

# *Bend It Like Beckham: Ethnic Identity and Integration*\*

Alberto Bisin<sup>†</sup>      Eleonora Patacchini<sup>‡</sup>      Thierry Verdier<sup>§</sup>      Yves Zenou<sup>¶</sup>

September, 2009  
PRELIMINARY DRAFT

## Abstract

We propose a theoretical framework to study the determinants of ethnic and religious identity formation. We distinguish formally between two distinct motivational processes for identity formation which have been proposed in the social sciences: *cultural conformity* and *cultural distinction*. Data on ethnic preferences and attitudes provided by the Fourth National Survey of Ethnic Minorities in the UK enables us to test the relative preponderance of these two motivating processes.

We find evidence consistent with intense ethnic and religious identity to be formed mostly as a *cultural distinction* mechanism. Consistently, we document that identity and socialization to an ethnic or religious minority are, other things equal, more intense in mixed neighborhood than in segregated neighborhoods.

**Key words:** Ethnicity, identity, intermarriage, cultural transmission

**JEL Classification:** A14, J15

---

\*This paper is part of the Polarization and Conflict Project CIT-2-CT-2004-506084 funded by the European Commission-DG Research Sixth Framework Programme. A different paper by the same authors circulated under the same title. We liked the title better than the paper; so we scrapped the second and kept the first. Thanks to Alberto Alesina, Bill Easterly and other participants to the Political Economy Summer Meetings of the NBER.

<sup>†</sup>New York University, Department of Economics, 269 Mercer Street, New York, NY 10003, USA. E-mail: alberto.bisin@nyu.edu

<sup>‡</sup>Universita' di Roma "La Sapienza," Facolta' di Scienze Statistiche, P.le Aldo Moro, 5 - 00185, Roma, Italy. E-mail: eleonora.patacchini@uniroma1.it

<sup>§</sup>PSE, 48 boulevard Jourdan, 75014 Paris, France and CEPR. E-mail: verdier@pse.ens.fr

<sup>¶</sup>IUI, The Research Institute of Industrial Economics, P.O. Box 55665, 102 15 Stockholm, Sweden, GAINS and CEPR. E-mail: yvesz@iui.se

*Bengali, bengali / Bengali, bengali / No no no / He does not want to depress you/ Oh no no no no / He only wants to impress you / Oh.. Bengali in platforms / He only wants to embrace your culture / And to be your friend forever.* [‘Bengali in Platform,’ Morissey, *Viva Hate*, 1988, Reprise/Wea]<sup>1</sup>

## 1 Introduction

In the last decades, immigration to advanced countries has become an important facet of globalization. While the direct economic effects of such trends have been debated for long by economists, there has been increased concerns with the implications of non-economic issues such as increased ethnic diversity in the countries receiving these migrants’communities<sup>2</sup>. In such a context, the issue of ethnic and religious identity has been at the forefront of the political debate in Europe as well as in the U.S. (see Alba (1990), (2005)). While diversity of social groups is considered as a source of benefits by multiculturalists, the persistence of migrants or minorities’ identities is however often perceived as a threat or source of frictions by natives and dominant groups. This is well illustrated in France by the recent intense debate around the wearing of the islamic burqa. In a presidential address to parliament on June 22 2009, President Sarkozy said that *‘It will not be welcome on the territory of the French Republic’* and supported banning the garment from being worn in public. Doing that, he faced intense critics who fear the burqa issue could stigmatize France’s Muslim population of about 4 millions. As Robert D. Putnam (2007) in his John Skytte Price Lecture mentions, this type of phenomenon emphasizes that *“One of the most important challenges facing modern societies, and at the same time one of our most significant opportunities, is the increase in ethnic and social heterogeneity in virtually all advanced countries”*.

It is then important to develop a better understanding of the mechanism of identity formation and of its determinants in particular. In fact, two opposing views characterize the theoretical analysis in the social sciences regarding identity formation;<sup>3</sup> see e.g., Putnam (2007). A first group of social scientists argues that ethnic identity is reduced by assimilation and blurring of groups’ boundaries. *Assimilation theories* in political science and sociology (Gordon (1964), Moghaddam and Solliday 1991), *contact theory* in social psychology (Allport 1954) are e.g., direct expressions of this line of thought. The basic premise is the idea that social contacts between groups and intensive social interactions help weaken group loyalties and group prejudices and hence lead to a more culturally homogeneous society. Through this process, minority groups adopt inclusive identities and integrate, progressively adopting the language, values and systems of the dominant group. Underlying this reasoning is the basic principle that group identity is driven by a motive for inclusiveness and *cultural conformity*. Economists capture the notion of *cultural conformity* through positive social interactions

---

<sup>1</sup>Thanks to Andrew Clark for Morissey’s quote.

<sup>2</sup>Alesina and La Ferrara (2005) provides a general discussion of the economic effects of increased ethnic diversity.

<sup>3</sup>The study of ethnic identity formation has a long theoretical and empirical tradition in social sciences: see e.g., Cross (1991), Phinney (1990), Ferdman (1995) in developmental psychology; Stryker (1980) in symbolic interactionist sociology; Tajfel (1981), Tajfel and Turner (1979), Turner et al. (1987) in social psychology; and Brewer (2001) in political psychology.

across individuals sharing the same characteristics, views and preferences<sup>4</sup>. Individuals therefore form inclusive identities to integrate so as to exploit social interactions.

The alternative view considers individuals motivated to retain their own distinctive cultural heritage and to identify with a ethnic/social group to enhance their psychological self-esteem by generating a sense of positive distinctiveness from individuals who are not part of the group (Abrams and Hogg 1988, Turner 1982). This sense of positive distinctiveness can be achieved through various cognitive and psychological mechanisms, from group solidarity to prejudice and negative stereotyping with respect to other groups. Negative attitudes towards members of other groups in turn reassert social identity of the group. Along these lines are expressed theories of *multiculturalism* (Glazer and Moynihan (1970), Taylor and Lambert (1996)), and *conflict* (Bobo (1999)).<sup>5</sup> Group identity formation, according to this view, is a sort of *cultural distinction* mechanism that allows individuals to reduce the psychological costs associated to cultural differences. In economic terms, the concept of *cultural distinction* can be motivated in terms of negative social interactions across individuals belonging to different identified groups.

To provide a conceptual frame to the issue of the determinants of identity formation, in this paper, we formulate a simple model of identity formation which accounts for both *cultural conformity* and *cultural distinction*. We can therefore formally establish the respective implications of these theoretical views of identity formation to guide an empirical investigation of, in particular, ethnic identity. In doing so, we show that distinguishing between *cultural conformity* and *cultural distinction* provides contrasting empirical implications on the way neighborhood segregation and identity formation interact in the process of ethnic integration. When *cultural conformity* is the main motivational process of identity formation, we show that neighborhood segregation and identity formation are likely to be complements for ethnic assimilation. On the contrary, when *cultural distinction* is the main process of identity formation, neighborhood segregation and identity formation tend to be substitutes for ethnic assimilation.<sup>6</sup>

*Cultural conformity* and *cultural distinction* are quite distinct theories of identity formation and have important different positive and normative implications for integration policies. Empirical evidence for *cultural distinction* would for instance suggest that intense and oppositional identities that give rise to ethnic conflicts might not necessarily be the result of the segregation of the neighborhood in which ethnic and racial minorities tend to live: policies favoring neighborhood integration would then not necessarily favor cultural integration, contrary to presumptions often exposed by social scientists and commentators.

---

<sup>4</sup>See Bernheim (1994) for a formal economic analysis of conformity.

<sup>5</sup>At a broader level this view is also related to *social identity theory* in social psychology (Tajfel 1981, Turner 1982).

<sup>6</sup>In economics, the distinction between *cultural conformity* and *cultural defense* is also related to the notion of cultural complementarity and cultural substitutability between socialization mechanism defined formally by Bisin-Verdier (2000); see also Patacchini-Zenou (2004) for an empirical investigation. Indeed, in Bisin-Verdier (2000), when family and society role models tend to be substitutes in the process of socialization, families with a relatively minoritarian cultural trait have larger incentives to spend resources to socialize their children to the trait and hence guarantee its persistence. Conversely, under cultural complementarity the more minoritarian is a family's cultural trait, the smaller are the family's incentives to socialize their children to the trait and hence to limit cultural assimilation.

The distinctive implications of *cultural conformity* and *distinction* can be tested empirically and can therefore inform on the main process of ethnic identity formation. To address more directly the issue of what motivates identity formation, in this paper we study ethnic and religious identity formation as a social phenomenon at the level of the neighborhood. In particular we consider the demographic characteristics of the neighborhood where agents reside, and most importantly its ethnic and religious composition, as underlying the identity formation mechanism. Furthermore, we link identity formation with homogamous marriages along ethnic and religious traits. It has been extensively documented in fact that interracial marriage is typically considered as a sign of inclination toward cultural assimilation (see, in particular, Al-Johar, (2005), Qian (1999), Meng and Gregory (2005), Lichter et al. (2007), Tucker and Mitchell-Kernan (1990)) and that marriage choices are at least in part determined by parents' preferences to socialize their children to their own (parents') trait (see Bisin, Topa, and Verdier (2004), and the evidence cited in Bisin and Verdier, (2000)).

The scarcity of empirical work examining the importance of ethnic preferences on individual behavior is partly due to the limited information and sample sizes on cultural variables. Our analysis exploits a unique UK data set, the Fourth National Survey of Ethnic Minorities (FNSEM), which over-samples ethnic minority groups and explicitly acknowledges the heterogeneity within the non-white population in terms of individual, demographic, family and socio-economic characteristics. Most importantly for our purposes, this survey asks a direct question about respondents' identification with their own ethnic group and additional (indirect) information about different dimensions of identity.(e.g. attitudes towards inter-marriage, importance of religion and other aspects of individual's ethnic preferences). In addition, the data can be merged with the 1991 Census, so that it is possible to obtain a detailed picture of each individual's residential neighborhood at a very high level of spatial disaggregation.

With FNSEM data we estimate our model of the joint determinants of ethnic and religious identity and homogamy both structurally and non-structurally. The model nests *cultural distinction* and *cultural conformity* as identity formation mechanisms. Our evidence is consistent with ethnic identity to be formed as a *cultural distinction* mechanism rather than due to *cultural conformity*. Ethnic identity appears to be formed in social contexts in which the minority ethnic trait is mostly "threatened" either directly by the actions of the majority group (e.g., through explicit acts of rejection or harassment), or indirectly simply by being exposed to the interaction with the majority norm of behavior in mixed neighborhood.

Our evidence for *cultural distinction* is consistent with two recent empirical studies addressing indirectly the issue of the process of identity formation by studying the link between identity and segregation. Notably, using a nationally representative sample of more than 90,000 students, from 175 schools, who entered grades 7 through 12 in 1994 in the US (the National Longitudinal Study of Adolescent Health), Fryer and Torelli (2005) find that "acting white" behaviors among blacks (i.e. the higher the test score the less popular a student is) are more developed in racially mixed schools.<sup>7</sup>

---

<sup>7</sup>Anthropologists have also observed that social groups seek to preserve their identity, an activity that accelerates when threats to internal cohesion intensify. Thus, groups may try to reinforce their identity by penalizing members for differentiating themselves from the group. The penalties are likely to increase whenever the threats to group cohesion intensify; for an early analysis of this issues, see Whyte (1943).

Also, Bisin, Topa, and Verdier (2004) document that religious socialization across U.S. states is more intense when a religious faith is a minority.<sup>8</sup>

We also address the issue of the alleged specificity of Muslim immigrants with regards to the strength of their identity and their (lack of) integration; an issue which recently surged at the center of the political debate in Europe (see, e.g., Gallis, 2005). We repeat our analysis on the restricted sample of Muslim respondents. The results are not qualitatively different from the ones found using the whole sample. This evidence suggests that the relationship between ethnic integration effort and ethnic neighborhood composition is not significantly different for Muslims with respect to other minorities. Evidence of slower integration for Muslims, both first and second generation, is however apparent in our analysis.

Beyond the large sociology and socio-psychology literature on ethnic identity formation, our work is related to a growing economic literature studying the evolution of culture and ethnic identity and its interactions with economic outcomes. Akerlof and Kranton (2000) consider identity formation as an explicit - more or less conscious - endogenous choice by individuals placed in a certain social context. Darity, Mason, and Stewart (2006) and Eaton, Eswaran and Oxoby (2009) provide evolutionary models discussing the relationship between identity formation and inter-racial interactions. In the specific context of African American communities of the Ante-Bellum American South, Bodenhorn and Ruebeck (2003) also underlines the endogeneity of racial identity. Chiswick (2006) emphasizes the role of ethnic specific human capital in minority groups' decisions to culturally assimilate or separate. Austen-Smith and Fryer (2005); Battu, Mwale and Zenou (2007), Bisin, Patacchini, Verdier and Zenou (2008) and Fang and Loury (2005) discuss the emergence persistence of "Oppositional" or "Dysfunctional" identities in marginalized social groups. In the context of migrant communities in Germany, Constant and Zimmermann (2008) and Constant, Gataullina, and Zimmermann (2009) analyse ethnic identity as the endogenous balance between commitment to and self-identification with the culture and society of the origin and the host country. Like us, this emerging literature recognizes the endogeneity and contextual character of ethnic and cultural identities formation. Our contribution however is to go further into opening the "black box" of identity formation, and to disentangle and identify in a specific empirical context, precise mechanisms of identity formation (ie. *cultural distinction* versus *cultural conformity*). which have distinct implications for cultural integration processes across social groups.

Before proceeding, we should briefly alert the reader to the methodological attitude of the paper. We proceed in steps, from a non-structural analysis of the data to a fully structured model estimation. More precisely, we start in section 2 with a non structural probit analysis of identity and homogamy in terms of ethnic composition (and different sets of controls). Section 3 then proposes a semi-structural analysis of homogamy and identity which, while disregarding some cross-equation restrictions as imposed in a fully structural approach, has power to flexibly and efficiently address the issue of *cultural distinction* vs. *cultural conformity*. Section 4 goes further by providing and

---

<sup>8</sup>Relatedly, Bisin and Verdier (2000) provide many examples of the resilience of ethnic and other cultural traits that can be explained by a similar mechanism, from the case of Orthodox Jews in Brooklyn to the case of aristocrats in France.

estimating a fully structural model of ethnic integration 4. Finally section 5 concludes.

## 2 Descriptive analysis of the data

Our analysis exploits the Fourth National Survey of Ethnic Minorities (FNSEM), which was collected in 1993/94 in the U.K. by the Policy Studies Institute (PSI). FNSEM over-samples ethnic minority groups and explicitly acknowledges the heterogeneity within the non-white population where the ethnic population is composed of six groups (Caribbean, Indian, Pakistani, African-Asian, Bangladeshi, and Chinese).<sup>9</sup> It also contains detailed information about respondents' identification with their own ethnic group (e.g. attitudes towards inter-marriage, importance of religion and other aspects of individual's ethnic preferences) as well as variables aiming at capturing the heterogeneity within the non-white population in terms of individual, demographic, family and socio-economic characteristics (see Modood et al., 1997, for details).

We enrich the analysis of ethnic identification, necessarily a self-reported subjective measure, with the study of marriage homogamy along ethnic lines. Homogamy, while more precisely measured, can in fact be considered a proxy measure of identity to which is conceptually strongly related.

