Religious Intermarriage and Socialization in the U.S.

Alberto Bisin NYU
Giorgio Topa NYU
Thierry Verdier DELTA, CEPR

December 2002*

Abstract

This paper presents an empirical analysis of a choice-theoretic model of cultural transmission. In particular, we use data from the General Social Survey to estimate the structural parameters of a model of marriage and child socialization along religious lines in the U.S.

The observed intermarriage and socialization rates are consistent with a strong preference of Protestants, Catholics and Jews for having children who identify with their own religious beliefs.

In contrast with the predictions obtained by various linear extrapolations from current intermarriage rates, the simulations of the model using the estimated parameters do not support the 'triple melting pot' hypothesis, nor the vanishing of American Jews. Depending on the initial conditions, we identify, in fact, two attractive stationary states of the distribution of the population by religious groups: one in which the population is composed of only Jews, and the other in which the population is composed of a majority of Protestants and a minority of the residual group, "Others", which includes mostly individuals with preferences for "no religion".

Keywords: Religion, Marriage, Socialization.
JEL: I2, Z1, D9

*The paper has benefited substantially from a spirited discussion during the presentation of some preliminary results at Minnesota. Many thanks go to M. Arellano, G. Becker, V.V. Chari, T. Conley, C. Flinn, A. Foster, J. Hotz, L. Iannaccone, M. Keane, Y. Kitamura, S. Levitt, C. Meghir, C. Mulligan, K. Munshi, A. Merlo, A. Moro, M. Rosenzweig, F. Vella, G. Violante, to many seminar audiences, and to Pierre Andre' Chiappori and two anonymous referees. J. Scheinkman's encouragement and help have been especially useful. Ryo Nakajima and Raquel Bernal provided excellent research assistance. Hospitality from the University of Wisconsin at Madison, the Federal Reserve Bank of Minneapolis, and the University of Minnesota at various stages is gratefully acknowledged.

1 Introduction

Since the 1950s a rich sociological literature has documented very low intermarriage rates along the religious dimension in the U.S. (see e.g. Landis (1949) and Thomas (1951)) and predicted, as a consequence, a low rate of assimilation of immigrants in the U.S. For instance,W. Herberg noticed in his classic (1955) study of inter-faith marriages that nothing seemed to suggest the assimilation of immigrants in the U.S. into a "melting pot" extending across the "three great faiths" (Protestant, Catholic, Jews). R. J. Kennedy (1952), facing similar evidence of low intermarriage rates in New Haven in the period 1870-1950, introduced the metaphor of the convergence to a "triple melting pot" along the religious dimension.[1]

More recently, though, a documented rise in intermarriage unions for Jews (see the Council of Jewish Federations's 1990 National Jewish Population Survey) has spurred an intense debate about the prediction of the vanishing of American Jews (see Dershowitz (1996) for a detailed report on the debate and Chiswick (1999) for a critical discussion of such predictions in the context of a human capital accumulation model of religion).

Most predictions of this sort are obtained by simple linear extrapolation of demographic and sociological trends, assuming constant intermarriage rates in the future. The most sophisticated sociological analyses of the dynamic implications of data on interfaith marriage rates account for the distribution of the population by religious group. These analyses assume that a member of a minority religious community finds it more difficult to meet a spouse who shares his/ her religious faith (Heer-Hubay (1975) and Johnson (1980) are examples of such analyses). By conditioning on the distribution of the population, these studies estimate, for members of each religious group, an unobserved component of their marriage choices, called 'intrinsic homogamy' or 'segregation effort', which drives homogamy rates. The studies then construct linear extrapolations of the dynamics of the distribution of the population maintaining these components constant.

But, if intrinsic homogamy or segregation effort are the result of the choices of individual agents in the marriage market, then they should depend on the distribution of the population by religious trait. For instance, individuals in a minority religious community might compensate for their status by segregating in marriage more intensely. In this case, extrapolating from estimated measures of intrinsic homogamy or segregation effort would underestimate the resilience of minority religious groups.

To evaluate the empirical relevance of the dependence of marriage choices on the distribution of the population by religious group, we construct a model of marriage

[1]In fact, various historical evidence on the behavior of immigrants in the U.S. seems to imply that their assimilation into a "melting pot" has progressed very slowly along several dimensions other than the religious dimension: Glazer-Moynihan (1963), for instance, in a celebrated study of the five major ethnic groups in New York City, document the "distinctive economic, political and cultural patterns" retained by such groups long after their immigration to the U.S. (see, also e.g., Gordon (1964)); Mayer (1979)).

segregation along religious lines (religious homogamy). Segregation effort is determined endogenously by the institutional characteristics of the marriage market and by the preference parameters of individuals by religious group. We estimate these parameters for U.S. survey data, over the period 1972-96, and simulate the dynamics of the distribution of the population by religious group at the estimated parameter values.

In light of our estimation and simulations, we conclude that linear extrapolations of intermarriage rates, even after conditioning on the distribution of the population, are severely misleading. The dependence of marriage rates on the distribution of the population by religious trait displays in fact substantial non-linearities. Once such non-linearities are taken into account, the simulations do not appear to support the triple melting pot hypothesis. Also, minorities do, in fact, segregate in marriage more intensely than majorities, and they socialize their children more strictly. As a consequence, Jews do not necessarily vanish by assimilation.[2] Depending on the initial conditions, the distribution of the population converges to either a stationary state composed of only Jews, or a stationary state composed of a majority of Protestants and a minority of the residual group, 'Others', which include individuals with no religious preference and members of different religious sects.

We proceed with a more detailed summary of the model, the estimation, and the results. Our analysis of the determinants of marriage rates across the religious dimension is based on Becker's early contributions on the economics of marriage (1973, 1974, 1981). Becker shows that positive assortative marriages, or 'marriage of the likes', arise as equilibria 'when such pairings maximize aggregate [...] output over all marriages, regardless of whether the trait is financial [...], biological [...], or psychological' (*A Treatise of the Family* (1981), page 70-1).[3] Many reasons can be given along these lines for the religious assortativeness of marriage in the U.S., not the least of which is that homogamous marriages are more stable, i.e., they have lower divorce rates (Becker et al. (1977), Heaton (1984), Lehrer-Chiswick (1993)). We evaluate one particular explanation of the assortativeness of marriage along the religious dimension – an explanation that emphasizes the link between marriage choices and the socialization of children to their parents' religious beliefs.

In our model, parents have a taste and a technology for transmitting their own religious faith to their children. Moreover, families that are homogamous with respect

[2] The sociological literature contains several instances of the underestimation of the resilience of minorities. The transformation of ethnic neighborhoods in the '60s has led, for example, several sociologists to extrapolate from the demographic trend and predict the rapid and complete assimilation of Orthodox Jews to American cultural values. Such predictions have proved counterfactual already in the '70s (see Mayer (1979) for a severely critical account of such predictions).

[3] Since then positive assortative mating has been documented for many traits, including intelligence, education, age, nonhuman wealth, ethnic origin, and several others (see e.g., Vandenberg (1972), Mancuso (1997), Mare (1991), Pencavel (1998)).

to their religious beliefs are endowed with a more productive technology to socialize children to such beliefs. Marriage choices are then motivated by the desire to socialize children, and will result in assortative marriage pairs along religious lines.

Intermarriage rates as well as socialization rates are therefore not only a consequence of social interactions of children but also in part a consequence of individual choices. Our empirical analysis confirms that the choice theoretic framework is important to fit the observed intermarriage and socialization rates in the U.S. An alternative model in which the marriage component is modelled as an economic decision problem and socialization is exogenously determined does not fit the data nearly as well. Nor does a model in which both marriage and socialization are exogenously determined.

The main structural parameters of the model are the 'relative intolerance' parameters, which are preference parameters defined as the perceived utility gains parents of religious group i derive from offspring of religion i rather than j. We estimate such parameters by matching the empirical frequencies of religious intermarriages (e.g., Protestant-Catholic, Protestant-Jews, etc.) and the empirical socialization rates with those implied by our model via a simple minimum distance procedure.

The observed marriage and socialization patterns are consistent with a strong preference by members of each religious group for having children who share their own religious trait. The estimated relative intolerance parameters are significant and in several cases asymmetric. For instance, the intolerance parameter of Protestants with respect to Catholics and the one of Catholics with respect to Protestant are not significantly different, while we estimate for Jews a much higher intolerance parameter with respect to Catholics than vice versa.

The socialization pattern in the data contains a bias in favor of the residual group, 'Others', for those children of all religious groups who are not directly socialized in the family. Such bias is consistent with relatively high rates of conversion from the three major religious faiths into the group of individuals with preference for "no religion" and for other religions.

The simulated choices of the individuals' marriage segregation effort and the parents' socialization effort for each religious group, based on the estimated structural parameters, are shown to be inefficient, but not severely so. The aggregate welfare loss due to the strategic interaction of the different religious groups is of the order of 2 percent of the segregation and socialization costs borne by the agents at equilibrium.

Moreover, marriage segregation and socialization effort share the same qualitative non-linear pattern in the simulations. When a group is a minority, marriage segregation and socialization efforts are increasing in the group's population share. This is because the estimated costs of socialization and marriage segregation are substantial for a minority. As a group grows towards being a majority, marriage segregation and socialization efforts become decreasing in the group's population share. This is because when a group population share is high, social interactions favor homogamy and socialization, independent of the explicit effort of individuals and parents.

As a consequence of such non-linearities, extrapolations from demographic and sociological trends are potentially severely misleading in their conclusions about the religious dynamics of the population. Although with our data such extrapolations in fact partially reproduce the triple melting pot prediction (revised to account for the supposed vanishing of the Jewish population due to their recent intermarriage behavior), our simulations based on the parameter estimates of the structural model paint a very different picture.

We estimate very high intolerance towards Catholics for all other religious denominations and, as a consequence, Catholics are never present in the stationary distributions. We also never see Jews and Protestants coexisting in the limit distribution.

We find two different stationary distributions of the population by religious trait, which are attractive for different initial conditions. One has a large (over 90 percent) majority of Protestants and a minority of the residual group, Others. The other is uniquely composed of Jews. Moreover, we show that the initial proportions of Protestants and Catholics, in large part, determine the dynamics of the share of Jews and of Others. In fact, Catholics have a very low estimated intolerance level towards Jews, while Protestants have lower intolerance towards Others than Jews. Thus, when Protestants are a high majority in the initial conditions, Jews tend to decline and Others gain a small but stable share of the population. When, instead, Catholics are well represented in the initial distribution, Jews are favored and their share rises.

More generally, our analysis is perhaps of some methodological value for empirical analyses of economies with social interactions.[4] Our implementation of the tests proposed by Vuong (1989) and Kitamura (2000) to compare non-nested models, for example, can be of general interest to evaluate the economic explanations of social phenomena vs. non choice-theoretic sociological analyses of the same phenomena. Also, we produce a general heuristic approach to the identification and computation difficulties of economic models in which social interactions give rise to multiple equilibria.

We only deal in this paper with marriage and socialization patterns along the religious dimension. We know of no other work in the economic and sociological literatures that aims at assessing, in a structural environment, the relevance of a cultural trait in the marriage market, and simulates the dynamics of such traits in the population. Religious traits offer a particularly appropriate set of observations for the general analysis of cultural traits (e.g., ethnicity, race, etc.) because *i)* religious traits are relatively well defined and measured (better e.g. than ethnicity); *ii)* they represent cultural traits that most families are keen to transmit to their children (see e.g., Glazer (1997)); and finally, because *iii)* the families' incentives to transmit their religious trait are not much obscured by related economic incentives, as religious beliefs have relatively minor effects per se in the U.S. on the agents' economic opportunities (but see Warren (1970)).

[4]See Brock-Durlauf (1999, 2001) and Glaeser-Scheinkman (2000, 2001) for general theoretical and empirical frameworks to study such economies.

The paper is organized as follows. Section 2 presents a brief description of intermarriage and socialization patterns in the U.S. Section 3 discusses the economic model. Section 4 presents the empirical methodology, and Section 5 the identification procedure. The estimation results are reported in Section 6. Section 7 performs tests of alternative model specifications to verify the endogeneity of marriage and socialization. Section 8 analyzes the dynamic implications of our estimates for the long-run distribution of the population by religious group. We then discuss several potential issues in our analysis (such as migration and unobserved heterogeneity) in Section 9. Section 10 concludes.

2 Intermarriage and Socialization in the U.S.

High homogamy rates along religious and ethnic dimensions in the U.S. are well documented. Using General Social Survey (GSS) data with U.S. states as geographic units, cumulated over the period $1972-96$, Figure 1 documents the sample probability that a member of a specific religious group marries homogamously. For all religious groups in the sample (Protestants, Catholics, Jews, and the residual Others), this probability is significantly higher than that implied by random matching (which would require all observations on the 45^o line).[5] Such marriage patterns, strongly positively assortative along the religious dimension, are characteristic of the whole of the U.S.

