ORIGINAL PAPER

Ethnic assortative matching in marriage and family outcomes: evidence from the mass migration to the US during 1900–1930

Ho-Po Crystal Wong[1]

Abstract Positive assortative matching in terms of traits such as ethnicity and race has been prevalent in marital formation. One possible explanation for this is that spouses in endogamous marriages possess complementary skills and tastes that increase marital surplus. This paper aims to estimate the effects of ethnic assortative matching on a variety of household outcomes by using the exogenous variation in immigrant flows in the USA during the period 1900–1930 to disentangle the selection effect of partners. The major finding is that the complementarities in home production from same ethnic marriage enhances investment in household public goods such as childrearing and home ownership and reduces the market labor supply of wives. The OLS estimates of the sizes of these effects appear to be substantially biased downward, indicating positive selection into intermarriage in terms of unobservable traits that increase marital surplus.

1 Introduction and literature review

Mate selection is not a random process and it is a well-documented stylized fact that in marriage, individuals tend to sort on traits such as ethnicity and race. This paper aims to

---

✉ Ho-Po Crystal Wong
  howong@mail.wvu.edu

[1] Department of Economics, West Virginia University, 1601 University Ave., Morgantown, WV 26506, USA

🖄 Springer

understand the economic explanation behind such human regularity. Conceivably same-ethnic marriage could be easier to sustain due to the common background of mates, which could play an important role in the labor supply, fertility, and other investments in household public goods of families. Meanwhile, interethnic marriage rates have often been conceived as a proxy for the degree of assimilation of ethnic groups in the host country (see Pagnini and Morgan 1990; Qian & Lichter 2007; Furtado and Trejo 2013). Understanding the family outcomes of ethnically endoga-mous marriages would provide important insights into the economic factors at play in the process of social assimilation as well as adaptation of immigrants to the host country and might also be associated with immigrations' ethnic self-identification (e.g., Duncan and Trejo 2005; Zimmermann et al. 2007; Constant and Zimmermann 2008; Chiswick 2009; Casey and Dustmann 2010).

As proposed in the seminal work of Becker (1973), one primary economic reason for marital sorting is the complementarity or substitution effects for traits of couples in marital production. Arguably same ethnicity is one important complement in marital production. Coordination in household production and decision making can be easier when spouses speak the same language and share a similar cultural background (Kalmijn et al. 2005). The common ground can be in form of similar tastes for food and interests, which would enhance the complementarity in leisure as well as the utility from public goods sharing and consequently the gain to marriage. Divergent cultural norms of ethnically mixed couples could hinder coordination and might lead to more family conflicts and disagreement. This paper seeks to estimate the economic effects of marital sorting by ethnicity.

The existing literature on intermarriage and endogamous marriage by race or ethnicity has been mostly descriptive. These studies primarily focus on examining the factors related to endogamous marriage over time and the extent to which ethnic groups differ in terms of the trends in endogamy and intermarriage. One major lesson we have learnt from these works is that individuals have a tendency to match themselves by ethnicity throughout human history and such trends are affected by factors such as education attainment of different ethnic groups, their sizes in the population, and the composition of local marriage market over time (Kalmijn 1998; Gilbertson et al. 1996; Qian 1997). Yet, the question of the extent to which ethnic assortative matching affects household outcomes has remained largely unexplored. Given such human regularity, examining the causal impact of ethnic endogamous marriage on household outcomes is important in understanding why endogamy persists in societies.

Empirical studies in the sociology literature show that intermarried couples tend to be matched positively by the educational dimension (Qian 1997; Chiswick and Houseworth 2011). Higher levels of education could enhance cultural assimilation and facilitate the interaction with people of different ethnic backgrounds (Pagnini and Morgan 1990; Qian et al. 2001). In addition to education, spouse selections are determined by many characteristics not observed by the researcher and yet might play an important role in household outcomes. One can easily imagine that spouses with different ethnic backgrounds might form a union because they possess some attributes attractive to each other. Imaginably, they might both be physically attractive or possess personality traits favorable to marriage (Lundberg 2010). In contrast, same ethnic couples might be less selective in these unobservable traits as they have additional

benefits from "mating of likes" compared to their intermarried counterparts. This implies positive selection into intermarriage.

Alternatively, for mate selection in ethnic groups with skewed sex-ratios in the marriage market, the same-ethnic mates in excess supply might have to accept less desirable mates outside their own ethnic groups.[1] This is particularly true in the period under study due to male-dominated migration. This force would lead to negative selection into inter-marriage. Theoretically, therefore, it is not clear which of the above two forces dominates in partner selection. One additional contribution of this paper is that it addresses empirically how mates in an ethnically diverse social environment select themselves into intermarriage.

While there is a solid theoretical foundation for assortative matching, empirically, it is a difficult task to estimate the causal effect of assortative matching on marital outcomes (see for instance Adserà and Ferrer 2014). Without proper instruments, it is extremely difficult to isolate this selection effect and disentangle the causal impact of endogamous marriage on family outcomes. The selection at play in intermarriage could give rise to substantial bias in estimating the effect of endogamous marriage using simple ordinary least squares (OLS) models.[2]

Most economic studies related to endogamous marriage take the preference for same-ethnic marriage as given and derive economic analyses based on this premise. Little attention has been directed toward understanding the favorable family outcomes resulted from endogamy itself. For instance, Angrist (2002) addressed the question of how sex ratios of ethnic groups affect the marriage and labor market outcomes of immigrants given their preference for endogamous marriage by ethnicity. Lafortune (2013) explored how a gender's scarcity within the same ethnic group affects educational investments based on the empirical fact that a large fraction of second-generation Americans tends to marry within their own ethnicity. This paper contributes to this line of literature by providing an economic basis for such preference in the context of the possible beneficial family outcomes from ethnic assortative matching.

Another strand of literature concentrates on marriages between *the natives and immigrants*. Furtado and Trejo (2013) provided a review on the literature on intermarriage. Most studies found that immigrants that marry natives tend to have better family outcomes than those who marry within their ethnicity, yet the difficulty in interpreting these findings lies in the selection into intermarriage. Meng and Gregory (2005) addressed this mate selection effect and estimated the wage premium for intermarriage. [3] They found that intermarried immigrants earn significantly higher than endogamously married immigrants (i.e., immigrants marrying immigrants). In contrast, Furtado (2009) found that children of immigrants that married natives are more likely to drop out of high school than immigrants that marry other immigrants after selection into intermarriage is accounted for. This paper focuses on the effects of endogamous marriage in

---

[1] I thank one anonymous referee for suggesting this theoretical possibility.

[2] In the same spirit as this paper, Frimmel et al. (2013) estimated the effect of assortative mating on marital stability in Austria by performing Cox proportional hazard models. One dimension of assortative mating they investigated is ethnicity. They found no evidence that a decrease in assortative mating by ethnicity to be associated with an increase in risk of divorce in Austria. But again, in their estimations, the non-random mate selection has not been accounted for, which potentially could confound their results.

[3] The local sex ratio of immigrants was used as an instrument for the probability of intermarriage between natives and immigrants.

ethnicity *among* (*first generation and second generation*) *immigrants* on family outcomes. Limiting the scope of analysis to the immigrants alone provides additional insight into the effect of cross-ethnic marriage not confounded by the factor of native premium.

Research that attempts to understand the reasons behind the vulnerability of intermarriage mostly comes from the sociology literature (Jones 1996; Dalmage 2000; Childs 2005; Bratter and King 2008). The most popular hypothesis is the "exogamy hypothesis," which suggests that the dissimilarity in values and norms from divergent cultures, lack of social support from family and friends, and the social stigma attached to intermarriages are the main reasons for which these marriages are relatively more fragile than their endogamous counterparts (Kalmijn et al. 2005; Dribe and Lundh 2012 and Kulu and Milewski 2008). However, none of these works examine the linkage between family outcomes associated with endogamous marriage and marital stability.

