Interethnic Marriage and the Labor Market Integration of Immigrants

Z. Eylem Gevrek†

Department of Economics
University of Arizona
Job Market Paper

This Version: October 18, 2009

Abstract

This study investigates the role of interethnic marriage in the economic integration of immigrants. I measure the economic performance of immigrants in terms of wages. Using data from the Dutch survey 'Social Position and Use of Public Utilities by Immigrants', I examine whether intermarried immigrants perform better in the labor market than their co-ethnic married counterparts. The empirical strategy accounts for both the potential endogeneity of intermarriage and selection into employment using a full-information maximum likelihood estimation procedure, the Discrete Factor Method. I find that intermarried immigrants earn more than their endogamously married counterparts. Marrying a native is associated with a wage premium of seven percent. The positive effect of intermarriage on wages is robust to changes in the estimation method and specification of the model. Moreover, the results indicate that the intermarriage premium varies across generations. Second-generation immigrants receive no gain from intermarriage.

JEL Classifications: J12, J61

† I am very grateful to my advisor Ronald Oaxaca, Price V. Fishback, Jonah B. Gelbach, Manuela Angelucci and Deniz Gevrek for their continued support and helpful discussions. I would also like to thank Tom Mroz, Alfonso Flores-Lagunes, Sonam Gupta and seminar participants at the IZA 2008 Summer School, the Society of Labor Economists (SOLE) 2009 and the European Economic Association (EEA) 2009 for their comments and suggestions. All remaining errors are my own. Contact Information: Department of Economics, University of Arizona, 401 McClelland Hall 1130 E. Helen Street, Tucson AZ 85721. Email: eylem@email.arizona.edu

1 Introduction

Social scientists consider interethnic marriage as one of the most important indicators of social and economic integration of immigrants for several reasons (Gordon 1964; Kalmijn 1998; Muttarak 2004). First, interethnic marriage helps accelerate the fading of cultural and social barriers between immigrants and natives. Second, a high level of interethnic marriage is associated with decreasing dissimilarities in labor market outcomes of immigrants and natives. Furthermore, Muttarak (2004) points out that intermarriage is not only an indicator of integration but also a primary cause of it. Although the economic integration of immigrants has been the subject of a large literature, research on the effect of intermarriage on immigrants' economic integration[1] is scarce (Kantarevic 2004; Meng and Gregory 2005; Meng and Meurs 2009).

This study aims to investigate the role of interethnic marriage in immigrants' economic integration in the Netherlands. To this end, I measure the economic performance of immigrants in terms of wages. There are at least two reasons to consider intermarriage as one of the important factors that affect the integration process of immigrants. First, having a native spouse gives access to native networks. These networks are influential in creating productive social capital that may promote immigrants' economic integration. Second, intermarried immigrants can increase their human capital accumulation, especially language proficiency and knowledge of the local labor market, through their native spouses, thus contributing positively to their labor market productivity.

Estimating the causal impact of intermarriage on wages is complicated for several reasons. First, intermarriage can be endogenous in the wage equation. Intermarried immigrants might be a selected sub-sample from the population of all married immigrants (Kantarevic, 2004). For instance, they may have unobserved characteristics, such as physical appearance and social skills that could also affect their labor market outcomes. Another source of endogeneity

[1] I use the terms 'integration' and 'assimilation' interchangeably throughout the article. Interethnic marriage is defined as the union of an immigrant with a native. These unions may be in the form of legal marriages as well as de facto relationships.

might be reverse causality. Marrying a native increases the human capital accumulation of an immigrant, thus improving the immigrant's position in the labor market, or it can be argued that the causal relationship is the opposite: favorable labor market outcomes may increase the probability of marrying a native. To endogenize the interethnic marriage decision, I use group size and sex ratio as instruments. Group size is the ratio of the number of female immigrants in the ethnic group to the total female population in the Netherlands. Sex ratio represents the number of female immigrants relative to number of male immigrants in the ethnic group. While the former captures the availability of potential spouses of the same ethnic group, the latter captures the competition for those potential spouses. The rationale for the use of these instruments as follows: The larger the size of an immigrant group, the more likely an immigrant will meet a partner from his own ethnic group and the higher the probability of marrying endogamously. The skewed sex ratios, i.e., more competition for co-ethnic spouses, encourage intermarriage. Another estimation issue is the selectivity of employment status. The sample selection problem arises as wages are observed only when individuals work. The econometric model used in the empirical analysis accounts for the potential biases resulting from the endogeneity of intermarriage and sample selection.

The data used come from the SPVA, 'Social Position and Use of Public Utilities by Immigrants'. This survey provides information on the socio-economic and socio-cultural position of the four largest ethnic minorities in the Netherlands: Turks, Moroccans, Surinamese and Antilleans. The sample used in the empirical analysis consists of first-and second-generation male immigrants.[2]

The contribution of this paper to the relevant literature is threefold. First, due to data limitations the literature on the intermarriage premium considers the unions between first- and second-generation immigrants as intermarriage. Even though second-generation immigrants are closer to the native population, they do not necessarily have the same human and

[2]I use the term 'first-generation' to refer to people who were born abroad. Second-generation is defined as those who were born in the Netherlands but have at least one foreign-born parent. Natives refer to individuals who were born in the Netherlands with both parents also born in the Netherlands.

social capital as the native population does. Therefore, treating second-generation immigrants as natives may affect the estimates of intermarriage premia. In this study, I am able to distinguish second-generation immigrants from natives as the data contain information on the parents' country of birth. Furthermore, this information allows me to test whether or not the intermarriage premium varies across generations.

Second, an important limitation of the previous literature is its failure to address two key empirical issues simultaneously. I control for both endogeneity of intermarriage and sample selection by using the discrete factor method of Mroz (1999). The discrete factor method (hereafter DFM) is a full-information maximum likelihood technique in that DFM does not impose a priori distributional assumptions on the unobserved heterogeneity, but instead it approximates the distribution of the unobserved heterogeneity with a discrete distribution where mass points and probabilities are estimated. Mroz (1999) shows that when the true distribution of error terms is joint normal, DFM performs as well as alternative estimators that assume normality and DFM is superior to normality-based estimators when the underlying distribution is non-normal.[3] Moreover, DFM outperforms 2SLS in the presence of weak instruments.

Third, an additional weakness of the previous literature is that the instruments, group size and sex ratio are measured at the time of survey under the assumption that immigrants have been exposed to the same structural factors at the time of marriage. However, this assumption is not valid if marriages took place a long time ago and structural factors changed over time. The data set I used contains information on the year of marriage. Taking advantage of this information, I construct the instruments by matching an immigrant's year of marriage with the corresponding group size and sex ratio in that year. Therefore, I am able to assess the influence of structural factors on intermarriage more accurately, making my identification strategy more credible.

[3]Though the DFM allows the distribution of unobserved heterogeneity to approximate a normal distribution if the data suggest it, this technique is general enough to include other distributions for unobserved heterogeneity.

To the best of my knowledge, there are only three studies that analyze the relationship between intermarriage and economic assimilation. Meng and Gregory (2005) find that intermarried immigrants assimilate faster and earn more than their endogamously married counterparts in Australia. They report a substantial intermarriage premium, around 20 percent, for immigrants from non-English-speaking countries. Meng and Meurs (2009) also provide evidence that intermarried immigrants have significantly higher earnings than their co-ethnic married counterparts in France. On the other hand, Kantarevic (2004) does not find any causal relationship between earnings and intermarriage in the U.S. He tests two competing hypotheses: the productivity hypothesis and the selection hypothesis. According to the productivity hypothesis, immigrants married to native-born spouses assimilate faster than comparable immigrants who are married to foreign-born spouses. In contrast, the selection hypothesis claims that the relationship between intermarriage and economic assimilation is spurious. The effect of intermarriage on the economic assimilation disappears after accounting for selection into intermarriage. Kantarevic's empirical findings support the selection hypothesis.

The main findings of this study indicate that intermarriage has a positive effect on immigrants' labor market outcomes. Intermarried immigrants earn more than their co-ethnic married counterparts. Marrying a native is associated with a wage premium of seven percent. When the effect of intermarriage is allowed to differ between first-and second-generation immigrants, the results suggest that intermarriage has a positive but statistically insignificant impact on earnings among second generation immigrants.

The paper is organized as follows: Section 2 presents the theoretical concepts that explain why labor market outcomes differ between immigrants married to natives and those married within their own ethnic group. Section 3 provides background information on the migration history and labor market performance of the four ethnic minorities in the Netherlands. Section 4 describes the data and variables used in the empirical work. Section 5 introduces the econometric model and discusses the identification strategy. Section 6 presents the results

from empirical analysis with a discussion on robustness checks while Section 7 concludes.

2 Theoretical Framework

Social Capital Theory places emphasis on the role of social capital on immigrants' labor market success (Aguilera 2002, 2005). It is argued that immigrants are disadvantaged in the labor market because they have a social network mostly comprising members of their own ethnic group whose knowledge of the labor market is not as good as that of natives. Marrying a native gives access to native networks. Having access to these networks could reduce job search costs and increase the probability of finding a job with higher wages as natives are better informed about job openings and they are better positioned to find jobs. Therefore, Social Capital Theory suggests that intermarried immigrants have better labor market outcomes than those who are endogamously married.

The Family Investment Hypothesis (FIH) focuses on the role of family structure in the assimilation of immigrants (Worswick 1999; Baker and Benjamin 1997; Blau et al. 2003). The FIH suggests that the labor market assimilation profile differs across family types: immigrant families (i.e. where both spouses are immigrants) versus mixed families (i.e. where only one spouse is an immigrant). Since immigrant families are assumed to be credit-constrained, borrowing and investing activities are separated among family members such that one partner foregoes her/his own investment in human capital and takes dead-end jobs to finance the human capital investment of the other partner.[4] On the other hand, immigrant partners in mixed families who are assumed not to be credit-constrained do not need to perform the borrowing function for their families as do their counterparts in immigrant families. Therefore, immigrant partners in mixed families can invest in host country-specific human capital or afford to wait for a job with a higher wage. Those different investment

[4]Although in the empirical literature the male partner is typically considered to be the primary worker who invests in host country-specific human capital while the female partner is assumed to be the secondary worker who undertakes labor market activities to finance current consumption and the human capital accumulation of the primary worker, Cobb-Clark and Crossley (2004) define the primary worker as the partner with greater labor market skills and experience regardless of gender.

behaviors across family types may explain why the labor market outcomes differ between immigrants married to immigrants and immigrants married to natives.

According to Classical Assimilation Theory, the process of assimilation begins with acculturation, i.e., adopting the culture of the native groups such as norms, practices, and values and learning the native language. Gordon (1964) suggests that acculturation is followed by structural assimilation which is defined as 'large scale entry into the cliques, clubs and institutions of the core society'. Gordon argues that structural assimilation weakens ethnic attachment and increases contact between immigrants and natives, leading to an increase in intermarriage. Intermarriage, in turn, gradually erases cultural and socioeconomic features that had previously distinguished the immigrant groups from natives and reduces the social distance between immigrants and natives. In this framework, intermarriage is regarded as the final stage of assimilation. Therefore, Assimilation theory asserts that intermarriage has a positive effect on the labor market outcomes of immigrants through weakening ethnic attachment that hampers economic assimilation.

The Productivity Hypothesis in the marriage premium literature argues that married men can accumulate more human capital than their single counterparts through greater specialization and division of labor within a household, leading to an increase in their labor market productivity (Becker 1973). The productivity hypothesis explains why intermarriage promotes economic integration of immigrants for several reasons. First, a native spouse can contribute to human capital accumulation of an immigrant, especially by improving his language proficiency. Proficiency in the host country's language is a crucial factor in the process of immigrants' economic assimilation and social integration. Chiswick and Miller (1995, 2002) find that fluency in the host country's language increases the likelihood of immigrants' being employed and is associated with higher earnings. Second, a native spouse can also provide knowledge about the host country's customs, culture and informal rules regulating the labor market, increasing an immigrant's competitiveness in the labor market.

