Intermarriage Pattern Among Swedish Immigrants In the US in 1900 and 1910

A Logit Analysis of Different Meanings of Factors Influencing Intermarriage Between Gender

by

Linda Schalling

Lschalling@hotmail.com

Abstract: Intermarriage patterns are often referred to as a step in the social assimilation. However, previous research on is ambiguous in its result concerning differences in intermarriage pattern between genders. According to the theory intermarriage patterns are both influences by preferences and opportunities. This thesis conducts a conditional fixed effects logistic regression to analyses the intermarriage patter among Swedish immigrants in the early twentieth century in the United States and provide insights in how the meaning of the local marriage market and English proficiency differ between men and women. The findings align with previous research on factors associates with intermarriage. Furthermore, the findings suggests that women's odds of intermarriage increases more when the local marriage market reduces than for men. Women also appears to have lower odds of intermarriage when they lack English proficiency than men. These results highlight the need for further research on other men and women's differences in factors affecting intermarriage to better understand historical and contemporary intermarriage patterns.

Acknowledgments

First, I would like to thank my supervisor for valuable insights, constructive feedback, and for reminding me to breath throughout the thesis writing process. Second, I want to thank all the friends I have gained during this program for going through this stressful period together, but ensuring a good mood and a lot of fika breaks. Last but not least, I would like to thank Jesper for putting up with and encouraging me through all the stress over the last couple of years.

Table of Contents

List of Tables

List of Figures

1 Introduction

In the 1800s, significant waves of European immigrants began moving to North America, continuing through the 1920s. During that time, migrating to North America was relatively easy, with few restrictions. The significant waves of migration during this period brought about substantial changes in society, impacting many people's lives and prompting extensive research into the economic and social integration of immigrants into their new communities (Bean et al., 2000; Gonsoulin & Fu, 2010; Kalmijn, 1998). Sweden was one of the largest exit countries in Europe, losing more than a fifth of their population to migration, and a large share of this population migrated to the United States. The Swedish economic condition incentivized them as a push factor and a pull from American opportunities. This has generated extensive research on the consequences and aftermath of mass migration, commonly known as the era of mass migration (e.g., Carlsson, 1976; Dribe, Eriksson & Helgertz, 2023).

Studying intermarriage patterns became crucial to understand how immigrants assimilated, as intermarriage between different groups is widely recognized as a key indicator of social assimilation (Qian & Lichter, 2007). Marriage patterns are influenced by both preference and opportunity. Individuals tend to prefer spouses with cultural similarities, often including shared values, opinions, and experiences. Nevertheless, individuals with cultural similarities often participate in the same activities and community, increasing exposure to these individuals (DiMaggio & Mohr, 1985; Kalmijn, 1998). Furthermore, English proficiency is often considered a barrier to intermarriage, and low English proficiency tends to decrease the likelihood of intermarriage. The opportunity to marry within a group depends on factors such as residential segregation, local marriage markets, and group size (Kalmijn, 1991, 1998). Individuals who strongly identify with a group are more likely to internalize the norms of endogamy, resulting in a greater likelihood of marrying within the group. This effect is particularly pronounced among racial and ethnic groups, where the internalization of endogamous norms is believed to be strong (Merton, 1941). Intermarriage can also be due to an unbalanced sex ratio within the ethnic group in the area. Thus, individuals might be forced to intermarry (Cretser & Leon, 1982).

Even though several studies look at men's and women's intermarriage patterns separately, it is difficult to statistically confirm differences between men's and women's intermarriage patterns. Some studies find differences between men's and women's intermarriage patterns (e.g., McCaa, 1993), and some see no difference between men and women (e.g., Pagnini & Morgan, 1990). Including a sex variable show a general distinction between male and female, but Spörlein, Schlueter, and van Tubergen (2014) introduce that interacting sex with other variables, such as sex ratio, can give additional information on gender differences. Thus, despite this research, the knowledge of gender differences is sparse, especially focusing on Sweden. Therefore, this study explores the intermarriage patterns of Swedish immigrants in the United States. Specifically, it will investigate the likelihood of intermarriage among Swedish immigrants in the United States at the beginning of the 1900s and explore different meanings of certain factors between genders. Hence, investigate the following research questions:

1. *To what extent is the local marriage market associated with the likelihood of intermarriage, and how does this vary between male and female first-generation Swedish immigrants aged 20-45 who got married after migrating to the United States?*

2. *How is English proficiency associated with the likelihood of intermarriage, and does this vary between male and female first-generation Swedish immigrants 20-45 who married after migrating to the United States?*

Focusing exclusively on Sweden is to gain specific insights into Swedish immigrants' marriage patterns. While intermarriage among Scandinavian immigrants has been studied commonly in some research (e.g., Peach, 1980), there is a lack of studies focusing exclusively on intermarriage among Swedish immigrants. Furthermore, some of the studies of Scandinavia as a group have also focused on limited areas, such as New York City (McCaa, 1993; Pagnini & Morgan, 1990) and New Haven (Peach, 1980a). Grouping Scandinavian countries fails to capture the unique intermarriage experiences of Swedish immigrants. Some studies exclusively on Sweden have examined intermarriage during other periods. For example, Beijbom (Beijbom, 1971) studied intermarriage developments in Chicago from 1846 to 1880, while Lindmark (1971) examined intermarriage in Illinois and Minnesota from 1914 to 1932. More recently, Dribe, Hacker, and Scalone (2018) studied Sweden separated from other Scandinavian countries and looked at immigrants all over the United States.

To answer the research questions, a quantitative approach is taken. A conditional fixed effects logistic regression model is used to analyze factors influencing the likelihood of intermarriage among Swedish immigrants in the United States. Thus, this study utilizes IPUMS 1900 and 1910 censuses in the United States to examine intermarriage patterns among Swedish immigrants. It will additionally include interactions between the sex and the marriage market[1] and between sex and English proficiency. This study aims to provide further insights into the factors that influenced intermarriage patterns among Swedish immigrants during the early 1900s in the United States, particularly concerning the increase in knowledge of gender differences in two factors commonly known to influence the likelihood of intermarriage. This will expand the literature on intermarriage among Swedish immigrants and highlight the importance of understanding the meaning of different factors between genders.

The thesis continues with a background of Swedish immigration to the United States to understand the drivers behind immigration and thus provide an understanding of the Swedish social assimilation process. Section three presents previous research on intermarriage in the United States, discusses contextual and individual factors related to intermarriage, and further explores the previous literature on gender differences in intermarriage patterns. Section four embeds the previous research with theory to provide better understanding and a theoretical framework for the analysis of factors associated with intermarriage and differences in the meaning of these factors. This is followed by a presentation the explaining the data set used and methodology. The empirical results are then presented in section six, showing that there are differences between men and women in the meaning of certain factors and their effect on intermarriage. The results discussed and contextualized in the discussion. Finally, section seven concludes the main findings and presents prospects for future research.

---

[1] Marriage market is the availability ratio of potential Swedish spouses of the opposite sex in the area. For further explanation, see section 4.2.

2   Background

Before presenting the background of Swedish immigration and intermarriage patterns in the United States, it is essential to define the term ethnical intermarriage. Ethnical technically refers to an individual's cultural belonging, thus one's heritage, whereas nationality is where that person has their national origin. In this study, ethnical and national origin are used interchangeably, and ethnical intermarriage refers to marriage between a Swedish-born individual and a non-Swedish-born individual, as USA IPUMS censused in 1900 and 1910 lacked variables to distinguish different ethnicities sufficiently. This section provides an outline for the Swedish migration to the United States to understand the background of the population being studied and the background to intermarriage in the United States during the shift to the twentieth century.

Social and economic turmoil characterized nineteenth-century Sweden, with farmers displaced from their land by capitalist agriculture and corporate forestry growth. In addition, unfavorable weather conditions led to poor harvests, resulting in significant food shortages and famine. Faced with starvation, many Swedes considered emigrating, leading to a period of large-scale migration from Sweden starting in the 1860s (Homer, 1964; Karadja & Prawitz, 2019). The open borders of the United States at the time enabled this mass migration, resulting in significant migration flows, mainly from Europe. Between 1850 and 1930, commonly known as the Era of Mass Migration saw, approximately 1.1 million Swedes, slightly over a fifth of the population, leaving their country for the United States. The Era of Mass Migration has spurred numerous studies investigating the reasons for migration, return migration, assimilation, and the impact on sending and receiving countries (e.g., Abramitzky, Boustan & Eriksson, 2012; Carlsson, 1976; Dribe, Eriksson & Helgertz, 2023).

Family migration characterized emigration from Sweden in the early years of migration in the mid-nineteenth century. With periods of famine and poverty, many took their whole family over the Atlantic to improve their living standards, attracted by the land available in the United States. Poor conditions pushed the migration during the 1960s and 70s for farmers, and emigrants mainly came from the countryside. Eurenius (2020) highlights the impact of a

family's emigration history, as individuals with relatives who had previously moved to the United States were more likely to emigrate themselves. The presence of friends or family who had already settled in the United States made it easier to migrate, and the ability to send money back home also played a role (Bohlin & Eurenius, 2010). As a result of family and friends aiding one another in the migration process, Swedish immigrants tended to concentrate in specific regions (see Figure 1).

Figure 1: Illustrates the distribution of Swedish immigrants based on the sample used in this study. Source: Ruggles et al. (2023), illustrations by the author.
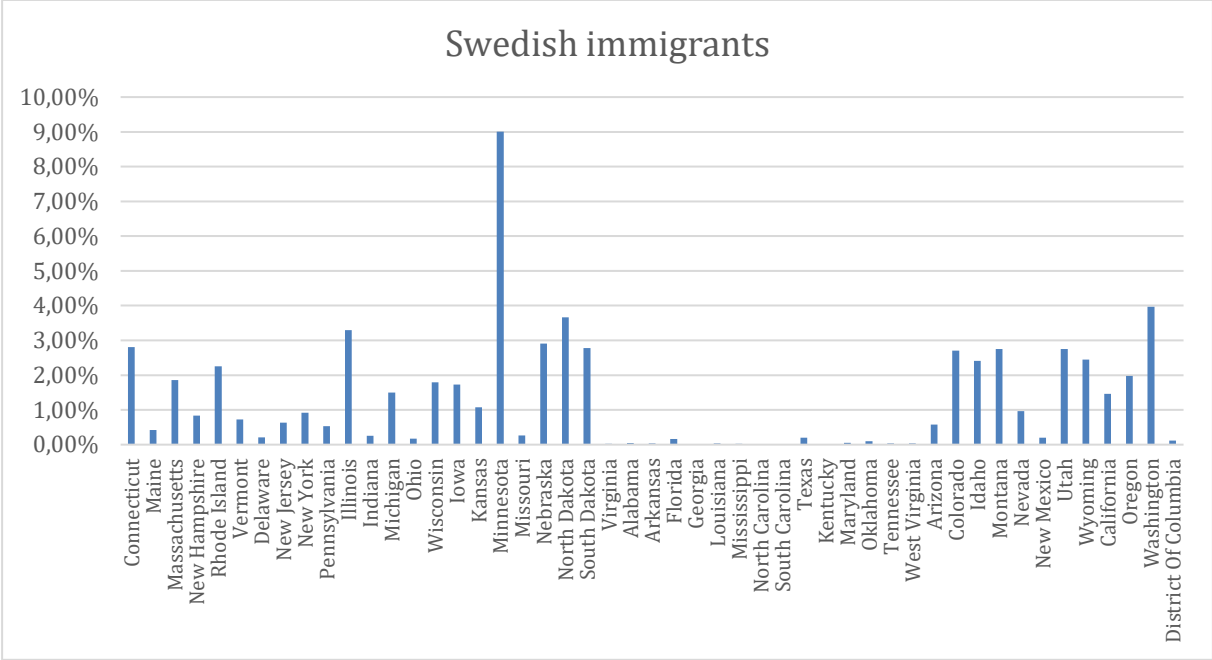
Beijbom (2010, p.65) points out that the social conditions in Sweden at this time have had great importance and impact and induced a gradual change towards an increasing share of young men and women migrating to improve their opportunities. It has also been discussed that some men emigrated to avoid military service (Carlsson, 1976). A large share of this was dissatisfaction with the Swedish voting rights system, which excluded more than four-fifth of the population who had the right to vote, especially among women who did not have voting right until 1921 when universal suffrage was implemented. Thus, many emigrants came from groups lacking civil rights. Furthermore, many migrated due to religious persecution and to avoid military service (Beijbom, 2010, pp.74–75). Thus, many Swedes left their country to escape restrictive social norms and expectations and seek greater freedom and opportunity. The United States offered greater social and economic development during this era, which aligns with Bohlin and Eurenius's (2010) findings that increasing living standards in Sweden