Religious socialization rates are also quite high in the U.S., and especially so for homogamous couples. This is documented, for the GSS data set, in Table 2. The probability that a Protestant parent has a Protestant child is, for instance, 92 percent in a homogamous marriage, while it is only 51 percent in a heterogamous marriage with a Catholic spouse.[6]

Can the high socialization rates associated with homogamous marriages actually explain the high homogamy rates that are observed in the U.S. ? This would be the case if parents valued having children who share their own individual beliefs. While several considerations other than the socialization of children affect actual marriage choices, substantial evidence points to the desire to socialize children as an important determinant of homogamy. Psychological studies of heterogamous couples consistently

[5]Table 1 reports marriage tables for 23 U.S. states, again from the same GSS data. For instance, in Tennessee the sample probability of homogamous marriage for Catholics is 66%, even though they represent only 4% of the sample population. This probability is 91% for Jews in Illinois, whose population represents less than 2% of the sample population of the state. Similarly, the probability that a Catholic marries a Protestant in South Carolina, where Protestants represent 87% of the population, is only 33%.

[6]The fact that homogamous marriages are more effective in socializing children along the religious dimension than heterogamous ones has been well documented in the sociological literature. For instance, Hoge and Petrillo (1978) and Ozorak (1989) report that children of mixed religious marriages have weaker religious commitments than those of homogamous ones; also, children of mixed religious marriages are less likely to conform to parental practices, such as church attendance or prescribed fertility behavior (see Heaton (1986); Hoge, Petrillo and Smith (1982)).

report the partners' concern about possible cultural attitudes of children when deciding to form a family (see, e.g., Mayer (1985), and Smith (1996)). Similarly, anthropological evidence points to the cultural identity of the children as a determinant of marriage choice (see e.g., Riesman-Szanton (1992)). Also, the documented fact that cohabitations are both much less fertile and less homogamous than marriages can be interpreted as evidence that homogamy matters mostly for fertile unions (see Rindefuss-Van den Heuvel (1990) for relative fertility of cohabitation, and Schoen-Weinick (1996) for homogamy rates). Finally, most major religious denominations severely regulate inter-marriages, often explicitly citing the difficulties of socialization as the main justification.[7]

Some indirect evidence from the GSS data set also supports our view that the desire for socialization explains in part the high homogamy rates along religious lines. In fact, homogamy rates are higher for young couples (less than 25 years of age at marriage), who are more fertile in expectation, and for effectively more fertile couples (with more than one child), who are also possibly more fertile in expectation when they get married (Table 3). Fertility rates are also higher for homogamous couples, for any religious group (except the residual group, Others), as to be expected if socialization drives in a relevant way homogamy rates (Table 4).

The analysis of marriage and socialization which follows will provide further evidence of the relationship between socialization and marriage along religious lines that we have suggested in this section.

3 The Model

The marriage and cultural transmission model we study is an extension of the model introduced by Bisin-Verdier (2000) to study the transmission of ethnic and religious traits.[8]

Parents have a taste and a technology for transmitting their own religious faith to their children. Compared to parents in heterogamous marriages, parents in homogamous marriages have a better technology to socialize their offspring to their own trait. As a consequence, homogamous unions have a higher value than heterogamous ones, and

[7] For example, the 1983 Code of Canon law for the Catholic Church says: "Without the express permission of the competent authority, marriage is forbidden between two baptized persons, one of whom was baptized in the Catholic Church ... and the other of whom is a member of a Church ... which is not in full communion with the Catholic Church" (801). Moreover, the permission cannot be granted unless the following condition is fulfilled: "the Catholic party declares that he or she is prepared to remove dangers of falling away from the faith and makes a sincere promise to do all in his or her power to have all children baptized and brought up in the Catholic Church."

[8] Early cultural transmission models are in Boyd-Richerson (1985) and Cavalli Sforza-Feldman (1981). See the discussion in Bisin-Verdier (2000) for a comparison.

The economic approach to the study of religion has been pioneered by Iannaccone; see his (1990) survey.

agents are willing to spend effort to segregate into a restricted marriage pool where they are more likely to meet prospective spouses of the same religious faith. The preference for socialization therefore drives the marriage choice of the agents.

We first introduce our modelling of the institutional structure of the marriage market (Section 3.1). We then introduce the socialization technology with which we assume families are endowed (Section 3.2). Finally, we study the decision of individuals regarding marriage segregation by religious group, and the decision of parents regarding the socialization of children (Section 3.3).

3.1 The Marriage Market

Let $i, j, k = 1, \ldots, n$ index different religious groups. All agents adhering to religious group i are ex-ante (before marriage, that is) identical. Fix a geographic unit of reference, e.g., a state in the United States. Let q^i denote the fraction of the population in the geographic unit (we do not keep track of the latter in the notation for simplicity) adhering to religious group i. Clearly, $\sum_{i=1}^n q^i = 1$. Let also $q = [q^1, \ldots, q^n]$. Let π^{ij} denote the probability that a member of the religious group i in the geographic unit is married to a member of the religious group j. Let α^i denote the probability that an agent of religion i in a geographic unit marries homogamously (with another member of religious group i) in a restricted (religiously segregated) marriage pool, where α^i is chosen by each agent of group i. Also let $\alpha = [\alpha^1, \ldots, \alpha^n]$.

The matching process can be defined as follows. Agents of an arbitrary religious group i first have a marriage draw in the restricted marriage pool. With probability α^i they are married there (all marriages in the restricted pool are homogamous). With probability $1-\alpha^i$ they are not married in the restricted pool, and hence they marry in the common pool, which is formed of all agents who are not married in their religious groups' respective restricted pools. Marriage in the common pool is by random matching, and hence, for instance, the probability that an agent of group i is married homogamously in the common pool (conditionally on not having found a marriage partner in the restricted pool) is $\frac{(1-\alpha^i)q^i}{\sum_{j=1}^n (1-\alpha^j)q^j}$.

We write then,

$$\pi^{ii} = \alpha^i + (1-\alpha^i)\frac{(1-\alpha^i)q^i}{\sum_{k=1}^n (1-\alpha^k)q^k} \quad i = 1, ..., n; \tag{1}$$

$$\pi^{ij} = (1-\alpha^i)\frac{(1-\alpha^j)q^j}{\sum_{k=1}^n (1-\alpha^k)q^k},\ i \neq j. \tag{2}$$

While very stylized, this marriage model does represent a rich statistical model of marriage, and hence does not impose many restrictions on the data *per se*.[9] Estimating

[9]Johnson (1980) also estimates a statistical model of religious intermarriage.

with the GSS data the probability of marrying homogamously in the restricted pool of each religious group i, α^i, for each U.S. state, so as to match the marriage tables (the realized homogamy and heterogamy rates in Table 1) produces an almost perfect fit. The p-value of the Sargan test for this estimate is 0.9966 (Figure 2 displays both empirical and simulated moments for the homogamy rates only). The restrictions imposed by our analysis on the data are those implied by our modelling of the decision to enter the restricted pool as made by rational agents, rather than those implied by the postulated institutional structure of the marriage market.

3.2 The Socialization Technology

Socialization and cultural transmission occur in the family and in society at large (as a consequence of social contact with peers, role models, etc.). Families are indexed by pairs ij, where i and j indicate the religious group of each parent. Let P_{ij}^i denote the probability that a child of a family of type ij has religious trait i. The socialization (cultural transmission) mechanism is as follows:

- A child from a religious homogamous family of type ii is directly socialized to the trait of the family with probability $\widetilde{\tau}^i \equiv \tau^i + m$, where τ^i is chosen by the parents while m is an exogenous probability independent of parents' effort (exogenous direct socialization). With probability $(1 - \widetilde{\tau}^i)$, he/she is not directly socialized by the family, and he/she picks a trait by matching in the population with a cultural parent who socializes the child to his/her own religious trait.

- A child in a heterogamous family of type ij, $i \neq j$, does not have a well-defined reference religious trait to be socialized to, and as a consequence he/she picks each parent's trait independently with some exogenous probability $\frac{m}{2}$. If not socialized by either parent, a child in an heterogamous family picks a trait by matching in the population with a cultural parent who socializes the child to his/her own religious trait.

This mechanism embeds two assumptions. One is that parents in homogamous unions have a better socialization technology than parents in heterogamous ones. The other is that children can acquire a given trait either through a "vertical" socialization process from their parents, or through an "oblique" socialization process from society at large. Substantial sociological evidence supports in fact such assumptions regarding the socialization mechanism. See Clark-Worthington (1987), Cornwall (1988), DeVaus (1983), Erickson (1992), Hayes-Pittelkow (1993), Hoge-Petrillo-Smith (1982), Ozorak (1989), for some direct evidence on religious socialization and Wilson (1987) for the sociological literature on adult role models.

The process of matching with cultural parents in the population, which determines the oblique socialization process, is not directed by agents' choices, and might in principle be biased in favor of some particular religious groups. This is essentially because of conversions. While Catholics and some Protestant denominations do actively proselytize (and Jews do not), it turns out that this phenomenon has a substantial role in our analysis only for the residual group, Others, which includes (in large part) individuals with preference for "no religion" and members of different religious sects.

Let Q^i denote the probability that any child, not socialized in the family, meets with a cultural parent of religious group i in the population, and hence is socialized to religion i. Obviously, $\sum_i Q^i = 1$. Note that random matching requires $Q^i = q^i$; when $Q^i > q^i$ conversions to religion i affect the matching process of children with cultural parents. The distance of such matching process from random matching will be estimated.

For any $i, j, k = 1, \ldots, n$, the socialization equations for homogamous families can be written as:

$$P_{ii}^i = \widetilde{\tau}^i + (1 - \widetilde{\tau}^i) Q^i \tag{3}$$

$$P_{ii}^j = \left(1 - \widetilde{\tau}^i\right) Q^j, \ i, j \text{ distinct} \tag{4}$$

and the socialization equations for heterogamous families as:

$$P_{ij}^i = \frac{1}{2} m + (1 - m) Q^i, \ i, j \text{ distinct} \tag{5}$$

$$P_{ij}^k = (1 - m) Q^k, \ i, j, k \text{ distinct} \tag{6}$$

3.3 Marriage Segregation and Socialization Choices

Each agent of religion i chooses α^i, the probability of being matched in the restricted pool (where all mates are of trait i). The cost associated with α^i, when the share of religious group i in the population is q^i, is $M(\alpha^i, q^i)$. The advantage of marrying homogamously is that it gives one the option of directly socializing one's children. The direct socialization rate in a homogamous marriage of religion i, τ^i, is chosen by parents. The cost associated with direct socialization τ^i, when the share of religious group i in the population is q^i, is denoted $S(\tau^i, q^i)$. The benefit of socialization derives from the fact that parents want their children to share their own religious faith. The value for a parent of type i of a child of type j, V^{ij}, is exogenously given and needs to be estimated. In this regard, we postulate $V^{ii} \geq V^{ij}$.

For clarity's sake, we analyze the model backwards. We first study the choice of direct socialization τ^i by homogamous parents of religion i. We then study α^i, the marriage choice of agents of religion i.

The direct socialization of one child of a homogamous family of type ii, τ^i, is the solution of the following maximization problem:

$$\max_{0\leq\tau^i\leq 1} P_{ii}^i V^{ii} + \sum_{j\neq i} P_{ii}^j V^{ij} - S\left(\tau^i, q^i\right) \tag{7}$$

subject to (3-6).

We do not write any explicit endogenous fertility problem for the agents. This is essentially because one extra optimization problem would make the model intractable. We assume that agents take as given a constant fertility rate of a marriage of type ij, denoted n^{ij}. Then, for an agent of type i, let $(n^{ij})^{\xi} W^{ij}$ denote the value of a marriage with an agent of type j, given that the marriage produces n^{ij} children. The parameter ξ denotes the dependence of the parents' preferences on the number of children in the marriage. For an agent of type i, the per child value of a marriage with a j spouse , W^{ij}, is the expected value of a child in such a marriage:

$$W^{ii} = P_{ii}^i V^{ii} + \sum_{j\neq i} P_{ii}^j V^{ij} - S\left(\tau^{*i}, q^i\right) \tag{8}$$

$$W^{ij} = P_{ij}^i V^{ii} + \sum_{k\neq i} P_{ij}^k V^{ik}, \;\; i\neq j \tag{9}$$

We are implicitly assuming that all children in a marriage are socialized to the same trait. This is just for simplicity and does not change any of our results.

When choosing the probability of being married in the restricted marriage pool, all agents take the composition of the common pool as given, as each agent is infinitesimal and hence does not affect the composition. In equilibrium the composition of the common pools will be required to be consistent with all of the agents' choices.