On the other hand, the Selection Hypothesis, an alternative to the Productivity Hypoth-

esis, claims that more productive men are self-selected into marriage; therefore the marriage premium vanishes when this selection issue is controlled for (Nakosteen and Zimmer 1987, Stratton 2002). In our settings, the Selection Hypothesis implies that intermarried immigrants do not become more assimilated after marriage; rather they were already more assimilated before marriage. Hence, the Selection Hypothesis predicts that intermarriage does not have a causal effect on the labor market outcomes.

3 Background

The Netherlands has been classified as a net-immigration country since the 1960s (Belevender and Veenman 2004). Foreigners, first-and second-generation immigrants, constituted 19.4% of the Dutch population in 2007.[5] Turks, Moroccans, Surinamese, and Antilleans are the four largest immigrant groups in the Netherlands and make up 66% of the immigrant population. First immigrants from Turkey and Morocco arrived in the Netherlands during the 1960s as a result of the Dutch government's attempt to alleviate the growing need for low-skilled workers in Dutch industry. Although the Dutch government stopped the labor recruitment program at the end of 1973, immigration from the above-named countries has continued as a result of increasing family reunification in the 1980s and marriage formation in the 1990s (Kalmijn and Van Tubergen 2007). Table 1 shows that Turk and Moroccan immigrants represented 2.3 % and 2% percent respectively of the Dutch population in 2007.

The first migration from Suriname and the Antilles, former colonies of the Netherlands, took place in the 1950s. More than 50,000 Surinamese came to the Netherlands due to political instability just before Suriname declared its independence from the Netherlands in 1975. The second immigration peak occurred in 1979 and 1980 shortly before the Dutch government introduced visa requirements to control the free movement of immigrants. At that time, approximately 30,000 Surinamese immigrated to the Netherlands. The prolonged

[5]In the publications of Statistics Netherlands, foreigners are defined as persons who are living in the Netherlands and of whom at least one parent is born abroad.

recession through the second half of the 1990s in the Antilles led to a large-scale Antillean immigration to the Netherlands (Bevelander and Veenman 2004). Table 1 indicates that the Antillean population in the Netherlands is much smaller than the Surinamese population; while 2% of the total population of the Netherlands are first-and second-generation Surinamese, the Antillean population represented only 0.8% of the Dutch population in 2007.

There are some specific differences between immigrants from Suriname and the Antilles (Caribbeans) and those from Turkey and Morocco (Mediterraneans). Because Suriname and the Antilles were former Dutch colonies, the Caribbean groups are more acquainted with the Dutch society or language than the Mediterranean groups who were not exposed to the Dutch culture and language before they immigrated to the Netherlands. Moreover, Turks and Moroccans are mostly Muslim as opposed to Caribbeans, especially Antilleans, who are mainly Christian. Kalmijin and Van Tubergen (2007) point out that Mediterraneans are a more closed group than Caribbeans in the Netherlands because of these differences in language and religion.

On average, the labor market position of immigrant workers lags behind that of the native Dutch. As far as employment levels are concerned, unemployment rates for the four main immigrant groups are several times as high as that for the native Dutch (Van Ours and Veenman 2005). Turks and Moroccans are the most disadvantaged, while the Caribbeans still have higher unemployment rates than the native Dutch. Bevelander and Veenman (2004) investigated the employment integration of ethnic minority males in the Netherlands from 1988 to 2002. They point out that after controlling for individual characteristics, both Caribbean and Mediterranean groups had lower employment chances in the years 1991, 1994, and 1998 compared to Dutch natives. However, Caribbeans had the same employment chances as the native Dutch in 1988 and 2002. The immigrant groups also are disadvantaged in terms of wages. Van Ours and Veenman (2005) indicate that the native Dutch have higher earnings than the aforementioned immigrant groups. The Antilleans most closely approximate the native Dutch and are followed by the Surinamese, the Turks and the Moroccans respectively.

In the literature there are several explanations for the disadvantageous labor market position of immigrants in the Netherlands. Studies focusing on the demand side of the labor market underline the existence of discrimination and prejudice in the Dutch labor market and the concentration of immigrants in particular industries that are affected more than others during recessions (Gras and Bovenkerk 1999; Kee 1995; Van Ours and Veenman 2005). Moreover, some studies emphasize the importance of institutional factors, such as policies on the labor market integration of immigrants (Reitzet et.al 1999). Van Ours and Veenman (2005) argue that the relatively poor employment position of ethnic minorities in the early 1980s in the Netherlands might be explained by the policy of the Employment Office that emphasized the preferences of employers rather than equal opportunity for workers.

On the other hand, studies focusing on the supply side of the labor market, in line with human capital theory, point out that immigrants usually lag behind the native Dutch in terms of education, country-specific skills, and language proficiency, resulting in a weak position in the labor market (Kee 1993). Applying a symmetric approach to the labor market in their analysis, Nieseng et al.(1994) suggest that half of the difference in employment chances between the native Dutch and ethnic minority groups in the Netherlands is attributable to differences in personal characteristics.

Another supply-side explanation emphasizes the impact of social capital on immigrants' labor market positions. Examining the role of human and social capital in immigrant employment and occupational status in the Netherlands, Van Tubergen and Kanas (2006) find that human capital has a positive impact on immigrants' labor market positions and the returns on host-country specific human capital are much higher than the returns on origin-country specific human capital. Although they find that immigrants who have more contacts with Dutch natives are less often unemployed and have a higher occupational status, this positive effect of social capital disappears considerably after human capital characteristics are controlled for. Conversely, Veenman (1998) suggests that in addition to education level, work experience and Dutch language proficiency, social contact with Dutch natives has a

positive impact on the employment integration of immigrants in the Netherlands.

4 Data and Descriptive Statistics

The data used in this study come from the SPVA, 'Social Position and Use of Public Utilities by Immigrants'. The SPVA is a large-scale, cross-sectional[6] immigrant-specific survey (Van Ours and Veenman 2003). The aim of this survey is to collect information on the socio-economic and socio-cultural position of the four largest ethnic minorities in the Netherlands: Turks, Moroccans, Surinamese and Antilleans. In the SPVA, the main respondents are heads of households who are interviewed through an extensive questionnaire. In addition, partners and offspring aged 11 years and older are interviewed with the shortened version of the main questionnaire. The SPVA surveys contain extensive information on immigrants' labor market outcomes, their migration history, education, and cultural attitudes.

The SPVA comprises random samples of the population in thirteen cities, including the four largest in the Netherlands.[7] The first survey was conducted in 1988 and repeated thereafter in 1991, 1994, 1998, and 2002. I restrict the sample to immigrants married after migration. The rationale for doing this is that immigrants who marry before they arrive in the Netherlands are not subject to the same structural factors as those who married in the Netherlands. For example, the sex ratio variable in the Netherlands in 1980 would be irrelevant for an immigrant who married in Morocco in 1980 and immigrated to the Netherlands in 1995. To apply this selection rule, I need information on the year of marriage and the year of immigration. This information was not collected for 1988 and 1991 surveys. Therefore, I make use of the surveys collected in 1994, 1998, and 2002.

The final sample includes both the first-and second-generation married male immigrants

[6]The SPVA contains a small panel, about 10% of the sample. I use the sub-samples of the SPVA surveys comprising those who were interviewed for the first time.

[7]For the 1994, 1998 and 2002 surveys, in addition to the four major cities in the Netherlands (Amsterdam, Rotterdam, The Hague, and Utrecht) Almere, Alphen aan de Rjin, Bergen op Zoom, Delft, Dordrecht, Eindhoven, Enschede, Hoogezand-Sappemeer, and Tiel are included.

aged between 25 and 60 years, married after migration and working full-time.[8] I use the term 'first generation' to refer to people who were born abroad and then immigrated to the Netherlands. Second-generation immigrants are defined as those who were born in the Netherlands but have at least one foreign-born parent. The lower and upper age categories chosen are based on the assumption that individuals older than 24 participate in the labor market since they have typically finished their studies by that age and that individuals older than 60 leave the labor market (Bevelander and Veenman 2004). After removal of cases with incomplete information, the final sample contains 1712 marriages.[9]

I am primarily interested in two types of marriages: (i) *Exogamous marriages*; in which the immigrant's partner is native Dutch;[10] and (ii) *Endogamous marriages*; in which the immigrant's partner or her parent(s) come from the same country as the immigrant. In 35 cases the information on the birthplace of parents-in-law is missing.[11] For these cases, I utilize information on the birthplace of the partner only. Table 2 presents the interethnic marriage rates among ethnic groups. The overall interethnic marriage rate in the sample is 11%. The Antilleans appear to have the highest percentage of intermarried immigrants while the Turks have the lowest intermarriage rate. The Caribbean groups, i.e., Surinamese and Antilleans, marry exogamously more often than the Mediterranean groups, i.e., Turks and Moroccans. There are several explanations for the high rates of intermarriage in Caribbean groups (Kalmijn and Van Tubergen 2007). First, as mentioned earlier, since Suriname and the Dutch Antilles were former colonies of the Netherlands, Caribbean groups are quite

[8] I restrict the analysis to men, because information on hourly wages is available only for the heads of households. Among Turks and Moroccans typically men and only a few women are heads of households. Ninety-six percent of employed males who are married after immigration and aged between 25 and 60 are full time workers. The full-time worker is defined as someone who reports that he is working 30 hours or more a week. The final sample includes the full-time workers only.

[9] It is important to note that the final sample consists of married immigrants only. I do not model the selection into marriage as it is quite difficult to find an instrument that is related to the probability of marriage but at the same time unrelated to the probability of intermarriage and does not have a direct impact on wages. Twenty percent of immigrants aged between 25 and 60 years are single.

[10] The final sample does not include exogamous marriages with the other ethnic groups for example, Surinamese men who are married to Turkish women. There are only 67 such marriages in the final sample.

[11] For those cases, some of the partners might be treated as native while they were actually second-generation immigrants.

familiar with the Dutch culture and language. On the other hand, the language background of the Mediterranean groups is different from that of the Caribbean groups. Turks and Moroccans were not exposed to the Dutch language before immigration. Immigrants who do not speak the language of the host country very well have fewer opportunities to interact with the native population and this would naturally decrease the propensity to intermarry. Second, religious affiliation is another important factor affecting the partner selection process. Unlike the Caribbean groups who are mainly Christian,[12] the Mediterranean groups are mostly Muslim. Islam stresses similarity in religious outlook as one of the most important traits in the partner selection process. Although in some circumstances, Muslim men are permitted to marry non-Muslim women, interfaith marriages are not recommended for the sake of religious compatibility between partners and the upbringing of children. A further explanation for the comparatively low rates of interethnic marriage in the Mediterranean groups is that among Turks and Moroccans, arranged marriages with the members of the same ethnic group are quite common. In this kind of marriage, partner selection is influenced by traditional family values.

Table 3 presents the definitions of the variables used in the empirical analysis. I use the logarithm of hourly wage as the measure of earnings. The hourly wage variable is constructed by the division of net monthly earnings from paid work by the monthly hours of work (weekly hours of work multiplied by four).[13] I distinguish four levels of education: (1) Primary education; (2) Lower secondary education; (3) Higher secondary education; and (4) University education. To control for ethnicity, I divide immigrant groups into two main categories: the Caribbean groups, i.e., Surinamese and Antilleans, and the Mediterranean groups, i.e., Turks and Moroccans, as they are homogenous in terms of religious, cultural and

[12]There are three main ethnic subgroups of the Surinamese population: (i)*East Indians* of whom the majority of this group are Hindu; (ii) *Indonesians* who are mostly Muslim; and (iii) *Creoles* who are predominantly Christian.