Finally, to address the main issue of this paper, the identification of *cultural distinction* versus *cultural conformity*, we need to study the variation of the respondents' identification with their own ethnic group across different residential neighborhoods as characterized by their ethnic composition. In fact, *cultural conformity* would induce a respondents' identification with their own ethnic group which grows stronger the higher the proportion of the group in the neighborhood. *Cultural distinction* would instead introduce an opposing force, motivating strong ethnic identity also in neighborhoods with relatively scarce presence of the ethnic group, as a form of distinction/defense from the majority. We then merge the FNSEM data with the 1991 Census in order to get valuable information of each individual's residential ward.<sup>10</sup>

### 2.1 Definition of the variables

The key variables in our analysis are *(i)* the ethnic composition of the residential neighborhood,  $q$ ; *(ii)* the intensity of the ethnic identity,  $\nu$ ; *(iii)*, the probability of homogamous marriage,  $\pi$ . They are described in turn.

*(i)* The ethnic composition of the neighborhood is observed at the level of the residential ward from the 1991 Census data. For each individual  $i$  we consider the percentage of ward inhabitants of her/his own ethnic group. It has been divided in seven classes,  $q_i \leq 2\%$ ,  $2\% < q_i \leq 5\%$ ,  $5\% < q_i \leq 10\%$ ,  $10\% < q_i \leq 15\%$ ,  $15\% < q_i \leq 25\%$ ,  $25\% < q_i \leq 33\%$ ,  $q_i \geq 33\%$ . As usual, the mean value of each interval is used in the regression analysis. Figure 1 reports the distribution of respondents over the ethnic composition of the neighborhood in which they live, ethnic group by ethnic group. It should be clear that variation in  $q$  does not proxy for ethnic group; that is, it is not the case e.g., that

---

<sup>9</sup>Black Africans were not included because the bulk of their immigration in the U.K. happened earlier. Furthermore, the survey only covers England and Wales.

<sup>10</sup>A UK Census ward contains on average 3,000-4,000 residents.

the respondents of distinct ethnic groups predominantly live in neighborhoods with specific ethnic composition.

[Insert Figure 1 here]

(ii) The survey contains a number of questions providing information on different dimensions of identity, in particular importance of religion, attitudes towards inter-marriage and the relevance of ethnicity in influencing the kind of school that people want for their children.<sup>11</sup> It also asks a direct question about ethnic identity.<sup>12</sup> We performed our analysis using separately the answers on each of these questions in turn. Identity, denoted by  $I$ , is coded as a dichotomous variable taking value 1 if the individual considers very important the role of religion in her/his life and 0 otherwise (*importance of religion*). It takes value 1 if the individual would personally mind if a close relative were to marry a white person and 0 otherwise (*inter-ethnic marriage*); it has value 1 if ethnicity has a very important or at least fairly important influence in choosing the school for a child and 0 otherwise (*school ethnic composition*); and finally it takes value 1 if the individual strongly agrees or agrees to the statement “In many ways, I think of myself as [respondent’s ethnic group]” and 0 otherwise (*ethnic group identification*). The variable measuring the intensity of ethnic identity,  $\nu$ , is then the probability that  $I = 1$ .

(iv) Homogamy  $H$  is a dummy variable taking value 1 if the respondent is married with a person of her/his own ethnic group and 0 otherwise. The variable  $\pi$  measures the probability that marriage is homogamous. Singles, somewhat consistently with the theoretical analysis of integration in Section 4, are assigned  $H = 0$ , that is, they are treated as non-homogamous.

An extensive set of control variables is also available. In addition to several individuals’ observable characteristics (i.e., education, age, sex, fertility choices, employment status, job qualification, household house ownership, macro-region of residence, time spent in the UK, a dummy indicating whether the respondent is born in the UK), the data set contains also control vector variables aiming at capturing the influence of the social environment (family, friends, neighbors) and workplace (language typically spoken in the family, with friends, at work, a dummy capturing instances of discrimination, and one indicating whether the marriage is arranged by the parents, the ward unemployment rate). Precise definitions of all these variables, as well as our sample descriptive statistics, can be found in the Data Appendix, Table A1. Excluding the individuals with missing or inadequate information on our target variables, we obtain a final sample of 1,565 individuals.

---

<sup>11</sup>The precise questions are the following ones: “How important is religion to the way you live your life? Is it not at all important, not very important, fairly important or very important?”; “Would you personally mind if a close relative were to marry a white person?; “If you were choosing a school for an eleven-year old child of yours, would your choice be influenced by how many (respondent’s ethnic group origin) children there were in the school? And, if “yes”, would it be a very important influence, a fairly important influence or a not very important influence?” .

<sup>12</sup>Specifically, in the FNSEM, the people interviewed are asked if they strongly agree, agree, neither agree or disagree, disagree, strongly disagree, with the statement: “In many ways, I think of myself as [respondent’s ethnic group]” .

## 2.2 Identity and homogamy: A probit

Empirically identifying cultural distinction from cultural transmission would require observing a negative relationship between identity (resp. homogamy) and the ethnic composition of the neighborhood as measured by  $q$ . Of course regression analysis is difficult to interpret because of the endogeneity of  $q$ : the ethnic composition of then neighborhood might in principle be an important factor affecting the respondents' residential decisions. In the bulk of the paper we shall deal with this endogeneity directly and indirectly, as well as by imposing more structure on the empirical analysis. It is nonetheless interesting to present the results of simple probit regressions, with identity  $I$  and homogamy  $H$  as dependent variables, allowing for non-linear dependence from ethnic composition  $q$ . We perform such a bivariate probit analysis with different (increasing) sets of controls variables and using alternative definitions of ethnic identity. "Importance of religion" is our measure of ethnic identity with the highest coverage. Table 1 shows the complete list of estimation results when "importance of religion" is used as a proxy for ethnic identity, whereas Table 2 collects the results on our target variables ( $q$  and  $q$  squared) when the estimation is run using the other measures of ethnic identity as alternative dependent variables.<sup>13</sup>

*[Insert Tables 1 and 2 here]*

The non-linearities, i.e. the quadratic terms, are clearly significative in almost all the regressions and document, for both identity and homogamy, a negative dependence from ethnic composition for values of  $q$  greater than 20%. Indeed, Figure 2a depicts the estimated (non-linear) effect of  $q$  on identity and homogamy when the influence of our most extensive set of controls has been purged out.<sup>14</sup>

*[Insert Figure 2a and 2b here]*

The non-linearities picked up by the probit models might in principle be due to the differential distribution of ethnic groups by neighborhood class.<sup>15</sup> This is not the case, in fact, as the same form of non-linear dependence in  $q$  results when the sample is restricted to Muslims (mostly Pakistanis, Bangladeshis, and Indians), as documented in Figure 2b.<sup>16</sup>

---

<sup>13</sup>In the probits we also experiment with a different measure of the ethnic composition  $q$ : the fraction of all-minority residents in the neighborhood. We are encouraged by the results to stick to measuring  $q$  as the fraction of the own ethnic group in the neighborhood. The fraction of all-minority residents in the neighborhood is however included among the controls.

<sup>14</sup>This graph is depicted using "importance of religion2 as a proxy for ethnic identity. The use of the other proxies leads to similar pictures.

<sup>15</sup>We thank Bill Easterly for raising this point in a critical discussion of the paper.

<sup>16</sup>We have also performed our analysis using a multidimensional measure of ethnic identity which summarizes the information contained in the available indicators. We have followed the standard approach in the sociological literature to derive quantitative information on sensitive topics using qualitative answers to a battery of related questions. This is a standard factor analysis, where the factor loadings of the different variables (questions) are used to derive the total score (multidimensional measure). The Cronbach- $\alpha$  measure is then used to assess the quality of the derived index. In our case, we obtain an  $\alpha$  equal to 0.86 ( $0 \leq \alpha \leq 1$ ) indicating that the different items incorporated in the index have

The analysis in this section is essentially descriptive; more structure is necessary to be confident in interpreting Figures 2a and 2b as evidence for cultural distinction. We start introducing structure in next section.

### 3 Conformity vs. distinction: Some structure

Consider a member of a minority ethnic or religious group. Let  $q$  denote the proportion of the minority in the reference neighborhood, where the member of the minority resides. Let homogamy be an index  $H \in \{0, 1\}$  with  $prob\{H = 1\} = \pi$ . Consider a member of a minority ethnic or religious group. It is psychologically costly to interact with individuals of the dominant majority; let these psychological costs depend on the marriage status of the minority member and be denoted  $C(H)$ .

It is natural to assume that such costs are lower in an homogamous than in an heterogamous marriage,  $\Delta C = C(0) - C(1) > 0$ . Minority members therefore put effort in finding a spouse of the same ethnic and religious background. Let effort be denoted  $\tau \in [0, 1]$ . The minority member first searches a spouse in a restricted pool of partners from his own community minority. The search intensity,  $\tau$ , determines the probability with which he finds his marital partner in the pool. With the residual probability  $1 - \tau$ , he remains unsuccessful and therefore goes to a common pool of partners including both minority and majority types. There he gets matched with a spouse of his community with probability  $q$ .

A minority agent married in a neighborhood with a fraction  $q$  of minority members has then a probability of marrying homogamously equal to

$$\pi(\tau, q) = \tau + (1 - \tau)q \tag{1}$$

The search intensity  $\tau$  is chosen by the agent, but it requires a utility cost,  $Z(\tau)$ , which is increasing and convex (in the same units of the psychological costs  $C(H)$ ); for analytical simplicity we assume

$$Z(\tau) = \frac{1}{2}\alpha\tau^2 \tag{2}$$

where  $\alpha$  is a measure of the relative cost of  $\tau$ .

A minority member's problem is:

$$\max_{\tau \in [0, 1]} -\pi(\tau, q) C(1) - [1 - \pi(\tau, q)] C(0) - Z(\tau).$$

The first order condition gives:

$$(1 - q)\Delta C = \alpha\tau$$

which after substitution simply provides the solution:

$$\pi = q + (1 - q)^2 \frac{\Delta C}{\alpha}. \tag{3}$$

---

considerable internal consistency. We find that even this aggregate measure do not depend (qualitatively) differently on  $q$ .

In general,  $\Delta C = C(0) - C(1)$  will be a function of  $q$  as well as of the identity of the minority's member, an index  $I \in \{0, 1\}$  with  $\text{prob}\{I = 1\} = \nu$ .<sup>17</sup> Furthermore, both  $q$  and  $\nu$  can be chosen, under some costs, by the minority member and hence they are endogenous. Suppose that,

$$\Delta C = \Delta C(q, I).$$
<sup>18</sup>

It is then straightforward in this context to formulate a precise definition of cultural conformity and distinction. We say that the minority member's preferences display

- i) *cultural distinction* if  $\frac{\partial \Delta C(q, I)}{\partial q} < 0$ ; that is, if costs  $\Delta C$  increase in the proportion of non-minority members  $1 - q$ ; and
- ii) *cultural conformity* if  $\frac{\partial \Delta C(q, I)}{\partial q} \geq 0$ ; that is, if costs  $\Delta C$  decrease in the proportion of non-minority members  $1 - q$ .

### 3.1 Empirical implementation

Let  $x$  denote a vector of exogenous variables which affect the minority member's choice of  $(q, \nu)$ . The endogeneity of  $(q, \nu)$  can then be abstractly, and essentially without loss of generality,<sup>19</sup> represented by a map

$$\begin{pmatrix} q \\ \nu \end{pmatrix} = F(x)$$

To be able to implement empirically the model, we assume that the exogenous variables in the vector  $x$  also affect the marriage effort choice by entering through the costs of effort:

$$Z(\tau) = \frac{1}{2}\alpha\tau^2 + (\gamma_{\tau} \cdot x)\tau$$

where  $\gamma_{\tau} \cdot x = \sum_{j=1}^K \gamma_{\tau, j} x_j$ .

The solution of the minority member writes then as:

$$\pi = q + \frac{1}{\alpha}(1 - q)^2 \Delta C(q, I) - \frac{1}{\alpha}(1 - q) (\gamma_{\tau} \cdot x)$$

Its empirical implementation requires therefore fitting

$$\Delta C : [0, 1] \times \{0, 1\} \rightarrow \mathbb{R} \text{ and } F : \mathbb{R}^K \rightarrow [0, 1]^2$$

with appropriate approximating polynomials.

Two important caveats are in order, however.

---

<sup>17</sup>We might interpret  $\nu$  directly as a measure of the intensity of identity.

<sup>18</sup>As long as,  $\Delta C$  is independent of  $\tau$  and so are the costs associated to the choice of  $q$  and  $\nu$ , equation (3) still holds.

<sup>19</sup>That is, under conditions sufficient to guarantee a unique choice of  $(q, \nu)$  for any  $x$ . In fact, we do not need to assume this. We can deal with the general case, in which equilibria are multiple and the equilibrium conditions are written as

$$G(\pi, q, \nu; x) = 0$$

along the lines of Bisin-Moro-Topa (2009). In fact, e.g., the example in Section 4.2 displays multiple equilibria.

1. The function  $F(x)$  is an equilibrium mapping. It will depend on deep parameters which determine  $\Delta C(q, I)$ .<sup>20</sup> In other words, the model implies cross-equation restrictions between  $\pi = q + \frac{1}{\alpha}(1 - q)^2 \Delta C(q, I) - \frac{1}{\alpha}(1 - q)(\gamma_\tau \cdot x)$  and  $\begin{pmatrix} q \\ \nu \end{pmatrix} = F(x)$ . By approximating  $\Delta C(q, I)$  via polynomials, we are implicitly disregarding these restrictions. At this stage, however, the objective of our empirical work is not to estimate the deep parameters of the model. But rather just to test if  $\frac{\partial \Delta C(q, I)}{\partial q} < 0$  against the alternative hypothesis that it is  $\geq 0$  (to test *cultural distinction* against *cultural conformity*).
2. The identity formation process of the members of a minority depends on the cultural characteristics of the minority itself (e.g., its cultural distance from the majority), but it also depends on the actions and predispositions of the majority (e.g., its racial attitudes). We do not have accurate data to distinguish these determinants of identity, though  $x$  contains controls of some relevance to this effect, e.g., a measure of the episodes of ethnic/religious harassment each respondent has been subject to.

### 3.2 Results

Results of the empirical analysis, reported in Figures 3a and 3b (whole sample and Muslim sample, respectively), clearly favor *cultural distinction*:  $\frac{\partial \Delta C(q, I)}{\partial q} < 0$  at both  $I = 0$  and  $I = 1$ .

[Insert Figure 3a and 3b here]

## 4 Ethnic integration: Structural models

In this section we complement the analysis of the previous section by structurally studying ethnic integration via marriage, location, and identity formation. Formally, this requires expliciting a model for  $\Delta C(q, I)$ . In fact we will study two distinct models, one in which we assume location is exogenous and one in which we assume identity is exogenous (e.g., determined by parental socialization). We race two different, extreme (semi-nested) models:

1. Marriage and identity (with exogenous location): people are dropped *tamquam tabula rasa* in a neighborhood and there they form their ethnic identity and look for a spouse.
2. Marriage and location (with exogenous identity): people are born with an identity (perhaps thanks to the family they are born into) and with this look for a neighborhood where to reside and a spouse.

