0.33 p.p.) and reduces segregation by only 1% (0.01).

To assess the role of general-equilibrium feedbacks—via house prices, skill prices, and neighborhood composition (potentially reflecting "White flight")—I re-compute outcomes under the welfare-maximizing policy (means-tested, ethnicity-conscious, 90% cap) while selectively shutting down these channels. I compare the full model, in which all three aggregates are endogenous, with four variants: (i) all three fixed at baseline; and versions that endogenize only (ii) house prices, (iii) skill prices, or (iv) neighborhood composition. Figure 5 summarizes the results. Holding all aggregates fixed yields welfare effects similar to the full model but predicts a much larger decline in segregation (about 27% vs. 22%), suggesting that partial-equilibrium calculations overstate desegregation. Turning the channels on one at a time shows that endogenizing house prices reduces welfare gains for both White British and non-White British households, endogenizing skill prices increases them, and endogenizing neighborhood composition raises White British welfare but lowers non-White British welfare compared to (i). For assimilation and segregation, fixing aggregates at baseline consistently implies greater integration—larger declines in both the unassimilated share and the segregation index. Importantly, allowing neighborhood composition to adjust alone does not reconcile partial- and general-equilibrium predictions; price feedbacks also matter. Overall,

Figure 5: Decomposition of Counterfactual Predictions



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.

partial-equilibrium analyses tend to overstate the policy's welfare benefits and its effects on integration – assimilation, and segregation.

To study transitional dynamics, I compare 30-year projections for the policies in Table 6, reported in Online Appendix Figure A.7. The top panels show that, under the welfare-maximizing, means-tested, ethnicity-conscious policy, White British households experience average welfare losses for roughly the first six years. This pattern suggests a role for temporary compensation or phased implementation to mitigate short-run welfare loss. By contrast, non-White British households do not incur short-run welfare losses and realize substantially

larger percentage gains. The bottom panels indicate that ethnicity-conscious policies deliver monotonic improvements in cultural assimilation and neighborhood desegregation relative to the no-policy baseline, whereas the ethnicity-neutral policy does not yield comparable gains.

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 previous research using a different research design (Chyn et al. (2025)): while MTO had little effect on racial desegregation, the race-based Gautreaux program markedly improved integration—seen in shifts toward whiter neighborhoods and higher intermarriage rates. This suggests that designing an effective desegregation policy without considering ethnicity may be challenging.

7 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 mirror-

ing 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)).

It bears emphasis that ethnicity-conscious policies require careful communication if they are to be implemented. As Fouka (2020) cautions, state-led integration initiatives can provoke backlash and even hinder assimilation when they are perceived as signaling prejudice or group superiority. Before adopting such policies, governments should clearly articulate the rationale, demonstrate fairness and effectiveness, and build broad-based support that underscores national solidarity across identity groups.

Also, it is important to note that, due to data limitations, this study cannot incorporate all determinants of integration, leaving several aspects for future research. Notably, it omits the effects of intergroup contact on racial prejudice and the role of diversity in public-goods provision and social cohesion. If suitable data become available, incorporating these elements would be a valuable extension. An additional avenue is to study how immigration policies can be designed to maximize 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.

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

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

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

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

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

Allport, G. W., K. Clark, and T. F. Pettigrew (1954). *The nature of prejudice*, Volume 2. Addison-wesley Reading, MA.

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.

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?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

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.

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.

Bisin, A. and T. Verdier (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic theory 97*(2), 298–319.

Blanchard, O. J. and L. F. Katz (1992). Regional evolutions. *Brookings Papers on Economic Activity* (1), 1–61.

Borusyak, K., P. Hull, and X. Jaravel (2022). Quasi-experimental shift-share research designs. *The Review of Economic Studies 89*(1), 181–213.

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.

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*.

Colas, M. (2025). Dynamic responses to immigration. *Working Paper*.

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

Drayton, E., P. Levell, and D. Sturrock (2025). The determinants of local housing supply in england. *IFS Working Paper 25/04*.

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.

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*.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020). Bartik instruments: What, when, why, and how. *American Economic Review 110*(8), 2586–2624.

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

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

Hu, Y. (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. (2025a). Bounding omitted variable bias using auxiliary data: With an application to estimate neighborhood effects. *Available at SSRN 3866876*.

Hwang, Y. (2025b). Conditional choice probability estimation with an imperfectly measured latent state. *Available at SSRN 3535098, Conditionally Accepted at Quantitative Economics*.

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.

Keane, M. P. and N. Wasi (2016). How to model consumer heterogeneity? lessons from three case studies on sp and rp data. *Research in Economics 70*(2), 197–231.

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. In S. J. Smith, M. Elsinga, L. Fox O'Mahony, S. E. Ong, and S. Wachter (Eds.), *International Encyclopedia of Housing and Home*, pp. 10–15. Oxford: Elsevier.

Lessem, R. (2018). Mexico–us immigration: Effects of wages and border enforcement. *The Review of Economic Studies 85*(4), 2353–2388.

Llull, J. and R. A. Miller (2018). Internal migration and work experience in dual labor markets. *Unpublished Manuscript*.

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. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics*, pp. 105–142. New York: Academic Press.

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 (2011). Table dc2201ew - ethnic group and religion. United Kingdom Census. Retrieved 2016-01-14.

Office for National Statistics (2022). Annual Survey of Hours and Earnings, 1997-2022: Secure Access. [data collection]. 21st Edition.

Office for National Statistics (2023). Median house prices by lower layer super output area: Hpssa dataset 46. Accessed: 2023-07-06.

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.

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].

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].

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.

ONLINE APPENDIX For
"Designing an Ethnic Minority Social Integration Policy" (Yujung Hwang)

A Additional Tables and Figures

Table A.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.

Table A.2: 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. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

I construct these indices by summing related proxy variables (re-ordered if needed) and normalizing the summed index to have a mean of 0 and a standard deviation of 1. I count Hinduism, Islam, Sikhism, and Buddhism as non-Western religions.