[13]It is also important to note that the currency used in the Netherlands before the Euro was the Dutch guilder. The Dutch guilder expired in January 28, 2002. Thus, the data on monthly earnings collected in the 1994 and 1998 waves were in Dutch guilders. For these survey years, the Dutch guilder is converted to Euro at the rate 1 euro=2.20371 Dutch Guilders. Individuals with net monthly earnings of less than 500 Euros are excluded from the analysis.

language background. In the SPVA, respondents were asked whether they have difficulty speaking Dutch. There are three possible answers: (1) Yes, always (2) Yes, sometimes (3) No, never. To measure language proficiency, I construct an indicator variable that takes on the value of 1 if the respondent's answer is 'No, never' and zero if his answer is 'Yes, always' or 'Yes, sometimes'.[14]

It is well-documented in the literature that the time spent in the host country, often referred to as years since migration, is an important factor influencing the assimilation process (Borgas 1994, 1999; Chiswick, 1991). Immigrants who have been in the host country for a longer time acquire more country-specific skills that affect their economic integration positively. I construct six categories for the years-since-migration variable: 0-5 years, 6-10 years, 11-15 years, 16-20 years and 21 years or more. Moreover, to control for possible variations in regional labor market conditions, I construct an indicator variable that takes on the value of 1 if the municipality in which the immigrant lives is large and zero otherwise.[15]

Next, I will discuss the construction of the two instruments used to endogenize the intermarriage decision. Sex ratio measures the number of group members of the opposite sex divided by the number of group members of the same sex aged 15-60.[16] Data are available for the period 1972-2002 on an annual basis.[17]

The sex ratio for an individual i from ethnic group e who married in year t can be

[14]Proficiency in the host country's language has been found to play an important role in the process of economic assimilation of immigrants (Chiswick 1991; Chiswick and Miller 1992,1995; Dustmann and Fabbri 2003). In my main analysis, I do not control for immigrants' Dutch language proficiency due to possible endogeneity between language fluency and earnings. In addition, the direction of causality between intermarriage and language fluency is ambiguous. While intermarriage increases language fluency, language fluency may also increase the likelihood of intermarriage (Meng and Gregory 2005). The complicated relationships among earnings, intermarriage and language fluency make the model more complex and necessitate a better dataset. In the Robustness Checks section, treating Dutch proficiency as exogenous in the wage equation, I examine how the results are affected when the indicator variable for speaking Dutch fluently is included in the wage equation.

[15]The binary variable Bigcity takes a value of 1 if the immigrant lives in one of the four major cities in the Netherlands, Amsterdam, Rotterdam, The Hague, and Utrecht, and zero otherwise.

[16]Examining the relationships between several measures of sex ratio and family formation, Fossett and Kiecolt (1991) find that the simple sex ratio based on the total population is highly correlated with that computed for unmarried population or narrow age ranges.

[17]The data are obtained from the database of Statistics Netherlands at http://www.cbs.nl/en-GB/default.htm

specified as:

$$SR_{it}^{e} = \frac{n_t^{fe}}{n_t^{me}} \tag{1}$$

where n_t^{fe} and n_t^{me} are the number of females and males aged 15-60 respectively in the ethnic group e in year t. To construct the sex ratio variable, I match the year in which the immigrant married with the relevant group-specific sex ratio for that year.

Group size represents the number of immigrants of the opposite sex relative to the total population of the opposite sex aged 15-60. The group size for individual i from ethnic group e who married in year t can be specified as:

$$GS_{it}^{e} = \frac{n_t^{fe}}{n_t^{f}} \tag{2}$$

where n_t^{fe} is the number of females aged 15-60 from ethnic group e in year t and n_t^{f} is the total number of females aged 15-60 in the Netherlands in year t. I construct this variable by matching an immigrant's year of marriage with the corresponding size of the immigrant group in that year. Because the data are not available for the pre-1972 period, for immigrants who married before 1972, I use the group size and sex ratio figures for 1972.[18] In the empirical analysis, I use the log transformation to reduce the degree of skewness in group size and sex ratio variables.

Figure 1 shows how the size of the immigrant groups has grown for the period 1972-2002 in the Netherlands, while Figure 2 presents the trend in sex ratios for each of the four groups for the same period. Figure 2 indicates that although the sex ratios for Turks and Moroccans were quite unbalanced at the beginning of the period, suggesting a dramatic shortage of female immigrants, they became more balanced over time. On the other hand, the Caribbean groups, i.e. Surinamese and Antilleans, initially have much more favorable sex ratios (close to one) than Turks and Moroccans do. Interestingly, the sex ratios for Surinamese turned out to be greater than one for the post-1980 period, indicating a shortage

[18]The marriage market might be geographically bounded. Since the data are not available at the province level, I measure sex ratio and group size at the nation level.

of men.

Table 4 presents the descriptive statistics by type of marriage, which lead to several main conclusions. First, intermarried immigrants earn more than their endogamously married counterparts. Second, intermarried and nonintermarried immigrants also differ in employment rate. In the intermarried sample, 72% of immigrants are employed while in the nonintermarried sample the employment rate is 54%. Third, the Caribbean groups, i.e., Surinamese and Antilleans, are more likely to intermarry than the Mediterranean groups, i.e., Turks and Moroccans. Fourth, intermarried immigrants are generally more educated, have spent more years in the Netherlands, and live in smaller cities compared to nonintermarried immigrants. Further, seventy-four percent of intermarried immigrants speak Dutch fluently, whereas the rate for co-ethnic married immigrants is only 38%. Finally, the generational composition of the intermarried sample is different from that of the non-intermarried sample. Second-generation immigrants seem more likely to marry exogamously than first-generation immigrants. To sum up, because the intermarried and nonintermarried immigrants differ in certain characteristics that might also affect their labor market outcomes, I control for these differences in the wage equation.

5 Econometric Framework

The objective of this paper is to investigate whether immigrants married to a native have better labor market outcomes than those married within their own ethnic group. To this end, the baseline earnings equation for immigrants is specified as follows:

$$ln(w_i) = X'_{1i}\beta_1 + \delta I_i + u_{1i} \tag{3}$$

where $ln(w_i)$ is the natural log of hourly wage for individual i; *intermarriage*, denoted by I, is a binary variable that takes on a value of one if the immigrant married a native and zero if the immigrant married within his own ethnic group. X_{1i} represents a vector of other

control variables,[19]and u_{1i} is an error term.

The term δ is the primary coefficient of interest and measures the intermarriage premium. Two important issues need to be addressed in order to obtain a consistent estimate of δ. First, intermarriage can be endogenous in the wage equation. It is possible that intermarried immigrants have unobserved characteristics, such as physical appearance or social skills that may also affect their labor market outcomes.[20] In order to take this endogeneity problem into account, I incorporate the intermarriage equation into the model. The intermarriage decision is specified in terms of the latent variable I_i^*:

$$I_i^* = X_{2i}'\beta_2 + u_{2i} \tag{4}$$

$$I_i = \begin{cases} 1 & \text{if } I_i^* \geq 0 \\ 0 & \text{if } I_i^* < 0 \end{cases}$$

where u_{2i} is an error term and X_{2i} is a vector of exogenous variables that affect the intermarriage decision. X_{2i} includes personal characteristics of individual i and instrumental variables that affect the intermarriage decision but do not have a direct wage effect. Exclusion restrictions through which the model is identified are discussed in the next section.

The other key empirical issue arises from the fact that wages are observed only for those who are working. To control for selection into employment status, the choice of employment status can be specified in terms of the latent variable E_i^*:

$$E_i^* = X_{3i}'\beta_3 + u_{3i} \tag{5}$$

$$E_i = \begin{cases} 1 & \text{if } E_i^* \geq 0 \\ 0 & \text{if } E_i^* < 0 \end{cases}$$

[19] X_{1i} includes age (and its squared term), educational level, years since migration, Caribbean, bigcity, and survey year indicators. See Table 3 for the definitions of control variables.

[20] For example, Hammermesh and Biddle(1994) show that physical appearance has a significant effect on earnings.

where the employment variable E_i equals one if the immigrant is employed at the time of survey and zero otherwise. X_{3i} is a vector of control variables that affect the choice of employment status and u_{3i} is an error term.

To account for both endogeneity and sample selection, this study uses the Discrete Factor Method (DFM) proposed by Mroz (1999). Although DFM is similar to standard full information maximum likelihood, it does not make a multivariate parametric assumption about the distribution of the error terms. Instead DFM imposes fewer assumptions on the distribution of correlated error terms. Using DFM, error terms can be decomposed into two parts: a component that is common to all equations and random variation.

$$\begin{aligned} u_{1i} &= \alpha_1 \upsilon + \epsilon_{1i} \\ u_{2i} &= \alpha_2 \upsilon + \epsilon_{2i} \\ u_{3i} &= \alpha_3 \upsilon + \epsilon_{3i} \end{aligned} \tag{6}$$

Under this specification, υ represents unobserved heterogeneity that results in the correlation across equations. Factor loadings (α) that can be interpreted as the coefficient of unobserved heterogeneity vary across equations. The ϵ_{1i}, ϵ_{2i} , ϵ_{3i} and υ are mutually independent and are independent of explanatory variables in the model. Moreover, ϵ_{2i} and ϵ_{3i} are assumed to follow a standard normal distribution and ϵ_{1i} is assumed to be distributed normally with mean zero and standard deviation σ.

DFM approximates the distribution of the unobserved heterogeneity with a discrete distribution. The distribution of υ is specified as follows:

$$P(\upsilon = \eta_k) = p_k \quad \text{for } k = 1....K \tag{7}$$

$$p_k \geq 0 \quad \forall k \quad \text{and} \quad \sum_{k=1}^{K} p_k = 1$$

where η_k are the points of support of the distribution and p_k represents the probability that υ takes on the value of each support point. The parameters of this distribution are estimated jointly with the other parameters describing the model. There is no standard theory regarding how to identify the optimal number of points of support in a finite sample. Most of the studies add points of support until the likelihood fails to improve significantly.[21] Mroz (1999) provides Monte Carlo evidence that this likelihood ratio test performs well when determining the number of points of support.