---

<sup>20</sup>It might be convenient in general to restrict the map  $F$  as follows:

$$\begin{aligned}
 F(x) &= f(X_q, X_\nu) \text{ for} \\
 f &: \mathbb{R}^2 \rightarrow [0, 1]^2 \text{ and } \begin{cases} X_q = \gamma_q x \\ X_\nu = \gamma_\nu x \end{cases}
 \end{aligned}$$

## 4.1 Endogenous identity model

For this, let us first define  $C(H, I)$  the psychological cost to interact with individuals of the dominant majority; depending on the marriage status  $H$  and the identity status  $I$  of the minority member with  $(H, I) \in \{0, 1\}^2$ . We assume that  $C(H, I) = (1 - IH)C(q)$ . This formulation basically captures the idea that a minority member faces a psychological cost  $C(q)$  of interaction with outsiders. This psychological cost can be reduced by identity and homogamy status acting as necessary complements to each other. This specification implies

$$\Delta C(q, I) = IC(q). \quad (4)$$

Assume location  $q$  is exogenous. The probability  $\nu = \text{prob}\{I = 1\}$  is modeled as a choice of the agent. The utility cost of developing identity  $\nu$ ,  $J(\nu)$  is increasing and convex (in the same units of the psychological costs  $C(q)$ ); for simplicity we assume

$$J(\nu) = \frac{1}{2}\nu^2. \quad (5)$$

A minority member's problem is:

$$\max_{\nu, \tau} -\pi(\tau, q)(1 - \nu)C(q) - [1 - \pi(\tau, q)]C(q) - Z(\tau) - J(\nu)$$

Given (3), (4), and (5), the first order condition of the problem are easily reduced to:

$$\nu = C(q)\pi \quad (6)$$

$$\pi = q + (1 - q)^2 \frac{C(q)}{\alpha} \nu \quad (7)$$

This is a simultaneous equation system in which  $(\nu, \pi)$  are the endogenous variables and  $q$  the exogenous variable, which is solved to:

$$\nu = \frac{qC(q)}{1 - \frac{1}{\alpha}(C(q))^2(1 - q)^2} \quad (8)$$

$$\pi = \frac{q}{1 - \frac{1}{\alpha}(C(q))^2(1 - q)^2} \quad (9)$$

## 4.2 Endogenous location model

Consider again the case in which a minority member not homogamously married in a neighborhood with a fraction  $q$  of minority members and whose identity is  $I$  faces additional psychological costs

$$\Delta C(q, I) = IC(q).$$

Assume however now that the probability of  $I = 1$ ,  $\nu$ , is exogenous, while location  $q$  is a choice of the agent. The utility cost of choosing to live in a neighborhood with composition  $q$ ,  $G(q)$  is increasing

and convex (in the same units of the psychological costs  $C(q)$ ),<sup>21</sup> for simplicity we assume

$$G(q) = \frac{1}{2}q^2 \quad (10)$$

The minority member's problem writes now as i

$$\max_{q, \tau} -\pi(\tau, q)(1 - \nu)C(q) - [1 - \pi(\tau, q)]C(q) - Z(\tau) - G(q)$$

Given (3), (4), and (10), the first order condition of the problem can be written as:

$$\begin{aligned} -[1 - \pi\nu]C'(q) + (1 - \tau)\nu C(q) - q &= 0 \\ (1 - q)\nu C(q) - \alpha\tau &= 0 \end{aligned}$$

or:

$$\begin{aligned} -(1 - \pi\nu)C'(q) + \nu\frac{1 - \pi}{1 - q}C(q) - q &= 0 \\ \nu C(q) - \alpha\frac{\pi - q}{(1 - q)^2} &= 0 \end{aligned}$$

This is a simultaneous equation system in which  $(q, \pi)$  are the endogenous variables and  $\nu$  the exogenous variable. The solution is rather involved. For our empirical application it is convenient to write the system in terms of  $\left(\frac{\nu}{\pi}\right)$  as a function of  $\pi, q$ . After some algebra one gets:

$$\nu = \alpha\frac{\pi - q}{(1 - q)^2 C(q)} \quad (11)$$

$$\pi = \frac{-B \pm \sqrt{B^2 - 4AD}}{2A}, \text{ where} \quad (12)$$

$$\begin{aligned} A &= \alpha(1 - q)C'(q) - \alpha C(q) \\ B &= \alpha C(q)(1 + q) - \alpha q(1 - q)C'(q) \\ D &= -(1 - q)^3 C(q)(C'(q) + q) - \alpha C(q)q \end{aligned} \quad (13)$$

The system has two solutions, in principle.

### 4.3 Conformity vs. distinction: A fully structural analysis

In the *cultural conformity* model the minority's psychological costs to interact with individuals of the dominant majority are decreasing in the proportion of the majority in the reference neighborhood, where the member of the minority resides,  $1 - q$ . The simplest formulation therefore has:

$$C(q) = cq$$

---

<sup>21</sup>We assume that neighborhoods with larger minority representation  $q$  are less preferable, other things equal (e.g., because of higher unemployment, less average income etc.). We do not need to make this assumption. We could estimate the sign of the dependence of costs from  $q$ . But all agents would live in segregated neighborhoods under the assumptions of our model if they were cheaper to reside in.

In the *cultural distinction* model the minority's psychological costs to interact with individuals of the dominant majority are increasing in the proportion of the majority in the reference neighborhood, where the member of the minority resides,  $1 - q$ . The simplest formulation therefore has:

$$C(q) = c(1 - q)$$

It is then straightforward to distinguish formally between the cultural substitution and cultural distinction models, that is, to nest the reduced form equations, by writing

$$\begin{aligned} C(q) &= c\psi(q), \text{ with} \\ \psi(q) &= \gamma_1 - \gamma_2 q \end{aligned}$$

so that

$$\gamma_1 = \gamma_2 = 1 \quad \text{distinction} \tag{14}$$

$$\gamma_1 = 0, \gamma_2 = -1 \quad \text{conformity}$$

Empirically, to discriminate between the two models it is sufficient to verify whether  $\gamma_2$  is greater or smaller than zero.

### 4.3.1 Empirical implementation

Both the endogenous location and the endogenous identity model are identified exploiting the nonlinearities induced by the choice problems and the functional form assumptions we impose.<sup>22</sup>

Adding the empirical implementation of costs

$$Z(\tau) = \frac{1}{2}\alpha\tau^2 + (\gamma_\tau \cdot x) \tau, \quad J(\nu) = \frac{1}{2}\nu^2 + (\gamma_\nu \cdot x) \nu, \quad G(q) = \frac{1}{2}q^2 + (\gamma_q \cdot x) q$$

respectively, to (9) and (13), we get a simultaneous equation system  $(\nu, \pi)$  for the endogenous identity model and one in  $(q, \pi)$  for the endogenous location model. The Appendix provides the details on the systems of equations we put to data as well as simple informal argument for identification of parameters  $\gamma_1, \gamma_2, c, \alpha, \gamma_\nu, \gamma_\tau$  in the case of endogenous identity.

Writing a likelihood and searching for the structural parameters  $(\gamma_1, \gamma_2, c, \alpha, \gamma_\nu, \gamma_\tau)$  is straightforward.<sup>23</sup> Because there are no a-priori arguments to select the variables to be included as determinants of the costs of identity and homogamy, we consider  $x_\nu = x_\tau = x$ .

We observe  $n$  (independent) bivariate Bernoulli trials with a pair of characteristics being studied at each trial. The probabilities of the outcomes vary over the trials.

$$prob\{I_i = x_i\} = \begin{cases} \nu_i & x_i = 1 \\ (1 - \nu_i) & x_i = 0 \end{cases} \quad \text{and} \quad prob\{H_i = y_i\} = \begin{cases} \pi_i & y_i = 1 \\ (1 - \pi_i) & y_i = 0 \end{cases}$$

<sup>22</sup>In particular there are no issues of *logical consistency*, as they might arise in models in which, for each agent,  $\nu$  and/or  $\pi$  depend on the realization of  $I$  and/or  $H$  (see Maddala, p. 118-9 and Wilde, 2000).

<sup>23</sup>Otherwise one can proceed via GMM as in Bisin-Topa-Verdier (2003).

Once the joint determination of  $I$  and  $H$  is explicitly accounted for by the forms of the probabilities  $\nu$  and  $\pi$ , the two random variables may be assumed as independent. Hence

$$prob\{I_i = x_i, H_i = y_i\} = \nu_i^{x_i} \times (1 - \nu_i)^{1-x_i} \times \pi_i^{y_i} \times (1 - \pi_i)^{1-y_i}$$

The likelihood function can be written as:

$$L = \prod_{i=1}^n \nu_i^{x_i} \times (1 - \nu_i)^{1-x_i} \times \pi_i^{y_i} \times (1 - \pi_i)^{1-y_i} \quad (15)$$

The maximization of the likelihood function (15) under the two (alternative) models (14) will then uncover which one of the two models of integration formation better fits our data.

### 4.3.2 Results

The estimates of the endogenous identity and endogenous location model are reported, respectively, in Tables 3a and 3b, under different set of controls (none, individual controls, contextual controls).<sup>24</sup>

[Insert Table 3a and 3b here]

Model selection statistics overwhelmingly favor the endogenous identity model over the endogenous location model. Following Burnham and Anderson (2002) we adopt the Akaike Information Criterion (AIC) as a statistical test for model selection. In fact, in our simple context, the AIC of non-nested models reduce to a likelihood ratio test. Likelihoods are reported in Table 2. It is apparent that the endogenous identity model (with exogenous location) fares best.<sup>25</sup>

We concentrate then on the *endogenous identity model* from now on. An unconstrained maximization of the likelihood of the endogenous identity model gives the results listed in Table 3a. A positive estimate of  $\gamma_2$  would be in line with the cultural distinction model, whereas a negative estimate would support the cultural conformity model. Table 3a shows that such an estimate is indeed positive, which indicates that a cultural distinction mechanism of identity formation is more consistent with the data.

We also pursue the estimation of the model (maximization of the likelihood function (15)) under the two (alternative) sets of constraints (14), distinction and conformity. Results are still in Table 3a. The maximized likelihood value is much higher when the constraints implied by the cultural distinction model (i.e.  $\gamma_2 = 1$ ) are imposed.<sup>26</sup> The parameter estimates for the cultural distinction model (i.e., the preferred one) are also reported in Table 3a.<sup>27</sup> As noted before, the effects of the exogenous regressors are not additively separable from the effect of  $q$ .

<sup>24</sup>The estimation has been performed using R programming language ([www.r-project.org](http://www.r-project.org)). The code is available upon request.

<sup>25</sup>For only one of the two solutions of the endogenous location model the estimate converges.

<sup>26</sup>A more rigorous statistical comparison is problematic in these cases because the distribution of the resulting likelihood ratio tests with inequality constraints is non-standard (Chernoff, H. (1954), Wilks S. S. (1938), Self, S.G. and K.-Y. Liang, (1987), Shaw, F. H. and C. J. Geyer (1997) ).

<sup>27</sup>Here the value of  $\gamma_2$  is set at 1 (see (14)).

To develop a better intuition about which aspect of the data drives this result, we study the endogenous identity model in more detail. Under *cultural distinction* the first order conditions of the model are:

$$\nu = c(1 - q)\pi \tag{16}$$

$$\pi = q + \frac{c}{\alpha}(1 - q)^3\nu \tag{17}$$

Various simple conclusions can be obtained from these equations. Consider first condition (16), expressing how identity formation  $\nu$  depends on the proportion  $q$  of minority members and on  $\pi$  the probability of homogamy. Clearly the larger the proportion  $q$  of minority members and the more segregated the neighborhood, the smaller  $C(q) = 1 - q$ , the psychological cost of interacting with majority people, and the smaller the incentives for identity formation as cultural distinction. On the other hand, the larger the probability  $\pi$  of homogamous marriage resulting from socialization efforts, the larger the expected benefits from identity formation and the more intense is identity.

Interestingly, marital segregation, as reflected by  $\pi$ , and neighborhood segregation, as reflected by  $q$ , are substitutes in terms of identity formation. In other words, the marginal effect of marital segregation on identity formation tends to be reduced the more segregated the neighborhood (i.e., the larger is  $q$ ).

In summary, under cultural distinction: *i) Identity is decreasing with neighborhood segregation and increasing with minority homogamy. ii) Marital segregation and neighborhood segregation are substitutes for identity formation.*

Consider now equation (17) characterizing the (endogenous) probability of homogamy as a function of identity  $\nu$  and neighborhood segregation  $q$ . The more intense identity formation, the larger the probability of homogamy. On the other hand, the effect of  $q$  on  $\pi$  is ambiguous and reflects two opposite effects. First, there is a direct effect related to the fact that the larger the proportion of minority people in the neighborhood, the larger the probability of finding a minority spouse in the common pool of potential partners. This effect is reflected in the first term  $q$  of (17). The second effect is illustrated by the second term  $[1 - q]^2 C(q)\nu$  and shows the effect of a change in  $q$  on the marginal incentives to marital segregation (i.e., the socialization effort  $\tau$ ). Indeed the more segregated the neighborhood (i.e., the larger is  $q$ ), the smaller the incentives to spend resources to find directly a partner in the restricted pool of minority spouses. First because social interactions with the majority people are less costly psychologically and there is less of a need for identity formation that can be effectively expressed in homogamous marriages. This is reflected by the term  $C(q)$ . Second, a larger proportion of minority people in the neighborhood reduces as well the incentives to make special efforts to find a spouse in a segregated marital pool, as minority people are already more likely to be found in the common marital pool. Both channels reduce therefore the incentives for socialization efforts  $\tau$ , which in turn tends to reduce the probability of homogamy  $\pi$ .

It is also interesting to observe that, for the choice of socialization effort and probability of homogamy, identity and neighborhood segregation (as reflected by  $q$ ) are substitutes. More precisely, the marginal effect of identity on minority homogamy tends to be reduced the more segregated the neighborhood (i.e., the larger is  $q$ ).

This discussion can be summarized as follows: *i) The probability of marital segregation is increasing in the intensity of identity while the effect of neighborhood segregation is ambiguous. ii) Identity and neighborhood segregation are substitutes for homogamy.*

Under *cultural conformity*, instead, the first order conditions of the model are:

$$\nu = cq\pi \tag{18}$$

$$\pi = q + \frac{c}{\alpha}q(1 - q)^2\nu \tag{19}$$

Again simple conclusions can be obtained from these equations. Consider first (18). In this case, the larger the proportion  $q$  of minority members and the more segregated the neighborhood, the higher  $C(q)$ . Similarly, the larger the probability  $\pi$  of homogamous marriage, the larger the expected benefits from identity, and hence the more intense is identity formation. Again the sign of the cross derivative is interesting. Marital segregation (as reflected by  $\pi$ ) and neighborhood segregation (as reflected by  $q$ ), under cultural conformity, are complements in terms of identity formation. In other words, the marginal effect of marital segregation on identity formation is larger, the more segregated the neighborhood (i.e., the larger is  $q$ ).

In summary, under cultural conformity: *i) Identity is increasing with neighborhood segregation and increasing with minority homogamy. ii) Marital segregation and neighborhood segregation are complements for identity formation.*

Consider now equation (19). The more intense is identity formation, the larger is the probability of homogamy. The effect of  $q$  on  $\pi$  is also ambiguous and reflects now three effects. First, there is as before the direct effect related to the fact that the larger the proportion of minority people in the neighborhood, the larger the probability of finding a minority spouse in the common pool of potential partners. This effect is illustrated by the first term  $q$  in (19). A second positive effect is illustrated by  $\nu C(q)$  in the second term  $[1 - q]^2 \nu C(q)$ . The larger  $q$ , the larger the conformity psychological gain of social interactions with other minority individuals in the neighborhood and the associated identity formation process that can be effectively expressed in homogamous marriages. This increases the incentives for homogamous marriages and the marginal incentives to marital segmentation (i.e., the socialization effort  $\tau$ ).