This paper contributes to an understanding of endogamy by investigating its effect on family outcomes after taking into account the selection bias from partner selection. I make use of the exogenous variation in immigration flows during the great migration period of the US from 1900 to 1930 to identify the causal effect of same-ethnic marriage on a variety of household outcomes including home ownership, childbearing as well as wives' labor supply. The drastic variation in immigration flows was primarily driven by warfare in Europe and changes in immigration policy in the US. Although cultural proximity and social distance between ethnic groups are important considerations in social assimilation, this paper primarily investigates the average effect of endogamous marriage, given the differences across ethnic groups.[4]

I find that same-ethnic marriage generates profound effects on investment in household public goods in the form of childrearing and home ownership and might have enhanced household specialization, as indicated by the lower labor force participation of endogamously married women. This is consistent with Becker's (1973) marital surplus hypothesis. These findings provide a possible economic explanation why endogamous marriage tends to be more stable. The estimates by simple OLS regressions appear to substantially underestimate the effect of endogamous marriage, indicating positive selection into intermarriage: Intermarried couples are compensated by other unobservable traits in mating that generate marital surplus. The results highlight the importance in addressing the selection issues in marital formation in estimating the causal effect of assortative matching on household outcomes.

The finding of this paper is relevant to a wide range of sociocultural setting where divergent cultural norms across different ethnic groups prevail. It also contributes to a better understanding of the effect of assortative matching on family structure generally. Traits that enhance complementarity in leisure or consumption are likely to play an increasingly central role in contemporary marriage given the decline in gender-based division of labor in households over recent decades. For instance, positive assortative matching in terms of traits like personality might have similar effects on household outcomes (see Lundberg 2010). The complementary effect of same-ethnicity in

---

[4] Studies relate cultural proximity of couples to marital outcomes can be found in Kalmijn et al. (2005) and Dribe and Lundh (2012).

marriage conceivably is not specific to marriage in a historical setting but is also applicable to contemporary ethnically diverse societies. For example, Huang and Zhou (2015) studied the relationship between inter-marriage and family outcomes by employing a quasi-natural experiment of a relaxation of the one-child policy that increases inter-ethnic marriage in China after 1984. They found that inter-ethnic couples are associated with poorer living conditions, low-skilled occupations, and worse education outcomes of children.

2 Theoretical framework

Under Becker (1973)'s assortative mating framework, holding other things equal, mating of likes is implied if such pairings maximize marital output. Suppose that the marital output function takes the form: $f(x, y)$ where x and y stand for a trait of men and women and let it be ethnicity and assume the function is twice differentiable. For simplicity, consider that there are only two ethnic groups: a and b. If a type x_a man and a type y_a woman match with each other and a type x_b man and a type y_b woman match with each other, endogamous marriage happens.

Endogamous marriage is optimal if:

$$\mathrm{f}(x_a, \ y_a) + \mathrm{f}(x_b, \ y_b) > \mathrm{f}(x_b, \ y_a) + \mathrm{f}(x_a, \ y_b) \tag{1}$$

This occurs when $\frac{\partial^2 f(x,y)}{\partial x \partial y} > 0$, i.e, when the arguments in the marital production function are complements. Arguably, ethnicity is one form of complements in marital production. Lifestyles of individuals are undoubtedly affected by the cultures they are brought up with. For example, people from the same ethnic group are likely to have more common interests, same religious beliefs, and similar tastes for food. They are more likely to agree on each other in terms of parenting practices. All these would enhance the gain to marriage. Theoretically, the extra marital surplus to marriage in the above is purely driven by the positive (or negative) assortative matching itself and such gain would not occur had spouses been matched differently. Therefore, holding other things constant, any individual in the marriage market would prefer a mate from the same ethnic group.

This framework however does not directly take into account the source of the additional marital surplus that arises out of the complementarity of traits in the marital output function.[5] Conceivably though, the complementarity could be in the form of public good sharing, such as home ownership and demand for children. One could imagine the marital output function to take the following form:

$$Z = \theta_{x_i y_j} f(q, c, \ h) \tag{2}$$

where $\theta_{x_i y_j} > 0$ and $\theta_{x_a y_a} + \theta_{x_b y_b} > \theta_{x_b y_a} + \theta_{x_a y_b}$; q, c, h correspond to home assets, children, and time devoted to home production.

---

[5] Lam (1988) incorporated the joint consumption of household public goods in the assortative matching model in marriage. The traits he focused on were wealth and market wages.

Fixing the level of household inputs at $\overline{q}, \overline{c}, \overline{h}$, we have:

$$\theta_{x_a y_a} f\left(\overline{q}, \overline{c}, \overline{h}\right) + \theta_{x_b y_b} f\left(\overline{q}, \overline{c}, \overline{h}\right) > \theta_{x_a y_b} f\left(\overline{q}, \overline{c}, \overline{h}\right) + \theta_{x_b y_a} f\left(\overline{q}, \overline{c}, \overline{h}\right) \quad (3)$$

Based on the above, same-ethnic couples relative to their intermarried counterparts produce a larger marital output for any given level of household inputs.[6] This would lower the probability of divorce or marriage disruption as a result of the higher marital surplus in same-ethnic marriages for any given level of inputs. This is consistent with empirical studies suggesting that endogamous marriages are more stable than intermarriages: racially and ethnically mixed couples tend to experience less stable marriage—their divorce rates are higher and the duration of marriages tends to be shorter (see Jones 1996; Kalmijn et al. 2005; Bratter and King 2008). Even though formal dissolution of marriages was uncommon in the period under study, for those couples that expect a higher likelihood of unsatisfactory marriage, conceivably, they would behave less cooperatively in time allocation and investment in household public goods (Lundberg and Pollak 1993). Marital stability could then induce a second order effect that further increases marital surplus through enhancement in investment in household public goods and intensified household specialization of labor, as same-ethnic couples are more committed to their marriage (see Johnson and Skinner 1986; Parkman 1992; England and Folbre 1999; Gray 1998; Stevenson 2007, 2008).

3 Econometric model

3.1 The three-step IV method

3.1.1 The ordinary least squares (OLS) regression

In order to estimate the effects of same-ethnic marriage, I consider the following regression model:

$$Y_{ijst} = X_i'\boldsymbol{\beta} + \psi E_i + \alpha_s + \sigma_t + \mu_j + _i \quad (4)$$

where Y_{ijst} is some outcome variable including home ownership and wives' labor supply and children in household i in ethnic group j, residing in state s in census year t. X_i is a vector of control variables including the age and age squared of the household head and his wife; dummies for literacy level, years in the US for the husband and the wife as well as whether each of them are second generation immigrants: a dummy

---

[6] The difference between inter-married and endogamous married couples is that, the marginal benefit from allocating resources to each unit of public good provision is higher for same-ethnic couples due to complementarity of same ethnicity in home production. So for any given amount of resources, same-ethnic couples would allocate more resources on investing in household public goods rather than on their own private consumption relative to intermarried couples. Conceptually, the resources could be in the form of time spent on housework, financial resources devoted to home assets and children. Arguably for mates that are otherwise identical, same-ethnicity matching is also more efficient because the total marital surplus generated by the households would be higher than if mates were to match across ethnicity.

variable that indicates the grandparents reside in household i and a dummy variable that indicates urban residence. E_i is an endogenous dummy variable that takes the value one if the marriage of household i is endogamous and zero otherwise. γ_s, σ_t and μ_j capture the state, year, and ethnicity fixed effects, respectively, and $_i$ is an i.i.d. noise term. Therefore, to identify the effect of endogamous marriage of family outcomes, E_i has to be instrumented for.

3.1.2 The three-stage procedure

The endogeneity of E_i is addressed by performing the following three-stage procedure. In the first stage, a Probit model is estimated using:

$$P\left(E_i = 1 \middle| X_i, \text{net}_{j,t}\text{Ethnic}_j', \text{State}_s, \text{Year}_t, \text{Ethnic}_j\right)$$
$$= G_i\left(X_i, \text{net}_{j,t}\text{Ethnic}_j', \text{State}_s, \text{Year}_t, \text{Ethnic}_j; \gamma\right) \tag{5}$$

where $\text{net}_{j,t}$ is the number of female immigrants minus the number of male immigrants over the male population of ethnic group j in year t.[7] This term aims to capture the "net supply" of female immigrants in a particular ethnic group to the marriage market as a percentage of the ethnic male foreign stock aged 20–45.[8] Therefore, it is the net flow of female immigrants as a percentage of the male stock. Ethnic is a vector of dummies that equal to 1 for the ethnic group to which the household head belongs and zero otherwise; θ_{1s}, ω_{1t} and ρ_{1j} capture the state, year and ethnic group fixed effects, respectively, in the first stage regression and φ_{1i} is the noise term.