Let A^{ii} denote the probability of an agent of type i marrying homogamously in the common pool, and let A^{ij} denote his/her probability of marrying an agent of type j in the common pool. Let $M\left(\alpha^i, q^i\right)$ denote marriage segregation costs. The marriage problem of an agent of type i is:

$$\max_{0\leq\alpha^i\leq 1} \pi^{ii}\left(n^{ii}\right)^{\xi} W^{ii} + \sum_{j\neq i} \pi^{ij}\left(n^{ij}\right)^{\xi} W^{ij} - M\left(\alpha^i, q^i\right) \tag{10}$$

subject to

$$\pi^{ii} = \alpha^i + (1-\alpha^i)A^{ii}$$

$$\pi^{ij} = (1-\alpha^i)A^{ij}, \;\; j\neq i$$

given A^{ii} and $A^{ij}, \forall j\neq i$.

In equilibrium,

$$A^{ii} = \frac{(1-\alpha^i)q^i}{\sum_{k=1}^{n}(1-\alpha^k)q^k}$$

and

$$A^{ij} = \frac{(1-\alpha^j)q^j}{\sum_{k=1}^{n}(1-\alpha^k)q^k}.$$

The reduced form equations of the structural model just introduced are reported and studied for clarity in Appendix 1.

4 The Empirical Implementation

We consider the following religious groups: Protestants, Catholics, Jews, and the residual group, 'Others'. The latter includes individuals with preference for "no religion" as well as individuals with preferences for other religious faiths. Moreover, we assume that the geographic unit of reference for marriage and socialization processes coincides with the state. In other words, we assume that the composition of the population by religious group, which is relevant for each agent in the marriage market and for each family when choosing the direct socialization levels, is the composition of the population in the U.S. state in which the agent or the family resides. The definition of both the religious groups and the geographic unit of reference in the analysis are arbitrary and determined essentially by the available data. We discuss the possible problems associated with such definitions, and the several attempts we made at assessing the robustness of our empirical results to different definitions of groups and geographical units, in Section 9.

In this section, we first make the model introduced in the previous section operational by introducing relevant assumptions and the necessary functional form parametrization. We discuss identification of the model parameters, and we briefly present the data we use in the estimation. We then construct a map from the parameters of the model, the composition of the population, and fertility rates by religious group, into intermarriage and socialization rates, to be matched with those that we observe in the data. (Some mathematical properties of this map are studied in Appendix 1.) Finally, we introduce an appropriate estimation procedure.

Index each state by s. Let $\Delta V^{ij} \equiv V^{ii} - V^{ij}$, for any i, j (obviously, $\Delta V^{ii} = 0$). ΔV^{ij} measures the perceived increment in utility for an agent of type i associated with having a child of type i rather than j: we refer to it as the "intolerance" of agents of type i towards group j.

We parametrize the cost functions by:[10]

$$S(\tau_s^i, q_s^i) \equiv [\sigma_\tau + \varepsilon_\tau(1-q_s^i)^2]\left[\lambda_\tau \frac{(\tau_s^i)^2}{2} + (1-\lambda_\tau)\left[\exp(\frac{\tau_s^i}{1-\tau_s^i}) - 1\right]\right] \quad (11)$$

$$M(\alpha_s^i, q_s^i) \equiv [\sigma_\alpha + \varepsilon_\alpha(1-q_s^i)^2]\left[\lambda_\alpha \frac{(\alpha_s^i)^2}{2} + (1-\lambda_\alpha)\left[\exp(\frac{\alpha_s^i}{1-\alpha_s^i}) - 1\right]\right] \quad (12)$$

We also parametrize the matching probabilities in the oblique socialization as:

$$Q_s^O = \frac{q_s^O + o}{1+o}; \; Q_s^i = \frac{q_s^i}{1+o}, \; i = P, C, J. \quad (13)$$

The parameter o represents the deviation away from pure random matching, in the matching process which determines the socialization of children in society at large. This deviation could be explained as the effect of the conversion rate into the residual group. We find no evidence in the data of such conversions into our main religious groups, Protestants, Catholics, and Jews.

Our crucial identifying assumptions are that (i) cost functions are not specific to any religious group, nor to the geographic units of reference (the states); (ii) the intolerance parameters, ΔV^{ij}, are naturally specific to the religious groups (i, j), but independent across states; and (iii) the socialization bias due to conversions, o, as well as the exogenous direct socialization rate, m, are constant across the geographic units of reference (the states).

The parameters of the model consist of the intolerance parameters ΔV^{ij}, for any i and $j \neq i$, the parameters which describe the cost functions, $(\sigma_c, \varepsilon_c, \lambda_c)$, with $c = \alpha, \tau$, the bias due to the conversions to 'Others' in the oblique socialization process, o, the exogenous direct socialization rate, m, and the preference for fertility, ξ. Let θ denote the vector of parameters.

We use data from the General Social Survey (GSS), covering the period $1972-96$, on the composition of marriages by religious affiliation of the spouses for each state; the composition of the population by religious group for each state; the socialization rates by religious affiliation of the spouses aggregated over the U.S.; and fertility rates by religious group of the spouses aggregated over the U.S. According to our notation, we have data on π_s^{ij}, for all i, j and s; q_s^i, for all i and s; $P_{ij}^k = \sum_s \omega_s^{ij} \cdot P_{ij,s}^k$, where ω_s^{ij} are sample weights representing the percentage of respondents in an ij marriage that live in state s, for all i, j, k; and n^{ij}, for all ij. (We do not have enough data to construct accurate empirical frequencies of socialization rates, $P_{ij,s}^k$, for each state s.)

Given the values of homogamous and heterogamous marriage unions, (W^{ii}, W^{ij} respectively) an equilibrium in the marriage market is a solution to the fixed point problem of (28-30) in Appendix 1 and (12). While an equilibrium always exists, there is

[10] The parametrization satisfies the convexity and Inada conditions which are necessary for our analysis; see Assumption 1 in Appendix 1.

no guarantee that the equilibrium is unique for general cost functions $M(\alpha^i, q^i)$ as in (12). Multiple equilibria are the consequence of the coordination problem implicit in our formulation of the marriage market. Suppose that under the parameters of the model two religious groups aim at segregating in the marriage market. The same segregation pattern can be achieved if agents of group i choose high α^i and agents of group j choose low α^j, as well as if vice versa it is group i that chooses a low α^i and group j chooses a high α^j. In the first case, agents of group j can segregate in the residual pool, which is composed mainly of agents of group j thanks to the high segregation effort of the other group, while in the second case it is agents of group i who can segregate in the residual pool. Such different segregation patterns have important distributional effects (the group segregating in the residual pool is favored, as the costs to enter the restricted pool are saved), but homogamy rates for the two groups can remain unaffected.[11]

Given q^i_s, for all i and s, and n^{ij} for all i, j, the structural model (represented by equations (21-30) in Appendix 1 and (11-13)) defines a mapping, $\widetilde{\Pi}(\theta)$, from θ into π^{ij}_s, for all i, j and s, and into P^k_{ij} for all i, j, k. We use a minimum distance estimation procedure, that matches the vector $\widehat{\Pi}$ of empirical moments $\left(\widehat{\pi}^{ij}_s, \widehat{P}^k_{ij}\right)$ from the data with the vector $\widetilde{\Pi}(\theta)$ of moments implied by the model for a given choice of θ. Formally, given a square weighting matrix Ω^*_N (where N denotes the total sample size), the minimum distance estimator $\widehat{\theta}$ minimizes

$$J_N(\theta) \equiv [\widehat{\Pi} - \widetilde{\Pi}(\theta)]^\top \Omega^*_N [\widehat{\Pi} - \widetilde{\Pi}(\theta)]. \tag{14}$$

Possible discountinuities of the map $\widetilde{\Pi}(\theta)$ may be problematic for various reasons. First, standard consistency proofs usually require continuity of the criterion to be minimized (and hence of $\widetilde{\Pi}(\theta)$). However, it is easy to show that *local* continuity at the global minimum of $J_N(\theta)$ is sufficient for consistency.

Second, in order to compute standard errors, one needs to ensure that $\widetilde{\Pi}(\theta)$ is locally smooth at $\widehat{\theta}$, and hence that the partial derivatives $\frac{\partial \widetilde{\Pi}(\theta)}{\partial \theta^\top}$ are well defined. We check that this requirement is satisfied in a neighborhood of $\widehat{\theta}$. As long as $\widehat{\theta}$ is indeed the global minimizer of the criterion, this is sufficient for local continuity as well.

Finally, discontinuities in $\widetilde{\Pi}(\theta)$ typically make it much harder for standard minimization algorithms to find the global minimum of $J_N(\theta)$. However, we use a Simulated Annealing algorithm that is especially well suited for problems where the objective function may have various discontinuities and/or several distinct local optima. A more detailed description of the data and the estimation methodology, as well as the Simulated Annealing algorithm, is contained in Appendix 2.

[11] In Figure 9 we plot a pair of best-reply functions for the α^i restricted pool marriage probabilities. We fix Jews' and Others' religious shares at their mean values and plot the Protestant and Catholic best-reply functions to each other's α^i, while keeping the marriage segregation probabilities for Jews and Others at their equilibrium levels. We repeat the exercise for different combinations of Protestant and Catholic religious shares. The plots clearly indicate the presence of non-convexities in the best-reply functions that generate multiple equilibria for at least some values of religious shares.

5 Identification

We are able to identify independently the intolerance of group j with respect to group i as well as the intolerance of group i with respect to group j out of data on intermarriages. This might be surprising, as, after all, the marriage unions between individuals of group i and j are also unions between individuals of group j and i. But in our model, realized intermarriage rates depend on the segregation efforts of the two groups, which depend non-linearly on the shares of the population by religious groups. Such non-linearities, together with the variation in the population distribution of religious groups across U.S. states, can be exploited to identify asymmetric intolerance parameters.

We can make this argument more precise by means of a simple example to illustrate our identification procedure. Consider an economy with only two religious groups, e.g., Catholics (C) and Protestants (P). Assume fertility rates are constant across all family types (so that the model is independent of fertility rates). Also, assume that the exogenous direct socialization rate, m, is zero, and that cost functions are quadratic: $S(\tau^i, q^i) = \frac{1}{2}\left(\tau^i\right)^2$, $M(\alpha^i, q^i) = \frac{1}{2}\left(\alpha^i\right)^2$. Using equations (19-20), in this special case we can solve for α_s^i in each state s:

$$\alpha_s^i = (1 - A_s^{ii})\left[(1 - q_s^i)\Delta V^{ij}\right]^2$$

In this example our estimation procedure would only need to match one moment, for instance π_s^{cc}, for each state s. In fact, given π_s^{cc} and q_s^c, we can solve for π_s^{cp} using $\pi_s^{cp} = 1 - \pi_s^{cc}$, then for π_s^{pc} using $\pi_s^{cp} q_s^c = \pi_s^{pc} q_s^p$, and finally for π_s^{pp} using $\pi_s^{pp} = 1 - \pi_s^{pc}$.

Writing equation (29-30) in implicit form, $A_s^{cc} = A^{cc}(q_s^c; \alpha_s^c, \alpha_s^p)$, $A_s^{pp} = A^{pp}(q_s^p; \alpha_s^p, \alpha_s^c)$, the model is reduced to three equations in each state s,

$$\pi_s^{cc} = \alpha_s^c + (1 - \alpha_s^c)A^{cc}(q_s^c; \alpha_s^c, \alpha_s^p)$$

$$\alpha_s^c = [1 - A^{cc}(q_s^c; \alpha_s^c, \alpha_s^p)] \cdot [(1 - q_s^c)\Delta V^{cp}]^2$$
$$\alpha_s^p = [1 - A^{pp}(q_s^p; \alpha_s^p, \alpha_s^c)] \cdot [(1 - q_s^p)\Delta V^{pc}]^2$$

and four unknowns: $\alpha_s^c, \alpha_s^p, \Delta V^{cp}$, and ΔV^{pc}. Therefore, the parameters $\Delta V^{cp}, \Delta V^{pc}$ cannot be identified independently with data on q_s^c and π_s^{cc} for a particular state s. But since we restrict the "intolerance" parameters $\Delta V^{cp}, \Delta V^{pc}$ to be independent of the state s and exploit the variability of the observations of q_s^c and π_s^{cc} across S states, we face $3S$ independent equations, and $2S + 2$ unknowns, ΔV^{cp}, ΔV^{pc}, and (α_s^c, α_s^p) for each state (for $S > 2$ states the system is therefore over-identified).[12]

Finally, if π_s^{cc} non-trivially depends on both α_s^p and α_s^c, then the system of equations is locally independent.

[12]Note that in this identification procedure we crucially exploit the assumed independence of cost functions from state and religious group, when solving for α^i.