The last effect of  $q$  on  $\pi$  is negative. As in the case of cultural distinction, it reflects simply the fact that a larger  $q$  reduces the incentives to make special efforts to find a spouse in a segregated marital pool., as minority people are already likely to be found in the common marital pool. This channel, captured by the expression  $[1 - q]^2$  in the second term  $[1 - q]^2 \nu C(q)$  of equation (19), decreases the incentives for socialization efforts  $\tau$ , and tends to reduce the probability of homogamy  $\pi$ .

Finally, identity and neighborhood segregation (as reflected by  $q$ ) interact in terms of the formation of homogamous minority marriages. The second cross derivative of  $\pi$  with respect to  $\nu$  and  $q$  has the sign of

$$\frac{d}{dq} \{ (1 - q)^2 C(q) \},$$

which is in general ambiguous. When  $C(0) = 0$ , namely when identity formation gains are very small

for small minority populations, it is easy to see that

$$\frac{d}{dq} \{(1 - q)^2 C(q)\} > 0$$

for small enough values of  $q$ . In that case, neighborhood segregation and identity are complements for homogamy.

We conclude that: *i) The probability of marital segregation is increasing in the intensity of identity while the effect of neighborhood segregation is ambiguous. ii) For small enough minority groups, identity intensity and neighborhood segregation are complements for homogamy.*

From the first order conditions of the cultural distinction and the cultural conformity model it appears clear that the distinctive characteristics of the cultural distinction model are:

1.  $\nu$  is decreasing in  $q$ , for  $q$  large; and
2.  $\frac{\partial^2 \pi}{\partial \nu \partial q} < 0$ .

The reader might want to conclude that these are the characteristics of cultural distinction we are identifying in the data.

### 4.3.3 On the endogeneity of $q$ once again

In the previous section we document that the data favor the endogenous identity over the endogenous location model. The endogenous identity model, in turn, postulate an exogenous  $q$ . Observe that  $q$  is exogenous for our purposes even if respondents do choose the neighborhood in which they reside as long as their choice does not depend on the ethnic composition of the neighborhood. In this section we offer some evidence to support the exogeneity of  $q$ . More precisely we present some facts about the ethnic identity of people whose residential choices are exogenous with respect of ethnic preferences and beliefs that seem to be still in line with our analysis. Indeed, the FSEM contains also detailed questions about the residential neighborhood and residential neighborhood preferences. Specifically, it is asked whether the residential neighborhood is good, neither good nor poor, poor for being with other people of the respondent's ethnic group and if, given a choice, the respondent would like to move out of this area or would prefer to stay there. The sample of people declaring that the residential area is poor for being with other people of their own ethnic group and that prefer to stay are those for which the residential choice can be considered as exogenous with respect of ethnic preferences and beliefs. However, we cannot run our analysis on this sub-sample because those detailed questions about neighborhood and the detailed questions about ethnic identity are asked to different individuals. This is a consequence of the FSEM design. Specifically, in the FSEM sample design, in each ethnic minority household, up to two adults were selected at random to answer questions about themselves. Because the number of questions to be asked would have made the interview too long if everybody had been asked all of the possible questions, two versions of the questionnaire were used.<sup>28</sup> They contain different level of detail on the different topics. Our analysis is run on the sample of individuals

---

<sup>28</sup>In single adult households, the questionnaire was assigned at random.

who answer to the questionnaire with detailed questions on self and culture. Nevertheless we can use the respondents to the other questionnaire to provide some facts about the relationship between neighborhood preferences and one indicator of ethnic identity. In fact, the alternative questionnaire asks whether the respondents strongly agree, agree, neither agree nor disagree, disagree or disagree strongly with the statement "people of Black-Caribbean and Asian origin should try to preserve as much as possible of their culture and way of life". Therefore, we select those individuals for which  $q$  can be considered as exogenous and inspect the distribution of those declaring to strongly agree to the previous statement, i.e. those with a strong ethnic identity, across neighborhood classes. Evidence in contrast with our analysis would be the findings of large fractions of people with high identity in more segregated neighborhood, thus pointing towards a cultural conformity behavioral mechanism. On the contrary, we find the picture illustrated in Table 4. Because of our criterion in identifying individuals with exogenous location, the majority of people (63%) live in highly mixed neighborhood ( $q < 1.99$ ). The interesting fact is that almost the 40% of them declare to have an extremely marked ethnic identity and this percentage decrease the more the neighborhood is segregated. Therefore even this exercise where the neighborhood can be taken as exogenous does not seem to show signs that strong ethnic identity are fostered in more segregated neighborhoods.

*[Insert Table 4 here]*

#### 4.3.4 Simulations

In this section we perform a series of simulations, using the estimates of our best model, the endogenous identity model restricted to cultural distinction, to predict ethnic integration, that is, identity and homogamy as a function of neighborhood composition and of "time in the U.K." (identified as "age" for second generation individuals). More precisely our analysis suggests that ethnic identities might be more intense in mixed rather than in segregated neighborhood. The additional question that we seek to investigate is the role played by neighborhood composition when the evolution of time of ethnic identity is considered, in particular distinguishing between first and second generations of immigrants.

Figures 4a and 4b show the predicted identity and homogamy values as a function of time spent in UK by generations. It appears a tendency toward ethnic integration which is particularly pronounced for the second generation, in particular for homogamy.

*[Insert Figure 4a and 4b here]*

We then simulate those patterns when setting neighborhood composition equal to the sample minimum, average and maximum level. The results are reported in Figures 5a and 5b, for identity and homogamy respectively. Although the starting values are always greater in mixed neighborhood (with the exception of homogamy for second generations), the pictures show a steeper decreasing trend over time in mixed neighborhood rather in segregated neighborhood. This is true for both identity and homogamy and both generations.

[Insert Figure 5a and 5b here]

## 5 Is Muslim identity different?

A large debate has recently spurred in the press about the alleged specificity of Muslim immigrants with regards to the strength of their identity and their (lack of) assimilation tendencies.<sup>29</sup>

Several of the ethnic groups for which we have data have in fact a significant Muslim population; notably Pakistani and Bangladeshi are predominantly Muslim, while Indians and African-Asian have substantial Muslim minorities. Furthermore, the FNSEM survey contains a question asking the respondent to identify his religious faith. In this section we exploit therefore our data to address directly the alleged Muslim specificity issue.

To this end we repeat our analysis on the restricted sample of Muslim respondents (roughly the 43 percent of the whole sample). We maintain however the distribution by ethnic group as the relevant neighborhood composition variable in the identity formation and socialization processes.

Tables 5a and 5b reports our results. They are not qualitatively different from those in Tables 3a and 3b. The endogenous identity with cultural distinction model remains the mechanism more likely to represent the observed evidence. Indeed, the estimate of  $\gamma_2$  is still positive and the maximized likelihood value with the constraints imposed by the cultural distinction model is higher than the value which is obtained when the likelihood is constrained accordingly to the cultural conformity mechanism of identity formation. Such a finding signals that the relationship between ethnic assimilation effort and ethnic neighborhood composition is not different for Muslims in respect of other minorities. In other words, we still find evidence in line with the possibility that ethnic identity and socialization effort are more intense in mixed rather than in segregated neighborhoods when only the Muslim identity is considered.

[Insert Table 5a and 5b here]

## 6 Discussion of results and policy implications

As noted few direct empirical analysis of identity formation exist in the economic literature. Two interesting exceptions are Manning and Roy (2007) and Constant, Gataullina and Zimmermann (2009). Manning and Roy (2007) study the UK Labour Force Survey data for 2001. Their results stand in contrast with ours as they find "no evidence of a culture clash in general, and one connected with Muslims in particular." More specifically, Manning-Roy (2006) adopt a measure of integration constructed from answers to the question "What do you consider your national identity to be? Please

---

<sup>29</sup>This position has been taken, in a rather extreme form, by several nationalist parties, e.g., the Lega in Italy, the Front National in France. But similar though less extreme positions have been taken by center-right parties essentially all over Europe. A clear example of the inflamed rhetoric that often accompanies this debate is Fallaci (2006). See also the discussion in Sheikh (2005) for muslims in the US.

choose as many or as few as apply.” Using this measure they document that a large fraction of those individual in the sample who are born in Britain actually report a British *national identity* and that such fraction is larger for third than for second generation immigrants.<sup>30</sup> The measure of integration adopted in our paper is however conceptually distinct, as it is constructed from questions regarding *importance of religion, attitude towards inter-marriage, and importance of racial composition in schools* (see i-iii) in Section 2.1). It is very well possible that integration in terms of national identity, as measured by Manning-Roy (2006), follows a very different pattern than the integration in terms of attitudes towards religion, marriage, and schooling, which we attempt at measuring in this paper. Consistently with this explanation, Constant, Gataullina and Zimmermann (2009) adopt a definition of integration which accounts for several cultural and religious factors, including social interactions, and finds significantly different integration patterns for Muslims than Christians in the German Socioeconomic Panel 2001, along the lines of our paper. No doubts, however, more data and more theoretical work are necessary to better understand and disentangle the many angles that characterize a complex concept like integration.

Our analysis suggests on the contrary that integration policies favoring the formation of mixed neighborhoods, fearing the effects of geographical segregation, are possibly minimally effective if not counterproductive. Integrationist policies, which include school busing, affirmative action in public schools and in the workplace, forced integration of public housing, and laws barring discrimination in housing and employment, consistently with the view that identity formation mechanisms are driven by *cultural distinction*, have often had limited effects and are even being at times opposed by the same minority groups in whose interest they have been pursued (see e.g., Jacoby, 1998, and Thernstrom and Thernstrom, 2002). J. Coleman, for instance, fifteen years after the Coleman Report (1966) which originally proposed busing, admitted that, “the assumption that busing would improve achievement of lower-class black children has now been shown to be fiction;” (cited in Jacoby, 1999).<sup>31</sup> But Moving to Opportunity (MTO) programs in the United States that relocates families from high- to low-poverty neighborhoods (and from racially segregated to mixed neighborhoods) also have had positive but arguably small effects (see, in particular, Ludwig, Duncan, and Hirschfield, 2001, and Kling, Ludwig, and Katz, 2005).<sup>32</sup> In Europe different integration policies and ambitious social programs have been implemented in urban areas where immigrants live but they also have had limited results. This is the case, for instance, for the creations of Zones of Educational Priority (ZEP) and for the rehabilitation of bleak housing projects in immigrant neighborhoods under the guise of urban policy (‘politique de la ville’) in France. Finally, even racially integrated schools have recently lost much of their appeal in African-American communities (see e.g., the ethnographic study of Gussin Paley, 1995, for schooling).

Far from supporting policies to establish segregated neighborhoods, in this paper, we simply document that the effect of mixed neighborhood on identity formation and socialization might be perverse,

---

<sup>30</sup>See also Modood et al. (1997).

<sup>31</sup>The failure of the busing and other civil right policies is certainly also due to the whites’ flying from de-segregated schools and neighborhoods.

<sup>32</sup>Similarly, the Toronto housing program where adults were assigned as children to different residential housing projects (Oreopoulos, 2003) did not give the expected results in terms of education outcomes.

because of *cultural distinction*.

# Appendix

1. In the case of endogenous identity a simple informal argument for identification of parameters  $\gamma_1, \gamma_2, c, \alpha, \gamma_\nu, \gamma_\tau$  follows. For each agent we observe  $q, x_\nu, x_\tau$  the realization of  $I \in \{0, 1\}$  and the realization of  $H \in \{0, 1\}$ . Since  $prob\{I = 1\} = \nu$  and  $prob\{H = 1\} = \pi$  we identify with  $\nu$  the fraction of agents with identity  $I = 1$  and with  $\pi$  the fraction of agents with homogamous marriages  $H = 1$ . The parameters  $\gamma_1, \gamma_2, c$  and  $\alpha$  are identified out of variations in  $q$ . Variation in  $x_\nu, x_\tau$  identifies instead  $\gamma_\nu, \gamma_\tau$ . This is the case even if a single vector  $x$  enters in the determination of  $\nu$  and  $\pi$ , that is,  $x_\nu = x_\tau = x$ . For given  $q, c, \alpha, \gamma_1, \gamma_2$ , in fact, in this case  $\gamma_\nu, \gamma_\tau$  solve

$$\begin{aligned} -\frac{\gamma_\nu + \frac{1}{\alpha}c\psi(q)(1-q)\gamma_\tau}{1 - \frac{c^2}{\alpha}(\psi(q))^2(1-q)^2} &= k_\nu \\ -\frac{\frac{1}{\alpha}c\psi(q)(1-q)^2\gamma_\nu + \frac{1}{\alpha}(1-q)\gamma_\tau}{1 - \frac{c^2}{\alpha}(\psi(q))^2(1-q)^2} &= k_\pi \end{aligned}$$

where  $k_\nu, k_\pi$  are the estimated coefficients of  $x$  in the equation for  $\nu$  and  $\pi$ , respectively, for given  $q$ . It is easy to check that one such solution exists.

2. The system of equations we put to data are as follows:

(a) Endogenous Identity: The first order conditions of the agent's problem writes as :

$$\begin{aligned} C(q)\pi - \nu - \gamma_\nu \cdot x_\nu &= 0 \\ C(q)(1-q)\nu - \alpha\tau - \gamma_\tau \cdot x_\tau &= 0 \end{aligned}$$

substituting  $\tau = \frac{\pi-q}{1-q}$  gives:

$$\begin{aligned} \nu &= C(q)\pi - \gamma_\nu \cdot x_\nu \\ \pi &= q + \frac{C(q)(1-q)^2\nu}{\alpha} - \frac{1-q}{\alpha}\gamma_\tau \cdot x_\tau \end{aligned}$$

substitution gives

$$\begin{aligned} \nu &= \frac{C(q)q}{1 - \frac{1}{\alpha}(C(q))^2(1-q)^2} - \frac{1}{1 - \frac{1}{\alpha}(C(q))^2(1-q)^2}\gamma_\nu \cdot x_\nu - \\ &\quad \frac{\frac{1}{\alpha}C(q)(1-q)}{1 - \frac{1}{\alpha}(C(q))^2(1-q)^2}\gamma_\tau \cdot x_\tau \\ \pi &= \frac{q}{1 - \frac{1}{\alpha}(C(q))^2(1-q)^2} - \frac{\frac{1}{\alpha}C(q)(1-q)^2}{1 - \frac{1}{\alpha}(C(q))^2(1-q)^2}\gamma_\nu \cdot x_\nu - \\ &\quad \frac{\frac{1}{\alpha}(1-q)}{1 - \frac{1}{\alpha}(C(q))^2(1-q)^2}\gamma_\tau \cdot x_\tau \end{aligned}$$

(b) Endogenous Location: The first order conditions of the agent's problem writes as :

$$\begin{aligned} -[1 - \pi\nu]C'(q) + (1 - \tau)\nu C(q) - q - \gamma_q \cdot x_q &= 0 \\ (1 - q)\nu C(q) - \alpha\tau - \gamma_\tau \cdot x_\tau &= 0 \end{aligned}$$

substituting  $\tau = \frac{\pi - q}{1 - q}$  gives:

$$\begin{aligned} -(1 - \pi\nu)C'(q) + \nu \frac{1 - \pi}{1 - q} C(q) - q - \gamma_q \cdot x_q &= 0 \\ \nu C(q) - \alpha \frac{\pi - q}{(1 - q)^2} - \frac{\gamma_\tau \cdot x_\tau}{1 - q} &= 0 \end{aligned}$$