Since each equation in the model includes an intercept and factor loading parameters are estimated, the location and scale of the distribution of υ are not identified. Therefore, I normalized the values of first and last mass points to 0 and 1 respectively. Further parametrization of points of support is specified as follows to make η_k fall on $[0, 1]$.

$$\eta_k = \frac{exp(\theta_k)}{1 + exp(\theta_k)} \quad \text{for } k = 2....K-1 \tag{8}$$

The probabilities are also parameterized such that they satisfy the nonnegativity conditions and the adding up constraint as follows:

$$p_k = \frac{\tau_k}{\sum_{k'=1}^{K} \tau_{k'}} \quad \text{where} \quad \tau_k = \begin{cases} exp(\psi_k) & \text{for } k = 1, 2....K-1 \\ 1 & \text{for } k = K \end{cases} \tag{9}$$

Conditional on υ, the joint density function of u_{1i} , u_{2i} and u_{3i} can be written as

$$f(u_{1i}, u_{2i}, u_{3i}|\upsilon) = \frac{1}{\sigma}\phi(\frac{u_{1i} - \alpha_1\upsilon}{\sigma})\phi(u_{2i} - \alpha_2\upsilon)\phi(u_{3i} - \alpha_3\upsilon) \tag{10}$$

where ϕ is the standard normal density function. If the cumulative distribution function

[21] See Picone et.al (2003); Mocan and Tekin (2002); Tekin (2007); and Van Ours (2007)

of υ is F(υ), then the unconditional joint distribution of errors is as follows:

$$f(u_{1i}, u_{2i}, u_{3i}) = \int f(u_{1i}, u_{2i}, u_{3i}|\upsilon)\, dF(\upsilon) \tag{11}$$

Since the cumulative distribution of υ is approximated by the discrete distribution specified above, the joint distribution of errors is

$$f(u_{1i}, u_{2i}, u_{3i}) = \sum_{k=1}^{K} p_k \frac{1}{\sigma}\phi(\frac{u_{1i} - \alpha_1\upsilon}{\sigma})\phi(u_{2i} - \alpha_2\upsilon)\phi(u_{3i} - \alpha_3\upsilon) \tag{12}$$

Mroz (1999) provides Monte Carlo evidence on the performance of the Discrete Factor Method relative to other estimation methods. When the true distribution of error terms is joint normal, discrete factor estimators perform as well as estimators that assume the error terms are normally distributed. DFM estimators also perform better than the normality-based estimators when the underlying distribution of error terms is not normal. In addition, Mroz (1999) demonstrates that the DFM outperforms 2SLS in the presence of weak instruments. The likelihood function associated with the model can be written as

$$\begin{aligned} L \;=\; & \prod_{i=1}^{N}\sum_{k=1}^{K} p_k \Phi(X'_{3i}\beta_3 + \alpha_3\eta_k)^{E_i}(1 - \Phi(X'_{3i}\beta_3 + +\alpha_3\eta_k))^{1-E_i}\Phi(X'_{2i}\beta_2 + \alpha_2\eta_k)^{I_i} \\ & \times(1 - \Phi(X'_{2i}\beta_2 + \alpha_2\eta_k))^{1-I_i}[\frac{1}{\sigma}\phi(\frac{lnw_i - X'_{1i}\beta_1 - \delta I_i - \alpha_1\eta_k}{\sigma})]^{E_i} \end{aligned} \tag{13}$$

where N is the sample size, K is the number of points of support to be used in approximating the distribution of the unobserved heterogeneity and Φ denotes the standard normal cdf.

5.1 Identification

The identification of the model is achieved through nonlinearities in the likelihood function and exclusion restrictions. In the literature of intermarriage, there are three general factors that determine intermarriage: individual preferences, influence of third parties and

opportunities to meet co-ethnics and members of other groups (Kalmijn 1998, Van Tubergen and Maas 2006). People search for potential spouses with certain socioeconomic and cultural characteristics. Therefore, intermarriages are, first of all, an outcome of people's preferences. Third parties such as family and religious community might also play an important role in the partner selection process. For example, parents might pressure their children to consider only prospective spouses with similar ethnic backgrounds or some religious groups may oppose interfaith marriages. Researchers have argued that in addition to preferences and third parties, levels of intermarriage between groups are affected by opportunities to meet co-ethnics and members of other groups. Demographic and structural factors such as the size of the group, sex ratio and residential segregation influence people's daily opportunities for meeting co-ethnics and those of other groups (Blau and Schwartz 1984). To endogenize the interethnic marriage decision, I use two important structural determinants of intermarriage as instruments: group size and sex ratio. A number of studies indicate that the size of the immigrant group is an important factor affecting intermarriage decisions (Qian and Lichter 2001; Blau and Schwartz 1984; Hwang et al. 1997). The larger the size of the immigrant group, the more likely for an immigrant to meet a partner from his own ethnic group, and the lower the probability of marrying exogamously. Moreover, Van Tubergen and Kalmijn (2007) point out that immigrants who are members of a larger group more strongly identify themselves with that group and could be more easily influenced by third parties which may also increase the chance of endogamy among the members of that group. In this study the group size represents the number of immigrants of the opposite sex relative to the total population of the opposite sex. It captures the availability of spouses of the same ethnic group. The sex ratio in the marriage market is another structural factor discussed in the literature. It is defined as the number of group members of the opposite sex divided by the number of group members of the same sex. The more skewed the sex ratio in an ethnic group, the more likely immigrants will look for potential partners outside their own ethnic group (Angrist 2002; Blau and Schwartz 1984). In other words, a shortage of marriageable co-ethnics of the

opposing sex promotes intermarriage.

The existing literature on the intermarriage premium also uses the structural factors of a marriage market (group size and sex ratio) as instruments. In the literature instruments are based on the measures of structural factors at the time of survey. However, it is quite possible that immigrants have been exposed to different demographic factors at the time of marriage because they married a long time ago. Figures 1-2 clearly show that the sex ratio and group size have changed over time. For example, a Turkish immigrant who was surveyed in the 2002 wave but married in 1980 had been subject to very different structural factors at the time of marriage than those at the time of survey.[22] Because the data set contains information on the year of marriage, I construct the instruments by matching an immigrant's year of marriage with the corresponding group size and sex ratio in that year. This allows me to examine the influence of structural factors, group size and sex ratio on intermarriage more accurately.

The credibility of the identification strategy hinges on the assumption that the instruments are valid. If instruments are valid, then: (i) they must be significant determinants of the interethnic marriage decision; and (ii) they must not be determinants of labor market outcomes. As discussed above, group size and sex ratio are both important factors affecting interethnic marriage decisions. Moreover, it is unlikely that an immigrant's wage depends on the ratio of the number of group members of the opposite sex to the number of group members of the same sex. On the other hand, one can argue that the size of an immigrant group might affect the development of networks among its members. If ethnic networks had any effect on labor market earnings of immigrants, then the group size variable might not be a valid instrument. To tackle this issue, I construct a new group size variable, *group size survey*. Unlike the group size instrumental variable measured at the time of marriage, the *group size survey* reflects the situation at the time of survey. To test whether the size of the

[22] Analyzing the determinants of ethnic intermarriage among first-generation immigrants in the Netherlands, Van Tubergen and Maas (2007) underscore the importance of this methodological problem. They argue that the contextual characteristics have to be measured at the time of marriage to examine their effects on intermarriage.

immigrant group affects labor market earnings of immigrants, I estimate the wage equation by including the new group size variable. The coefficient of the new group size variable is not statistically significant at conventional levels. Therefore, even if there is a correlation between these two group size variables,[23] *group size instrument* and *group size survey*, the insignificant coefficient of the *group size survey* variable in the wage equation ensures that the *group size instrument* is validly excluded from the wage equation. To identify the employment equation, I use the number of children as an exclusion restriction. I assume that the number of children affects the probability of being employed but does not have a direct effect on the wage.

6 Results

I begin estimating the wage equation with OLS, treating intermarriage as exogenous. Table 5 presents the OLS results of the wage equation. These results may be viewed as a benchmark that sheds light on whether the effect of intermarriage on earnings persists after other observed factors are controlled for. First, I regress log hourly wages on the indicator of intermarriage status to estimate the raw intermarriage premium. In the second specification, I control for age, education level, years since migration, place of residence, *Bigcity*, ethnic groups, *Caribbean*, and survey year effect. In the third specification, I add a binary variable indicating immigrant generation, *Second Generation*, and an interaction variable which is created by the product of intermarriage and second generation dummies to test whether the intermarriage premium varies across generations.

The following conclusions can be drawn from Table 5. First, the raw intermarriage premium is 18 percent. However, the descriptive statistics in Table 4 suggest that intermarried and non-intermarried immigrants differ in certain characteristics that might also affect their earnings. After controlling for these differences, I find that the variable of primary interest,

[23] The correlation coefficient between the two variables is 0.08, 0.17, 0.12, and 0.18 for Turks, Moroccans, Antilleans and Surinamese respectively.

intermarriage, has a positive and significant coefficient. Second, most of the other variables have the expected signs and are statistically significant. Educational level has a positive impact on earnings. The Caribbean groups, i.e., Surinamese and Antilleans, earn on average six percent more than the Mediterranean groups, i.e. Turks and Moroccans. In line with the earlier studies, immigrants who have been in the Netherlands for a longer period have higher earnings. The estimated coefficients on survey-year indicator variables indicate that hourly earnings are approximately 24% and 10% higher in 2002 and in 1998 respectively than those in 1994.[24] Finally, in the third specification the coefficients of second generation and interaction variables are not statistically significant.

Table 6 presents the DFM estimation results.[25] First, I estimate the model ignoring sample selection and only accounting for the endogeneity of intermarriage. The first two columns of Table 6 report parameter estimates from this endogeneity-only model. The estimation results show that the main coefficient of interest, the coefficient of the indicator for intermarriage, has the expected positive sign and is statistically significant at the 1% level. Marrying a native is associated with a wage premium of seven percent. The remaining parameter estimates of the wage equation are mostly of the expected sign and significance. Hourly earnings are an increasing and concave function of age. Education and years since migration have a positive and statistically significant effect on wages. The Caribbean groups have significantly higher earnings than the Mediterranean groups. Most importantly, the estimated coefficients of the identifying instruments, sex ratio and group size, have the expected signs and are statistically significant at the 1% level, suggesting that the larger the size of an immigrant group and the less skewed the sex ratio, the more likely immigrants are to marry endogamously.

[24] Consumer Price Index (1994=100) in the Netherlands is 108.2 and 122.5 in 1998 and 2002 respectively (Statline, Statistics Netherlands).

[25] The DFM results are based on three points of support. I do not reject the model with three points of support in favor of the model with four points of support. The likelihood ratio test suggested by Mroz (1999) with two degrees of freedom at the 25 % significance level is used to determine whether or not the additional mass point should be added. It is also important to note that when I use four points of support, the estimated intermarriage premium is very similar to that based on three points of support.

The last three columns of Table 6 report parameter estimates from the DFM accounting for both endogeneity of intermarriage and sample selection. The introduction of sample selection into the model hardly affects the estimated intermarriage premium. Most of the other coefficients are consistent across the two models. Compared to the reference group (primary school graduates), college graduates have a higher probability of being intermarried. A comparison of the first and fourth columns of Table 6 reveals that when selection into employment is accounted for, the estimated coefficient of the caribbean variable turns out to be statistically significant, suggesting Caribbean groups are more likely to be intermarried than Mediterranean groups. The number of children serves as an exclusion restriction in the employment equation. The coefficient of the children variable is negative and statistically significant at the 1% level, indicating that the likelihood of being employed decreases with the number of children. A possible explanation for this result could be as follows. The Netherlands has had a child benefit system since the 1940s.[26] All residents with dependent children, regardless of their income, are entitled to child benefits. The amount of the child benefit depends on the number of children in the family and their ages. The higher the number of children an immigrant has the higher the child benefit he could receive. Therefore, the extra income obtained from the child benefit may exert a negative effect on the employment decision of an immigrant. Table 7 reports the parameters describing the unobserved heterogeneity. The estimated mass points and the associated probabilities indicate that the most of mass of the distribution of the unobserved heterogeneity is given to two points: 0 with probability 0.46 and 1 with probability 0.52. These results suggest that estimation methods assuming normality may not be appropriate, as the distribution of unobserved heterogeneity does not approximate a normal distribution.

To test whether the intermarriage premium differs between first-and second-generation immigrants, I include an indicator for immigrant generation and its interaction with the intermarriage variable in the wage equation. Table 8 presents the DFM estimation results

[26]The Social Insurance Bank (Sociale Verzekeringsbank, SVB) implements the arrangement and payment of the child benefit. For more information see the following link: $http://www.svb.nl$

of this specification. Intermarried first-generation immigrants earn 6.7 percent more than their co-ethnic married counterparts. The point estimate of the sum of the coefficients on intermarriage and interaction variables is positive, though not significantly so, implying that there is no intermarriage premium for second-generation immigrants.[27]. This finding is in line with my expectations. As second-generation immigrants are closer to the native population in terms of social and human capital, they may not gain from their Dutch spouses as much as first-generation immigrants.

