The Role of Own-Group Density and Local Social Norms for Ethnic Marital Sorting: Evidence from the UK*

Alexander Vickery[a], and Dan Anderberg[a,b,c]

a) Royal Holloway University of London; b) Institute for Fiscal Studies, London

c) CESifo, Munich

May 8, 2019

Abstract

We present a structural marriage market model where individuals differ in ethnicity and qualifications, and where marital choices are affected by social conformity preferences. The model is estimated using White, Black and Asian individuals born in the UK between 1965 and 1989, and is identified from regional demographic variation. We find strong preferences for marital sorting both on ethnicity and qualifications. Black and Asian individuals are more likely to marry intra-ethnically in regions where the own ethnicity share is relatively large. We further find evidence of significant social conformity preferences, implying substantial variation in marital social norms. Using the estimated model, we make predictions for a set of more recent cohorts, born between 1990 and 2006, whose marital choices are still to be completed. Due to their increased population shares, the proportions of Black and Asian individuals marrying within their own ethnic group are expected to increase and this effect is amplified by endogenously changing equilibrium social norms.

Keywords: Marriage, Ethnicity, Assortative Mating, Conformity, Social Norms

JEL Classification: D10, J11, J12, J15

---

*We are very grateful to comments from Frederic Vermeulen and Monica Costa Dias, and from participants at the Workshop on Labour and Family Economics at York, the Society of Economics of the Household conference in Paris, the Applied Microeconomics Summer School at Lancaster, the Applied Microeconomics Summer School at Warwick, and at seminars at Royal Holloway.

I Introduction

As western economies are growing increasingly ethnically diverse through immigration, how minorities integrate is crucial not only to their own experiences of their host countries, but also to public opinion on immigration. Indeed, the very concept of integration remains controversial and a variety of indicators and measures have been proposed (Ager and Strang, 2004). Generally, integration is recognized as a two-way process involving "the development of a sense of belonging in the host community with some renegotiation of identity by both newcomers and hosts"(Phillimore, 2012). A traditionally held view is that marriages between ethnic-minority immigrants and native-born individuals is the ultimate proof of integration, stirring the ethnic melting pot and diminishing the significance of group differences (Blau et al., 1984). From a social acceptance perspective, inter-ethnic marriages have been seen as the breaking of the "last taboo" in ethnic and race relations (Qian, 2005). The role of inter-ethnic marriages raises interesting questions about the dynamics of the integration process. In particular, as ethnic minorities grow in relative size do inter-ethnic marriages become more or less frequent? What role in this process is played by changing marital opportunities and by evolving social norms respectively?

Economics research on marriage patterns, originating with the seminal work of Becker (1973), has focused heavily on the notion of "assortative mating", primarily on education and income. There have been strong justifications for such a focus: theory predicts that the economic benefit of marriage reflects intra-household specialization that exploits partners' relative economic opportunities and contributions (Becker, 1973; Lam, 1988), and the degree of assortative mating in these dimensions have important implications for inequality across households (Greenwood et al., 2014).

But the sharing of a common culture and identity may also contribute to the value of marriage, rationalizing observed assortative mating on, for instance, race, religion, and nationality. Indeed, marrying across ethnic lines remains a rare event. In his survey of US marriage trends, Fryer (2007) notes that inter-racial marriages account for approximately 1 percent of White marriages, 5 percent of Black marriages, and 14 percent of Asian marriages. Similarly, the empirical marriage pattern suggests that members of large dominant religious groups – Protestants, Catholics and Jews – have strong preferences for having their children identify with their own religious beliefs, thereby creating strong incentives for intra-group marriages (Bisin et al., 2004).

The seminal work of Choo and Siow (2006) (henceforth, CS) has provided a workhorse model for empirically implementing Becker's transferable utility model of the marriage market. In their framework, individuals are of different discrete "types". The utility surplus generated by a particular marriage depends systematically on a couple's type-profile and, additionally, on the partners' idiosyncratic preferences over partner type. The CS empirical framework therefore recasts the individual's marriage choice as a discrete (multinomial logit) choice problem within a market equilibrium – aggregating the choices of everyone – that determines how the systematic surplus is shared. The CS model has, over the past decade, been implemented and extended in a variety of directions and applied in a variety of contexts.[1]

Methodologically, the CS model has strong parallels to the literature on discrete choice with social interactions (Brock and Durlauf, 2001) where the utility or payoff an individual receives from a given action depends on the choices of everyone else in their reference group. The social interactions literature has been instrumental in empirically operationalizing the notion of social influence and endogenous norms of behaviour (Glaeser and Scheinkman, 2014). Marital choices, in particular in the context of inter-ethnic marriages, may also reflect endogenous social norms and a desire of individuals to conform with the behaviour of others: marrying inter-ethnically may be less of a taboo when more individuals from your own group do the same. It is then natural to combine the CS framework with the social interactions framework particularly in the context of inter-ethnic marriages. This is the direction taken in the current paper: we will study marriage choices within and across ethnic boundaries in the UK using a CS-style model extended to incorporate endogenous conformity preferences.

Such an extension requires a strong source of identification. The exogenous inputs in a CS model are the population supplies. Using the observed marital patterns, the CS approach then recovers how systematic marital surplus depends on couples' type-profiles. Indeed, in its most

[1] Key contributions include *inter alia* Galichon and Salanié (2015), Choo (2015), Dupuy and Galichon (2014), Chiappori et al. (2017), Graham (2013), and Brandt et al. (2016).

basic form with a single marriage market and an unrestricted marital surplus structure, the CS model is exactly identified.[2] Hence, any attempt to extend the CS model and identify additional features or parameters will require either restrictions on the marital surplus structure or richer exogenous population variation (Galichon and Salanié, 2015).

In this paper we follow the latter route. In particular, we will exploit that the post-war waves of first-generation immigrants that established the Black and Asian minorities in the UK had a distinct geographical settlement pattern. For instance, the vast majority of the Black immigrants arriving to the UK during this period settled in the London region. In contrast, the arriving Asian immigrants often settled in the East- and West-Midlands, Yorkshire, and also in and around London. This initial geographical settlement pattern of the arriving first generation immigrants in turn determined the geographical distribution of the second generation who are the focus of the current study. Our study will focus on the marriage patterns of UK-born individuals differing in two dimensions. First with respect to White, Black or Asian ethnicity.[3] And second, between low and high educational attainment, where "low" former is defined as having a GCSE as highest qualification and where "high" is defined as having an A-level qualfication or above. As the GCSE qualification is obtained at the minimum school-leaving age of 16, our classification effectively distinguishes between individuals without and with post-compulsory schooling respectively.

In order to exploit the geographical variation in population supplies, we will adopt a multi-market approach where each region is treated as a separate marriage market. The value of using multi-market approach for estimating a CS-style model was recently demonstrated by Chiappori et al. (2017). In our framework an individual's type is given by her ethnicity-qualification profile. In terms of preference structure as the CS framework assumes transferable utility, we specify the form of the joint marital utility and assume it has two key components (i) a "principal

[2]Intuitively, suppose there are N types of men and women. The unrestricted marital surplus function then also has N^2 terms corresponding to all possible couple type-profiles. These N^2 free parameters are then identified from the N^2 observed marriage frequencies.

[3]There is no single agreed international definition of ethnicity and race or of the distinction between them. In general, ethnicity has many dimensions which include or combine nationality, citizenship, race, colour, language, and religion. Our broad categorization of ethnicity into White, Black and Asian people is a condensed version of the categorization used by the Office of National Statistics.

preferences" component, common across all marriage markets, that varies with a couple's type profile (as in CS), and (ii) an endogenous social norm component that depends the strength of conformity preferences and on how common that particular marriage choice is locally among individuals of the own type. It should be noted that we are not the first to incorporate social preferences in a CS model. Indeed, our modelling of conformity preferences is analogous to Mourifié (2017) and Mourifié and Siow (2017). However, we are first to use a multi-market approach to identify such a model and to apply it to study inter-ethnic marriages.

Our paper makes three contributions. First, we highlight that the ethnic composition of UK-born individuals varies substantially across regions, and we show that this variation is closely related to the settlement pattern of the first generation of immigrants arriving to the UK between the mid-1950s and the mid-1970s. We document the marriage patterns within our estimating cohorts of UK-born individuals born between 1965 and 1989, both in the aggregate and across regions. A key stylized fact highlighted by our analysis is that Black and Asian individuals are more likely to be married within their own ethnicity in regions where the density of their own ethnicity is relatively high.

Second, we show that a CS model with conformity preferences is identified using a multi-market approach that exploits variation in population supplies and we estimate our model using the aforementioned sample of UK-born individuals and eleven regions. We find that the principal preferences exhibit significant complementarity in ethnicity as well as in academic qualifications. In general we find that complementarities in ethnicity are far stronger than in qualifications. We also find evidence of strong conformity preferences. As regions differ in population supplies, so do relative marriage frequencies and, by implication, so do social marriage norms. Stated differently, a marital choice that is effectively a "taboo" in one region may be less so in a different region. We show that the estimated model naturally fits the regional pattern with respect to inter-ethnic marriage frequencies: in areas with relatively larger ethnic minority groups, Whites naturally more frequently marry inter-ethnically, but critically, individuals from the ethnic minorities themselves less frequently marry Whites.

Third, in order to explore the evolution of inter-ethnic marriages in an increasingly diverse society and specifically the potential role of evolving social norms in this process, we use the estimated model to predict marriage patterns among UK-born individuals born between 1990 and 2006. In these recent cohorts the ethnic Asian minority in particular and the Black minority to a lesser extent are substantially larger than in the estimating cohorts. In order to separate out the effect of the changing demographic composition from changing social norms we contrast the model's predictions for the marriage patterns of these young cohorts for the case with endogenously evolving social norms and to a case with fixed norms. We find that, as the ethnic minorities grow in terms of their shares of the population, Whites will become more likely to marry inter-ethnically. However, Blacks and Asians will themselves become less likely to marry inter-ethnically. These effects are amplified by the endogenously evolving social norms.

The rest of the paper is organized as follows. Section II gives a brief overview of post-war immigration into the UK. Section III describes the data that we use and the marriage pattern among the estimating cohorts. Section IV set up the model and outlines the identification of social preferences while Section V describes how we estimate the model. Section VI presents the estimation results while section VII highlights our predictions for the more recent cohorts. Section VIII concludes.

II UK Post-War Immigration

Following the conclusion of the second world war the UK had a non-White population of around 30,000 people. By the end of the 20^{th} century this figure was over 3 million.[4] As a result, parts of the UK now have large ethnic minority populations. For instance, parts of London, Liverpool, Birmingham and Manchester – areas that were previously home to the White working class – are now composed predominantly of individuals from ethnic minority backgrounds. The current ethnic minority population is mixture of first-, second- and beyond-generation immigrants.

One of the key features of the settlement of ethnic minority immigrants arriving to the UK between the 1950s and the 1970s was its particular geographical pattern. As will be outlined below, the spatial distribution of the early arriving immigrants was largely driven by labour

[4]See Hansen (2000) for a comprehensive outline of post-war immigration.

demand at the time. The concentrated settlements implied that ethnic minorities have been highly prevalent in some areas for the past half century, whereas other areas remained almost entirely ethnic White. Our interest will be in UK-born individuals and hence in the second- and beyond generations. We will argue that the geographical distribution of the UK-born ethnic minorities mirrors the early location choices of the first generation and hence the labour demand and industry structure in the decades following the end of the war. In this sense, when we study the marital choices of UK-born individuals, born 1965-1989 (marrying in most cases only in the last 20 years), we will treat the geographical distribution of ethnic minorities as exogenously given.

Arrival of the First Generation

The post-war rise in immigration can be attributed to a combination of government policy and labour demand. In 1948, the British Nationality Act granted individuals from the British Empire the freedom to live and work in the UK. These extensive rights to migrate to the UK were in place until the early 1960s when, in response to a perceived heavy influx of immigrants, regulation was tightened (see below), permitting only those with government-issued employment vouchers, limited in number, to settle.

Meanwhile, the Labour government, formed after a landslide victory in 1945, attempted to eradicate the war-time austerity via an extensive nationalization and social policy, promising full employment, fair wages, and homes for all. It was quickly recognised that the reconstruction of the British economy required a large influx of immigrant labour. Initially the appeal for new workers was aimed primarily at White Europeans. Immigrants came from all over Europe, including refugees from the Communist regimes in Eastern Europe and the Soviet Union, displaced persons from refugee camps throughout Europe, along with a substantial numbers of Irish and Italian workers. In terms of the aggregate numbers this inflow of White immigrants was stationary until the early 1960s and declined thereafter until the mid 1980s.

Labour was also, for the first time, sought from outside of Europe – mainly from the Caribbean and from the Indian subcontinent. The symbolic starting point of this mass migration to the "mother country" was the journey of the *SS Empire Windrush* from Kingston, Jamaica, to Tilbury, Essex, in June 1948, carrying close to 500 West Indians intent of starting new lives in the UK. This began a large wave of migration from the Caribbean - now referred to as the "Windrush generation".˙The majority of immigrants from the Indian subcontinent arrived to Britain during the 1950s and 1960s. After Britain relinquished its Indian empire in 1947 and the turbulent partition of India into what is known today as the Republic of India and the Islamic Republic of Pakistan, many individuals migrated in order to depart from the unrest but also to take advantage of the monetary reward.

The timing and composition of post-war immigration is highlighted in Figure 1.

Fig 1 here.

It is worth noting that immigration, regardless of ethnic group, spiked significantly during the period 1960-1962. This spike is likely a result of the Commonwealth Immigrants Act 1962 which required those individuals whose passports were not issued directly by the UK government to be subject to immigration control. This was therefore an anticipatory response by those individuals whom the policy change would affect directly. Overall, the figure shows that, in the 35 years following the end of the war, in addition to a stable inflow of immigrants, there was significant ethnic variation within this flow.

Geographical Dispersion

As the post-war immigration was driven primarily by a shortage of labor, both skilled and unskilled immigrants settled in areas where the shortages were perceived to be the largest. The traditional mill towns like Bradford and Leeds in Yorkshire and Manchester in the North West of England were seen as attractive places by Asian immigrants as many of them had been involved in the textile industry prior coming to the UK. In addition, the industrial towns in the East and West Midlands (e.g. Coventry) and London's East end (e.g. Woolwich) were inviting for the high skilled Asian workers to supply labor to the manufacturing, car production and food processing sectors.

In contrast, the Black ethnic immigrants were heavily concentrated in London - in particular, areas such as Brixton and Notting Hill. These Black Caribbean immigrants were typically invited to fill labor shortages in London's hospitals, transportation, and to help with railway development, and are widely recognized as being a major contributing factor to the regeneration of London's post-war economy.

III Data

Our primary aim is to characterize the marital choices, based on ethnicity and educational attainment, of UK-born individuals. This requires us to look at a set of cohorts who (i) exhibit a substantial and geographically varied presence of second-generation ethnic minorities, and (ii) are old enough to have made their marital choices. To this end, we focus on the birth cohorts 1965-1989. We further need to choose a relevant geography to represent different marriage markets.

We will use data on individuals, aged 25 or above, from the Quarterly Labour Force Survey (QLFS) – the largest household study in the UK – for the survey years 1996-2015. This means that we will use data on individuals who are adults when surveyed. This raises the question of potential migration between areas. If all internal migration occurred *prior* to entering the marriage market it would not pose any issue for our empirical approach: all individuals would be observed in their marriage-relevant area. Indeed, empirically, the by far highest rates of internal migration is observed among individuals aged below 25. This applies for instance to migration of young individuals for purposes of university studies and labour market entry. Also, the size of areas have a large impact on the size of internal migration: while internal migration is substantial between local authorities (e.g. between London boroughs), our choice of geography is based on the larger Statistical Regions (UK Nomenclature of Territorial Units for Statistics, NUTS1), specifically Wales, Scotland, and the nine statistical regions of England.[5] At this geographical level, the rate of internal migration is substantially lower.[6]

[5]We will omit Northern Ireland as the proportions of ethnic minorities there are too small for any meaningful analysis.

[6]Based on the most recent data (ONS, 2016), the rate of internal migration between local authorities is 4.5 percent per year. The migration between areas used in our geography is substantially lower. For instance, the annual in- and outflow migration for Scotland is, in the latest figures, about 0.7 percentage points. In terms of age, the rate of internal migration peaks at age 19, the main age at which people leave home for study. There is another smaller peak at age 22; in many cases this will reflect graduates moving for employment, further study,

Sample Population

The QLFS allows us to characterize each individual's ethnicity and educational attainment. Moreover, as the survey interviews all adults in each household, it allows us to measure the same characteristics for partners. As above, we work with three broad ethnic groups: White (W), Black (B), and Asian (A). Educational attainment is broadly classified into two groups: "Low" (L) and "High" (H). We classify an individual as having "low" educational attainment if their highest academic qualification is a GCSE (General Certificate of Secondary Education) or no academic qualification at all. The GCSE is the first tier of academic qualifications in the UK, obtained at the end of the academic year in which the individual turns 16 (which also corresponds to the end of compulsory education for the cohort in question). In contrast, we classify an individual as having "high" educational attainment if their highest academic qualification is an A-level (Advanced Level) or higher including a university degree. The A-level degree is obtained at the age of 18 – after two years of post-compulsory schooling – and is the standard qualification for entry to university. Hence our "low/high" classification of educational attainment broadly partitions the population into those who left at the minimum school-leaving age and those who undertook post-compulsory studies and we use the qualification notations L and H and GCSE(-) and A-Level(+) interchangeably.