1. (a) substitution gives

$$\begin{aligned} \nu &= \alpha \frac{\pi - q}{(1 - q)^2 C(q)} + \frac{\gamma_\tau x_\tau}{(1 - q) C(q)} \\ A\pi^2 + B\pi + D &= 0 \end{aligned}$$

$$A = \alpha(1 - q)C'(q) - \alpha C(q)$$

where  $B = \alpha C(q)(1 + q) - \alpha q(1 - q)C'(q) + (C'(q)(1 - q)^2 - (1 - q)C(q))\gamma_\tau x_\tau$

$$D = -(1 - q)^3 C(q)(C'(q) + q + \gamma_q x) - \alpha C(q)q + (1 - q)C(q)\gamma_\tau x_\tau$$

Thus,

$$\nu = \alpha \frac{\pi - q}{(1 - q)^2 C(q)} + \frac{\gamma_\tau x_\tau}{(1 - q) C(q)} \quad \text{and} \quad \pi = \frac{-B \pm \sqrt{B^2 - 4AD}}{2A}$$

## References

- [1] Abrams D. and M. Hogg (1988). *Social identifications: A social psychology of intergroup relations and group processes*. London: Routledge.
- [2] Alba, R.D. (1990), *Ethnic Identity: The Transformation of White America*, New Haven: Yale University Press.
- [3] Alba R.D. (2005), "Bright vs. Blurred Boundaries: Second Generation Assimilation and Exclusion in France, Germany and the United States", *Ethnic and Racial Studies*, 28(1), 20-59.
- [4] Allport G.W. (1954) *The Nature of Prejudice*, Cambridge, MA, Addison-Wesley.
- [5] Akerlof, G.A. and R.E. Kranton (2000), "Economics and Identity," *Quarterly Journal of Economics* 115, 715-753.
- [6] Alesina, A. and E. La Ferrara (2005), "Ethnic Diversity and Economic Performance," *Journal of Economic Literature* 43, 762-800.
- [7] Al-Johar, D. (2005), "Muslim Marriages in America: Reflecting New Identities," *The Muslim World* 95, 557-574.
- [8] Austen-Smith, D. and R.D. Fryer, Jr (2005), "An Economic Analysis of 'Acting White'," *Quarterly Journal of Economics* 120, 551-583.
- [9] Battu, H., McDonald M., & Zenou, Y. (2007). "Oppositional Identities and the Labor Market". *Journal of Population Economics*, 20(3), 643-67..
- [10] Bernheim, B.D. (1994), "A Theory of Conformity," *Journal of Political Economy* 102, 841-877.
- [11] Bisin A., Moro A. and G. Topa (2006). "The Empirical Content of Models with Multiple Equilibria," *2006 Meeting Papers 660, Society for Economic Dynamics*.
- [12] Bisin, A. and T. Verdier (2000), "Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits," *Quarterly Journal of Economics* 115, 955-988.
- [13] Bisin, A. and T. Verdier (2001), "The Economics of Cultural Transmission and the Dynamics of Preferences," *Journal of Economic Theory* 97, 298-319.
- [14] Bisin, A., Topa, G. and T. Verdier (2004), "Religious intermarriage and socialization in the United States," *Journal of Political Economy* 112, 615-664.
- [15] Bisin, A., Patacchini, E., Verdier, T. and Y. Zenou (2008), "Are Muslims Immigrants Different in terms of Cultural Integration?", *Journal of the European Economic Association*, forthcoming.
- [16] Bisin, A., Patacchini, E., Verdier, T., & Zenou, Y. (2008). "Formation and Persistence of Oppositional Identities", PSE Mimeo.

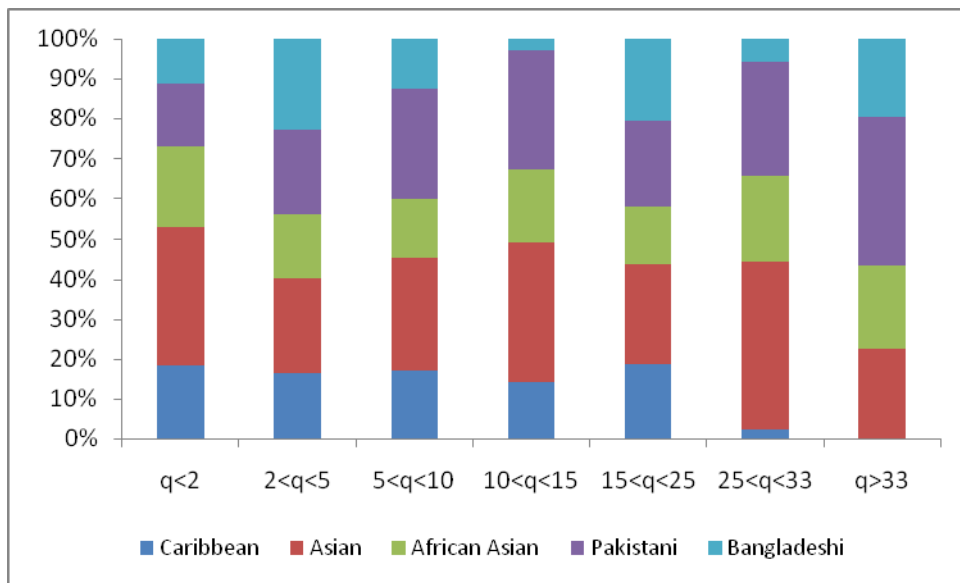
- [17] Bobo L.D. (1999), "Prejudice as a Group Position: Microfoundations of a Sociological Approach to Racism and Racial Relations, *Journal of Social Issues* 55, 445-472.
- [18] Bodendorn H., and C. Ruebeck (2003), "The Economics of Identity and the Endogeneity of Race", NBER WP No 9962
- [19] Boyd, R. and P. J. Richerson (1985), *Culture and the Evolutionary Process*, Chicago: University of Chicago Press.
- [20] Brewer M. (2001), "The Many Faces of Social Identity ; Implications for Political Psychology", *Political Psychology*, 22, 115-125.
- [21] Burnham, K. P. and D. R. Anderson (2002), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, Springer-Verlag.
- [22] Cavalli-Sforza, L.L. and M.W. Feldman (1981), *Cultural Transmission and Evolution*, Princeton: Princeton University Press.
- [23] Chernoff, H. (1954), "On the Distribution of the Likelihood Ratio", *The Annals of Mathematical Statistics*, Vol. 25, No. 3 pp. 573-578.
- [24] Chiswick, C.U. (2009). The Economic Determinants of Ethnic Assimilation. *Journal of Population Economics*, forthcoming.
- [25] Darity, W. Jr., Mason, P.L., & Stewart, J.B. (2006). The Economics of Identity: The Origin and Persistence of Racial Identity Norms. *Journal of Economic Behavior and Organization*, 60(3), 283-305.
- [26] Cross W.E. (1991), *Shades of Black: Diversity in African-American Identity*, Philadelphia: Temple University Press.
- [27] Constant, A.F., & Zimmermann, K.F. (2008). Measuring Ethnic Identity and Its Impact on Economic Behavior. *Journal of the European Economic Association*, 6(2-3), 424-33.
- [28] Constant, A.F., Gataullina, L., & Zimmermann, K.F. (2009). Ethnosizing Immigrants. *Journal of Economic Behavior and Organization*, 69(3), 274-87.
- [29] Eaton C., M. Eswaran and R. Oxoby (2009), "Us and Then: The Origin of Identity, and its Economic Implications". Mimeo
- [30] Fallaci O., (2006), *The Force of Reason*, New York, Rizzoli.
- [31] Fang H., and G. Loury (2005), "Dysfunctional Identities Can be Rational", *American Economic Review*, 95, pp. 104-111.
- [32] Fryer, R.G. Jr. and P. Torelli (2005), "An Empirical Analysis of 'Acting White'," NBER Working Paper No. 11334.

- [33] Gallis, P. (Coordinator) (2005), "Muslims in Europe: Integration Policies in Selected Countries," CRS Report for Congress, Congressional Research Service, The Library of Congress.
- [34] Glazer N. and D.P. Moynihan (1970), *Beyond the Melting Pot: The Negroes, Puerto Ricans, Jews, Italians and Irish of New York City*, Cambridge, MA: The MIT Press.
- [35] Gordon M.M. (1964), *Assimilation in American Life: The Role of Race, Religion and National Origins*, New York: Oxford University Press.
- [36] Gussin Paley, V. (1995), *Kwanzaa and Me: A Teacher's Story*, Cambridge, MA: Harvard University Press.
- [37] Jacoby, T. (1998), *Someone Else's House: America's Unfinished Struggle for Integration*, New York: Free Press.
- [38] Jacoby, T. (1999), "Beyond Busing," *The Wall Street Journal*, July 21st.
- [39] Kling, J.R., Ludwig, J. and L.F. Katz (2005), "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment," *Quarterly Journal of Economics* 120, 87-130.
- [40] Lichter D., J.B. Brown, Z. Qian and J. H. Carmalt (2007), marital Assimilation Among Hispanics: Evidence of Declining Cultural and Economic Incorporation?, *Social Science Quarterly*, 88, 745-765.
- [41] Ludwig, J., Duncan, G.J. and Hirschfield, P. (2001), "Urban Poverty and Juvenile Crime: Evidence from a Randomized Housing-Mobility Experiment," *Quarterly Journal of Economics* 116, 655-679.
- [42] Manning A. and S. Roy (2007), "Culture Clash or Culture Club? The Identity and Attitudes of Immigrants in Britain," CEP Discussion Papers dp0790, Centre for Economic Performance, LSE
- [43] Meng, X. and R.G. Gregory (2005), "Intermarriage and the Economic Assimilation of Immigrants," *Journal of Labor Economics* 23, 135-176.
- [44] Moghaddam F.M and E.A. Solliday (1991) " Balanced Multiculturalism and the Challenge of Peaceful Coexistence inPluralistic Societies", *Psychology and Developing Societies*, 3, 51-71.
- [45] Modood, T., R. Berthoud, J. Lakey, J. Nazroo, P. Smith, S. Virdee, S. Beishon (1997), *Ethnic Minorities in Britain: Diversity and Disadvantage*, London: Policy Studies Institute.
- [46] Oreopoulos, P. (2003), "The Long-Run Consequences of Living in a Poor Neighborhood," *Quarterly Journal of Economics* 4, 1533-75.
- [47] Phinney, J.S. (1990), "Ethnic Identity in Adolescents and Adults: Review of Research," *Psychological Bulletin* 108, 499-514.

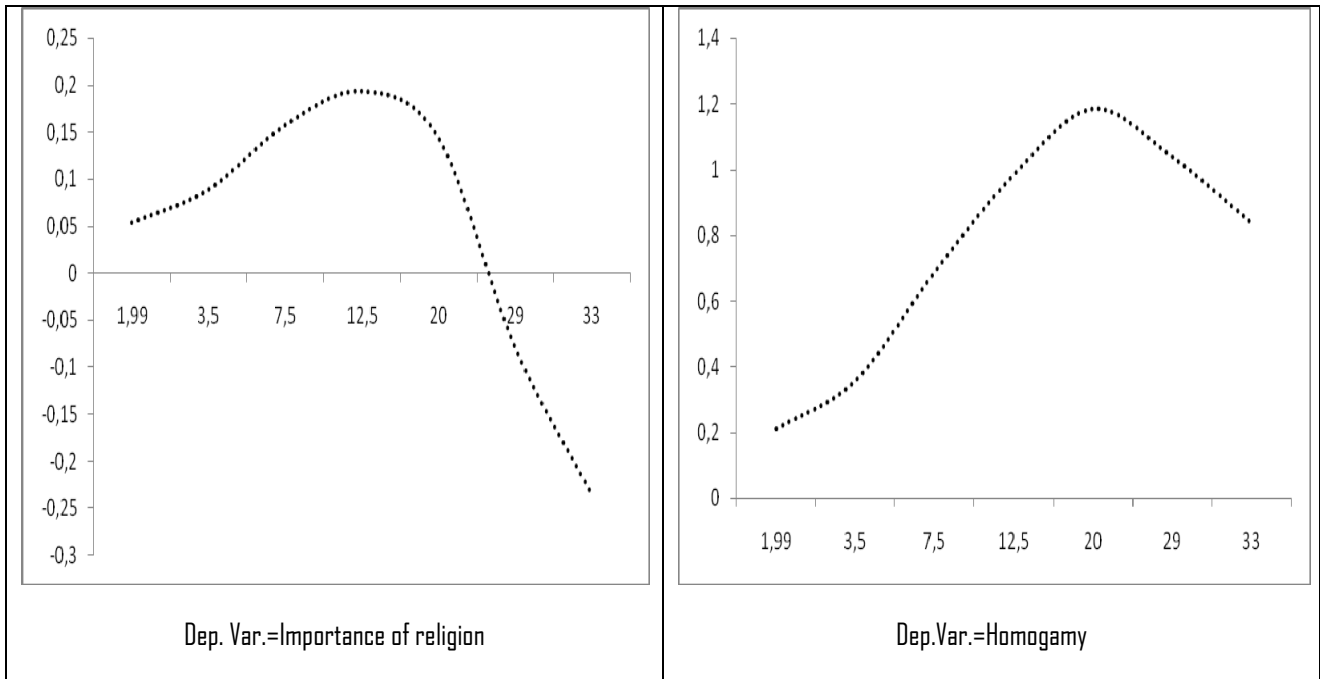
- [48] Putnam R. (2007) "E Pluribus Unum: Diversity and Community in the Twenty First Century; The 2006 Johan Skytte Prize Lecture". *Scandinavian Political Studies*, 30, 137-174.
- [49] Qian, Z. (1999), "Who Intermarries? Education, Nativity Region, and Interracial Marriage 1980 and 1990," *Journal of Comparative Family Studies* 30, 579-597.
- [50] Self, S.G. and K.-Y. Liang, (1987), "Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests Under Nonstandard Conditions", *Journal of the American Statistical Association*, Vol. 82, No. 398, pp. 605-610
- [51] Shaw, F. H. and C. J. Geyer (1997), "Estimation and Testing in Constrained Covariance Component Models", *Biometrika*, Vol. 84, No. 1 , pp. 95-102
- [52] Sheikh C. (2005) "The Effect of Religiosity on Ethnic Identity Among Second-Generation Muslim Americans" Paper presented at the annual meeting of the American Sociological Association Philadelphia [http://www.allacademic.com/meta/p21451\\_index.html](http://www.allacademic.com/meta/p21451_index.html).
- [53] Stryker S. (1980), *Symbolic Interactionism: A Social Structural Version*, Menlo Park, CA: Benjamin Cummings
- [54] Tajfel H.(1981), *Human Groups and Social Categories: Studies in Social Psychology*, Cambridge, England: Cambridge University Press.
- [55] Tajfel H. and J.C.Turner (1979), "An Intergrative Theory of Intergroup Conflict", in W.G. Austin &S. Worchel (Eds.) *The Social Psychology of Intergroup Relations* (pp. 33-47), Monterey , CA Brooks/Cole.
- [56] Taylor D.M.and W.E. Lambert (1996), "The Meaning of Multiculturalism in a Culturally Diverse Urban American Area", *Journal of Social Psychology*, 136, 727-740.
- [57] Thernstrom A. and S. Thernstrom (2002), *Beyond the Color Line: New Perspectives on Race and Ethnicity in America*, New York: Hoover Press and the Manhattan Institute.
- [58] Tucker, M. B. and C. Mitchell-Kernan (1990), "New Trends in Black American Interracial Marriage: The Social Structural Context," *Journal of Marriage and the Family* 52, 209-218.
- [59] Turner J.C. (1982), "Towards a Cognitive Redefinition of the Social Group" in H. Tajfel (ed.) *Social Identity and Intergroup Relations* (pp. 15-44) Cambridge England: Cambridge University Press.
- [60] Turner J.C., M.A. Hogg, P.J. Oakes, S.D Reicher and M.S. Wethrell (1987), *Rediscovering the Social Group: A Self-categorization Theory*, Oxford England: Blackwell.
- [61] Wilson, W. J. (1987), *The truly Disadvantaged: The Inner City, the Underclass, and Public Policy*, Chicago: University of Chicago Press.