The construction of this instrument set is based on the fact that migration during the period under study was highly male-dominated. The numerator captures the compositional effect of the flow of immigrants while the denominator captures the size of the ethnic marriage market. An increase in female immigrants as a proportion of the male foreign stock in the same ethnic group should increase the likelihood for a man to marry endogamously.

To ensure the estimates are insensitive to the construction of the exclusive instruments based on the immigration flows, I construct an alternative set of instruments that are similar to (5). The first stage regression is alternatively given by:

$$P\left(E_i = 1 \middle| X_i, \text{Immigrant}_{j,t}\text{Ethnic}_j', \text{State}_s, \text{Year}_t, \text{Ethnic}_j\right)$$
$$= G_i\left(X_i, \text{Immigrant}_{j,t}\text{Ethnic}_j', \text{State}_s, \text{Year}_t, \text{Ethnic}_j; \gamma\right) \tag{6}$$

where $\text{Immigrant}_{j,t}$ denotes the total number of immigrants divided by the total population aged 15–40 for females and 20–45 for males, of ethnic group j in year t.

---

[7] The immigration flows in the numerators of the instruments are calculated by female and male immigrants of all ages. The reason for not confining the calculation to a certain age group is because the data on the immigrants by age were not very consistently grouped in the sample years.

[8] The foreign born plus the native born of foreign or mixed parentage are referred to as the foreign stock by the US Census Bureau.

θ_{2s}, ω_{2t} and ρ_{2j} capture the state, year, and ethnic group fixed effects in specification 2. ϕ_{2i} is the noise term.

Instead of focusing on the supply of women in the marriage market, this instrument set captures the flow of immigrants as a percentage of the foreign stock in a given ethnic group. A larger immigrant flow as a percentage of the foreign stock in the country should increase the supply of mates overall. However, if sex-biased migration prevails, as was the case in the period under study, it would actually increase competition in the marriage market for same-ethnic mates. The impact of the immigrant flows on ethnically endogamous marriages thus depends on the age, sex, and marital status composition of the local ethnic groups as well as those the arriving immigrants. Since detailed information on the demographic composition of the immigrants is not available, the interaction of the instruments with the ethnic group dummies allows for group- specific effects for these instruments.

The fitted value of G, denoted as $\hat{G}$ is obtained estimating the above Probit estimation. In the second stage, an OLS model is performed by regressing E on $\hat{G}$ and X, State, Year, Ethnic . The fitted value of E_i in the second stage is denoted as $\hat{E}$. In the third stage, the two-stage least squares (2SLS) of equation (4) is performed using $\hat{E}$, X, State, Year, Ethnic as instruments for E. One major advantage of using this three-stage procedure over the conventional two-stage least squares (2SLS) model in estimating the effect of endogamous marriage, which is a binary variable, is that the estimates are more efficient when the probit first stage model can better approximate the first stage conditional expectation function (see Newey 1990; Wooldridge 2001; Angrist and Pischke 2009).[9]

Note that this IV method captures the local average treatment effect (LATE) of endogamous marriage on the subpopulation whose marital outcomes are affected by the variation in the immigration flows.[10] This means that the estimates are not necessarily applicable to the immigrants that were married prior to their arrivals in the US. Yet this is a subpopulation of great interests, as it sheds light on immigrants' spousal choice in an ethnically diverse social setting and its economic effects on families, which is important in understanding the process of assimilation.

3.2 The instruments

The justification for the validity of these instruments is that the variation in the immigrant flows over decades and across ethnic groups is primarily driven by conditions outside the US and a set of immigration policies uncorrelated with the family outcomes of the immigrants. For instance, the large-scale migration of Italians to the US during this period is mainly a result of the poor land management and disease in South Italy (see Foerster 1924). Secondly, the immigrants from Europe dramatically decreased when the US entered the first World War (see Figure 1), which is an exogenous political event. Finally, a variety of Immigration Acts set forth in the 1920s imposing quotas on immigrants based on their country of origin. For example, the first per centum immigration quota law went into effect in 1921. With an aim to

---

[9] See Wooldridge (2002; 623) for further discussion on the advantages of this Probit IV estimator and on the procedure.
[10] The local average treatment effect of the IV method was first discussed by Imbens and Angrist (1994).
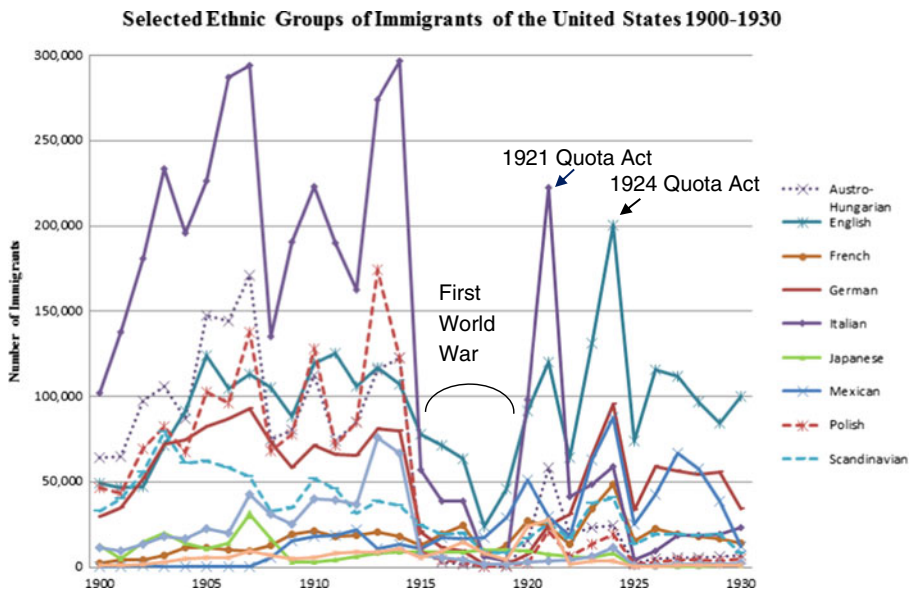
Fig. 1 Immigration flows of great migration in the United States 1900–1930. Note: Data are from Annual Report of the Commissioner General of Immigration, Department of Labor (1900, 1930)

reduce the number of immigrants from Southern and Eastern Europe, the law restricted the admission of aliens in any one year to 3 % of the number of foreign-born persons of each nationality residing in the US in 1910, as determined by the 1910 census in which most of the immigrants came from Northern and Western Europe (Table 1). The 1924 National Origins Act further reduced the annual quota of admissible aliens from 3 to 2 % (see Bloch 1929).[11] Also, the quota limit law imposed a more severe restriction on immigrants from countries of South and East Europe and East Asia, but less on North and West Europe (see the Annual Report of the Commissioner General of Immigration in the fiscal year ended June 30, 1924). Immigrants from the Western Hemisphere such as those born in Mexico were exempted from the quota laws (Table 2).

Figure 1 presents the immigration flows by selected ethnicities in the US during 1900–1930. Overall, the number of immigrant arrivals increased sharply at the beginning of the First World War and dropped dramatically subsequently. The influx of European immigrants resumed to a normal level after the end of First World War but was artificially suppressed in the mid-1920s due to the passage of a variety of quota acts on immigration.[12] The immigration trends of the Mexican and Japanese were less affected by the First World War and the quota acts, which aim to limit the number of immigrants in mid-1920s from Europe, particularly Eastern Europe.

As shown clearly in Fig. 1, the two quota acts drastically reduced the number of Italian, English, and Spanish immigrants subsequent to the passage of these two quota acts. The effects of these restrictive immigration legislation were most strongly felt a

---

[11] Also the quota limit law put a more severe restriction of immigrants from countries of South and East Europe and East Asia, but less on North and West Europe (see the Annual Report of the Commissioner General of Immigration in the fiscal year ended June 30, 1924).

[12] Angrist (2002) also used this exogenous variation in immigrant flows to instrument for the sex ratios of ethnic groups.

