DISCUSSION PAPER SERIES

No. 8054

BEND IT LIKE BECKHAM: ETHNIC IDENTITY AND INTEGRATION

Alberto Bisin, Eleonora Patacchini, Thierry Verdier and Yves Zenou

PUBLIC POLICY

CE PR

Centre for Economic Policy Research

www.cepr.org

BEND IT LIKE BECKHAM: ETHNIC IDENTITY AND INTEGRATION

Alberto Bisin, New York University and NBER
Eleonora Patacchini, Università di Roma "La Sapienza" and CEPR
Thierry Verdier, Paris School of Economics and CEPR
Yves Zenou, Stockholm University, Research Institute of Industrial Economics and CEPR

This Discussion Paper is issued under the auspices of the Centre's research programme in **PUBLIC POLICY**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

ABSTRACT

Bend It Like Beckham: Ethnic Identity and Integration*

We propose a theoretical framework to study the determinants of ethnic and religious identity along two distinct motivational processes which have been proposed in the social sciences: cultural conformity and cultural distinction. Under cultural conformity, ethnic identity is reduced by neighborhood integration, which weakens group loyalties and prejudices. On the contrary, under cultural distinction, ethnic minorities are more motivated in retaining their own distinctive cultural heritage the more integrated are the neighborhoods where they reside and work. Data on ethnic preferences and attitudes provided by the Fourth National Survey of Ethnic Minorities in the UK enables us to test the relative significance of these two identity processes. We find evidence consistent with intense ethnic and religious identity mostly formed as a cultural distinction mechanism. Consistently, we document that ethnic identities are more intense in mixed than in segregated neighborhoods.

Alberto Bisin
Department of Economics
New York University
Room 705
269 Mercer Street
New York, NY 10003
USA

Email:
alberto.bisin@nyu.edu

Eleonora Patacchini
Faculty of Statistics
Universita di Roma 'La Sapienza'
Box 83 - RM62
P.le Aldo Moro 5
00185 Roma
ITALY

Email:
eleonora.patacchini@uniroma1.it

Thierry Verdier
Paris-Jourdan Sciences Economiques
48 Boulevard Jourdan
75014 Paris
FRANCE

Email: verdier@pse.ens.fr

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=113317

Yves Zenou
Department of Economics
Stockholm University
10691 Stockholm
SWEDEN

Email: yves.zenou@ne.su.se

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=126027

Submitted 2 October 2010

Bengali, bengali / Bengali, bengali / No no no / He does not want to depress you/ Oh no 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][1]

1   Introduction

In the last decades, immigration into western countries has become an important facet of globalization. This phenomenon has induced increased concerns regarding the implications of non-economic issues such as the rise of ethnic diversity in the host countries.[2] While the diversity of social groups can be considered as a source of benefits, the persistence of ethnic minority' identities is, however, often perceived as a threat or source of frictions by natives (see Alba, 1990, 2005). This is well illustrated, for instance, by the recent passionate debate around the wearing of the Islamic burqa in France. As Putnam (2007) put it, in his John Skytte Price Lecture, "*the increase in ethnic and social heterogeneity in virtually all advanced countries is one of the most important challenges facing modern societies, and at the same time one of our most significant opportunities*".

It is therefore of paramount importance to have a precise understanding of the mechanism of identity formation and its determinants. In fact, two opposing views characterize the theoretical analysis in the social sciences regarding identity formation.[3] A first group of social scientists argue that ethnic identity is reduced by assimilation and by the blurring of groups' boundaries. *Assimilation theories*, in political science and sociology (Gordon, 1964; Moghaddam and Solliday 1991) and *contact theory* in social psychology (Allport, 1954) are the prominent manifestations of this line of thought. The basic premise is the idea that social contacts between natives and ethnic minorities 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 principle that group identity is driven by a motive for inclusiveness and *cultural conformity*, positive social interactions across individuals sharing the same characteristics, views and preferences.[4]

The alternative view represents ethnic minorities as motivated in keeping their own distinctive cultural heritage, in identifying themselves with an ethnic/social group to enhance their psychological self-esteem and generate a sense of positive *distinctiveness* from individuals who are part of that group (Abrams and Hogg, 1988; Turner 1982). Various cognitive and psychological mechanisms, from group solidarity to prejudice and negative stereotypes with respect to other groups, are deemed responsible in achieving such positive distinctiveness. Negative attitudes towards members of other groups, in turn,

---

[1] Thanks to Andrew Clark for Morissey's quote.

[2] Alesina and La Ferrara (2005) provide a general discussion of the economic effects of increased ethnic diversity.

[3] The study of ethnic identity formation has a long theoretical and empirical tradition in social sciences with 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.

[4] See Bernheim (1994) for a formal economic analysis of conformity.

consolidate the social identity of the group. These ideas compose the core of theories of *multiculturalism* (Glazer and Moynihan, 1970; Taylor and Lambert, 1996), and *conflict* (Bobo, 1999).[5] According to this view, group identity formation is a sort of *cultural distinction* mechanism that allows individuals to reduce the psychological costs associated with 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 framework on these issues, we propose a simple model of identity formation that accounts for both *cultural conformity* and *cultural distinction*. We can therefore formally analyze the implications of each of these opposite views of identity formation and pursue an empirical investigation guided by the theoretical analysis. We show that *cultural conformity* and *cultural distinction* provide 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, neighborhood segregation and identity formation are likely to be complements for ethnic assimilation. On the contrary, when *cultural distinction* is at work, neighborhood segregation and identity formation tend to be substitutes for ethnic assimilation.[6] 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. In that case, neighborhood mixing policies would not necessarily favor cultural integration, contrary to presumptions often exposed by social scientists and commentators.

The different implications of *cultural conformity* and *distinction* can be tested empirically and can therefore inform us on the main process of ethnic identity formation. To address more directly the issue of what motivates identity formation, 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, more importantly, its ethnic and religious composition. Furthermore, we link identity formation with homogamous marriages along ethnic and religious traits. It has indeed been extensively documented 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 trait (Bisin, Topa, and Verdier, 2004, and the evidence cited in Bisin and Verdier, 2000).

To be more precise, our analysis exploits a unique UK dataset, the Fourth National Survey of Ethnic

---

[5] At a broader level, this view is also related to the *social identity theory* in social psychology (Tajfel, 1981; Turner, 1982).

[6] In economics, the distinction between *cultural conformity* and *cultural distinction* is also related to the notion of cultural complementarity and cultural substitutability between socialization mechanisms. This has been defined formally by Bisin and Verdier (2000). Indeed, in Bisin and Verdier (2000), when family and role models tend to be substitutes in the process of socialization, families with a relatively minoritarian cultural trait have larger incentives to spend resources socializing their children to their trait in order to guarantee its persistence. Conversely, under cultural complementarity, the more minoritarian is a family's cultural trait, the lower are the family's incentives to socialize their children to the trait and hence to limit cultural assimilation.

Minorities (FNSEM). This survey, which over-samples ethnic minority groups, asks a direct question about the respondents' identification with their own ethnic group and provides 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 survey data can be merged with the 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 the 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 neighborhoods.

Our evidence for *cultural distinction* is consistent with several empirical studies studying the link between identity and segregation. Fryer and Torelli (2010), on data from the National Longitudinal Study of Adolescent Health, find that "acting white" behavior among blacks is more developed in racially mixed schools.[7] Also, Bisin, Topa, and Verdier (2004) document that, in General Social Survey data, religious socialization across U.S. states is more intense when a religious faith is in minority.[8] Finally, Munshi and Wilson (2008) combine data from the U.S. census and the National Longitudinal Survey of Youth to identify a negative relationship across counties in the Midwest of the United States between ethnic fractionalization in 1860 and the probability that individuals have professional jobs or migrated out of the county by 2000.

We also address the issue of the alleged specificity of Muslim immigrants with regard to the strength of their identity and their (refusal of) integration; an issue which has recently surged at the center of the political debate in Europe (see, e.g., Gallis, 2005). To this end we repeat our analysis on the restricted sample of Muslim respondents only. 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 than for the other ethnic 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, 2010) consider identity forma-

---

[7]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).

[8]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.

tion as an explicit − more or less conscious − endogenous choice by individuals exposed to 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 underline the endogeneity of racial identity. Chiswick (2009) 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, et al. (2008b) 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) analyze ethnic identity as the endogenous balance between commitment to and self-identification with the culture and society of the origin and the host country. Battu and Zenou (2010), Nekby and Rödin (2010), and Manning and Roy (2010) find that language proficiency in the host country, isolation and segregation, years spent in the host country, intermarriage, and education acquired before immigration are significantly correlated with ethnic identity.

Like us, this emerging literature recognizes the endogeneity and contextual character of ethnic and cultural identity formation. Our contribution is, however, to try and open the "black box" of identity formation, to disentangle and identify in a specific empirical context the precise mechanisms of identity formation (i.e., *cultural distinction* versus *cultural conformity*), which have distinct implications for cultural integration processes across social groups.

Before proceeding, we should briefly alert the reader of the methodological choice of the present paper. We proceed in steps, from a non-structural analysis of the data to a fully structured model estimation. By doing so, we make apparent that our results are only in part determined by the structural (functional-form) identification obtained by putting directly to data the theoretical model. On the contrary, our structural results are suggested by more flexible statistical approaches to the data. More precisely, we start in Section 2 with a non structural probit analysis of identity and homogamy in terms of ethnic composition. Section 3 then proposes a semi-structural analysis of homogamy and identity. While disregarding some cross-equation restrictions as imposed in a fully structural approach, thus approach has the ability to address the issue of *cultural distinction* versus *cultural conformity*. Section 4 goes further by providing and estimating a fully structural model of ethnic integration. Section 5 focuses on Muslims only. Finally, Section 6 concludes and discusses some policy implications.

2 Descriptive analysis of the data

The Fourth National Survey of Ethnic Minorities (FNSEM) was collected in 1993/94 in the U.K. by the Policy Studies Institute (PSI). It over-samples ethnic minority groups, distinguishing explicitly six of them: Caribbean, Indian, Pakistani, African-Asian, Bangladeshi, and Chinese.[9] The survey contains

---

[9]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.

detailed information about the 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 can in fact be considered an "objective" measure of identity, which is conceptually strongly related to our subjective measure.

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. To this end, we merge the FNSEM data with the 1991 Census in order to get valuable information of each individual's residential ward.[10]

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 ethnic identity, ν; and *(iii)*, the probability of homogamous marriage, π. 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 the same 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. It should be clear that variation in q does not proxy for ethnic group; that is, it is not the case that, for example, 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, the importance of religion, the attitudes towards inter-marriage, and the relevance of ethnicity in influencing the kind of school people want for their children.[11] It also asks a direct question about ethnic identity.[12] We perform our analysis using separately the answers on each

---

[10] A UK Census ward contains on average 3,000-4,000 residents.

[11] 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 (respondenr'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?" .

[12] 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]".

6

of these questions. Identity, denoted by I, is coded as a dichotomous variable taking value 1 if the individual considers as 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 takes 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*). 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, ν, is then the probability that $I = 1$, for each aspect of these different measures of identity.

(iii) Homogamy H is a dummy variable taking value 1 if the respondent is married to a person of her/his own ethnic group, and 0 otherwise. The variable π 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 or not), the data set also contains control 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 1,559 individuals.

2.2 Identity and homogamy: A probit

Empirically identifying cultural distinction from cultural transmission would require observing a negative relationship between ethnic identity (resp. homogamy) and the ethnic composition of the neighborhood as measured by q. A negative sign would indicate cultural distinction while a positive sign would mean cultural conformity. Naturally, a regression analysis is difficult to interpret because of the endogeneity of q: the ethnic composition of the neighborhood might, in principle, be an important factor affecting the respondents' residential decisions. In the course of this paper, we shall deal with this endogeneity problem, both directly and indirectly, by imposing more structure on the empirical analysis. It is nonetheless interesting to present the results of a simple probit regression, looking at the correlation between ethnic identity (measured either directly by I or indirectly by homogamy H) and ethnic composition (measured by q), allowing for a non-linear relationship. We perform such a bivariate probit analysis with different (increasing) sets of control variables and using alternative definitions of ethnic identity. Table 1 shows the complete list of estimation results when "importance of religion" is used as a proxy for ethnic identity, whereas Table 2 displays the same results when the three other measures of ethnic identity (*ethnic group identification, school ethnic composition,*

inter-ethnic marriage) are used.[13]

[*Insert Tables* 1 *and* 2 *here*]

We find that there is a positive and significant relationship between ethnic identity and ethnic neighborhood composition q. We also find significant non-linearities, i.e., the quadratic term in q is negative and significant. In particular, both identity and homogamy appear to be negatively related to ethnic composition for values of q greater than 20%. This suggests a *cultural distinction* mechanism. Figure 2a confirms this result by depicting 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.[14] The non-linearities picked up by the probit models might, in principle, be due to the differential distribution of ethnic groups by neighborhood class.[15] This is not the case, in fact, since the same form of non-linear dependence in q is obtained when the sample is restricted to Muslims only (mostly Pakistanis, Bangladeshis, and Indians), as documented in Figure 2b.[16] These non-linearities suggest that ethnic identity is weak in relatively homogenous neighborhoods and strong in mixed neighborhoods.