- [62] Wilks S. S. (1938), "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses", *The Annals of Mathematical Statistics*, Vol. 9, No. 1, pp. 60-62
- [63] Wooldridge, J. (2002), *The Econometric Analysis of Cross Section and Panel Data*, Cambridge: MIT Press.
- [64] Whyte, W.F. (1943), *Street Corner Society*, Chicago: University of Chicago Press.

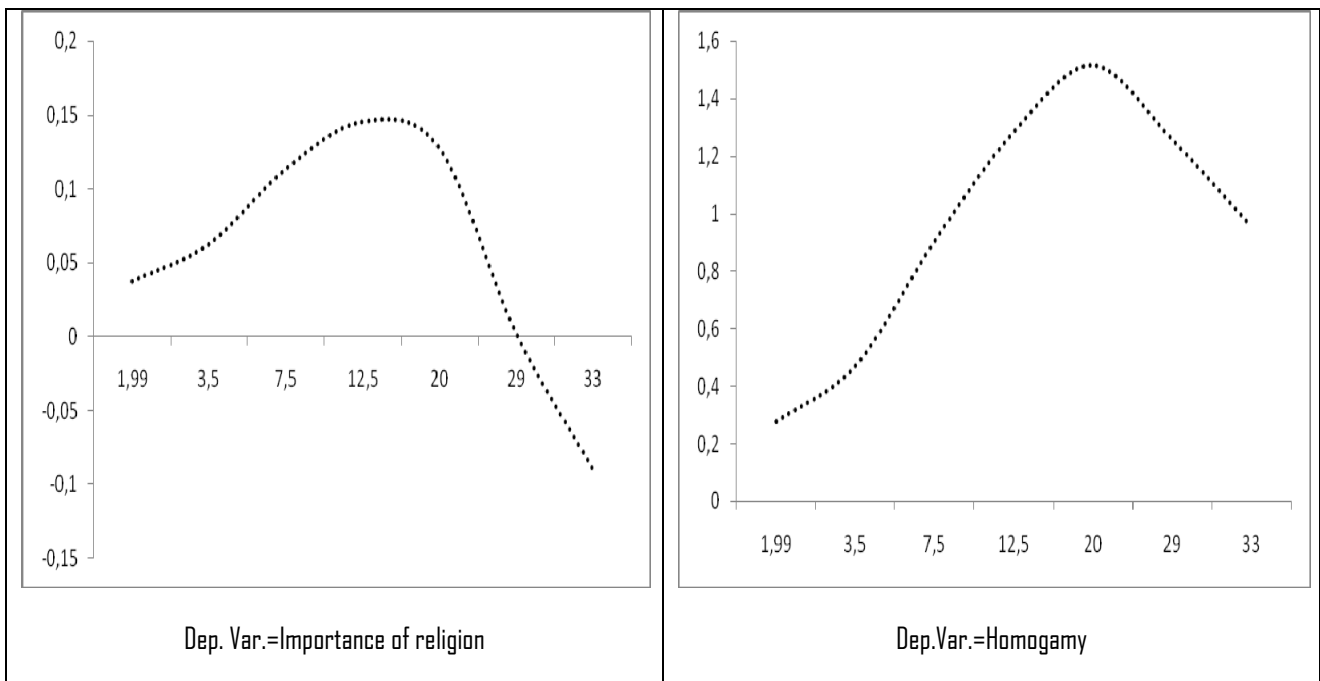
Figure 1: Ethnic composition by neighborhood class



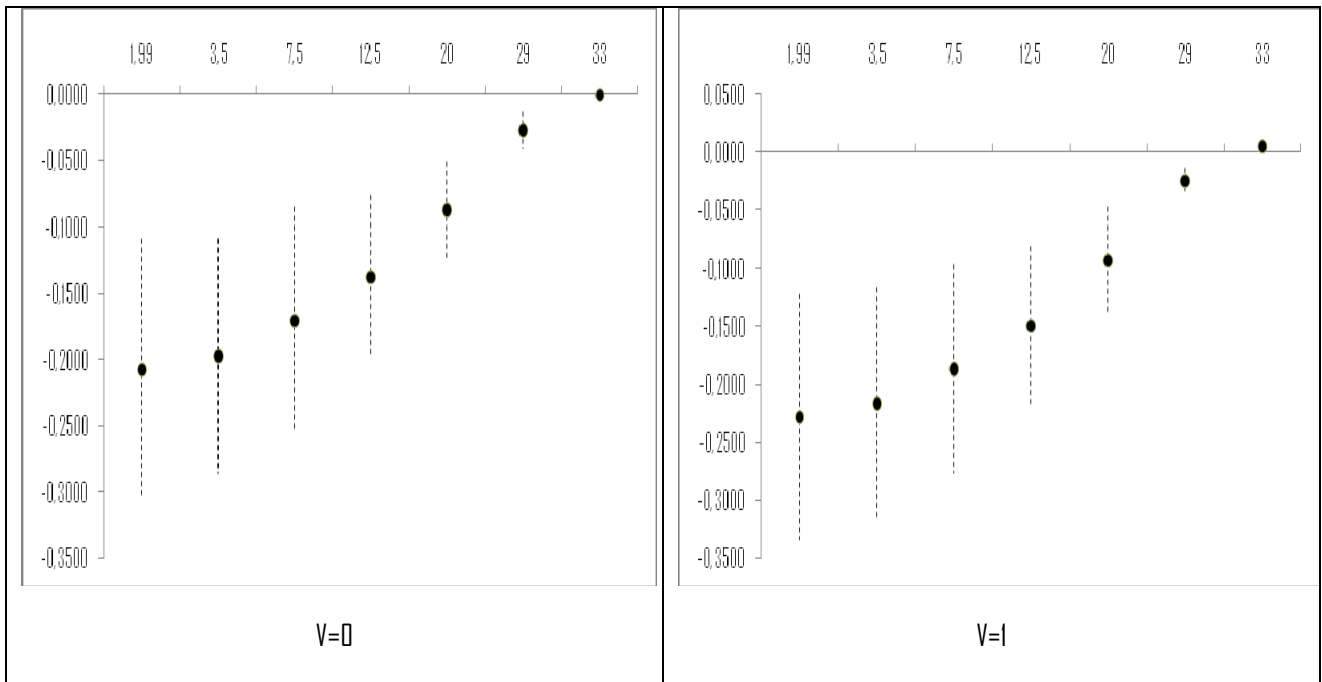
**Figure 2a: Non linear effect of neighborhood ethnic composition on identity and homogamy- whole sample**



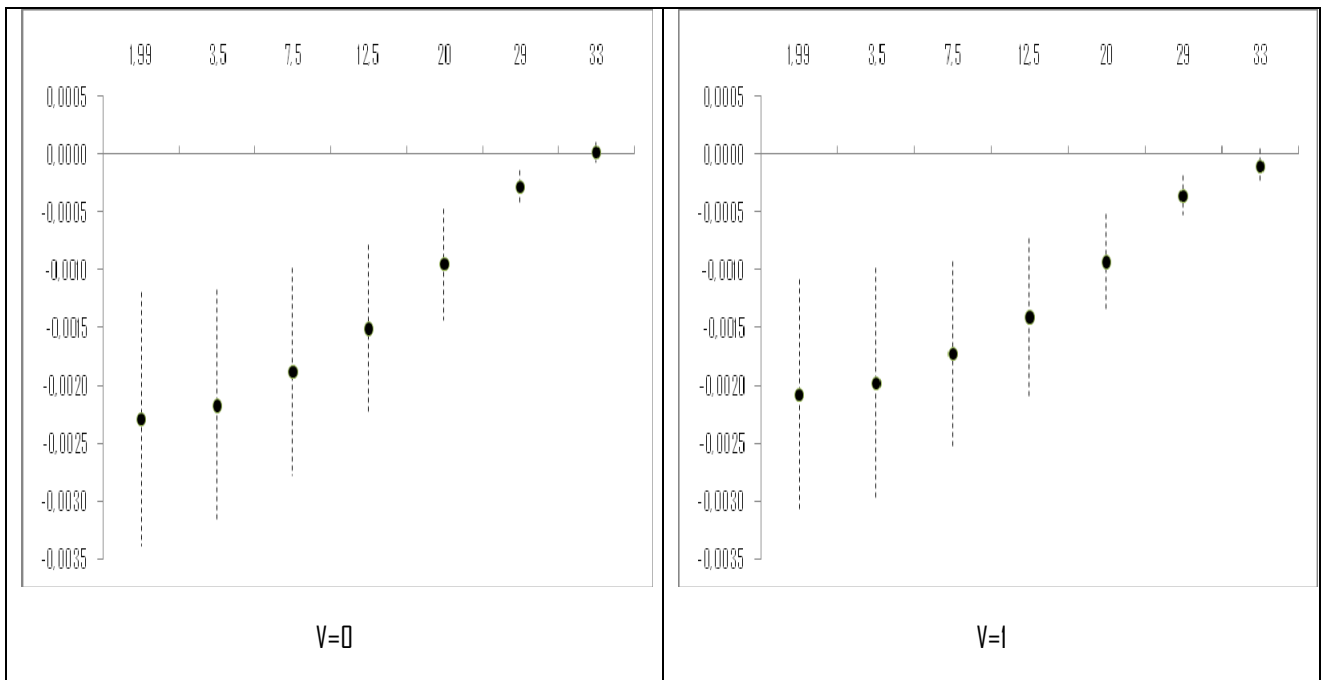
**Figure 2b: Non linear effect of neighborhood ethnic composition on identity and homogamy - Muslim sample**



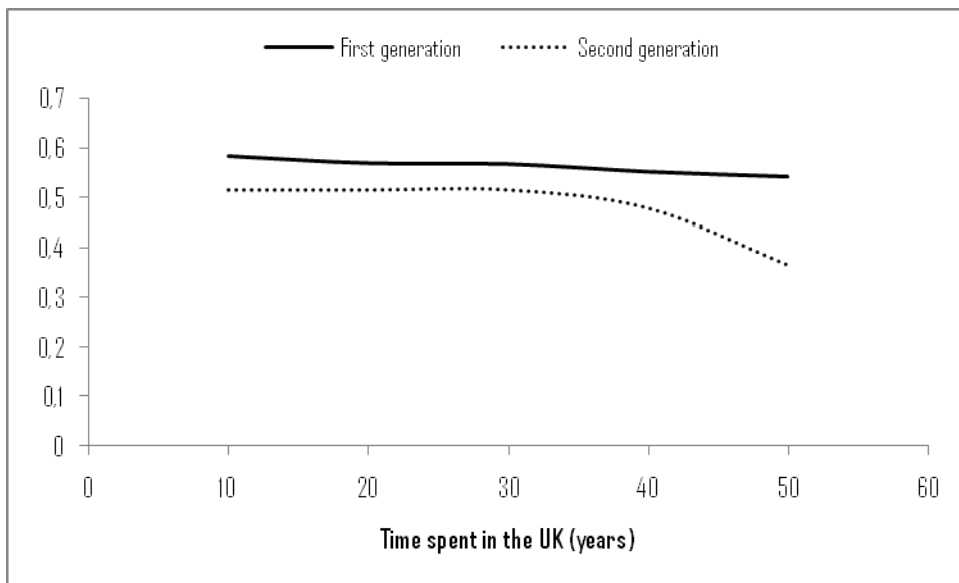
**Figure 3a: Partial derivative of costs as a function of q –whole sample**



**Figure 3b: Partial derivative of costs as a function of q –Muslim sample**

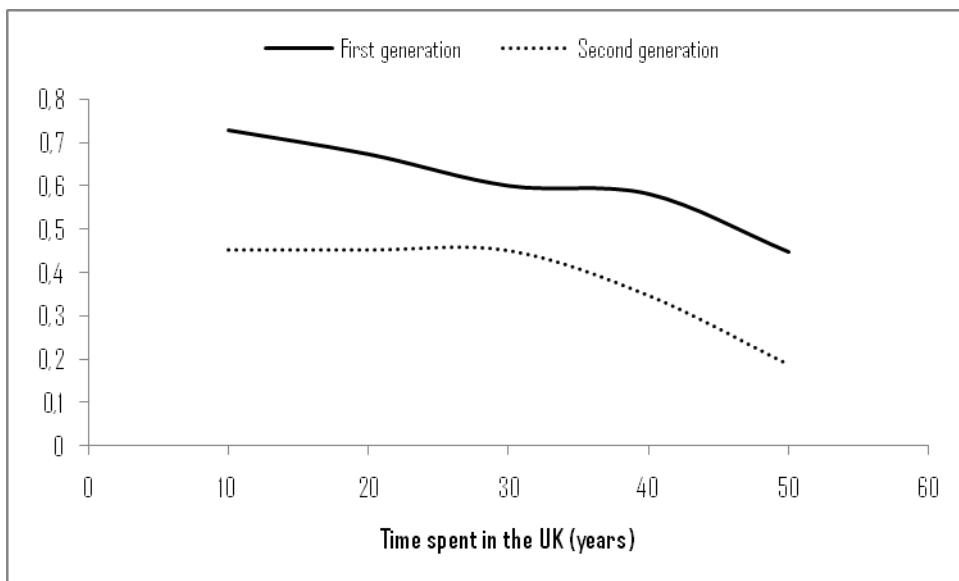


**Figure 4a Predicted Identity as a function of time spent in the UK**

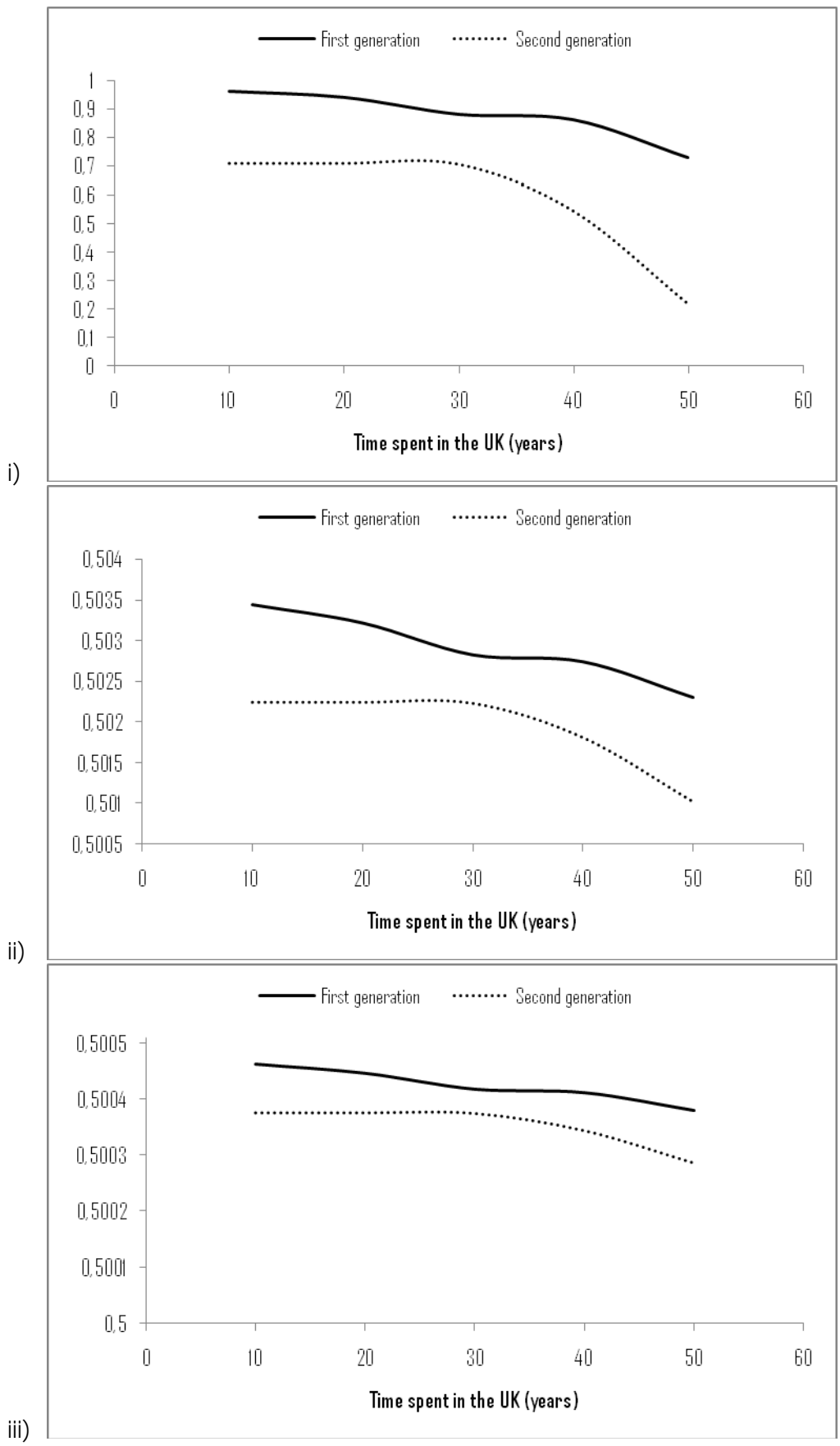


Notes to Figures 4a to 7b. Endogenous identity model, cultural distinction, controls at sample average