Table A.3: Preset Parameters

Parameter	Description	Value	Source
β	Time Discount	0.95	Kennan and Walker (2011), 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

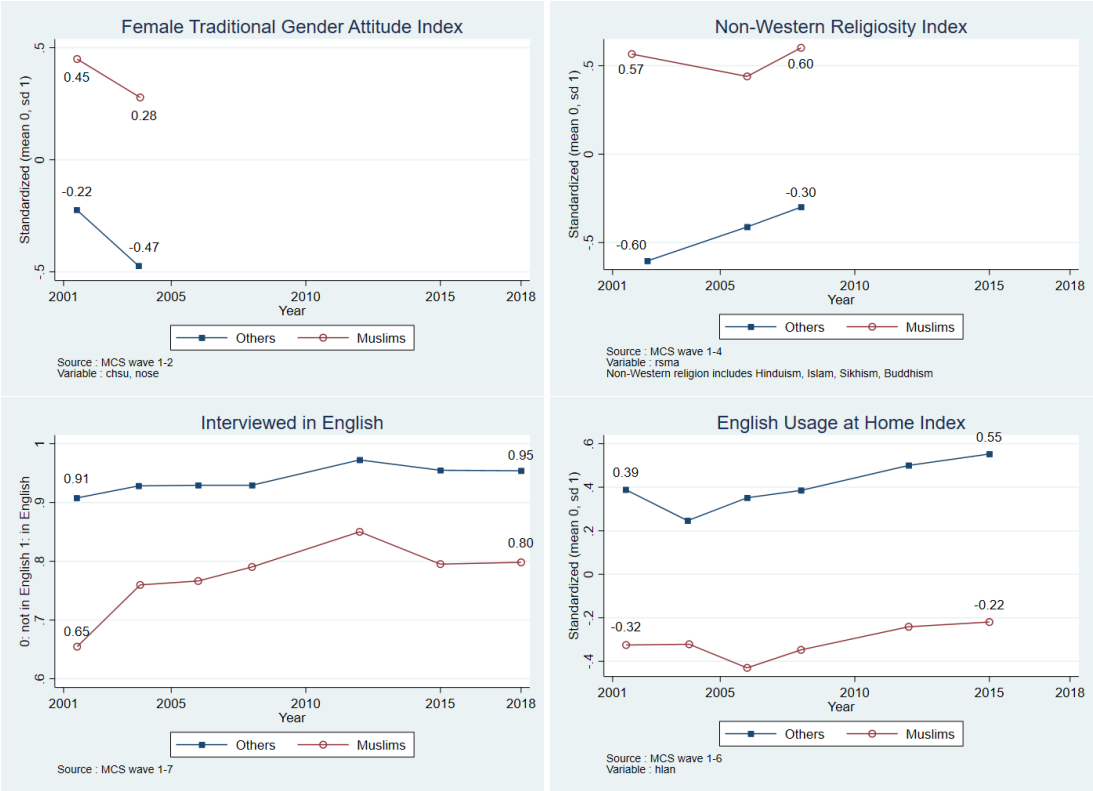
Note: Marsden (2015) used rent and housing price data from the year 2014. When computing the expenditure share on housing, I imputed rent for homeowners using the rent-to-house price ratio 0.037 reported in Marsden (2015). The housing expenditure did not include utilities.

Table A.4: 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.

Figure A.1: Evidence of Time-Varying Cultural Traits



Note: This figure shows the raw means of proxy variables for cultural traits from MCS data. 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 A.1). The gender attitude questions were asked only in the first two waves, and for male respondents, the responses to these questions in wave two are mostly missing. So, I only report the female traditional gender attitude index changes over time. For the non-Western religiosity index, I count Hinduism, Islam, Sikhism, and Buddhism as non-Western religions. I construct this index by using a variable on the frequency of religious meeting attendance. Note 'Others' is defined as non-White British, non-Muslims.

Table A.5: 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.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.6: 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.7: 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 A.8: 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

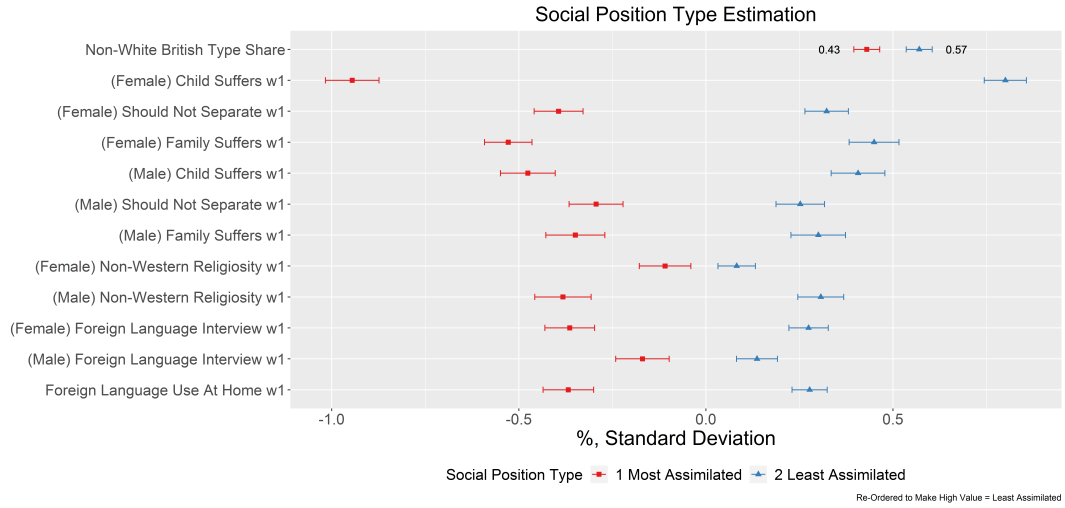
Note: Standard errors are clustered at the household level and are in parentheses. In every regression, a quadratic age polynomial, individual fixed effect, year fixed effect, and labor market fixed effect are controlled. The unit for labor income is 2015 £. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.9: Main Parent (Male) Characteristics in the First Interview