In order to focus on individuals who have had the necessary time to complete full time education and reach a marriage age, we drop individuals below age 25. Pooling the 20 years of the QLFS we then obtain a sample of 193,546 individuals born between 1965-1989 and aged 25 or above when surveyed. Table 1 provides a description of the sample by gender. As the QLFS also identifies cohabiting couples, we treat these as married. Hence it should be noted being "married" in our context includes both formal marriage and live-in partnerships.

Table 1 here.

On average, the individuals in our sample are in their mid-thirties and, on average, born in the first half of the 1970s. Table 1 also confirms the well-known stylized fact that, among the

or returning to their home address.

UK-born, educational attainment is higher among the ethnic minorities than among Whites. We further observe that Black individuals are less likely to be married when compared to Whites and Asians. This finding, which is well-documented in the US (Ellwood and Crane, 1990; Saluter, 1994; Brien, 1997; Seitz, 2009) is, in general, attributed to fundamental differences in population supplies amongst ethnic groups. The table further distinguishes between those married with a UK-born partner and those married with a non-UK-born partner. Non-UK-born partners are particularly common among the Asian population, especially among the low qualified (Charsley et al., 2012).

Geographical Variation in Ethnicity and Qualifications

In order to illustrate how the regional variation in the presence of UK-born ethnic minorities developed, panel (a) of Figure 2 first illustrates the share of Whites by region not among our estimating cohorts, among individuals born between 1950-1964. This shows how those born in the UK through this earlier period were – in practically every region – by and large exclusively White. Only London and the West Midlands had, by this time, seen some early emergence of a UK-born non-White local population.

Figure 2 here.

The remaining three panels show the proportion White, Black and Asian respectively among the UK-born individuals within our estimating cohorts. These show how a substantial Black minority had been established in the London region in particular, and how a sizeable Asian minority had emerged in London, the East- and West Midlands, Yorkshire and Humberside, and in the North West. As our estimating sample includes only UK-born individuals, this illustrates how the geographical variation of ethnicity among our sample population reflects the settlement patterns of the first generation discussed in Section II. This gives us some reassurance that the location choices of the second generation individuals were, in effect, determined exogenously by their parents.

While the period from the mid-1960s was the first time that Blacks and Asian constituted a sizeable share of the UK-born cohorts, the Asian share in particular has continued to grow. This is shown in Figure 3 which depicts the proportions of ethnic Black and Asian by birth cohort. The Black ethnic population was fairly steady between one and two percent of the UK-born population in the birth cohorts from 1960 to 1990, their proportion has started increasing more rapidly in the cohorts born after 1990. In contrast, the Asian ethnic population as a share of all UK-born individuals has increased steadily since the mid-1960s and has since the mid-1960s exceeded the Black ethnic UK-born population.

Figure 3 here.

Figure 4 highlights that there is also a strong geographical variation in the distribution of educational attainment. London has the highest proportion of individuals with the highest educational attainment, and with the exception of Scotland, the most highly educated individuals are concentrated in the southern regions of England. The figure further shows that the qualification rate is generally higher among females than among males, with the gender gap in the rate of holding a high qualification being around 2 percentage points in most areas.

Figure 4 here.

Empirical Marital Choices

While our main interest will be in assortative mating, we also need to account for the fact that some observed individuals are not married. A further complication is that some individuals who are married are married to partners who are not UK-born. As noted above (and also shown below), such marriages are particularly common among the Asian population. To account for this, we will for our purposes define "marital status" to have three categories: single, married to a UK-born partner, and married to a non-UK-born partner where, as noted above, among married we include cohabiting individuals.[7] Figure 5 shows the distribution of marital status by gender, educational attainment and ethnicity for our aggregate population.

Figure 5 here.

[7]Our classification is based on current marital status as, for married individuals, we need to measure the characteristics of their partners. This means that we classify divorced, separated and widowed individuals as single.

Two key features with respect to ethnicity stand out. First, the rate of singlehood is substantially higher among Blacks than among Whites and Asians, both for men and for women and also for low- and high educated individuals. Second, the proportion of individuals who are married to non-UK-born partners is particularly high among Asians. In contrast, only a very small proportion of White individuals – irrespective of gender and education – marry partners who are born outside the UK. With respect to educational attainment, it is interesting that, for both the Black and the White ethnic groups, being more educated also makes you more likely to be married, whereas the converse is true for Asians.

Figure 6 here.

Consider next marital sorting on ethnicity and educational attainment in the aggregate population. Figure 6 shows for each ethnicity and educational attainment, and separately for men and women, the distribution of partner type.[8] Given that the White ethnic group is a large majority group, relatively unsurprisingly, White individuals – both men and women – predominantly marry other Whites. More interestingly the marriage partner choices by Black individuals is more ethnically varied. Among married Black males, being married to a White partner is as common as being married to a Black partner. Among married Black women, being married to a White male is also common, but not quite as common as being married to a Black male. For individuals from the Asian ethnic group we observe that both males and females are most likely to marry a partner who is also Asian. However about one fifth of both Asian males and Asian females who are married to UK-born partners are married to White partners.[9] Finally, Figure 6 show that marriages between Blacks and Asians are exceptionally rare.

In the following we will use as measure of assortative mating the proportion of married individuals with a given set of characteristics whose partners share the same characteristics.[10] Table

[8] For this we restrict the sample to those married to UK-born partners for two reasons. First, doing so directly ties in with the modelling approach below where we model the supply of UK-born prospective partners, but not the non-UK-born supply. Second, doing so avoids having to classify in particular the qualifications of the non-UK-born partners.

[9] As noted above, marriages to non-UK-born spouses are particularly common among Asians, typically to spouses from the own country of origin, thus lowering the proportion married to White partners among all married Asians.

[10] A number of measures of assortative mating have been proposed in the literature. The proposed measures generally compare the observed marriage pattern to some counterfactual reference allocation, most commonly

2 provides these empirical proportions (with standard errors) in the aggregate sample population of married individuals with UK-born partners for each ethnicity/qualification characteristic by gender. For instance the table shows that, among married white males with UK-born partners, 99.5 percent are married to White partners. In contrast, only 39 percent of married Black males are married to Black partners. This is far below the 79 percent Asian married men who are married to Asian partners. The ethnic sorting among married women is naturally similar (though higher for married Black women than for married Black men). The table also shows that about 60 percent of low qualified married individuals are married to low qualified partners, whilst around 70 percent of high qualified married individuals (slightly higher for males and slightly lower for females) have high qualified partners.

Table 2 here.

The above reported empirical marriage patterns were all in the aggregate. But of course one of the central aspects of the current analysis is that marriage outcomes differ across regions. Consider first marital status.

Figure 7 here.

Figure 7 shows the distribution of marital status by region and gender. In the figure the regions have been ordered in ascending order in terms of proportion Whites, thus starting with London as the most ethnically diverse region through to Wales as the most homogenously ethnically White. There are two noteworthy features. First, the proportion single is particularly high in London compared to the other regions.[11] Second, the areas with relatively larger ethnic

random matching (Liu and Lu, 2006 and Eika et al., 2018). A key problem with such measures is that the reference point is very sensitive to variation in the supply at low levels in particular. Intuitively, the likelihood of marriages within a very small minority group under random matching is vanishingly small, but increases fast as their share of the population increases. As our focus is indeed on the increase in the share of ethnic minorities, such measures tend to fairly mechanically indicate sharp decreases in assortative mating due to sharp increases in the probability of intra-group marriages in the random matching reference allocation. A further problem with such measures is that they, by taking the married population as given, ignore the extensive marriage margin. This was recently noted by Dupuy and Weber (2018) who use a Choo-Siow-style model to explore the role of changing marital patterns for the evolution of income inequality. These issues further motivate our choices of highlighting actual and predicted choice frequencies directly rather than using summarizing measures of sorting.

[11] The observed lower proportion married in the London is consistent with data from the 2011 census (ONS, 2012, Table KS103EW).

minority groups are generally also the areas with larger proportions of people married to non-UK-born partners.

Consider next marital sorting, specifically with respect to ethnicity. Figure 8 shows the proportion of married individuals whose partners share the same ethnicity, by region and gender, and by ethnicity. Naturally, the proportion of married Whites whose partners are also White approaches 100 percent in areas where the Black and Asian ethnic minorities are very small. The more central feature highlighted by the figure is how the marriage behaviour of members of the ethnic minorities varies systematically across regions. Specifically, the figure suggests that married ethnic minority individuals are more likely to be married to a same-ethnicity partner in areas where the ethnic minorities are comparatively large. Specifically, married Asians are more likely to have Asian partner if they are from London, the East- and West Midlands, Yorkshire or the North West, than if they are from elsewhere. Similarly, married Black individuals from London are particularly likely to have Black partners.

Figure 8 here.

In order to test the robustness of this stylized observation Table 3 reports on the estimation of a set of simple probit models. In each reported regression we use the relevant subsample of individuals married to UK-born partners, and regress a dummy for the individual being intra-ethnically married on the proportion of marriage market peers who are of the own ethnicity. As we have defined an individual's relevant marriage market as comprising all UK-born people, born between 1965-1989, in the own region, we use our estimating sample also to characterize the proportion of own ethnicity in the own region.

We present regressions by ethnicity and gender, in each case using three different specifications. Consider first specification (i) which is a simple probit regression without any further controls. We already know from Table 2 that more than 99 percent of all Whites males married to UK-born partners have White partners. The estimated marginal effect of the proportion of own ethnicity suggests e.g. that when the proportion Whites drops by 20 percentage points – which is close to the difference between the most homogenously White regions and the London region – the frequency of intra-ethnicity marriages reduces by 1.3 percentage points. As we will see in Figure 11 below, this is consistent with the intra-ethnicity marriage frequency for White males being very close to 100 percent in the most homogenously White regions and about 98 percent in the London region. A very similar result holds for White females. Among Blacks married males, we know from Table 2 that only about 40 percent have Black partners. The estimated coefficient suggests that this fraction increases substantially with the proportion of Blacks in the local population. Indeed, as we will see in Figure 11 below, the proportion of married Black males who are intra-ethnically married increases from less than 20 percent in the areas with the smallest Black populations to about 60 percent in London. Again, a similar result holds for married Black females. Finally, for Asians we see a similar patterns as for Blacks though slightly less pronounced: for both males and females we find that the proportion intra-ethnically married increases with the share of the Asians in the local population. This is consistent with Figure 11 below which shows that the rate of intra-ethnic marriages Asians increases from around 0.5 in the areas with the smallest Asian population shares to around 0.9 in London.

In specification (ii) we instrument for the share of the own ethnicity in the own area drawing on the logic that the settlement pattern of the postwar immigrants was a strong determinant of the local ethnic compositions of our estimating cohorts. To this end we construct our instrument as the share of the own ethnicity among individuals, in the own region, born between 1940 and 1960. Note that in constructing the instrument, we include both UK- and non-UK-born individuals in order to include first generation immigrants, capturing that these older cohorts would naturally be parents to the estimating cohorts. Instrumenting for the share of the own ethnicity in the local marriage market has little effect on the estimates for Whites and Blacks. For Asians the estimated effect weakens and in the case of women loses statistical significance.

In specification (iii) we introduce the own qualification level as an additional regressor in order to explore whether high qualified individuals are more or less prone to marry intra-ethnically. For White males, the negative estimated marginal effect indicates that high qualified married White men are 0.3 percentage points – or more than 50 percent – more likely to have a non-White partner compared to low qualified married White men. High qualified White women also appear to be more likely to have non-White partners, though the gap between high and low qualified is smaller than for men and not statistically significant. Among Black males, being high qualified is associated with a higher probability of being having a Black partner. The estimated effect is positive also for Black women, but not statistically significant. This is contrary to the Asians for whom being high qualified is associated with lower probability of being intra-ethnically married.

Table 3 here.

IV The Model

Our model builds on Choo and Siow (2006) (CS), whose seminal work showed how the static, frictionless, transferable utility equilibrium model of the marriage market could be estimated as a set of discrete choice problems that collectively identify the marital surplus structure. The key innovation in CS was to assume that all individuals belong to some discrete set of observable types and that individuals have preferences over partner type with these preferences having both a deterministic and a random component. This setup allowed CS to recast the frictionless marriage market model with transferable utility as a set of discrete choice problems connected via equilibrium constraints.

As noted in the introduction the CS modeling framework has been subsequently generalized in a variety of directions. Our model is closest in spirit to the generalization by Mourifié and Siow (2017) who include "peer effects" in the CS model. The notion that peers may influence individuals' marital choices is one that has previously been explored in the literature. For instance, Adamopoulou (2012) uses panel data on friendship networks and shows that direct friends influence individuals' partnership formation choices. But peer effects also encompass wider social norms, including what are considered as socially acceptable conventions or, conversely, what behaviours are considered to be taboos. What makes social norms particularly interesting and challenging to study is that they are endogenous equilibrium concepts that can vary across groups and geographical areas. Indeed, one of our innovations is to implement a marriage market model with social preferences in a multi-market setting where social norms vary across regions (Burke and Young, 2011).

Setup

Consider an economy consisting of a continuum $\mathcal{I}$ of men and a continuum $\mathcal{J}$ of women. We can normalize the size of the overall population to unity. The individuals, males and females, differ in observable "type", where any individual's type is observable to the researcher. The set of possible types, denoted X, is finite with N elements, and is the same for men and women. In our application, a type will be an ethnicity-qualification combination, with three ethnicity groups and two qualification levels. A typical type is denoted x. The individuals in the economy are further partitioned into groups. Let G denote the (finite) set of groups, with typical element g. Let $h^{g,k}(x)$ denote the probability mass function describing the population distribution across groups $g \in G$, gender $k = m, f$, and types $x \in X$.

Each individual faces a choice between marrying and remaining single ("option 0"). Marriage between male $i \in \mathcal{I}$ of type x_i and female $j \in \mathcal{J}$ of type x_j in group g generates a "principal" systematic utility denoted $\zeta^g(x_i, x_j)$, where the mapping $\zeta^g(\cdot, \cdot)$ can be represented as a $N \times N$ matrix. Within marriage, utility is assumed to be transferable. The principal utility from remaining single is normalized to zero.

The model contains two additional utility terms. Following CS, male i has a set of additive random preferences over partner type and over singlehood, with his idiosyncratic utility from marrying a partner of type x_j denoted $\varepsilon_i(x_j)$ and over remaining single denoted $\varepsilon_i(0)$. Also following CS, these random utilities are all taken to be i.i.d. extreme value distributed both across partner types and across individuals. Note that individuals in this economy do not hold preferences over specific individuals of the opposite gender, only over their types.[12] In line with this, let $\omega^{g,m}_{x_j|x_i}$ denote the proportion of males who belong to group $g \in G$ and are of type $x_i \in X$ who marry partners of type $x_j \in X$. Similarly let $\omega^{g,f}_{x_i|x_j}$ denote the proportion of women who belong to group $g \in G$ and are of type $x_j \in X$ who marry partners of type $x_i \in X$. Let $\omega^{g,m}_{0|x_i}$ and $\omega^{g,f}_{0|x_j}$ denote the proportions who remain single.

Following Brock and Durlauf (2001) we will model social preferences as preferences for conformity with choices within a reference group. Specifically, we assume that the relevant reference

[12]This assumption on preferences effectively rules out sorting on any unobserved personal characteristics (Galichon and Salanié, 2015).

group of a given individual is the set of individuals of the same type and gender in the own group. We then assume that the utility of a particular choice depends on the proportion of individuals in the individual's reference group making the same choice. Hence if male i of type x_i in group g marries a type x_j woman, the utility from doing so will depend on $\omega^{g,m}_{x_j|x_i}$. Correspondingly, from the female's side, the marriage utility will depend on $\omega^{g,f}_{x_i|x_j}$. We assume that these social marriage preferences augment the principal preferences additively in log form with coefficients ϕ^m_1 and ϕ^f_1 respectively.[13] These are scalar parameters capturing the strength of male and female social preferences over marital choices respectively. Hence we model the *joint systematic utility* from a marriage between a male of type x_i and a female of type x_j in group g as

$$Z^g(x_i, x_j) = \zeta^g(x_i, x_j) + \phi^m_1 \log\left(\omega^{g,m}_{x_j|x_i}\right) + \phi^f_1 \log\left(\omega^{g,f}_{x_i|x_j}\right). \tag{1}$$

Social preferences can also affect the utility from being single, shifting the systematic utility from (the normalized) zero to

$$U^g(x_i, 0) \equiv \phi^m_0 \log\left(\omega^{g,m}_{0|x_i}\right), \quad \text{and} \quad V^g(0, x_j) \equiv \phi^f_0 \log\left(\omega^{g,f}_{0|x_j}\right), \tag{2}$$

for males of type x_i and females of type x_j in group g respectively. Here the scalars ϕ^m_0 and ϕ^f_0 capture the strength of male and female social preferences for singlehood.

Establishing some terminology at this stage will be helpful. We will refer to ϕ^k_1 and ϕ^k_0 as "social preferences". These are common to everyone of gender $k = m, f$. However, as the individuals have different reference groups, the equilibrium values of the utility terms $\phi^k_1 \log\left(\omega^{g,k}_{x'|x}\right)$ and $\phi^k_0 \log\left(\omega^{g,k}_{0|x_i}\right)$ will vary across individuals. We will refer to these utility components as (local equilibrium) "social norms".