6.1 Robustness Checks

I test the robustness of the results to change in the estimation techniques. First, I estimate the wage equation using two-stage least squares (2SLS) to account for the possible endogeneity of intermarriage. The first column of Table 9 presents the first stage estimation results while the second column presents the second stage results for 2SLS estimation of the wage equation. The estimated coefficients of the instruments have the expected sign and are statistically significant at the 1% level. The F-statistic of excluded instruments implies that instruments are jointly significant.[28] The joint and individual significance of the instruments in the first stage regression ensures that the instruments are relevant. The results of overidentification testing, the Hansen J statistic, indicate that the null hypothesis that all instruments are uncorrelated with the error term can not be rejected. In the wage equation, the main coefficient of interest, the coefficient of intermarriage, is positive and statistically significant at the 10% level. The estimated coefficient implies that intermarried immigrants on average earn 19% more than their endogamously married counterparts. [29]

Second, as an alternative to the DFM to account for the endogeneity of intermarriage

[27]The t-statistics for the null hypothesis that the coefficients on intermarriage and interaction variables sum to zero is 0.947.

[28]Stock, Wright and Yogo (2002) suggest that the first-stage F statistic can be used to test whether instruments are weak or not. 2SLS inference is reliable when the first-stage F statistic is large. As a rule of thumb, for one endogenous regressor an F statistic less than 10 is cause for concern.

[29]After accounting for the endogeneity of intermarriage, I observe an increase in the intermarriage premium suggesting that unobservable characteristics that are positively correlated with the likelihood of intermarriage have a negative effect on earnings.

and selection into employment, I use a three-equation model (TEM) proposed by Wooldridge (2002).[30] The estimation results of the TEM appear in Table 10. The model is identified using the same exclusion restrictions as the DFM model. The first column in table 10 presents probit estimation of the employment equation. The coefficient of children is negative and statistically significant at the 1 percent level. The second and third columns of Table 10 report the results for 2SLS estimation of the structural wage equation. The coefficients of the instruments, sex ratio and group size, have the expected signs and are both individually and jointly significant. The instruments also pass the Hansen J overidentification test. In the wage equation, the coefficient of intermarriage is positive and statistically significant at the 5 percent level. Intermarried immigrants earn on average 23 percent more than their co-ethnic married counterparts. Consistent with the DFM results, the coefficient of inverse mills ratio is not statistically different from zero in the wage equation, suggesting that there is no evidence of sample selection bias. In sum, 2SLS and TEM estimation results indicate that the positive effect of intermarriage on earnings is robust to change in the estimation technique.

Table 11 presents OLS, 2SLS, TEM and DFM estimation results for the wage equation. Given that there is no evidence of sample selection bias, it is instructive to compare the

[30]The model consists of the following three equations:

$ln(w_i) = x'_{1i}\delta_1 + \alpha I_i + \epsilon_{i1}$ Wage Equation (1)

$I_i = x'_{2i}\delta_2 + \epsilon_{i2}$ Intermarriage Equation (2)

$E_i = 1(x'_{3i}\delta_3 + \epsilon_{i3} > 0)$ Selection Equation (3)

The first equation is the structural equation of interest. $ln(w_i)$ is the log of hourly wage for individual i, I_i is a potentially endogenous variable in equation 1 that takes on a value of one if the immigrant married a native and zero if the immigrant married within his own ethnic group. The second equation is a linear projection for the endogenous variable I_i and the third equation is the selection equation, where E_i equals one if the immigrant is employed and zero otherwise. The estimation strategy of the three-equation model can be summarized as follows. First, the selection equation is estimated as a probit model. After obtaining the estimated coefficients $\hat{\delta}_3$, I calculate the inverse mills ratios $\hat{\lambda}_{i3} = \frac{\phi(x'_{3i}\hat{\delta}_3)}{\Phi(x'_{3i}\hat{\delta}_3)}$. Second, I estimate the wage equation below by 2SLS, using instruments ($\hat{\lambda}_{i3}$, x_{2i}).

$ln(w_i) = x'_{1i}\delta_1 + \alpha_1 I_i + \gamma_{i3}\hat{\lambda}_{i3} + \nu_i$ (4)

It is important to note that all exogenous variables appear in the selection equation and are used as instruments in estimating equation (4) by 2SLS. The identification requires that at least two elements of x_{2i} not be in x_{1i} (i.e., I should have at least one instrument for I_i and another exogenous variable that determines selection). The hypothesis of no selection problem can be tested using the usual 2SLS t-statistic for the coefficient of the estimated inverse mills ratio, $\hat{\gamma}_{i3}$, in equation (4). The hypothesis of no selection problem requires that the coefficient of $\hat{\lambda}_{i3}$ is statistically insignificant.

DFM results to the results employing 2SLS. Several points are worth highlighting. First, the intermarriage premium from the 2SLS method is more than twice as large as that predicted by the DFM and only marginally significant at the 10% level. The DFM leads to an increase in efficiency. Although the main coefficient of interest changes significantly when DFM is employed, most of the other parameter estimates are close to those from the 2SLS.

One would argue that immigrants married to natives may be more likely to be employed than those married within their own ethnic group. Using DFM, I reestimate the model in which the employment equation includes the intermarriage variable. The coefficient of the intermarriage variable in the employment equation is positive but not statistically significant at the conventional levels. Furthermore, the estimated coefficient of the intermarriage variable in the wage equation is unaffected when I control for the intermarriage variable in the employment equation.[31]

To address the possibility that group size is not a valid instrument, I reestimate the model with sex ratio as the only instrument. I find that the intermarriage premium is statistically significant and very similar to that reported in Table 6. Moreover, the coefficient of sex ratio variable is highly significant in the intermarriage equation.

The literature suggests that lack of host country language proficiency is punished in the labor market (Chiswick 1991; Chiswick and Miller 1992, 1995; Dustmann and Fabbri 2003). Treating the Dutch proficiency variable as exogenous in the wage equation, I reestimate the model using DFM. The impact of intermarriage remains positive and statistically significant at the 1% level. Dutch proficiency has a positive and significant effect on the earnings.[32] It is also important to note that when language proficiency is included in the wage equation, the coefficient of the Caribbean indicator variable decreases in magnitude and significance, indicating that the Caribbean groups have a better Dutch proficiency than the Mediterranean

[31]The coefficient of intermarriage in the wage equation is 0.073 and is statistically significant at the 1% level.

[32]In the wage equation, the estimated coefficient of intermarriage is 0.071(s.e=0.021), while the coefficient of Dutch proficiency is 0.048(s.e=0.019).

groups.[33] This finding is consistent with the fact that the official language in Suriname and the Antilles is Dutch.[34] I also examine the sensitivity of results if second-generation wives are treated as natives.[35] The estimated intermarriage premium is slightly higher than that reported in Table 6.[36]

Meng and Gregory (2005) argue that intermarried immigrants might have higher quality labor-market attributes at the time of arrival than those who are endogamously married. If this is the case, the intermarriage premium may be a reward for better labor market quality at the time of arrival rather than reflecting economic assimilation. Although the empirical strategy I have followed accounts for the possible endogeneity of intermarriage, it is worth distinguishing the assimilation effect from the cohort quality effect at the time of arrival to test whether intermarriage results in a faster assimilation process.

Using the basic premise of the Borjas (1985) methodology, I estimate the following wage equation to distinguish the assimilation effect from the cohort quality effect at the time of arrival:

$$ln(w_{it}) = X'_{it}\pi_1 + Y'_{it}\pi_2 + \varphi_1 YSM_{it} + \varphi_2 YSM_{it}I_{it} + C'_{it}\lambda_1 + C'_{it}I_{it}\lambda_2 + \varepsilon_{it} \qquad (14)$$

where $ln(w_{it})$ is the natural log of the hourly wage for individual i at time t; X_{it} represents the vector of variables including age (and its squared term), education level, ethnic groups, and place of residence. The vector of survey year indicator variables Y_{it} is included to capture period effects. I assume that the period effects on earnings are the same for both intermarried and co-ethnic married immigrants. The variable YSM represents the number of years the immigrant has been in the Netherlands. Intermarriage, denoted by I_{it}, is a binary variable that takes on a value of one if the immigrant married a native and zero if the immigrant

[33] In the wage equation, the coefficient of the Caribbean indicator variable is 0.038 (s.e=0.019).

[34] Although the official language is Dutch; other languages spoken in Suriname include Suriname Javanese, Sarnami Hindustani, and several Amerindian and Creole languages. In the Antilles, Papiamento, Spanish, Creole English and other local languages are also spoken.

[35] There are only 46 second-generation wives in the final sample.

[36] The coefficient of intermarriage is 0.075 and statistically significant at the 1 percent level.

married within his own ethnic group. C_{it} is a set of indicator variables representing different immigration arrival cohorts: immigrants arriving in 1973-1980, 1981-1990, and 1991-2002. Immigrants arriving before 1973 are used as the reference group. The variables $YSM_{it}I_{it}$ and $C'_{it}I_{it}$ refer to interaction between the intermarriage indicator variable and YSM_{it} and C_{it} respectively. ε_{it} is an error term.

The coefficient vector λ_2 measures whether cohort quality effects differ between intermarried and co-ethnic married immigrants, while φ_2 captures the difference in the assimilation process across two groups. The estimation results of Equation (14) are reported in Table 12. The results are based on the sample of first-generation immigrants only. The coefficients of cohort indicator variables and their interaction with intermarriage are not statistically significant, suggesting that intermarried immigrants do not have better initial labor market quality than their nonintermarried counterparts. Like Meng and Gregory (2005), I find that the intermarriage premium increases with time spent in the Netherlands, even after controlling for the cohort quality effect at the time of arrival. The estimated coefficient on $YSM_{it}I_{it}$ implies that intermarried immigrants earn 0.2% more than their co-ethnic married counterparts for every additional year that they spend in the Netherlands. The steeper years since migration-earnings profile for intermarried immigrants provides evidence that intermarriage promotes immigrants' economic integration.

7 Conclusion

This study examines the relationship between interethnic marriage and the labor market integration of immigrants. Using the data from the Dutch Survey, I investigate whether intermarried immigrants perform better in the labor market than their co-ethnic married counterparts. Social Capital Theory, Productivity Hypothesis, and Assimilation Theory suggest that intermarriage promotes the economic and social integration of immigrants by increasing their knowledge about the host country's culture, language and rules regulating

the labor market. Also, having a native spouse gives access to native networks. The networks may enhance social capital accumulation, which positively influences economic success of an immigrant. In line with these theoretical concepts, the empirical findings indicate that intermarriage has a positive effect on immigrants' labor market outcomes. After controlling for the possible biases resulting from the endogeneity of intermarriage and selection into employment, I find that intermarried immigrants earn, on average, seven percent more than their endogamously married counterparts. The positive effect of intermarriage on earnings is robust to changes in the estimation method and specification of the model.

Unlike previous studies that treat second-generation immigrants as natives, this study distinguishes second-generation immigrants from natives, as the data set contains information on the parents' country of birth. This information allows me to test whether or not the intermarriage premium varies across generations. I expect that second-generation immigrants do not gain from their Dutch spouses as much as first-generation immigrants, because second-generation immigrants who attended school in the Netherlands and have more skills in the Dutch language are more likely to be closer to Dutch culture than first-generation immigrants. Consistent with my expectations, the empirical findings show that there is no intermarriage premium for second-generation immigrants.