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

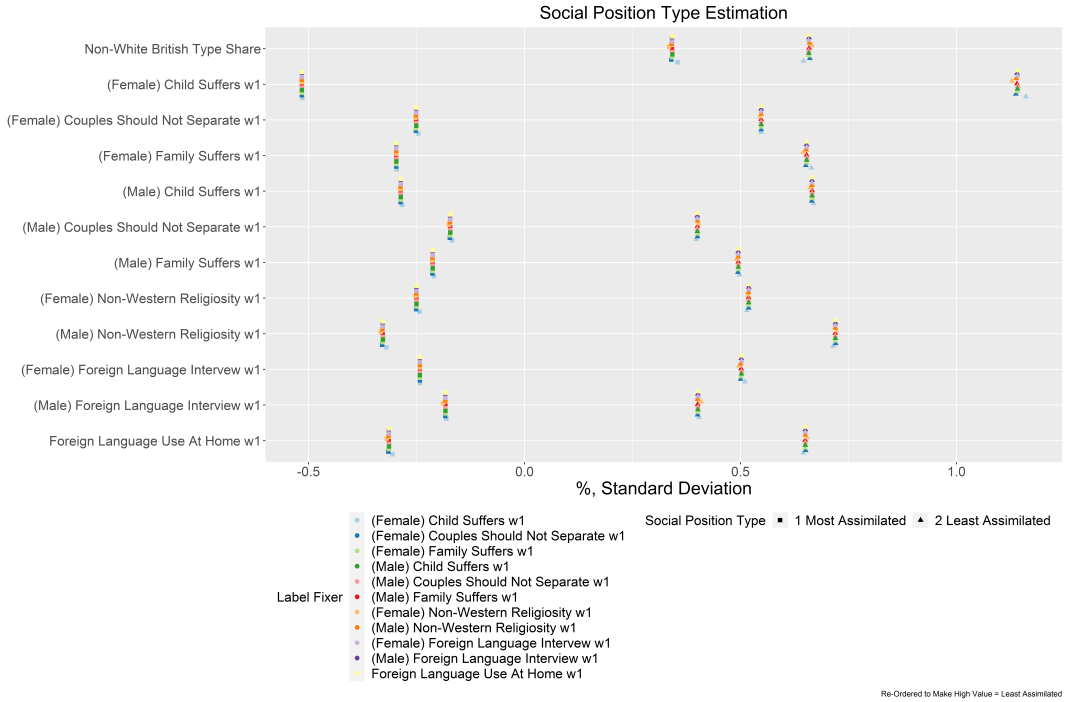
Source: MCS wave 1-7
Note: This table shows the average and standard deviations of male sample characteristics in the first interview. The standard deviation of each variable is in parentheses. By construction, White British sample does not include any Muslims. I do not report the immigration-related statistics for the White British sample to protect the confidentiality, as they are small in number.

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



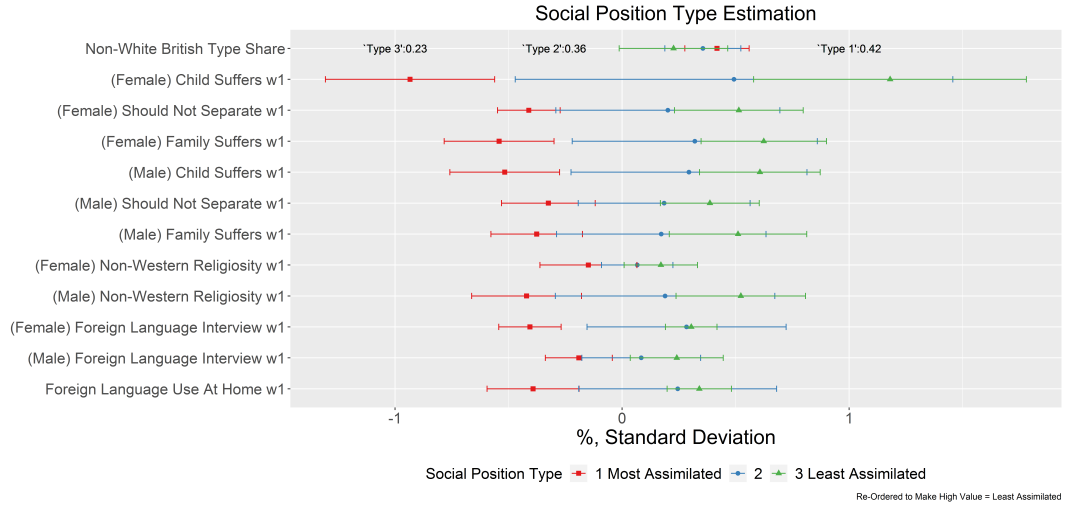
Source: MCS wave 1
Note: This figure shows the initial social position (assimilation status) type share estimates and the average survey responses of each type, assuming two types. Whiskers indicate the 95% confidence intervals, which were computed from bootstrapping 100 times.

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



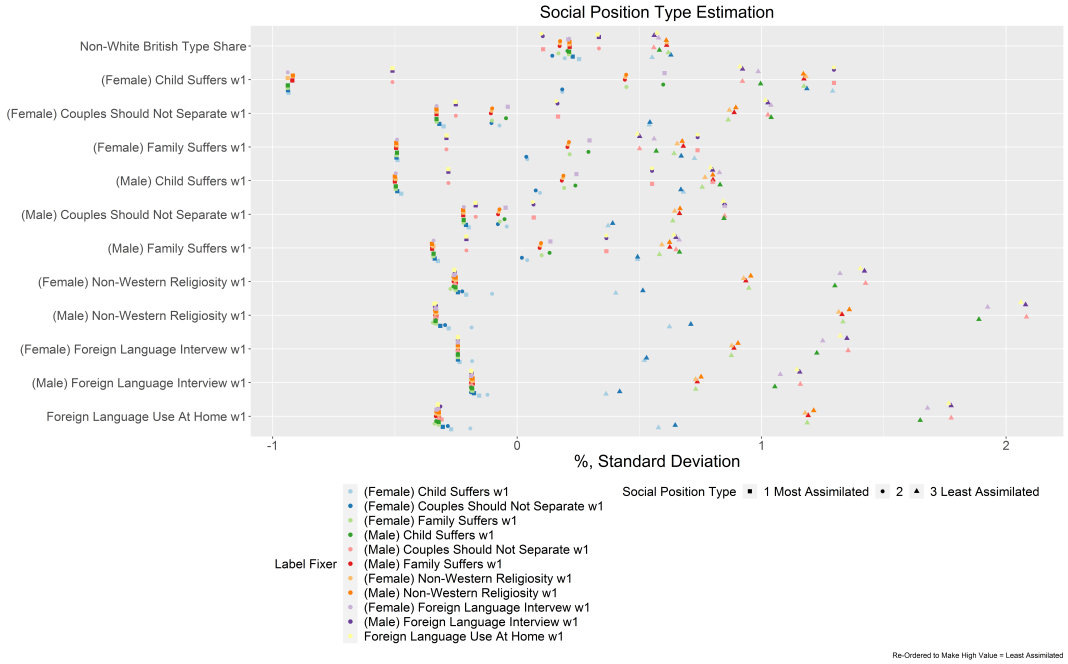
Source: MCS wave 1
Note: This figure shows the robustness of initial social position type estimates by different choices of an anchoring variable. I assumed two social position types.