led to a decline in emigration in the long run. The rise of animal husbandry created more jobs in the agricultural sector, while internal labor immigration to the growing industrial sector reduced competition for rural employment in Sweden. As a result, the incentives for emigration decreased with the country's development and increasing opportunities.

A letter from a person who had emigrated to the United States says in a letter the following:

> *" A large percentage of those who emigrated at a mature age would if the motherland could give them the same benefits as America, return. But as it is now: an intolerant, in every way privileged state church, the heavy conscription, the high taxes, the class distinction unworthy of our enlightened time, the hopelessness, no matter how one strives, of ever being able to acquire land of one's own — a home of one's own. This and other things mean that the Swedes here do not have the slightest thought of returning to their homeland. We have had enough of slavery. As long as the current state of society is dominant in Sweden, the compatriots will continue to emigrate."* (Emigrationsutredningen VII, 1908, p.170, authors translation).

Other letters state similarly that due to the circumstances in Sweden with problems of having a home or making money, they would rather stay in the United States. Planning on permanently staying acquire efforts to assimilate with the new country. However, this does not entail that all were lucky enough to do well in the United States; some even state that they had it worse in the United States. Beijbom (2006) has extensively studied the experience of emigrating women from Sweden. The condition for Swedish women was incredibly restrictive, with no civil rights or access to education unless she could get it privately. They also had limited livelihood opportunities as many occupations were in the male domain. Still, in nineteenth-century Sweden, women's role was mainly to raise children and home care. Industrialization and emigration led to increasing numbers of young women moving to urban areas to work outside the home, such as in the textile industry. However, the conditions were poor. For example, even though they worked ten to twelve-hour shifts, they could barely make it on the pay. Beijbom (2006) further writes that gender equality came sooner in the emigration movement than in Swedish society, and it was attractive for many young women to emigrate. However, it was not all glamorous after emigrating. A difficulty immigrants faced when migrating was the English language. It became a compulsory subject in the Swedish school in 1962, meaning few immigrants knew the language when arriving in the

United States. Some American guides and basic knowledge was given out in articles and newspaper when immigration was widespread (Barr, 2015).

Based on letters of Swedish immigrants from the second half of the nineteenth-century state, they mainly migrated from Sweden due to push factors in Sweden; they were driven out of the country due to poverty and despair (Emigrationsutredningen VII, 1908). Thus, as many felt they could not return to Sweden, they had to assimilate into the new culture and thus found partners, got married and started a family in the United States. Several letters from the United States and experiences brought up in Emigrationsutredningen show that several Swedish immigrants in the United States either got married or had a relative who got married in the United States, both with other Swedes and spouses of different nationalities (Emigrationsutredningen VII, 1908).

3 Previous Research

Ethnic intermarriage is widely acknowledged as a significant indicator of social assimilation (Bean et al., 2000; Gonsoulin & Fu, 2010; McCaa, 1993). According to Kennedy (1944), intermarriage can serve as both a measure of assimilation level in a society and a mechanism for encouraging assimilation. However, Marcson (1950) challenges the idea that intermarriage is an accurate measure of assimilation and argues against this assumption. He suggests that if intermarriage were genuinely linked to assimilation, intermarriage rates would be higher among the oldest ethnic groups and lower among the most recently arrived groups. Instead, Marcson proposes that factors such as higher education, middle-class status, professional occupations, second and third-generation status, and urban nonfarm residence facilitate intermarriage. Several studies address how intermarriage relates to social assimilation and its consequences. For example, Rodríguez-García (2015) explores how intermarriage impacts social integration in various countries worldwide. The author contends that intermarriage fosters integration by creating social ties between ethnic and racial groups. While intermarriage is linked to increased social integration in many countries, its impact varies depending on the context and can result in cultural assimilation and loss of cultural identity. Research on intermarriage can adopt either a macro-level perspective, examining the societal impacts of intermarriage, or a micro-level viewpoint, exploring the consequences of intermarriage on social integration. Furthermore, such research can also, as this study will, focuses on investigating the factors that influence intermarriage. This section presents the relevant previous literature on marriage patterns in the context of intermarriage.

3.1 Contextual Factors

Rodríguez-García (2015) points out that marriage patterns depend on different contexts. According to Peach (1980a), higher levels of residential segregation between ethnic groups were associated with lower intermarriage rates studying European immigrants in the United States during the twentieth century. Conversely, lower levels of residential segregation were

associated with higher intermarriage rates. The study provided insights into urban societies' complex interplay between residential segregation, intermarriage, and socio-economic factors. It is theoretically assumed that urban areas should operate as a breeding ground for an ethnic melting pot as large urban areas was an area that encourages intermarriage (Norman, 1976). However, Beijbom (1971) found while studying Swedish immigrants in Chicago in the second half of the 1800s that few Swedes married outside their ethnical group.

The size and composition of the group surrounding an individual are frequently identified as significant contextual factors that play a crucial role in intermarriage patterns. McCaa (1993) studied the marital conditions of all immigrants in New York City using the PUMS 1900, 1910, 1960, and 1980. structural factors, such as the size and composition of the ethnic groups in the city, influence the likelihood of intermarriage between people of different ethnicities in New York City. According to Spörlein, Schlueter, and van Tubergen (2014), who explore the factors contributing to ethnic intermarriage from 1880 to 2011, intermarriage is more likely to increase over time when the relative size of an immigrant group decreases; sex ratios become more imbalanced. Cultural factors also play a role, as an origin group's exogamous behavior in the past exerts long-term effects, and exogamous practices increase over time when the prevalence of early marriage customs declines. The study finds that some determinants of intermarriage have different longitudinal and cross-sectional effects. Several studies support these results that the marriage market of potential spouses from one's own ethnic group decreases the intermarriage rates in that area. Dribe, Hacker, and Scalone (2018) found that during the early 1900s, the size of immigrant groups within a community appeared to have important implications for intermarriage. The greater the number of individuals with the same nationality in a local area, the greater the opportunity to marry, especially endogamously. The diversity of a community was related to the tendency toward exogamy. Overall, the study highlights the complex factors influencing intermarriage patterns among immigrants in the early 20th century and their integration into host societies. Similarly, Chiswick and Houseworth (2011) and Choi and Tienda (2017) found that the availability of potential spouses from one's own ethnic group and the size of the group in the area affect intermarriage rates at the end of the 1900s. Both Dribe, Hacker, and Scalone (2018) found in 1910 and Choi and Tienda (2017) in the late 1900s found that living in an urban area increases the odds of intermarriage and the geographical area is significantly associated with the odds of intermarriage. Furthermore, Peach (1980a) suggests that structural influences play a role in

determining level parameters. Specifically, the distribution of ethnic groups across regions in the United States and residential segregation within areas impact marriage tendencies.

Thus, findings on the group size, sex ratio and availability of potential partners within the same ethnical group in intermarriage highlight the contextual factors for understanding intermarriage patterns. No studies have been found to contradict the presented studies' results regarding the availability of potential spouses. As well as the influence of residence (urban/rural) and regional variables. Regarding urban/rural residence, the results are mixed, where Beijbom (1971) found in the second half of the 1800s, even though living in Chicago, few Swedes married endogamously. However, exogamy was more uncommon during the 1800s. Dribe, Hacker, and Scalone (2018), studying 1910, found that in smaller urban areas, the odds of intermarriage are higher than in rural areas for women. Still, there is no significant difference between higher-populated urban and rural areas. However, there does not appear to be a significant difference in the odds of intermarriage between urban and rural men.

3.2   Individual Factors

In addition to the contextual factors, an individual's factors affect the choice of spouse. Studying individuals' ethnicities has shown that, in general, endogamy is more common among new immigrants. Furthermore, it appears that in the case of intermarriage, it was more common within old immigrant groups and within new immigrant groups (Pagnini & Morgan, 1990). In their study of intermarriage patterns in 1979, Alba and Golden (1986) found that ethnic background influences assortative mating, although to a lesser degree than in previous cohorts. Studies such as Pagnini and Morgan (1990) examined intermarriage patterns during the turn of the twentieth century and emphasized the significance of ethnic background in assortative mating.

McCaa (1993) found that individual characteristics, such as education and income, affected the choice of intermarriage. Similarly, Choi and Tienda (2017) found that highly educated are more likely to intermarry. Peach (1980a) also highlighted the impact of socioeconomic factors on intermarriage rates, with individuals from higher socio-economic backgrounds more likely to intermarry than those from lower socio-economic backgrounds. Dribe, Hacker, and Scalone (2023) examine intermarriage patterns among United States immigrants in the early 20th

century using 1910 census data. The characteristics they found to increase the probability of intermarriage were English-speaking ability, literacy, and elevated socioeconomic status. Examining intermarriage patterns among United States immigrants from 1880 to 1990, Wildsmith, Gutmann, and Gratton (2003) highlight the significant assimilation and intermarriage levels observed among Swedish immigrants attributed to their relatively high education and occupational status levels.

Chiswick and Houseworth (2011) analyze interethnic marriages among immigrants in the United States using data from the 1980 United States Census. The study identifies several factors that influence intermarriage. The results show that the probability of intermarriage increases with more extended residence in the United States and younger age at arrival. Furthermore, they study the effect of having the ability to speak English. They use the distance from an immigrant's mother tongue and English as a proxy for English skills for the immigrant. Similarly to Dribe, Hacker, and Scalone (2023), Chiswick and Houseworth (2011) found that lower English proficiency reduces ethnic intermarriages. In addition, Beijbom (1971) found that workers were less likely to intermarry compared to businessmen, office and shop employees, who had a higher tendency to intermarry, especially if the work required English proficiency.