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

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

3   Conformity versus distinction: Some structure

Consider a member of an ethnic or religious group.[17]   As in the previous section, let q denote the proportion of ethnic minorities in the neighborhood where this ethnic person resides. Let homogamy be an index $H \in \{0,1\}$, with $\Pr\{H = 1\} = \pi$ (i.e. $H = 1$ means that the ethnic minority is married to someone from the same ethnic group while $H = 0$ indicates heterogamy). There is a psychological cost of interacting with individuals from the majority (dominant) group. We assume

---

[13]In the probit estimations, we measure q as the fraction of *own ethnic group* in the neighborhood. The fraction of *all minority residents* in the neighborhood has been, however, included among the controls.

[14]This graph is depicted using "importance of religion" as a proxy for ethnic identity. The use of the other proxies leads to similar pictures.

[15]We thank Bill Easterly for raising this point in a critical discussion of the paper.

[16]We have also performed our analysis using a 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 Crombach-α measure is then used to assess the quality of the derived index. In our case, we obtain an α equal to 0.86 ($0 \le \alpha \le 1$) indicating that the different items incorporated in the index have considerable internal consistency. We find that even this aggregate measure do not depend (qualitatively) differently on q.

[17]We only consider here two groups: the ethnic minority and the majority groups. Our theoretical analysis can easily be extended to more than one ethnic group.

8

that this (psychological) cost depends on the marriage status of the minority member and is denoted by $C(H)$. We further assume that such costs are lower in an homogamous than in an heterogamous marriage, i.e.

$$\Delta C = C(0) - C(1) > 0 \tag{1}$$

Indeed, all ethnic minorities need to interact with natives, for example when their kids go to the same school and have the same social activities, or when they need to find a job. The *unit* cost of interacting with natives is the same for all ethnic minorities. However, we postulate that, when a minority individual gets married with someone from the same ethnic group, he/she does not need to socially interact as much with natives (he/she puts children in an "ethnic" school, works with people from the same ethnic group, etc.). As a result, the *total cost* of socially interacting with natives is lower for an individual in an homogamous marriage than for someone who is married to a native (heterogamy) since the latter spends more time with natives (he/she needs to meet the parents in law, to put his/her kids in the majority school, etc.). In our framework, C denotes the total cost of interacting with a native and it satisfies (1), that is $C(0) > C(1)$.

In general, ΔC is also a function of q as well as of the strength of identity of the ethnic minority; we denote this strength of identity by an index $I \in \{0, 1\}$ with $\Pr\{I = 1\} = \nu$.[18] Indeed, the composition of the neighborhood as well as the strength of identity have an obvious impact on this cost differential since families living in more "ethnic" neighborhoods and with stronger ethnic identities are more ensured that their ethnic identity will be passed on to their kids. We thus have $\Delta C = \Delta C(q, I)$.[19] It is then straightforward to formulate a precise definition of cultural conformity and distinction:

The preferences of an ethnic minority individual display

cultural conformity *if the differential cost ΔC decreases with the proportion of non-minority members $1 - q$, i.e.$\frac{\partial \Delta C(q,I)}{\partial q} > 0$;*

cultural distinction *if the differential cost ΔC increases with the proportion of non-minority members $1 - q$, i.e.$\frac{\partial \Delta C(q,I)}{\partial q} < 0$.*

In the cultural conformity assumption, the minority's psychological costs of interacting with individuals from the majority group are *decreasing* in the proportion of whites living in the neighborhood where the minority resides. In the cultural distinction assumption, we have the opposite.[20] To be more precise, when minorities are more exposed to the majority group (i.e., when $1 - q$ increases), the difference in interaction costs with whites between an homogamous and heterogamous minority person is reduced with cultural conformity. Indeed, in that case, minorities tend to assimilate to the majority norm and their marital status have less impact on interaction costs. However, when we consider cultural distinction, this cost differential tends to increase because minorities are now rejecting the "white" norm and homogamous minorities interact much less with the majority group.

---

[18] C is therefore also a function of both q and I.

[19] Remember that in this section both q and I are exogenous. We will relax these assumptions below.

[20] See our discussion in the Introduction.

Minority members put effort in finding a spouse of the same ethnic background. Let this effort be a continuous variable denoted by $\tau \in [0,1]$. The minority member first searches a spouse in a restricted pool of partners from his/her own community minority. The search intensity, τ, determines the probability with which he/she finds his/her marital partner in the pool. With the residual probability $1-\tau$, he/she remains unsuccessful and therefore goes to a common pool of partners that includes both minority and majority types. There, he/she gets matched with a spouse of his/her community with probability q. As a result, an ethnic minority individual living in a neighborhood with a fraction q of minority members has a probability of marrying homogamously equals to

$$\pi(\tau, q) = \tau + (1 - \tau)q \tag{2}$$

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

$$Z(\tau) = \frac{1}{2}\alpha\tau^2 \tag{3}$$

where α is a measure of the relative cost of τ.

A minority member's problem is thus:[21]

$$\max_{\tau \in [0,1]} \left\{ -\pi(\tau, q)C(1) - [1 - \pi(\tau, q)]C(0) - \frac{1}{2}\alpha\tau^2 \right\}$$

In other words, each minority individual chooses the homogamy effort τ that minimizes the expected cost of interacting with whites, i.e.

$$\min_{\tau \in [0,1]} \left\{ -\pi(\tau, q)\Delta C(q, I) + C(0) + \frac{1}{2}\alpha\tau^2 \right\}$$

Using (2), the first order condition leads to:

$$\tau^* = \frac{(1 - q)\Delta C(q, I)}{\alpha} \tag{4}$$

The probability of an homogamous marriage is thus equal to

$$\pi^* = q + \frac{1}{\alpha}(1 - q)^2 \Delta C(q, I) \tag{5}$$

Observe that if there is cultural distinction, then $\partial\tau^*/\partial q < 0$, i.e., the higher is the proportion of the same ethnic group living in the neighborhood, the lower is the homogamy effort while the opposite sign is true ($\partial\tau^*/\partial q > 0$) with cultural conformity if the elasticity of q with respect to ΔC is large enough. However, the sign of $\partial\pi^*/\partial q$ tends to be ambiguous with either cultural distinction or cultural conformity. This is what we want to investigate empirically now.

---

[21] For notational simplicity, we don't put q and I as arguments of the cost function $C(.)$.

3.1 Empirical implementation

Based on this simple model, we would like to test equation (5), i.e. the relationship between π (homogamy) and q (neighborhood ethnic composition). To be able to empirically implement the model, we introduce x, a vector of exogenous variables (like e.g. age, education, etc.) which affects both q (neighborhood composition) and ν (identity strength), and the cost $Z(\tau)$. The relationship between (q, ν) and x can be abstractly, and without loss of generality,[22] represented by a map

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

Concerning the cost $Z(\tau)$, we assume that:

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

where $\gamma_\tau x = \sum_{j=1}^K \gamma_{\tau j} x_j$. Rewriting the first-order condition (5) with this extension in terms of x leads to:

$$\pi^* = q + \frac{1}{\alpha}(1-q)^2 \Delta C(q, I) - \frac{1}{\alpha}(1-q)(\gamma_\tau x) \tag{6}$$

The empirical implementation of this equation requires therefore fitting

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

with appropriate approximating polynomials.

Two important caveats are, however, in order. First of all, the function $F(x)$ is an equilibrium mapping. It will depend on deep parameters that determine $\Delta C(q, I)$.[23] In other words, the model implies cross-equation restrictions between $\pi = q + \frac{1}{\alpha}(1-q)^2\Delta C(q,I) - \frac{1}{\alpha}(1-q)(\gamma_\tau x)$ 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 *cultural conformity* (i.e. $\frac{\partial \Delta C(q,I)}{\partial q} > 0$) against *cultural distinction* (i.e. $\frac{\partial \Delta C(q,I)}{\partial q} < 0$).

Second, the identity formation process of the members of a minority group depends on the cultural characteristics of the minority itself (e.g., his/her cultural distance to the majority), but also on the actions and predispositions of the majority (e.g., their racial attitudes). We do not have accurate data to be able to distinguish between these determinants of identity, even though x contains some controls of some relevance, e.g., a measure of the episodes of ethnic/religious harassment each respondent has been subject to.

---

[22] That is, under conditions sufficient to guarantee a unique choice of (q, ν) 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, and Topa (2009).

[23] It might be convenient in general to restrict the map F so that $F(x) = f(X_q, X_\nu)$, for $f : \mathbb{R}^2 \to [0,1]^2$ and $\begin{cases} X_q = \gamma_q x \\ X_\nu = \gamma_\nu x \end{cases}$.

Let us now describe more precisely our empirical strategy here. Let the deep parameters of the model be denoted by θ, a vector. Let θ be decomposed so that $\theta = [\theta_1, \theta_2, \theta_3]$. Without loss of generality, let's assume that $\Delta C(q, \nu)$ only depends on parameters $[\theta_1, \theta_2]$, (α, γ_τ) are part of θ_1, and that $\begin{pmatrix} q \\ \nu \end{pmatrix} = F(x)$ only depends on parameters $[\theta_2, \theta_3]$.

As stated above, the objective of our empirical work here is not to estimate the deep parameters of the model, θ but rather to test if the sign of $\frac{\partial \Delta C(q, \nu)}{\partial q}$ is negative (i.e. cultural distinction) or positive (i.e. cultural conformity). In that case, it is sufficient to estimate (6), approximating $\Delta C(q, \nu)$ as follows via a second order polynomial:

$$\Delta C(q, \nu) \approx \beta_1 + \beta_2 q + \beta_3 q^2 + \beta_4 \nu + \beta_5 \nu^2 + \beta_6 q\nu + \beta_7 q^2 \nu + \beta_8 q\nu^2 + \beta_9 q^2 \nu^2$$

3.2 Results

Results of the empirical analysis are reported in Figures 3a and 3b (for the whole sample and the Muslim sample only, respectively). For both samples, the partial derivative of the interaction costs with respect to q is always negative,

$$\frac{\partial \Delta C(q, I)}{\partial q} < 0, \text{ for both } I = 0 \text{ and } I = 1.$$

This result is clearly evidence in favor of the *cultural distinction* mechanism: integration, as manifested by a relatively high q, leads to significant difference in interaction costs between homogamous and heterogamous individuals.

[*Insert Figures* 3a *and* 3b *here*]

4 Ethnic integration: Structural models

We now complement the analysis of the previous section by structurally studying ethnic integration via marriage, location, and identity formation. Formally, this requires developing a model for $\Delta C(q, I)$ so that q and I are chosen and are not exogenously given (as in the previous section). In fact, we will study two distinct models, one in which location is exogenous and one in which identity is exogenous. To be more precise, we develop two different, extreme (semi-nested) models:

Marriage and identity (with exogenous ethnic composition of the location, q): people are dropped in a neighborhood and then form their ethnic identity (choose ν) and look for a spouse (choose τ).

Marriage and location (with exogenous identity I): people are born with an identity and then look for a neighborhood where to reside (choose q) and a spouse (choose τ).