Figure A.4: Initial Social Position Type Estimates (Assuming Three Types)



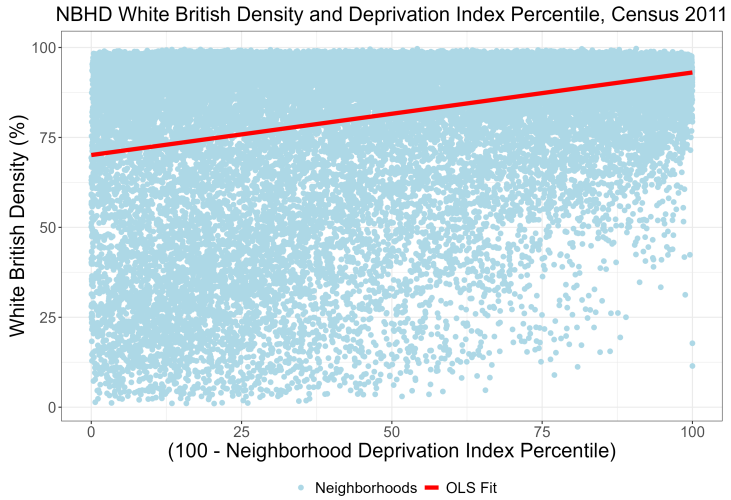
Source: MCS wave 1
Note: This figure shows the initial social position type share estimates and the average survey responses of each type, assuming three social position types. Whiskers show the 95% confidence intervals which were computed from bootstrapping 100 times.

Figure A.5: Robustness of Initial Social Position Type Estimates by Choice of a Different Anchoring Variable (Assuming Three Types)



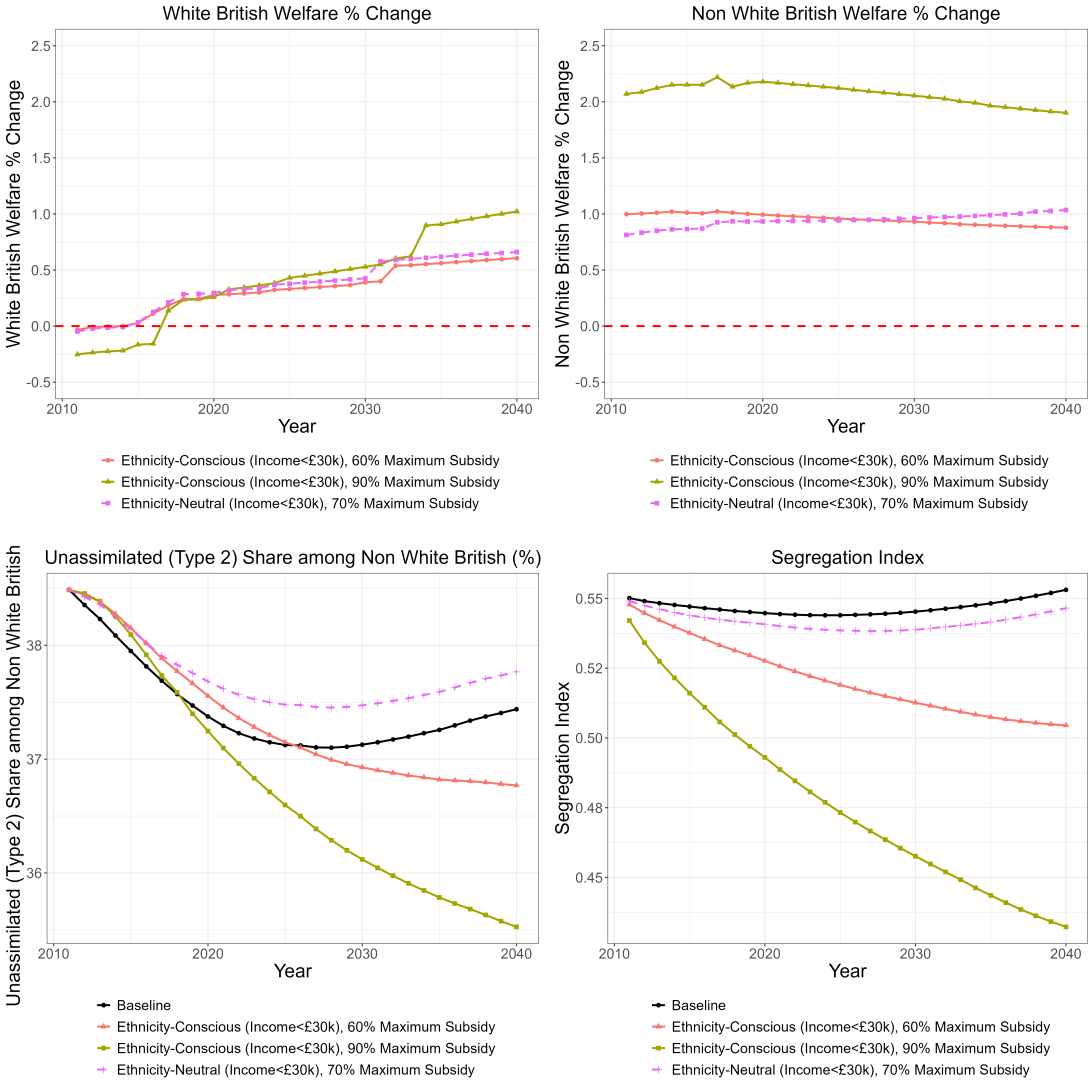
Source: MCS wave 1
Note: This figure shows the robustness of initial social position type estimates by diffferent choice of an anchoring variable. I assumed three social position types.

Figure A.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.

Figure A.7: 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.

B Details on the Estimation

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

B.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 (see references cited in Meghir and Pistaferri (2011)), 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* [27].

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.[28]

[27] 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.

[28] As explained in Meghir and Pistaferri (2011), 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

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

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

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

– 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*[29] 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.[30] Other exogenous regressors X_{it} include labor market X time dummy and quadratic age polynomials.

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

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

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

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

ϕ: PDF for standard normal distribution, Φ: CDF for standard normal

identification available is to use the bounding method, as discussed in Meghir and Pistaferri (2011).

[29] 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.

[30] 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.

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

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

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 $r_{it(2)}$,

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

, 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.