As utility is transferable in marriage, utility in marriage can be shared in any way between the partners. As shown by Chiappori et al. (2017), under the assumed additive random preferences, the marriage market equilibrium in group g will entail an equilibrium allocation of $Z^g(x_i, x_j)$

[13] We are assuming that social preferences enter utility in log form; this is contrary to Brock and Durlauf (2001). In principle, there is no obvious economic ex ante justification for either the linear or log specification. The log specification allows a semi-closed form solution.

into a part obtained by the male, denoted $U^g(x_i, x_j)$, and a part obtained by the female, denoted $V^g(x_i, x_j)$.[14]

The equilibrium utility of male i of type x_i and group g from marrying a type x_j woman would then be $U^g(x_i, x_j) + \varepsilon_i(x_j)$ while his utility from remaining single would be $U^g(x_i, 0) + \varepsilon_i(0)$. Correspondingly, the equilibrium utility of female j of type x_j and group g from marrying a type x_i man would then be $V^g(x_i, x_j) + \varepsilon_j(x_i)$ while her utility from remaining single would be $V^g(0, x_j) + \varepsilon_j(0)$. With the random preferences being extreme value distributed, it follows that the choice frequencies of males take the standard logit form. Specifically,

$$\omega^{g,m}_{x_j|x_i} = \frac{\exp\left(U^g(x_i, x_j)\right)}{\sum_{x' \in X \cup \{0\}} \exp\left(U^g(x_i, x')\right)}, \qquad \text{for } x_j \in X \cup \{0\}. \tag{3}$$

Similarly, for females,

$$\omega^{g,f}_{x_i|x_j} = \frac{\exp\left(V^g(x_i, x_j)\right)}{\sum_{x' \in X \cup \{0\}} \exp\left(V^g(x', x_j)\right)}, \qquad \text{for } x_i \in X \cup \{0\}. \tag{4}$$

Solving for the Equilibrium

Consider the marriage market equilibrium in group g. Taking the ratios of the frequencies in (3) and (4) and adding we obtain that, for any marriage profile $(x_i, x_j) \in X \times X$,

$$\begin{aligned} \log\left(\frac{\omega^{g,m}_{x_j|x_i}}{\omega^{g,m}_{0|x_i}} \frac{\omega^{g,f}_{x_i|x_j}}{\omega^{g,f}_{0|x_j}}\right) &= \zeta^g(x_i, x_j) + \phi^m_1 \log\left(\omega^{g,m}_{x_j|x_i}\right) + \phi^f_1 \log\left(\omega^{g,f}_{x_i|x_j}\right) \\ &\quad -\phi^m_0 \log\left(\omega^{g,m}_{0|x_i}\right) - \phi^f_0 \log\left(\omega^{g,f}_{0|x_j}\right), \end{aligned} \tag{5}$$

where we used that $U^g(x_i, x_j) + V^g(x_i, x_j) = Z^g(x_i, x_j)$ and (1) and (2). Rearranging and using that, in equilibrium, the market is balanced so that $\omega^{g,m}_{x_j|x_i} h^{g,m}(x_i) = \omega^{g,f}_{x_i|x_j} h^{g,f}(x_j)$ we obtain that

$$\omega^{g,m}_{x_j|x_i} = \left[\exp \zeta^g(x_i, x_j) \left(\omega^{g,m}_{0|x_i}\right)^{1-\phi^m_0} \left(\omega^{g,f}_{0|x_j}\right)^{1-\phi^f_0} \left(\frac{h^{g,f}(x_j)}{h^{g,m}(x_i)}\right)^{1-\phi^f_1}\right]^{\frac{1}{2-\phi^m_1-\phi^f_1}}, \tag{6}$$

[14] While the idiosyncratic utility components $\varepsilon_i(x_j)$ and $\varepsilon_j(x_i)$ could, in principle, also be shared a key aspect of the equilibrium – following from the assumption that any individual only holds preferences over partner type and the frictionless competitive environment – is that these will fully accrue to the individual.

for males and

$$\omega_{x_i|x_j}^{g,f}=\left[\exp\zeta^g\left(x_i,x_j\right)\left(\omega_{0|x_i}^{g,m}\right)^{1-\phi_0^m}\left(\omega_{0|x_j}^{g,f}\right)^{1-\phi_0^f}\left(\frac{h^{g,m}\left(x_j\right)}{h^{g,f}\left(x_i\right)}\right)^{1-\phi_1^m}\right]^{\frac{1}{2-\phi_1^m-\phi_1^f}}, \tag{7}$$

for females.

This offers a way to solve for the equilibrium in group g conditional on the set of parameters and the population supplies. Specifically, we can use that, for any group $g \in G$, adding-up holds for all types

$$\omega_{0|x_i}^{g,m}+\sum_{x_j\in X}\omega_{x_j|x_i}^{g,m}=1 \text{ for all } x_i\in X, \quad \text{and} \quad \omega_{0|x_j}^{g,f}+\sum_{x_i\in X}\omega_{x_i|x_j}^{g,f}=1 \text{ for all } x_j\in X. \tag{8}$$

Substituting for the marriage frequencies using (6) and (7) we can then solve numerically for the equilibrium singles rates, $\omega_{0|x_i}^{g,m}$ and $\omega_{0|x_j}^{g,f}$, of all male and female types simultaneously. We use a simple Newton method that solves the equilibrium swiftly. This procedure is nested within the estimation, thus solving the equilibrium in each group at every trial value of the parameters.

Identification

The general model is not identified for two fundamental reasons. First, from CS we know that, under the assumption of no social preferences, $\phi_0^k=\phi_1^k=0$ for $k=m,f$, the N^2 terms in the joint marriage utility $\zeta^g(\cdot,\cdot)$ is exactly identified from the same number observable marriage rates in group g. Hence any identification of social preferences requires imposing an assumption on the principal preferences. The natural assumption to impose is that $\zeta^g(\cdot,\cdot)$ does not vary across groups.

Assumption 1. (Common principal preferences) $\zeta^g(x_i,x_j)=\zeta(x_i,x_j)$ for all $g\in G$ and all $(x_i,x_j)\in X\times X$ for some common joint utility function $\zeta(\cdot,\cdot)$.

The assumption of common principal preferences reduces the number of parameters in the model to N^2+4. These can conveniently be collected in a parameter vector $\boldsymbol{\theta}$,

$$\boldsymbol{\theta}\equiv\left[\zeta\left(x_1,x_1\right),...,\zeta\left(x_1,x_N\right),...,\zeta\left(x_N,x_1\right),...,\zeta\left(x_N,x_N\right),\phi_1^m,\phi_0^m,\phi_1^f,\phi_0^f\right]'. \tag{9}$$

Second, the general strength of the social preferences is not identified. This can be seen by noting that (5) can be rewritten as

$$\begin{aligned}\zeta\left(x_i,x_j\right) \quad = \quad & \left(1-\phi_1^m\right)\log\left(\omega_{x_j|x_i}^{g,m}\right)+\left(1-\phi_1^f\right)\log\left(\omega_{x_i|x_j}^{g,f}\right) \\ & -\left(1-\phi_0^m\right)\log\left(\omega_{0|x_i}^{g,m}\right)-\left(1-\phi_0^f\right)\log\left(\omega_{0|x_j}^{g,f}\right).\end{aligned} \tag{10}$$

Specifically, dividing through by a constant would rescale $\boldsymbol{\theta}$ but leave the equation intact, meaning that the economy with the same population supplies but rescaled parameters would have the same equilibrium. This result, which is a version of the reflection problem noted by Manski (1993), implies that a normalization is required. A natural assumption to impose is that one of the social preferences parameters is zero.

Assumption 2. (Normalization) The social preferences for singlehood is zero among women, $\phi_0^f=0$.

Having imposed the assumption of common principal preferences, relative marriage frequencies will vary across groups for two reasons. First, and fundamentally, they will vary due to variation in relative population supplies across groups. Second, social preferences will vary endogenously with the group-specific relative choice frequencies. This means that having variation in relative population supplies across groups is required for there to be variation across groups also in the observable relative choice frequencies and hence for identification.[15]

Assumption 3. (Variation in population supplies) At least two groups have population distributions that are not the same: $h^{g,k}(x)\neq h^{g',k}(x)$ for some $g,g'\in G$, some gender $k=m,f$, and some type $x\in X$.

Remark 1. It should be noted that any difference in the population supplies, for instance a shifting of mass between just two types, will generally affect all equilibrium marriage and singles rates. For a formal comparative statistics analysis of the CS model, see Decker et al. (2013).

Using Assumptions 1 - 3 we can now prove identification

[15]It is easy to verify that the equilibrium within each group g is invariant to population size.

Proposition 1. ***(Identification)*** *Suppose $N \geq 2$ and $|G| \geq 2$, and that Assumptions 1 and 3 hold. Then $\boldsymbol{\theta}$ is identified from observable marriage and singles rates.*

Proof. See **Appendix A**

Proposition 1 sets out the minimum conditions for identification of the model parameters, including the social preferences parameters. In our application we have both more than two types and two groups. Specifically, we will have six ethnicity-qualification types and eleven groups (regions). This also means that we introduce some further components to the model specification. In particular, we will allow for group-specific fixed effects, δ_g, assumed to enter additively in the marriage utility (1) for all $g \in G$ (except for a reference group). These group fixed effects will be identified from level differences in marriage rates across groups.

Furthermore, as noted in Section III, a non-negligible number of individuals are observed to be married to partners who were not born in the UK. Hence we model this as a further option that is always available to any individual (that is, without modelling the supply of such potential partners). Denoting this option as -1, we model it as having a type- and gender-specific systematic utility – $\xi(x_i, -1)$ for males and $\xi(-1, x_j)$ for females – and an individual extreme value distributed random utility component. This thus adds a further $2N$ parameters that are identified from the observable rates at which men and women of different types marry partners born outside of the UK.

In the following section we introduce a simple estimator used to recover the parameter vector.

V Estimation

An individual's "type" in our setting is given by their ethnicity and qualification profile. There are three possible ethnicities $Z \equiv \{W, B, A\}$ and two qualification levels $Q \equiv \{L, H\}$. Hence an individual's type is a pair $x = (z, q) \in X \equiv Z \times Q$. For instance a type $x = (W, L)$ is a White low-qualified individual. Our set of groups G are the eleven regions – Wales, Scotland and the nine statistical regions of England – as outlined in Section III. The population distributions $h^{g,k}(x)$ are taken to be known, given by the observable relative frequencies of ethnicity-qualification types by region and gender. From the QLFS data we observe a sample with $N^{g,k}(x)$ individuals of type x and gender k from region g (which may or may not be proportional to the population distribution). Hence we observe marital choice frequencies by gender, region and type.

Let $\boldsymbol{\theta}$ denote the vector of parameters (as defined in (9) but augmented with the group fixed effects and the type-specific utilities from marrying a non-UK-born partner). Estimating the parameters involve solving the equilibrium, across all groups, at a trial value of the parameter vector $\boldsymbol{\theta}$, then adjusting $\boldsymbol{\theta}$ to obtain the best possible fit. The criterion used here is maximum likelihood.

Using the notation previously defined let $\omega^{g,k}_{x'|x}(\boldsymbol{\theta})$ denote the predicted choice behavior of individuals of gender k in group g at the parameter vector $\boldsymbol{\theta}$. Let $\widehat{\omega}^{g,k}_{x'|x}$ denote the corresponding observed behaviour. We then have that the likelihood of the observed choices among gender k in group g can be written as,

$$L^{g,k}(\boldsymbol{\theta}) = \prod_{x\in X}\prod_{x'\in X\cup\{0,-1\}}\left[\omega^{g,k}_{x'|x}(\boldsymbol{\theta})\right]^{N^{g,k}(x)\widehat{\omega}^{g,k}_{x'|x}}. \tag{11}$$

The overall log likelihood, $\ln L(\boldsymbol{\theta})$, is then obtained by taking the log of (11) and summing over genders and regions. One of the benefits of using a maximum likelihood estimator is that we can consequently use likelihood ratio tests to test parameter restrictions. For instance, this allows us to test the power, if any, that the social preferences, have in explaining marriage patterns.

VI Results

Specification Choice

A preferred specification is selected by estimating a set of alternatives. In order to statistically compare the alternative specifications likelihood ratio tests and mean squared error (MSE) goodness of fit measures are used. The four alternative specification estimated are:

(i) "Baseline specification": No regional fixed effects ($\delta_g = 0$ for all $g \in G$) or social preferences ($\phi^m_1 = \phi^f_1 = \phi^m_0 = \phi^f_0 = 0$).

(ii) "Regional effects specification": Includes regional fixed effects, δ_g for $g \in G$, but no social preferences ($\phi_1^m = \phi_1^f = \phi_0^m = \phi_0^f = 0$).

(iii) "Marriage social preferences specification": Includes regional fixed effects, δ_g for $g \in G$, and includes social preference over marriage, ϕ_1^m and ϕ_1^f, but not over singlehood ($\phi_0^m = \phi_0^f = 0$).

(iv) "Full social preferences specification": Includes regional fixed effects, δ_g for $g \in G$, and includes social preferences parameters for both marriage, ϕ_1^m and ϕ_1^f, and for singlehood for men ϕ_0^m but with $\phi_0^f = 0$ for women as a normalization.

The model fit measures are presented in Table 4 where each column corresponds to a specification outlined above. As each specification adds one or more parameters relative to the previous one, the likelihood ratio test statistic and associated p-value in each of columns (ii) to (iv) reports the results from testing whether the constrained version in the previous column is rejected.

Table 4 here.

Comparing specifications (i) and (ii) shows that the inclusion of regional specific effects are important for the model to fit the data. This is consistent with the observation in Figure 7 that the proportion married is substantially lower in some regions – London in particular – than in others.

Specification (iii) includes gender-specific social preferences for marriage, ϕ_1^m and ϕ_1^f, but not for singlehood. The likelihood ratio test confirms that, relative to specification (iii), the constrained model that excludes ϕ_1^m and ϕ_1^f is statistically rejected. Hence social preferences with respect to marriages help fit the data.

Finally, specification (iv) includes a full set of (normalized) social preferences. Specifically, as ϕ_0^f is normalized to zero (2), it adds ϕ_0^m relative to specification (iii). In this case however, the likelihood ratio tests does not reject the constraint imposed in specification (iii) relative to specification (iv). This thus indicates that there is no gender differences in social preferences with respect to singlehood. As an alternative measure of fit, we further provide the mean squareed error (MSE). This measure takes provides the (unweighted) mean of $\left[\omega_{x'|x}^{g,k}(\boldsymbol{\theta}) - \widehat{\omega}_{x'|x}^{g,k}\right]^2$ across all cells, defined by gender k, region g, own type x and marital choices x'. This measure confirms that adding regional fixed effects and social preferences over marriage (but not singlehood) improves the model fit.

Based on the above results we select (iii) as our preferred specification and report parameter estimates and model fit below for this particular specification.

Model Fit

We start by showing how the estimated model fits the aggregate moments presented in Section III. Consider first marital status by gender and type, where marital status is either "single","married to UK-born partner", or "married to non-UK-born partner". Figure 9 shows the model-predicted distribution of marital status, directly corresponding to the empirical distribution shown in Figure 5.

Figure 9 here.

Similarly, Figure 10 shows the model-predicted distribution of partner type by own type among married individuals, directly corresponding to the empirical distribution shown in Figure 6.

Figure 10 here.

Overall, the estimated model matches the aggregate distributions of marital status and partner type very closely. Correspondingly, the differences between empirical proportions reported in Table 2 and the model-predicted counterparts are all less than a percentage point. Thus the model naturally replicates the assortative mating both on ethnicity and on qualifications in the aggregate.

Equally import for our purposes is how the model fits the local marriage patterns, in particular how marriage choices vary with the local ethnic composition. Figure 11 plots, for each ethnicity and gender and across regions, the observed and model-predicted proportions of marriages that are intra-ethnic against the share of own ethnicity in the local population.

Figure 11 here.

The two top panels highlight how, trivially, for White males and females the proportion of intra-ethnic marriages approaches unity as the share of Whites in the local population approaches unity. The figures however also show that the model provides a realistic prediction for the one area (London) where the share of Whites in the local population is substantially lower (82 percent). Turning to Blacks, the two middle panels highlight how the proportion of Blacks who are married to Blacks increases with the share of Blacks in the local population, and how the model correctly predicts this relationship. The bottom two panels do the same for Asians, again highlighting how both in the data and in the model predictions, the proportion of marriages that are intra-ethnic increases with the share of the own ethnicity in the local population.

Parameter Estimates

In this section we present the estimated parameters from the preferred specification. Recall that the estimated model has the following set of parameters: (i) 36 principal type-profile-based marriage utilities, denoted $\zeta(x_i, x_j)$, (ii) two gender-specific social preferences over marriage, ϕ_1^k for $k = m, f$, (iii) ten (after one normalization) regional marriage utility fixed effects, δ_g, and (iv) twelve gender and type-specific utilities from marrying a non-UK-born partner, $\xi(x_i, -1)$ for men and $\xi(-1, x_j)$ for women. Hence the preferred specification has a total of 60 parameters.

Principal Type-Profile Preferences and Complementarities

Table 5 presents the estimated principal type-profile-based joint marriage utilities. The preference parameters are fairly precisely estimated with exceptions relating to some very rare marriages between Blacks and Asians. As the model assumes transferable utility assortative mating is fundamentally driven by complementarities in the joint marital utility function. Hence it is instructive to explore the ubiquity and strength of complementarities with respect to ethnicity and with respect to qualifications. A complication when doing so is that, when looking for instance at complementarities in ethnicity, this can occur for alternative qualification profiles etc.

Table 5 here.