Age and age at marriage are common individual control variables when looking at intermarriage. Spörlein, Schlueter, and van Tubergen (2014) find that an individual's age in the study decreases the odds of intermarriage. Choi and Tienda (2017) looked at different age groups' impact on intermarriage during the early 2000s. They found that individuals in the study between 25 and 34 were more likely to intermarry than those under 25. In contrast, individuals over the age of 34 were less likely to intermarry than those under 25, indicating a non-linear relationship between age and intermarriage. Other studies, such as Dribe, Hacker, and Scalone (2018), use age at marriage instead to study whether the age when getting married impacts the odds of intermarrying. On the other hand, studying age looks at how the odds of intermarriage change between cohorts.

There is evidence that race is an influential factor in intermarriage. Some studies look at the individual as non-white (e.g., Spörlein, Schlueter & van Tubergen, 2014), while others divide individuals into several categories based on origin (e.g., Jacobs & Labov, 2002). Instead of studying intermarriage between individuals of different nationalities, it is also common to study between different races (for a summary of studies during the study period, see, for

example, Porterfield, 1982). Kennedy (1944, 1952) has also done extensive research on intermarriage during the beginning of the 1900s between ethnicities, where she found that intermarriage generally is channeled by religious barriers, where individuals within the same religion tend to intermarry. This was later criticized for being simplistic and neglecting groups, among other things (e.g., Peach, 1980b). However, Lindmark's (as cited in Norman, 1976) studied intermarriage among Swedish immigrants between 1914 and 1932, and he found that Swedes rarely intermarried outside of their religious faith. Rosenfeldt (2008) states that racial endogamy was high during the mass migration era, and race in the marriage market was, and still is the most powerful division. Thus, during the early 1900s, racial intermarriage was rare.

Previous research has found several individual-level characteristics that influence the choice of a spouse for intermarriage. Among these factors, ethnicity plays a significant role. Another individual factor that influences intermarriage is education. Highly educated individuals are more inclined to intermarry compared to those with lower educational attainment. Similarly, socioeconomic status also plays a role in intermarriage rates, and individuals from higher socioeconomic backgrounds are likelier to engage in intermarriage. Furthermore, in the early twentieth century, English-speaking ability was a significant factor in intermarriage among immigrants in the United States. Finally, race and religion play an essential role in intermarriage. Proficiency in English was associated with higher intermarriage rates. Age or age at marriage is commonly included as a control.

3.3   Gender Differences in intermarriage patterns

Studying patterns for men and women separately has generated ambiguous results. Choi and Tienda (2017) studied intermarriage patterns at the end of the 1900s and compared gender differences in mate selection behavior. Their results suggest that there are gender differences in mate selection behaviors, with men being more likely to intermarry. Furthermore, Pagnini and Morgan (1990) investigate the assortative mating patterns of European immigrants and native whites by ethnicity and generation, utilizing two data sets: a national sample from the 1910 public use sample census and a sample of marriages registered in New York City from 1908 to 1912. The study found that the pattern of ethnic intermarriage was comparable for both genders. Dribe, Hacker, and Scalone (2023) also found no consistent evidence of gender

differences in these intermarriage patterns. McCaa (1993) also studied intermarriage patterns in the early 1900s and found that gender played a role in intermarriage patterns, with men being more likely to marry outside their ethnic group than women. Furthermore, he states that gender has a structural influence on the overall likelihood of marriage. Understanding the social forces that operate through gender dynamics in the marriage market requires thorough investigation and can offer valuable insights. Similarly, Chiswick and Houseworth (2011) found that men were more likely to intermarry at the end of the 1900s. The issue of ambiguity concerns studies irrespective of time, which suggests that the differences are not due to the fact that they have changed over time.

Spörlein, Schlueter, and van Tubergen (2014) included a sex ratio variable in which they included an interaction with gender as the availability of a spouse of the opposite sex within one's origin group differs between men and women. The study indicates that as the gender balance within a community becomes more imbalanced, the odds of exogamy increase over time. Specifically, if the sex ratio decreases by one unit, there is a 14 percent increase in probability for males. The corresponding results suggest a more significant increase of 29 percent for females. In addition, in a study of linguistic intermarriage, Stevens and Schoen (1988) found that men and women weigh culturally specific traits, such as language, and economically relevant characteristics, differently. Hence, reviewing the literature, it has been found that previous studies have either found that there are no differences in the intermarriage pattern between men and women or that men are more likely to intermarry. This indicates a need to look deeper into the gender differences in intermarriage patterns and study gender differences in the factor affecting intermarriage.

4 Theoretical Framework

Intermarriage is a multifaceted and complex phenomenon influenced by various social, economic, and cultural factors. Swedish migration, driven by a combination of push and pull factors, resulted in a significant portion of migrants choosing to settle down in the United States permanently. This decision entails establishing roots and planning for the future in the United States. As individuals build their lives in this new country, marriage becomes pivotal. A part of marriage is individuals' options of marrying someone from their own ethnic background or from a different origin, influenced by various determining factors. Consequently, these factors shape marriage patterns and impact the decision between intermarriage and non-intermarriage. Although people often marry within their own ethnic group, it is common for individuals to marry someone from a different group. This mixing of ethnic and racial groups, known as miscegenation, happens everywhere and involves interbreeding and intermarriage. All groups interbreed, blending different ethnic and racial backgrounds (Marcson, 1950). Kalmijn (1991) outlines three social forces influencing marriage patterns: partner preference for certain characteristics, social group influence, and marriage market constraints. The following section embeds the research with theory to create a conceptual framework of marriage patterns in the context of intermarriage among Swedish immigrants in the United States.

4.1 Partner Preferences

Kalmijn (1998) emphasizes that preferences for cultural similarity in spouses stem from shared values, opinions, tastes, and knowledge. Such similarities lead to mutual validation of behavior and worldviews, increased participation in joint activities, and enhanced mutual understanding. However, it should be noted that preferences for socioeconomic and cultural resources do not necessarily result in homogamy (choosing a partner of similar social or educational status) or endogamy (marrying within one's own ethnic group) when it comes to social characteristics. Homogamy and endogamy may unintentionally arise due to individual

14

preferences for certain resources in a partner. Nonetheless, social characteristics such as education, occupation, race, and ethnicity serve as symbolic indicators that individuals use to project their identities, influencing their choice of spouse.

The cultural theory posits that cultural factors, such as beliefs, values, and norms, significantly shape social behavior. This theory suggests that immigrants prefer spouses of the same nationality as they share characteristics such as language, experience, and norms and are thus more likely to marry someone of the same origin (Baily, 1980). Language is particularly singled out as it lays the foundation for communication between two people, and an individual is more likely to intermarry with someone from a different origin, the better their English skills (Chiswick & Houseworth, 2011). In contrast, cultural pluralism theory suggests that immigrants can maintain their cultural identity while participating in American society, and intermarriage can bridge cultures (Gans, 1979). When people from different backgrounds get married, they may have different ways of adapting to each other's cultures, but there is usually trade between different characteristics (Baily, 1980; Berry, 1997). A lack of proficiency in English may hinder communication and the ability to navigate cultural differences within a marriage. Thus, individuals with limited English skills may be more prone to choose partners from their own cultural background due to shared language and familiarity.

Additionally, the homogamy theory highlights the role of status and prestige in partner preferences. When discussing gendered preferences in marriage, it often comes down to the benefits derived from the division of paid and domestic labor. Men have an advantage in productive work due to prevailing gender differences in earnings. As a result, women's time is often spent on household labor, maximizing productivity. This leads to an exchange of resources between men and women in terms of paid and domestic labor. Similar arguments can be made regarding status and prestige. Suppose family status primarily relies on the husband's occupation; there is an exchange of male prestige and female qualities in other aspects, such as social class background, physical attractiveness, and cultural involvement affecting the marital preference (Kalmijn, 1998). The significance of socioeconomic resources in partner selection stems from a desire to marry someone financially secure, irrespective of one's financial situation. On the other hand, the influence of cultural resources in partner selection arises from a preference to marry someone with a similar cultural background and interests (Kalmijn, 1998). For horizontally differentiated groups, such as ethnic groups,

endogamy is probably more the result of preferences for cultural similarity and not so much the result of competition for economically attractive spouses (Kalmijn 1991a).

The theory posits that the exchange of resources between men and women influences gendered preferences in marriage. It suggests that factors such as socioeconomic resources and cultural background play crucial roles in partner selection, reflecting desires for financial security and cultural compatibility.

4.2   Marriage Market Constraints

The ability of individuals to find their ideal life partner is greatly influenced by the presence of potential partners within their local social circles (Kalmijn, 1998; Schwartz, 2013). This factor plays a significant role in shaping mate-selection behavior. Blau (1977) proposed that the correlation between group size and inmarriage is due to the weakened influence of ingroup norms on intermarriage within small groups. According to his theory, the structural limitations imposed by smallness make it easier for intermarriage to occur. The chances to marry endogamous are higher when one meets people within the group more often and interacts with group members daily. The chance to encounter a member of one's group also depends on how a group is geographically dispersed. Groups concentrated in specific regions of the country generally have more opportunities to marry endogamy than those not. However, relaxing the assumption of an even geographic distribution may lead to new problems. For instance, controlling for geographic segregation implicitly assumes that people base their decision to live in a given area on factors independent of ingroup preferences, which is not always realistic (Blau & Schwartz, 1997). Structural theory, as highlighted by Blau and Schwartz (1997). suggests that minority groups are more likely to engage in intermarriage than larger groups. This is attributed to the fact that individuals belonging to smaller groups have more opportunities to interact with individuals outside their own group, thus reducing social distance.

The patterns of endogamy in marriages are influenced by various factors at both the individual and group levels, but structural arrangements shape opportunities for social interactions (Kalmijn, 1998). The likelihood of marrying within one's group increases when frequent encounters and daily interactions with group members occur. Additionally, the

16

geographic dispersion of a particular group affects the chances of encountering someone from the same group. Concentrated groups residing in specific regions tend to have more opportunities for endogamous marriages compared to dispersed groups (Lieberson & Waters, 1988). However, accounting for uneven geographic distribution presents its own challenges. Assuming that people make residential decisions independently of their in-group preferences is often unrealistic, but their geographical position likely affects preferences and opportunities (Blau & Schwartz, 1997).

The marriage market plays a significant role in influencing intermarriage rates as it creates opportunities; however, individual preferences also play a crucial role in determining whether someone chooses to intermarry. Exploring variations in preferences and understanding how changes in the marriage market influence the decision to intermarry can provide valuable insights into the gender differences observed in some research on intermarriage patterns.

4.3   Hypotheses

The first and second hypotheses are based on the theory that the availability of potential partners within one's group affects the likelihood of intermarriage, which has also been confirmed by previous research (e.g., Dribe, Hacker & Scalone, 2018). Furthermore, Spörlein, Schlueter, and van Tubergen (2014) found by interacting with sex among European countries that the probability of intermarriage increases more for women than for men when the availability ratio of spouses within one's group decreases.

H1: The probability of intermarriage is hypothesized to increase in the available marriage market of Swedish-born spouses.

H2: Females are hypothesized to be more affected by a smaller available marriage market than males, thus increasing the probability of intermarriage more for females than males.