4.1 Marriage and identity

As in Section 3.1, we define the psychological cost of interacting with individuals from the majority group by $C(H, I, q)$. In other words, this cost for an ethnic minority depends on his/her marriage status H $(= 1$ if homogamous and 0 otherwise), his/her identity status I $(= 1$ if having a strong identity and 0 otherwise) and q, the percentage ethnic minorities where he/she resides. We have $(H, I) \in \{0, 1\}^2$ and $q \in [0, 1]$. Observe that the probability $\nu = \Pr\{I = 1\}$ is here modeled as a choice of the individual. In terms of notation, we use $\nu \in [0, 1]$ rather than I. We assume that identity and homogamy act as complements to each other so that either $\nu = \Pr\{I = 1\}$ *and* $H = 1$ or $1 - \nu = \Pr\{I = 0\}$ *and* $H = 0$. In other words, individuals with strong identity are necessary married to someone from the same ethnic group $(H = 1)$ while people with weak identity are necessary married to whites $(H = 0)$. For simplicity, we use the following explicit function:

$$C(H, \nu, q) = (1 - \nu H)C(q)$$

which implies that either $C(1, \nu, q) = (1 - \nu)C(q)$ or $C(0, \nu, q) = C(q)$. This means that a strong identity (ν) with an homogamous marriage $(H = 1)$ leads to a lower cost of interacting with whites than a weak identity with an heterogamous marriage. Our explanation is as before. Individuals married to someone from the same ethnic group and having a strong identity do not interact very much with whites and thus have a lower interacting cost than someone more "integrated" to the majority group. This specification implies that

$$\Delta C(q, \nu) = \nu C(q) \tag{7}$$

In this section, we assume that location q is exogenous and individuals choose their identity and homogamy efforts, ν and τ. The utility cost of developing identity ν is denoted by $J(\nu)$ and is increasing and convex (in the same units of the psychological costs $C(q)$). For simplicity, we assume

$$J(\nu) = \frac{1}{2}\nu^2 \tag{8}$$

As a result, a minority member's problem is:

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

Indeed, as before, the ethnic minority minimizes his/her expected cost of interacting with whites, that is:

$$\min_{\nu, \tau} \left\{ -\nu\pi(\tau, q)C(q) + C(q) + \frac{1}{2}\alpha\tau^2 + \frac{1}{2}\nu^2 \right\}$$

The first order conditions of this problem are given by:

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

$$\tau^* = \frac{q(1 - q)C^2(q)}{\alpha - (1 - q)^2 C^2(q)}$$

13

where α is assumed to be large enough for $\nu^* > 0$ and $\tau^* > 0$. These first order conditions can easily be reduced to:

$$\nu = \pi \, C(q) \tag{9}$$

$$\pi = q + (1-q)^2 \frac{C(q)}{\alpha} \nu \tag{10}$$

This is a simultaneous equation system in which (ν, π) are the endogenous variables and q the exogenous variable. We can explicitly calculating π^* by using (9). We obtain:

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

4.2 Marriage and location

In this section, we still assume (7). The choice of identity ν is now exogenous while location q is chosen. The cost of living in a neighborhood with a percentage q of ethnic minorities is denoted by $G(q)$ and is increasing and convex (in the same units of the psychological costs $C(q)$).[24] For simplicity, we assume

$$G(q) = \frac{1}{2}q^2 \tag{12}$$

The minority member's problem can now be written as:

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

As before, the ethnic minority minimizes his/her expected cost of interacting with whites, that is:

$$\min_{q,\tau} \left\{ -\nu\pi(\tau,q)C(q) + C(q) + \frac{1}{2}\alpha\tau^2 + \frac{1}{2}q^2 \right\}$$

The first order conditions of this problem are given by:

$$-\nu(1-\tau)C(q) + C'(q)\,(1-\nu\pi) + q = 0$$
$$-\nu(1-q)C(q) + \alpha\tau = 0$$

which, using (2), are equivalent to:

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

This is a simultaneous equation system in which (q, π) are the endogenous variables and ν the exogenous variable. The solution is rather involved. For our empirical application, it is convenient to write

---

[24] We assume that, other things equal, neighborhoods with higher q are less preferable (because, for example, 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.

the system in terms of ν and π (even if ν is exogenous here) as a function of π and q. After some algebra one obtains:

$$v^* = \alpha \frac{(\pi - q)}{(1 - q)^2 C(q)} \tag{13}$$

$$\pi^* = \frac{-B \pm \sqrt{B^2 - 4AD}}{2A} \tag{14}$$

where

$$A = \alpha \left[(1 - q) \, C'(q) - C(q) \right]$$

$$B = \alpha \left[C(q) \, (1 + q) - q \, (1 - q) \, C'(q) \right]$$

$$D = -\alpha q C(q) - \left[C'(q) + q \right] (1 - q)^3 C(q)$$

The system has two solutions, in principle.

4.3   Conformity versus distinction: A full structural analysis

In the *cultural conformity* model, the minority's psychological costs of interacting with individuals from the majority group are decreasing in the proportion $1 - q$ of whites residing in the neighborhood where the minority lives. The simplest formulation therefore has:

$$C(q) = c \, q \tag{15}$$

In the *cultural distinction* model the minority's psychological costs to interact with individuals from the majority group are increasing in the proportion $1 - q$ of whites living in the neighborhood where the minority resides. The simplest formulation thus has:

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

Given (7), this implies that $\Delta C(q, \nu) = \nu c q$ and thus $\partial \Delta C(q, \nu)/\partial > 0$ for the *cultural conformity* model, and $\Delta C(q, \nu) = \nu c (1 - q)$ and thus $\partial \Delta C(q, \nu)/\partial < 0$ for the *cultural distinction* model. This is exactly the definitions we gave in Section 3.

We can then plug the value of $C(q)$, given by either (15) or (16), into the marriage and identity model (Section 4.1) and the marriage and location model (Section 4.2). It is therefore crucial to distinguish between the cultural substitution and cultural distinction models. Formally, we can nest the reduced form equations by writing:

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

so that

$$\begin{array}{lll} \text{when} & \gamma_1 = \gamma_2 = 1 & \text{there is cultural distinction} \\ \text{when} & \gamma_1 = 0 \text{ and } \gamma_2 = -1 & \text{there is cultural conformity} \end{array} \tag{17}$$

Empirically, to discriminate between the two models, it is sufficient to verify whether γ_2 is greater or smaller than zero.

4.4 Empirical implementation

Both the marriage and identity, and the marriage and location models are identified exploiting the non-linearities induced by the choice problems and the functional form assumptions we impose. As in Section 3.1, for the empirical implementation of costs, we have to add the vector of individuals' characteristics x to obtain:

$$
\begin{aligned}
Z(\tau) &= \frac{1}{2}\alpha\tau^2 + (\gamma_\tau\, x_\tau)\ \tau \\
J(\nu) &= \frac{1}{2}\nu^2 + (\gamma_\nu\, x_\nu)\ \nu \\
G(q) &= \frac{1}{2}q^2 + (\gamma_q\, x_q)\ q
\end{aligned}
$$

where x_j $(j = \tau, \nu, q)$ are the exogenous determinants of the cost and regressors in the empirical implementation, and $\gamma_\tau\, x_\tau$, $\gamma_\nu\, x_\nu$, $\gamma_q\, x_q$ represent the vector notation: $\gamma_\tau\, x_\tau = \sum_{j=1}^{N} \gamma_{\tau j} x_j$, $\gamma_\nu\, x_\nu = \sum_{j=1}^{N} \gamma_{\nu j} x_j$, and $\gamma_q\, x_q = \sum_{j=1}^{N} \gamma_{qj} x_j$.

By using these cost formulations and solving each program, we can extend equations (9) and (10) and obtain a simultaneous equation system in ν and π for the marriage and identity model. Likewise, for the marriage and location model, we extend equations (13) and (14) and obtain a simultaneous equation system in q and π. The Appendix provides the details on the systems of equations we put to data (equations (26)-(27) and equations (28)-(29)) as well as simple informal argument for the identification of parameters $\gamma_1, \gamma_2, c, \alpha, \gamma_\nu, \gamma_\tau$.

Writing a likelihood function and searching for the structural parameters $(\gamma_1, \gamma_2, c, \alpha, \gamma_\nu, \gamma_\tau)$ is straightforward.[25] 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. We have:

$$
\Pr\{I_i = y_i\} = \begin{cases} \nu_i & y_i = 1 \\ 1 - \nu_i & y_i = 0 \end{cases}
\qquad \text{and} \qquad
\Pr\{H_i = z_i\} = \begin{cases} \pi_i & z_i = 1 \\ 1 - \pi_i & z_i = 0 \end{cases}
$$

Once the joint determination of I and H is explicitly accounted for by the forms of the probabilities ν and π, the two random variables may be assumed to be independent. Hence

$$
\Pr\{I_i = y_i,\ H_i = z_i\} = \nu_i^{y_i} (1 - \nu_i)^{1 - y_i}\, \pi_i^{z_i} (1 - \pi_i)^{1 - z_i}
$$

The likelihood function can then be written as:

$$
L = \prod_{i=1}^{n} \nu_i^{y_i} \times (1 - \nu_i)^{1 - y_i} \times \pi_i^{z_i} \times (1 - \pi_i)^{1 - z_i} \tag{18}
$$

The maximization of the likelihood function (18) under the two (alternative) models (17) will then uncover which one of the two models of integration formation better fits our data.[26]

---

[25] Otherwise one can proceed via GMM as in Bisin, Topa and Verdier (2004).

[26] The estimation has been performed using R programming language (www.r-project.org).

16

4.5 Results

The maximum likelihood estimation results of the endogenous identity (i.e. marriage and identity) and the endogenous location (i.e. marriage and location) are reported in the first two columns of Table 3, using the most extensive set of controls.[27] The last two columns report the results of our analysis when restricting the attention to Muslims only, which we will discuss in Section 5. Following Burnham and Anderson (2002), we adopt the Akaike Information Criterion (AIC) as a test between non-nested models for model selection. Different competing models can be ranked according to their AIC, with the one having the lowest AIC being the best. In the general case, $AIC = 2k - 2\ln(L)$ where k is the number of parameters in the statistical model, and L is the value of the maximized likelihood function. In our simple context (in which k does not vary between models), the AIC reduces to a likelihood comparison: the preferred model is the one with the highest maximized likelihood (i.e. lowest AIC). The maximized likelihoods are reported in the last row of Table 3. The evidence overwhelmingly favors the endogenous identity model over the endogenous location model. Indeed, the endogenous identity model shows a maximized likelihood value that is much higher than the one of the endogenous location model (-1288.972 and -1369.067 respectively). It is thus apparent that the endogenous identity model (with exogenous location) performs best.[28]

[*Insert Table* 3 *here*]

Let us now focus our attention on the parameter estimates. As noted before, the effects of the exogenous regressors are not additively separable from the effect of q. Our test of the cultural distinction against the cultural conformity mechanism is based on the estimate of γ_2. As stated in (17), a positive estimate of γ_2 would be in line with the cultural distinction model whereas a negative estimate would support the cultural conformity model. Table 3 shows that such an estimate is clearly positive, which indicates that a cultural distinction mechanism of identity formation is consistent with our data.[29]

To develop a better intuition about which aspect of the data drives this result, we study the endogenous identity model (i.e. the better performing model) in more detail. Under *cultural distinction*, the first order conditions of the model are given by:[30]

$$\nu = \pi\,c(1-q) \tag{19}$$

$$\pi = q + \frac{c(1-q)^3}{\alpha}\nu \tag{20}$$

Various simple conclusions can be obtained from these equations. Consider first condition (19), expressing how identity formation ν depends on the proportion q of minority members and on π

---

[27] The estimation has also been performed under different (increasing) sets of controls (as in Table 1). The results on our target parameters remain qualitatively unchanged.

[28] For only one of the two solutions of the endogenous location model the estimation converges.

[29] A more rigorous statistical comparison which would imply to compare the unconstrained with the constrained likelihood under the two sets of constraints $\gamma_2 > 0$ and $\gamma_2 < 0$ is problematic in these cases because the distribution of the resulting likelihood ratio tests with inequality constraints are non-standard (see Chernoff, 1954; Wilks, 1938; Self and Liang, 1987; Shaw and Geyer, 1997).

[30] We replace $C(q) = c(1-q)$ in (9) and (10).

the probability of homogamy. Clearly, the larger is the proportion of minority members and the more segregated the neighborhood is, the lower is $C(q) = c(1 - q)$, the psychological cost of interacting with the majority group, and the smaller are the incentives for identity formation as a cultural distinction. On the other hand, the higher is the probability π of homogamous marriage resulting from socialization effort, the larger are the expected benefits from identity formation and the more intense is the identity.

Interestingly, marital segregation, as reflected by π, 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 is (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 (20) characterizing the (endogenous) probability of homogamy as a function of identity ν and neighborhood segregation q. The more intense is the identity formation, the higher is the probability of homogamy. On the other hand, the effect of q on π 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 (20). The second effect is illustrated by the second term $(1 - q)^2 C(q)\nu$ and indicates the impact of a change in q on the marginal incentives to marital segregation (i.e., the socialization effort τ). Indeed, the more segregated is the neighborhood (i.e., the larger is q), the smaller are the incentives to spend resources of finding directly a partner in the restricted pool of minority spouses. First, because social interactions with the majority people are less costly, there is less of a need for identity formation that can be effectively expressed in an homogamous marriage. This is reflected by the term $C(q)$. Second, a larger proportion of minority people in the neighborhood also reduces the incentives to make special effort 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 τ, which, in turn, tends to reduce the probability of homogamy π.

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 given by:

$$\nu = \pi \, cq \tag{21}$$

$$\pi = q + \frac{c}{\alpha} q(1 - q)^2 \nu \tag{22}$$

Again, simple conclusions can be drawn from these equations. Consider first (21). In that case, the larger the proportion q of minority members and the more segregated the neighborhood, the higher $C(q)$. Similarly, the larger the probability π 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 π) 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 is (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 (22). The more intense is identity formation, the larger is the probability of homogamy. The effect of q on π 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 (22). 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 τ).