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

  * Run nonlinear least squares regression to estimate $\sigma^2_\omega, \rho_{\omega\eta}, \sigma^2_m$ using the fol-

lowing three-moment conditions. The derivation is presented below.

$$E\left[r_{it(2)}|P_{it}=1,P_{it-1}=1\right]=\sigma_{\omega}\rho_{\omega\eta}E\left[\lambda(\alpha_{it})\right].$$
$$E\left[r_{it(2)}^{2}|P_{it}=1,P_{it-1}=1\right]=\sigma_{\omega}^{2}\left(1+\rho_{\omega\eta}^{2}E\left[\alpha_{it}\lambda(\alpha_{it})\right]\right)+2\sigma_{m}^{2}.$$
$$E\left[r_{it(2)}(r_{it-1(2)}+r_{it(2)}+r_{it+1(2)})|P_{it+q}=1,q=-2,-1,0,1\right]$$
$$=\sigma_{\omega}^{2}\left(1+\rho_{\omega\eta}^{2}E\left[\alpha_{it}\lambda(\alpha_{it})\right]\right).$$

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.

$$\omega_{it(2)}\sim N(0,\sigma_{\omega}^{2}).$$
$$P(P_{it}=1)=P(X_{it}\beta+\eta_{it(2)}>0).$$
$$F(\eta_{it(2)})=\frac{1}{1+exp(-\eta_{it(2)})}.$$
$$(\Phi^{-1}(F(\omega_{it(2)})),\Phi^{-1}(F(\eta_{it(2)}))\sim N\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1&\rho_{\omega\eta}\\\rho_{\omega\eta}&1\end{pmatrix}\right).$$

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

Selection term :

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

$$= \sigma_\omega \rho_{\omega\eta} \frac{\phi(\Phi^{-1}(F(\alpha_{it})))}{1-\Phi(\Phi^{-1}(F(\alpha_{it})))} \quad (\because \text{truncated normal})$$
$$= \sigma_\omega \rho_{\omega\eta} \frac{\phi(\Phi^{-1}(F(\alpha_{it})))}{1-F(\alpha_{it})}.$$

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 $logm_{it(2)} \underset{i.i.d.}{\sim} (0, \sigma_m^2)$.

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 $\lambda(\alpha) \equiv \frac{\phi(\Phi^{-1}(F(\alpha)))}{1-F(\alpha)}$.

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

$$E\left[\omega^2_{it(2)} | \eta_{it(2)} > \alpha_{it(2)}\right] = E\left[\omega^2_{it(2)} | \Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it(2)}))\right]$$
$$= E\left[E\left[\omega^2_{it(2)} | \Phi^{-1}(F(\omega_{it(2)}))\right] | \Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it(2)}))\right]$$
$$= \sigma^2_\omega E\left[\Phi^{-1}(F(\omega_{it(2)}))^2 | \Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it(2)}))\right]$$
$$= \sigma^2_\omega E\left[E\left[\Phi^{-1}(F(\omega_{it(2)}))^2 | \Phi^{-1}(F(\eta_{it(2)}))\right] | \Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it(2)}))\right]$$

$$= \sigma_\omega^2 E\Big[Var\Big(\Phi^{-1}(F(\omega_{it(2)}))|\Phi^{-1}(F(\eta_{it(2)}))\Big) +$$
$$E\Big[\Phi^{-1}(F(\omega_{it(2)}))|\Phi^{-1}(F(\eta_{it(2)}))\Big]^2 \Big|\Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it(2)}))\Big]$$
$$= \sigma_\omega^2 \Big(1 - \rho_{\omega\eta}^2 + \rho_{\omega\eta}^2 E\left[\Phi^{-1}(F(\eta_{it(2)}))^2 | \Phi^{-1}(F(\eta_{it(2)})) > \Phi^{-1}(F(\alpha_{it(2)}))\right]\Big)$$
$$= \sigma_\omega^2 \left(1 + \rho_{\omega\eta}^2 E\left[\alpha_{it(2)} \frac{\phi(\Phi^{-1}(F(\alpha_{it(2)})))}{1 - F(\alpha_{it(2)})}\right]\right)$$
$$= \sigma_\omega^2 \Big(1 + \rho_{\omega\eta}^2 E\left[\alpha_{it(2)} \lambda(\alpha_{it(2)})\right]\Big).$$

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

$$E\left[r_{it(2)} | P_{it(2)} = 1, P_{it-1(2)} = 1\right] = \sigma_\omega \rho_{\omega\eta} E\left[\lambda(\alpha_{it(2)})\right].$$
$$E\left[r_{it(2)}^2 | P_{it(2)} = 1, P_{it-1(2)} = 1\right] = \sigma_\omega^2 \Big(1 + \rho_{\omega\eta}^2 E\left[\alpha_{it(2)} \lambda(\alpha_{it(2)})\right]\Big) + 2\sigma_m^2.$$
$$E\left[r_{it(2)}(r_{it-1(2)} + r_{it(2)} + r_{it+1(2)}) | P_{it+q(2)} = 1, q = -2, -1, 0, 1\right]$$
$$= \sigma_\omega^2 \Big(1 + \rho_{\omega\eta}^2 E\left[\alpha_{it} \lambda(\alpha_{it(2)})\right]\Big).$$

B.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 13). 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 $f(i)$.[31] 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 B.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 B.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.*

$$\Delta logw_{it(2)} = (logR_{l(j_{it})t} - logR_{l(j_{it})t-1}) + \tag{B.3}$$

[31] 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.

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

– 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 following relationship.

$$logR_{l(j_{it})t-1} - logR_{l(j_{it-1})t-1} = (logR_{l(j_{it})0} - logR_{l(j_{it-1})0}) \tag{B.4}$$
$$+ \underbrace{\sum_{q=2}^{t-1} \left(logR_{l(j_{it})q} - logR_{l(j_{it})q-1}\right)}_{\text{estimated in step 1.2}} - \underbrace{\sum_{q=2}^{t-1} \left(logR_{l(j_{it-1})q} - logR_{l(j_{it-1})q-1}\right)}_{\text{estimated in step 1.2}}.$$

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

$$logR_{l(j_{it})0} - logR_{l(j_{it-1})0} = (logR_{l(j_{it})0} - logR_{l_0 0}) - (logR_{l(j_{it-1})0} - logR_{l_0 0}),$$

where l_0 is a base labor market. I run a constrained linear regression using equation B.3 including *both* labor market stayers and switchers with $(L-1)^2+2$ constraints (L=4). [32]

– Step 1.4. Normalize the sequence $R_{l0}, \cdots, R_{lt}$ using $R_{l_0 0} = 1$ 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

[32] 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 B.4 are equal to one.

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 squares 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 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.[33] 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.

$$logR_{lt} = logA_{lt} + (\zeta - 1)logH_{lt} + log\ \zeta. \tag{B.5}$$

$logH_{l(j)t}$ 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,

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

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

The estimate table F.1 is in the Online Appendix.