The third and fourth hypotheses are intertwined with the first research question, driven by the underlying theory that an individual's preference for a spouse is influenced, at least in part, by the desire for perceived similarity with a potential partner. Having English proficiency increases communication with potential spouses that are non-Swedish. Furthermore, the study

by Stevens and Schoen (1988) suggested that men and women weigh characteristics such as language differently.

H3: Having English proficiency is hypothesized to be associated with a higher probability of intermarriage.

H4: It is hypothesized that the association between English proficiency and intermarriage will vary between males and females.

5   Methodology

To answer the research questions of whether there are gender differences in how the marriage market and the ability to speak English are associated with intermarriage, a quantitative method is used. Previous studies have rarely focused on Sweden exclusively and rarely explored different meanings of certain factors between genders. The choice to intermarry is affected by many individual and contextual factors; thus, factors at both levels are included in the model. A binary fixed effects logistic regression model is utilized to estimate the probability of how the contextual factor marriage market and individual factor ability to speak English are associated with intermarriage[2]. The study is limited to studying gender differences in these two factors.

5.1   Data

The dataset used for the analysis is a 25 percent random sample retrieved from IPUMS USA full count data (1790-1940) (Ruggles et al., 2023). The full count dataset is compiled from census microdata at the individual and household level in the United States and is a collaboration with Ancestry.com (IPUMS USA, 2023).

---

[2] In reality, the binary regression model provides odds ratios interpreted as the odds that an outcome (intermarriage) will occur given a particular exposure relative to the base outcome (no intermarriage).

In this study, micro-level data based on censuses from 1900 and 1910 are limited to individuals between 20 and 45 years old (full sample). This sample was not restricted to solely include Swedish immigrants as all ethnicities were needed to calculate the local marriage market variable (see section 5.2.). The sample used for the regression analysis (restricted sample) consists of individuals whose birthplace is Sweden and living in the United States during the censuses in 1900 and 1910. Furthermore, this sample only includes married individuals in the census who married after immigrating to the United States.

	Full sample		Restricted sample	
Census year	Freq.	Percent	Freq.	Percent
1900	7,390,088	44.31	11,294	62.27
1910	9,288,379	55.69	6,842	37.73
Total	16,678,467	100.00	18,136	100.00

Ethnicity is defined using birthplace, as previously discussed. In addition, other variables from the census include the years an individual has lived in the United States and nuptiality variables such as marital status and years in current marriage. Marriages taking place in the United States for the restricted sample are identified by restricting to individuals whose duration of marriage was less than their years in the United States and limiting to individuals in a marital relationship. Additionally, the spouse's birthplace is used to distinguish endogamy from exogamy couples.

5.2   Variables

This section specifies the describes the variables, summarized in Table 1. The variables included are compiled from the previous research and shown to influence the probability of intermarriage. The outcome variable in this study is intermarriage. It is constructed by using the spouse's birthplace and assigning all individuals in the sample married to a spouse with a non-Swedish birthplace 1 and those with a Swedish-born spouse 0.

The first variable of interest is the marriage market. As the theory suggests, the marriage market is a determining factor of whom an individual will marry. It shows the availability of potential spouses within one's group among all potential spouses. The marriage market variable is adapted from Goldman, Westoff, and Hammerslough (1984). However, it is in this study constructed using the enumeration district as the geographical variable. The variable is created by taking the share of Swedish born of the opposite sex, as the dataset only covers heterosexual marriages in an enumeration district, divided by the total number of individuals of the opposite sex in that enumeration district. Thus, the total number of Swedish-born men divided by the total amount of men in a specific district gives the marriage market for the women living in that enumeration district (*d*) and vice versa (see Equation 1). This variable will serve as a proxy for the local marriage market where the person lives. One is marriage market value is constructed for women and one for men in each enumeration district.

$$Marriage\ market\ female_d = \sum \frac{Swedish\ born\ male_d}{Totat\ male_d} \qquad (1)$$

The marriage market (MM) variable is calculated on married and single individuals; as Dribe, Hacker, and Scalone (2018) pointed out, using the entire population rather than just the single population of marriageable age is considered more relevant.

The second variable of interest is the ability to speak English (SE), a categorical variable indicating the respondent's English proficiency. This variable is used to answer the research question of whether the ability to speak English significantly is associated with intermarriage and distinguish differences in its association between females and males. The variable does not include different levels of English, as this was not included in the 1900 and 1910 censuses. In Dribe, Hacker, and Scalone (2018).

The variables female (FEM) and age (AGE) are included to control for individual characteristics. Female is a dummy variable where 0 stands for male and 1 for female, and age

is a continuous variable ranging from 20 to 45[3]. Furthermore, the squared age (AGE2) is included to account for non-linearity in the variable. The variable years in the United States (YUS) are included in the analysis to capture the effect of assimilation on the likelihood of intermarriage. The underlying theory suggests that individuals who have spent more time in the United States are likely to be more assimilated or integrated into the culture, which may increase their likelihood of intermarriage. In other words, longer duration in the United States could lead to adopting cultural norms and behaviors that facilitate intermarriage. Additionally, including the squared term of years in the United States (YUS2) accounts for the possibility of a non-linear relationship between the duration of stay in the United States and the likelihood of intermarriage. By including the squared term, the analysis considers the potentially diminishing returns in the effect of assimilation. This allows for a more flexible and accurate representation of the relationship between the duration of stay and intermarriage likelihood.

Furthermore, occupational score (OS) is a continuous variable indicating the occupational score of the respondent. We include this variable as a proxy of socioeconomic status associated with intermarriage. This is not a perfect measure of socioeconomic status as many women during this period did not work after marriage, as also pointed out by Dribe, Hacker, and Scalone (2018). It is merely included to control for socioeconomic status, but interpretation should be made with caution. The variable was divided into five categories, 1-20, 21-40, 41-60, 61-80, and missing (no reported occupation). Including the missing category is necessary because a significant number of women did not report their occupations. It is crucial to include this group of individuals in the study to avoid excluding a substantial portion, especially women, which would undermine the purpose and significance of the study.

---

[3] see section 5.2.1 for a discussion of this limitation.

Urban is a dummy variable indicating whether the respondent lives in a rural (0) or urban (1) area. We include this variable as an indicator variable to control for any urban/rural differences in intermarriage rates. Finally, a year dummy is included to control for influence from the different census years, coded 0 for 1900 and 1 for 1910.

Table 1: Overview of the variables

Variable	Abbvariation	Unit
Intermarriage	IM	Yes/No
Marriage Market	MM	Share of Swedish men/women in each enumeration district
Speaks English	SE	Yes/No
Female	FEM	Male/Female
Age	AGE	Years
Years in the US	YUS	Years
Occupational Score	OC	Category
Urban/Rural	UR	Urban/Rural
Census year	YEAR	1900/1910

Source: (Ruggles et al., 2023)

5.2.1   Representativity

The representativeness concerns how well the full sample represents the population in each enumeration district and how well the restricted sample represents Swedish immigrants who

married in the United States. Due to limitations with Stata, the sample size used in the study was limited to 25 percent of IPUMS USA full count data; thus, only a sample of Swedish immigrants in the United States during this period is included. Because of random sampling, their representativity in each enumeration district should represent the actual proportion of Swedish men/women compared to the overall number of men/women in that district. Thus, the sampling constraint potential measurement errors due to the smaller sample size relative to the entire population are unlikely to be systematic. However, there could have been individuals who lived in the United States before or between the census years, which are not represented. Thus, this study's results only apply to 1900 and 1910. In addition, individuals could potentially have lived in a different enumeration district once selecting a partner, hence a different marriage market. However, given the substantial sample sizes, it is presumed that this measurement error is negligible.

To ensure maximum inclusivity of individuals in the study, the age range was limited to 20 to 45. This age range was also implemented as individuals within this range are assumed to be actively selecting a partner. However, it is essential to note that this age range limitation may limit the generalizability of the results only cover individuals within this age range. However, as the analysis is made on a sample, generalizations should always be cautiously made.

5.3   Logistic Regression Model

The method for this study is commonly used when studying intermarriage patterns when the outcome variable is binary (e.g., Dribe, Hacker & Scalone, 2018; Jacobs & Labov, 2002; Spörlein, Schlueter & van Tubergen, 2014). A conditional fixed effects logit regression is used to analyze the likelihood of intermarriage between Swedish-born individuals and non-Swedes. Logistic regression is a statistical method used to model the probability of a binary outcome. It is commonly used to analyze dichotomous variables, which is a variable that can take two distinct values (1/0) (Best & Wolf, 2013). Thus, the model predicts the odds of marrying another Swedish-born individual relative to marry an individual outside of one's ethnical group. Another commonly used model for studying intermarriage is the log-linear model (e.g., Alba & Golden, 1986). However, since the dependent variable in this study is dichotomous (intermarried/not intermarried), the binary logit model is used.

Using, for example, the linear probability model (LPM) for analyzing dichotomous outcomes raises several concerns, as Best and Wolf (2013) discussed. Firstly, the LPM assumes a linear relationship between independent variables and the probability of the outcome. However, with dichotomous outcomes, the relationship is non-linear, resulting in poor model fit and inaccurate predictions. Secondly, the LPM assumes constant error variance across all levels of independent variables. However, with dichotomous outcomes, the error term correlates with the dependent variable, violating the assumption of homoscedasticity and leading to inefficient estimation and biased standard errors. Additionally, the LPM can produce predictions outside the range of probabilities $(0 \leq Pr(y = 1) \geq 1$. Lastly, the LPM assumes normally distributed errors, but dichotomous outcomes typically lead to non-normally distributed residuals. When analyzing dichotomous outcomes, a non-linear model such as the logistic regression will avoid violating the assumptions above and provide more appropriate and accurate estimates and predictions for binary outcomes.

This study utilizes the conditional fixed effects logit regression model as the Hausman test suggests that the fixed effects model is more appropriate than random effects (see 6.3 for further discussion). Using county-fixed effects incorporates county-specific effects, capturing unobserved heterogeneity among counties, leading to more robust standard errors. Important to note is that the county-fixed effects capture time-invariant effects on the county level, but it does not capture time-variant effects. This approach acknowledges the potential correlation or heteroscedasticity within counties and provides more robust standard errors (Wilson & Lorenz, 2015). Intermarriage rates can vary substantially across different counties. Fixed effects models explicitly control for group-level heterogeneity by estimating separate intercepts for each county. This approach accounts for unobserved characteristics that vary across counties and may be correlated with the outcome variable. Including fixed effects allows for capturing county-specific factors that may influence intermarriage rates but remain constant over time. Fixed effects models yield more efficient estimates than random effects models because they exploit within-group variation. Additionally, fixed effects models are generally robust to omitted variable bias, as they control for time-invariant unobserved heterogeneity at the county level. However, it's worth noting that fixed effects models have some limitations. They cannot estimate the effects of time-varying predictors, as these models rely on within-group variation (Pforr, 2014).