In the context of the more general model, an identification argument can be sketched along the following lines. From (27) we can write the optimal τ^i as

$$\tau_s^i = S_1^{-1}\left[\sum_h q_s^h \Delta V^{ih} \,|\, \sigma_\tau, \varepsilon_\tau, \lambda_\tau\right],\ \forall i, s. \tag{15}$$

The parameters m and o are identified from the equations for the socialization rates, (23-26). The observed π_s^{ij} pin down the α_s^i and the A_s^i. Then one can use the FOCs for α_s^i, (28) and (15), to write down a list of non-linear equations in the observed (or estimated) $(\alpha_s^i, A_s^{ii}, n^{ij}, q_s^i, m, o)$ and the remaining unknown parameters $(V^{ii}, \Delta V^{ij}, \sigma_\tau, \varepsilon_\tau, \lambda_\tau, \sigma_\alpha, \varepsilon_\alpha, \lambda_\alpha, \xi)$. These constitute nS equations in n^2+7 parameters (where n is the number of religious groups).[13] The order condition for identification is satisfied in our case, with $n=4$ and $S=23$, and the rank condition can be checked locally.

5.1 Multiplicity of Equilibria

Because of the possibility of multiple equilibria, our identification procedure must jointly identify the parameters of the model and the equilibrium selection. In terms of the minimum distance criterion, for our estimate to be consistent we need to find the value of θ that minimizes the lower envelope of the multiple $J_N(\theta)$ surfaces generated by the different equilibria.

No standard procedure for identification is available in the face of multiple equilibria.[14] We use therefore an heuristic approach to locally identify the equilibrium selection and the parameters of the model. In the course of the minimization of $J_N(\theta)$, for each candidate value $\overline{\theta}$, we let the algorithm randomly pick several distinct starting values for the iteration that yields the equilibrium, in order to try to generate several possible equilibria. We then compute J_N for each of these equilibria and use the lowest value as the value $J_N(\overline{\theta})$ for that particular value of θ. Since we cannot compute all possible equilibria for each θ due to computational limitations, this procedure is at least a step in the direction of searching for the equilibrium selection, as well as the parameter estimate, which minimize the criterion $J_N(\theta)$.

6 Estimation Results

Table 5 presents the estimation results of the structural marriage and socialization model introduced in Section 3 with the GSS data.

[13] In the actual estimation we restrict V^{ii} to be the same across all groups i, therefore the number of parameters is reduced to $n(n-1)+8$.

[14] But, see Dagsvik-Jovanovic (1994). Moro (2000) has ingeniously introduced a procedure which allows the local identification of a specific model of statistical discrimination with multiple equilibria. Moro's procedure cannot be simply adopted in our set-up.

The model fits the intermarriage data quite well, whereas it fits the socialization data less well. The p-value of the Sargan test of the over-identifying restrictions is quite high (0.11) when one considers the intermarriage moments alone, but drops to about 0.017 when one considers the socialization moments as well. In order to get a visual impression of the fit, Figure 3 compares the empirical homogamous marriage frequencies, $\widehat{\pi}^{ii}$, to those generated by the model, $\widetilde{\pi}^{ii}(\theta)$, at the estimated parameter values. Table 6 compares the empirical socialization frequencies with those implied by the model. We are able to match the homogamous socialization rates quite well, whereas we do less well in matching the heterogamous ones.

The low empirical frequencies of religious intermarriages are the consequence of a strong estimated preference by members of each group for having children who share their own religious faith, that is of high intolerance parameters.[15] We estimate significant positive intolerance parameters (with the exception of the parameter describing attitudes towards Jews of the residual group, Others). The most striking estimates are those describing the intolerance parameters of Jews, which are about four times as high as those of any other religious group.[16]

The parameter estimates for the cost functions reveal a strong dependence of both socialization costs and marriage costs on the proportion of one's religious group in the state, $\varepsilon_\tau, \varepsilon_\alpha$. The more a given religious faith is a minority in the population of reference, the harder it is to socialize one's children to that particular faith or to segregate in marriage. Figure 4 plots the estimated cost functions for each group. The religious share q^i is set equal to the median share of the different groups across states: 0.667 for Protestants, 0.245 for Catholics, 0.014 for Jews and 0.068 for Others. For instance, because of the small median share of Jews in the population, the cost of directly socializing a child with probability $\frac{1}{2}$, is about twice as large for Jews as for Catholics. Figure 5 plots the cost surfaces as a function of both τ (or α) and q. Costs are increasing in τ (or α) and decreasing in q.

The matching probabilities Q_s^i in the oblique socialization are biased in favor of the

[15]Since cost functions are assumed to be independent of the specific religious group, intolerance levels can be meaningfully compared across groups. In particular, the implicit unit of measure can be identified as follows. The estimated socialization cost function of any religious group i, when the group represents half of the population, $q^i = \frac{1}{2}$, can be easily computed and takes value 1 for $\tau^i = .12$. Therefore, both costs and intolerance levels are measured as multiples of the cost of increasing the probability of socializing a child, for a homogamous family of a religious group that constitutes half the population, by 12 percentage points. While such a cost is not pinned down by the analysis, if we roughly identify it with the opportunity cost of spending two hours with the child per week for 10 years, it lies in the $20,000 - 40,000$ dollars range (for hourly wages in the $20 - 40$ dollars range).

[16]Following the rough computations in the previous footnote our estimates of the intolerance parameters for the Jews imply a relative value of a child with maintained Jewish identity in the $10 - 20$ million dollars range.

A better sense of what it means for an intolerance level to be high can be derived indirectly from the simulation of the effects of the estimated intolerances on dynamics of the distribution of the population across religious groups (Section 7).

residual group Others. The estimated bias parameter o induces a sizeable distortion: the implied probability of becoming 'Other' in society at large, once direct family socialization fails, exceeds on average the share of Others in the population by about 16 percentage points. Such bias can be accounted for by conversions. The residual group Others includes a majority of individuals with no religious preference (over 70 percent on average in the U.S. of our residual group in the sample) and a minority which includes major religious faiths not largely represented in the U.S. (Islam, Buddhism, Hinduism), and religious sects. Proselytizing activities (broadly speaking) could account for the high conversion rate implied by our analysis of the socialization data, at least in the case of individuals with no religious preferences and individuals belonging to religious sects.

Parents' preferences are not very sensitive to the number of children in a given marriage. The estimated ξ is very close to zero, implying that parents seem to care only about their average child. Since we do not explicitly model endogenous fertility, this result must be interpreted with particular caution.

The estimated choices of direct socialization of homogamous families, τ^i, which is the differential probability of direct socialization with respect to the exogenous direct socialization rate, m, and the choice of marriage segregation in the restricted pools, α^i, are quite instructive about the implications of our results for socialization and for the marriage market. Figure 6 (resp., Figure 7) presents the estimated τ^i (resp., α^i), for Protestants, Catholics, and Jews as a function of q^i.[17]

For Protestants and Catholics, the direct socialization levels of homogamous families when fully minority ($q^i = 0$) are significantly positive: τ^i is greater than .3 in both cases, and m is .35, giving a probability of direct socialization for homogamous families when they are minorities above 75 percent. If socialization costs were independent of q^i, socialization levels would decrease with q^i. As a result of the estimated strong dependence of socialization costs on q^i, direct socialization for both Protestants and Catholics first increases and then decreases, peaking at about $q^i = .5$. Jews socialize much more than Catholics and Protestants in the whole relevant range of q^i. For example, when Jews are a small minority, the probability of direct socialization levels for homogamous Jewish families is roughly 90 percent.

Similar considerations hold for the estimated marriage segregation probabilities in the restricted pool. As we have mentioned earlier, the marriage game which determines marriage segregation levels in the restricted pools exhibits multiple equilibria, for the estimated parameters. As a consequence, the estimated equilibrium marriage segrega-

[17] These plots are constructed by fixing the religious shares for two groups (e.g., Jews and Others) and letting the religious share of a specific group (e.g. Protestants) increase and the share of the residual group (Catholics) decrease in order to satisfy $\sum_{i=1}^{n} q^i = 1$. For Protestant and Catholics these are calculated by keeping the proportion of the population of Jews and Others fixed at their means; while for Jews we report calculations relative to the case in which the proportion of Catholics is fixed at its highest (resp. lowest; resp. mean) level, the proportion of Protestants is fixed at its lowest (resp. highest; resp. mean) level, and finally, the proportion of Others is fixed at its mean level.

tion levels are a discontinuous selection of the equilibrium set. For Protestants and Catholics the marriage segregation level first increases (because of the high socialization and marriage costs) and then decreases as a function of q^i. When a religious group is a small minority (i.e. when their fraction in the population is close to zero), marriage segregation in the restricted pool is about 65 percent for Catholics, and about 55 percent for Protestants. Catholics have a higher α^i than Protestants for most of the range of q^i. Jews' marriage segregation does not display much variation in the relevant range of the proportion of Jews in the population; the marriage segregation level is very high in the whole range (including when fully minority), above 80 percent.

Finally, the out-of-sample simulations of homogamous marriage probabilities, π^{ii}, implied by our parameter estimates as a function of religious share q^i, are plotted in Figure 8. When q^i is close to zero, the probability of homogamous marriage is well above that implied by random matching (which lies on the 45 degree line) – for Protestants it is around 0.55, for Catholics it is about 0.65, and for Jews it is above 0.8. This is due to the strictly positive socialization and marriage segregation levels implied by our estimates. The simulated π^{ii} is increasing in q^i and becomes close to the probability implied by random matching only when the share of religious group i in the population approaches 90 percent.[18]

As marriage segregation costs depend on the populations shares, the equilibrium in the marriage market is not necessarily efficient.[19] The same homogamy rates can in fact be generated by different segregation practices of the different religious groups, and the equilibrium does not necessarily pick the structure of segregation practices, the configuration of α^i for each group i, which minimizes the aggregate costs across groups. We have computed the efficient configuration of α^i for each state, defined as the configuration which minimizes costs under the constraint that it generate the same homogamy and heterogamy rates across groups (we report the computations for California, Illinois, Texas and New York in Table 7 as an illustration). The results indicate that the equilibria we estimate are close to efficiency in the sense that the reduction in aggregate segregation costs in going from the equilibrium to the efficient segregation practices are quite small, less than 2 percent in the highest case (Texas). However, efficient segregation practices might involve a large redistribution of the marriage segregation costs by religious group. For instance, and again in Texas, the efficient segregation practice requires a sizeable

[18] It is worth noting that the discontinuities in α^i generate only small jumps in the implied homogamous marriage probabilities π^{ii}. Multiple equilibria correspond, in fact, to different equilibrium distributions of segregation costs across the different religious groups without affecting the implied homogamy and heterogamy rates much. The segregation of one group in its own restricted pool has in fact a positive externality on the other groups as they gain higher implied homogamy rates without the need of segregating in their own restricted marriage pools.

[19] On the other hand, the socialization choice is efficient as long as individuals, as we assumed, only care about their children, and not about their whole dynasty.

increase in the marriage segregation of Others in their own restricted marriage pool and a corresponding reduction for Protestants. As a result, moving from the equilibrium to the efficient marriage segregation practices increases segregation costs for Others by 14 percentage points, and decreases the same costs for Protestants, a much larger fraction of the population, by 11 percentage points.

7 Are Marriage and Socialization Endogenous?

The model of marriage and socialization we estimate is based on the behavioral assumption that marriage and socialization are endogenously determined as economic decisions of agents who have preferences for children with their own religious attitudes. In this section we aim at assessing the relevance of economic behavior to explain the observed socialization and marriage rates. To this end, we conduct some tests on our baseline estimates, and compare the performance of our model to several alternative specifications that make different behavioral assumptions.

Results are reported in Table 8 (the first column reproduces our baseline estimate to simplify comparisons). Column 2 reports parameter estimates for a model in which marriage segregation choices are endogenous but socialization is exogenous. In particular, it reports the results of the estimation of the direct socialization effort, τ^i, assuming it varies across religions but not across states. Column 3 examines instead a model in which both marriage and socialization are exogenous. Both α^i and τ^i are estimated to vary across religions but not across states. Finally, column 4 reports estimation results for a model in which the value of a homogamous marriage ($W^{ii} - W^{ih}$) is exogenous and independent of own group's share in the geographic state. This is an attempt to capture alternative explanations of the high prevalence of homogamous marriages, where the benefits of homogamy are intrinsic in homogamous marriage unions and therefore constant across states. Examples of some alternative explanations are that the benefit of homogamous marriages resides in the possible advantages (consumption value) of sharing the same cultural representation of life and society, or in the homogamous marriages' inherent stability with respect to divorce (see Becker et al. (1977), Heaton (1984) and Lehrer-Chiswick (1993)).

The rankings of the Sargan test of the over-identifying restrictions suggests that all three alternative models do not fit the data nearly as well as our baseline model: p-values vary between 0.02 and 0.0017, compared with 0.11 in our estimate. However, a formal test comparing our baseline model with the alternative specifications we are interested in is not straightforward as the models are non-nested. We adopt therefore a procedure to compare non-nested models first introduced by Vuong (1989) and further developed by Kitamura (2000).