B.3 Step 3-1. Calibrate Human Capital Stocks

Before conducting the CCP estimation, I calibrate individual human capital stocks so that the only unobserved state is social position. The production function and income process parameter estimates serve as the foundation for this calibration. 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 stock levels from observed wages. 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.

B.4 Step 3-2. 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

(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 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%.[34] The White British neighborhoods are defined as neighborhoods whose White British resident share exceeds 80.5%.[35] 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[36] 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

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

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

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

neighborhood college graduate share $log(ColEd_{jt}+0.1)$, neighborhood housing quality index $hsqual_j$[37], control function $controlfun_{jt}$, log distance between origin and destination $log(distance(j, j_{it(1)})+1)$, whether the neighborhood is different from the origin neighborhood $1(j \neq j_{it(1)})$, and the interaction between whether the neighborhood is different from the origin and neighborhood characteristics $1(j \neq j_{it(1)}) \times \{log(R_{jt}), log(hp_{jt}), log(WHB_{jt}+0.1), log(MUS_{jt}+0.1), -log(IMD_j), log(ColEd_{jt}+0.1), hsqual_j, controlfun_{jt}\}$. For White British households, the household characteristics include the female log human capital stock $log(h_{it(1)})$, female age Age_{it}, and the time index t. For non-White British households, I add a dummy variable whether the household is a least assimilated social position type $1(u_{it(1)} = \overline{u})$ to a household characteristics vector of White British. So in total, the number of regressors is 72 for White British households and 90 for non-White British households.

The female work CCP is also estimated separately for White British and non-White British sample. For White British households, the regressors are one vector for a constant, log of human capital stock $log(h_{it(2)})$, female age Age_{it}, log of median house price $log(hp_{jt})$, log of skill price in the labor market $log(R_{jt})$, log of neighborhood Muslim density, $log(MUS_{jt}+0.1)$, 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

[37] The construction of the housing quality index is explained in Section C.

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 B.3.
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.

$$P^{(m)}(u_{it} = s|\Omega_{it}) = \frac{L_{it}(u_{it} = s)}{\sum_{s'} L_{it}(u_{it} = s')}. \tag{B.6}$$

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 :

$$\text{per-period likelihood} \qquad L_{it} = L_{it}^{j} L_{it}^{P} L_{it}^{Y}. \tag{B.7}$$

$$\text{likelihood from neighborhood choice} \qquad L_{it}^{j} \equiv l(j_{it}|\Omega_{it}). \tag{B.8}$$

$$\text{likelihood from female work choice} \qquad L_{it}^{P} \equiv l(P_{it}|\Omega_{it}). \tag{B.9}$$

$$\text{likelihood from proxies} \qquad L_{it}^{Y} \equiv l(Y_{it}|u_{it}). \tag{B.10}$$

$$\text{transition likelihood} \qquad L_{it}^{u} \equiv l(u_{it+1}|\Omega_{it}). \tag{B.11}$$

$$\tag{B.12}$$

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

B.5 Step 3-3. 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. [38] 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

[38] 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)).

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. [39] I assume the household moves to a different neighborhood $\tilde{j}_{it+1(2)}$ along the two choice paths[40]. States along the choice paths 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.

$$p_{j't} \equiv P(j_{it(1)} = j' | \Omega^o_{it(1)}), \quad p_{p't} \equiv P(P_{it(2)} = p' | \Omega^o_{it(2)}). \tag{B.13}$$

$$p_{u't} \equiv P(u_{it+1(1)} = u' | \Omega^o_{it(3)}). \tag{B.14}$$

$$\psi_{j't} \equiv \gamma - log p_{j't}, \quad \psi_{p't} \equiv \sigma_p \gamma - \sigma_p log p_{p't}, \quad \psi_{u't} \equiv \sigma_u \gamma - \sigma_u log p_{u't}. \tag{B.15}$$

$\gamma \approx 0.5772$ 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*

[39] 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.

[40] So this condition also identifies the social distance moving cost parameter C_2. I choose $\tilde{j}_{it+1(2)}$ to be the $(J-1)$-th neighborhood in the choice set.

representations, equation B.16 - B.18, and B.24 - B.26, using the finite dependence path in Box 1.

$$p_{jt} = \frac{exp(\widetilde{v}_{jt(1)}(\Omega_{it(1)}))}{\sum exp(\widetilde{v}_{j't(1)}(\Omega_{it(1)}))}. \tag{B.16}$$

$$p_{pt} = \frac{exp(\widetilde{v}_{pt(2)}(\Omega_{it(2)})/\sigma_p)}{\sum exp(\widetilde{v}_{p't(2)}(\Omega_{it(2)})/\sigma_p)}. \tag{B.17}$$

$$p_{ut} = \frac{exp(\widetilde{v}_{ut(3)}(\Omega_{it(3)})/\sigma_u)}{\sum exp(\widetilde{v}_{u't(3)}(\Omega_{it(3)})/\sigma_u)}. \tag{B.18}$$

where

$$\widetilde{v}_{jt(1)}(\Omega_{it(1)}) = E\Bigg[\Pi^1_j(\Omega_{it(1)}; j) + \gamma^P_1 log(R_{lt}) + \psi^{(1)}_{pt} \tag{B.19}$$

$$+\psi^{(1)}_{ut} + \beta\{MC(j, j_{it+1(1)}) + \psi^{(1)}_{jt+1}\}\Bigg]. \tag{B.20}$$

$$\widetilde{v}_{pt(2)}(\Omega_{it(2)}) = E\Bigg[\Pi^2_j(\Omega_{it(2)}; p) + \psi^{(2)}_{ut} + \beta\{\psi^{(2)}_{jt+1} + \Pi^2_j(\Omega_{it+1(2)}; 1-p) + \psi^{(2)}_{pt+1}\}\Bigg]. \tag{B.21}$$

$$\widetilde{v}_{ut(3)}(\Omega_{it(3)}) = E\Bigg[\Pi^3_u(\Omega_{it(3)}; u) + \beta\{C_2(|u - \overline{u}_{jt+1(1)}|) \times 1(j_{it+1(1)} \neq j_{it(3)})$$

$$+\psi^{(3)}_{jt+1} + \gamma^P_{1u} P^F_{it+1(2)} + \psi^{(3)}_{pt+1} + \Pi^3_u(\Omega_{it+1(3)}; u_{it+2(1)}) + \psi^{(3)}_{ut+1}\}\Bigg]. \tag{B.22}$$