The logit model expresses the association between a predictor variable X_i and the conditional probability that the outcome variable Y_i equals one. The function can be represented as:

$$P(Y_i = 1) = \frac{e^{\beta_0 + \beta_1 X_i}}{1 + e^{\beta_0 + \beta_1 X_i}} \qquad (2)$$

Where $P(Y_i = 1)$ is the conditional probability that the outcome variable equals one (Sommet & Morselli, 2017), the logit function is the natural logarithm of the odd, so Equation 2 can be converted to a post-logit transformation logistic regression which can be presented as:

$$logit\ (odds) = \beta_0 + \beta_1 X_i \qquad (3)$$

The logit(odds) are the log-odds, thus corresponding to the logit of the conditional probability that the outcome variable equals one divided by the probability that it equals zero (Sommet & Morselli, 2017).

This study aims to identify whether there are any gender differences in how the marriage market and the ability to speak English associated with intermarriage. Thus, these two predictor variables are added to interaction with females. To explore potential gender differences in the relationship between these two variables and intermarriage, two models are formulated, one with each interaction. The county fixed effects are also added,d and following Wilson & Lorenz (2015), the first model for this study is specified as follows:

$$logit\ odds\ (IM_i)$$
$$= \beta_0 + \beta_1 SEX_i + \beta_2 MM_{id} + \beta_3\ SE_i + \beta_4 AGE_i + \beta_5 AGE2_i + + \beta_6 YUS_i$$
$$+ \beta_7 YUS2_i + \beta_8 OC_i + \beta_9 UR_i + \beta_{10} YEAR_i + \gamma_c$$

Where i represent the individual and d the enumeration district in which the individual lives, the initial model presented above serves as a foundational model that excludes any interactions, allowing us to observe the outcomes independently of their combined effects by starting with this simplified approach, we can gain insight into the individual impacts of variables before introducing the complexity of interactions into the analysis. This step helps establish a baseline understanding of the model's outcomes and provides a solid starting point for further exploration and refinement.

26

$$logit\ odds\ (IM_i)$$
$$= \beta_0 + \beta_1 MM_{id} * SEX_i + \beta_2 SE_i + \beta_3 AGE_i + \beta_4 AGE2_i + +\beta_5 YUS_i$$
$$+ \beta_6 YUS2_i + \beta_7 OC_i + \beta_8 UR_i + \beta_9 YEAR_i + \gamma_c$$

In the second model above, female and marriage markets interact. Finally, in the third model below, female and English speaking interacted:

$$logit\ odds\ (IM_i)$$
$$= \beta_0 + \beta_1 MM_i + \beta_2 SE_i * SEX_i + \beta_3 AGE_i + \beta_4 AGE2_i + +\beta_5 YUS_i$$
$$+ \beta_6 YUS2_i + \beta_7 OC_i + \beta_8 UR_i + \beta_9 YEAR_i + \gamma_c$$

In this model, we are regressing the binary dependent variable intermarriage on a number of independent variables. The model includes the interaction between the marriage market (MM) and females (FEM) (the first model) and between speaking English (SE) and female (FEM) (the second model). Then, age (AGE), age squared (AGE2), years in the United States (YUS), the years in the United States squared (YUS2), occupational score (OC), urban/rural dummy variable (UR), and a year dummy (YEAR) are included as controls. The term γ_c represents the county fixed effects. To ensure the model's reliability, a sensitivity analysis was conducted where the marriage market was divided into four levels of marriage market; low, medium-low, medium-high, and high. To complement this, several econometric tests were conducted to assess its robustness, which is discussed in section 6.3.

6 Empirical Analysis

6.1 Descriptive Statistics

Before looking at the regression output and analyzing the potential differences between males and females in the variables of interest, it is helpful to get a descriptive overview to understand how the characteristics differ between males and females. This section describes the restricted sample; thus, the sample used for the analysis consists only of Swedish immigrants between 20 and 45 years old who married after migration. Table 2 presents the number of intermarriages or not intermarried men and women separately. This shows expectingly that more individuals in this sample did not intermarry, and these shares between not intermarried and intermarried were quite similar between men and women. A slightly larger share of men was intermarried, around six percentage units.

Table 2: Share of intermarriage by sex

	Male	**Female**	**Total**
No Intermarriage	73,95%	79,86%	77,45%
Intermarriage	26,05%	20,14%	22,55%
Total	100%	100%	100%

Table 3 displays an overall description of the characteristics of male and female intermarried and not intermarried individuals. The mean marriage market is similar for males and females, and interestingly the marriage market appears to, on average, be larger where not intermarried individuals lived and similar for both males and females. However, as shown by the statistics in Table 4, the standard deviation for the male marriage market is 0.09, which is high. Another interesting observation is that for males, those who are not intermarried, on average, have English proficiency compared to intermarried individuals. The opposite and more

28

expected shows for women where it is more common to intermarry when having English ability than those who do not. The ages are relatively similar between the sexes, but as Table 3 shows, not intermarried individuals seem to be slightly higher than intermarried individuals. Furthermore, intermarried individuals have, on average, lived fewer years in the United States than not intermarried individuals. The occupational score shows, as expected, a low average for males and a high value for females, as many females did not report an occupation and thus are in the fifth category. On average, individuals in the sample live in an urban area. However, the standard deviation in Table 4, which is about 0.5 regardless of sex, indicates a large variation in the variable. Finally, the year dummy indicated that the observations were relatively evenly distributed between 1900 and 1910.

Table 3: Descriptive statistics of male and female, intermarried and not intermarried

Variables	**Male**		**Female**	
	Intermarried (n=1848)	Not Intermarried (n=5324)	Intermarried (n=2033)	Not Intermarried (n=8282)
Marriage market	0.04	0.07	0.04	0.07
Speaks English	0.87	0.91	0.95	0.85
Age	36.03	38.22	34.15	36.73
Year in the US	4.7	8.97	4.67	7.84
Occupational Score	1.91	1.82	4.57	4.89
Urban	0.55	0.59	0.53	0.56
Year dummy	0.48	0.39	0.39	0.37

Table 4: Descriptive statistics over the variables, male and female, separately.

Variable	**Male** n= 7,172				**Female** n=10,315			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Marriage market	0.06	0.09	0	0.65	0.07	0.1	0	0.7
Speaks English	0.9	0.31	0	1	0.87	0.34	0	1
Age	37.65	5.44	20	45	36.22	6.24	20	45
Year in the US	7.87	6.16	0	41	7.21	6.83	0	40
Occupational Score	1.84	1.07	1	5	4.83	0.79	1	5
Urban	0.58	0.49	0	1	0.55	0.5	0	1
Year dummy	0.41	0.49	0	1	0.38	0.48	0	1

6.2 Results

The output in Table 5 shows the results from the fixed effects logit regression with counties as the group variable. The first column presents the regression results from Model 1 on the restricted sample, that is, a random 25 percent sample of Swedes between 20 and 45 who got married after migration to the United States. The second column includes interaction between females and the marriage market, and the third column shows the results when an interaction between females and English proficiency is included. The results are presented in odds ratios for more intuitive and straightforward interpretations, which is the exponentiated log odds. If the odds ratio equals 1, it suggests that there is no association between the predictor variable and the probability of the outcome variable. In other words, when exposed to that predictor, the odds of the outcome occurring are the same for both outcome groups (intermarried versus not intermarried). A value higher than 1 indicates that the odds of intermarriage are higher than not intermarried when exposed to a predictor, and below 1 implies the opposite. Important to note that all interpretations are when holding other variables constant.

Twenty-four counties consisting of 36 observations were automatically excluded by Stata from the analysis because they had all positive or negative outcomes. In logistic regression, it is essential to have variation in the outcome variable within each group or category. If a group has only positive or negative outcomes, it does not provide any information for estimating the model parameters and may result in computational issues or biased estimates.

Table 5: The results from the county fixed effects logit regression model with intermarriage as the dependent variable in all regression

VARIABLES	Model 1 intermarriage	Model 2 intermarriage	Model 3 intermarriage
Male		Reference group	
Female	1.09	**1.19****	**0.38*****
	(0.09)	(0.10)	(0.06)
Marriage market	**0.005*****	**0.01*****	**0.005*****
	(0.00)	(0.01)	(0.00)
Female*Marriage market		**0.13*****	
		(0.08)	
English Speaker			
No		Reference group	
Yes	**1.89*****	**1.89*****	1.09
	(0.13)	(0.13)	(0.10)
Female*English speaker			**3.22*****
			(0.46)
Age	**0.85*****	**0.85*****	**0.85*****
	(0.03)	(0.03)	(0.03)
Age^2	**1.00*****	**1.00*****	**1.00*****
	(0.00)	(0.00)	(0.00)
Years in the US	**0.80*****	**0.80*****	**0.80*****
	(0.01)	(0.01)	(0.01)
Years in the US^2	**1.01*****	**1.01*****	**1.01*****
	(0.00)	(0.00)	(0.00)
Occupational Score			
1-20		Reference group	
21-40	**0.85*****	**0.84*****	**0.87****
	(0.05)	(0.05)	(0.05)
41-60	**1.37****	**1.37****	**1.44****
	(0.21)	(0.20)	(0.22)
61-80	1.96	1.93	2.10
	(1.57)	(1.55)	(1.69)
Missing	**0.45*****	**0.45*****	**0.46*****
	(0.04)	(0.04)	(0.04)
Rural		Reference group	
Urban	1.00	1.00	1.00
	(0.05)	(0.05)	(0.05)
Year	**1.62*****	**1.62*****	**1.60*****
	(0.07)	(0.07)	(0.07)
Observations	17,451	17,451	17,451
Number of counties	112	112	112
Log-likelihood	-7774.76	-7769.85	-7739.22
Overall p	0.00	0.00	0.00
Pseudo R^2	0.1149	0.1155	0.1190

Notes: Standard errors are presented in parentheses and *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The result from Model 1, presented in Table 5, suggests that the marriage market, ability to speak English, age, years in the United States, and occupational score are significantly associated with the odds of intermarriage. Contrary to previous findings, this exploration of Swedish immigrants in the United States reveals no notable disparity between men and women regarding the probability of intermarriage (Model 1). Furthermore, looking at Model 2, when including the interactive effect of the female and marriage market, women appear to significantly have higher odds of intermarriage than men.

The odds ratio of 0.005 indicates that when the marriage market increases by one unit, the odds of intermarriage decrease by 0.005 times; thus, the odds decrease by 99.5 percent. This demonstrates a significant reduction in the odds of intermarriage as the share of potential partners within one's own group increases, confirming the first hypothesis (H1). Furthermore, the first model indicates that the ability to speak English significantly increases the odds of intermarriage. The odds ratio of 1.89 suggests that individuals who speak English have 89 percent higher odds of intermarriage compared to those who do not, confirming the third hypothesis (H3).

Model 1 further shows that Age exhibits a statistically significant negative association with intermarriage. As individuals' age increases, the likelihood of intermarriage decreases. This indicates that younger individuals in the sample have a higher likelihood of intermarriage. The odds ratio for age is 0.85, indicating that the odds of intermarriage decrease by 15 percent with each additional year of age. Years in the United States exhibit a statistically significant negative association with intermarriage. As the number of years spent in the United States increases, the likelihood of intermarriage decreases. The odds ratio for years in the United States is 0.80, suggesting that for every additional year, an individual has lived in the United States, the odds of intermarriage decrease by approximately 21 percent. In other words, individuals who have spent more years in the United States are less likely to intermarry compared to those with fewer years in the United States.