Due to data limitations, I restrict my analysis to immigrant men. However, the sociological literature suggests that there are large gender differences in social network composition. Men tend to have more diverse and extensive networks and more able to use them as instrumental resources (Moore 1990; Marsden 1987). As Social Capital Theory implies that intermarried immigrant women receive more gain from a Dutch spouse than their male counterparts, it would be interesting to test whether the intermarriage premium varies by gender. To get a better understanding of the mechanisms that explain why the marriage pattern of immigrants is such an important element in the integration process, it would also be interesting to include the sample of immigrants married to immigrants from other ethnic groups into analysis.

References

Aguilera, M.A.(2002), "The impact of social capital on labor force participation:Evidence from the 2000 Social capital benchmark Survey," *Social Science Quarterly*, 83: 853-878.

——— (2005), "The Impact of Social Capital on the Earnings of Puerto Rican Migrants," *The Sociological Quarterly*, 46: 569-92.

Angrist, J.(2002), " How Do Sex Ratios Affect Marriage And Labor Markets? Evidence From America's Second Generation," *Quarterly Journal of Economics,* 107(3): 997-1038.

Baker, M., and Benjamin D.(1997), "The Role of the Family in Immigrants' Labor-Market Activity: An Evaluation of Alternative Explanations, "*American Economic Review,* 87(4): 705-27.

Blau, P.M., and Schwartz, J.E.(1984), "Croscutting social circles: Testing a macrostructural theory of intergroup relations", New York: Academic Press.

Blau, F.D., Kahn, L.M., Moriarty, J.Y., and Souza, A. (2003), "The Role of the Family in Immigrants' Labor-Market Activity: An Evaluation of Alternative Explanations: Comment," *American Economic Review,* 93(1): 429-447.

Becker, Gary S. (1973),"The Role of the Family in Immigrants' Labor-Market Activity: An Evaluation of Alternative Explanations," *Journal of Political Economy,* 81(4): 813-46.

Bevelander, P., and Veenman, J.(2003), "Variation in Perspective: The Employment Success of Ethnic Minority Males in the Netherlands, 1988-2002", *International Migration*, 42(4): 36-64.

Borjas, George J.(1985), "Assimilation, Changes in Cohort Quality, and Earnings of Immigrants," *Journal of Labor Economics*, 3: 463-489.

——— (1994), "The economics of immigration," *Journal of Economic Literature*, 32(4): 1667-1717.

——— (1995), "Assimilation, Changes in Cohort Quality revisited: what happenned to immigrant earnings in the 1980s?," *Journal of Labor Economics*, 13(2): 201-245.

——— (1999), "Economic Assimilation," *Handbook of labor Economics*, Volume3: Elsevier Science B.V., 1721-1740.

Chiswick, Barry R. (1978), "The effect of Americanization on the earnings of foreign-born men," *Journal of Political Economy*, 86(5): 897-921.

——— (1991), "Speaking, reading, and earnings among low-skilled immigrants," *Journal of Labor Economics*, 9: 149-70.

Chiswick, Barry R., and Miller, Paul W. (1992), "Language in the labor market: The immigrant experience in Canada and the United States," *Immigration, language and ethnic issues: Canada and United States*, ed. Barry R.Chiswick, 229-296. Washington, DC: American Enterprise Institute.

——— (1995), "The Endogeneity between Language and Earnings: International Analyses," *Journal of Labor Economics*, 13(2): 246-88.

Cobb-Clark, D. and Crossley Thomas F.(2004),"Revisiting the family investment hypothesis," *Labour Economics*, 11(3): 373-393.

Dustmann, C.,and Fabbri, F. (2003), "Language Proficiency and Labour Market Performance of Immigrants in the UK," *The Economic Journal*, 113(July): 695-717.

Gras, M. and Bovenkerk F. (1999), "Migrants and Ethnic minorities in the Netherlands: discrimination in access to employment," *Migrants, Ethnic minorities and the Labor market: Integration and Exclusion in Europe*, London, MacMillan Press.

Fossett, M.A., and Kiecolt, K.J. (1991), "A Methodological Review of the Sex Ratio: Alternatives for Comparative Research," *Journal of Marriage and the Family*, 53(4): 941-57.

Gordon, Milton M. (1964), *Assimilation in American Life,* New York, Oxford University.

Hamermesh, D., and Biddle, J. (1994), "Beauty and the Labor Market," *The American Economic Review*, 84(5): 1174-94.

Hwang, S., Saenz, R., and Aquirre, B. (1997), "Structural and assimilationist explanations of Asian American intermarriage," *Journal of Marriage and the family*, 59: 758-772.

Kalmijn, M.(1998), "Intermarriage and homogamy:Causes, patterns, trends, " *Annual Review*

of Sociology, 24: 395-421.

Kalmijn, M. and Van Tubergen, F.(2007), "Ethnic Intermarriage in the Netherlands: Confirmations and Refutations of Accepted Insights.", *European Journal of Population Studies*, 22(4): 371-97.

Kantarevic, J.(2004), "Interethnic marriages and economic assimilation of migrants.", *IZA Discussion Paper*, no.1142, May.

Kee, P.(1995), "Native-Immigrant Wage Differentials in the Netherlands: Discrimination?," *Oxford Economic Papers*, 47(2): 302-17.

Marsden, P.V. (1987), "Core Discussion Networks of Americans," *American Sociological Review*, 52: 122-31.

Meng, X.,and Gregory, Robert G.(2005), "Intermarriage and Economic Assimilation of Immigrants," *Journal of Labor Economics*, 23(1): 135-176.

Meng, X., and Meurs, D. (2009), "Intermarriage, language, and economic assimilation process: A case study of France," *International Journal of Manpower*, 30(2): 127-144.

Moore, G. (1990), "Structural Determinants of Men's and Women's Personal Networks," *American Sociological Review*, 55: 726-735.

Mroz, T.(1999), "Discrete factor approximations in simultaneous equation models:Estimating the impact of a dummy endogenous variable on a continuous outcome", *Journal of Econometrics*, 92: 233-74.

Muttarak,R.(2004), "Marital Assimilation: Interethnic Marriage in Britain," *Australian Population Association Conference*, 15-17 September 2004.

Mocan, N.H., and Tekin, E.(2002), "Nonprofit Sector and Part-Time Work: An Analysis of Employer-Employee Matched Data of Child Care Workers", *Review of Economics and Statistics*, 85(1): 38-50.

Nieseng, W., Van Praag, B.M.S., and Veenman, J.(1994), "The unemployment of ethnic minority groups in the Netherlands", *Journal of Econometrics*, 61(1): 173-96.

Nakosteen, R.A., and Zimmer, M.A. (1987), "Marital Status and Earnings of Young Men:

A model with Endogenous Selection", *The Journal of Human Resources*, 22(2): 248-268.

Picone, G.A., Sloan, F.A., Chou, S., and Taylor, D.H. (2003), "Does Higher Hospital Cost Imply Higher Quality of Care?", *Review of Economics and Statistics*, 85(1):51-62

Stratton, S.L. (2002)"Examining the Wage Differential for Married and Cohabiting Men,", *Economic Inquiry*, 40(2): 199-212.

Stock, J.H., Wright, J.H., and Yogo, M.(2002), "A Survey of Weak Instruments and weak Identification in Generalized Method of Moments", *Journal of Business and Economic Statistics*, 20(4):518-29.

Tekin, E.(2007), "Single Mothers working at Night: Standard Work and Child Care Subsidies", *Economic Inquiry*, 45(2): 233-50.

Qian, Z. and Lichter, D. (2001), "Measuring marital assimilation: intermarriage among natives and immigrants", *Social Science Research*, 30: 289-312.

Van Ours, J. C. and Veenman, J.(2003), "The educational attainment of second-generation immigrants in the Netherlands", *Journal of Population Economics*, 16: 739-53.

——— (2005), "The Netherlands; Old emigrants-Young immigrant country", *European Migration: What do we Know?* , pp.173-196, Oxford: Oxford University Press.

Van Ours, J.C.(2007), "The Effects of Cannabis Use on Wages of Prime-age Males", *Oxford Bulletin of Economics and Statistics*, 69(5): 619-34.

Van Tubergen, F. and Kanas, A.M.(2006), "The Impact of Human and Social Capital on I immigrants' Employment and Occupational Status", Unpublished paper.

Van Tubergen, F. and Maas, I.(2007), "Ethnic Intermarriage among Immigrants in the Netherlands: An Analysis of Population Data", *Social Science Research*, 36: 1065-86.

Worwick, C. (1999), "Credit constraints and the labour supply of immigrant families in Canada", *Canadian Journal of Economics*, 32(1): 152-70.

Wooldridge, J. M. (2002), *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

Table 1
Number of Ethnic Minorities in the Netherlands
in 1990, 2002 and 2007

	1990		2002		2007	
Natives(%)	12,668,000	(81.6)	13,140,336	(81.6)	13,187,586	(80.6)
Turks(%)	206,000	(4.2)	331,000	(2.1)	368,600	(2.3)
Moroccans(%)	168,000	(3.4)	284,000	(1.8)	329,493	(2)
Surinamese(%)	237,000	(4.8)	315,000	(2)	333,504	(2)
Antilleans(%)	81,000	(1.7)	125,000	(0.8)	129.965	(0.8)
Total Population(%)	14,892,574	(100)	16,105,285	(100)	16,357,992	(100)

Source: Statline, Statistics Netherlands. Natives refer to individuals who were born in the Netherlands with both parents also born in the Netherlands.

Table 2
Descriptive Statistics of the Marital Behavior
of the Four Ethnic Groups

	Endogamy		Exogamy	
Turks(%)	699	(45.7)	31	(17.1)
Moroccans(%)	593	(38.7)	30	(16.6)
Surinamese(%)	188	(12.3)	63	(34.8)
Antilleans(%)	51	(3.33)	57	(31.5)
Total(%)	1531	(100)	181	(100)

Source: SPVA surveys 1994, 1998, and 2002.

Table 3
Definition of Variables

Variable Name	Definition
Intermarriage	binary variable,=1 if the immigrant married to a native and zero if the immigrant married within his own ethnic group.
Log(hourly wage)	the logarithm of hourly wage.
Employment	binary variable,=1 if the immigrant is employed at the time of the survey.
Age	age of the immigrant at the time of the survey.
Primary	binary variable,=1 if the immigrant has a primary school diploma or no degree.
Lower secondary	binary variable,=1 if the immigrant is a lower secondary school graduate.
Higher secondary	binary variable,=1 if the immigrant is a higher secondary school graduate.
University	binary variable,=1 if the immigrant has a college degree or more.
YSM1	binary variable,=1 if duration of residence in the Netherlands is between 0-5 years.
YSM2	binary variable,=1 if duration of residence in the Netherlands is between 6-10 years.
YSM3	binary variable,=1 if duration of residence in the Netherlands is between 11-15 years.
YSM4	binary variable,=1 if duration of residence in the Netherlands is between 16-20 years.
YSM5	binary variable,=1 if duration of residence in the Netherlands is 21 years or more.
Second Generation	binary variable,=1 if the individual is a second-generation immigrant.
Caribbean	binary variable,=1 if the immigrant is from Suriname or the Antilles (the Caribbeans) and =0 if the immigrant is from Turkey or Morocco (the Mediterraneans).
Bigcity	binary variable,=1 if the immigrant lives in one of the four major cities in the Netherlands, i.e. Amsterdam, Rotterdam, The Hague, and Utrecht.
Survey1994	binary variable,=1 if the immigrant was surveyed in the 1994 wave.
Survey1998	binary variable,=1 if the immigrant was surveyed in the 1998 wave.
Survey2002	binary variable,=1 if the immigrant was surveyed in the 2002 wave.
Children	the number of children the immigrant has.
Dutch Proficiency	binary variable,=1 if the immigrant has no difficulty speaking Dutch.
Sex Ratio	the logarithm of the number of women for each man in the ethnic group aged 15-60.
Group Size	the logarithm of the number of female immigrants in the ethnic group relative to the total female population in the Netherlands aged 15-60.