The procedure consists of determining the distance of each model from the true distribution that generates the data, where the distance is measured by the Kullback-

Leibler Information Criterion (KLIC). The test statistic, constructed as the difference between the minimum KLIC for each of the two models to be compared, coincides with the expected log-likelihood ratio (Vuong (1989)). The expected log-likelihood ratio, when appropriately normalized by its asymptotic variance σ^2_∞, is distributed as a standard Normal under the null hypothesis that the two models are equivalent. It diverges to $\pm\infty$ if one of the models is better than the other.[20] Kitamura (2000) extends this framework to the case of models that are identified via moment conditions. An information theoretic procedure is used to construct the non-parametric analog to Vuong (1989)'s test statistic, defined as:

$$\widehat{D}_n \equiv \sqrt{n}\left[\exp\left(-I(\widehat{\theta}_B)\right) - \exp\left(-I(\widehat{\theta}_A)\right)\right]/\widehat{\sigma}^2,$$

where $I(\widehat{\theta}_j)$ is the KLIC distance from the truth for model j.[21] Subscript B refers to the baseline model, while subscript A refers to each of the three alternative models. The null hypothesis (H_0) is that the two models are equivalent, whereas the alternative (H_1) is that one model is better than the other (in the KLIC sense). Specifically, one chooses a critical value c from the standard Normal for a given significance level. If the value of the test statistic is higher than c one rejects H_0 in favor of model B being *worse* than model A; if the test statistic is lower than $-c$, one rejects H_0 in favor of model B being *better* than model A; finally, if $|\widehat{D}_n| \leq c$, one cannot discriminate between the two competing models.

The results of the test are reported at the bottom of Table 8. Our model performs better, in the KLIC sense, than each of the three alternative specifications examined here. The p-values of the tests are practically zero.[22] It is interesting to note that the ranking of the four models is different from that implied by comparing the Sargan specification test results. In particular, the model in column 4 performs best among the alternative specifications. The estimates reported in column 4 of Table 8 refer to a model in which the subjective benefits from marrying homogamously stay constant across states and are estimated to match the observed homogamous and heterogamous marriage rates. The good fit of this model suggests that the benefits of homogamy might also contain components which are intrinsic to homogamous marriages and therefore invariant with respect to the distribution of the population by religious group.

[20] Note that, unlike in the literature on non-nested models that uses Cox-type tests (see Cox (1961), Singleton (1985), Smith (1992)), in this approach it is not assumed that one of the models is the true one, but, rather, that both models are somewhat misspecified.

[21] Technically, Kitamura's procedure is applied to Generalized Method of Moments (GMM) estimators, whereas we perform minimum distance estimation. However, we can apply his test in our context by bootstrapping. In particular, we sample $N = 1,000$ times with replacement from the empirical distribution that generated our empirical moments in order to create N artificial samples for our moment conditions. Details of our construction and of Kitamura's test can be found in an Appendix available from the authors upon request.

[22] In order to properly compare our baseline model to the model in column 4, we re-estimated our model to match only the empirical intermarriage rates $\widehat{\pi}$, since the alternative model does not have implications for the socialization rates P^k_{ij}.

We take our analysis of alternative models to suggest that endogenous socialization and marriage segregation are indeed an important part of marriage and socialization mechanisms with respect to religious trait.

8 Long-Run Dynamics of the Distribution of Religious Groups

Given the distribution of the population by religious group at some time t, the marriage and socialization mechanisms we estimated determine the distribution of the population in the successive generation, at time $t+1$.

The difference equation ruling the dynamics of the distribution of religious traits in the population is the following:

$$q^i_{t+1} = \frac{N_t}{N_{t+1}} \sum_j q^j_t \sum_h \pi^{jh}_t \frac{n^{jh}}{2} P^i_{jh,t}, \tag{16}$$

where N_t denotes the total number of adults at time t. The evolution of N_t can be obtained by studying the evolution of the number of adults for each religious group, N^i_t:

$$N_t = \sum_i N^i_t. \tag{17}$$

$$N^i_{t+1} = \sum_j N^j_t \sum_h \pi^{jh}_t \frac{n^{jh}}{2} P^i_{jh,t} \tag{18}$$

It is also easy to derive the equations which determine the stationary states of the distribution of the population by religious group (see Appendix 2 for the derivation).[23]

8.1 Simulations of the Dynamics

Using the estimated structural parameters and the empirical religious composition of several U.S. states as initial conditions, we simulate the evolution of the distribution of the population by religious group, q^i_s, over time. In particular, we use the current composition of California, Illinois, New York and Texas as initial conditions to illustrate the dynamic paths implied by our estimates. Results are reported in Figure 11.[24]

[23] The number and local stability properties of stationary states are studied, for a simple version of this economy, by Bisin-Topa-Verdier (2000).

[24] We have replicated the simulations with parameter values chosen at various edges of the confidence intervals (we do not report all such simulations, but see Figure 12 for some relevant examples). It turns out that increasing or decreasing the intolerance parameters of Catholics, Jews and Others, as well as the conversion bias, by two standard deviations has no qualitative effects on the dynamics. But, changing the intolerances of Protestants within their confidence interval does change the basin of attraction of the two stationary states. We conclude that the results of the simulations are robust with respect to variations of the estimated parameters in their confidence interval.

We find two different stationary distributions of the population by religious trait, which are attractive for different sets of initial conditions: one has a large majority of Protestants (about 90 percent) and a minority of the residual group, Others (about 10 percent); while the other is uniquely composed of Jews. The stationary state in which only Jews are represented is attractive for instance for the initial composition of the populations of Illinois and New York. The stationary state composed of Protestants and Others, is attractive for the initial conditions of California and Texas. The population settles into a stationary distribution in at most 45 periods (a period should be interpreted as a generation, that is $25 - 30$ years).

The dependence of the dynamics on the initial conditions is interesting and complex. For instance, even though the initial proportion of Jews is higher in California than in Illinois, their share rises exponentially in Illinois, while it declines quickly in California. This is because in Illinois Catholics are well represented in the initial distribution (about 40 percent) and Catholics have an estimated low intolerance level towards Jews. Protestants, on the other hand, have lower intolerance towards Others than Jews (by a factor of four), and as a consequence, when Protestants are a large majority in the initial conditions, Jews tend to decline and Others gain a small but stable share of the population; this happens for instance for the initial conditions represented by the present composition of California and Texas.

The transitions towards the stationary states are also quite interesting. There appear to be cycles of temporary dominance by several groups, before either Protestants or Jews establish themselves as majorities. With the initial conditions of Illinois, for instance, while the share of Jews rises exponentially, so does the share of Others for a little less than 40 generations before dropping quickly. Also, in Illinois, Catholics enjoy early success, reaching over 50 percent after about 20 generations before declining. In New York, Protestants and Catholics decline steadily, while Others rise before declining. This is strong evidence that the dynamics of the distribution across religious groups are determined by the specific non-linear interactions between the choices of marriage segregation and direct family socialization, which at times generate substitution across religious groups and at times complementarities.

The relative success of Others in the simulations is made even more striking when noticing that the average fertility rate of this group is below reproduction (less than 2), and is particularly low for homogamous marriages (less than 1.7). Naturally, though, the dynamics of the distribution of the population is influenced in an important way by the estimated oblique socialization bias due to conversions, which favors Others. While, in fact, Others are represented in the stationary state which is attracting from the initial composition of the population in California and Texas even if the socialization bias is reduced by two standard errors (Figure 12), this is not true if the bias is arbitrarily put to zero (Figure 13).

Catholics are never present in the stationary distributions. This is because we estimate very high intolerance levels towards Catholics for all other religious denominations, including Others; see Table 5.

Even increasing the fertility rate of Catholics to account for the Hispanic Catholic migration does not generate a transition toward a path converging to a stationary state in which Catholics are present. Rather, accounting for such migration flows into the Catholic population has the effect of favoring the Jews in the long run. In particular, we simulate an increase of the Catholic population of 80 and 90 percentage points for 3 generations from the initial conditions of, respectively, California and Texas (Figure 14). This is roughly consistent with state forecasts of the inflow of the Hispanic Catholic population in California and Texas, based on Census data.[25] The transitional dynamics of the distribution by religious groups, when we account for the migration inflows, are similar to those we observe when starting the system from the initial condition of Illinois, and the limit population is composed of only Jews (this is true even if we project the same migration rates for 10 generations; see Figure 14). This is consistent with the properties of the path of the distribution of the population by religious group indicated above, where a higher proportion of Catholics in the initial conditions is what makes the stationary state, composed of Jews only, attractive.[26]

Our simulations are in striking contrast with the triple melting pot prediction as well as the predictions concerning the vanishing of the U.S. Jewish population derived from the National Jewish Population data. This is not an artifact of the data, but rather is due to our methodology of performing the simulations of the dynamics at the estimated deep preference and institutional parameter values of the structural model rather than extrapolating from the observed or implied behavioral marriage and socialization rules. (As we argued in the Introduction, such a methodology accounts for possible nonlinearities in the way marriage and socialization rules depend on the distribution of the population by religious groups.)

We have, in fact, also simulated the dynamics of the distribution of the population by religious group by extrapolation under two alternative assumptions: *i)* the average marriage and socialization rates in the U.S. are constant (and set equal to the observed rates); *ii)* the behavioral rules for marriage segregation and socialization of the different groups are constant (and use the estimated values of α^i and $\tau^i + m$, for each i, in the alternative model with exogenous socialization and marriage introduced in Section 6 and Table 8, column 3).

Results for the initial conditions of California, Illinois, New York, and Texas are

[25]The data is publicly available at http://www.lao.ca.gov/1998_calfacts/98calfacts_demographics.html/ for California, and at http://www.county.org/cip/presentations/demog98/sld025.html/ for Texas.

[26]Computations which we do not report show that a migration rate of Hispanics accounting for a 10 percentage point increase in the growth of the Catholic population for three generations is sufficient to switch to the new stationary state with Jews only, in California. In Texas, on the other hand, an increase of about 20 percentage points is necessary for the dynamics of the distribution of the population to converge to such stationary state.

reported in Figures 15 and 16. In accordance with the triple melting pot prediction, revised to account for the recent increase in the intermarriage rates of Jews, the simulations obtained with constant marriage and socialization rates predict a limit population composed mainly of Catholics and Protestants, the major religious groups in the initial conditions. A statistical artifact of random matching is that such groups have low intermarriage rates. The Jewish population in fact vanishes, while the conversion bias guarantees a small population of Others in the limit (Protestants account for 66 percent of the unique stationary distribution in this case, Catholics for 24 percent, and Others for 10 percent, see Figure 15). If on the other hand, we extrapolate from the constant estimated behavioral rules for marriage segregation and socialization of the different groups we already find evidence against the triple melting pot: Catholics and Protestants cannot co-exist in the limit, and which group is represented in the stationary distribution depends on the initial conditions. But in all these simulations, Jews still vanish (and Others maintain a 10 percent share of the population; see Figure 16).

It is worth noting again that our simulation exercises are necessarily based on the assumption that the parameters estimated are stable and therefore constant over time. Since we estimate the deep preference and technology parameters of the marriage and socialization model, stability is less severe an assumption here, than in the case in which behavioral rules are directly estimated and the simulations of the population composition dynamics are obtained by linear extrapolations from such rules (as is the case in Figures 15 and 16).[27] Still, it is important to exercise caution in interpreting the results of the simulations. A time period in the simulation is a generation, and, therefore, stability of the parameters over the forty or so generations that the distribution of the population takes to reach a stationary state is impossible to maintain. As a consequence, the simulations we report are aimed at illustrating the implications of our estimation results, and should not be interpreted as direct forecasts of the future prevalence of the different religious denominations.

9 Discussion

In order to empirically estimate the marriage and socialization processes along the religious dimension in the U.S. we assume that the geographical unit of reference coincides with the state. Moreover, while our analysis treats marriage as an economic decision of each agent, motivated by his/her preferences regarding the socialization of children, we treat the distribution of agents by state as fixed. In other words we do not consider the agents' moving decisions as endogenous. This is done for obvious data limitations, as

[27]Notice that we are assuming constant differential fertility rates across religious groups and types of marriages. Our simulations therefore do depend on the stability of fertility rates.

the GSS survey only records the residence of the individual respondent at the moment of the survey, and if the respondent ever moved in the past. The endogeneity of moving decisions might be problematic for our estimates, in principle, if these decisions were motivated in part by marriage and socialization and if they caused some unobserved heterogeneity which could otherwise explain our results. We turn therefore to addressing these issues in the context of our analysis of marriage and socialization by religious traits in the U.S.