Having an occupational score in the group 21-40 significantly decreases the odds of intermarriage compared to the reference group 1-20. However, being in the group of occupational scores between 41-60 significantly increases the odds of intermarriage compared to the reference group. As previously mentioned, there is also a group of individuals who do not have an occupational score, likely because they are not working, and this group has a

lower probability of intermarriage compared to those in the reference group. The odds ratio indicates that those who do not have an occupational score have 55 percent lower odds of intermarriage than the reference group. There does not appear to be any significant difference in the odds of intermarriage between living in an urban or rural area.

The year dummy indicates that there are significantly higher odds of intermarriage in 1910 than in 1900; with an odds ratio of 1.63, the relative odds of intermarried are 63 percent higher in 1910 than in 1900. The control variables do not substantially deviate throughout the three models, which suggests stability in the model and variables despite adding interactive variables.

In Model 2, when examining gender differences in the marriage market, the results show significant gender differences. In comparison to men (used as the baseline), women have lower odds of intermarriage when the marriage market is larger. Specifically, the odds of intermarriage decrease more for women than men as the marriage market increases. The odds ratio of 0.13 suggests that, for the interaction between the female and marriage market, the odds of intermarriage are significantly lower for females compared to males if the availability of potential spouses within the ethnic group and enumeration district (marriage market) increases. The results support the second hypothesis that females are hypothesized to be more affected by a smaller available marriage market than males.

In Model 3, the output suggests that there are gender differences in the association between having English proficiency and intermarriage, confirming the fourth hypothesis (H4). The odds for an English-speaking man to intermarry are 1.09 times higher than a non-English-speaking man. However, there is not a significant difference between English-speaking and non-English-speaking men. The odds for a non-English speaking woman to intermarry are 0.38 times the odds of a non-English speaking man to intermarry. This suggests that non-English speaking women have significantly lower odds of being intermarried than non-English speaking man. The odds for an English-speaking woman to intermarry are 1.33 (0.38*1.09*3.22) times higher than that of a non-English-speaking man. The odds of an English-speaking woman being intermarriage are 1.22 (0.38*3.22) times the odds of a non-English-speaking woman.

All models exhibit an overall p-value 0.000, indicating a significant relationship between the predictors and the response variable. This low p-value provides strong evidence against the

null hypothesis, that there is no relationship or association between the predictors and the response variable and confirms the statistical significance of the models. The number of observations and counties included in each model are also reported, and the log-likelihoods for each model are also presented in Table 5. However, it is essential to note that these results are based on the specific sample and variables in the model and may not necessarily generalize to other populations or periods. Other factors not included in the model may also be relevant for explaining changes in intermarriage.

6.3   Robustness Check

To detect potential multicollinearity, a VIF test is performed. Furthermore, a sensitivity analysis is presented in the results sections looking at subsamples and different marriage market categories. Collinearity refers to the degree to which independent variables in a statistical model are correlated with each other. High levels of collinearity can lead to instability and inaccurate estimates in a statistical model. Table 6 in Appendix A shows the VIF value for each independent variable. It is important to consider that the VIF value for age and age squared is exceptionally high (120.47 and 121.34 respectively). This is expected as the squared age is generated from the variable age. Typically, a VIF above ten should lead to excluding one of those variables to avoid collinearity. However, it is not worrying in this case and both variables are left in the model. Comparing the VIF-values of the other variables when the squared age is included and when it is not showing that the inclusion of the squared age has no considerable effect on the other variables, nor the regression result (Appendix B Table 7). Years in the United States and the squared years in the United States slightly increase, but no major changes appear; thus, the squared age is left in the model. It shows that some variables in the model (e.g., years in the United States and the years in the United States squared) have high VIF values, indicating high levels of collinearity. This could be a concern for the stability and accuracy of the statistical model. However, since they are below 10, they are kept in the model.

The Hausman test compares fixed effects and random effects logit models. These models differ in how they handle unobserved heterogeneity. The fixed effects logit model includes individual-specific effects, while the random effects logit model assumes these effects are random and uncorrelated with the independent variables. The Hausman test checks whether

the random effects model is appropriate by testing if the individual-specific effects are uncorrelated with the independent variables. If the test rejects this assumption, it suggests that the fixed effects model is preferable. If the statistic exceeds a critical value, the null hypothesis is rejected in favor of the alternative hypothesis that the fixed effects model should be used. The Hausman test helps choose the appropriate model by considering the trade-off between consistency and efficiency in handling unobserved heterogeneity (Amini et al., 2012). Thus, the Hausman test on Model 1 gives a p-value of 0.036 (<0.05), the null hypothesis is rejected, and the fixed effect is the more appropriate model.

Additionally, we examine the inclusion of interaction variables in Models 2 and 3 by looking at the pseudo-R squared values in Table 5[4]. When the interaction between the female and marriage market is included (Model 2), the pseudo-R squared increases to 0.1155, thus a slight increase compared to the first model, which had a pseudo-R-square. Including the interaction between female and English proficiency also increases the pseudo-R squared to 0.1190. Furthermore, a Wald test was performed on the respective interactive variables parameters, which for the parameters connected to the interaction between female and marriage market gave a p-value of 0.004 (<0.05), indicating that the null hypothesis that the coefficients are jointly zero can be rejected. The p-value for the Wald test on the parameters connected to the interaction between females and English proficiency similarly provided a p-value of 0.000 (<0.05), thus rejecting the null hypothesis. Hence, both interactive variables significantly improve the respective models (Appendix C Table 8). Finally, to test if the models are improved by including the interactions, Akaike's (1973) information criterion (AIC) is used, which is presented in Appendix D (Table 9). A lower value indicates a better-fitting model (Long & Freese, 2006). As both the AIC for Models 2 and 3 are lower than the

---

[4] Important to note is that pseudo-R^2 is not equivalent to the R^2 in a linear regression model. Pseudo-R^2 values can only be meaningfully interpreted and compared within the same dataset, outcome variable, and type of pseudo-R^2, indicating that the current model exhibits a better fit than the previous model for predicting the outcome (Long & Freese, 2006).

base Model 1, it suggests that including the interactions gives the model a better fit than not including them.

In Table 10 of Appendix E, an alternative modeling approach is presented together with the output from Models 1-3, as shown in Table 5. This approach involves replacing the marriage market variable with a categorical variable consisting of four categories: low (<0.011), medium-low (0.011-0.027), medium-high (0.027- 0.067), and high (>0.067). Each category represents a specific percentile range within the dataset, with approximately 25 percent of the observations allocated to each group. This categorization allows for further examination of the robustness of the model by comparing the outcomes with the results obtained in Model 3. Such an analysis helps assess whether the observed relationships and patterns in the model hold across different representations of the variable, thereby testing the stability and generalizability of the results. The results show that the odds for a man living in an area with a medium-low marriage market to intermarry is 0.46 times that of a man in an area with a low marriage market, and the odds decrease as the marriage market increases. The odds of women living in an area of a low marriage market to be intermarriage are 1.27 times higher than that of a man in the same marriage market category. This further indicates the previous results that the odds for women in lower marriage markets to be intermarriage are significantly higher than for men.

6.4   Discussion

In this section, the obtained empirical results will be discussed and contextualized. The analysis aims to provide a comprehensive understanding of the findings by placing them within their relevant context. It will also discuss the significance and implications of the results, providing a deeper understanding of their implications. The objectives of this study were to explore different meanings of certain factors between genders in factors shown in previous literature affecting the odds of intermarriage. The study is limited to studying Swedish immigrants in the United States with data retrieved from IPUMS USA 1900 and 1910 censuses.

This study supports the theory and aligns with previous studies' results that the increased availability of potential partners within the same ethnic group reduces the likelihood of intermarriage with other ethnicities among Swedish immigrants. This is likely because, as the

theory predicts, individuals often prefer spouses which whom they share similarities. Thus, two individuals are expected to prefer an individual with the same origin if there is an availability of such. This study further expanded this to study gender differences in this relationship, and the results show significant gender differences in this relationship, with women being more sensitive to changes in the marriage market. The interaction between females and the marriage market substantially impacts the likelihood of intermarriage. Women appear to have experienced an increased likelihood of intermarriage compared to males, particularly in areas with a lower share of the opposite sex born in Sweden. The results align with what Spörlein, Schlueter, and van Tubergen (2014) found and support the hypothesis suggesting that these differences are also prevalent among Swedish immigrants. This interaction effect highlights the importance of considering both the individual's sex and the local marriage market conditions when examining intermarriage patterns.

This result could be linked to the fact that when there is limited availability of potential partners within the same social group or community where an individual resides, women may be more sensitive to this constraint. As previously stated, during the period under study, a significant proportion of women did not participate in the workforce or work in a limited number of occupations, which further restricted their opportunities to meet potential partners. In contrast, men's meeting platforms might be more diverse due to their engagement in employment outside of their immediate community. This difference in meeting opportunities could influence the likelihood of intermarriage between genders.

That the results show that having at least some proficiency in English increases the odds of intermarriage compared to those who do not speak any English; however, it is essential to consider the limitations of the questionnaires used in 1900 and 1910, which only inquired about whether respondents could speak English or not. This creates a gray area where individuals with limited English proficiency may classify themselves either as English speakers or non-English speakers. Consequently, this ambiguity can impact the validity of the results. With this in mind, the empirical results show that both theory and previous research have shown that English can be a barrier to exogamous marriages, confirming the third hypothesis (H3). In reflection of the fourth hypothesis (H4), the result suggests that there is a gender difference in the odds of intermarriage between men and women who cannot speak English. It shows that non-English speaking women decrease the odds of being intermarried by 62 percent compared to the odds of a non-English speaking man being intermarried. These

results suggest that not speaking English is a bigger barrier for women than for men regarding the odds of intermarriage.

Although the initial model contradicted previous research by suggesting no significant gender difference in intermarriage odds, subsequent analyses in Models 3 and 4 present evidence that there may indeed be gender differences in the factors affecting intermarriage, these models demonstrate significant findings regarding the predictive odds of intermarriage based on factors such as the availability of potential spouses within the ethnic group and enumeration district, as well as the barrier posed by English proficiency. The results suggest that women are more sensitive to changes in the local marriage market than men. Additionally, the findings indicate that the barrier of English proficiency has a stronger impact on the odds of intermarriage for women. The cultural theory suggests that a lack of proficiency in English may hinder communication and the ability to navigate cultural differences within a marriage. Thus, these results indicate that women with no English skills may be more prone to choose partners from their own origin background due to shared language and familiarity (Baily, 1980; Berry, 1997).

As the results in the previous section show, the control variable age implies that the older the individuals in the sample are, the lower odds of intermarriage. This can be linked to Choi and Tienda's (2017) findings that individuals in the study between 25 and 34 were more likely to intermarry than those under 25. In contrast, individuals over the age of 34 were less likely to intermarry than those under 25, suggesting that age increases the odds of intermarriage to a certain age, but then there is a reversed relationship. However, when studying a non-linear relationship in individuals age when limited to 20–45-year-olds, there is no significant non-linearity that can be found. Nevertheless, it is possible that looking at different age groups would provide more insight into this aspect; however, it was not the aim of this study.