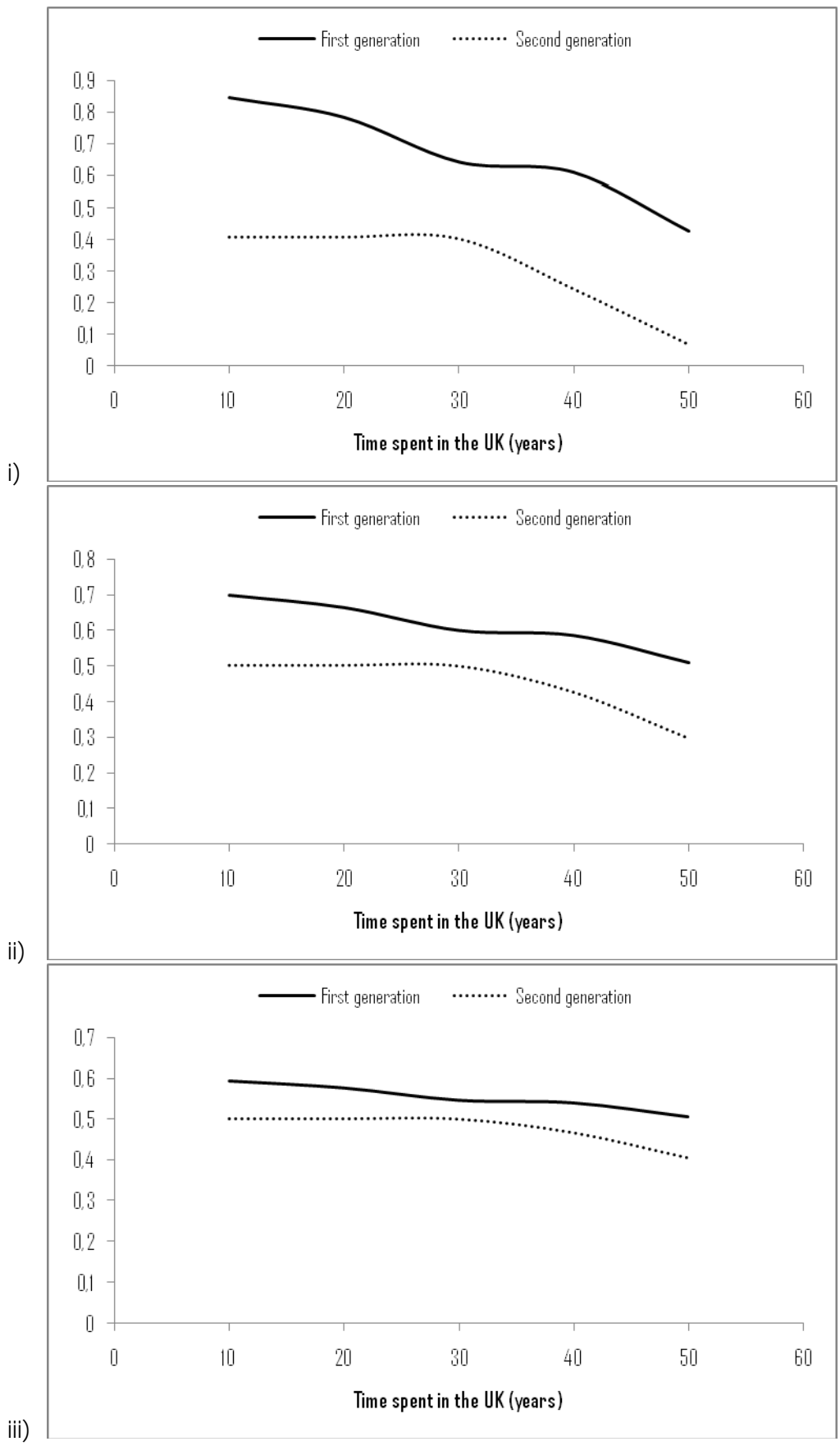
**Figure 4b Predicted Homogamy as a function of time spent in the UK**



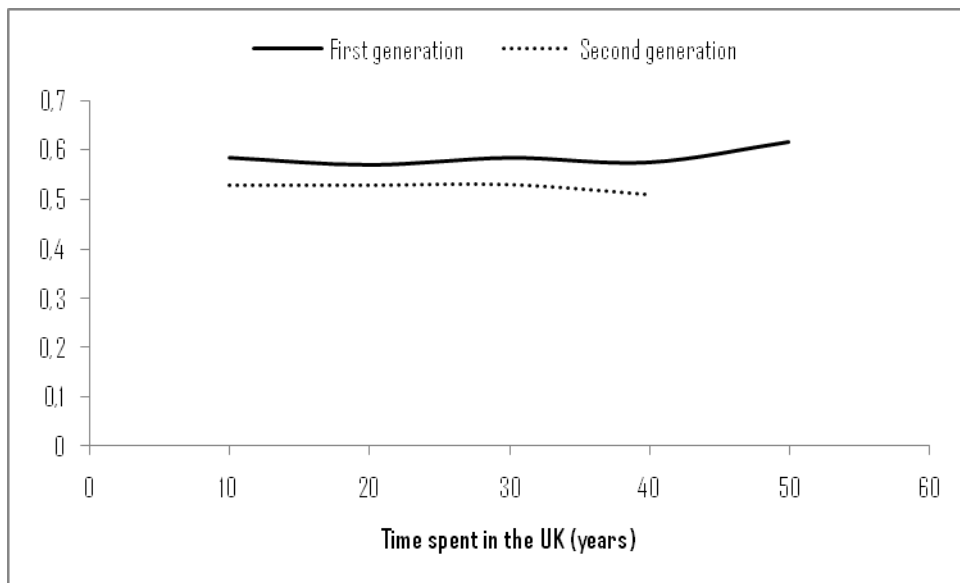
**Figure 5a Predicted Identity as a function of time spent in the UK - i) at minimum q, ii) at average q, iii) at maximal q**



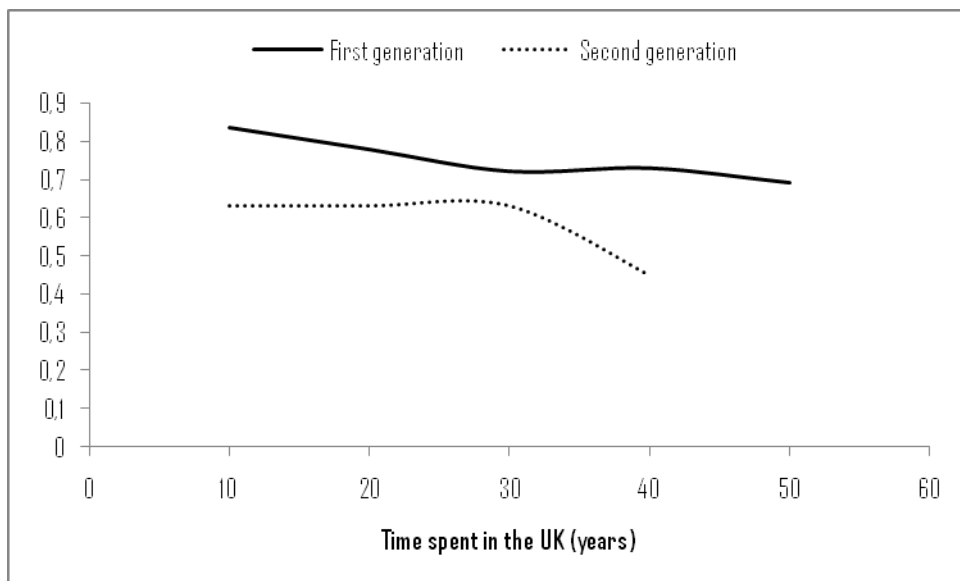
**Figure 5b Predicted Homogamy as a function of time spent in the UK – i) at minimum q, ii) at average q, iii) at maximal q**



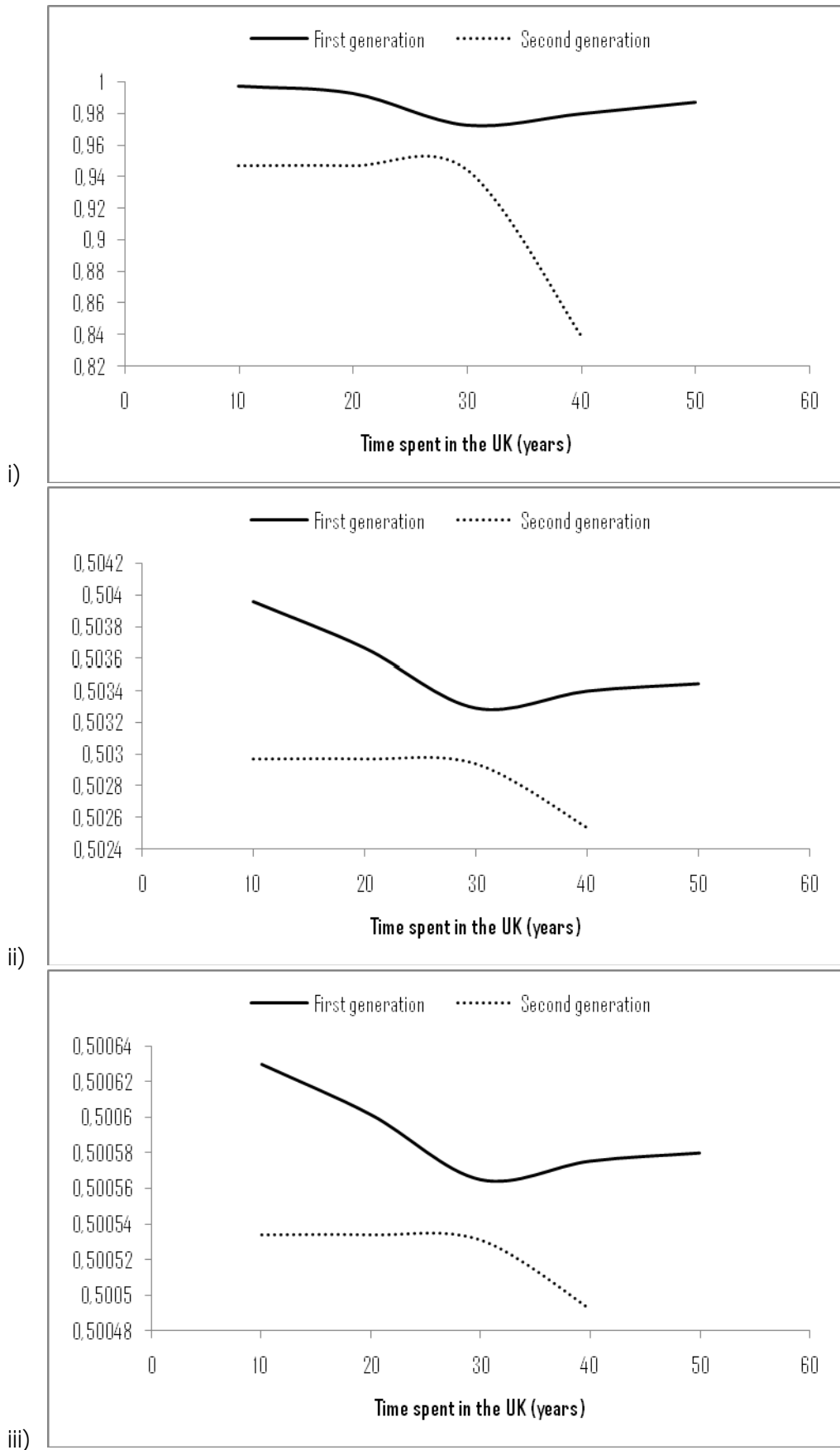
**Figure 6a Predicted Identity as a function of time spent in the UK- Muslim sub-**



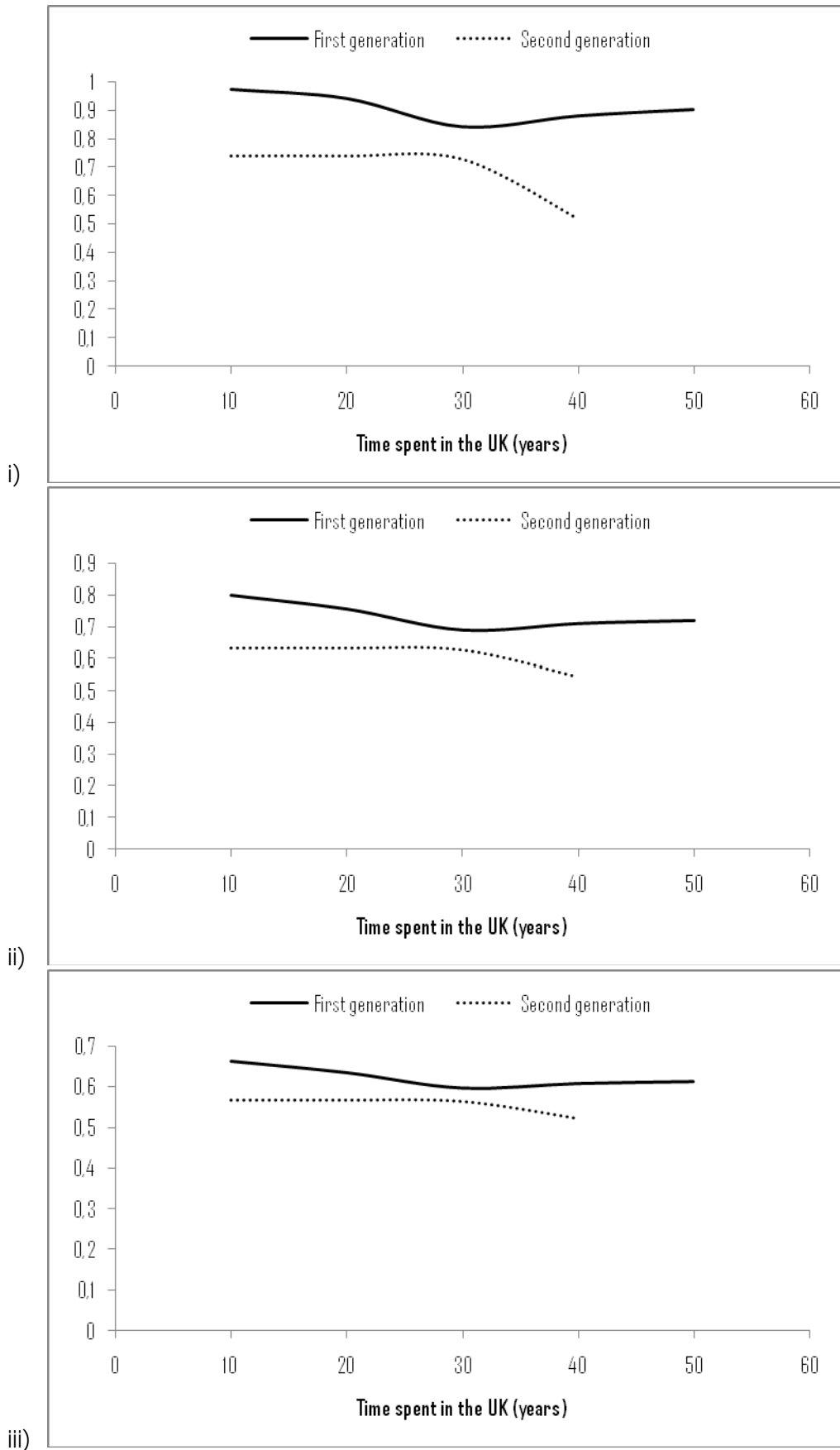
**Figure 6b Predicted Homogamy as a function of time spent in the UK- Muslim sub-**