The last effect of q on π 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 (22), decreases the incentives for socialization efforts τ, and tends to reduce the probability of homogamy π.

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

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

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

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

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 models, it clearly appears that the distinctive characteristics of the cultural distinction model are:

1. ν 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.

4.6 Simulations

We now perform a series of simulations, using the estimates of our best model, the endogenous identity model, to predict *ethnic integration*. We study the latter by looking at the relationship between the identity or the homogamy status of an ethnic minority and his or her *time spent in the UK*. We will also look at the role played by the neighborhood ethnic composition where this individual lives in shaping such a process. We distinguish between first and second generations of immigrants.[31]

Our analysis so far suggests that ethnic identities might be more intense in mixed rather than in segregated neighborhoods. The additional questions that we seek to investigate regard the implications of our structural estimates: *(i)* Do they imply that ethnic identity and homogamy decrease with the time a person spends in the UK? And, is such a tendency more marked for second generation immigrants? *(ii)* Do they imply that minorities living in more integrated (mixed) neighborhoods show a stepper decreasing trend in the ethnic variables than those living in segregated neighborhoods? And, is there any difference in the implied assimilation pattern between first and second generation of immigrants for a given neighborhood ethnic composition? We address such questions while controlling for a variety of other individual and contextual characteristics, i.e., our simulations are run on the model that includes our extensive set of controls.

We begin by showing the integration patterns of ethnic minorities over time. Figures 4a and 4b show the predicted identity and homogamy values as a function of time spent in the UK for different generations. We can see that there is a tendency toward ethnic integration, which is particularly pronounced for the second generation, in particular for identity.

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

We then simulate those patterns when setting neighborhood composition equals to the sample minimum, average and maximum levels. The results are reported in Figures 5a and 5b, for identity and homogamy variables, respectively. Interestingly, the pictures remains largely unchanged for the different neighborhood ethnic compositions. It only appears a less steep decreasing trend of homogamy in more segregated neighborhoods.

---

[31] "Time spent in the UK" is measured as "years since arrival" " for first generation immigrants and as "age" for second generation immigrants. Second generation immigrants are defined as ethnic minority individuals who are born in the UK.

Such findings thus show evidence of an integration patterns of immigrants, which is particularly marked for second generation individuals and when ethnic identity is the ethnic dimension considered. The residential neighborhood ethnic composition, however, does not seem to play a crucial role in shaping such a pattern, neither for the first nor for the second generation of immigrants.

Therefore, our analysis seems to suggest not only that ethnic identities might be more intense in mixed rather than in segregated neighborhoods, but also that living in more mixed neighborhoods does not ease the integration pattern.

5 Is Muslim identity different?

A large debate has recently emerged in the popular press about the alleged specificity of Muslim immigrants with regard to the strength of their identity and their (lack of) assimilation tendencies.[32]

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 direct question asking the respondent to identify his or her 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 (713 individuals, roughly 45 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. Descriptive statistics of our variables on this sub-sample can be found in the Data Appendix (Table A1).

Our maximum likelihood results are contained in the last two columns of Table 3. They are not qualitatively different from those in the first two columns, which are obtained on the whole sample and discussed in the previous section. The endogenous identity with cultural distinction model remains the mechanism that is more likely to represent the observed evidence. Indeed, a comparison between the maximized likelihood values between the endogenous identity and the endogenous location model clearly favors again the former one (-402.489 versus -446.066, respectively). The positive (and statistically significant) estimate of γ_2 points to the fact that cultural distinction is the mechanism that seems to fit better our data. Such a finding indicates that the relationship between ethnic assimilation effort and ethnic neighborhood composition is not different for Muslims than for other minorities. In other words, we still find evidence in line with the possibility that ethnic identity and socialization

---

[32] This position has been taken, in a rather extreme form, by several nationalist parties, e.g., the British National Party in England, the Lega in Italy, the Front National in France, Sverige Demokraterna in Sweden. Similar though less extreme positions have also 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 and Bisin et al. (2008a) for Muslims in the UK.

effort are more intense in mixed rather than in segregated neighborhoods when only the Muslim subsample is considered.

As for the whole sample, we now perform some simulations on the Muslims sub-sample of our best model, the endogenous identity model, to predict ethnic integration, that is, identity and homogamy as a function of "time spent in the UK", and the role played by the neighborhood ethnic composition in shaping such a process. The results are reported in Figures 6a and 6b, and 7a and 7b. Evidence of slower integration for Muslims is apparent from these figures, in particular for ethnic identity. More interestingly, although the second generation shows levels of identity and homogamy always lower than the first generation, we find for second generation Muslims a tendency *to increase rather than decrease the strength of their ethnic identity and homogamy over time.* The almost non-existent role of neighborhood ethnic composition in shaping these patterns remain confirmed in this sub-sample of Muslims.

[*Insert Figures* 6a, 6b, 7a, *and* 7b *here*]

6 Discussion and policy implications

In this paper, we find that: (*i*) in mixed neighborhoods, ethnic identity is much stronger than in homogenous neighborhoods, suggesting *cultural distinction* among ethnic minorities in England; (*ii*) there is a tendency toward ethnic integration, which is particularly pronounced for the second generation, in particular for ethnic identity; (*iii*) there are no significant differences in the integration patterns in neighborhoods with different ethnic compositions. Although Muslims, and in particular the second generation, seem to be more reluctant to integrate, our findings (*i*) and (*iii*) remain unchanged.

In light of these results, our analysis suggests that integration policies favoring the formation of mixed neighborhoods, fearing the effects of geographical segregation, are possibly minimally effective if not counterproductive. In this respect, our results could explain why the different integration policies implemented in the US and in Europe seem to have small effects because of the possible perverse effects of integration policies which might trigger more intense ethnic identities. Indeed, mixing 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,[33] 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), consistently with the view that identity formation mechanisms are driven by *cultural distinction*,

James Coleman, for instance, fifteen years after the Coleman Report in 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).[34] Moving to Opportunity (MTO) programs in the United States that relocate families from high- to low-poverty neighborhoods

---

[33] See Lang (2007) for an overview of these policies in the U.S.

[34] The failure of the busing and other civil right policies is certainly also due to the whites' flight from de-segregated schools and 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).[35] 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. In Europe different integration policies and ambitious social programs have been implemented in urban areas where immigrants live but they have had also 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.[36] 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).

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, because of *cultural distinction.*

---

[35]There are, of course, other complementary mechanisms that could explain why the MTO programs had small effects. It is indeed possible that it is the loss of social networks that makes this policy relatively inefficient. A recent book (De Souza Briggs et al., 2010) evaluating the MTO programs shows that poor black families who move to richer areas tend to mostly interact with their old friends (i.e. their old social networks) from their old "bad" neighborhoods. Based on interviews and case studies, the authors claim that this may explain the relative small effects of this policy. Another explanation for the relative failure of MTO programs given by Quigley and Raphael (2008) is that families move from very poor to poor neighborhoods and thus the interactions in the new neighborhood is not that beneficial. Using data on adolescents in the US (AddHealth data), Weinberg (2010) show that individuals associate with people whose behaviors and characteristics are similar to their own and this tendency is stronger in large groups. This suggests that the MTO types of policies can have small effects because the people who move to richer areas only interact with similar types of people and not necessary with those who could generate positive externalities.

[36]See, for example, Benabou, Kramarz and Prost (2009) for an evaluation of the ZEP and Brubaker (2001) who compares the different ways of assimilating ethnic minorities in France, Germany, and the US.

References

[1] Abrams D. amd 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, 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] Akerlof, G.A. and R.E. Kranton (2010), *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*, Princeton: Princeton University Press.

[7] Alesina, A. and E. La Ferrara (2005), "Ethnic Diversity and Economic Performance," *Journal of Economic Literature* 43, 762-800.

[8] Al-Johar, D. (2005), "Muslim Marriages in America: Reflecting New Identities," *The Muslim World* 95, 557-574.

[9] Austen-Smith, D. and R.D. Fryer, Jr (2005), "An Economic Analysis of 'Acting White'," *Quarterly Journal of Economics* 120, 551-583.

[10] Battu, H., McDonald M., and Y. Zenou (2007), "Oppositional Identities and the Labor Market," *Journal of Population Economics* 20, 643-67.

[11] Battu, H. and Y. Zenou (2010), "Oppositional Identities and Employment for Ethnic Minorities. Evidence for England," *Economic Journal* 120, F52-F71.

[12] Benabou, R., Kramarz, F. and C. Prost (2009), "The French Zones d'Education Prioritaire: Much Ado About Nothing?" *Economics of Education Review* 28, 345-356.

[13] Bernheim, B.D. (1994), "A Theory of Conformity," *Journal of Political Economy* 102, 841-877.

[14] Bisin A., Moro A. and G. Topa (2009), "The Empirical Content of Models with Multiple Equilibria," Unpublished manuscript, New York Universiy.

[15] 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.

[16] Bisin, A., Topa, G. and T. Verdier (2004), "Religious intermarriage and socialization in the United States," *Journal of Political Economy* 112, 615-664.

[17] Bisin, A., Patacchini, E., Verdier, T. and Y. Zenou (2008a), "Are Muslims Immigrants Different in terms of Cultural Integration?" *Journal of the European Economic Association* 6, 445-456.

[18] Bisin, A., E. Patacchini, T. Verdier, and Y. Zenou (2008b), "Formation and Persistence of Oppositional Identities," Unpublished manuscript, Paris School of Economics.

[19] 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.

[20] Bodendorn H., and C. Ruebeck (2003), "The Economics of Identity and the Endogeneity of Race," NBER WP No 9962.

[21] Brewer M. (2001), "The Many Faces of Social Identity ; Implications for Political Psychology," *Political Psychology* 22, 115-125.

[22] Brubaker, R. (2001), "The Return of Assimilation? Changing Perspectives on Immigration and its Sequels in France, Germany, and the United States," *Ethnic and Racial Studies* 24, 531-548.

[23] Burnham, K. P. and D. R. Anderson (2002), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, Berlin: Springer-Verlag.

[24] Chernoff, H. (1954), "On the Distribution of the Likelihood Ratio," *Annals of Mathematical Statistics* 25, 573-578.

[25] Chiswick, C.U. (2009), "The Economic Determinants of Ethnic Assimilation," Journal of Population Economics 22, 859-880.

[26] Constant, A.F., L. Gataullina, and K.F. Zimmermann (2009), "Ethnosizing Immigrants," *Journal of Economic Behavior and Organization* 69, 274-87.

[27] Constant, A.F. and K.F. Zimmermann (2008), "Measuring Ethnic Identity and Its Impact on Economic Behavior," *Journal of the European Economic Association* 6, 424-33.

[28] Cross W.E. (1991), *Shades of Black: Diversity in African-American Identity*, Philadelphia: Temple University Press.

[29] Darity, W. Jr., P.L. Mason and J.B. Stewart (2006), "The Economics of Identity: The Origin and Persistence of Racial Identity Norms," *Journal of Economic Behavior and Organization* 60, 283-305.

[30] De Souza Briggs, X., Popkin, S.J. and J, Goering (2010), *Moving to Opportunity. The Story of an American Experiment to Fight Ghetto Poverty*, Oxford: Oxford University Press.

[31] Eaton, B.C., M. Eswaran and R.J. Oxoby (2009), "Us and Them: The Origin of Identity, and its Economic Implications," Unpublished manuscript, University of Calgary.

[32] Fallaci O., (2006), *The Force of Reason,* New York, Rizzoli.

[33] Fang H., and G. Loury (2005), "Dysfunctional Identities Can be Rational," *American Economic Review* 95, 104-111.

[34] Ferdman, B.M. (1995), "Cultural Identity and Diversity in Organizations: Bridging the Gap between Group Differences and Individual Uniqueness, In: M.M. Chemers, S. Oskamp, and M. A. Costanzo (Eds.), *Diversity in Organizations: New Perspectives for a Changing Workplace*, Thousand Oaks, CA: Sage, pp. 37-61.

[35] Fryer, R.G. Jr. and P. Torelli (2010), "An Empirical Analysis of 'Acting White'," *Journal of Public Economics* 94, 380-396.

[36] Gallis, P. (Coordinator) (2005), "Muslims in Europe: Integration Policies in Selected Countries," CRS Report for Congress, Congressional Research Service, The Library of Congress.

[37] Glazer N. and D.P. Moynihan (1970), *Beyond the Melting Pot: The Neogroes, Puerto Ricans, Jews, Italians and Irish of New York City*, Cambridge, MA: MIT Press.

[38] Gordon M.M. (1964), *Assimilation in American Life: The Role of Race, Religion and National Origins*, New York: Oxford University Press.

[39] Gussin Paley, V. (1995), *Kwanzaa and Me: A Teacher's Story*, Cambridge, MA: Harvard University Press.

[40] Jacoby, T. (1998), *Someone Else's House: America's Unfinished Struggle for Integration*, New York: Free Press.