Table 4
Descriptive Statistics

	Total Sample (N=1712)		**Endogamous Marriage (N=1531)**		**Exogamous Marriage (N=181)**	
	Mean	**Std. Dev.**	**Mean**	**Std.Dev.**	**Mean**	**Std. Dev.**
Log(hourly wage)	2.06	0.29	2.04	0.26	2.22	0.40
Employment	0.56	0.49	0.54	0.50	0.72	0.44
Age	38.4	9.06	38.2	8.99	40.3	9.47
Education Level						
Primary	0.52	0.50	0.56	0.49	0.22	0.41
Lower secondary	0.22	0.41	0.20	0.40	0.30	0.46
Higher secondary	0.17	0.38	0.17	0.37	0.22	0.42
University	0.08	0.28	0.06	0.24	0.25	0.43
Years since migration	19.7	9.32	19.2	8.72	24.5	12.4
YSM1 (0-5 years)	0.06	0.25	0.06	0.25	0.08	0.28
YSM2 (6-10 years)	0.12	0.33	0.13	0.33	0.08	0.28
YSM3 (11-15 years)	0.13	0.33	0.13	0.34	0.06	0.24
YSM4 (16-20years)	0.18	0.39	0.19	0.39	0.10	0.30
YSM5 ($>$ 20 years)	0.49	0.50	0.47	0.50	0.66	0.47
Second-generation	0.06	0.24	0.05	0.21	0.19	0.39
Caribbean	0.21	0.41	0.15	0.36	0.66	0.47
Bigcity	0.66	0.47	0.70	0.46	0.36	0.48
Survey-year Indicators						
Survey1994	0.24	0.43	0.24	0.42	0.28	0.45
Survey1998	0.47	0.50	0.47	0.49	0.42	0.50
Survey2002	0.28	0.45	0.28	0.45	0.29	0.46
Dutch Proficiency	0.42	0.49	0.38	0.48	0.74	0.44
Children	2.43	1.79	2.51	1.82	1.69	1.35

Table 5
Results from OLS Estimation of the Wage Equation

Log(hourly wage)	**1**	**2**	**3**
Intermarriage	0.179**	0.064*	0.058+
	(0.036)	(0.030)	(0.033)
Age		0.011	0.014
		(0.009)	(0.009)
Age^2/100		−0.011	−0.014
		(0.011)	(0.011)
Education Level			
Lower secondary		0.039**	0.038**
		(0.018)	(0.018)
Higher secondary		0.108**	0.107**
		(0.022)	(0.022)
University		0.323**	0.319**
		(0.030)	(0.030)
Years since migration			
YSM2 (6-10 years)		0.116**	0.114**
		(0.034)	(0.034)
YSM3 (11-15 years)		0.116**	0.112**
		(0.035)	(0.035)
YSM4 (16-20years)		0.212**	0.205**
		(0.033)	(0.033)
YSM5 (>20 years)		0.221**	0.205**
		(0.033)	(0.035)
Caribbean		0.062**	0.061**
		(0.020)	(0.020)
Survey2002		0.246**	0.244**
		(0.021)	(0.021)
Survey1998		0.097**	0.097**
		(0.018)	(0.018)
Bigcity		0.018	0.017
		(0.016)	(0.016)
Second Generation			0.048
			(0.040)
Intermarriage×Second Generation			-0.005
			(0.074)
R^2	0.045	0.374	0.376
Observations	971	971	971

Notes: The dependent variable is the logarithm of hourly wage. Robust standard errors are given in parentheses. **, * and + indicate that the estimated coefficients are statistically significant at the 1%, 5% and 10% levels respectively. Constants are not reported. The reference categories for education, years since migration, and survey year indicator variables are primary education, duration of residence between 0 and 5 years, and Survey1994 respectively.

Table 6
Parameter Estimates for the Discrete Factor Method

	Accounting for Endogeneity		**Accounting for Endogeneity & Selectivity**		
	Intermarriage	**Wage**	**Employment**	**Intermarriage**	**Wage**
Intermarriage	-	0.071**	-	-	0.072**
		(0.022)			(0.021)
Age	−0.027	0.017*	0.147**	−0.007	0.016+
	(0.071)	(0.008)	(0.035)	(0.047)	(0.010)
Age^2/100	0.047	−0.019+	−0.229**	0.023	−0.018+
	(0.087)	(0.011)	(0.043)	(0.057)	(0.011)
Education Level					
Lower secondary	0.300+	0.027	0.164+	0.217+	0.027
	(0.168)	(0.022)	(0.089)	(0.125)	(0.022)
Higher secondary	0.269	0.104**	0.381**	0.148	0.102**
	(0.192)	(0.020)	(0.098)	(0.144)	(0.020)
University	0.741**	0.316**	0.509**	0.650**	0.316**
	(0.183)	(0.024)	(0.142)	(0.148)	(0.024)
Years since migration					
YSM2 (6-10 years)	−0.686*	0.095**	0.387**	−0.572**	0.094*
	(0.286)	(0.043)	(0.158)	(0.196)	(0.043)
YSM3 (11-15 years)	−0.827**	0.101*	0.438**	−0.738**	0.086*
	(0.293)	(0.043)	(0.158)	(0.216)	(0.043)
YSM4(16-20years)	−0.411	0.194**	0.538**	−0.660**	0.181**
	(0.272)	(0.043)	(0.155)	(0.200)	(0.042)
YSM5 (>20 years)	−0.096	0.206**	0.593**	−0.249	0.197**
	(0.247)	(0.041)	(0.158)	(0.177)	(0.040)
Caribbean	0.267	0.053**	0.701**	0.555**	0.050**
	(0.273)	(0.018)	(0.096)	(0.144)	(0.018)
Survey2002	-	0.242**	0.511**	-	0.242**
		(0.024)	(0.096)		(0.024)
Survey1998	-	0.093**	0.347**	-	0.093**
		(0.023)	(0.083)		(0.023)
Bigcity	-	0.007	-0.026	-	0.018
		(0.016)	(0.071)		(0.016)
Sex ratio	−1.956**	-	-	−0.914**	-
	(0.597)			(0.169)	
Group size	−0.671**	-	-	−0.309**	-
	(0.148)			(0.096)	
Children	-	-	−0.055**	-	-
			(0.022)		
Observations	971	971	1712	1712	971
Log likelihood	-221.25		-1385.52		

Notes: The estimates are based on the discrete factor method with three points of support. Standard errors are given in parentheses. **, * and + indicate respectively 1%, 5% and 10% significance levels. Constants are not reported. The reference categories for education, years since migration, and survey year indicator variables are primary education, duration of residence between 0 and 5 years, and Survey1994 respectively.

Table 7
DFM Parameters

	Accounting for Endogeneity	
σ	0.227^{+}	
	(0.119)	
Factor Loadings		
α_1	0.057^{*}	
	(0.029)	
α_2	0.106^{*}	
	(0.052)	
	Mass Points	**Probability Weight**
1st support	0	0.508
2nd support	0.472	0.040
3rd support	1	0.450
	Accounting for Endogeneity and Sample Selection	
σ	0.217^{+}	
	(0.113)	
Factor Loadings		
α_1	0.054^{*}	
	(0.026)	
α_2	0.080^{*}	
	(0.038)	
α_3	0.006	
	(0.010)	
	Mass Points	**Probability Weight**
1st support	0	0.468
2nd support	0.479	0.002
3rd support	1	0.528

Table 8
Parameter Estimates for the Discrete Factor Method

	Employment	**Intermarriage**	**Wage**
Intermarriage	-	-	0.067**
			(0.024)
Age	0.146**	-0.019	0.019*
	(0.035)	(0.046)	(0.009)
Age^2/100	−0.226**	0.038	−0.020$^+$
	(0.043)	(0.057)	(0.011)
Education Level			
Lower secondary	0.184*	0.288*	0.042$^+$
	(0.089)	(0.125)	(0.022)
Higher secondary	0.376**	0.239	0.106**
	(0.098)	(0.145)	(0.020)
University	0.507**	0.720**	0.320**
	(0.141)	(0.148)	(0.024)
Years since migration			
YSM2 (6-10 years)	0.332*	−0.536**	0.100*
	(0.158)	(0.194)	(0.043)
YSM3 (11-15 years)	0.335*	−0.682**	0.100*
	(0.158)	(0.216)	(0.043)
YSM4(16-20years)	0.466**	−0.623**	0.188**
	(0.156)	(0.199)	(0.043)
YSM5 (>20 years)	0.479**	−0.288	0.190**
	(0.159)	(0.175)	(0.042)
Caribbean	0.680**	0.602**	0.058**
	0.097	(0.144)	(0.018)
Survey2002	0.545**	-	0.246**
	(0.096)		(0.024)
Survey1998	0.376**	-	0.099**
	(0.083)		(0.023)
Bigcity	-0.077	-	0.024
	(0.0719)		(0.016)
Second Generation	-	-	0.050
			(0.040)
Intermarriage×Second Generation	-	-	-0.022
			(0.051)
Sex ratio	-	−0.887**	-
		(0.169)	
Group size	-	−0.348**	-
		(0.097)	
Children	−0.047*	-	-
	(0.022)		
Observations	1712	1712	971
Log likelihood			-1382.34
Intermarriage+Intermarriage×Second Generation=0			0.045
			(0.047)

Notes: The estimates are based on the discrete factor method with three points of support. Standard errors are given in parentheses. **, * and $^+$ indicate respectively 1%, 5% and 10% significance levels. Constants are not reported. The reference categories for education, years since migration, and survey year indicator variables are primary education, duration of residence between 0 and 5 years, and Survey1994 respectively.

Table 9
Results from 2SLS Estimation

	Intermarriage	**Wage**
Intermarriage	-	0.187^{+}
		(0.108)
Age	-0.020^{+}	0.013
	(0.011)	(0.009)
Age^2/100	0.028^{*}	-0.014
	(0.014)	(0.011)
Education Level		
Lower secondary	0.007	0.038^{*}
	(0.025)	(0.019)
Higher secondary	0.013	0.106^{**}
	(0.025)	(0.022)
University	0.139^{**}	0.303^{**}
	(0.032)	(0.034)
Years since migration		
YSM2 (6-10 years)	-0.070	0.125^{**}
	(0.047)	(0.036)
YSM3 (11-15 years)	-0.084^{+}	0.126^{**}
	(0.048)	(0.037)
YSM4(16-20years)	-0.042	0.216^{**}
	(0.046)	(0.034)
YSM5 (>20 years)	0.034	0.218^{**}
	(0.045)	(0.035)
Caribbean	0.141^{**}	0.031
	(0.027)	(0.035)
Survey2002	-0.097^{**}	0.257^{**}
	(0.029)	(0.023)
Survey1998	-0.092^{**}	0.108^{**}
	(0.025)	(0.021)
Bigcity	-0.138^{**}	0.036
	(0.019)	(0.022)
Instruments		
Sex ratio	-0.241^{**}	-
	(0.033)	
Group size	-0.139^{**}	-
	(0.022)	
Observations	971	971
F-statistic of excluded instruments	26.19	
(Prob$>$ F)	(0.000)	
Hansen J statistic	0.009	
(Prob$>$ $\chi^2(1)$)	(0.926)	

Notes: Robust standard errors are given in parentheses. **, * and $^{+}$ indicate respectively 1%, 5% and 10% significance levels. The reference categories for education, years since migration, and survey year indicator variables are primary education, duration of residence between 0 and 5 years, and Survey1994 respectively.