Small religious communities are often concentrated in religiously homogeneous city neighborhoods or counties; for instance the Orthodox Jews of Boro Park, Brooklyn, New York, or the Amish of Lancaster County, Pennsylvania. The relevant marriage pools of the members of such "enclaves" are not the state, as our model postulates, but rather the enclave itself. However, this is less of a problem for our estimates than it might appear. Consider the Jews of Boro Park as an illustration (see Mayer (1979) for a sociological analysis of this community). Modelling Boro Park as a geographical unit of our analysis would imply considering the Jews living in this neighborhood as a majority in which marriages are homogamous and children are socialized with minimal effort by families. But while it is true that living in Boro Park essentially ensures homogamy and socialization, we claim it is incorrect to conclude that this is achieved with minimal effort on the part of families. This argument in fact disregards the effort and costs associated with moving into such small closed-knit communities, and the cost associated with not moving out of them (see e.g., Borjas (1995) for an analysis of such costs for ethnic communities). In the analysis of this paper, instead, the Jews of Boro Park are a small minority in New York State; it is costly for them to marry homogamously and socialize their children, exactly because it requires some form of segregation. Their high homogamy and socialization rates in the data therefore are attributed to the intensity of their preferences for transmitting their own religious faith to their children, and hence correctly contribute to the estimation of high rates of intolerance.

A potentially more serious problem, which is a consequence of not considering the endogeneity of moving decisions, is due to the unobserved heterogeneity regarding the intensity of religious preferences. Small religious communities could, for instance, have more intense religious preferences because those individuals with limited religious identity and attachment to the "land of the fathers" have left over the generations. In this case, our estimation procedure would erroneously attribute high marriage and socialization rates of minorities to high intolerance rates of the religious denomination as a whole rather than to the religious intensity of the minority itself. We would then over-estimate the intolerance levels of religious groups for which minorities have a relatively higher intensity of religious preferences. This is because *i)* our identification procedure requires identical preferences within religious groups, and therefore it disregards any heterogeneity in religious intensity; and because *ii)* it is exactly the homogamy and socialization rates of minorities, and hence their intolerance levels, which determine in our estimation procedure the intolerance levels attributed to the whole religious group (homogamy and socialization rates of majorities are due mostly to random matching).

Within the limitations of the data, we have attempted to address these issues, using data on "church attendance" from the GSS survey over the period $1972-96$ as a measure of religious intensity (for Others, we obviously exclude individuals who express no religious preference). While there is some evidence of unobserved heterogeneity with regards to religious intensity, it turns out that it is in larger communities that members display higher religious intensity, rather than in smaller communities. Therefore, if anything, we are in effect under-estimating intolerance levels in our analysis. In particular, for Jews, Catholics and Others, average attendance does not depend on the religious share of the state of residence (the GSS does not record the state of origin); this is true for both movers (the subset of agents who moved across states prior to the GSS interview) and non-movers separately. For Protestants instead, attendance is positively correlated with the share of Protestants in the residence state, both for movers and non-movers: see Figure 17 and Table 10a. By grouping agents into low and high intensity and then looking at the distribution of movers by religious share in the residence state one finds essentially the same pattern (see Figure 18 and Table 10b). For Protestants, the ratio of high intensity to low intensity movers is higher for states in which Protestants are a majority; the same is true for Jews, but no such relationship appears in the data for Catholics or Others.

Because the estimates of the intolerance levels of the different religious groups are so central in our analysis we have further pursued the analysis of the implication of unobserved heterogeneity with respect to religious intensity. We have estimated our structural model of marriage and socialization under the assumptions that, given a baseline intolerance level for religion i, ΔV^i, the intolerance level of minorities (defined as communities in the lowest quartile of the distribution of q^i) is either $0.75 \cdot \Delta V^i$ or $1.25 \cdot \Delta V^i$. Consistent with the analysis we just reported of the correlation of average attendance and religious share, we find that the model in which the intolerance level of minorities is smaller ($0.75 \cdot \Delta V^i$) fares better, in the $KLIC$ sense, than the model in which it is larger.[28] Either model does not fit the data quite as well as our baseline model in which intolerance levels are assumed invariant across religious shares, that is both for minorities and majorities.[29] We conclude therefore that this study of religious intensity, albeit necessarily coarse, supports our baseline model which restricts the preferences of individuals of the same religious group to be identical in terms of intensity. If a bias due to the unobserved heterogeneity of religious intensity is present in our analysis it leads us to underestimate the importance of intolerance levels in determining the high observed homogamy and socialization rates.[30]

[28]The Kitamura test statistic is -14.625, with a p-value very close to zero.

[29]The Kitamura test statistic when we test our baseline model against the two specifications where minorities are less (more) intolerant is -7.796 (-20.219), with p-values practically zero.

[30]It is worth mentioning that we have also estimated our baseline model using data for non-movers only (see Table 9 and Appendix 2 for details): the point estimates are quite similar to those obtained

Alternative explanations of the evidence regarding the high socialization and marriage segregation levels of small religious communities can be developed, that rely on various other possible instances of unobserved heterogeneity across agents. High homogamy rates of minorities could also be rationalized, for instance, if small religious communities are more homogeneous in some dimension along which marriage would be assortative. In this case, in fact, high religious homogamy rates in small religious communities would be a statistical artifact of the assortativeness of marriages along dimensions other than religious faith. Natural examples might consist of race and education levels. It is well known for instance that individuals tend to marry spouses with a similar education level (resp. of the same race). Hence, if small religious communities are more homogeneous in terms of education (race) than larger communities, we would observe disproportionately high religious homogamy rates in small religious communities. While the correlation of the coefficient of variation of race between the members of religious group i and the population share of religion i is zero or negative for all groups except Others, the correlation of the coefficient of variation of education between the members of religious group i and the population share of religion i is in fact positive for Protestants and Catholics (.5 and .37, respectively). We conclude that the homogeneity of education levels (but not of race) could contribute to explain the socialization and marriage behavior of minorities. Considering religious faith and education levels as joint determinants of the assortativeness of marriage rates is potentially very important (see the concluding section).

Another dimension in which our analysis of marriage homogamy and socialization by religious denomination in the U.S. could be problematic consists of our definition of the religious groups, which aggregates several potentially very different sub-groups. The residual group, Others, is obviously the most heterogeneous in terms of beliefs and cultural characteristics, but substantial heterogeneity also characterizes Protestants and even Jews (perhaps less so Catholics). Our assumption that preferences are identical within any religious groups requires therefore some scrutiny: our analysis might be mistakenly aggregating over different preference parameters of, say, Black and White Protestants, or Catholics of Hispanic and non-Hispanic ethnic origin. To address this issue we re-estimate our model after expanding the religious groups to include Black and White Protestants separately.[31] The structural estimation results in this case with five religious groups are reported in Table 11.

---

using the whole sample.

[31]We have also re-estimated the model treating separately Black and White Protestants, as well as Hispanic and non-Hispanic Catholics. We do not discuss here such estimates (reported however in Table 12 for the interested reader) because they are too imprecise due to the small sample size in several cells (the overall number of usable observations is reduced to $8,147$ in 16 states). However, most of the point estimates agree with our baseline estimates: interestingly, the distinction between Hispanic and non-Hispanic Catholics seems to matter very little.

We interpret here the results, even though a word of caution is necessary since the point estimates have fairly large standard errors (this is due to the relatively large number of moments used in the estimation, and the relatively small sample size of several cells). First, Black and White Protestants appear in fact as two quite distinct groups, with elevated estimated intolerance levels for each other. This is especially so for White Protestants towards Blacks rather than vice versa. Second, the elevated intolerance levels of Jews and the small intolerance levels of Catholics we estimated in our baseline model with 4 groups hold with the extension to 5 groups; similarly, the conversion parameter in favor of the residual group, Others, remains stable.

Since the intolerance parameters are difficult to interpret *per se*, we find it useful to further assess the robustness of our analysis by running the simulations of the dynamics of the distribution of the population by religious trait for the estimated model with 5 groups. We use as initial conditions of the simulations, as in Section 8.1, the religious shares in our data for New York, Illinois, Texas and California: the simulations are reported in Figure 19a. In all simulations Black Protestants do not survive. This is possibly in part a consequence of the higher estimated intolerance towards them by White Protestants. In fact, for all initial conditions a unique stationary state appears, composed of a majority of White Protestants and a minority of Others. In contrast to our baseline estimates with 4 groups, a stationary state with a majority of Jews does not appear in the version with 5 groups. However, if we re-run the simulations using only a slight perturbation of the parameter values, the stationary state with a majority of Jews re-emerges and the dynamics of the distribution of the population converge to it from the initial conditions of New York and Illinois. It turns out that the critical parameter for the survival of Jews in the simulations is the intolerance of White Protestants towards Others, ΔV^{WO}: it is sufficient to reduce it from 453 to 445, well inside its 99% confidence interval, to observe a majority of Jews in the stationary state with the initial conditions of New York (see Figure 19b). A further reduction of ΔV^{WO} to 265 (within the 99% confidence interval) produces in the limit a majority of Jews for the initial conditions of Illinois (Figure 19c). We conclude that our baseline analysis of marriage and socialization with 4 groups is relatively robust to the extension to 5 groups, i.e. to separating Protestants into Blacks and Whites. The race distinction is not irrelevant for religious segregation though, and, while we cannot pursue it further because of data limitations, our analysis suggests the presence of interesting phenomena related to the interaction of race and religion in marriage and socialization.

10 Conclusions

We concentrated our analysis on a simple dimension of assortativeness in marriage: religious homogamy. It is well documented, however, that marriages are assortative in several other dimensions, including education (see for instance the survey of Mare (1991)) and race. The possible correlation between education, as well as other characteristics of spouses, and religious faith can, therefore, have important implications for our analysis of marriage and socialization. As we noticed, for instance, the relative homogeneity with respect to education of small religious communities could at least in part explain the observed high socialization and homogamy rates of minorities.

We have conducted a preliminary analysis of a model in which assortative marriage can occur along both the education and religious dimensions. This indicates that, overall, our socialization-based interpretation of intermarriage rates is robust to the inclusion of a preference for educated spouses. The many interesting interactions between the preferences for education and for religious homogamy require an independent treatment, though. The same, we argued, can be said for the interaction of religious homogamy, socialization and race.

Similarly, a more detailed analysis of the endogenous determinants of fertility is bound to be of great value, especially with regards to the simulations of the dynamics of the distribution of the population. For instance, a significantly higher estimated sensitivity of parental preferences with respect to fertility might have an important effect on the long run distribution of the population by religious groups.

Also, we have dealt only marginally with the issue of conversions. Our analysis indicates that the issue is of substantial relevance for the dynamics of the relative shares of the religious groups. Substantial literature has stressed the role of conversions in marriage for religious socialization (see Iannaccone (1990)).

Finally, data limitations did not allow us to consider several other important issues related to marriage and socialization: the effects of mobility in the determination of the relevant marriage pools, gender asymmetries in socialization, the hierarchical representation of different religious group in the social psychology of the U.S., to cite only some of such issues.

REFERENCES

Barron, Milton, "The Church, the State, and Intermarriage," in M. Barron (Ed.), *The Blending American: Patterns of Intermarriage*, Chicago, 1972.

Becker, Gary S., "A Theory of Marriage: Part I," *Journal of Political Economy*, LXXXI (1973), 813-46.

Becker, Gary S., "A Theory of Marriage: Part II," *Journal of Political Economy*, LXXXII (1974), 11-26.

Becker, Gary S., *A Treatise on the Family* (Cambridge, MA: Harvard University Press, 1981).

Becker, Gary S., Landes, Elizabeth M., and Michael, Robert T., "An Economic Analysis of Marital Instability," *Journal of Political Economy* LIIIV (1977), 1141-87.

Bisin, Alberto, and Thierry Verdier, "The Economics of Cultural Transmission and the Dynamics of Preferences: The N-Trait Case", NYU, 2000.

Bisin, Alberto, and Thierry Verdier, "Beyond the Melting Pot: Cultural Transmission, Marriage and the Evolution of Ethnic and Religious Traits," *Quarterly Journal of Economics*, (2000) forthcoming.

Borjas, George J., "Ethnicity, Neighborhoods, and Human Capital Externality," *American Economic Review*, LXXXV (1995), 365-90.

Boyd, Robert, and Peter Richerson, *Culture and the Evolutionary Process* (Chicago, IL: University of Chicago Press, 1985).

Brock, William and S. Durlauf, "How Can Statistical Mechanics Contribute to Social Sciences ?," *Proceedings of the American Academy of Sciences*, 96 (1999), 10582-84.

Brock, William and S. Durlauf, "Discrete Choice and Social Interactions," forthcoming in *Review of Economic Studies*, 2001.

Cavalli-Sforza, Luigi, and Marcus Feldman, *Cultural Transmission and Evolution: A Quantitative Approach* (Princeton, NJ: Princeton University Press, 1981).

Chiswick, Carmel U., "The Economics of Jewish Survival," mimeo presented at the *International Conference on Economics of Judaism and Jewish Observance*, 1999.