[41] Jacoby, T. (1999), "Beyond Busing," *The Wall Street Journal*, July 21st.

[42] 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.

[43] Lang, K. (2007), *Poverty and Discrimination*, Princeton: Princeton University Press.

[44] 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.

[45] 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.

[46] Maddala, G.S. (1983), *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.

[47] Manning A. and S. Roy (2010), "Culture Clash or Culture Club? The Identity and Attitudes of Immigrants in Britain," *Economic Journal* 120, F72-F100.

[48] Meng, X. and R.G. Gregory (2005), "Intermarriage and the Economic Assimilation of Immigrants," *Journal of Labor Economics* 23, 135-176.

[49] Moghaddam F.M and E.A. Solliday (1991), "Balanced Multiculturalism and the Challenge of Peaceful Coexistence in Pluralistic Societies", *Psychology and Developing Societies* 3, 51-71.

[50] 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.

[51] Munshi, K. and N. Wilson (2007): "Identity, Parochial Institutions, and Occupational Choice: Linking the Past to the Present in the American Midwest," NBER Working Paper 13717.

[52] Nekby, L. and M. Rödin (2010), "Acculturation Identity and Employment among Second and Middle Generation Immigrants," *Journal of Economic Psychology* 31, 35-50.

[53] Oreopoulos, P. (2003), "The Long-Run Consequences of Living in a Poor Neighborhood," *Quarterly Journal of Economics* 4, 1533-75.

[54] Patacchini, E. and Y. Zenou (2009), "Neighborhood Effects and Parental Involvement in the Intergenerational Transmission of Education," Unpublished manuscript, Stockholm University.

[55] Phinney, J.S. (1990), "Ethnic Identity in Adolescents and Adults: Review of Research," *Psychological Bulletin* 108, 499-514.

[56] 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.

[57] Qian, Z. (1999), "Who Intermarries? Education, Nativity Region, and Interracial Marriage 1980 and 1990," *Journal of Comparative Family Studies* 30, 579-597.

[58] Quigley, J.M. and S. Raphael (2008), "Neighborhoods, economic self-sufficiency, and the MTO program," *Brookings-Wharton Papers on Urban Affairs*, 1-46.

[59] 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* 82, 605-610.

[60] Shaw, F. H. and C. J. Geyer (1997), "Estimation and Testing in Constrained Covariance Component Models," *Biometrika* 84, 95-102.

[61] 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.

[62] Stryker S. (1980), *Symbolic Interactionism: A Social Structural Version*, Menlo Park, CA: Benjamin Cummings.

[63] Tajfel H.(1981), *Human Groups and Social Categories: Studies in Social Psychology*, Cambridge: Cambridge University Press.

[64] Tajfel H. and J.C.Turner (1979), "An Intergrative Theory of Intergroup Conflict," In: W.G. Austin and S. Worchel (Eds.), *The Social Psychology of Intergroup Relations*, Monterey:, CA Brooks/Cole, pp. 33-47.

[65] 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.

[66] 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.

[67] 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.

[68] Turner J.C. (1982), "Towards a Cognitive Redefinition of the Social Group," In: H. Tajfel (Ed.), *Social Identity and Intergroup Relations*, Cambridge: Cambridge University Press, pp. 15-44.

[69] 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: Blackwell.

[70] Weinberg, B. (2010), "Social Interactions with Endogenous Associations," Unpublished manuscript, Ohio State University.

[71] Wilde, J. (2000), "Identification of Multiple Equation Probit Models with Endogenous Dummy Regressors," *Economics Letters* 69, 309-312.

[72] Wilks S. S. (1938), "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses," *Annals of Mathematical Statistics* 9, 60-62.

[73] Whyte, W.F. (1943), *Street Corner Society*, Chicago: University of Chicago Press.

APPENDIX

1. The marriage and identity model

Let us consider the model where both the strength of identity and the type of marriage are choice variables. In the extended model, the ethnic minority minimizes his/her expected cost of interacting with whites, that is:

$$\min_{\nu,\tau} \left\{ -\nu\pi c\left(\gamma_1 - \gamma_2 q\right) + c\left(\gamma_1 - \gamma_2 q\right) + \frac{1}{2}\alpha\tau^2 + \left(\gamma_\tau\, x_\tau\right)\tau + \frac{1}{2}\nu^2 + \left(\gamma_\nu\, x_\nu\right)\nu \right\}$$

where the cost $C(q) = c\psi(q)$, with $\psi(q) = \gamma_1 - \gamma_2 q$. In the cultural conformity model, $\gamma_1 = 0$ and $\gamma_2 = -1$, while, in the cultural distinction model, $\gamma_1 = \gamma_2 = 1$. Solving this program leads to:

$$\nu = \left[c\left(\gamma_1 - \gamma_2 q\right)\right]\pi - \gamma_\nu\, x_\nu \tag{23}$$

$$\tau = \frac{\nu(1-q)\left[c\left(\gamma_1 - \gamma_2 q\right)\right] - \gamma_\tau\, x_\tau}{\alpha} \tag{24}$$

Now using the fact that $\tau = (\pi - q)/(1 - q)$ (from (2)), equation (24) can be written as:

$$\pi = q + \frac{\nu}{\alpha}(1-q)^2\left[c\left(\gamma_1 - \gamma_2 q\right)\right] - \frac{(1-q)}{\alpha}\gamma_\tau\, x_\tau \tag{25}$$

Observe that (23) and (25) are equivalent to (9) and (10) for the extended model. We would like to express ν and π as a function of q only. By combining these equations, we obtain:

$$\nu^* = \frac{\alpha\left[c\left(\gamma_1 - \gamma_2 q\right)\right]q - (1-q)\left[c\left(\gamma_1 - \gamma_2 q\right)\right]\gamma_\tau\, x_\tau - \alpha\gamma_\nu\, x_\nu}{\alpha - \left[c\left(\gamma_1 - \gamma_2 q\right)\right]^2(1-q)^2} \tag{26}$$

$$\pi^* = \frac{\alpha q - (1-q)\gamma_\tau\, x_\tau - \left[c\left(\gamma_1 - \gamma_2 q\right)\right](1-q)^2\gamma_\nu\, x_\nu}{\alpha - \left[c\left(\gamma_1 - \gamma_2 q\right)\right]^2(1-q)^2} \tag{27}$$

We need therefore to identify $\gamma_1, \gamma_2, c, \alpha, \gamma_\nu$ and γ_τ. A simple informal argument for the identification of these parameters is as follows. For each ethnic minority, we observe q (the proportion of ethnic minorities living in his/her neighborhood), x_ν (the individual's characteristics links to the cost of identity choice) and x_τ (the individual's characteristics links to the cost of the type of marriage choice), the realization of $I \in \{0,1\}$ (i.e. whether he/she has a strong identity or not), and the realization of $H \in \{0,1\}$ (i.e. whether he/she is married to someone from the same ethnic group). Since $\Pr\{I = 1\} = \nu$ and $\Pr\{H = 1\} = \pi$, we identify ν with the fraction of individual with identity $I = 1$ and π with the fraction of agents with homogamous marriage $H = 1$. The parameters γ_1, γ_2, c and α are identified out of variations in q, from equations (26) and (27). Variations in x_ν and x_τ identify instead γ_ν and γ_τ. This is even the case if a single vector x enters in the determination of ν and π, that is, $x_\nu = x_\tau = x$. For given $q, c, \alpha, \gamma_1, \gamma_2$, in fact, in this case γ_ν and γ_τ solve:

$$k_\nu = -\frac{\alpha\gamma_\nu + c\left(\gamma_1 - \gamma_2 q\right)(1-q)\gamma_\tau}{\alpha - (1-q)^2 c^2\left(\gamma_1 - \gamma_2 q\right)^2}$$

$$k_\pi = -\frac{c\,(\gamma_1 - \gamma_2 q)\,(1-q)^2 \gamma_\nu + (1-q)\gamma_\tau}{\alpha - (1-q)^2 c^2 \,(\gamma_1 - \gamma_2 q)^2}$$

where k_ν and k_π are the estimated coefficients of x in the equations (26) and (27) for ν and π, respectively, for a given q. It is easy to check that one such solution exists.

2. The marriage and location model

The ethnic minority minimizes his/her expected cost of interacting with whites, that is:

$$\min_{q,\tau} \left\{ -\nu \pi c\,(\gamma_1 - \gamma_2 q) + c\,(\gamma_1 - \gamma_2 q) + \frac{1}{2}\alpha\tau^2 + (\gamma_\tau\, x_\tau)\,\tau + \frac{1}{2}q^2 + \left(\gamma_q\, x_q\right) q \right\}$$

The first order conditions of this individual's problem are:

$$-\nu c\left[(1-\tau)\,(\gamma_1 - \gamma_2 q) - \pi\gamma_2\right] - \gamma_2 c + q + \gamma_q\, x_q = 0$$
$$-\nu\,(1-q)\,c\,(\gamma_1 - \gamma_2 q) + \alpha\tau + \gamma_\tau\, x_\tau = 0$$

which, using $\tau = \frac{\pi - q}{1 - q}$ (from (2)), are equivalent to:

$$(1 - \pi\nu)\,c\gamma_2 + \nu\left(\frac{1-\pi}{1-q}\right)c\,(\gamma_1 - \gamma_2 q) - q - \gamma_q\, x_q = 0$$
$$\nu c\,(\gamma_1 - \gamma_2 q) - \alpha\frac{(\pi - q)}{(1-q)^2} - \frac{\gamma_\tau\, x_\tau}{1 - q} = 0$$

Writing this system of equations in terms of ν and q leads to:

$$\nu = \frac{\alpha\,(\pi - q) + (1 - q)\,\gamma_\tau\, x_\tau}{(1-q)^2\,c\,(\gamma_1 - \gamma_2 q)}$$

and π is implicitly determined by the following equation

$$A_e\pi^2 + B_e\pi + D_e = 0$$

where

$$A_e = -\alpha\,(\gamma_1 + \gamma_2 - 2q\gamma_2)$$

$$B_e = \alpha\,(\gamma_1 - \gamma_2 q)\,(1 + q) + \alpha q\,(1-q)\,\gamma_2 - (1-q)\,[\gamma_1 - \gamma_2 q + \gamma_2\,(1-q)]\,\gamma_\tau\, x_\tau$$

$$D_e = (1 - q)\,(\gamma_1 - \gamma_2 q)\,\gamma_\tau\, x_\tau - \alpha q\,(\gamma_1 - \gamma_2 q) - \left(q + \gamma_q\, x_q - c\gamma_2\right)(1-q)^3\,(\gamma_1 - \gamma_2 q)$$

As a result,

$$\nu^* = \frac{\alpha\,(\pi - q) + (1 - q)\,\gamma_\tau\, x_\tau}{(1-q)^2\,c\,(\gamma_1 - \gamma_2 q)} \tag{28}$$

$$\pi^* = \frac{-B_e \pm \sqrt{B_e^2 - 4A_e D_e}}{2A_e} \tag{29}$$

To identify $\gamma_1, \gamma_2, c, \alpha, \gamma_\nu$ and γ_τ, we can again use the same informal argument as we did above for the marriage and identity model.