Hence consider first complementarity with respect to ethnicity. Fix a male-female qualification profile $(q_i, q_j) \in Q \times Q$ and a pair of ethnicities $z, z' \in Z$, $z \neq z'$, and define

$$\eta\left(z, z'; q_i, q_j\right) \equiv \zeta\left(z, q_i; z, q_j\right) + \zeta\left(z', q_i; z', q_j\right) - \zeta\left(z, q_i; z', q_j\right) - \zeta\left(z', q_i; z, q_j\right), \tag{12}$$

as the measure of complementarity in ethnicity within z and z' conditional on (q_i, q_j). For instance, consider two low-qualified males – one White and one Black – and two high qualified females – one White and one Black. Based on the principal preferences alone, if they marry, they would sort by ethnicity if $\eta\left(W, B; L, H\right) > 0$. As such, this measures the strength of preferences towards sorting that would need to be overcome either by social norms or by idiosyncratic preferences for sorting not to occur. As there are four possible male-female qualification profiles and three possible ethnicity combinations, there are there a total of twelve ethnicity complementarity measures.

Similarly, consider complementarity with respect to qualification. Now fix a male-female ethnicity profile $(z_i, z_j) \in Z \times Z$ define,

$$\eta\left(q, q'; z_i, z_j\right) \equiv \zeta\left(z_i, q; z_j, q\right) + \zeta\left(z_i, q'; z_j, q'\right) - \zeta\left(z_i, q; z_j, q'\right) - \zeta\left(z_i, q'; z_j, q\right), \tag{13}$$

where $q, q' \in Q$, $q \neq q'$ which, since there are only two elements in Q, in this case must simply be low and high qualification respectively. For instance, consider two White males – one low and one high qualified – and two Black females – one low and one high qualified. Based on the principal preferences, if they marry, they would sort by qualification if $\eta\left(L, H; W, B\right) > 0$. As there are nine possible ethnicity profiles, there are there a total of nine qualification complementarity measures.

Table 6 here.

Table 6 presents all ethnicity complementarity measures in Panel A and all qualification complementarity measures in Panel B. All measures in Panel A are positive, indicating assortative mating on ethnicity for every qualification profile and ethnicity combination. The most noteworthy feature of Panel A is that $\eta\left(W, A; q_i, q_j\right) > \eta\left(W, B; q_i, q_j\right)$ for every male-female qualification profile (q_i, q_j); this suggests that the impetus towards ethnical sorting embedded in the principal preferences is uniformly stronger between Asians and Whites than among Blacks and Whites. The estimated ethnic complementarities between Blacks and Asians are also very large – reflecting how rare such marriages are even in areas, such as London, where both ethnicities have a strong presence– but the estimates are also highly imprecise.

Similarly, all measures in Panel B are positive indicating assortative mating on qualifications for every ethnicity profile. The numbers on the main diagonal – thus involving intra-ethnic marriages – are of similar magnitude, though slightly larger for Asians, suggesting that stronger principal preferences for educational sorting among Asians than among Whites and Blacks. It is also noteworthy that, comparing the principal preferences for sorting on ethnicity in Panel A are much stronger than are the corresponding principal preferences for (within-ethnicity) sorting on education in Panel B. For inter-ethnic marriages, complementarities in qualifications appear stronger in marriages involving Whites and Asians, but weaker in marriages between Whites and Blacks. This suggests that inter-ethnic marriages that occur between Whites and Asians tend to be more strongly sorted on qualifications than are inter-ethnic marriages between Whites and Blacks. The measures of complementarity for marriages between Blacks and Asian, while large, are again highly imprecise.

Social Preferences

In the estimation of alternative specifications in Table 4, we found no evidence of (gender-differential) social preferences for singlehood. Hence our preferred specification only includes social preferences for marriage, ϕ_1^m and ϕ_1^f. The estimated values of these parameters are provided in Panel A of Table 7.

Table 7 here.

Both coefficients are positive, of similar magnitude, and not statistically significantly different. Henceforth we will use $\sigma^g(x_i, x_j)$ to denote the joint marital utility arising from local social norms regarding marriages of type profile (x_i, x_j), defined as

$$\sigma^g(x_i, x_j) \equiv \phi_1^m \log\left(\omega^{g,m}_{x_j|x_i}\right) + \phi_1^f \log\left(\omega^{g,f}_{x_i|x_j}\right). \tag{14}$$

The estimates thus indicate that a marrying couple experience a surplus penalty if their particular type profile occurs infrequently within their relative reference groups. As a result, the social norms induce individuals to make choices that conform with the choices of others in their individual reference groups. As the principal preferences $\zeta(x_i, x_j)$ tend to lead to assortative mating – making cross-group marriages relatively rare – social norms tend to reinforce type-based sorting.

It is worth noting that, while the social preferences, ϕ_1^m and ϕ_1^f, are common to all men and women respectively, the equilibrium social norms vary both across partner type-profile (x_i, x_j) and across regions $g \in G$. One might intuitively expect social norms to be more favourable towards inter-ethnic marriages in areas where a given ethnic minority is larger. However, that is not necessarily the case. To see this, consider for instance marriage between Whites and Blacks. In areas - London in particular – where Blacks are more prevalent, on the one hand, the proportion of Whites marrying Blacks is indeed larger. But on the other hand, the proportion of Black marrying Whites is lower than in areas with fewer Blacks. As a result, the social norms towards White-Black inter-ethnic marriages in an area such as London may be either more or less favourable than in an area with a lower presence of Blacks in the local population. Indeed, noting from Table 7 that the marriage social preferences are very similar for men and women, it follows from (14) that the local social norm will become less favourable if the inter-marriages within the minority group decreases faster in proportional terms than the inter-marriages increases in proportional terms in the majority group.

In order to visualize the location and spread of social norms relative to the underlying principal preferences Figure 12 plots the distribution of both the principal, type-profile dependent joint marital utility $\zeta(x_i, x_j)$, and the distribution of the corresponding social norms. As there are $N = 6$ types in total, there are $N^2 = 36$ principal utility terms. These are marked in blue squares and organized into nine panels, each panel corresponding to a couple ethnicity profile. Within each panel there are four rows, each corresponding to a qualification profile. For marriage profiles where both partners belong to the same ethnic group (the panels on the lead diagonal), there is a distinctive 'U-shaped' pattern in $\zeta(x_i, x_j)$. This reflects the principal incentive for marital sorting on qualifications. For ease of comparison, the figure uses the same scale in each panel. This means that a few the marital utility terms for marriages between Blacks and Asian are "below the scale" and hence not shown.

Figure 12 here.

For any given marriage type-profile (x_i, x_j) the equilibrium local social norm in region g with respect to this particular type profile is given by $\sigma^g(x_i, x_j)$ as defined in (14). The distribution of these local social norms across regions, for each type-profile, illustrated in red dots, with a vertical Black line indicating the mean value.

A few observations are worth noting. First, for inter-ethnic marriages involving White women, social norms generally work less against marriages to Black men than against marriages to Asian men. In contrast, for inter-ethnic marriages involving White men, social norms are of similar magnitude with respect to marriages to Black and Asian women (though more variable for the former). Second, the figure illustrates that the spread in social norms across regions can be substantial. Third, for every ethnicity profile (except the very rare marriages between Blacks and Asians), the utility terms are generally higher for marriages where the partners share the same qualification level – reflected in the plotted terms being U-shaped. But this pattern is less pronounced for inter-ethnic marriages, particularly between Whites and Blacks.

Region-Specific Preferences

Panel B of Table 7 presents the estimated region-specific marriage utility parameters, δ_g for $g \in G$ As discussed in Section IV one δ_g must be normalized and used as reference group. We use the North East region of England for this purpose, normalizing its parameter value to zero. Therefore the remaining parameters can be interpreted as the difference in the systematic joint utility from a marriage in region g relative to one in the North East of England.

Note specifically, that since δ_g does not vary with partner type-profile it mainly acts as preference shift term for marriage versus singlehood. Indeed, the estimated δ_g-terms, while precisely estimated, are with one exception relatively small compared to the other systematic marriage utility terms illustrated in Figure 12. The one exception is London δ_{London} which is strongly negative, reflecting the larger proportion single in that region (see Figure 7).

Preferences for Non-UK-Born Partners

Finally, for completeness, Panel C of Table 7 provides the estimated preferences for marrying a non-UK-born partner, by gender and own type. These parameters are also precisely estimated and with the expected ethnic pattern. For White and Black men and women the estimated preferences are negative reflecting that marrying a non-UK-born partner is a rare choice. In constrast, for Asians the estimated preferences are positive for low qualified men and women and zero for high qualified, reflecting the empirical finding tha marrying a non-UK-born partner is a common choice among Asians, particularly among those with low qualifications.

VII Predicted Future Ethnic Marital Sorting

In this final section we will use our estimated model to predict future ethnic integration through marriage in the UK. Specifically, we will use the parameters from the estimated model alongside the observed demographic shifts – both ethnicity shares and in qualifications – to predict marriage patterns among individuals born between 1990 and 2006. We will simply refer to these younger individuals as the "recent cohorts" and we present a demographic characterization of these cohorts in Appendix B.[16] As the ethnic minority groups have continued to grow in recent cohorts relative to the estimating cohorts, the aim of the exercise is to explore whether the model will predict how these population trends can be expected to affect the future marital mixing and, in this sense, the process of integration. In doing so we will also explore what role endogenously evolving social norms are expected to play in this process.

Figure 3 above outlined how the aggregate share of Asians in particular, and the share of Blacks to a lesser extent, grew between our estimating- and the recent cohorts. Table 8 describes the relative aggregate type ethnicity-qualification type frequencies, by gender, in the

[16] Noting the some of these recent cohorts have yet to complete their education, Appendix B also provides details of how we impute predicted distributions of completed education by gender, ethnicity and region.

estimating cohorts and the recent cohorts respectively. The table highlights both the changing ethnic composition and the increase in qualifications. In Appendix B we further show the underlying geographical variation, highlighting that the aggregate trend towards more ethnic diversity permeates practically all regions.

Table 8 here.

In line with our discussion above, we will continue to use the share of intra-ethnic marriages as our basic measure of marital sorting. Before showing the model-predicted marital sorting in the recent cohorts, it is instructive to outline some of the mechanisms by which the demographic shifts can be expected to affect future ethnic marital sorting.

From Figure 11 we know that for both Asians and Blacks, as an empirical stylized fact, the share of marriages that are intra-ethnic increases with the own ethnicity's share of the local population. Conversely, for Whites, as the ethnic majority group, the larger is the share of non-Whites the larger is the proportion who marry inter-ethnically. The same figure shows that the estimated model naturally replicates these empirical patterns. As a consequence, in terms of projections for the recent cohorts, as the shares of Asians and Blacks grow relative to the share of Whites, the model will naturally predict that the ethnic minorities will become less likely to marry inter-ethnically whilst Whites will becomes more likely to do so. These direct supply effects will in turn have an ambiguous effect on social norms – see equation (14) – making them relatively more favourable towards inter-ethnic marriages among the majority Whites but less so among the ethnic minorities themselves. Hence evolving social norms may either boost future integration through inter-ethnic marriages or reduce it.

With respect to qualifications, our estimated model did not uncover any marked differences in the strength of primitive preferences for ethnic marital sorting across qualification levels (see Table 6). Consequently, the increase in the rate of holding high qualifications between the estimating and recent cohorts can be expected to have at most a minor effect on the future proportion of inter-ethnic marriages.

In Figure 13 we plot predicted shares of intra-ethnic marriages, by ethnicity and gender, across regions against the own ethnicity share. Indeed, each subfigure presents three sets of predictions. The first set, using square markers, illustrates the model predictions for the estimating cohorts and are thus carried forward from Figure 11 as benchmark. The second set, using diamond-shaped markers, illustrates the model predictions for the recent cohorts. Notably, these predictions include predicted equilibrium future social norms. Specifically to highlight the role of the evolving social norms, we also present a third set of predictions, using triangular markers. This set of predictions is computed holding each social norm term $\sigma^g(x_i, x_j)$ – that is, the norm for each region $g \in G$ and each partner type profile $(x_i, x_j) \in X \times X$ – fixed at the corresponding estimated value for the estimating cohorts.

Looking first at the Asians (panels e and f), we see that, due to their increasing population shares there is a general movement to the right between the estimating and the recent cohorts. For instance, in the estimating cohorts, the share of Asians was below 4 percentage points in the majority of areas. In contrast, in the recent cohorts only four out of the eleven regions now have less than 4 percent Asians. Similarly, in London the share of Asians nearly doubled from just below 10 percent to 19 percent. The predicted direct effect (holding social norms constant) for the rate of intra-ethnic marriages among the Asians is most clearly positive for women where, as the Asians' share of the local population increases, the predicted share of intra-ethnic marriages grows steadily. This predicted direct effect is then further boosted by evolving social norms to the extent that the model predicts that that proportion of intra-ethnic marriages among Asian women in the recent cohorts will be 80 percent or higher in every region. A similar prediction holds for the Asian men.

Turning to the Blacks, Figure 11 (panels c and d) shows that the growth of their population share has been much more modest in most areas: in the estimating cohorts, their share was below 2 percent in all regions except London. In the more recent cohorts, Blacks' share of the local population is still less than 2 percent in all but two areas. Consequently, in terms of predicted shares of intra-ethnic marriages among Black, only modest increases – amplified by changing social norms – are predicted in most areas. For London, there has been a sharp increase in the share of Blacks in the local population – from just over 8 percent in the estimating cohorts to over 18 percent in the recent cohorts. But for London, the predicted share of intra-ethnic marriages among Blacks was high – around 80 percent – already for the estimating cohorts, and is predicted to further increase only relatively marginally in the recent cohorts.

Finally, for the Whites, their population shares were 95 percent or higher in all areas except two in the estimating cohorts. In contrast, in the recent cohorts the share of Whites is at 95 percent or higher only in four areas. As a result, their predicted local rates of intra-ethnic marriage decrease, and this is further amplified by predicted evolving social norms. Nevertheless, the predicted rate of intra-ethnic marriage remains above 95 percent for both men and women in all regions, including London.

Table 9 here.

Table 9 aggregates the predictions over regions to show predicted shares of marriages that are intra-ethnic by gender, ethnicity and qualification level both within the estimating cohorts and the recent cohorts. This confirms that the Whites are predicted to increase their rate of marrying inter-ethnically while Asians in particular are predicted to increase their rate of marrying intra-ethnically.

Finally, if we aggregate over all marriages in the estimating and the recent cohorts respectively, we find that the predicted fraction of all marriages (between UK-born partners) that are inter-ethnic doubles from about 1.3 percentage points to about 2.6 percentage points. From Table 9 it is clear that most of this aggregate increase in inter-ethnic marriages comes from changes in marital behaviour for both low- and high-qualified individuals rather than from the increase qualifications between the two cohorts. Hence overall as a proportion of all marriages, inter-ethnic marriages are predicted to increase. However, this is still consistent with a predicted decrease in the rates of inter-ethnic marriages among both Blacks and Asians.

VIII Conclusions

As Western economies become more ethnically diverse, will inter-ethnic marriages act as a force for long-term integration, breaking down barriers between natives and immigrants? The UK is a natural example to consider due to its distinct history of post-war Black and Asian immigration.

That history created a particular and persistent regional variation in ethnic composition that has carried over to the second and beyond generations of immigrants. Most of the second generation immigrants who were born through the 1960s and 1970s have now gone on to marry and we can study how their marital choices reflect the ethnic composition of their local populations. For example, London is the UK region with by far the largest population shares of both Blacks and Asians. It is then striking that the rate of inter-ethnic marriage is lower among Blacks and Asians in London compared to other regions of the UK; this suggests that higher ethnic minority densities may not foster higher rates of inter-ethnic marriage among members of the minority groups. Furthermore marrying inter-ethnically can also be argued to go against social norms, raising the question of whether, as the ethnic minorities grow in relative size, social norms can be expected to change and contribute to the integration process.

To answer these question we have set up and estimated a structural model of the marriage market, building on the workhorse model of Choo and Siow (2006), and extended with endogenous social norms modelled as conformity preferences. Using that the geographical distribution of UK-born ethnic minorities reflects the settlement patterns of the first generation post-war immigrants, we found that Blacks and Asians are systematically less likely to marry inter-ethnically if they live in regions where their own ethnicity is comparatively larger. The estimates suggest strong ethnic complementarities – stronger than for qualifications – reflected in strong assortative mating on ethnicity. We also found that allowing for endogenous social norms provided a better fit to the data.

We used the estimated model to predict marital patterns among a set of more recent cohorts in order to consider whether more integration through marriage can be expected going forward. Here we found that, although inter-ethnic marriages can be expected to grow as a share of all marriages, members of the ethnic minority groups are predicted to become less likely to marry inter-ethnically, a result that is also amplified by evolving endogenous social norms. In this sense, the model does not suggest that the Black and Asian populations will become increasingly integrated with the White majority group through the formation of marriages.

A caveat to the current approach is of course that it assumes stable preferences: at the heart of the model is a set of primitive preference parameters that describe how the systematic joint marriage utility relate to partner type profile, and these preferences are assumed not to change. Viewed from this perspective, the results presented above can be argued to show that future integration through marriage will require that a change of fundamental preferences. Indeed, the recent work by Merlino et al. (2019), using quasi-random variation in ethnic exposure during childhood, show that exposure to Black peers when young lead whites to have more relationships with Blacks as adults. The authors argue that their results reflect underlying effects on attitudes which, translated to our setting, is more akin to a change in the primitive preferences than to a change social norms. That suggests that the evolution of both population supplies and of preferences and attitudes will play a key role in shaping the future patterns of inter-ethnic marriages. It also highlights the research challenge of modelling changing population supplies and changing of preferences and attitudes jointly as determinants of the evolution of the ethnic marital mixing.

References

Adamopoulou, E. (2012), 'Peer effects in young adults' marital decisions'. Universidad Carlos III de Madrid.

Ager, A. and Strang, A. (2004), 'Indicators of integration'. Home Office: Development and Practice Report No.28.

Becker, G. (1973), 'A theory of marriage: Part 1', *Journal of Political Economy* **81**(4), 813–846.