Table 10

The Three-Equation Model: Accounting for Endogeneity and Sample Selection

	Employment	**Intermarriage**	**Wage**
Intermarriage	-	-	0.230*
			(0.121)
Mills ratio (λ)	-	0.469*	0.074
		(0.229)	(0.131)
Age	0.176**	0.038	0.021
	(0.035)	(0.026)	(0.016)
Age^2/100	−0.266**	−0.054	-0.026
	(0.043)	(0.038)	(0.023)
Education Level			
Lower secondary	0.147$^+$	0.045	0.046*
	(0.088)	(0.033)	(0.023)
Higher secondary	0.365**	0.093$^+$	0.121**
	(0.096)	(0.052)	(0.035)
University	0.456**	0.22**	0.316**
	(0.136)	(0.061)	(0.040)
Years since migration			
YSM2 (6-10 years)	0.394**	0.052	0.143**
	(0.154)	(0.070)	(0.050)
YSM3 (11-15 years)	0.410**	0.046	0.145**
	(0.155)	(0.071)	(0.050)
YSM4(16-20years)	0.501**	0.115	0.236**
	(0.151)	(0.079)	(0.051)
YSM5 (>20 years)	0.542**	0.20**	0.236**
	(0.153)	(0.082)	(0.049)
Caribbean	0.657**	0.260**	0.046
	(0.111)	(0.070)	(0.045)
Survey2002	0.548**	0.050	0.283**
	(0.099)	(0.073)	(0.052)
Survey1998	0.366**	0.010	0.127**
	(0.084)	(0.053)	(0.040)
Bigcity	-0.085	−0.157**	0.039$^+$
	(0.070)	(0.021)	(0.022)
Sex ratio	-0.108	−0.290**	-
	(0.106)	(0.041)	
Group size	-0.135	−0.194**	-
	(0.088)	(0.029)	
Children	−0.075**	−0.047**	-
	(0.024)	(0.012)	
Observations	1712	971	971
F-statistic of excluded instruments		16.36	
(Prob$>$ F)		(0.000)	
Hansen J statistic		0.027	
(Prob$>\chi^2(2)$)		(0.986)	

Notes: Robust standard errors are given in parentheses. **, * and $^+$ indicate that the estimated coefficients are statistically significant at the 1%, 5% and 10% levels respectively. The reference categories for education, years since migration, and survey year indicator variables are primary education, duration of residence between 0 and 5 years, and Survey1994 respectively..

Table 11
Wage Equation Parameter Estimates

	OLS/2SLS/TEM			Discrete Factor Method	
	OLS	**2SLS**	**TEM**	**Endogeneity**	**Endogeneity &Selectivity**
Intermarriage	0.064*	0.187+	0.230*	0.071**	0.072**
	(0.030)	(0.108)	(0.121)	(0.022)	(0.021)
Age	0.011	0.013	0.021	0.017*	0.016+
	(0.009)	(0.009)	(0.016)	(0.008)	(0.010)
Age^2/100	−0.011	-0.014	-0.026	−0.019+	−0.018+
	(0.011)	(0.011)	(0.023)	(0.011)	(0.011)
Education Level					
Lower secondary	0.039**	0.038*	0.046*	0.027	0.027
	(0.018)	(0.019)	(0.023)	(0.022)	(0.022)
Higher secondary	0.108**	0.106**	0.121**	0.104**	0.102**
	(0.022)	(0.022)	(0.035)	(0.020)	(0.020)
University	0.323**	0.303**	0.316**	0.316**	0.316**
	(0.030)	(0.034)	(0.040)	(0.024)	(0.024)
Years since migration					
YSM2 (6-10 years)	0.116**	0.125**	0.143**	0.095**	0.094*
	(0.034)	(0.036)	(0.050)	(0.043)	(0.043)
YSM3 (11-15 years)	0.116**	0.126**	0.145**	0.101*	0.086*
	(0.035)	(0.037)	(0.050)	(0.043)	(0.043)
YSM4(16-20years)	0.212**	0.216**	0.236**	0.194**	0.181**
	(0.033)	(0.034)	(0.051)	(0.043)	(0.042)
YSM5 (>20 years)	0.221**	0.218**	0.236**	0.206**	0.197**
	(0.033)	(0.035)	(0.049)	(0.041)	(0.040)
Caribbean	0.062**	0.031	0.046	0.053**	0.050**
	(0.020)	(0.035)	(0.045)	(0.018)	(0.018)
Survey2002	0.246**	0.257**	0.283**	0.242**	0.242**
	(0.021)	(0.023)	(0.052)	(0.024)	(0.024)
Survey1998	0.097**	0.108**	0.127**	0.093**	0.093**
	(0.018)	(0.021)	(0.040)	(0.023)	(0.023)
Bigcity	0.018	0.036	0.039+	0.007	0.018
	(0.016)	(0.022)	(0.022)	(0.016)	(0.016)
Observations	971	971	971	971	971

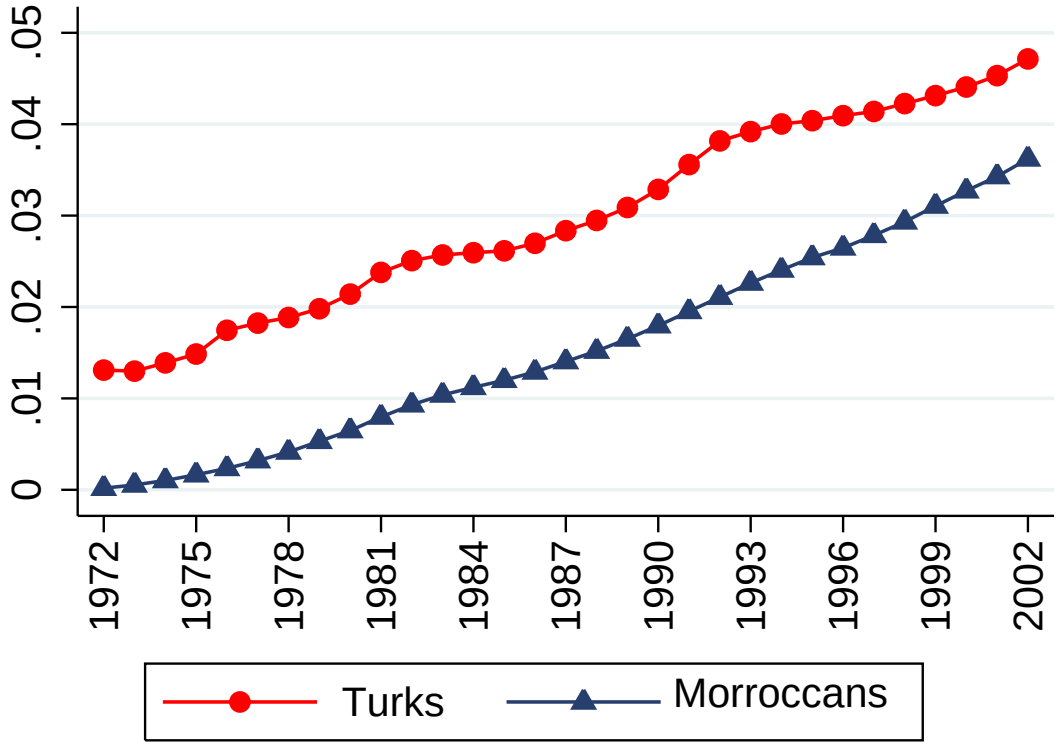
Notes: Columns 1, 2, 3, 4, and 5 reproduce the wage equation estimates from Tables 5, 9, 8, 6, and 6 respectively. The dependent variable is the logarithm of hourly wage. Standard errors are given in parentheses. **, * and + indicate that the estimated coefficients are statistically significant at the 1%, 5% and 10% levels respectively. The reference categories for education, years since migration, and survey year indicator variables are primary education, duration of residence between 0 and 5 years, and Survey1994 respectively.

Table 12
Results from the Wage Equation (14)

	Log(hourly wage)
Age	0.022*
	(0.009)
Age^2/100	−0.029*
	(0.012)
Education Level	
Lower secondary	0.033+
	(0.018)
Higher secondary	0.106**
	(0.022)
University	0.310**
	(0.031)
Caribbean	0.067**
	(0.020)
Survey2002	0.229**
	(0.030)
Survey1998	0.107**
	(0.021)
Bigcity	0.016
	(0.016)
Cohort arrives 1973-1980	-0.011
	(0.036)
Cohort arrives 1981-1990	-0.020
	(0.054)
Cohort arrives 1991-2002	-0.081
	(0.073)
Cohort arrives 1973-1980×Intermarriage	0.001
	(0.062)
Cohort arrives 1981-1990×Intermarriage	-0.056
	(0.067)
Cohort arrives 1991-2002×Intermarriage	-0.052
	(0.068)
Years since migration	0.006+
	(0.003)
Years since migration×Intermarriage	0.002+
	(0.001)
R^2	0.364
Observations	889

Notes: The dependent variable is the logarithm of hourly wage. Robust standard errors are given in parentheses. **, * and + indicate that the estimated coefficients are statistically significant at the 1%, 5% and 10% levels respectively. Constants are not reported. The reference categories for education, years since migration, and survey year indicator variables are primary education, duration of residence between 0 and 5 years, and Survey1994 respectively.

Figure 1
Group Size



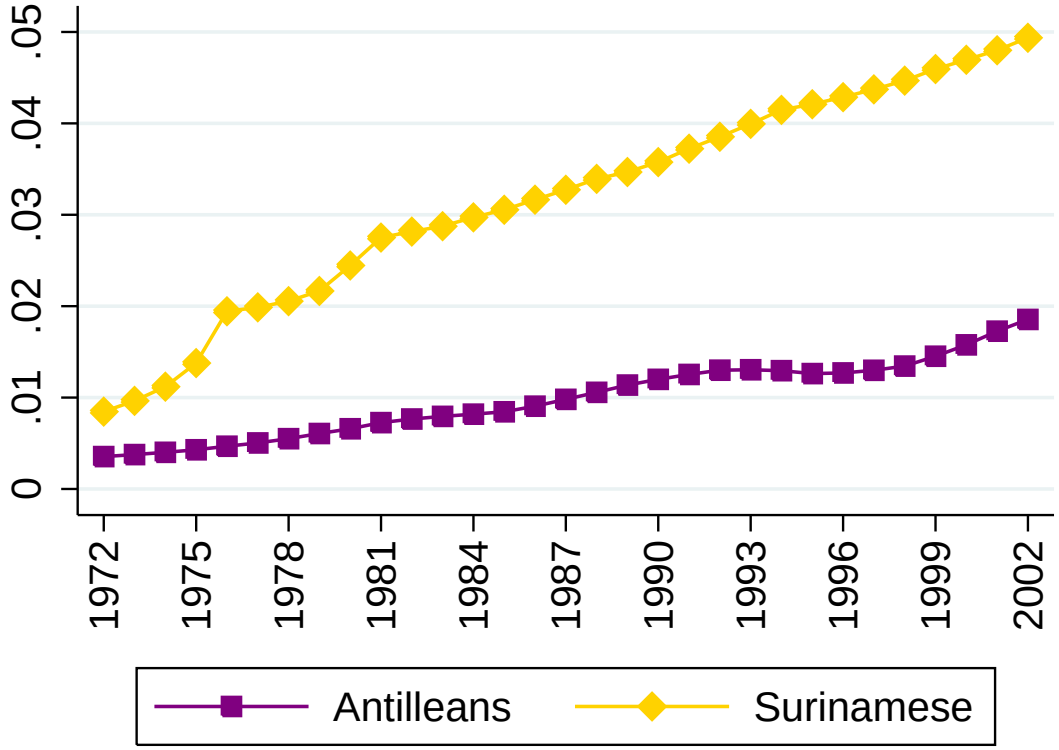


Figure 2
Sex Ratio