Table A1: Description of data

Variable	Explanation of the variable	Whole sample n.obs: 1,559		Muslim sample n.obs: 713	
		Mean	*St.dev.*	*Mean*	*St.dev.*
Main variables					
Importance of religion	In the text	0.61	0.49	0.79	0.41
Homogamy	In the text	0.92	0.27	0.97	0.17
Ward density of own ethnicity	In the text	13.60	10.76	15.04	11.23
Individual controls					
Age at arrival	Respondent's age in years at arrival in the UK	21.15	10.38	20.49	9.63
Age	Respondent's age in years	41.37	13.09	39.44	12.84
Female	Dummy variable taking value one if the respondent is female.	0.47	0.50	0.46	0.50
Born in the UK	Dummy variable taking value one if the respondent is born in the UK	0.10	0.30	0.06	0.26
Arranged Marriage	Dummy variable taking value one if the husband/wife of the respondent has been chosen by the parents.	0.37	0.48	0.53	0.50
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	0.09	0.28
Children	Dummy variable taking value one if the respondent has children.	0.91	0.28	0.93	0.25
Years since arrival	Number of years since respondent's arrival in the UK.	20.22	11.42	18.96	10.63
British degree	Dummy variable taking value one if the respondent has a UK degree.	0.21	0.41	0.09	0.28
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	0.08	0.26
Foreign education	Dummy variable taking value one if the respondent has a qualification achieved abroad.	0.28	0.45	0.26	0.44
Employed	Dummy variable taking value one if the respondent is employed.	0.47	0.50	0.29	0.45
Manager	Dummy variable taking value one if the respondent is a manager.	0.03	0.16	0.01	0.12
Employee	Dummy variable taking value one if the respondent is an employee	0.57	0.50	0.42	0.49
House owner	Dummy variable taking value one if the household owns (or is buying) the accommodation	0.76	0.43	0.67	0.47
Contextual controls					
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	0.07	0.25
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	0.45	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	0.39	0.49
Ward density of all ethnic groups	Percentage of residents of any ethnic group in the ward	33.63	21.01	38.66	20.27
Ward unemployment rate	Ward unemployment rate	14.38	5.24	16.60	4.51

Figure 1: Ethnic composition by neighborhood class
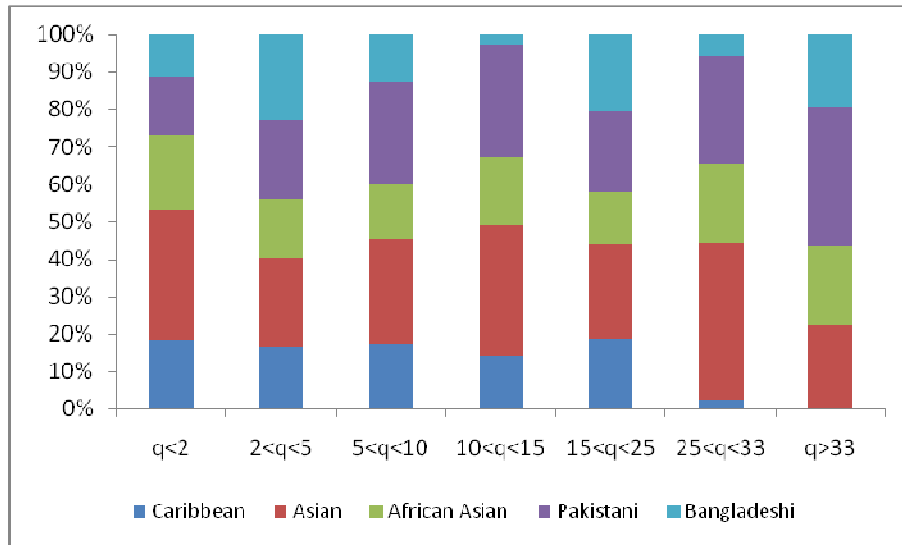
Notes. The Asian group comprises Indians and Chineses

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

Dep. Var.= Ethnic Identity

Dep. Var.= Homogamy
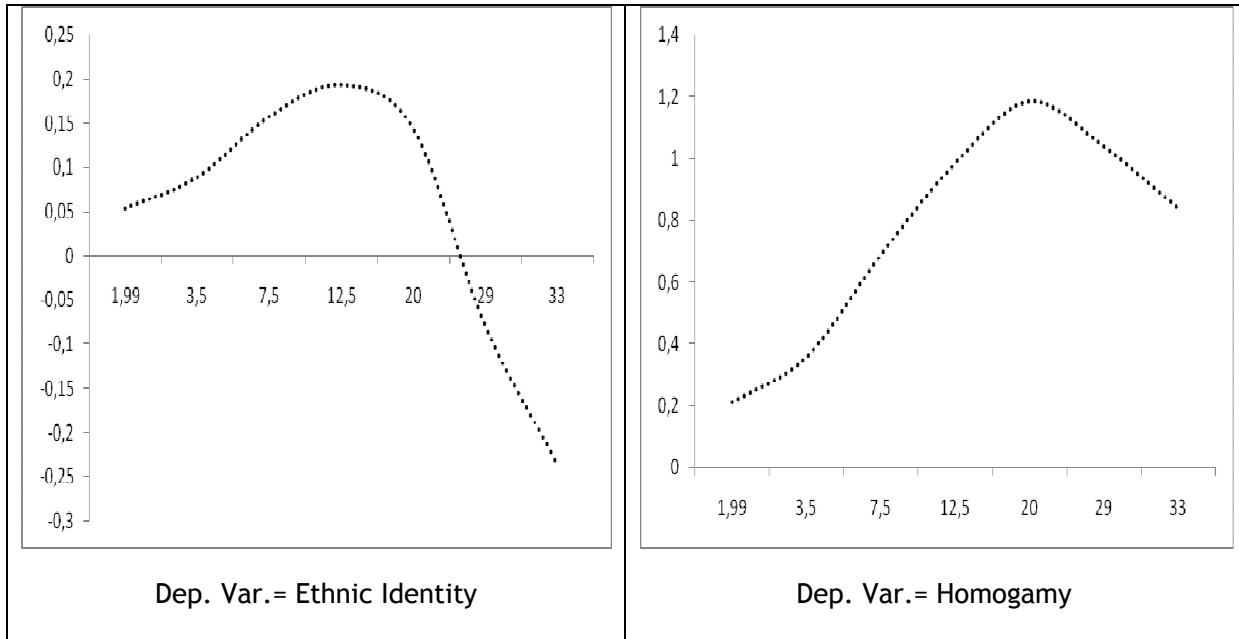
**Figure 2b: Non linear effect of neighborhood ethnic composition on identity and homogamy
Muslim sample**

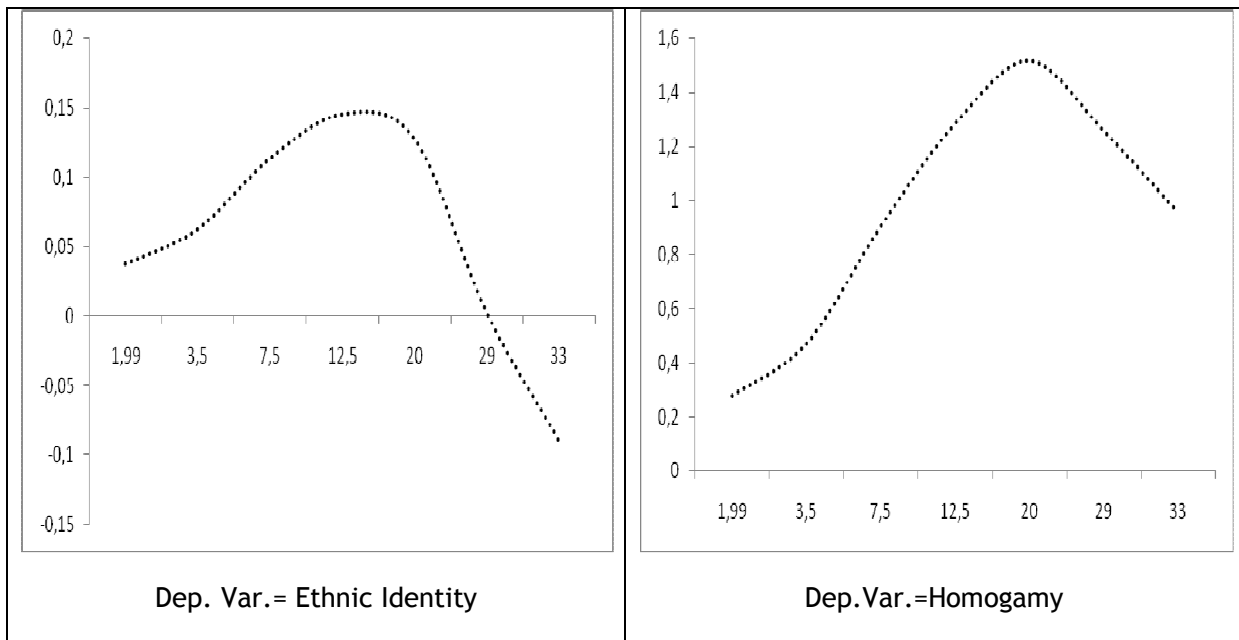Dep. Var.= Ethnic Identity

Dep.Var.=Homogamy

Figure 3a: Partial derivative of costs as a function of q
Whole sample

Ethnic Identity = 0   Ethnic Identity = 1
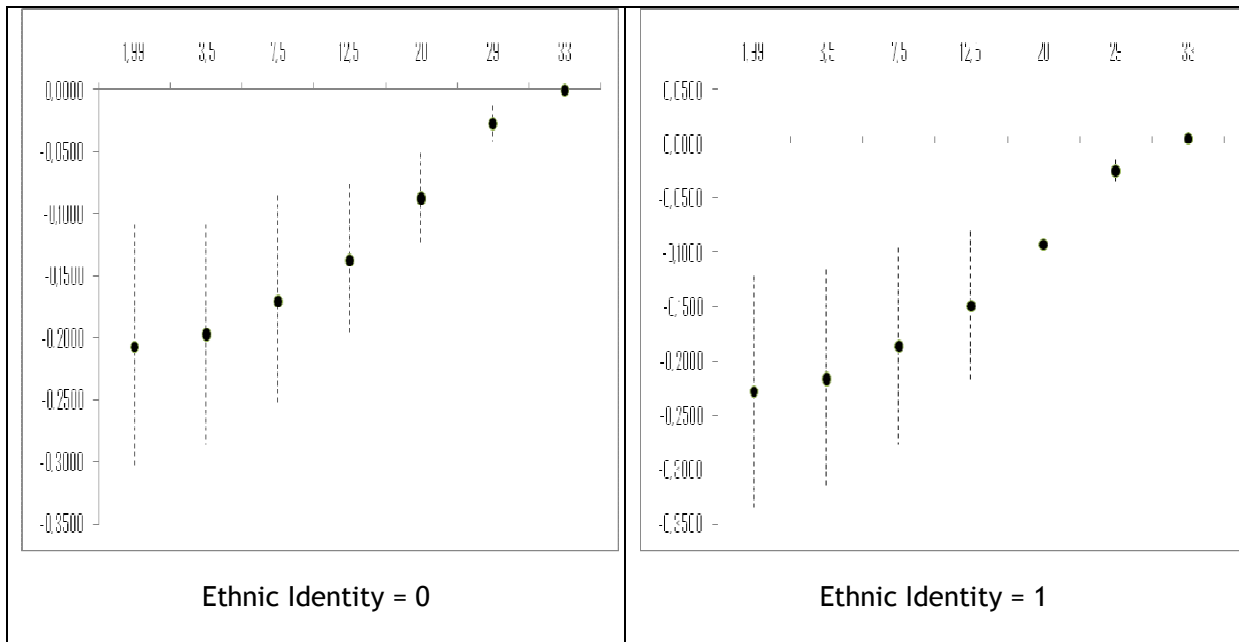
Figure 3b: Partial derivative of costs as a function of q
Muslim sample

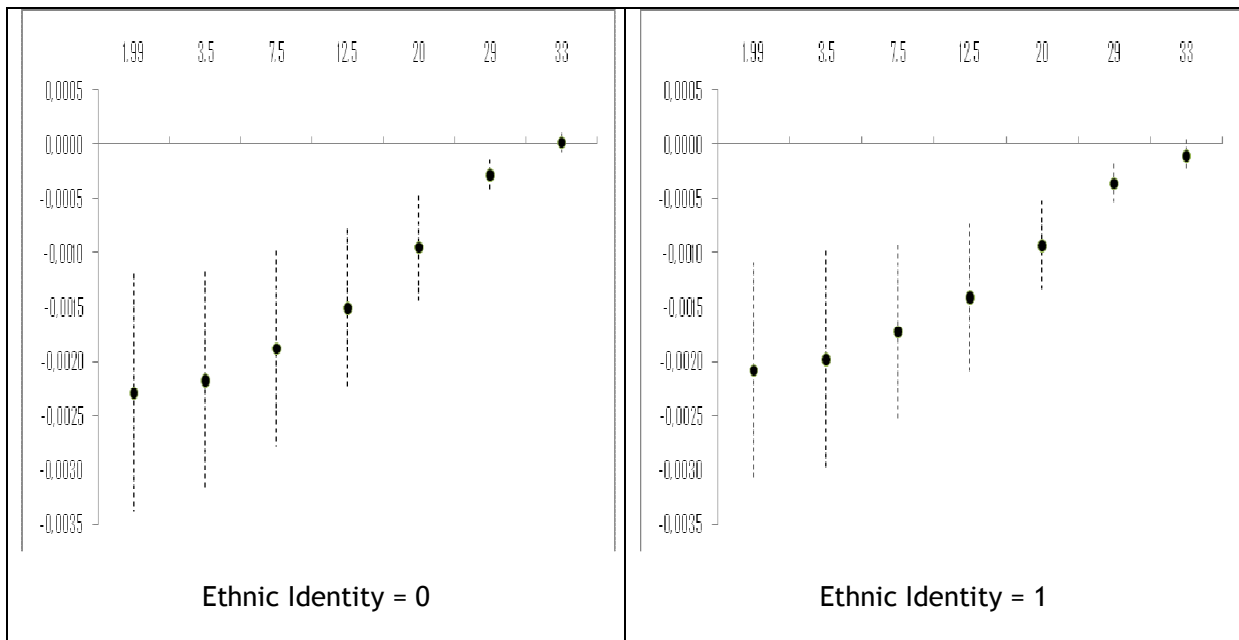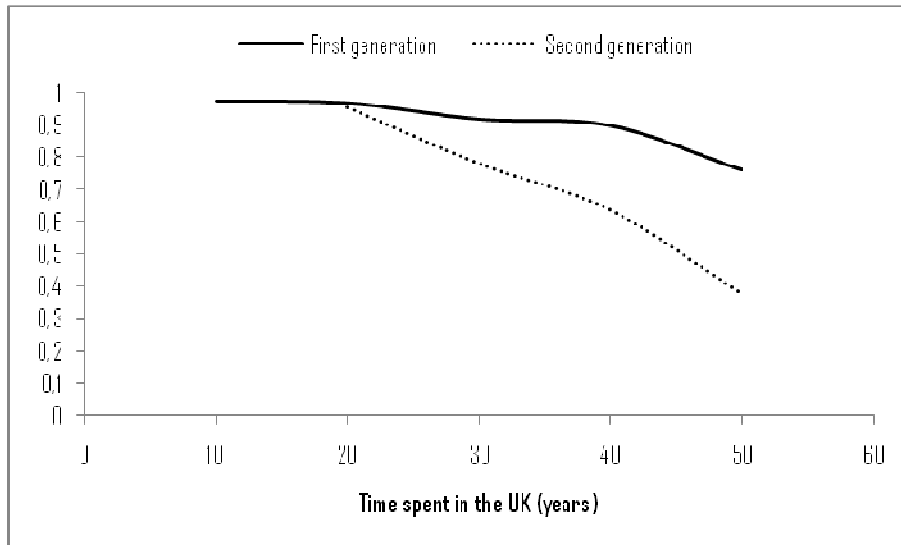Ethnic Identity = 0   Ethnic Identity = 1