Bisin, A., Topa, G. and Verdier, T. (2004), 'Religious intermarriage and socialization in the United States', *Journal of Political Economy* **112**(3), 615–664.

Blau, P., Beeker, C. and Fitzpatrick, K. M. (1984), 'Intersecting social affiliations and intermarriage', *Social Forces* **62**(3), 585–606.

Brandt, L., Siow, A. and Vogel, C. (2016), 'Large shocks and small changes in the marriage market', *Journal of the European Economic Association* **14**(6), 1437–68.

Brien, M. (1997), 'Racial differences in marriage and the role of marriage markets.', *Journal of Human Resources* **32**(4), 741–778.

Brock, W. and Durlauf, S. (2001), 'Discrete choice with social interactions', *The Review of Economic Studies* **68**(2), 235–260.

Burke, M. A. and Young, H. P. (2011), Social norms, *in* A. Bisin, J. Benhabib and M. Jackson, eds, 'The Handbook of Social Economics', Vol. 1, Elsevier, chapter 8, pp. 311–338.

Charsley, K., Storer-Chuch, B., Benson, M. and Van Hear, N. (2012), 'Marriage-related migration in the UK', *International Migration Review* **46**(4), 861–890.

Chiappori, P.-A., Salanié, B. and Weiss, Y. (2017), 'Partner choice, investment in children, and the marital college premium', *American Economic Review* **107**(8), 2109–21067.

Choo, E. (2015), 'Dynamic marriage matching: An empirical framework', *Econometrica* **83**(4), 1373–1423.

Choo, E. and Siow, A. (2006), 'Who marries whom and why?', *Journal of Political Economy* **114**(1), 175–201.

Decker, C., Lieb, E., McCann, R. and Stephens, B.-K. (2013), 'Unique equilibria and substitution effects in a stochastic model of the marriage market', *Journal of Economic Theory* **148**(2), 778–792.

Dupuy, A. and Galichon, A. (2014), 'Personality traits and the marriage market', *Journal of Political Economy* **122**(6), 1271–1319.

Dupuy, A. and Weber, S. (2018), 'Marital patterns and income inequality', IZA DP No. 11572.

Eika, L., Mogstad, M. and Zafar, B. (2018), 'Educational assortative mating and household income inequality', *Journal of Political Economy* . Forthcoming.

Ellwood, D. and Crane, J. (1990), 'Family change among black Americans: What do we know?', *Journal of Economics Perspectives* **4**(4), 65–84.

Fryer, R. (2007), 'Guess who's been coming to dinner? Trends in interracial marriage over the 20th century', *Journal of Economic Perspectives* **21**(2), 71–90.

Galichon, A. and Salanié, B. (2015), 'Cupid's invisible hand: Social surplus and identification in matching models', *Columbia University* .

Glaeser, S. and Scheinkman, J. (2014), 'Measuring social interactions', *in S. Durlauf and P. Young (eds.) "Social Dynamics", MIT Press* .

Graham, B. (2013), 'Comparative static and computational methods for an empirical one-to-one transferable utility matching model', *Advances in Econometrics* **31**, 153–181.

Greenwood, J., Guner, N., Kocharkov, G. and Santos, C. (2014), 'Marry your like: Assortative mating and income inequality', *American Economic Review* **104**(5), 348–353.

Hansen, R. (2000), *Citizenship and Immigration in Post-war Britain: The Institutional Origins of a Multicultural Nation*, Oxford University Press.

Lam, D. (1988), 'Marriage markets and assortative mating with household public goods: Theoretical results and empirical implications', *Journal of Human Resources* **23**(4), 482–487.

Liu, H. and Lu, J. (2006), 'Measuring the degree of assortative mating', *Economics Letters* **92**, 317–322.

Manski, C. (1993), 'Identification of endogenous social effects: The reflection problem', *The Review of Economic Studies* **60**(3), 531–542.

Merlino, L. P., Steinhardt, M. F. and Wren-Lewis, L. (2019), 'More than just friends? School peers and adult interracial relationships', *Journal of Labor Economics* . Forthcoming.

Mourifié, I. (2017), 'A marriage matching function with flexible spillover and substitution patterns', *University of Toronto - Working Paper* .

Mourifié, I. and Siow, A. (2017), 'The Cobb-Douglas marriage matching function: Marriage matching with peer and scale effects', *Working Paper* .

ONS (2012), '2011 Census: Key statistics for local authorities in England and Wales'. Office for National Statistics, Statistical Bulletin.

ONS (2016), 'Internal migration, England and Wales: Year ending June 2015'. Office for National Statistics, Statistical Bulletin.

Phillimore, J. (2012), 'Implementing integration in the UK: lessons for integration, theory, policy and practice', *Policy and Politics* **40**(4), 525–545.

Qian, Z. (2005), 'Breaking the last taboo: interracial marriage in America', *Contexts* **4**(4), 33–37.

Saluter, A. (1994), 'Marital status and living arrangements: March 1993.', *U.S. Bureau of the Census, Current Population Reports* pp. 20–478.

Seitz, S. (2009), 'Accounting for racial differences in marriage and employment.', *Journal of Labor Economics* **27**(3), 385–437.

Appendix A: Proof of Identification

By Assumption 3, we have access to at least two groups with non-identical relative population supplies. Without loss of generality, we can assume that these are $g = 1, 2$ and we can ignore any further groups as two groups will be sufficient for identification. With two groups, $N \geq 2$ types, and a common principal marriage utility $\zeta^g(x_i, x_j) = \zeta(x_i, x_j)$ for every possible marriage profile $(x_i, x_j) \in X \times X$ (Assumption 1), we have $2N^2$ matching equations of the form (10). We can collect these matching equations in matrix form by defining the following matrices (each of size $N^2 \times 2$),

$$M^{g,m} \equiv \begin{bmatrix} \log\left(\omega^{g,m}_{x_1|x_1}\right) & -\log\left(\omega^{g,m}_{0|x_1}\right) \\ \vdots & \vdots \\ \log\left(\omega^{g,m}_{x_N|x_1}\right) & -\log\left(\omega^{g,m}_{0|x_1}\right) \\ \vdots & \vdots \\ \log\left(\omega^{g,m}_{x_1|x_N}\right) & -\log\left(\omega^{g,m}_{0|x_N}\right) \\ \vdots & \vdots \\ \log\left(\omega^{g,m}_{x_N|x_N}\right) & -\log\left(\omega^{g,m}_{0|x_N}\right) \end{bmatrix} \text{ and } M^{g,f} \equiv \begin{bmatrix} \log\left(\omega^{g,f}_{x_1|x_1}\right) & -\log\left(\omega^{g,f}_{0|x_1}\right) \\ \vdots & \vdots \\ \log\left(\omega^{g,f}_{x_1|x_N}\right) & -\log\left(\omega^{g,f}_{0|x_N}\right) \\ \vdots & \vdots \\ \log\left(\omega^{g,f}_{x_N|x_1}\right) & -\log\left(\omega^{g,f}_{0|x_1}\right) \\ \vdots & \vdots \\ \log\left(\omega^{g,f}_{x_N|x_N}\right) & -\log\left(\omega^{g,f}_{0|x_N}\right) \end{bmatrix}, \tag{A1}$$

and the $N^2 \times 1$ column vector,

$$B^g \equiv \begin{bmatrix} \log\left(\frac{\omega^{g,m}_{x_1|x_1}}{\omega^{g,m}_{0|x_1}} \frac{\omega^{g,f}_{x_1|x_1}}{\omega^{g,f}_{0|x_1}}\right) & \cdots & \log\left(\frac{\omega^{g,m}_{x_N|x_1}}{\omega^{g,m}_{0|x_1}} \frac{\omega^{g,f}_{x_1|x_N}}{\omega^{g,f}_{0|x_N}}\right) & \cdots & \log\left(\frac{\omega^{g,m}_{x_1|x_N}}{\omega^{g,m}_{0|x_N}} \frac{\omega^{g,f}_{x_N|x_1}}{\omega^{g,f}_{0|x_1}}\right) & \cdots & \log\left(\frac{\omega^{g,m}_{x_N|x_N}}{\omega^{g,m}_{0|x_N}} \frac{\omega^{g,f}_{x_N|x_N}}{\omega^{g,f}_{0|x_N}}\right) \end{bmatrix}'. \tag{A2}$$

The $2N^2$ matching equations can then be stacked and written succinctly as $A\boldsymbol{\theta} = B$ where

$$A \equiv \begin{bmatrix} I_{N^2} & M^{1,m} & M^{1,f} \\ I_{N^2} & M^{2,m} & M^{2,2} \end{bmatrix}, \qquad \text{and} \qquad B \equiv \begin{bmatrix} B^1 \\ B^2 \end{bmatrix}, \tag{A3}$$

where I_n is the $n \times n$ identity matrix, and where $\boldsymbol{\theta}$ is as defined in (9). Note that the matrix A has dimension $2N^2 \times \left(N^2 + 4\right)$ and contains observable marriage and singles proportions. Similarly, B is $2N^2 \times 1$ and contains observable marriage and singles proportions.

Consider then the rank of the matrix A. Due to the indeterminacy of the scale of the parameters and the need for a normalization (Assumption 2), A has at most one less than full rank, $rank\,(A) \leq N^2+3$. Under Assumption 3, that is, groups $g=1,2$ have different relative population supplies A generically achieves the maximum rank, $rank\,(A) = N^2+3$. Hence it is possible to select N^2+3 independent equations from the $2N^2$ available matching equations, reducing A to some A_0 and B to a corresponding B_0, and obtaining an invertible system $A_0\boldsymbol{\theta}_0 = B_0$, and recovering the normalized $\boldsymbol{\theta}_0$ as $\boldsymbol{\theta}_0 = A_0^{-1}B_0$ where $\boldsymbol{\theta}_0$ is N^2+3 non-normalized parameters. Note that if the two groups had identical relative population supplies, the lower halves of the matrix A and B will be identical to their upper halves, reducing the rank of A to N^2 and identification would be lost. #

Tables

Table 1: Descriptive Statistics of the QLFS Sample

	Part A: Males					
	White		**Black**		**Asian**	
	Low Qual	**High Qual**	**Low Qual**	**High Qual**	**Low Qual**	**High Qual**
Age in Years	35.91	35.31	36.81	36.45	34.49	33.81
	(5.27)	(5.30)	(5.51)	(4.94)	(5.09)	(4.89)
Cohort	71.22	72.08	70.37	71.09	73.00	73.92
	(5.17)	(5.26)	(5.19)	(5.13)	(5.19)	(5.08)
Married -	0.64	0.67	0.32	0.36	0.28	0.38
UK-Born Partner	(0.48)	(0.48)	(0.47)	(0.48)	(0.45)	(0.49)
Married -	0.03	0.06	0.07	0.15	0.47	0.31
Non-UK-Born Partner	(0.16)	(0.23)	(0.47)	(0.36)	(0.50)	(0.46)
Obs.	44,903	49,316	447	579	963	1,737
	Part B: Females					
	White		**Black**		**Asian**	
	Low Qual	**High Qual**	**Low Qual**	**High Qual**	**Low Qual**	**High Qual**
Age in Years	35.26	34.45	36.60	36.12	33.88	33.16
	(5.35)	(5.41)	(5.26)	(5.47)	(5.35)	(4.96)
Cohort	71.71	73.10	70.35	71.43	73.19	74.97
	(5.23)	(5.44)	(5.11)	(5.43)	(5.36)	(5.24)
Married -	0.55	0.66	0.14	0.20	0.21	0.35
UK-Born Partner	(0.50)	(0.47)	(0.34)	(0.40)	(0.41)	(0.48)
Married -	0.02	0.04	0.08	0.12	0.52	0.31
Non-UK-Born Partner	(0.14)	(0.19)	(0.27)	(0.33)	(0.50)	(0.46)
Obs.	45,815	55,456	519	845	1,179	2,043

Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born between 1965-1989, living in England, Scotland or Wales, and aged 25 or above when observed, and with available information on gender, ethnicity, educational attainment, marital status. Standard deviations in parenthesis.

Table 2: Aggregate Empirical Assortative Mating

	Panel A: Males				
	Ethnicity			Qualification	
	White	**Black**	**Asian**	**Low**	**High**
Prop.	0.995	0.390	0.790	0.582	0.746
	(0.000)	(0.026)	(0.013)	(0.003)	(0.002)
Obs.	61,760	352	933	29,127	33,918
	Panel B: Females				
	Ethnicity			**Qualification**	
	White	**Black**	**Asian**	**Low**	**High**
Prop.	0.994	0.578	0.768	0.663	0.675
	(0.000)	(0.032)	(0.014)	(0.003)	(0.002)
Obs.	61,844	237	955	25,542	37,494

Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born between 1965-1989, living in England, Scotland or Wales, and aged 25 or above when observed, and with available information on gender, ethnicity, educational attainment, marital status. Standard errors in parenthesis.

Table 3: The Effect of Own Ethnicity Share on the Probability of Being Intra-Ethnically Married

	Panel A: White					
	Males			Females		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
Prop. Own. Ethn.	0.066***	0.064***	0.062***	0.074***	0.074***	0.073***
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
High Qual.			-0.003***			-0.001
			(0.001)			(0.001)
Obs.		61,760			61,844	
Est. Spec.	Probit	IV Probit	Probit	Probit	IV Probit	Probit
	Panel B: Black					
	Males			Females		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
Prop. Own. Ethn.	4.234***	4.185***	4.223***	3.683***	3.548***	3.712***
	(0.579)	(0.582)	(0.576)	(0.772)	(0.788)	(0.770)
High Qual.			0.092*			0.049
			(0.049)			(0.066)
Obs.		352			237	
Est. Spec.	Probit	IV Probit	Probit	Probit	IV Probit	Probit
	Panel C: Asian					
	Males			Females		
	(i)	(ii)	(iii)	(i)	(ii)	(iii)
Prop. Own. Ethn.	3.234***	2.120**	3.271***	2.200***	0.330	2.234***
	(0.531)	(0.588)	(0.530)	(0.539)	(0.526)	(0.530)
High Qual.			-0.008			-0.109***
			(0.028)			(0.033)
Obs.		933			955	
Est. Spec.	Probit	IV Probit	Probit	Probit	IV Probit	Probit

Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born between 1965-1989, living in England, Scotland or Wales, aged 25 or above and married to a UK-born partner, and with available information on gender, ethnicity, educational attainment, marital status. The proportion of the local population who are of the own ethnicity ("Prop. Own Ethn.) is measured using the same data but without conditioning on being married. Specification (i) includes no further controls. Specification (ii) instruments for the own ethnicity proportion with the proportion of own ethnicity in the local population among all (UK- and non-UK-born) individuals born 1940-1960 as observed in the QLFS 1996-2015. Specification (iii) includes a dummy variables for own qualification level being high (A-Level+) as control.

Table 4: Alternative Empirical Model Specifications

	Specification			
	(i)	**(ii)**	**(iii)**	**(iv)**
Log Likelihood Value	-242,775	-242,029	-242,020	-242,020
LR Test Statistic		1,490.5	19.2	0.2
p-value		<0.001	<0.001	0.920
Mean Squared Error (MSE)×100	0.349	0.339	0.329	0.329
Regional Effects (δ_g, $g \in G$)	N	Y	Y	Y
Marriage Social Preferences (ϕ_1^k, $k = m, f$)	N	N	Y	Y
Singlehood Social Preferences (ϕ_0^k, $k = m, f$)	N	N	N	Y

Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born between 1965-1989, living in England, Scotland or Wales, and aged 25 or above when observed, and with available information on gender, ethnicity, educational attainment, marital status. The LR test statistic and associated p-value reported for each specification (ii) - (iv) tests whether the constrained model in the previous column is statistically rejected.

Table 5: Estimates of Principal Joint Marital Utility by Ethnicity-Qualification Profile

		Female Type					
		White, GCSE(-)	White, A-Level(+)	Black, GCSE(-)	Black, A-Level(+)	Asian, GCSE(-)	Asian, A-Level(+)
Male Type	White, GCSE(-)	0.81*** (0.04)	0.69*** (0.05)	-4.51*** (0.28)	-4.42*** (0.27)	-3.84*** (0.26)	-4.20*** (0.24)
	White, A-Level(+)	0.22*** (0.05)	1.71*** (0.03)	-4.61*** (0.30)	-3.29*** (0.22)	-4.27*** (0.31)	-2.04*** (0.17)
	Black, GCSE(-)	-3.52*** (0.21)	-3.41*** (0.22)	-2.39* (0.19)	-2.38*** (0.20)	-5.55*** (0.74)	-4.76*** (0.52)
	Black, A-Level(+)	-3.35*** (0.21)	-2.87*** (0.20)	-2.52*** (0.23)	-1.00* (0.14)	-32.59 (769)	-4.00*** (0.41)
	Asian, GCSE(-)	-4.20*** (0.24)	-3.86*** (0.24)	-5.47*** (0.74)	-32.4 (734)	0.15 (0.13)	-0.39*** (0.11)
	Asian, A-Level(+)	-4.95*** (0.25)	-2.67*** (0.16)	-33.01 (736)	-33.3 (737)	-0.49*** (0.11)	0.73*** (0.07)

Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born between 1965-1989, living in England, Scotland or Wales, aged 25 or above when observed, and with available information on gender, ethnicity, educational attainment, marital status. Based on the preferred specification (iii) in Table 4. Standard errors in parenthesis.