Table 1 Number of immigration and population during 1900–1930 by ethnicity

Ethnicity	1900–1909		1910–1919						1920–1929			
	Number of new immigrants	Sex ratio of new immigrants	Foreign stock in 1930	New immigrants as a percent of foreign stock	Foreign stock in 1910	New immigrants as a percent of foreign stock	Number of new immigrants	Sex ratio of new immigrants	Number of new immigrants	Sex ratio of new immigrants	Foreign Stock in 1920	New immigrants as a percent of foreign stock
Dutch	72,786	1.897	187,468	30.52	129,526	56.19	57,209	1.325	88,328	1.735	169,052	52.25
English	841,028	1.266	3,026,469	35.99	3,968,168	21.19	1,089,374	1.199	857,327	1.179	3,382,234	25.35
Finnish	129,941	1.932	150,865	11.96	106,023	122.6	18,038	0.843	75,600	1.634	130,161	58.08
Former Austro-Hungarian	1,037,390	2.948	1,336,989	12.31	1,178,532	88.02	164525	1.022	525,236	1.957	1,300,004	40.40
French	92,398	1.411	198,713	120.9	186,039	49.67	240,260	1.335	172,436	1.311	239,972	71.86
German	656,363	1.475	2,955,749	16.33	3,812,805	17.21	482,810	1.252	409,725	1.314	3,052,458	13.42
Greek	175,432	21.01	149,225	46.43	70,671	248.2	69,280	1.467	263,560	7.185	145,482	181.2
Italian	1,982,418	3.744	1,949,808	27.65	1,002,132	197.8	539,178	1.522	1,290,899	2.602	1,371,958	94.09
Japanese	142,536	5.493	102,236	39.33	109,900	129.7	40,207	0.780	77,257	0.709	112,641	68.59
Mexican	23,991	1.994	525,518	91.49	190,740	12.58	480,804	2.307	173,663	1.503	343,316	50.58
Polish	792,250	2.268	1,472,592	5.41	867,828	91.29	79,706	0.894	600,055	1.750	1,222,457	49.09
Portuguese	63,144	1.454	136,800	33.93	67,106	94.10	46,421	2.410	78,413	1.648	113,504	69.08
Romanian	68,409	10.668	127,534	10.55	37,293	183.4	13,454	1.030	68,279	4.416	88,501	77.15
Russian	328,233	2.967	1,543,034	2.82	882,490	37.19	43,590	1.155	419,308	2.739	1,403,129	29.88
Scandinavian	511,020	1.603	1,541,554	14.80	1,362,615	37.50	228,185	1.852	284,440	1.717	1,463,723	24.76
Spanish	44,218	4.731	57,533	111.6	23,281	189.9	64,201	4.992	85,197	4.633	48,697	175.0
Turkish	11,643	26.08	45,629	3.52	59,144	19.69	1,606	1.365	9,230	16.28	17,759	51.97

Data on the number of new immigrants are from Annual Report of the Commissioner General of Immigration, Department of Labor (1910, 1929). Ethnic populations are estimated by 1910, 1920, and 1930 Census of Population, Public Use Microdata, 1 % sample. The number of immigrants consists of immigrants of all ages. Foreign stock consists of males aged 20–45 and females aged 15–40 in each ethnic group

Table 2 Single immigrants during 1910–1919 and 1920–1929 by ethnicity and gender

Ethnicity	1910–1919			1920–1929		
	Single male new immigrants	Single female new immigrants	Sex ratio of single new immigrants	Single male new immigrants	Single female new immigrants	Sex ratio of single new immigrants
Dutch	37,453	17,127	2.19	20,381	12,767	1.6
English	108,573	99,208	1.09	141,665	133,666	1.06
Finnish	34,837	21,729	1.6	5621	6892	0.82
Former Austro-Hungarian	34,044	33,008	1.03	10,157	10,321	0.98
French	63,452	44,361	1.43	93,348	66,365	1.41
German	153,054	110,539	1.38	192,189	152,699	1.26
Greek	151,881	19,040	7.98	22,334	14,672	1.52
Italian	87,643	28,422	3.08	31,847	15,430	2.06
Japanese	18,722	4824	3.88	9540	3998	2.39
Mexican	64,360	33,819	1.9	193,826	73,906	2.62
Polish	231,868	159,075	1.46	21,632	24,990	0.87
Portuguese	28,304	18,448	1.53	19,807	7493	2.64
Romanian	14,626	4788	3.05	4227	3328	1.27
Russian	69,963	14,604	4.79	10,132	5504	1.84
Scandinavian	150,925	81,853	1.84	118,594	60,761	1.95
Spanish	45,217	8273	5.47	33,107	5828	5.68
Turkish	5294	271	19.54	663	254	2.61

The sex-ratio is defined as the number of male individuals/ the number of female individuals. Notes: Data are from Annual Report of the Commissioner General of Immigration, Department of Labor [1910, 1929]

year after its implementation in 1920 and 1924. These changes in immigration law provide excellent exogenous variation of the immigrant flows to the US, which no doubt produced significant effects on the local marriage market condition and the occurrence of endogamous marriages (Table 3).

This nationwide variation in immigration flows is also adopted by Angrist (2002). He provided evidence that regional variation could lead to biases in the estimates. The reason is that locational choice *within* the US is likely to be responsive to local economic conditions, which could vary across states.[13] Another concern using regional variation is that the state of residence of the immigrants reported in Census does not necessarily match with the state where the couples met. Bodnar (1985; 175) pointed out that newcomers during the great migration period were overall very mobile, "spatially they were constantly on the move as were most residents of urban-industrial America in the nineteenth and early twentieth centuries." One well-known pattern is that most new

---

[13] Chiappori et al. (2002) also pointed out the potential problem of reverse causality in the relationship between regional sex ratios and economic conditions. States that specialize in industries dominated by males will attract more men through migration, and thus, driving up the sex-ratio in these states. The same reasoning can be applied to the inter-state locational choice of immigrants.

Table 3 Single foreign stock in 1910, 1920 and 1930 by ethnicity and gender

Ethnicity	1910			1920			1930		
	Single male foreign stock aged 20–45	Single female foreign stock aged 15–40	Sex ratio of single foreign stock	Single male foreign stock aged 20–45	Single female foreign stock aged 15–40	Sex ratio of single foreign stock	Single male foreign stock aged 20–45	Single female foreign stock aged 15–40	Sex ratio of single foreign stock
Dutch	24,953	26,070	0.96	31,089	28,067	1.11	31,807	29,894	1.06
English	858,754	972,055	0.88	680,814	779,228	0.87	581,919	628,568	0.93
Finnish	31,613	19,650	1.61	29,199	22,660	1.29	32,434	35,685	0.91
Former Austro-Hungarian	268,230	167,569	1.60	211,195	201,707	1.05	225,953	287,234	0.79
French	43,716	37,312	1.17	39,680	39,508	1.00	30,303	30,380	1.00
German	737,765	807,646	0.91	553,338	579,499	0.95	484,846	452,442	1.07
Greek	51,116	2106	24.27	79,508	4417	18.0	46,000	11,484	4.01
Italian	268,186	100,195	2.68	280,262	171,645	1.63	332,552	397,150	0.84
Japanese	60,384	2402	25.14	24,797	5661	4.38	20,132	13,019	1.55
Mexican	52,939	21,060	2.51	75,363	49,522	1.52	107,194	76,384	1.40
Polish	199,138	138,290	1.44	194,954	183,498	1.06	240,579	33,7291	0.71
Portuguese	14,236	9734	1.46	18,910	15,234	1.24	27,960	19,601	1.43
Romanian	8719	6419	1.36	17,072	12,797	1.33	20,780	27,240	0.76
Russian	171,202	153,748	1.11	246,979	233,437	1.06	279,136	330,816	0.84
Scandinavian	382,348	325,985	1.17	344,970	313,601	1.10	329,183	284,212	1.16
Spanish	8936	3405	2.62	21,209	2,664	7.96	15,238	6938	2.20
Turkish	23,373	3908	5.98	5,963	2,557	2.33	10,289	5032	2.04

The estimates are computed using by 1910, 1920, and 1930 Census of Population, Public Use Microdata, one percent sample

arrivals were heavily concentrated at a few ports and in the eastern seaboard and the North Central States while avoiding the South, with the exception being the Mexicans in the Southwest and Scandinavians in the rural areas of the North Central and Northwestern states (see Hutchinson 1956; 22). These locations were more likely to be the starting points of the immigrants nationwide, and some of them subsequently moved to other locations and became thinly scattered elsewhere.[14]

4 Data

The data for this study come from the 1910, 1920, and 1930 Census IPUMS, 1 % sample. The data contain information regarding the birthplaces of the respondents and their parents as well as the number of years they have resided in the US if they are first generation immigrants. It also provides basic information with respect to the work status of the respondents. The information on the number of children in the household and home ownership is also recorded. The sample includes only observations without missing values in all the variables utilized in the analysis.[15] I match the data Census data to the annual data on the immigration flows of different ethnic groups from 1900–1929, which come from the Annual Report of the Commissioner General of Immigration prepared by the Department of Commerce and Labor. These are administrative records for the number of immigrants from different origins.[16] Details on grouping and defining the ethnicity are provided in Appendix C.