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

Notes to Figures 4a to 5b. Endogenous identity model, second generation immigrants in the sample are older than 10 years

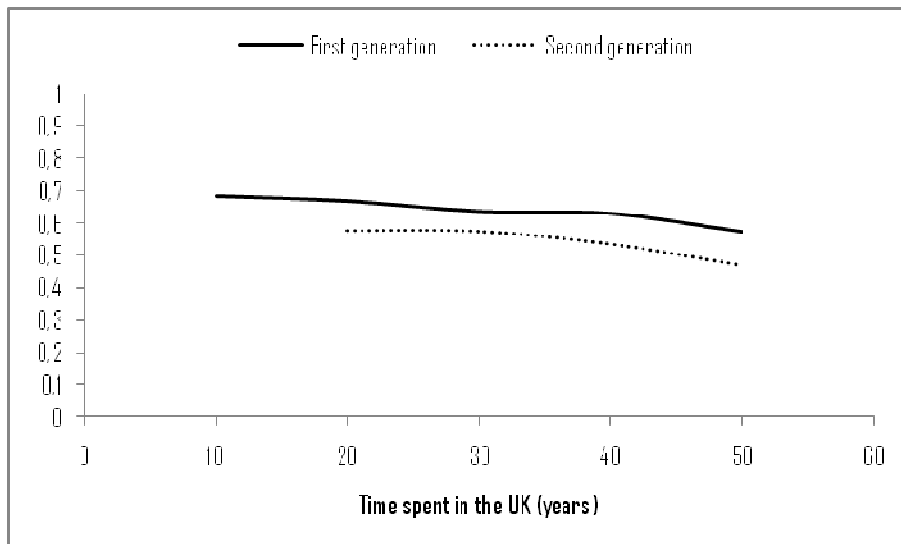**Figure 4b Predicted Homogamy as a function of time spent in the UK**

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
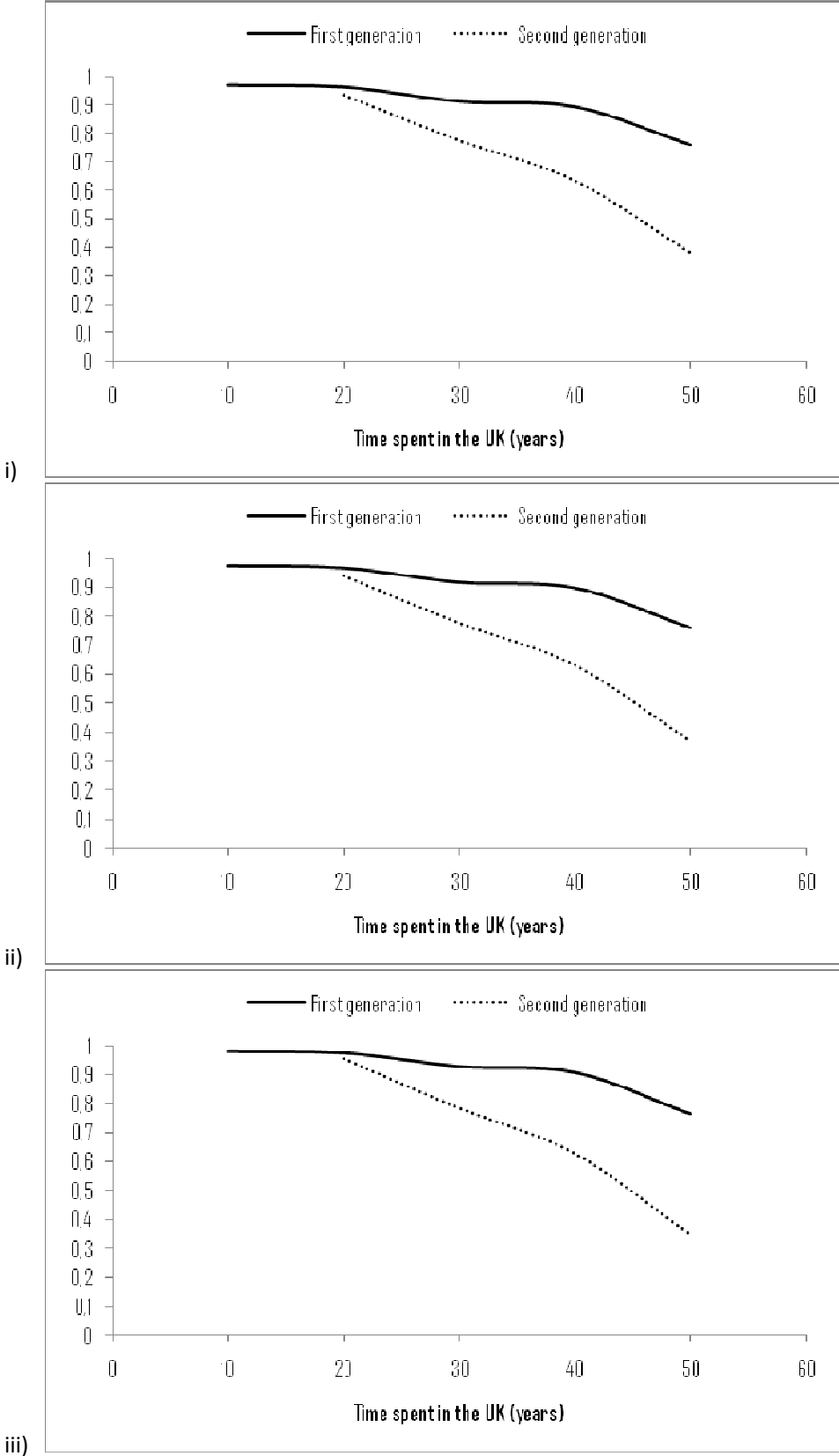

i)

ii)

iii)

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
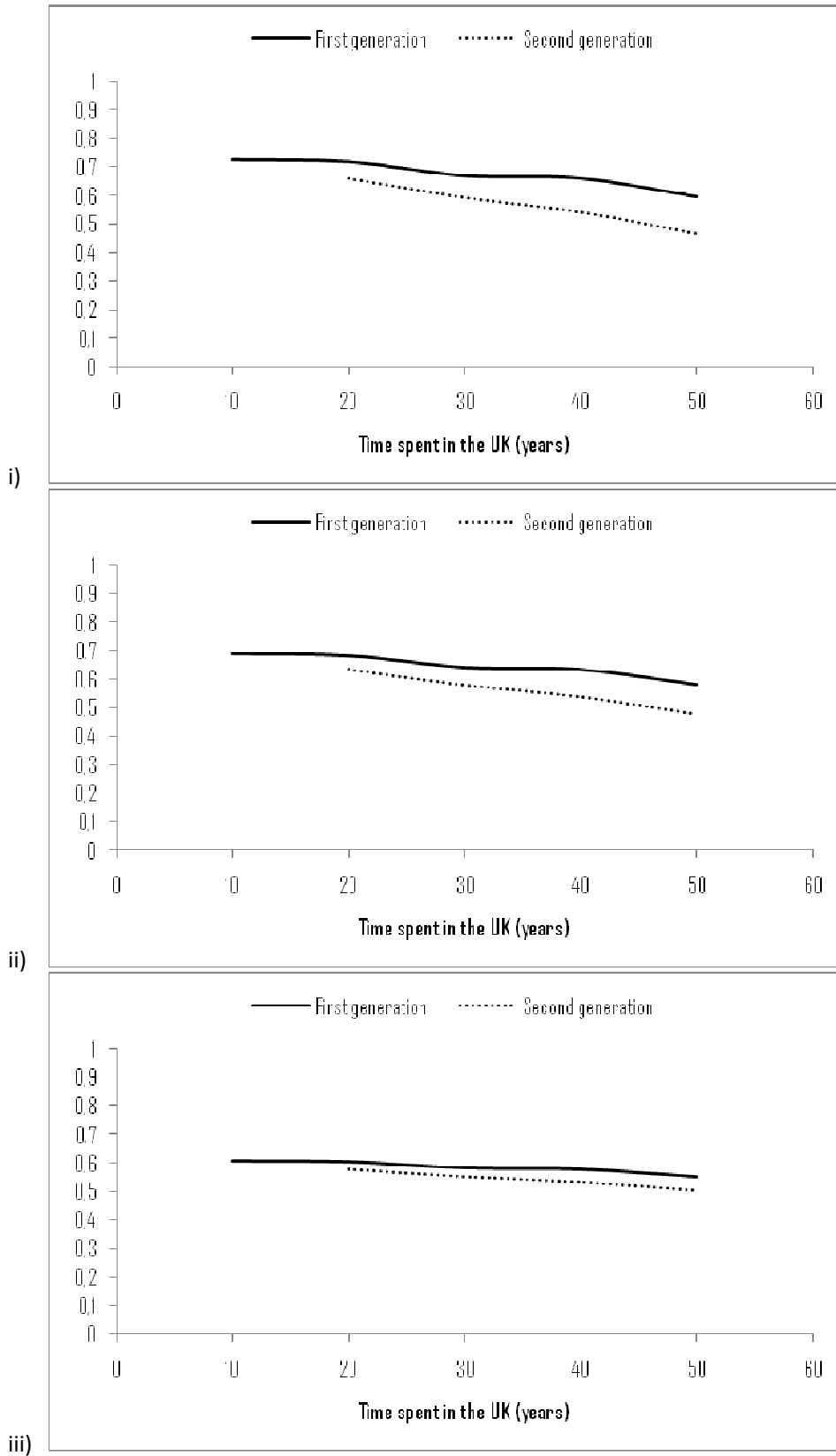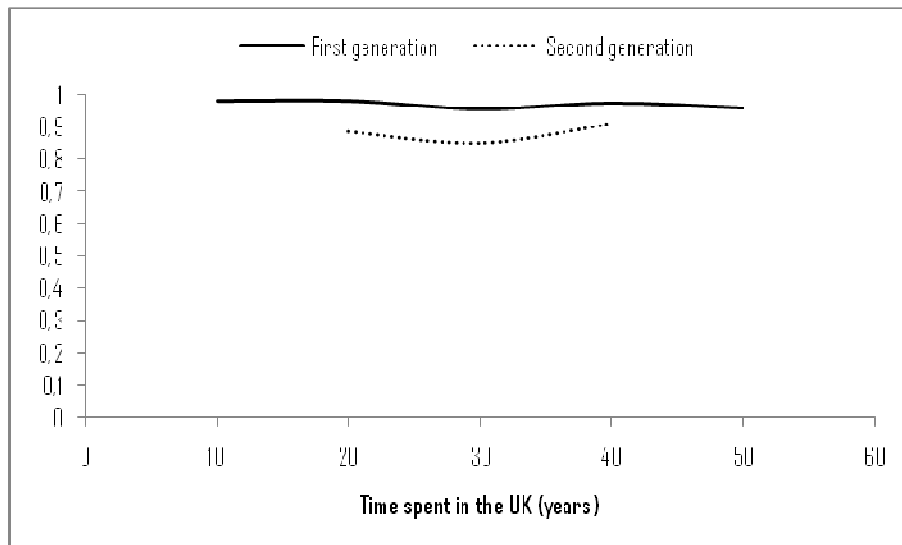

i)

ii)

iii)