Table 6: Complementarities in Ethnicity and Qualifications

	Panel A: Ethnicity Complementarity		
Qual. Profile	**White, Black**	**White, Asian**	**Black, Asian**
Low, Low	6.45***	8.99***	8.78***
	(0.27)	(0.29)	(1.88)
High, Low	7.34***	9.98***	64.16
	(0.37)	(0.38)	(177.6)
Low, High	7.58***	10.70***	36.12**
	(0.22)	(0.42)	(21.24)
High, High	6.86***	7.15***	36.99***
	(0.18)	(0.21)	(0.32)

	Panel B: Qualification Complementarity		
	Wife Ethnicity		
Husb. Ethnicity	**White**	**Black**	**Asian**
White	1.60***	1.23***	2.60***
	(0.02)	(0.28)	(0.23)
Black	0.37***	1.52***	27.79
	(0.14)	(0.36)	(92.49)
Asian	1.94***	26.71	1.77***
	(0.21)	(1036)	(0.06)

Notes: Panel A presents the estimates of $\eta\left(z, z'; q_i, q_j\right)$ while Panel B presents the estimates of $\eta\left(q, q'; z_i, z_j\right)$. See text for definitions. Based on the preferred specification (iii) in Table 4. Standard errors in parenthesis.

Table 7: Estimates of Remaining Preference Parameters

Panel A: Social Preferences					
		ϕ_1^m	ϕ_1^f		
		0.47***	0.38***		
		(0.03)	(0.03)		
Panel B: Regional-Specific Preferences					
δ_{London}	$\delta_{\text{West Midl.}}$	$\delta_{\text{East Midl.}}$	$\delta_{\text{South East}}$	$\delta_{\text{North West}}$	$\delta_{\text{East Engl.}}$
-0.99***	0.10***	0.26***	0.11***	-0.13***	0.14***
(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
$\delta_{\text{Yorks.\&Humb.}}$	$\delta_{\text{South West}}$	$\delta_{\text{North East}}$	δ_{Wales}	δ_{Scotland}	
0.12***	0.03	0 (ref)	0.02	-0.22***	
(0.04)	(0.04)	-	(0.04)	(0.04)	
Panel C: Preferences for Non-UK-Born Spouse					
Males ($\xi(x_i, -1)$)					
White, GCSE(-)	White, A-Level(-)	Black, GCSE(-)	Black, A-Level(-)	Asian, GCSE(-)	Asian, A-Level(-)
-2.59***	-1.57***	-2.17***	-1.16***	0.60***	-0.01
(0.03)	(0.02)	(0.19)	(0.12)	(0.08)	(0.06)
Females ($\xi(-1, x_j)$)					
White, GCSE(-)	White, A-Level(-)	Black, GCSE(-)	Black, A-Level(-)	Asian, GCSE(-)	Asian, A-Level(-)
-2.79***	-1.93***	-2.09***	-1.42***	0.72***	-0.03
(0.03)	(0.02)	(0.15)	(0.10)	(0.06)	(0.04)

Notes: Based on the preferred specification (iii) in Table 4. Standard errors in parenthesis.

Table 8: The Distribution of Ethnicity and Qualifications in the Estimating and Recent Cohorts

		Birth Cohorts			
		1960-1989		**1990-2006**	
		Males	**Females**	**Males**	**Females**
White	GCSE(-)	0.458	0.433	0.173	0.158
	A-Level(+)	0.503	0.524	0.739	0.754
	Total	0.961	0.957	0.912	0.912
Black	GCSE(-)	0.005	0.005	0.003	0.002
	A-Level(+)	0.006	0.008	0.022	0.022
	Total	0.011	0.013	0.025	0.024
Asian	GCSE(-)	0.010	0.011	0.007	0.006
	A-Level(+)	0.018	0.019	0.056	0.058
	Total	0.028	0.030	0.063	0.064

Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born between 1965-2006, living in England, Scotland or Wales, and with available information on gender, ethnicity, and educational attainment.

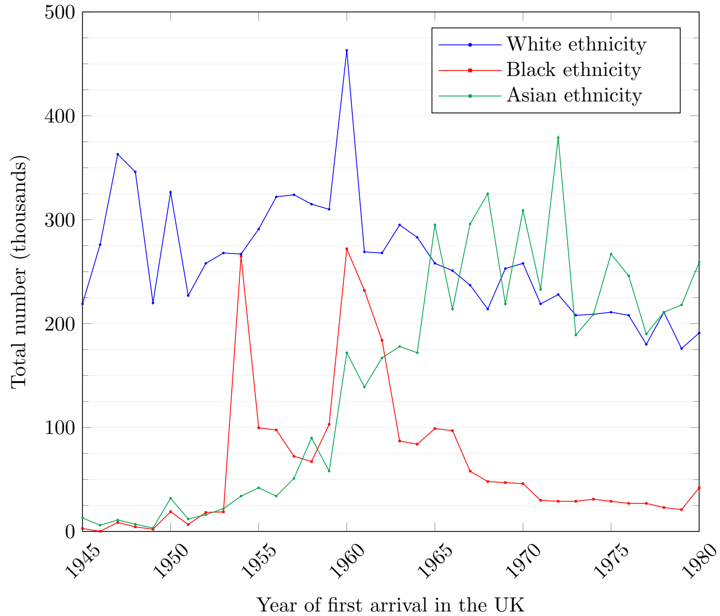
Table 9: Predicted Shares of Intra-Ethnic Marriages by Gender, Ethnicity, and Qualification in the Estimating and Recent Cohorts

	Panel A: Males					
	White		**Black**		**Asian**	
	Low Qual.	**High Qual.**	**Low Qual.**	**High Qual.**	**Low Qual.**	**High Qual.**
Born 1965-1989	0.994	0.993	0.551	0.592	0.837	0.734
Born 1990-2006	0.979	0.978	0.796	0.868	0.946	0.882
Difference	-0.015	-0.015	0.245	0.276	0.109	0.148
	Panel B: Females					
	White		**Black**		**Asian**	
	Low Qual.	**High Qual.**	**Low Qual.**	**High Qual.**	**Low Qual.**	**High Qual.**
Born 1965-1989	0.997	0.993	0.320	0.435	0.800	0.794
Born 1990-2006	0.986	0.973	0.674	0.794	0.936	0.931
Difference	-0.011	-0.020	0.354	0.359	0.136	0.137

Notes: Predictions for "estimating cohorts" (born 1965-1989) and for "recent cohorts" (born 1990-2006) based on parameter estimates from the preferred specification (iii) in Table 4. See Appendix B for a description of the demographic composition of the recent cohorts.

Figures

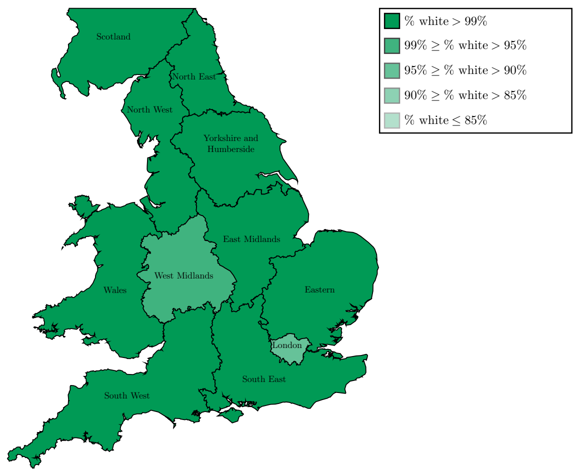
Figure 1: Number of Arriving Immigrants by Ethnic Group and Year of Arrival



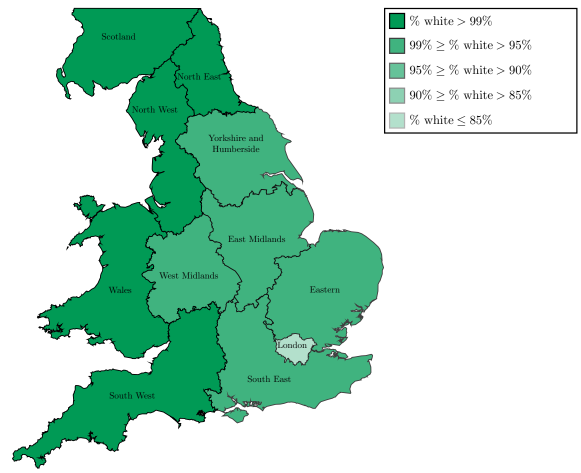
Notes: Authors' calculations based on the sample of all non-UK born individuals observed in the QLFS 1996-2015, living in England, Scotland or Wales, with information on ethnicity and year or arrival. Population weights were used to obtain estimated gross number of arriving individuals by ethnicity and year of arrival.

Figure 2: Ethnic Composition by Region

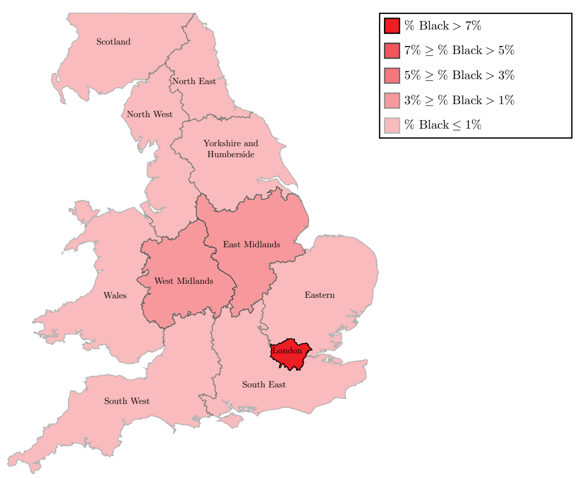
(a) Prop. Whites, Born 1950 - 1964



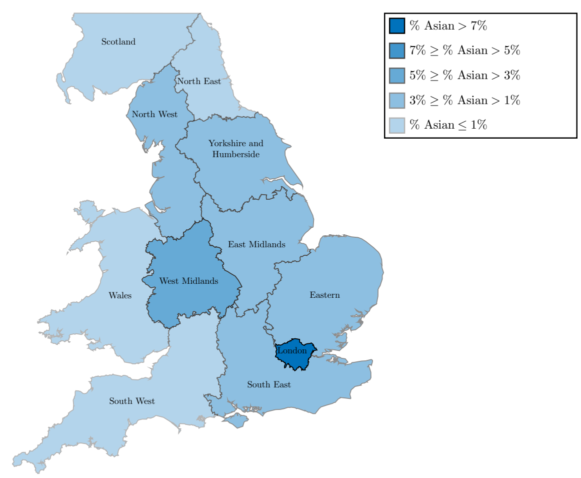
(b) Prop. Whites, Born 1965 - 1989



(c) Prop. Blacks, Born 1965 - 1989

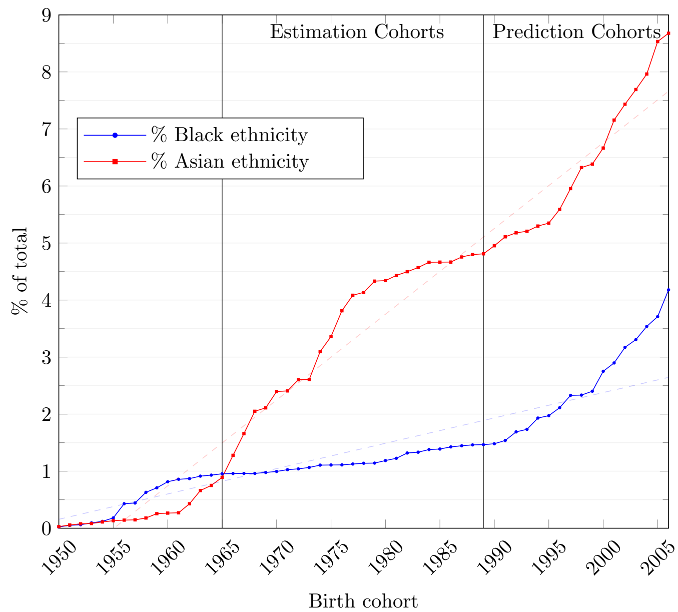


(d) Prop. Asians, Born 1965 - 1989



Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born in the indicated cohorts, living in England, Scotland or Wales, and aged 25 or above when observed, and with available information on gender, ethnicity, educational attainment, marital status.

Figure 3: Ethnicity Proportions by Birth Cohort



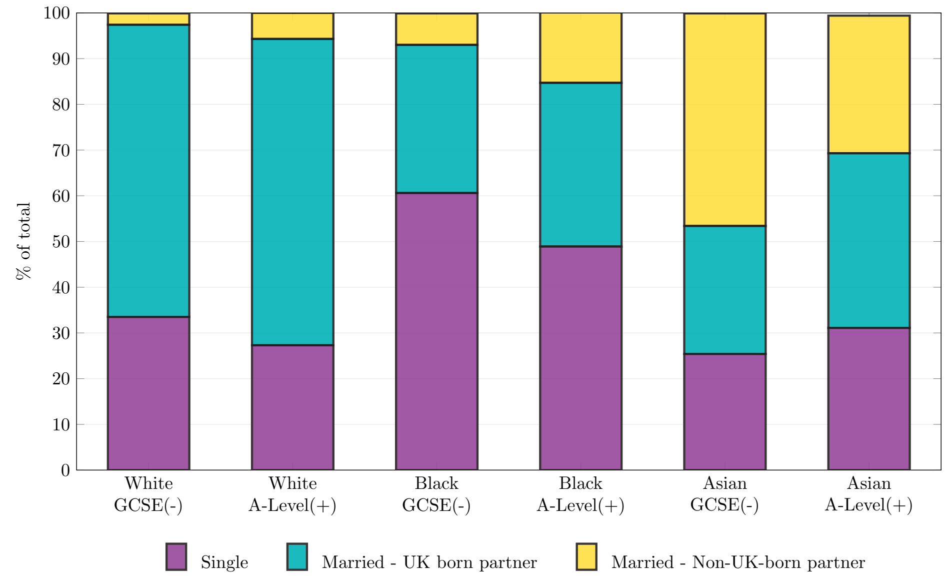
Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born between 1950 and 2006, living in England, Scotland or Wales, and with available information on ethnicity.

Figure 4: Proportion of Individuals with High Educational Attainment by Region and Gender

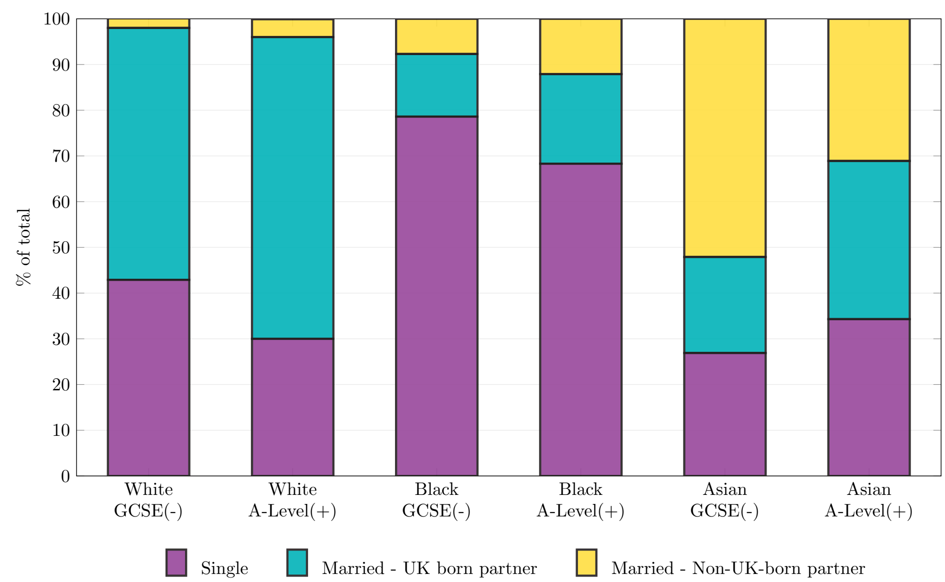
Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born between 1965-1989, living in England, Scotland or Wales, and aged 25 or above when observed, and with available information on gender, ethnicity, educational attainment, marital status.