A major reason for using the data from the Annual Report of the Commissioner General of Immigration as opposed to the counts in the Census data based on the years of residence in the USA for the first generation immigrants is that, the latter excludes those immigrants that subsequently returned to their home country. This is a concern because return migration is likely to be endogenous to the local economic conditions and individual economic outcomes. For instance, Bodnar (1985) suggested that quite some immigrants that did not do well in the US returned to their home country.

I define the household head to be the husband based on the traditional gender roles in the family.[17] I restrict my analysis to foreign and second generation households heads aged 20–45 with wives aged 15–40. The mate selection of individuals in these age ranges is more likely to be influenced by the influx of new immigrants. This restriction increases the likelihood of appropriately assigning the immigration flows to marriages at the time of their formation. To isolate the potential effects of assimilation and American nativity on family outcomes (Baker and Benjamin 1997; Meng and

---

[14] Therefore, it is likely that for the majority of the immigrants in the period under study, their marital choice was not confined to the availability of same-ethnic mates within their residing states as reported in the Census data but was affected by the flows of immigrants nationwide, whose arrival locations were heavily concentrated in states where their ethnic groups tended to congregate. However, caution should be taken when using nationwide immigration flows as instruments for local marriage market conditions as the nationwide variation might not affect local marriage market conditions in different geographic location equally.

[15] The results are virtually the same without this treatment.

[16] During 1895–1903, the data was collected by the Bureau of Immigration in the Department of Treasury. It was then transferred to the Bureau of Immigration and Naturalization in the Department of Commerce and Labor between 1906 and 1913 and subsequently became the Bureau of Immigration in the Department of Labor.

[17] The results are very similar using the wife as the household head.

Gregory 2005), which might confound the sorting effect of immigrants, I exclude natives from my sample, who are defined to be individuals that were born in the US with American born parents). Also, practically it is impossible to identify the ethnicity of natives in the data. All the regressions and summary statistics are weighted by the household weight.

Table 4 displays the summary statistics by generation of immigrants. Overall, the second-generation population is associated with higher literacy and home ownership rates but they tend to have fewer children. Tables 5 and 6 illustrates the overall decline in endogamous marriage over time. These trends suggest that it is highly possible that the First World War and the quota act affected the marriage market in the US through an exogenous reduction in the supply of foreign mates that could potentially match with the immigrants in the US.

One drawback in using these historical data is that the demographic variables in Census IPUMS in the period under study are rather limited. The dependent variables in this study consist of household public goods variables namely home ownership and households with at least one child and at least two children, respectively, as well as wives' labor supply. Several measures are used to capture the socioeconomic characteristics of the households in the sample. The quadratics in ages for both spouses is used

Table 4 Summary statistics by generation of immigrants

	Husbands aged 20–45		Wives aged 15–40	
	Foreign born	Second generation	Foreign born	Second generation
Age	35.0	33.7	31.0	30.1
	(5.77)	(5.95)	(5.59)	(5.76)
Number of children	2.46	1.89	2.48	1.94
	(1.91)	(1.70)	(1.93)	(1.72)
Number of children >=1	0.847	0.783	0.849	0.653
	(0.360)	(0.412)	(0.358)	(0.476)
Number of children >=2	0.647	0.518	0.782	0.532
	(0.478)	(0.500)	(0.406)	(0.499)
In the labor force	0.99	0.99	0.07	0.05
	(0.10)	(0.10)	(0.26)	(0.22)
Own house	0.30	0.37	0.29	0.38
	(0.46)	(0.48)	(0.45)	(0.48)
Can read and write	0.88	0.99	0.81	0.99
	(0.33)	(0.10)	(0.39)	(0.11)
Grandparents in household	0.05	0.08	0.05	0.08
	(0.22)	(0.27)	(0.21)	(0.27)
Urban dummy	0.79	0.67	0.80	0.69
	(0.41)	(0.47)	(0.40)	(0.46)
N	49,469	27,913	43,911	33,471

Standard errors are in parentheses. The means are weighted by the household weight; Data source: 1910, 1920 and 1930 Census of Population, Public Use Microdata, 1 % sample

Table 5 Summary statistics by endogamous and exogamous marriages

	Endogamous marriages	Exogamous marriages		Endogamous marriages	Exogamous marriages
Age of husband	34.6	34.0	Husband can read and write	0.904	0.980
	(5.86)	(5.91)		(0.295)	(0.139)
Age of wife	30.68	30.28	Wife can read and write	0.867	0.976
	(5.69)	(5.62)		(0.340)	(0.152)
Number of children	2.359	1.798	Husband in the labor force	0.990	0.990
	(1.891)	(1.639)		(0.101)	(0.100)
Number of children >=1	0.836	0.771	Wife in the labor force	0.063	0.064
	(0.371)	(0.420)		(0.243)	(0.246)
Number of children >=2	0.624	0.496	Grandparents in household	0.058	0.075
	(0.485)	(0.500)		(0.234)	(0.263)
Own house	0.328	0.323	Urban dummy	0.742	0.779
	(0.469)	(0.468)		(0.438)	(0.415)
N	63,490	13,892			

Standard errors in parentheses. The means are weighted by the household weight. Data source: 1910, 1920 and 1930 Census of Population, Public Use Microdata, 1 % sample

to capture the life cycle effect on household outcomes. I include a dummy variable that indicates the presence of parents of the spouses. The presence of grandparents in the family conceivably could affect household outcomes, for instance, they can take care of the children in the households and might affect the labor supply of the wife, the demand for children, and home space.

In addition, I take into account whether households' location is urban. Households residing in rural areas tend to have more children and a higher home ownership rate. Wildsmith et al. (2003) suggest that rural push toward intermarriage can be found in the Irish and Italian groups. I also create 17 dummy groups for first-generation household heads and wives having resided in the US for 1 to 5 years, 6 to 10 years and so on as well as a dummy variable that indicates whether the household head and wife second-generation immigrants. Foreign-born men that have resided in the US for a longer period or second-generation immigrants have a higher likelihood to marry outside of their own ethnicity as they have been more exposed to the local culture and are more likely to have a better command of English that could facilitate intermarriage and cultural assimilation. Since education level of individuals is not available in the sample period, the literacy level of spouses is included to account for its effect on household outcomes and labor supply. The literacy skills are coded as a series of mutually exclusive dummy variables that include "illiterate (cannot read or write)," "cannot read, can write," "cannot write, can read," and "literate (read and write)." I also include a dummy variable that indicates that there are grandparents residing in the household.