Furthermore, the odds ratio for years in the United States suggests that the odds of intermarriage decrease with the years an individual has lived in the United States. This contradicts the findings by Chiswick and Houseworth (2011). However, it is important to note that their study focused on the 1980's census and did not include Sweden. Therefore, comparisons should be approached with caution. It is possible that intermarriage was relatively uncommon up until 1910 (McCaa, 1993), making earlier marriages more likely to be endogamous and, thus, the longer an individual has lived in the United States, the lower

odds of intermarriage. Examining temporal changes in this variable could shed light on whether this hypothesis holds.

Among the control variables, urban stands out the most as the results indicate no significant difference between urban and rural areas. However, this could be because it differs depending on the size of the urban population. For example, Dribe, Hacker, and Scalone (2018) found some notable distinctions between living in urban areas and rural areas among women. However, the results among men suggested no significant difference for urban and rural areas on the odds of intermarriage; neither did higher populated urban areas for women imply a significant difference in odds of intermarriage compared to rural areas. This is partially in line with the results of this study, where no significant differences in the likelihood between urban and rural areas could be found among Swedish immigrants.

To ensure internal validity, a robustness check was conducted to examine the stability and consistency of the results. This involved assessing the variance inflation factor (VIF) to identify potential multicollinearity issues among the predictor variables. The low VIF values indicated no significant problems with multicollinearity, enhancing the internal validity of the analysis. It is important to consider the limitations of external validity due to the sample used in this study, which represents only 25% of the full population count. Therefore, interpretations and conclusions should be made cautiously, recognizing the potential limitations in generalizing the findings to the broader population. In summary, while internal validity was addressed through robustness checks and VIF analysis, external validity considerations should be considered due to the limited sample size. Despite these limitations, this study offers valuable insights into gender differences in intermarriage factors during the study period. It provides a historical context that can inform contemporary studies in this area.

Furthermore, it is important to acknowledge that the model is constrained by the available variables from the censuses used in this study. While it provides valuable insights into gender differences in the factors influencing the likelihood of intermarriage, including additional variables could further enhance the model's explanatory power. Nevertheless, the findings of this study have relevance for modern research as they shed light on historical gender differences in various factors associated with intermarriage. To conduct a more comprehensive analysis of the meaning of different factors influencing intermarriage among Swedish immigrants in the United States, it may be important to consider various factors, such as differences in the spouse's origin. Specifically, it is essential to explore which

nationalities Swedish immigrants tend to intermarry with, as this can provide valuable insights into the dynamics of intermarriage patterns within this population. By examining the specific nationalities Swedish immigrants are more likely to intermarry with, a deeper understanding of the interplay between culture, immigration, and intermarriage can be gained. For example, English proficiency may not be a barrier to intermarriage between a Swede and a Norwegian. Thus, this could explain why it is a greater barrier for women if they more often intermarry with non-Scandinavian individuals. However, additional research is needed to investigate this statistically.

Studying historical intermarriage patterns remains relevant despite the passage of time, as the fundamental theories underlying these patterns persist. As the previous literature throughout the late 1800s and 1900s show, the factors influencing intermarriage are similar. However, it is crucial to investigate historical and contemporary patterns, considering that longitudinal factors can differ. For instance, in today's globalized world, Swedish immigrants are likelier to possess relatively good English proficiency. Still, Spörlein, Schlueter, and van Tubergen's (2014) study suggest that during 1880-2011 looking at English at a cross-sectional level, an individual of an English-speaking origin group increased the odds of intermarriage. However, English proficiency may not hold the same level of relevance in increasing the likelihood of intermarriage among Swedish immigrants today. Therefore, while acknowledging the enduring significance of basic theories regarding intermarriage patterns, it remains important to incorporate both historical and contemporary perspectives in studying intermarriage.

7   Conclusion

Intermarriage is historically and contemporary an important step in the social assimilation of immigrants. Thus, analyzing intermarriage patterns give an understanding of the social assimilation among different immigration groups. This study aimed to explore and gain insight into the factors associated with intermarriage among Swedish immigrants at the beginning of the twentieth century and explore the meanings of certain factors between genders. This analysis of intermarriage patterns among Swedish immigrants in the United States has provided a foundation to answer the two research questions:

1. *To what extent is the local marriage market associated with the likelihood of intermarriage, and how does this vary between male and female first-generation Swedish immigrants aged 20-45 who got married after migrating to the United States?*

2. *How is English proficiency associated with the likelihood of intermarriage, and does this vary between male and female first-generation Swedish immigrants 20-45 who married after migrating to the United States?*

This study aligns with the theory and previous research that a negative association exists between the odds of intermarriage and the local marriage market. Furthermore, contrary to previous research, there is no notable disparity between men and women regarding the probability of intermarriage. However, when considering the interactive effect of sex and the marriage market, women have significantly higher odds of intermarriage than men when the local marriage market decreases. A potential cause of this could be that women generally had a smaller community than men. Thus, when the local marriage market decreased, women tended to be forced to intermarry to a larger extent than men, who possibly had a larger community due to their work.

The analysis also reveals that the ability to speak English significantly increases the odds of intermarriage. Individuals who speak English have 89% higher odds of intermarriage compared to those who lack the skill. However, there are limitations in these findings as there could be a grey zone of people of have limited English skills answering both yes and no to the

question of whether they can speak English, which is important to consider. Furthermore, when exploring the interaction between sex and English proficiency, the empirical result suggests that lacking English proficiency is associated with lower odds of intermarriage among women than men. This has been linked to the theory that when looking for a potential spouse, an individual, there is a trade-off between characteristics, and the results suggest that women tend to prefer endogamy when they lack English proficiency.

Although the results in this study are not exhaustive, they provide further insights into intermarriage patterns among Swedish immigrants in the early 1900s and the meanings of certain factors between genders. Future research could explore a further analysis of the meaning of different factors associated with gender, such as exploring these differences in relation to the origin of the individual's spouse, to further enhance our understanding of intermarriage patterns. Comparative studies across different immigrant groups and periods would also contribute to a more comprehensive understanding of these intermarriage patterns in more contemporary settings as well. Overall, this study adds to the body of knowledge on intermarriage and highlights the importance of considering gender differences in individual preferences and opportunities in intermarriage patterns.

8 References

Abramitzky, R., Boustan, L. P. & Eriksson, K. (2012). Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration, *American Economic Review*, vol. 102, no. 5, pp.1832–1856

Akaike, H. (1973). Maximum Likelihood Identification of Gaussian Autoregressive Moving Average Models, *Biometrika*, vol. 60, no. 2, pp.255–265

Alba, R. D. & Golden, R. M. (1986). Patterns of Ethnic Marriage in the United States*, *Social Forces*, vol. 65, no. 1, pp.202–223

Amini, S., Delgado, M. S., Henderson, D. J. & Parmeter, C. F. (2012). Fixed vs Random: The Hausman Test Four Decades Later, in B. H. Baltagi, R. Carter Hill, W. K. Newey, & H. L. White (eds), *Essays in Honor of Jerry Hausman*, Vol. 29, [e-book] Emerald Group Publishing Limited, pp.479–513, Available Online: https://doi.org/10.1108/S0731-9053(2012)0000029021 [Accessed 22 May 2023]

Baily, S. L. (1980). Marriage Patterns and Immigrant Assimilation in Buenos Aires, 1882-1923, *The Hispanic American Historical Review*, vol. 60, no. 1, pp.32–48

Barr, E. (2015). Swedes in Canada: Invisible Immigrants, University of Toronto Press

Bean, F. D., Lee, J., Batalova, J. & Leach, M. (2000). Immigration and Fading Color Lines in America, *The American People: Census*, pp.302–331

Beijbom, U. (1971). Swedes in Chicago : A Demographic and Social Study of the 1846-1880 Immigration, [e-journal], Available Online: http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-127750 [Accessed 13 April 2023]

Beijbom, U. (2006). Utvandrarkvinnor: svenska kvinnoöden i Amerika, Stockholm: Norstedt

Beijbom, U. (2010). Amerika, Amerika!, Natur & Kultur

Berry, J. W. (1997). Immigration, Acculturation, and Adaptation, *Applied Psychology*, vol. 46, no. 1, pp.5–34

Best, H. & Wolf, C. (2013). Logistic Regression, in H. Best & C. Wolf (eds), *The SAGE Handbook of Regression Analysis and Causal Inference*, SAGE

Blau, P. M. (1977). Inequality and Heterogeneity: A Primitive Theory of Social Structure, Vol. 7, Free Press New York

Blau, P. M. & Schwartz, J. E. (1997). Crosscutting Social Circles: Testing a Macrostructural Theory of Intergroup Relations, New Brunswick, N.J: Transaction Publishers

Bohlin, J. & Eurenius, A.-M. (2010). Why They Moved — Emigration from the Swedish Countryside to the United States, 1881–1910, *Explorations in Economic History*, vol. 47, no. 4, pp.533–551

Carlsson, S. (1976). Chronology and Composition of Swedish Emigration to America, in *From Sweden to America: A History of the Migration*, Minneapolis, Minn: Univ. of Minnesota Press [u.a.], pp.114–48

Chiswick, B. R. & Houseworth, C. (2011). Ethnic Intermarriage among Immigrants: Human Capital and Assortative Mating, *Review of Economics of the Household*, vol. 9, no. 2, pp.149–180

Choi, K. H. & Tienda, M. (2017). Marriage-Market Constraints and Mate-Selection Behavior: Racial, Ethnic, and Gender Differences in Intermarriage, *Journal of Marriage and Family*, vol. 79, no. 2, pp.301–317

Cretser, G. A. & Leon, J. J. (1982). Intermarriage in the U.S., *Marriage & Family Review*, vol. 5, no. 1, pp.3–15

DiMaggio, P. & Mohr, J. (1985). Cultural Capital, Educational Attainment, and Marital Selection, *American Journal of Sociology*, vol. 90, no. 6, pp.1231–1261

Dribe, M., Eriksson, B. & Helgertz, J. (2023). From Sweden to America: Migrant Selection in the Transatlantic Migration, 1890–1910, *European Review of Economic History*, vol. 27, no. 1, pp.24–44

Dribe, M., Hacker, J. D. & Scalone, F. (2018). Becoming American: Intermarriage during the Great Migration to the United States, *The Journal of Interdisciplinary History*, vol. 49, no. 2, pp.189–218

Emigrationsutredningen VII. (1908). Emigrationsutredningen. Bil. 7, Utvandrarnes egna uppgifter : upplysningar inhemtade genom Emigrationsutredningens agenter äfvensom bref från svenskar i Amerika / redigerade av Emigrationsutredningen, [e-journal], Available Online: https://gupea.ub.gu.se/handle/2077/49503 [Accessed 12 April 2023]

Eurenius, A.-M. (2020). A Family Affair: Evidence of Chain Migration during the Mass Emigration from the County of Halland in Sweden to the United States in the 1890s, *Population Studies*, vol. 74, no. 1, pp.103–118

Gans, H. J. (1979). Symbolic Ethnicity: The Future of Ethnic Groups and Cultures in America, *Ethnic and Racial Studies*, vol. 2, no. 1, pp.1–20