Figure 5: Distribution of Marital Status by Own Type and Gender

(a) Males



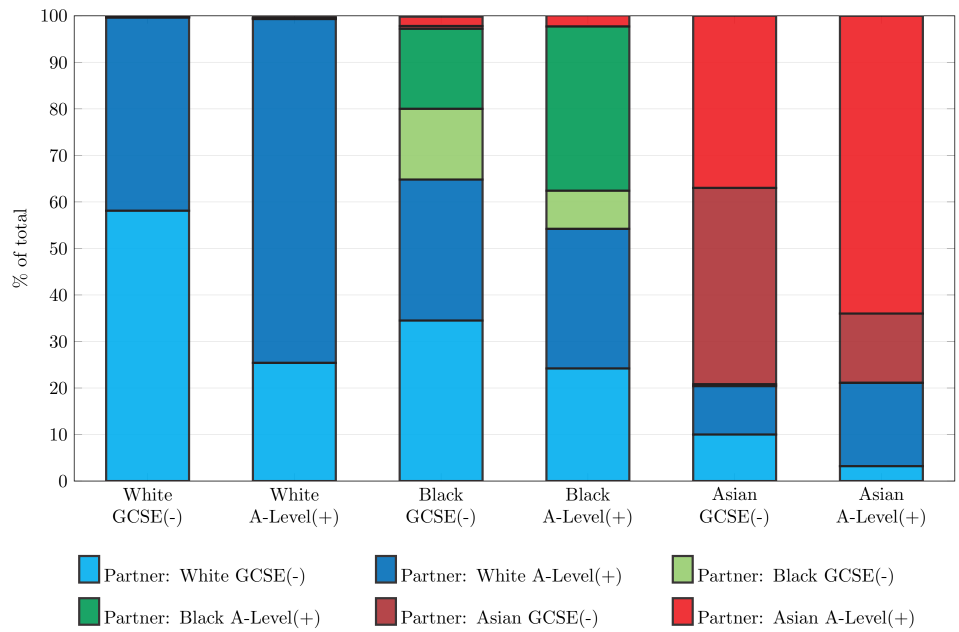
(b) Females



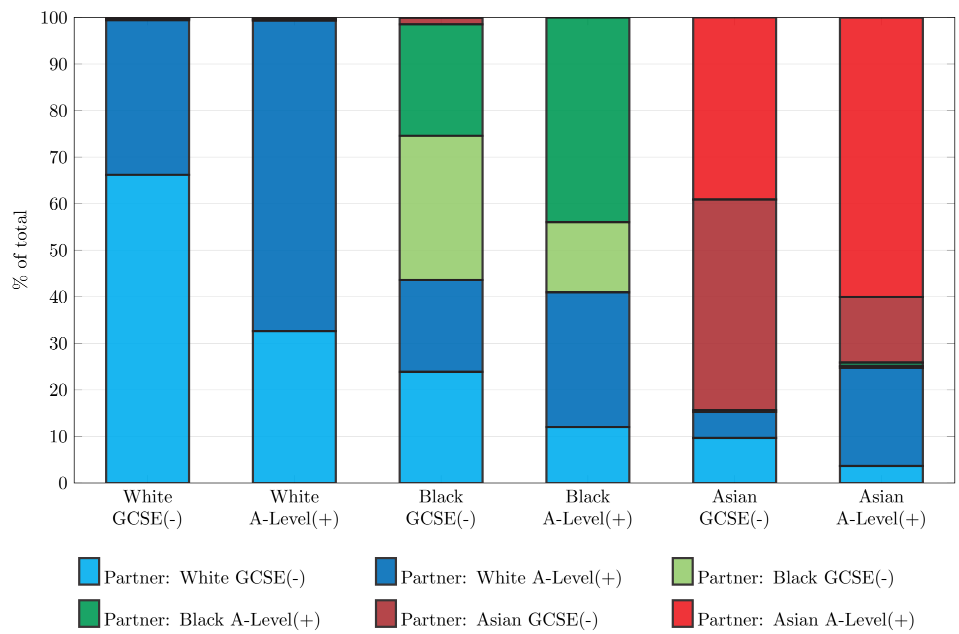
Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born between 1965-1989, living in England, Scotland or Wales, and aged 25 or above when observed, and with available information on gender, ethnicity, educational attainment, marital status.

Figure 6: Distribution of Partner Type by Own Type and Gender

(a) Males

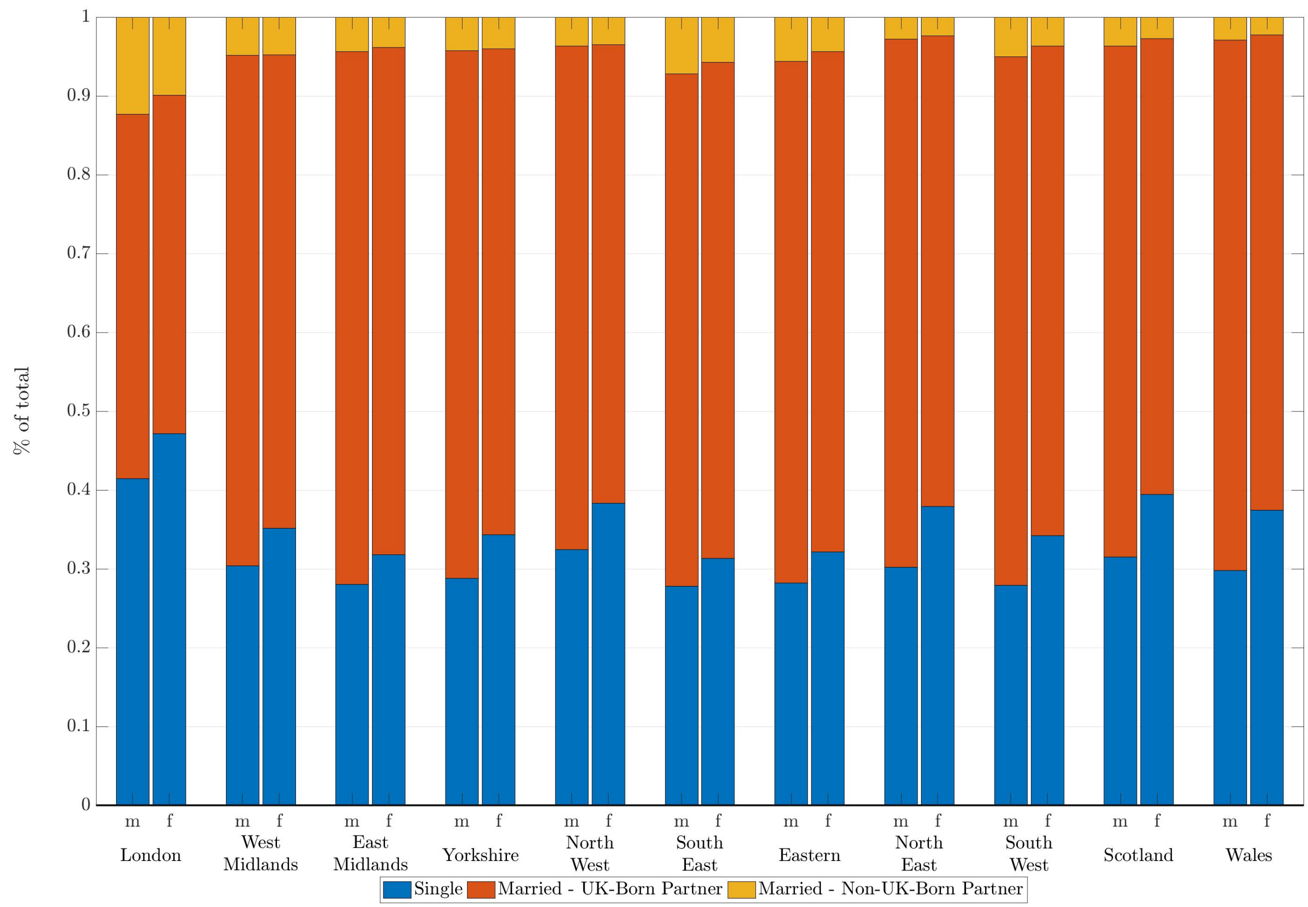


(b) Females



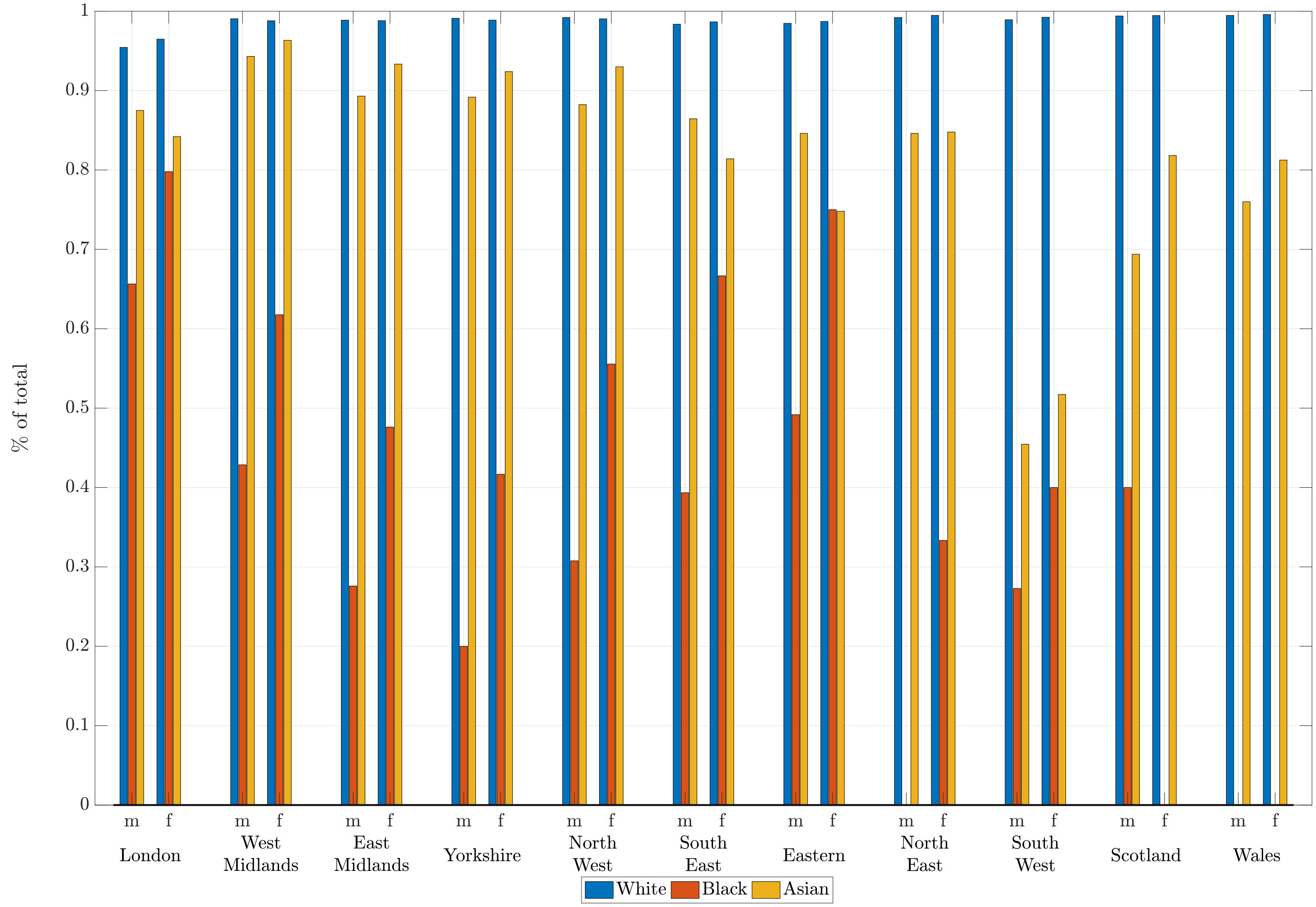
Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born between 1965-1989, living in England, Scotland or Wales, aged 25 or above, married to a UK-born partner at the time of the survey, and with available information on gender, ethnicity, and educational attainment.

Figure 7: Distribution of Marital Status by Region and Gender



Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born between 1965-1989, living in England, Scotland or Wales, and aged 25 or above when observed, and with available information on gender, ethnicity, educational attainment, marital status.

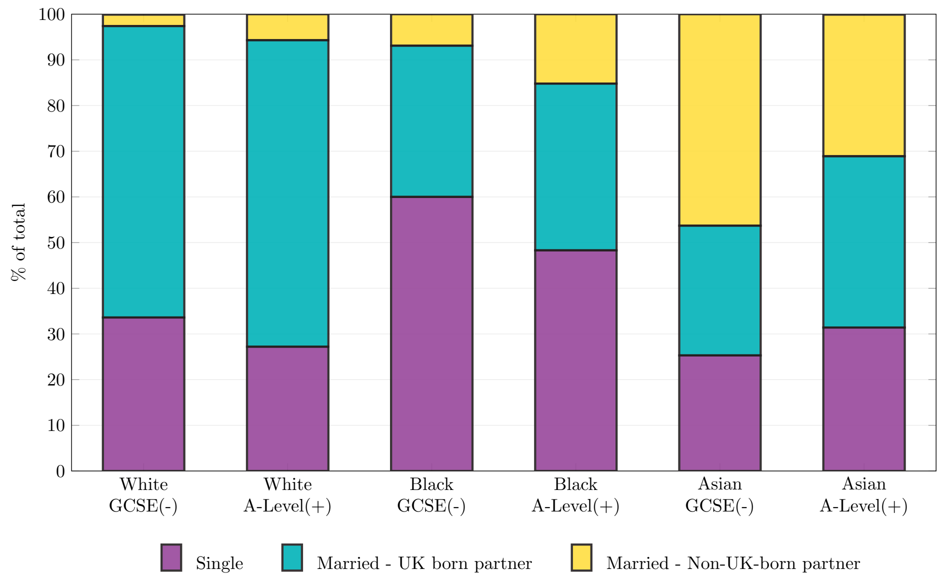
Figure 8: Proportion Intra-Ethnic Marriages by Region and Gender



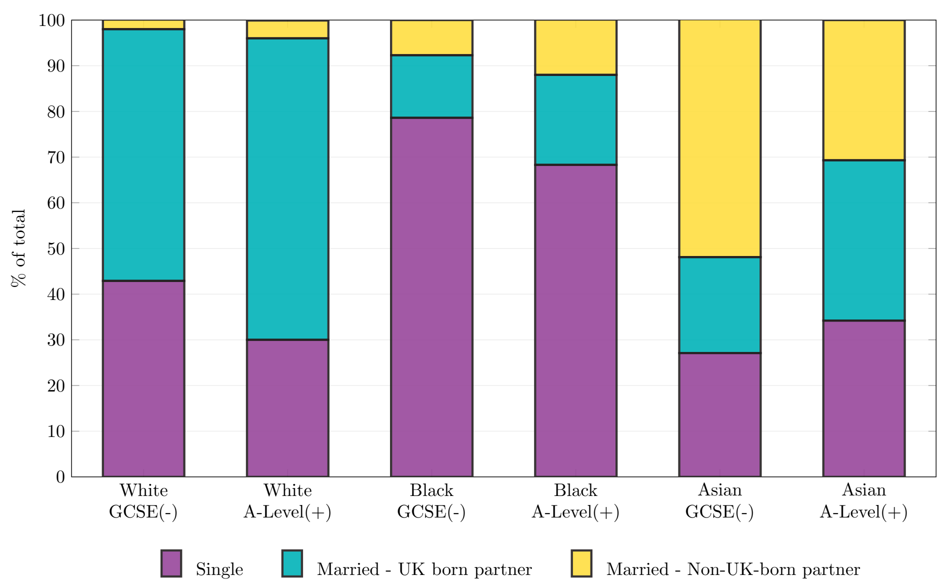
Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born between 1965-1989, living in England, Scotland or Wales, aged 25 or above, married to a UK-born partner at the time of the survey, and with available information on gender, ethnicity, and educational attainment.

Figure 9: Predicted Distribution of Marital Status by Own Type and Gender

(a) Males

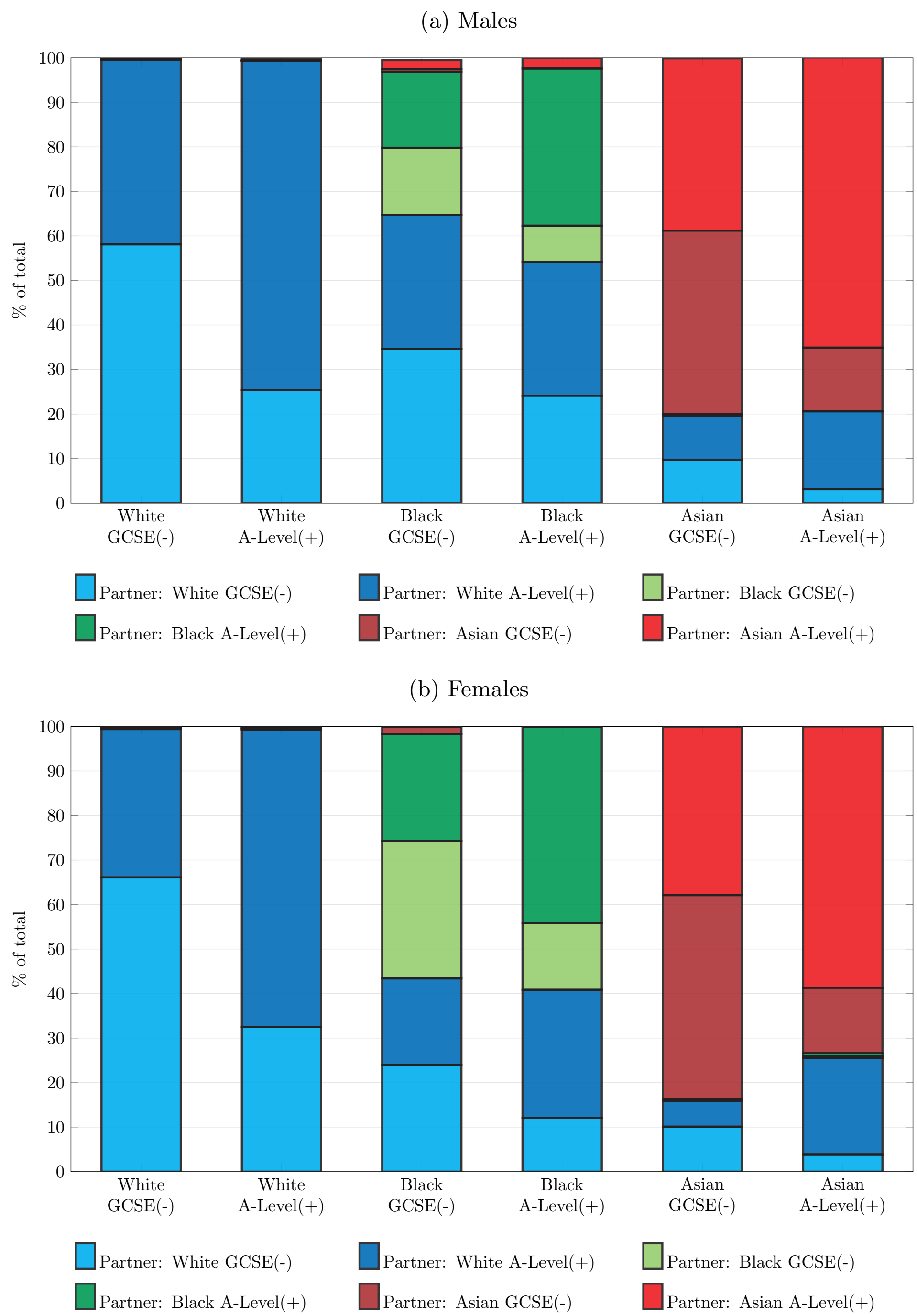


(b) Females



Notes: Predictions based on the preferred model specification (iii) in Table 4.