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

Notes to Figures 6a to 7b. Endogenous identity model, second generation immigrants in the muslim sub- sample are older than 10 years and younger than 40 years

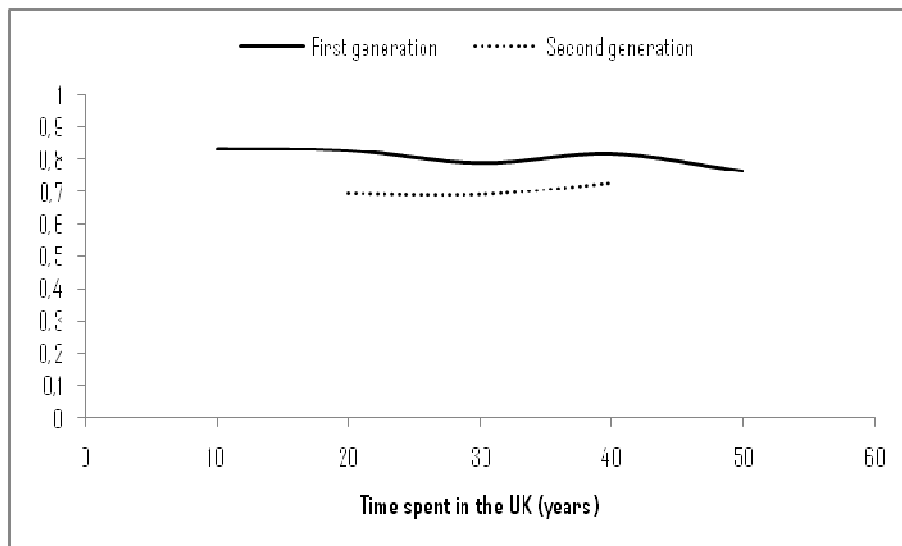**Figure 6b Predicted Homogamy as a function of time spent in the UK- Muslim sample-**

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 sample
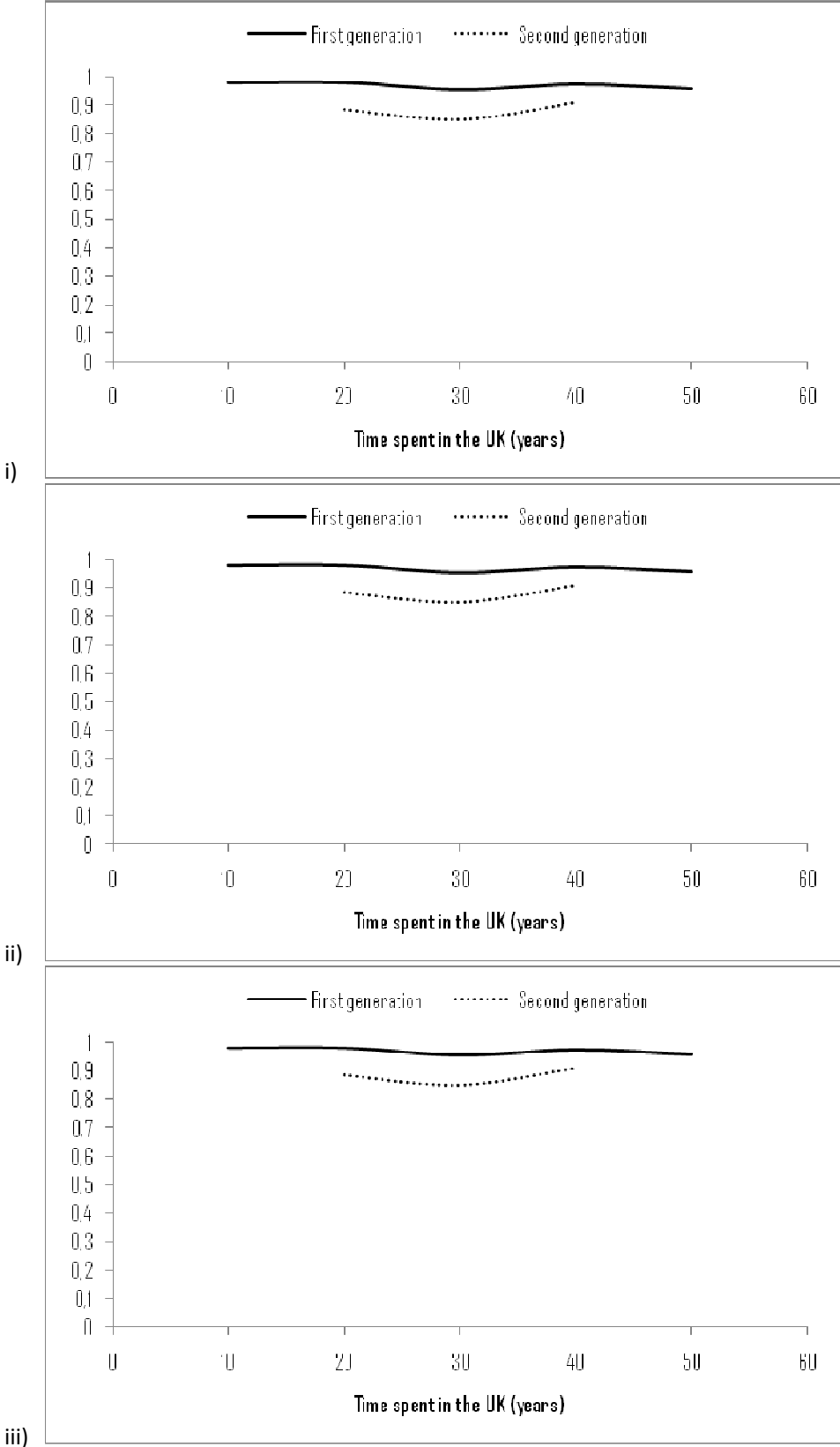

i)

ii)

iii)

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 sample

i)

ii)
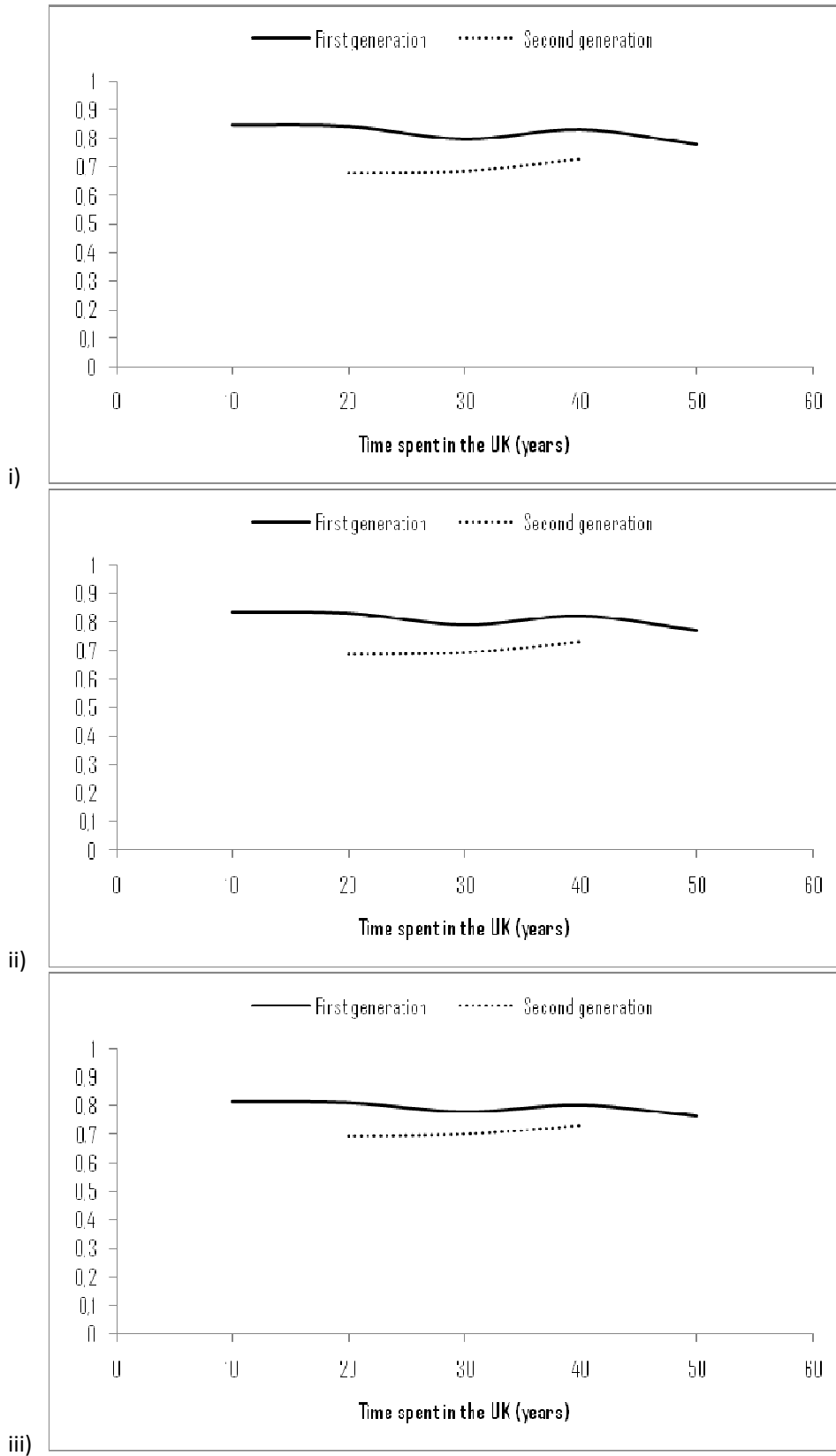

iii)

Table 1: Ethnic Identity, Homogamy and Ethnic Neighborhood Composition

	(1)	(2)	(1)	(2)	(1)	(2)
	(1): Dep. Var.: Ethnic identity measured by *importance of religion*					
	(2): Dep. Var.: Homogamy					
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)
Individual controls						
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)
Contextual controls						
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)
N.Obs.	1559		1559		1559	
Likelihood	-1342.0755		-1152.1474		-1115.9003	

Notes. Bivariate probit model estimation results. Marginal effects and robust standard errors in parentheses are reported. *** 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	0.1069***	0.0166	0.1310***	0.0138	0.1131***
	(0.0210)	(0.0268)	(0.0205)	(0.0315)	(0.0220)	(0.0344)
Ward density of own ethnic groups^2	-0.0004	-0.0019**	-0.0005	-0.0026***	-0.0002	-0.0026**
	(0.0006)	(0.0008)	(0.0006)	(0.0010)	(0.0006)	(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***	0.1120***	0.0431***	0.1393***	0.0356**	0.1156***
	(0.0157)	(0.0268)	(0.0161)	(0.0313)	(0.0172)	(0.0342)
Ward density of own ethnic groups^2	-0.0011**	-0.0020**	-0.0011**	-0.0028***	-0.0011**	-0.0028***
	(0.0005)	(0.0008)	(0.0005)	(0.0010)	(0.0005)	(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***	0.0999***	0.0437**	0.1247***	0.0346*	0.1027***
	(0.01562)	(0.0268)	(0.0175)	(0.0316)	(0.0187)	(0.0346)
Ward density of own ethnic groups^2	-0.0007	-0.0016*	-0.0010**	-0.0023**	-0.0008	-0.0024**
	(0.0005)	(0.0008)	(0.0005)	(0.0010)	(0.0005)	(0.0011)
Control set for all specifications						
Individual controls	*no*	*no*	*yes*	*yes*	*yes*	*yes*
Contextual controls	*no*	*no*	*no*	*no*	*yes*	*yes*

Notes. Bivariate probit model estimation results. Marginal effects and robust standard errors in parentheses are reported.
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Maximum likelihood results
-Structural approach-

	Whole sample		Muslim sample	
	Endogenous identity model	*Endogenous location model*	*Endogenous identity model*	*Endogenous location model*
c	0.8934*** (0.0706)	0.4352 *** (0.0733)	0.0355 (0.1293)	0.2162 *** (0.0188)
α	-0.3485 (0.2963)	0.7161 *** (0.2152)	2.8455 *** (0.2656)	1.6525 *** (0.0570)
γ_1	-1.6156 *** (0.1310)	-1.6272 *** (0.1855)	1.5372*** (0.4439)	1.6548 *** (0.1065)
γ_2	1.0006 *** (0.2435)	0.7334*** (0.1786)	0.9828** (0.4105)	0.8321 *** (0.0764)
Control set				
Individual variables	*yes*	*yes*	*yes*	*yes*
Contextual variables	*yes*	*yes*	*yes*	*yes*
Likelihood value	-1288.972	-1369.097	-402.489	-446.066