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:

$$v_{jt(1)}(\Omega_{it(1)}) = E\Bigg[\Pi^1_j(\Omega_{it(1)}; j) + \gamma^P_1 log(R_{lt}) + \psi^{(1)}_{pt}$$

$$+\psi^{(1)}_{ut} + \beta\{MC(j, j_{it+1(1)}) + \psi^{(1)}_{jt+1} + V_{t+1(2)}(\Omega_{it+1(2)})\}\Bigg].$$

$$v_{pt(2)}(\Omega_{it(2)}) = E\Bigg[\Pi^2_j(\Omega_{it(2)}; p) + \psi^{(2)}_{ut} + \beta\{\psi^{(2)}_{jt+1} + \Pi^2_j(\Omega_{it+1(2)}; 1-p) + \psi^{(2)}_{pt+1} + V_{t+1(3)}(\Omega_{it+1(3)})\}\Bigg].$$

$$v_{ut(3)}(\Omega_{it(3)}) = E\Bigg[\Pi^3_u(\Omega_{it(3)}; u) + \beta\{C_2(|u - \overline{u}_{jt+1(1)}|) \times 1(j_{it+1(1)} \neq j_{it(3)})$$

$$+\psi^{(3)}_{jt+1} + \gamma^P_{1u} P^F_{it+1(2)} + \psi^{(3)}_{pt+1} + \Pi^3_u(\Omega_{it+1(3)}; u_{it+2(1)}) + \psi^{(3)}_{ut+1} + \beta V_{t+2(1)}(\Omega_{it+2(1)})\}\Bigg].$$

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 gen-

der norm cost for working female in a term $\Pi_j^2(\Omega_{it+1(2)}; P^F_{it+1(2)} = 1)$ 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.

$$\widehat{v}_{jt(1)}(\Omega_{it(1)}) = E\Bigg[\widehat{\Pi}^1_j(\Omega_{it(1)}; j) + \widehat{\gamma}^P_1 log(\widehat{R}_{lt}) + \widehat{\psi}^{(1)}_{pt} \tag{B.23}$$

$$+\widehat{\psi}^{(1)}_{ut} + \beta\{\widehat{MC}(j, j_{it+1(1)}) + \widehat{\psi}^{(1)}_{jt+1}\}\Bigg]. \tag{B.24}$$

$$\widehat{v}_{pt(2)}(\Omega_{it(2)}) = E\Bigg[\widehat{\Pi}^2_j(\Omega_{it(2)}; p) + \widehat{\psi}^{(2)}_{ut} + \beta\{\widehat{\psi}^{(2)}_{jt+1} + \widehat{\Pi}^2_j(\Omega_{it+1(2)}; 1-p) + \widehat{\psi}^{(2)}_{pt+1}\}\Bigg]. \tag{B.25}$$

$$\widehat{v}_{ut(3)}(\Omega_{it(3)}) = E\Bigg[\widehat{\Pi}^3_u(\Omega_{it(3)}; u) + \beta\{\widehat{C}_2(|u - \widehat{\overline{u}}_{jt+1(1)}|) \times 1(j_{it+1(1)} \neq j_{it(3)})$$

$$+\widehat{\psi}^{(3)}_{jt+1} + \widehat{\gamma}^P_{1u} P^F_{it+1(2)} + \widehat{\psi}^{(3)}_{pt+1} + \widehat{\Pi}^3_u(\Omega_{it+1(3)}; u_{it+2(1)}) + \widehat{\psi}^{(3)}_{ut+1}\}\Bigg]. \tag{B.26}$$

To estimate the flow utility $\widehat{\Pi}^1_j(\Omega_{it(1)}; j)$ 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,
$N_{it(s)} = \{1, log(R_{lt}), log(hp_{jt}), log(WHB_{jt} + 0.1), log(MUS_{jt} + 0.1),$
$log(-IMD_j), log(ColEd_{jt}), hsqual_j, controlfun_{jt}\}$.

$$\text{individual characteristics regression} \qquad u_{it(s)} = N_{it(s)}\beta_1 + \epsilon_{it(s)}.$$

$$\text{neighborhood average prediction} \qquad \widehat{u}_{jt(s)} = N_{it(s)}\widehat{\beta}_1.$$

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.

$$p_{jt} = \frac{exp(\hat{v}_{jt(1)}(\Omega_{it(1)}))}{\sum exp(\hat{v}_{j't(1)}(\Omega_{it(1)}))}. \tag{B.27}$$

$$p_{pt} = \frac{exp(\hat{v}_{pt(2)}(\Omega_{it(2)})/\sigma_p)}{\sum exp(\hat{v}_{p't(2)}(\Omega_{it(2)})/\sigma_p)}. \tag{B.28}$$

$$p_{ut} = \frac{exp(\hat{v}_{ut(3)}(\Omega_{it(3)})/\sigma_u)}{\sum exp(\hat{v}_{u't(3)}(\Omega_{it(3)})/\sigma_u)}. \tag{B.29}$$

C 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[41].

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

Table C.1 shows the first stage regression results from the regression equation C.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, col-

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

Table C.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 parentheses. 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.
$^{*}p < 0.05$, $^{*}\dot{p} < 0.01$, $^{*}\ddot{p} < 0.001$

lege 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 $\widehat{\beta}_1 \times \text{NBHD House Char}_{jt}$ for 19 years, so I run a slightly modified regression equation C.2 to compute the control function residual, $\text{ControlFun}_{jt} = \widehat{\epsilon}_{jt}$.

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

D More Details on the Counterfactual Analysis

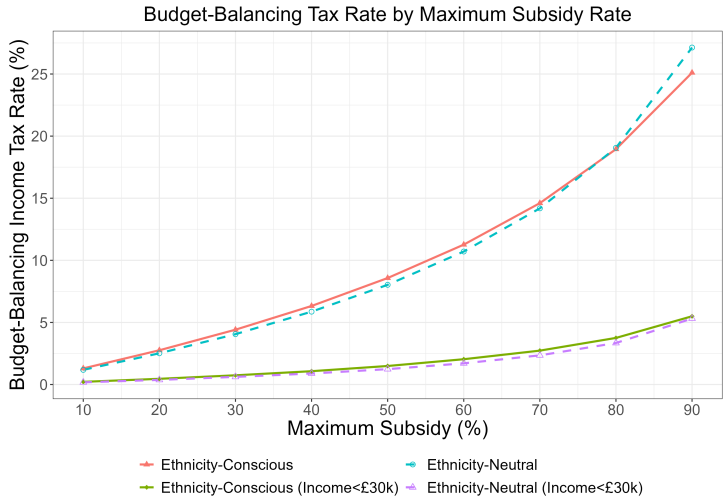
This section reports counterfactual outcomes for 36 policies—nine subsidy caps crossed with non–means-tested vs. means-tested and ethnicity-conscious vs. ethnicity-neutral designs.