**Figure 7a Predicted Identity as a function of time spent in the UK - i) at minimum q, ii) at average q, iii) at maximal q –Muslim sub-sample**



**Figure 7b Predicted Homogamy as a function of time spent in the UK – i) at minimum q, ii) at average q, iii) at maximal q- Muslim sub-sample**



## LIST OF TABLES

**Table 1: Ethnic Identity, Homogamy and Ethnic Neighborhood Composition**

	(1): Dep. Var.: Ethnic identity measured by importance of religion		(2): Dep. Var.: Homogamy			
	(1)	(2)	(1)	(2)	(1)	(2)
Ward density of own ethnic groups	0.0461*** (0.0156)	0.1081*** (0.0265)	0.0436*** (0.0166)	0.1355*** (0.0312)	0.0292* (0.0176)	0.1112*** (0.0339)
Ward density of own ethnic groups^2	-0.0013*** (0.0004)	-0.0019** (0.0008)	-0.0014*** (0.0005)	-0.0027*** (0.0010)	-0.0011** (0.0005)	-0.0026** (0.0011)
Age at arrival			0.0055 (0.0049)	0.0288*** (0.0099)	0.0066 (0.0049)	0.0287*** (0.0102)
Age			0.4078* (0.2396)	0.0062 (0.1792)	0.3926* (0.2311)	0.0973 (0.1781)
Age^2			-0.0069* (0.0040)	-0.0008 (0.0027)	-0.0066* (0.0039)	-0.0021 (0.0026)
Female			0.0913 (0.0882)	0.0272 (0.1411)	0.0950 (0.0896)	0.1132 (0.1524)
Born in the UK			-6.5322* (3.5163)	-0.2273 (2.9235)	-6.4069* (3.3959)	-1.7753 (2.9432)
Arranged marriage			0.4712*** (0.0960)	1.0410*** (0.3494)	0.3992*** (0.0982)	0.9717*** (0.3413)
Discrimination			0.1050 (0.1312)	-0.0803 (0.1874)	0.0733 (0.1310)	-0.1211 (0.1877)
Children			0.0562 (0.1537)	0.0008 (0.2007)	0.0434 (0.1585)	0.0429 (0.2036)
Years since arrival			-0.0294 (0.0196)	0.0486 (0.0363)	-0.0245 (0.0203)	0.0295 (0.0343)
Years since arrival^2			0.0005 (0.0004)	-0.0014* (0.0007)	0.0003 (0.0004)	-0.0009 (0.0007)
British degree			0.0550 (0.1133)	-0.6510*** (0.1475)	0.1555 (0.1139)	-0.6238*** (0.1546)
British high education			-0.4770*** (0.1285)	-0.0833 (0.1612)	-0.3747*** (0.1315)	-0.1422 (0.1669)
Foreign education			-0.1913* (0.0978)	0.0639 (0.1717)	-0.1211 (0.0995)	-0.0479 (0.1739)
Employed			-0.2961*** (0.0999)	0.1892 (0.1572)	-0.2842*** (0.1106)	0.1101 (0.1746)
Manager			-0.1726 (0.2373)	-0.0637 (0.2829)	-0.1385 (0.2312)	0.0278 (0.2953)
Employee			-0.0834 (0.0933)	-0.3205* (0.1695)	-0.0792 (0.0952)	-0.3517** (0.1773)
House owner			-0.0946 (0.1093)	0.5326*** (0.1577)	0.0049 (0.1181)	0.4821*** (0.1707)
English spoken at home					-0.2996** (0.1360)	-0.6528*** (0.1849)
English spoken at work					0.0745 (0.1232)	0.3815** (0.1904)
English spoken with friends					-0.1136 (0.1102)	0.3006* (0.1772)
Ward density of all ethnic groups					-0.0019 (0.0034)	0.0171*** (0.0063)
Ward unemployment rate					0.0438*** (0.0118)	-0.0171 (0.0191)
South					-0.1248 (0.0948)	-0.0089 (0.1375)
Observations	1559		1559		1559	
Likelihood	-1342.0755		-1152.1474		-1115.9003	

Notes. Bivariate probit model estimation results. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 2: Ethnic Identity, Homogamy and Ethnic Neighborhood Composition**

(1): Dep. Var.: Ethnic identity measured by ethnic group identification (2): Dep. Var.: Homogamy						
	(1)	(2)	(1)	(2)	(1)	(2)
Ward density of own ethnic groups	0.0151 (0.0210)	0.1069*** (0.0268)	0.0166 (0.0205)	0.1310*** (0.0315)	0.0138 (0.0220)	0.1131*** (0.0344)
Ward density of own ethnic groups^2	-0.0004 (0.0006)	-0.0019** (0.0008)	-0.0005 (0.0006)	-0.0026*** (0.0010)	-0.0002 (0.0006)	-0.0026** (0.0011)
(1): Dep. Var.: Ethnic identity measured by school ethnic composition (2): Dep. Var.: Homogamy						
	(1)	(2)	(1)	(2)	(1)	(2)
Ward density of own ethnic groups	0.0419*** (0.0157)	0.1120*** (0.0268)	0.0431*** (0.0161)	0.1393*** (0.0313)	0.0356** (0.0172)	0.1156*** (0.0342)
Ward density of own ethnic groups^2	-0.0011** (0.0005)	-0.0020** (0.0008)	-0.0011** (0.0005)	-0.0028*** (0.0010)	-0.0011** (0.0005)	-0.0028*** (0.0011)
(1): Dep. Var.: Ethnic identity measured by inter-ethnic marriage (2): Dep. Var.: Homogamy						
	(1)	(2)	(1)	(2)	(1)	(2)
Ward density of own ethnic groups	0.0420*** (0.01562)	0.0999*** (0.0268)	0.0437** (0.0175)	0.1247*** (0.0316)	0.0346* (0.0187)	0.1027*** (0.0346)
Ward density of own ethnic groups^2	-0.0007 (0.0005)	-0.0016* (0.0008)	-0.0010** (0.0005)	-0.0023** (0.0010)	-0.0008 (0.0005)	-0.0024** (0.0011)
Control set for all specifications						
Individual variables	no	no	yes	yes	yes	yes
Contextual variables	no	no	no	no	yes	yes

Notes. Bivariate probit model estimation results. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 3a: Maximum likelihood results**  
**-Structural approach-**  
(1) Unconstrained model, (2) Cultural defence model

	(1)	(2)	(1)	(2)	(1)	(2)
c	0.0658*** (0.0136)	0.5408* (0.3001)	0.9058*** (0.1079)	0.4542*** (0.1867)	0.8033*** (0.1176)	0.5890*** (0.0910)
alpha	-5.4493 (3.0135)	2.0174 (7.3598)	2.1743*** (0.4494)	1.0347 (0.8250)	1.2111*** (0.2114)	0.9553*** (0.2454)
gamma1	13.9730*** (1.9789)	-4.9059 (21.1685)	2.9126*** (0.3914)	-2.0702 (1.5951)	-2.3301*** (0.2691)	-1.8608*** (0.3686)
gamma2	0.2819 *** (0.0462)	-	0.0156 (0.0253)	-	0.3900*** (0.1159)	-
Control set						
Individual variables	no	no	yes	yes	yes	yes
Contextual variables	no	no	no	no	yes	yes
<b>Constrained likelihood</b>						
<i>Cultural defense model</i>		-2031.612		-1970.366		-1955.342
<i>Cultural conformity model</i>		-2033.426		-2019.004		-2027.065

**Table 3b: Maximum likelihood results**  
**-Structural approach-**  
(1) Unconstrained model, (2) Cultural defence model

	(1)	(2)	(1)	(2)	(1)	(2)
c	3.6349*** (0.0604)	10.5065 (12.0523)	-2.2563 (1.3993)	0.5377 (0.3007)	0.6768*** 0.0145	0.6905*** 0.0981
alpha	2.4432*** (0.0114)	8.5981 (218.3171)	0.9979*** (0.0042)	2.2542** (1.1085)	1.7376*** (0.0211)	1.5077*** (0.1128)
gamma1	-0.1224** (0.0609)	1.6419 (18.5896)	1.9496** 0.7959	0.7528* (0.3402)	0.5209*** (0.0109)	1.5581*** 0.0461
gamma2	0.9123*** (0.0479)	-	-0.0015 0.0031	-	1.5025*** (0.0233)	-
Control set						
Individual variables	no	no	yes	yes	yes	yes
Contextual variables	no	no	no	no	yes	yes
<b>Constrained likelihood</b>						
<i>Cultural defense model</i>		-8564.976		-9830.328		-9938.35
<i>Cultural conformity model</i>		-12268.60		-10096.61		-13160.74

**Table 4: Distribution of people with exogenous q by neighbourhood type and ethnic identity**

q	1.99	3.5	7.5	12.5
Residents	0.63	0.17	0.11	0.04
of which				
High identity residents	0.37	0.47	0.40	0.25

**Table 5a: Maximum likelihood results****-Structural approach-**

(1) Unconstrained model, (2) Cultural defence model

MUSLIM SUB-SAMPLE

	(1)	(2)	(1)	(2)	(1)	(2)
c	0.3211 (0.3038)	0.1509 (0.2603)	0.1905*** (0.0237)	0.0831*** (0.0084)	0.9416 *** (0.1978)	0.6750*** (0.0957)
alpha	5.3463 (54.6523)	1.6258 (6.3470)	2.3255*** (0.4012)	2.3991*** (0.3192)	1.3533*** (0.3565)	1.3666*** (0.2933)
gamma1	-14.6787 (78.6984)	-15.2402 (31.0822)	1.6581*** (0.1827)	1.6810*** (0.4018)	-2.2249*** (0.4624)	-1.5582*** (0.5604)
gamma2	-0.00179 (1.4587)	-	0.4935*** (0.0308)	-	0.1522** (0.0696)	-
Control set						
Individual variables	no	no	yes	yes	yes	yes
Contextual variables	no	no	no	no	yes	yes
<b>Constrained likelihood</b>						
<i>Cultural defense model</i>		-811.745		-794.6653		-791.9213
<i>Cultural conformity model</i>		-823.684		-832.4669		-815.3603

**Table 5b: Maximum likelihood results****-Structural approach-**

(1) Unconstrained model, (2) Cultural defence model

MUSLIM SUB-SAMPLE

	(1)	(2)	(1)	(2)	(1)	(2)
c	4.43097*** (0.5091)	4.6382 (92.4344)	-1.2623*** (0.0439)	1.4321*** (0.1990)	0.2976*** (0.0751)	0.9186*** (0.1781)
alpha	16.3128*** (0.9763)	23.7373 (13483.0034)	3.1102*** (0.1262)	2.3504*** (0.5479)	3.8396*** (0.0492)	1.3373*** (0.2689)
gamma1	2.0658*** (0.5572)	-18.4402 (9812.577)	2.0234*** (0.0562)	0.3239 (0.3194)	0.3921*** (0.0408)	1.8843*** (0.0100)
gamma2	-2.5708*** (0.5727)	-	-1.4948*** (0.06141)	-	1.6885*** (0.0899)	-
Control set						
Individual variables	no	no	yes	yes	yes	yes
Contextual variables	no	no	no	no	yes	yes
<b>Constrained likelihood</b>						
<i>Cultural defense model</i>		-8104.676		-2854.586		-2882.496
<i>Cultural conformity model</i>		-11050.31		-3231.863		-3462.099



**DATA APPENDIX**  
**Table A1: Description of data**

		n.obs: 1,559	
<i>Variable</i>	<i>Explanation of the variable</i>	<i>Mean</i>	<i>St.dev.</i>
<i>Main variables</i>			
Importance of religion	In the text	0.61	0.49
Homogamy	In the text	0.92	0.27
Ward density of own ethnicity	In the text	13.60	10.76
<i>Individual controls</i>			
Age at arrival	Respondent's age in years at arrival in the UK	21.15	10.38
Age	Respondent's age in years	41.37	13.09
Female	Dummy variable taking value one if the respondent is female.	0.47	0.50
Born in the UK	Dummy variable taking value one if the respondent is born in the UK	0.10	0.30
Arranged Marriage	Dummy variable taking value one if the husband/wife of the respondent has been chosen by the parents.	0.37	0.48
Discrimination	Dummy variable taking value one if the respondent had been insulted or threatened in the last year for reasons to do with race or colour.	0.10	0.30
Children	Dummy variable taking value one if the respondent has children.	0.91	0.28
Years since arrival	Number of years since respondent's arrival in the UK.	20.22	11.42
British degree	Dummy variable taking value one if the respondent has a UK degree.	0.21	0.41
British high education	Dummy variable taking value one if the respondent has a UK O-level (or equivalent) or above qualification.	0.16	0.37
Foreign education	Dummy variable taking value one if the respondent has a qualification achieved abroad.	0.28	0.45
Employed	Dummy variable taking value one if the respondent is employed.	0.47	0.50
Manager	Dummy variable taking value one if the respondent is a manager.	0.03	0.16
Employee	Dummy variable taking value one if the respondent is an employee	0.57	0.50
House owner	Dummy variable taking value one if the household owns (or is buying) the accommodation	0.76	0.43
<i>Contextual controls</i>			
English spoken at home	Dummy variable taking value one if English is the language normally spoken at home with family members (who are older) by the respondent.	0.12	0.33
English spoken with friends	Dummy variable taking value one if English is the language normally spoken with friends (outside work) by the respondent.	0.51	0.50
English spoken at work	Dummy variable taking value one if English is the language normally spoken at work by the respondent.	0.48	0.50
Ward density of all ethnic groups	Percentage of residents of any ethnic group in the ward	33.63	21.01
Ward unemployment rate	Ward unemployment rate	14.38	5.24