Clark, C. A., and A. Worthington, "Family Variables Affecting the Transmission of Religious Values from Parents to Adolescents: A Review," *Family Perspective*, XXI (1987), 1-21.

Cornwall, Marie, "The Influence of Three Agents of Religious Socialization," in *Religion and Family Connection: Social Science Perspectives*, Darwin L. Thomas, ed. (Provo, UT: Brigham Young University Press, 1988).

Cox, D.R., "Tests of Separate Families of Hypotheses", in *Proceedings of the Fourth Berkeley Symposioum in Mathematical Statistics and Probability,* (Berkeley, CA: University of California Press, 1961)

Dagsvik, John, and Boyan Jovanovic, "Was the Great Depression a Low-Level Equilibrium ?," *European Economic Review*, 38 (1994), 1711-1729.

DeVaus, D. A., "The Relative Importance of Parents and Peers for Adolescent Religious

Orientation: An Australian Study," *Adolescence*, LXIX (1983), 147-58.
Erickson, Joseph A., "Adolescent Religious Development and Commitment: A Structural Equation Model of the Role of Family, Peer Group, and Educational Influences," *Journal for the Scientific Studies of Religion*, XXXI (1992), 131-52.
Glaeser, Edward and Jose' Scheinkman, "Non Market Interactions," mimeo for the 2000 World Meeting of the Econometric Society, 2001.
Glaeser, Edward and Jose' Scheinkman, "Measuring Social Interactions," mimeo, Princeton University.
Glazer Nathan and Daniel P. Moynihan, *Beyond the Melting Pot* (Cambridge, MA, MIT Press, 1970).
Glazer, Nathan, *We Are All Multiculturalists Now* (Cambridge, MA: Harvard University Press, 1997).
Goffe, William L., "SIMANN: A Global Optimization Algorithm using Simulated Annealing", *Studies in Nonlinear Dynamics and Econometrics,* I (1996), 169-176.
Goffe, William L., G.D. Ferrier and J. Rogers, "Global Optimization of Statistical Functions with Simulated Annealing", *Journal of Econometrics,* LX (1994), 65-99.
Gordon Milton M., *Human Nature, Class, and Ethnicity*, (New York, Oxford University Press, 1964)
Hayes, Bernadette C., and Yvonne Pittelkow, "Religious Belief, Transmission, and the Family: An Australian Study," *Journal of Marriage and the Family*, LV (1993), 755-66.
Heaton, Tim, "Religious Homogamy and Marital Satisfaction Reconsidered," *Journal for the Scientific Study of Religion*, XXV (1984), 248-58.
Heaton, Tim, "How Does Religion Influence Fertility? The Case of Mormons," *Journal for the Scientific Study of Religion*, XXVIII (1986), 283-99.
Heckman, James, 'The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process,' in Charles Manski and Daniel McFadden (Eds.), *Structural Analysis of Discrete Data*, MIT Press, 1981.
Herberg, William, *Protestant-Catholic-Jew*, (New York: Doubleday, 1955).
Heer, David M., "Intermarriage," in Stephan Thernstrom, Ann Orlov and Oscar Handlin (Eds.), *Harvard Encyclopedia of American Ethnic Groups*," (Cambridge, MA: Harvard University Press, 1980).
Heer, David M., and Charles A. Hubay, Jr., "The Trend of Interfaith Marriages in Canada, 1922 to 1972," in S. Parvez Wakil (Ed.), *Marriage, Family and Society: Canadian Perspectives*, Toronto (1975).
Hoge, Dean R., and G. H. Petrillo, "Determinants of Church Participation and Attitudes among High School Youth," *Journal for the Scientific Study of Religion*, XVII (1978), 359-79.
Hoge, Dean R., G. H. Petrillo, and Eric I. Smith, "Transmission of Religious and Social Values from Parents to Teenage Children," *Journal of Marriage and the Family*, XLIV (1982), 569-80.
Iannaccone, Laurence, "Religious Practice: A Human Capital Approach," *Journal for*

the Scientific Study of Religion, XXIX (1990), 297-314.
Johnson, Robert Alan, *Religious Assortative Marriage in the United States* (New York, NY: Academic Press, 1980).
Jovanovic, Boyan, "Observable Implications of Models with Multiple Equilibria," *Econometrica*, 57 (1989), 1431-1437.
Landis, Judson, "Marriages of Mixed and Non-Mixed Religious Faith," *American Sociological Review*, 14 (1949), 401-07.
Kennedy, Ruby Jo Reeves, "Single or Triple Melting Pot ? Intermarriage Trends in New Haven," *American Journal of Sociology*, 49 (1952), 331-39.
Kitamura, Yuichi, "Comparing Misspecified Dynamic Econometric Models Using Nonparametric Likelihood", mimeo (2000), University of Wisconsin.
Lehrer, Evelyn, and Carmel Chiswick, "The Religious Composition of Unions: Its Role as a Determinant of Marital Instability,"*Demography*, XXX (1993), 385-404.
Mancuso, David, *Some Implications of Marriage and Assortative Mating by Schooling for the Earnings of Men*, (Ph.D dissertation, Stanford University, 1997)
Mare, Robert D., "Five Decades of Educational Assortative Mating", *American Sociological Review,* 56(1), (1991), 15-32.
Mayer, Egon, *From Suburb to Shetl: The Jews of Boro Park* (Philadelphia, PE: Temple University Press, 1979).
Mayer, Egon, *Love and Tradition: Marriage Between Jews and Christians* (New York, NY: Plenum Press, 1985).
Moro, Andrea, "The Effect of Statistical Discrimination on Black-White Wage Inequality: Estimating a Model with Multiple Equilibria," mimeo, University of Minnesota.
Ozorak, Elizabeth W., "Social and Cognitive Influences on the Development of Religious Beliefs and Commitment in Adolescence,"*Journal for the Scientific Study of Religion*, XXVIII (1989), 448-63.
Pencavel John; "Assortative Mating by Schooling and the Work Behavior of Wives and Husbands", *American Economic Review Papers and Proceedings* , (1998), 88(2), 326-329
Riesman, Paul, and David L. Szanton, *First Find Your Child a Good Mother: The Construction of the Self in Two African Communities*, Rutgers University Press, 1992.
Rindfuss, Ronald R., and Audrey Van den Heuvel, "Cohabitation: A Precursor to Marriage or an Alternative to Being Single ?,"*Population and Development Review*, XVI (1990), 703-26.
Sander, William, "Catholicism and Intermarriage in the United States,"*Journal of Marriage and the Family*, LV (1993), 1037-41.
Schoen, Robert, and Robin M. Weinick, "Partner Choice in Marriages and Cohabitations,"*Journal of Marriage and the Family*, LV (1993), 408-14.
Singleton, Kenneth J., "Testing Specifications of Economic Agents' Intertemporal Optimum Problems in the Presence of Alternative Models", *Journal of Econometrics,* 30 (1985), 391-413.
Smith, Roger C., *Two Cultures One Marriage* (Berrien Springs, MI: Andrews University

Press, 1996).
Smith, Richard J., “Non-Nested Tests for Competing Models Estimated by Generalized Method of Moments”, *Econometrica,* 60 (1992), 973-980.
Thomas. John L., “The Factor of Religion in the Selection of Marriage Mates,” *American Sociological Review*, 16 (1951), 487-91.
Vandenberg, Steven G., “Assortative Mating, or Who Marries Whom,” *Behavior Genetics*, II (1972), 127-157.
Vuong, Quang H., “Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses”, *Econometrica,* 57 (1989), 307-333.
Warren, Brian L., “Socio-economic Achievement and Religion,” *Sociological Inquiry*, XL (1970), 130-155.
Wilson, William J., *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy* (Chicago, IL: The University of Chicago Press, 1987).

Appendix 1
The Structural Model

The structural model is introduced in Section 3. We assume the following.

Assumption 1 *The cost functions, $M(\alpha^i, q^i)$ and $S(\tau^i, q^i)$, are differentiably strictly convex, and satisfy the Inada conditions for interiority:* $\forall q^i \in (0,1)$,

$$\lim_{\alpha^i \to 1} M(\alpha^i, q^i) = \infty, \quad \lim_{\tau^i \to 1} S(\tau^i, q^i) = \infty$$

Assumption 1 guarantees that the first order conditions,

$$\frac{\partial S}{\partial \tau^i}\left(\tau^i, q^i\right) = (1 - Q^i)V^{ii} - \sum_{j \neq i} Q^j V^{ij} \tag{19}$$

$$\frac{\partial M}{\partial \alpha^i}(\alpha^i, q^i) = \left(1 - A^{ii}\right)\left(n^{ii}\right)^{\xi} W^{ii} - \sum_{j \neq i} A^{ij}\left(n^{ij}\right)^{\xi} W^{ij} \tag{20}$$

relative respectively to the socialization and the marriage problems (7) and (10), are necessary and sufficient for a solution.

After some algebraic manipulations, and taking explicit account of the geographic state index, the socialization and marriage model introduced in the previous section is simply represented by the system of equations in the following Table:

The Structural Equations

$$\pi_s^{ii} = \alpha_s^i + (1-\alpha_s^i)A_s^{ii}, \quad \forall i \tag{21}$$

$$\pi_s^{ij} = (1-\alpha_s^i)A_s^{ij}, \quad \forall i \neq j \tag{22}$$

$$P_{ii,s}^i = \widetilde{\tau}_s^i + (1-\widetilde{\tau}_s^i)Q_s^i \tag{23}$$

$$P_{ii,s}^j = \left(1-\widetilde{\tau}_s^i\right) Q_s^j, \ i,j \text{ distinct} \tag{24}$$

$$P_{ij,s}^i = \frac{1}{2}m + (1-m)\, Q_s^i, \ i,j \text{ distinct} \tag{25}$$

$$P_{ij,s}^k = (1-m)\, Q_s^k, \ i,j,k \text{ distinct} \tag{26}$$

$$\frac{\partial S}{\partial \tau_s^i}\left(\tau_s^i, q_s^i\right) = \sum_j Q_s^j \Delta V^{ij}, \quad \forall i \tag{27}$$

$$\frac{\partial M}{\partial \alpha_s^i}(\alpha_s^i, q_s^i) = \left[(1-A_s^{ii})(n^{ii})^{\xi}\widetilde{\tau}_s^i - m\sum_{j\neq i} A_s^{ij}(n^{ij})^{\xi}\right]\sum_h q_s^h \Delta V^{ih} + \tag{28}$$

$$+\frac{m}{2}\sum_{j\neq i} A_s^{ij}(n^{ij})^{\xi}\Delta V^{ij} + \left[\sum_{j\neq i} A_s^{ij}\left((n^{ii})^{\xi} - (n^{ij})^{\xi}\right)\right]\sum_h q_s^h V^{ih} - (1-A_s^{ii})(n^{ii})^{\xi} S\left(\tau_s^i, q_s^i\right)$$

$$A_s^{ii} = \frac{(1-\alpha_s^i)q_s^i}{\sum_{k=1}^n (1-\alpha_s^k)q_s^k}, \quad \forall i \tag{29}$$

$$A_s^{ij} = \frac{(1-\alpha_s^j)q_s^j}{\sum_{k=1}^n (1-\alpha_s^k)q_s^k}, \quad \forall i \neq j \tag{30}$$

We can now prove the following:

Theorem 1 *The solution of equations (21-30) and (11-13), given q_s^i, for all i and s, and n^{ij} for all i,j, defines a mapping, $\widetilde{\Pi}(\theta)$, from θ into π_s^{ij}, for all i,j and s, and into P_{ij}^k for all i,j,k. Under Assumption 1, such a map is an upper-hemi-continuous correspondence, and is smooth except at points of discontinuity. Moreover, it is sufficient for costs to be independent of q_s^i for the map to be a continuously differentiable function.*

Proof. Fix q_s^i, for any s, and n^{ij} for all i, j. Since problem (7) is convex and (13) is smooth, its solution is represented by a continuously differentiable function mapping θ into τ^i. Since problem (10) is convex, its solution is also represented by a continuously differentiable function mapping A_s^{ij}, for any j and any s, and θ into α^i. Substituting such mapping into (29-30), a fixed point problem is constructed, which represents the solutions for α_s^i, A_s^{ij}, for all i, j, s. This fixed point problem has a solution, as the conditions for the Brouwer Fixed Point theorem are satisfied.

Upper-hemi-continuity of the map which represents a solution of (21-30) and (11-13) is immediately guaranteed. The map is smooth except at discontinuity points as a consequence of the smoothness of the system of equations (21-30) and (11-13).