Figure D.1 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 budget-balancing tax rate increases at a much steeper rate when the maximum subsidy rate exceeds 70%.

Figure D.2 shows the simulated outcomes under various subsidy policies. The top two panels show the average welfare changes[42] 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 on average stemming from changing market prices

[42] Welfare is measured by the value function $V_{t(2)}(\Omega_{it(2)})$.

Figure D.1: Budget-Balancing Linear Income Tax Rate for Different Subsidy Policies



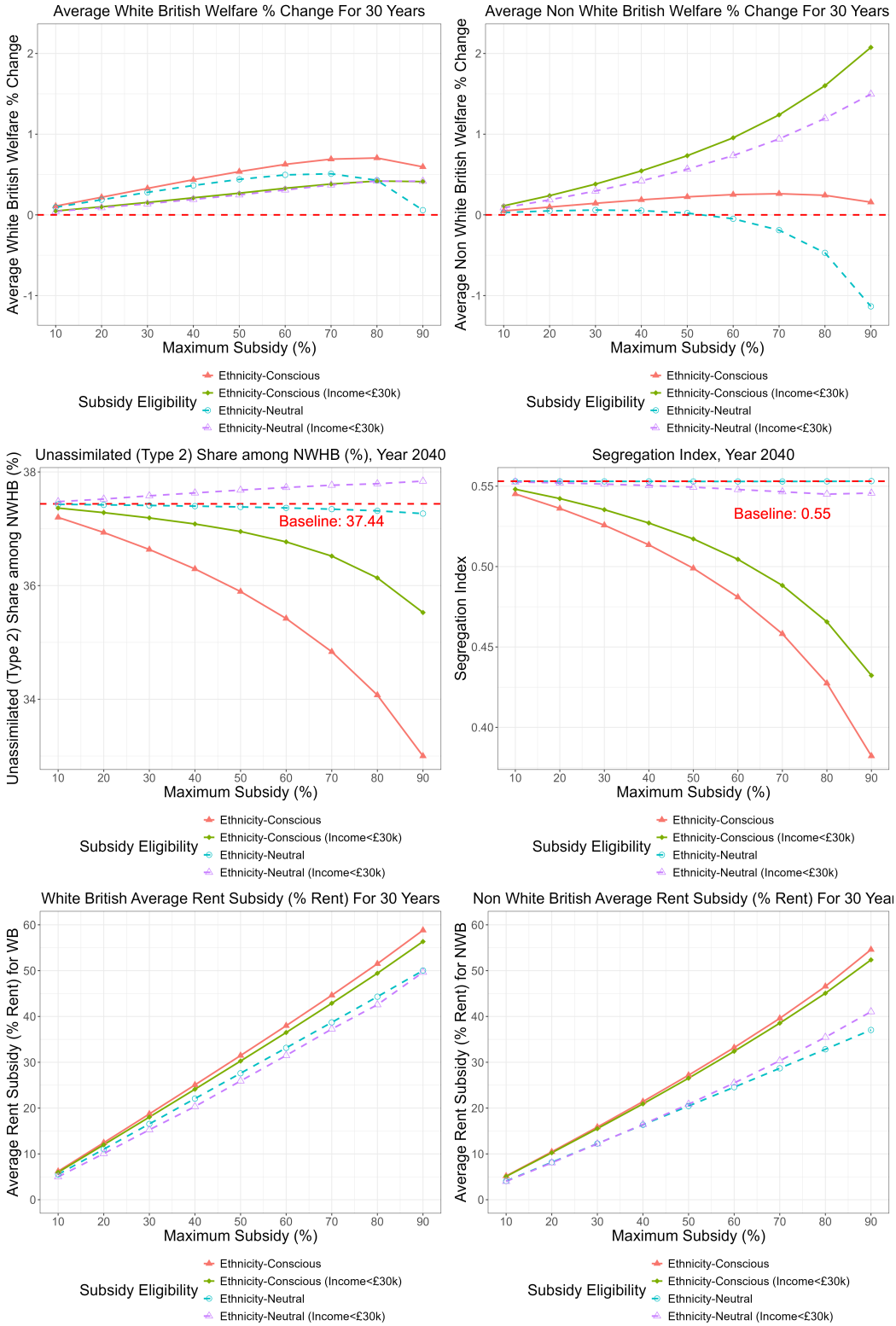
following desegregation and ethnic minorities' assimilation. The only case when average welfare loss happens is for non-White British under the ethnicity-neutral policy with the maximum subsidy rate greater than 50%. If we restrict the focus to only means-tested one, 90% maximum subsidy rate policy maximizes both White British and non-White British welfare.

The middle two panels show the outcomes 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 ethnic minorities' 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 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. 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 6), so it becomes slightly more generous than the current UK housing subsidy policy.

Figure D.2: 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.

E 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.

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,[43] while probability estimates appear without parentheses and standard errors are shown inside parentheses.

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

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

Female Gender Attitude Divorse ('nose') Wave 1

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

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

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

[43] 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

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

Male Gender Attitude Divorce ('nose') Wave 1

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

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

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

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

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

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

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

Female Interviewed in Foreign Language Wave 1

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

Male Interviewed in Foreign Language Wave 1

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

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

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

F Labor Market Skill Price

Table F.1: Log Skill Price Estimates

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

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

G Estimates for Counterfactual

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

G.1 Average Neighborhood Characteristics

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

Table G.1: Average Neighborhood Characteristics Estimates

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

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

G.2 State Variable Distribution Estimates

In this section, I report auxiliary estimates used to conduct counterfactual simulations. To simulate changes in the aggregate state variables, I need estimates for the distribution of various state variables. The first column reports the multinomial logit coefficients. The second

and third columns report the linear regression coefficients as well as the standard deviation of the residuals.

Table G.2: Distributional Parameter Estimates for Aggregate Variables

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

Source: Millennium Cohort Study, Wave 1-7
Note: This table shows the population distribution estimates for counterfactual simulation. For social position results, I run multinomial logit regressions. For human capital stock, I assume it follows a normal distribution, and I run an OLS regression.