Goldman, N., Westoff, C. F. & Hammerslough, C. (1984). Demography of the Marriage Market in the United States, *Population Index*, vol. 50, no. 1, pp.5–25

Gonsoulin, M. & Fu, X. (2010). Intergenerational Assimilation by Intermarriage: Hispanic and Asian Immigrants, *Marriage & Family Review*, vol. 46, no. 4, pp.257–277

Homer, D. (1964). The Rockford Swedish Community, *Journal of the Illinois State Historical Society (1908-1984)*, vol. 57, no. 2, pp.149–155

Jacobs, J. A. & Labov, T. G. (2002). Gender Differentials in Intermarriage Among Sixteen Race and Ethnic Groups, *Sociological Forum*, vol. 17, no. 4, pp.621–646

Kalmijn, M. (1991). Shifting Boundaries: Trends in Religious and Educational Homogamy, *American Sociological Review*, vol. 56, no. 6, pp.786–800

Kalmijn, M. (1998). Intermarriage and Homogamy: Causes, Patterns, Trends, *Annual Review of Sociology*, vol. 24, no. 1, pp.395–421

Karadja, M. & Prawitz, E. (2019). Exit, Voice, and Political Change: Evidence from Swedish Mass Migration to the United States, *Journal of Political Economy*, vol. 127, no. 4, pp.1864–1925

Kennedy, R. J. R. (1944). Single or Triple Melting-Pot? Intermarriage Trends in New Haven, 1870-1940, *American Journal of Sociology*, vol. 49, no. 4, pp.331–339

Kennedy, R. J. R. (1952). Single or Triple Melting-Pot? Intermarriage in New Haven, 1870-1950, *American Journal of Sociology*, vol. 58, no. 1, pp.56–59

Lieberson, S. & Waters, M. C. (1988). From Many Strands: Ethnic and Racial Groups in Contemporary America, Russell Sage Foundation

Lindmark, S. (1971). Swedish America, 1914-1932: Studies in Ethnicity with Emphasis on Illinois and Minnesota

Long, J. S. & Freese, J. (2006). Regression Models for Categorical Dependent Variables Using Stata, Second Edition, Stata Press

Marcson, S. (1950). A Theory of Intermarriage and Assimilation, *Social Forces*, vol. 29, no. 1, pp.75–78

McCaa, R. (1993). Ethnic Intermarriage and Gender in New York City, *The Journal of Interdisciplinary History*, vol. 24, no. 2, pp.207–231

Merton, R. K. (1941). Intermarriage and the Social Structure: Fact and Theory, *Psychiatry*, vol. 4, no. 3, pp.361–374

Norman, H. (1976). Swedes in North America, in H. Runblom & H. Norman (eds), *From Sweden to America: A History of the Migration*, Minneapolis, Minn: Univ. of Minnesota Press [u.a.], pp.228–90

Pagnini, D. L. & Morgan, S. P. (1990). Intermarriage and Social Distance Among U.S. Immigrants at the Turn of the Century, *American Journal of Sociology*, vol. 96, no. 2, pp.405–432

Peach, C. (1980a). ETHNIC SEGREGATION AND INTERMARRIAGE∗, *Annals of the Association of American Geographers*, vol. 70, no. 3, pp.371–381

Peach, C. (1980b). Which Triple Melting Pot? A Re-examination of Ethnic Intermarriage in New Haven, 1900–1950, *Ethnic and Racial Studies*, vol. 3, no. 1, pp.1–16

Pforr, K. (2014). Femlogit—Implementation of the Multinomial Logit Model with Fixed Effects, *Stata Journal*, vol. 14, pp.847–862

Porterfield, E. (1982). Black-American Intermarriage in the United States, *Marriage & Family Review*, vol. 5, no. 1, pp.17–34

Qian, Z. & Lichter, D. T. (2007). Social Boundaries and Marital Assimilation: Interpreting Trends in Racial and Ethnic Intermarriage, *American Sociological Review*, vol. 72, no. 1, pp.68–94

Rosenfeld, M. J. (2008). Racial, Educational and Religious Endogamy in the United States: A Comparative Historical Perspective, *Social Forces*, vol. 87, no. 1, pp.1–31

Ruggles, S., Flood, S., Sobek, M., Brockman, D., Cooper, Richards, S. & Schouwiler, M. (2023). IPUMS USA: Version 13.0, Available Online: https://usa.ipums.org [Accessed 31 March 2023]

Schwartz, C. R. (2013). Trends and Variation in Assortative Mating: Causes and Consequences, *Annual Review of Sociology*, vol. 39, no. 1, pp.451–470

Sommet, N. & Morselli, D. (2017). Keep Calm and Learn Multilevel Logistic Modeling: A Simplified Three-Step Procedure Using Stata, R, Mplus, and SPSS., *International Review of Social Psychology*, vol. 30, pp.203–218

Spörlein, C., Schlueter, E. & van Tubergen, F. (2014). Ethnic Intermarriage in Longitudinal Perspective: Testing Structural and Cultural Explanations in the United States, 1880–2011, *Social Science Research*, vol. 43, pp.1–15

Stevens, G. & Schoen, R. (1988). Linguistic Intermarriage in the United States, *Journal of Marriage and Family*, vol. 50, no. 1, pp.267–279

Wildsmith, E., Gutmann, M. P. & Gratton, B. (2003). Assimilation and Intermarriage for U.S. Immigrant Groups, 1880–1990, *The History of the Family*, vol. 8, no. 4, pp.563–584

Wilson, J. R. & Lorenz, K. A. (2015). Fixed Effects Logistic Regression Model, in J. R. Wilson & K. A. Lorenz (eds), *Modeling Binary Correlated Responses Using SAS, SPSS and R*, [e-book] Cham: Springer International Publishing, pp.225–246,

Available Online: https://doi.org/10.1007/978-3-319-23805-0_11 [Accessed 16 May 2023]

9 Appendix A

Table 6: A collinearity test showing the VIF values for the independent variables and the mean VIF value.

Variable	VIF (including age^2)	VIF (excluding age^2)
Female	3.48	3.48
Marriage market	1.05	1.05
Speaks English	1.05	1.05
Age	**120.47**	1.34
Age^2	**121.34**	
Year in the US	8.07	7.99
Year in the US^2	7.59	7.37
Occupational Score	3.48	3.48
Urban	1.07	1.07
Year dummy	1.09	1.09
Mean VIF	26. 87	

10 Appendix B

Table 7: Comparison in output results between the model inclusion and exclusion of age square

VARIABLES	Model 1 Intermarriage	Model 1[+] Intermarriage
Male	Reference group	
Female	1.09	1.10
	(0.09)	(0.08)
Marriage market	**0.005***	**0.004***
	(0.00)	0.00
English Speaker		
No	Reference category	
Yes	**1.89***	**1.89***
	(0.13)	(0.13)
Age	**0.85***	**0.98***
	(0.03)	0.03
Age^2	**1.00***	
	(0.00)	
Years in the US	**0.80***	**0.80***
	(0.01)	0.01
Years in the US^2	**1.01***	**1.01***
	(0.00)	0.00
Occupational Score		
1-20	Reference category	
21-40	**0.85***	**0.85***
	(0.05)	(0.05)
41-60	**1.37**	**1.38**
	(0.21)	(0.20)
61-80	1.96	1.91
	(1.57)	(1.58)
Missing	**0.45***	**0.45***
	(0.04)	(0.04)
Rural	Reference group	
Urban	1.00	1.00
	(0.05)	(0.05)
1900	Reference group	
1910	**1.62***	**1.62***
	(0.07)	(0.04

Observations	17,451	17,451
Number of counties	112	112
Log-likelihood	-7774.76	-7782.85
Overall p	0.00	0.00

11   Appendix C

Table 8: Wald test on the inclusion of interactive variables.

(1) [intermarriage]1.female = 0 (2) [intermarriage]1.female#c.marriagemarket = 0	(1) [intermarriage]1.female = 0 (2) [intermarriage]1.SE = 0 (3) [intermarriage]1.SE#1.female = 0
chi2(2) = 10.88	chi2(3) = 130.51
Prob > chi2 = 0.0044	Prob > chi2 = 0.0000

12 Appendix D

Table 9: Results from Akaike's information criterion (AIC)

Model	AIC
Model 1	15575.5
Model 2	15567.7
Model 3	15506.4

13 Appendix E

Table 10: Robustness check

VARIABLES	Model 1 intermarriage	Model 2 intermarriage	Model 3 intermarriage	Model 4 intermarriage
Male			Reference group	
Female	1.09	1.19**	0.38***	1.27**
	(0.09)	(0.10)	(0.06)	(0.13)
Marriage market	0.005***	0.01***	0.005***	
	(0.00)	(0.01)	(0.00)	
Female*Marriage market		0.13***		
		(0.08)		
Low (<0.011)				Reference group
Medium-low (0.011-0.027)				0.46***
				(0.04)
Medium-high (0.027-0.067)				0.40***
				(0.03)
High (>0.067)				0.30***
				(0.03)
Female*Medium-low (0.011-0.027)				0.95
				(0.10)
Female*Medium-high (0.027- 0.067)				0.92
				(0.10)
Female*High (>0.067)				0.68***
				(0.08)
English Speaker				
No			Reference category	
Yes	1.89***	1.89***	1.09	1.84***
	(0.13)	(0.13)	(0.10)	(0.13)
Female*English speaker			3.22***	
			(0.46)	
Age	0.85***	0.85***	0.85***	0.98***
	(0.03)	(0.03)	(0.03)	(0.00)
Age^2	1.00***	1.00***	1.00***	1.00***
	(0.00)	(0.00)	(0.00)	(0.00)

Years in the US	**0.80*****	**0.80*****	**0.80*****	**0.80*****
	(0.01)	(0.01)	(0.01)	(0.01)
Years in the US^2	**1.01*****	**1.01*****	**1.01*****	**1.01*****
	(0.00)	(0.00)	(0.00)	(0.00)
Occupational Score				
1-20				
21-40	**0.85*****	**0.84*****	**0.87****	**0.86****
	(0.05)	(0.05)	(0.05)	(0.05)
41-60	**1.37****	**1.37****	**1.44****	**1.36***
	(0.21)	(0.20)	(0.22)	(0.21)
61-80	1.96	1.93	2.10	1.93
	(1.57)	(1.55)	(1.69)	(1.58)
Missing	**0.45*****	**0.45*****	**0.46*****	**0.44*****
	(0.04)	(0.04)	(0.04)	(0.04)
Rural		Reference group		
Urban	1.00	1.00	1.00	0.99
	(0.05)	(0.05)	(0.05)	(0.05)
1900		Reference group		
1910	**1.62*****	**1.62*****	**1.60*****	**1.65*****
	(0.07)	(0.07)	(0.07)	(0.07)
Observations	17,451	17,451	17,451	17,451
Number of counties	112	112	112	112
Log-likelihood	-7774.76	-7769.85	-7739.22	-7609.90
Overall p	0.00	0.00	0.00	0.00

Notes: Model 1-3 represent the regression output from the results, Model 4$^+$ is included as an additional robustness test of the models. Standard errors are presented in parentheses and *** p<0.01, ** p<0.05, * p<0.1.