Continuous differentiability of the map for cost functions independent of q_s^i is not trivial, but follows from an extension of the argument provided in Bisin-Verdier (2000) for $i = 1, 2$. ■

Appendix 2
Data and Methodology

The data for the empirical exercise come from the General Social Survey (GSS).[32] In the survey, each respondent reports his/her religious affiliation and that of his/her spouse.[33] Respondents are considered a representative sample of the religious affiliations of individuals in each state of residence. In the empirical implementation, we consider 4 religious groups (Protestants, Catholics, Jews, and the residual group, Others), $i, j = P, C, J, O$. The GSS survey does distinguish between individuals who prefer "no religion" and those with religious faiths other than Protestant, Catholic, and Jewish. The dimension of the sample though forces us to aggregate the last two groups of individuals into one group, which we call 'Others'.[34]

The total number of respondents is 35,284. We eliminate respondents who are not married at the time of the survey, divorcees, or those for whom we lack information about own or spouse religion. This leaves us with 16,722 observations. The 23 states we consider are the following: California, Colorado, Connecticut, Florida, Georgia, Illinois, Indiana, Maryland, Massachusetts, Michigan, Minnesota, Missouri, New Jersey, New York, North Carolina, Ohio, Oregon, Pennsylvania, South Carolina, Tennessee, Texas, Virginia, and Wisconsin. For the remaining states we do not have sufficient data to estimate all the q_s^i religious shares. This final filter brings the number of observations down to 13,790.

We also estimated the model excluding respondents who changed residence at some time before the survey, as we do not know the state in which they resided when they got married. This reduces the number of observations to 7,286, and reduces the number of states for which we have sufficient data to 15. While the fit is much worse and the estimates quite imprecise in this case, the point estimates are very similar to those with the whole sample. This estimate is reported for comparison in Table 9.

Since respondents are the sampling unit in the GSS survey, we construct our measure of the religious composition of marriages, and our measure of the religious composition of the population, based on *respondents* rather than marriages.[35] We report in Table 1

[32]Most of the data is publicly available at: http://www.icpsr.umich.edu/ with the exception of the state geocodes for all respondents, which are only available upon request from NORC. The GSS is a nearly annual national survey of U.S. residents that focuses on attitudes, perceptions, and social trends, in addition to more conventional socio-economic and demographic characteristics of respondents. An incomplete list of topics covered over the years includes class, religion, politics, sex, and health issues. The cumulative dataset covers the period 1972-1996 (the most recent wave, 1998, has just been incorporated).

[33]The exact text of the question in the GSS is: "What is your religious preference? Is it Protestant, Catholic, Jewish, some other religion, or no religion?"

[34]Overall, individuals with no religious preference are the majority of the our residual group 'Others', accounting for about 77 percent of 'Others' on average over the U.S., with a maximum of 86 percent in Tennessee and a minimum of 60 percent in Maryland.

[35]Independent measures of religious affiliation by state do exist, e.g., the ARDA dataset. Such

the shares of each religious group as well as the sample probability that a member of each religious group i in our subdivision marries homogamously or with a member of religious group $j \neq i$, by state.

The data on socialization come from a special module of the GSS, collected in 1988, that asks respondents to report the religious identity of their parents. Unfortunately, there are only 1,232 useful observations in this module. Therefore it is only possible to estimate the empirical frequencies of the offspring's religious choices given the parents' identities for the whole U.S. and not for individual states.[36]

The structural parameters θ are estimated via minimum distance, by matching the vector $\widehat{\Pi}$ of empirical moments $\left(\widehat{\pi}_s^{ij}, \widehat{P}_{ij}^k\right)$ from the data with the vector $\widetilde{\Pi}(\theta)$ of moments $\left(\widetilde{\pi}_s^{ij}(\theta), \widetilde{P}_{ij}^k(\theta)\right)$ implied by the model for a given choice of θ (see Theorem 1 in the text).

Formally, given a square weighting matrix Ω_N^* (where N denotes the total sample size), the minimum distance estimate $\widehat{\theta}$ minimizes

$$J_N(\theta) \equiv [\widehat{\Pi} - \widetilde{\Pi}(\theta)]^\top \Omega_N^* [\widehat{\Pi} - \widetilde{\Pi}(\theta)]. \tag{31}$$

Each empirical probability $\widehat{\pi}_s^{ij}$ is estimated by computing the empirical frequency of marriage ij in state s:

$$\widehat{\pi}_s^{ij} \equiv \frac{\#(i \text{ married } j)_s}{\#(i)_s} = \frac{1}{B_s q_s^i} \sum_{b=1}^{B_s} X_b^{ij},$$

where the subscript b denotes an individual observation, B_s is the number of observations in state s, and X_b^{ij} is a dummy variable that is equal to one if individual respondent b is of religion i and married to a j person, and zero otherwise.[37]

Each empirical probability $\widehat{P}_{ij}^k$ is estimated by computing the empirical frequency of socialization to group k of children in families of type ij, aggregating over all states:

$$\widehat{P}_{ij}^k \equiv \frac{\#(i \text{ children})}{\#(ij \text{ parents})}$$

Thus the only source of randomness in the estimation is the sampling error present in each sample moment $\widehat{\Pi}$.

---

measures grossly underestimate the members of each denomination as a fraction of the population, as they are based on church membership. We have replicated our estimates with the ARDA measures of religious affiliation. While the model's fit is worse, the qualitative structure of the parameter estimates remains unchanged.

[36]Moreover, since the individual respondents are the sampling units, the distribution by religion of the parents of the respondents we observe is not representative of the distribution of the population of the parents. While this is obviously problematic in principle, various attempts at correcting the distortion have not resulted in significantly different estimates.

[37]The unit of observation in the data is an individual respondent, not an individual marriage.

We do not use all of the available moment conditions in the estimation. In particular, we only match the moments π_s^{ij} for ij = PP, PC, PJ, CC, CJ, JJ. The reason we omit the residual moment conditions is that they are linearly dependent on the others. By definition of the probabilities π_s^{ij} and q_s^i, the following linear restrictions hold in the population for each state s:

$$\begin{aligned} \pi_s^{ij} q_s^i &= \pi_s^{ji} q_s^j, \quad \forall i \neq j \\ \sum_j \pi_s^{ij} &= 1, \quad \forall i. \end{aligned} \tag{32}$$

In the estimation, we can therefore omit any 10 of the 16 available marriage moment conditions for each state (given four religious groups) since they are linearly dependent. A subset of these restrictions (eq. (32)) do not hold exactly in the data, though, because of sampling error. Then, the choice of which moment conditions to omit makes a difference in the estimation. We have omitted the moment conditions for the groups for which sampling error is likely to be more prevalent due to small sample size, i.e. Jews and Others. Also, for any possible couple ij only 3 socialization moments are considered, as

$$\sum_k P_{ij}^k = 1.$$

The optimal weighting matrix Ω_N^* used in the minimum distance criterion (31) is $\Omega_N^* = \widehat{S}^{-1}$, where S is the covariance matrix of the vector of moments $[\widehat{\Pi} - \widetilde{\Pi}(\theta)]$. We assume that the individual Π moments that we do include in the estimation are uncorrelated across religions (i.e., that the sampling error associated with the estimation of each $\widehat{\Pi}_s^{ij}$ is uncorrelated with that of any $\widehat{\Pi}_s^{kj}$, $i \neq k$). On the other hand, $\widehat{\Pi}_s^{ij}$ and $\widehat{\Pi}_s^{ij'}$, $j \neq j'$, are negatively correlated according to a multinomial distribution. Therefore, only the within-religion $V\left(\widehat{\Pi}_s^{ij}\right)$ and $Cov\left(\widehat{\Pi}_s^{ij}, \widehat{\Pi}_s^{ij'}\right)$ terms of S are non-zero, and can be easily estimated using the properties of multinomial distributions.

The estimation procedure requires that the map $J_N(\theta)$ be locally smooth in a neighborhood of the estimated value of the parameters θ. The map $J_N(\theta)$ inherits the properties of $\widetilde{\Pi}(\theta)$, which is only upper-hemi-continuous (Theorem 1), in general, because of the possibility of multiple equilibria (Figure 10 illustrates such multiplicity).

Figure 10 reports the values of the $J_N(\theta)$ criterion in the neighborhood of our point estimates, moving each parameter value θ_l one at a time (in the Figure we only report a selection of the most typical patterns). As the Figure shows, our estimates have in fact the property that the selection is locally smooth, even though they are sometimes close to critical points of the marriage market equilibrium set that generate the displayed discontinuities in the minimum distance criterion.

The properties of the estimator $\widehat{\theta}$ are then the standard ones of minimum distance estimators. In particular, $\widehat{\theta}$ is consistent[38] and asymptotically normal, with variance equal

[38] It is easy to show that local smoothness of $J_N(\theta)$ is sufficient to ensure consistency of our estimator $\widehat{\theta}$.

to $V \equiv \left\{DS^{-1}D^{\top}\right\}^{-1}$,where $D^{\top} \equiv p\lim\left\{\frac{\partial\widetilde{\Pi}(\theta)}{\partial\theta^{\top}}|_{\theta=\theta_0}\right\}$ is a matrix of partial derivatives evaluated at the true value θ_0.[39]

The $\widetilde{\Pi}(\theta)$ moments implied by the model are computed as follows. For a given value of the parameters θ, a choice of ΔV^{ij}, $i \neq j$ and o, together with the religious shares q^i_s, pins down the socialization probabilities τ^i_s through equations (27) and (13). Given these, and V^{ii}, we can compute (W^{ii}_s, W^{ih}_s) for an individual of type i in state s. Conditional on a set of (W^{ii}_s, W^{ih}_s), n^{ij} and q^i_s, equation (28) defines a mapping $f : A \to \alpha$. Likewise, equations (29) and (30) define a mapping $g : \alpha \to A$. Therefore, we need to find a fixed point of the mapping $A = g \circ f(A)$. This, in turn, yields a choice of α, the restricted pool matching probabilities. Then, the equilibrium values of α and A determine the vector of theoretical moments $\widetilde{\Pi}(\theta)$ implied by the model, through equations (21 - 26).

Finally, the $J_N(\theta)$ criterion is minimized by using a simulated annealing algorithm. Simulated annealing performs a random search over the parameter space, and accepts not only downhill moves but also uphill moves. The probability of accepting an upward move depends positively on a "temperature" parameter that decreases as the search progresses. At the beginning of the search, the algorithm is allowed to make large upward moves, and thus searches over the whole parameter space. As the temperature drops, the algorithm concentrates on more promising regions, but the random nature of the search still allows it to escape local minima.[40]

This algorithm is explicitly designed to find a global minimum of functions that may present multiple local optima and/or discontinuities. Due to computational limitations, we cannot compute all possible equilibria for each θ. To identify the equilibrium selection jointly with the parameters, in the face of possibly multiple equilibria, we adapt the algorithm so that, in the course of the minimization of $J_N(\theta)$, for each candidate value $\overline{\theta}$, it randomly picks several distinct starting values for the iteration that yields the equilibrium. This procedure effectively searches for several possible equilibria, computes the criterion J_N for each of these equilibria, and picks the lowest value as the value $J_N(\overline{\theta})$ for that particular value of θ. As a robustness check, given our estimates $\widehat{\theta}$, we are able to compute all the possible multiple equilibria for the set of α and A, and consequently for $\widetilde{\Pi}$. We then evaluate the minimum distance criterion $J_N(\theta)$ over all possible equilibria, to check that our estimate $\widehat{\theta}$ is indeed a local minimizer of $J_N(\theta)$ over the entire equilibrium set.

[39]In our context, the distribution of religious shares is a discrete-time process. However, we only make use of data at a single time t. Therefore, the problem of initial conditions in estimating discrete time-discrete data stochastic processes (see Heckman (1981)) does not arise.

[40]For a more detailed description of the algorithm, see Goffe (1996) and Goffe et al. (1994). We are very grateful to Bill Goffe for providing us with his MATLAB simulated annealing routines.

Appendix 3

The Stationary States of the Distribution by Religious Groups

At a stationary state, by definition, $q_t^j = q^j$ is constant over time, as well as $\pi_t^{jh} = \pi^{jh}$, $P_{jh,t}^i = P_{jh}^i$. Also, at a stationary state

$$\frac{N_t^i}{N_{t+1}^i} = \kappa_i \text{ constant over } t \tag{33}$$

and

$$\frac{N_t}{N_{t+1}} = \sum_i \frac{N_t^i}{N_{t+1}} = \sum_i \kappa_i q^i = \gamma. \tag{34}$$

We can now write the equations that define a stationary state of the dynamics of the distribution of religious traits in the population as follows:

$$\begin{aligned}
\frac{1}{\kappa_i} &= \sum_j \frac{q^j}{q^i} \sum_h \pi^{jh} \frac{n^{jh}}{2} P_{jh}^i, \ \forall i \\
q^i &= \gamma \sum_j q^j \sum_h \pi^{jh} \frac{n^{jh}}{2} P_{jh}^i, \ \forall i \\
\gamma &= \sum_i \kappa_i q^i.
\end{aligned}$$

The unknown stationary state parameters are (q^i, κ_i, γ).