To account for the time trend and the permanent state effect on the outcome variables, I add year and state dummies in my estimations. The ethnicity of household

Table 6 Percentage of endogamous marriage for first and second generation men marrying non-natives

Ethnicity	First generation						Second generation					
	1910	N	1920	N	1930	N	1910	N	1920	N	1930	N
Dutch	81.3	160	83.2	197	70.1	166	64.3	101	57.9	133	62.7	185
English	91.0	2108	88.9	1554	87.1	1369	73.0	2715	64.3	2101	60.1	2005
Finnish	95.7	207	91.9	247	94.7	150	85.7	7	95.0	20	73.4	79
Former Austro-Hungarian	93.4	2323	88.9	2623	85.5	1650	71.8	273	67.9	563	62.9	1012
French	74.5	94	62.0	121	57.4	94	9.782	92	16.67	138	15.5	116
German	87.0	2346	74.6	1113	80.6	1063	78.6	3943	69.3	3246	64.0	3201
Greek	68.6	35	75.6	201	73.5	325	–	0	0	4	0	4
Italian	97.3	2148	96.6	3337	94.9	3661	66.3	80	66.8	238	77.4	872
Japanese	97.0	133	99.3	288	99.1	211	–	0	100	2	100	28
Mexican	98.8	255	99.0	628	98.9	927	97.7	86	94.4	72	97.7	173
Polish	96.1	1766	90.9	2802	86.7	2379	82.7	179	73.0	471	75.9	835
Portuguese	96.9	131	97.6	205	92.5	187	62.5	16	73.9	46	73.3	101
Romanian	75.0	84	63.1	225	61.0	231	0	1	28.6	7	29.5	44
Russian	92.5	2189	87.1	3288	79.6	2520	78.7	108	72.0	397	69.5	1062
Scandinavian	89.0	1459	85.1	1090	84.3	807	764	635	71.9	1104	63.3	1402
Spanish	65.0	20	81.5	54	71.3	97	0	6	0	2	41.4	15
Turkish	93.1	72	75.0	32	78.1	114	–	0	–	0	100	1
Total		15,530		18,005		15,951		8242		8544		11,135

Author's calculation using 1910, 1920 and 1930 Census of Population, Public Use Microdata, one percent sample Public Use Microdata, 1 % sample

heads is used to capture for the differences in marital and market behavior that vary across ethnic groups. The analysis comprises seventeen ethnic groups: "Dutch," "English," "Finnish," "Former Austro-Hungarian," "French," "German," "Greek," "Italian," "Japanese," "Mexican," "Polish," "Portuguese," "Romanian," "Russian," "Scandinavian," "Spanish," and "Turkish." The groupings are based on geographic proximity and languages of countries.[18]

The next section presents the estimates of the first-stage and second-stage regressions using alternative econometric models. Section 6 provides robustness checks for alternative specifications and examines whether the results are sensitive to exclusion of ethnic groups that might raise concern for the estimates. Overall, I find that the magnitudes of the effects of endogamous marriage without accounting for selection into inter-marriage are substantially underestimated, and the findings are robust to alternative specifications and samplings. Section 7 concludes.

---

[18] The changes in national boundaries during the First World War permit consistent separation of ethnic groups in the former Austria-Hungary. My results are insensitive to exclusion of the "Former Austro-Hungarian" group.

5 Results and discussion

5.1 Immigration flows and endogamous marriage

Based on the reasoning provided in the theoretical framework, holding all other attributes constant, an individual seeking a mate in the marriage market would prefer a spouse in the same ethnic group over one outside his or her ethnic group. Yet in the marriage market, the availability of same-ethnic mates could be very scarce. Theoretically, individuals therefore search for a same-ethnic mate until the marginal cost of such search exceeds its marginal benefit. A large influx of immigrants that are dominated by one sex for any ethnicity would increase the search cost for getting a same ethnic mate. This would increase the likelihood of marriage across ethnic lines.

Table 7 presents the first stage results in OLS using different sets of instruments.[19] Columns (1) and (2) provide estimates using the instruments in equations (5) and (6), respectively. I denote the two sets of instruments as IV(1) and IV(2). I also provide estimates using the ethnic sex-ratios of the immigrant flows in column (3).[20] Column 4 applies IV(1) using single immigrants and single foreign stocks only. The first-stage regressions suggest that the F-statistics for all the instrument sets are well above the conventional threshold for weak instruments. It is noteworthy that the effect of the immigration flow as a percentage of the local ethnic population (IV(2)) on endogamous marriage of each group appears to be quite different from the net flow of women as a percentage of the male stock (IV(1)). One conjecture is that the direction of how the influx of immigrants affects endogamous marriage largely depends on the gender composition of immigrants. Wildsmith et al. (2003) also found that states with large proportion of men in the ethnic population led to exogamy for men.

Table 3 displays the number of single immigrants by gender and the sex ratios of singles by nativity during the periods 1910–1919 and 1920–1929.[21] There are more single men than female immigrants migrating to the US during the period under study. So on the net, the number of immigrants as a percentage of the male population imposes a negative effect on endogamous marriage as the new male immigrants compete with the existing immigrants and the second generation for same-ethnic mates. In contrast, the net flow of female immigrants as a percentage of the male population produces a positive effect for the majority of ethnic groups. By comparing the results in columns (1) and (4) in Table 7, we can also observe that the estimated coefficients are larger using IV(1) based on singles only. This suggests that the marriage market of the US was more responsive to the influx of single female immigrants.

---

[19] To examine the validity of the instruments, I obtained the F-statistics in the first stage regression using OLS. But actual estimates are performed by a three-stage procedure as stated in section 3 with the first stage being a Probit model. The estimates of the Probit first stage evaluated at the sample means are very close to the OLS estimates.

[20] The F-statistics in Table 8 for the first stage regression using this ethnic-specific sex ratio IV and in a specification assuming constant effect of sex ratio presented in Table B1 of Appendix B shows that sex ratios of arriving immigrants as instruments are weaker instruments than IV(1) and IV(2), as it only captures the effect of the gender composition of immigrants on endogamous marriage but not the size effect of immigrants on the foreign stock in the marriage market. Yet it is reassuring that the main results are robust to these alternative instrument sets. These instruments are not employed at the same time due to collinearity for some ethnic groups.

[21] This information is unavailable for the period 1900–1909.

Table 7 First stage regression results

Independent Variables	Coefficients of IV(1) OLS (1)		Coefficients of IV(2) OLS (2)		Coefficients of IV(3) OLS (3)		Coefficients of IV(1) OLS (single only) (4)	
Dutch	0.077	(0.135)	−0.150	(0.128)	0.008	(0.062)	−0.019	(0.084)
English Immigrant	0.250	(0.518)	0.130	(0.081)	0.363***	(0.124)	−136.1	(1433.8)
Finnish Immigrant	0.145***	(0.036)	−0.120***	(0.023)	−0.075***	(0.026)	0.152***	(0.058)
Former Austro-Hungarian Immigrant	0.044***	(0.016)	−0.081***	(0.016)	−0.003	(0.007)	3.561	(2.458)
French Immigrant	−0.070	(0.204)	0.067	(0.059)	−0.079	(0.420)	0.004	(0.101)
German Immigrant	−1.255***	(0.304)	0.578**	(0.234)	0.323***	(0.053)	1.129***	(0.228)
Greek Immigrant	0.041***	(0.021)	−0.059**	(0.024)	−0.007*	(0.004)	0.024	(0.025)
Italian Immigrant	0.073***	(0.006)	−0.090***	(0.006)	−0.043***	(0.005)	0.382***	(0.070)
Japanese Immigrant	0.139***	(0.013)	−0.283***	(0.021)	−0.033***	(0.004)	0.326***	(0.049)
Mexican Immigrant	−0.177***	(0.017)	0.192***	(0.015)	0.094***	(0.011)	−0.127***	(0.018)
Polish Immigrant	0.073***	(0.019)	−0.089***	(0.014)	−0.017*	(0.010)	0.070**	(0.034)
Portuguese Immigrant	0.488	(0.328)	−0.170***	(0.048)	0.038	(0.028)	0.235	(0.330)
Romanian Immigrant	−0.010	(0.020)	−0.004	(0.031)	0.005	(0.006)	0.327***	(0.080)
Russian Immigrant	0.070**	(0.035)	−0.143***	(0.034)	−0.002	(0.007)	0.090	(0.064)
Scandinavian Immigrant	0.195	(0.158)	−0.184***	(0.057)	−0.082**	(0.062)	−0.639	(0.647)
Spanish Immigrant	0.200*	(0.109)	−0.114	(0.092)	−0.028	(0.184)	1.045	(1.404)
Turkish Immigrant	0.070	(0.118)	−0.145	(0.170)	0.002	(0.002)	0.072	(0.102)
N	77,403		77,403		77,403		53,631	
F-statistics (OLS) for inclusion of the instruments	27.44		28.64		23.23		10.50	

***variable is statistically significant at 1 % level; **variable is statistically significant at 5 % level; *variable is statistically significant at 10 % level. Robust standard errors are in parentheses