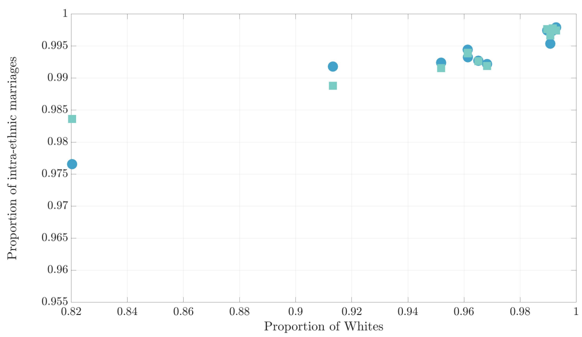
Figure 10: Predicted Distribution of Partner Type by Own Type and Gender



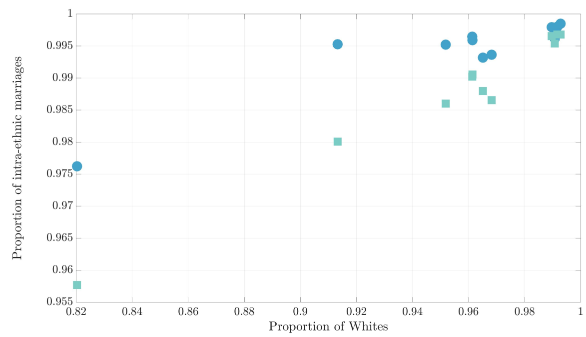
Notes: Predictions based on the preferred model specification (iii) in Table 4.

Figure 11: Observed and Predicted Shares of Intra-Ethnic Marriages across Regions by Ethnicity and Gender

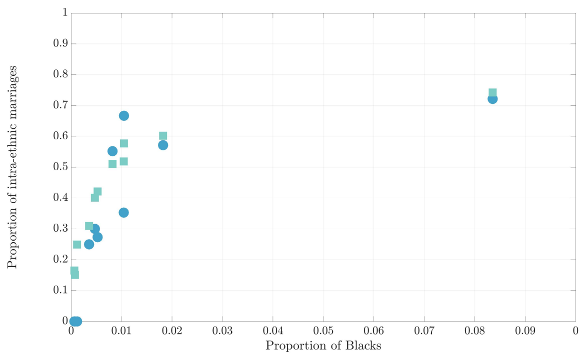
(a) White Males



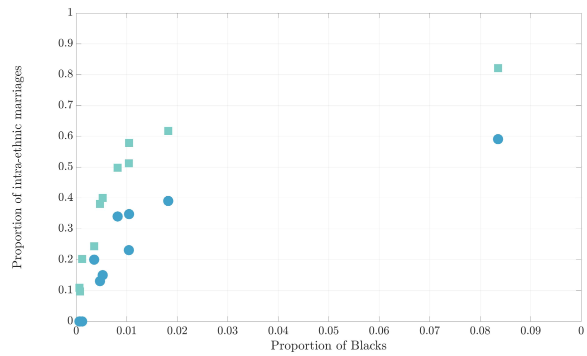
(b) White Females



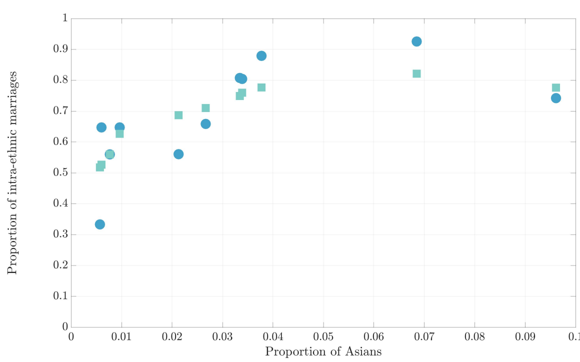
(c) Black Males



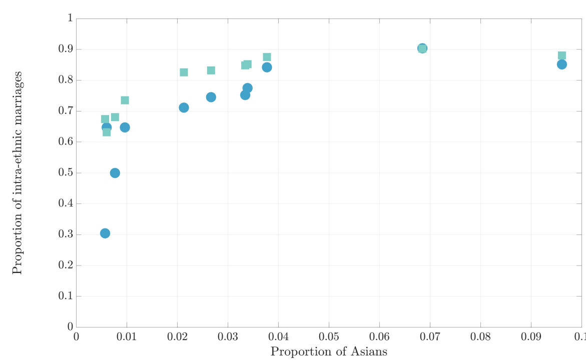
(d) Black Females



(e) Asian Males

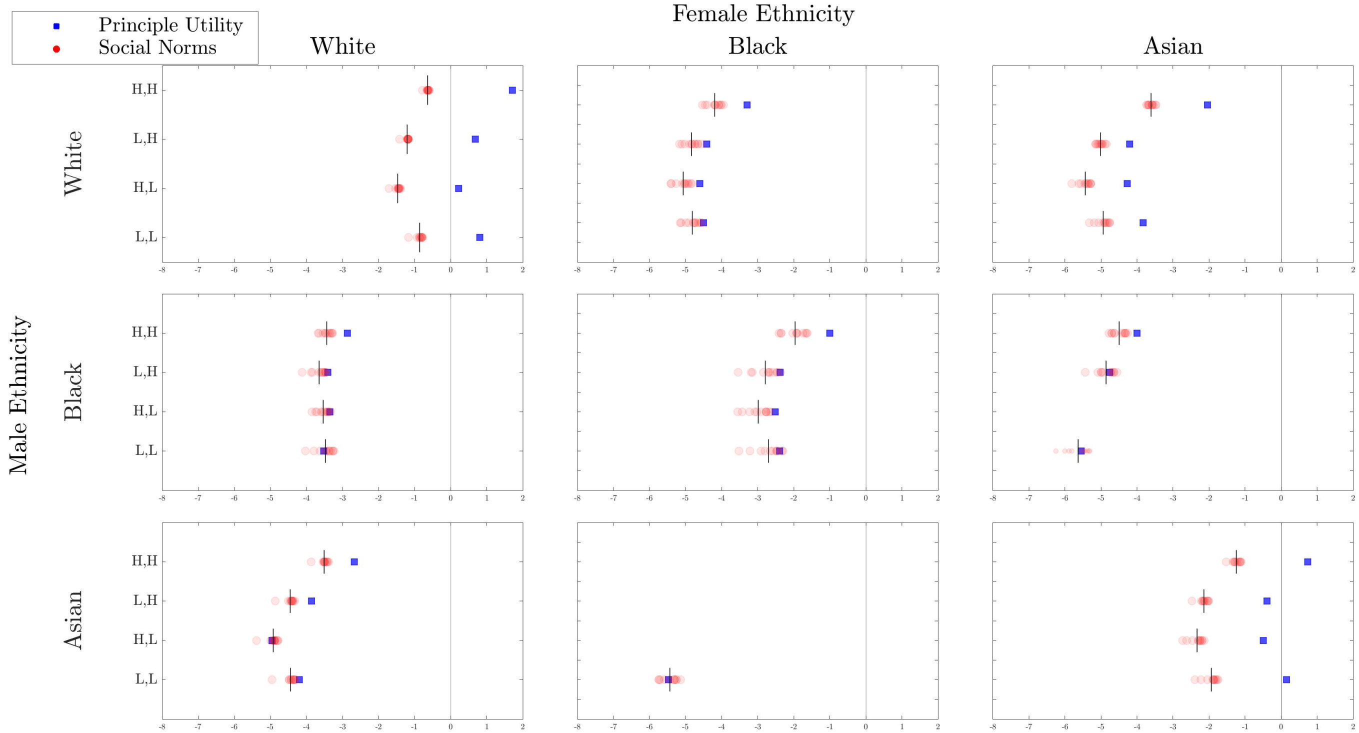


(f) Asian Females



Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born between 1965-1989, living in England, Scotland or Wales, aged 25 or above, married to a UK-born partner at the time of the survey, and with available information on gender, ethnicity, and educational attainment. Predictions based on the preferred model specification (iii) in Table 4.

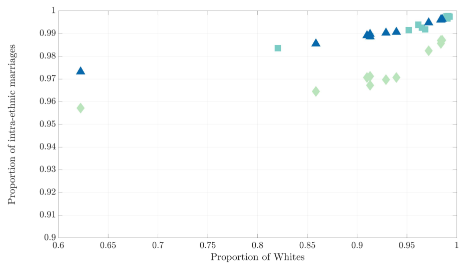
Figure 12: Principal Marriage Utility and Distribution of Social Norms by Marriage Type-Profile



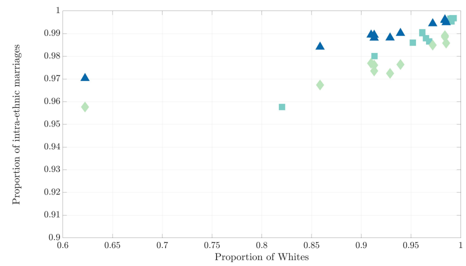
Notes: The "principal utility" depicts the estimated $\zeta(x_i, x_j)$ for each partner type profile $(x_i, x_j) \in X \times X$ while the "social norms" depicts the estimated distribution of $\phi_1^m \log\left(\omega_{x_j|x_i}^{g,m}\right) + \phi_1^f \log\left(\omega_{x_i|x_j}^{g,f}\right)$ across regions $g \in G$. The estimates are based on the preferred specification (iii) in Table 4.

Figure 13: Predicted Shares of Intra-Ethnic Marriages across Regions by Ethnicity and Gender in Recent Cohorts Compared to Estimating Cohorts

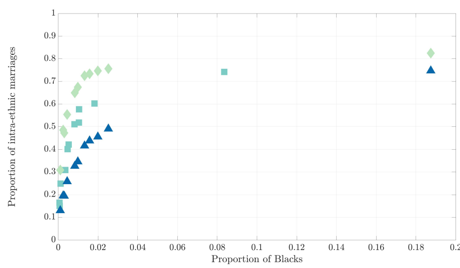
(a) White Males



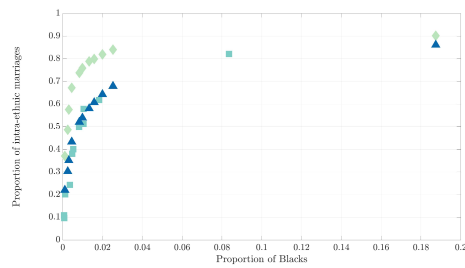
(b) White Females



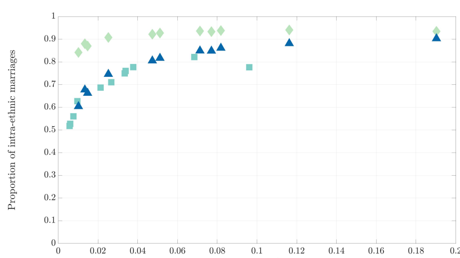
(c) Black Males



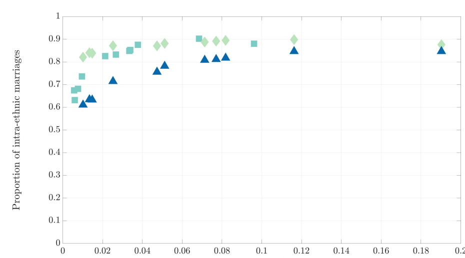
(d) Black Females



(e) Asian Males



(f) Asian Females



Notes: All predictions are based on parameter estimates from the preferred specification (iii) in Table 4. Predictions for the “estimating cohorts born 1965-1989 correspond to Figure 11. Predictions for the “recent cohorts are based on the demographic distribution in cohorts born 1990-2006 described in Appendix B.

Appendix B: The 1990-2006 Cohorts (Not for Publication)

In this appendix we provide details of the construction the type distribution in the "recent cohorts" by region. To this end we use the QLFS 1996-2015 and keep all individual born in the UK between 1990-2006. As ethnicity is directly observable, we can directly characterize the ethnic distribution by gender and region for these cohorts. However, as many of the individuals included in this sample had not completed their education by the time they were observed, we need to impute the proportion with a high qualification by gender, region, and ethicity in these recent cohorts.

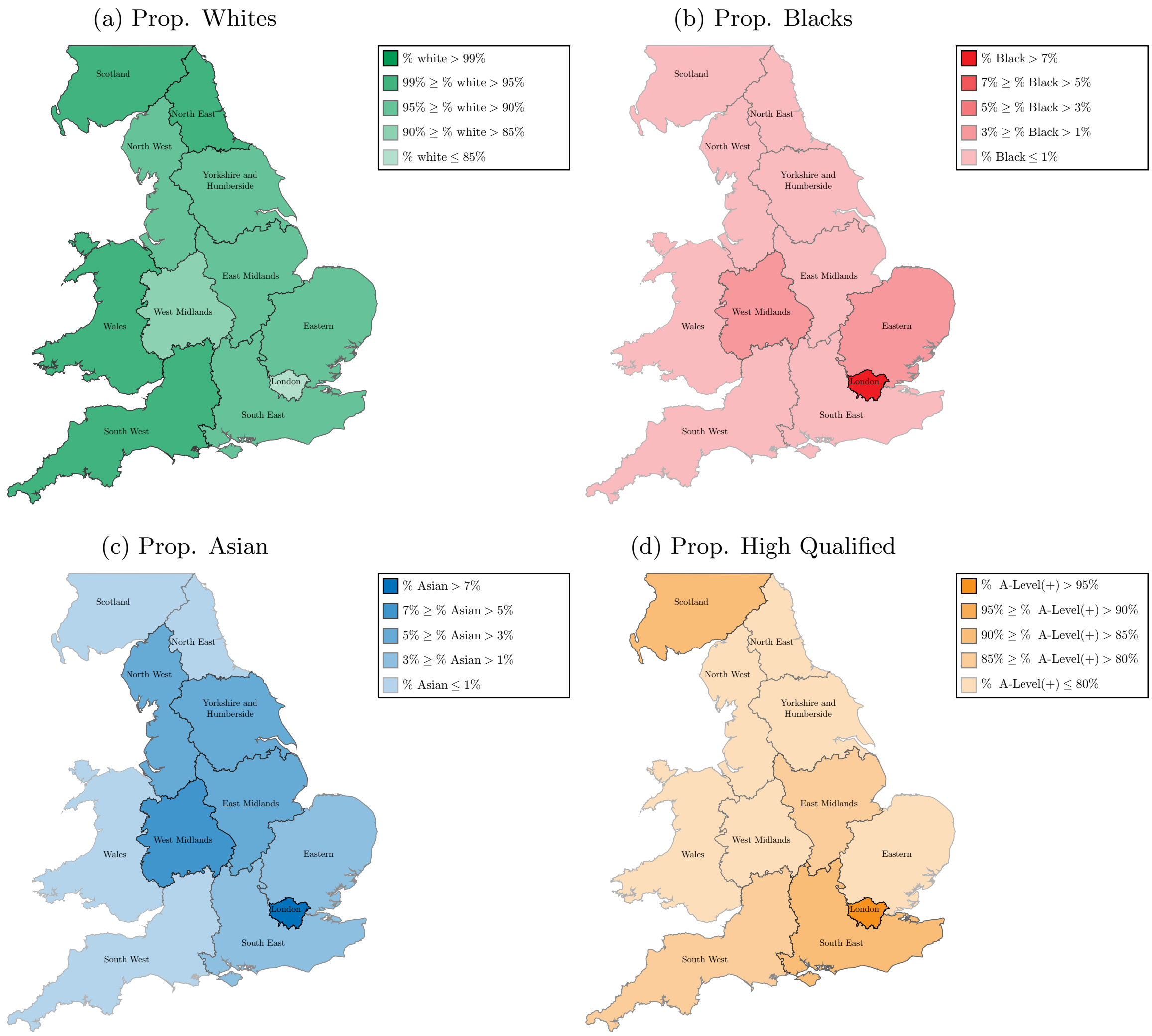
To do so, we estimate a linear probability model for holding a high qualification (A-level+) using the estimation cohort sample observed in the QLFS 1996-2015, born in the UK and aged 25 or higher when observed. The estimated specification models the educational attainment (dummy for having an A-level+ qualification) of individual i, of gender $k = m, f$, living in region $g \in G$, and of ethnicity $e \in \{W, B, A\}$ and birth cohort c_i, as

$$q_{ikge} = \alpha + \beta_k + \gamma_g + \delta_e + \kappa_{kge} c_i + \varepsilon_{ikgec}. \tag{B1}$$

The model thus includes gender-, region- and ethnicity fixed-effects, and models a linear growth in rate of holding an A-level+ qualification that is also gender-, region, and ethnicity-specific. Based on the estimated model we then impute an average qualification rate by gender, region and ethnicity for the recent cohorts using the distribution of c_i within the cell.

Figure B.1 (panels a-c) shows the empirical ethnic distribution across regions in the recent cohorts. Panel d shows the imputed rate of holding a high qualification by region. Note that the scale in panel d is not the same as in Figure 4 reflecting the substantial increase in the high qualification rate between the estimating- and recent cohorts.

Figure B.1: Demographic Composition by Region Cohorts 1990-2006



Notes: The sample consists of all UK-born individuals observed in the QLFS 1996-2015, born between 1990 and 2006 cohorts, living in England, Scotland or Wales, and with available information on ethnicity.