To provide quantitative interpretations for the first stage results, take Italians as an example. The OLS estimates of IV(1) suggest that after controlling for ethnic fixed effects, if the net supply of female Italian immigrants (i.e., female immigrants minus male immigrants) amounts to 20 % of the Italian male foreign stock in the US, endogamous marriage among Italians would increase by 1.5 % points. IV(2) suggests that if the Italian immigrants amounts to 20 % of the Italian foreign stock, endogamy among Italians would be reduced 0.018 due to gender-biased migration towards males, which makes the marriage market more competitive as proportionately more Italian males arrived the US. These effects appear small but note that the ethnic immigration flows were very large relative to the stock, especially during the period 1900–1920 before the series of quota acts was in place. Table 1 shows that for most of the ethnic groups under study, the new immigrants as a fraction of the foreign stocks were well above 20 %. The estimates using immigrant sex ratios as instruments in column (3) support the gender-biased migration argument.[22] When the sex-ratio of arriving Italian immigrants increases by 50 % (for instance from a balanced sex ratio to 1.5), the probability of Italian endogamous marriage would go down by two percent points. Column (4) constructs IV(1) based on single immigrants alone. The magnitude of the effect of net single females as a percentage of the single Italian population is much larger than the estimate using total number of male and female immigrants. A net supply of single Italian females amounting to a 20 % of the total single Italian foreign stock would increase endogamous marriage among Italians by 7.6 % points. This indicates that as expected, single immigrants produce very profound effects on the marriage market for immigrants.

Overall, the marital choice of immigrants responds to the changes in the availability of same-ethnic mates nationwide in an expected fashion, particularly for those ethnic groups in which immigration flows, and their gender composition produce statistically significant effects on endogamous marriage, with Germans and Mexicans as the exception. The likelihood of endogamous marriage for the German immigrant group is statistically not responsive to the change in the immigrant flow as a percentage of the German stock but it responds negatively to the change in net supply of women. Interestingly, Table 3 shows that the sex ratio and immigration flows for Germans follow an exceptional pattern. In the census years 1910 and 1920, the single sex ratios of the German population are below one, which suggests that there was actually a surplus of female German mates in the marriage market during these years. This explains why for the Germans, the probability of endogamous marriage went down when the net supply of female immigrants increased and rose when the immigration

---

[22] Angrist (2002) argued that the sex ratios of immigrants could affect marital outcomes. But given that the divorce rate was very low: 0.9–1.6 during 1910–1930, (Source: Centers for Disease Control, National Center for Health Statistics) and importantly divorce was extremely difficult to obtain and could only be granted by proof of fault (such as insanity, drunkenness, physical abuse, imprisonment for a crime, cruelty of treatment, and desertion), it is more plausible that the effect of sex ratios on marital outcomes operated through its effect on endogamous marriage rather than through its changes in remarriage prospects of spouses in the marriage market. A spouse that was unhappy with a marriage could not obtain a divorce unilaterally and remarry. This made divorce as a threat point unlikely. In particular, couples that entered into marriage in that period would not conceive the possibility of divorce when divorces were fault-based only. Therefore, once couples had entered into marriages, it is hard to imagine that the sex-ratios in the marriage market could alter family outcomes determined within marriage because the marriage market conditions became irrelevant to couples that were already married when the possibility of remarriage and divorce was almost out of reach.

rate went up, as immigration flows were biased towards males. The exceptional response of Germans found in the results of IV(1) and IV(2) could be related to that the most female immigrants from Germany arrived were married so that they might not have contributed to the net supply of single women in the marriage market. And secondly, the German immigrant group might be culturally better compatible with the rest of the old immigrant groups, which could explain for its nonresponsiveness to the instrument set IV(2). The differences in the estimated coefficients across ethnic groups also indicate the importance of modeling the effects of immigration flows as ethnic-specific, as the underlying gender composition of ethnic populations varied. This modeling allows the ethnic marriage markets to respond differently to the changes in immigration flows and their sex composition.[23] As for the Mexicans, the mechanism through which immigration flows affects the marriage market in the US might differ from other ethnic groups possibly because of Mexico's proximity to the US and the difference in the nature of migration of Mexicans.[24]

5.2 Main results

A casual examination of Table 5 indicates that the mating criteria and family outcomes for endogamous and exogamous couples are quite different. Even before taking into account the selection effect, an endogamous household on average has 0.56 more children in the family and is 12.8 % more likely to have at least two children. The wife in an intermarried household is 10.9 % points more likely to be able to read and write and 7.6 % points for the husband in a mixed-ethnic marriage. This is not unexpected given that a higher level of literacy can facilitate communication with other ethnic groups. The above suggests that without properly dealing with the selection problem in marital formation, the estimated effects of endogamous marriage is likely to be biased due to the existence of unobserved attributes the researchers cannot account for such as physical appearance, the respondents' ability to adapt to different cultures and in our study the exact education level of spouses.

Table 8 provides estimates of the effect of endogamous marriage on a variety of household outcomes using the instruments in equations (5) and (6) based on a variety of econometric models. The standard errors of all the estimates are clustered at the ethnicity-census year level. I denote the three-stage procedure as three-stage least square (3SLS). The discussion that follows are primarily based on the results in 3SLS using IV(1), denoted as 3SLS(1). This set of instruments captures both the sex composition and size effect of immigrants. Nonetheless, the direction of the bias is the same across alternative sets of instruments. The OLS and the 2SLS results using IV(1) as instruments are also presented for comparison purpose. Column (1) shows that the OLS regressions substantially underestimate the magnitude of the effect of

---

[23] I also estimate results using an alternative specification that the effects of the instruments to be constant among ethnic groups (i.e., no interaction terms with the ethnic dummies). The results are presented in Appendix B. The estimates are nonetheless not sensitive to this alternative econometric specification.

[24] Massey et al. (2002) suggested that the motivation of migrating to the US among the earliest Mexicans was generally not to relocate permanently but to work temporarily in the US to support families in Mexico (see also Rosenblum et al. 2012). If for example, proportionately more female Mexicans remain in or migrate to the US on a permanent basis when the economic condition in Mexico is bad and increase in the net supply of female Mexicans could reduce the tendency for endogamous marriage among the Mexicans in the US.

Table 8 Estimates of the effects of endogamous marriage

	OLS	2SLS(1)	3SLS(1)	3SLS(2)	3SLS(3)	3SLS(1): single only
Dependent Variables	(1)	(2)	(3)	(4)	(5)	(6)
Home ownership	0.033***	0.216	0.277***	0.284***	0.291***	0.286***
	(0.004)	(0.168)	(0.072)	(0.073)	(0.071)	(0.107)
Wife's labor supply	−0.006***	−0.287***	−0.139***	−0.139***	−0.143***	−0.161***
	(0.002)	(0.099)	(0.048)	(0.048)	(0.048)	(0.061)
Number of children >=1	0.034***	0.185**	0.145**	0.144**	0.148**	0.196**
	(0.005)	(0.075)	(0.067)	(0.067)	(0.067)	(0.088)
Number of children >=2	0.060***	0.286**	0.227**	0.226**	0.218**	0.282*
	(0.007)	(0.115)	(0.109)	(0.110)	(0.109)	(0.147)
N	77,403	77,403	77,403	77,403	77,403	53,631

***variable is statistically significant at 1 % level; **variable is statistically significant at 5 % level; *variable is statistically significant at 10 % level. Robust standard errors clustered at the ethnicity-year level are in parentheses

endogamous marriage except for the effect on husbands' labor supply, which is statistically insignificant across all econometric models.

After accounting for this selection effect into intermarriage, the 3SLS estimate suggests that endogamously married couples are 0.277 more likely to own a house compared to 0.033 using OLS. Endogamously married couples are also more likely to have children relative to their intermarried counterparts. The 3SLS(1) estimates indicate that they are 18.5 and 28.6 % points more likely to have at least one child and at least two children, respectively (as opposed to 3.4 and 6 % points using OLS). The latter is a particularly important cutoff. As shown in the summary statistics in Table 4, the average number of children is 2.4 and 1.8 for endogamous and intermarried couples, respectively. These results are consistent with the complementarity effect of ethnic assortative sorting and indicate positive selection into intermarriage.

In addition, endogamously married wives are 13.9 % points less likely to participate in the labor market (−6 % points by OLS). This is in line with the conjecture that endogamous marriage encourages household specialization of labor, which generates additional marital surplus, holding other things constant. The higher marital surplus could have made endogamous marriage more stable and increased the incentives for wives to specialize in home production.[25]

Column (5) estimates the effects of endogamous marriage using sex-ratios of the arriving immigrants. The estimates are largely in line with using IV(1) and IV(2). Column (6) constructs IV(1) based on single immigrants alone.[26] Ideally, single immigrants would better capture the effect of immigrants on the marriage market

---

[25] I also estimated the effect of endogamous marriage on the labor force participation of husbands. None of the results are statistically significant. This is not surprising given that the labor force participation rate of husbands is 99 % during the sample period.

[26] The instrument set is the number of single female minus the number of single males over the single male population aged 20–45 in ethnic group j in year t.

conditions. However, data on single immigrants are unavailable from 1900–1909; therefore, the sample period is limited to Census years 1920 and 1930 using singles as instruments. The effects of endogamous marriage tend to be larger using the instruments based on single immigrants but reassuringly the sign and the direction of the bias are consistent with the rest of the 3SLS estimates.

What could explain for such sizeable selection effects in intermarriage? One factor is related to the social setting of America during 1900–1930, in which immigrants were highly ethnically diverse. Except the immigrants that were from English speaking countries, most of them did not share a common language. And thus, intermarried couples were naturally highly selective. These large influxes of immigrants were from many different countries and their lifestyles largely differed, not to mention the language barriers that existed among these ethnic groups.

This argument can be further elaborated by considering the following thought experiment: mates are randomly assigned in the marriage market. Those who get intermarried were very unlikely to be able to communicate effectively with each other, share similar food tastes, and have similar parental practices, especially among the first generation immigrants and those whose home countries were culturally very different from America. It is very likely that marriages from these random matches are less stable and generate less marital surplus compared to those that were drawn to marry a mate from the same ethnic group. As a result, these intermarried couples would be less committed to their marriage, which is consistent with the findings of this paper: *intermarried* spouses have a lower probability of home ownership and having children and a higher probability for the wife to participate in the labor market. The existing finding from the sociology literature also suggests that people select positively into intermarriage (for example Qian 1997; Furtado and Theodoropoulos 2011; Chiswick and Houseworth 2011). Some measurable attributes examined include education. These individuals were likely to possess human capital that enables them to better adapt to new cultures and facilitate inter-ethnic communication (Furtado 2012). These skills were likely to be linked with wealth and thus home ownership. The omission of these variables results in sizable downward bias in the effect of endogamous marriage.

To understand how the effect of endogamous marriage might be affected by the possibility of "generation endogamy" (Pagnini and Morgan 1990), I add controls for the generation composition of spouses to equation (4) namely "both first generation," "the head being first generation and the wife being second generation," "the head being second generation and the wife being first generation."[27] Table 9 provides the estimated results using 3SLS(1). The estimates are similar to the original ones even though the compositional dummies appear to capture part of the effects of endogamous marriage. In particular, relative to the couples in which both are second generation immigrants, families comprise spouses that are both first-generation immigrants are less likely to own homes and have children. Also, their wives are more likely to participate in the labor market. These findings are related to that first generation immigrants were in general less wealthy than the second generation immigrants.

---

[27] "Both second generation" is the omitted group.

Table 9 Estimates of the effects of endogamous marriage with generation composition controls

	3SLS(1)			
	Home ownership	Wife's labor supply	Number of children >=1	Number of children >=2
Independent Variable	(1)	(2)	(3)	(4)
Endogamous marriage	0.225***	−0.1242**	0.107*	0.217**
	(0.060)	(0.042)	(0.054)	(0.092)
Head 1st generation; wife 1st generation	−0.207***	0.080***	−0.101***	−0.161***
	(0.017)	(0.016)	(0.018)	(0.027)
Head 1st generation; wife 2nd generation	−0.082***	0.052***	−0.033***	−0.062***
	(0.012)	(0.010)	(0.013)	(0.017)
Head 2nd generation; wife 1st generation	−0.095***	0.016***	−0.050***	−0.106***
	(0.015)	(0.006)	(0.012)	(0.022)
N	77,403	77,403	77,403	77,403

***variable is statistically significant at 1 % level; **variable is statistically significant at 5 % level; *variable is statistically significant at 10 % level. Robust standard errors clustered at the ethnicity-year level are in parentheses

6 Robustness checks

To examine the extent to which the estimates might be sensitive to the groupings and alternative specifications, I perform a variety of robustness checks as presented in Table 10. Specification (1) is the benchmark model using 3SLS(1). Specifications (2) and (3) confine the sample to the household heads being first generation and second generation, respectively. The results show that endogamous marriage positively affects home ownership headed by first-generation immigrants but produces no statistically significant effect for households in which the husbands are second generation immigrants. Endogamous marriage appears to reduce wives' labor supply among households headed by husbands that are second generation immigration only. In addition, same-ethnic marriage increases the probability of having children for households headed by first generation immigrants. The effect is particularly pronounced for the probability of having 2 children or more (0.362 higher relative to the intermarried counterparts). Endogamous marriage appears to produce positive effect on the probability of having children at least one child for households headed by second generation immigrants but not on the probability of having more than two children.

Specification (4) further confines the sample to first generation heads and wives only.[28] Noticeably, it is a very selected sample and thus the results should be interpreted with caution. Endogamous marriage is found to increase the probability of home ownership and having more than 2 children in the households among the first generation immigrants. These effects are more sizeable than the original estimates using the

---

[28] I have not confined the sample to second generation heads and wives only for the reason that the instruments constructed by the variation in first generation immigration flows might not be valid when the sample is limited to matching among second-generation immigration alone.

Table 10 Robustness checks

| | 3SLS(1) | Effect of endogamous marriage | | | | | | | |
| | Benchmark | First generation heads only | Second generation heads only | First generation heads and wives | Household heads age 20–40 only | Household heads age 20–35 only | No English | Excluding Germans | European Only |
Independent Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Home ownership	0.277***	0.274***	−0.034	0.380***	0.216***	0.191***	0.320***	0.384***	0.321***
	(0.072)	(0.079)	(0.126)	(0.128)	(0.060)	(0.066)	(0.082)	(0.092)	(0.082)
Wife's labor supply	−0.139***	−0.065	−0.121*	−0.00002	−0.154***	−0.161***	−0.114**	−0.162***	−0.113***
	(0.048)	(0.041)	(0.065)	(0.052)	(0.045)	(0.043)	(0.047)	(0.050)	(0.046)
Number of children >=1	0.145**	0.116*	0.167***	0.121	0.133**	0.102	0.193**	0.283***	0.086
	(0.067)	(0.058)	(0.062)	(0.074)	(0.068)	(0.086)	(0.079)	(0.093)	(0.073)
Number of children >=2	0.227**	0.362***	0.161	0.452***	0.183*	0.062	0.313**	0.420***	0.100
	(0.109)	(0.097)	(0.131)	(0.104)	(0.107)	(0.115)	(0.123)	(0.158)	(0.095)
N	77,403	49,486	27,878	39,473	64,282	41,609	62,613	58,654	74,558

***variable is statistically significant at 1 % level; **variable is statistically significant at 5 % level; *variable is statistically significant at 10 % level. Robust standard errors clustered at the ethnicity-year level are in parentheses

full sample possibly due to a higher degree of assimilation among the second generation immigrants. However, its effect on wives' labor supply becomes insignificant.

As the partner selection of older men is less likely to be affected by the immigrant flows, specifications (5) and (6) further narrow the sample to households headed by men aged 20–40 and 20–35, respectively. For men aged 20–40, the effects of same-ethnic marriage on home ownership and fertility drop but the estimates remain sizeable and significant. The effect on wives' labor supply is somewhat larger than using the full sample. Further limiting the sample to men aged 20–35 shows a similar response but the effects on fertility become statistically insignificant. These results are not very surprising given that home ownership, fertility, and wealth tend to increase with the age of the household heads.

English (including Irish people) and Germans are considered to be "old immigrants" and shaped the cultures of the US. These two ethnic groups might better resemble the natives and benefit less from same-ethnic marriages compared to other ethnic groups in the sample. English-speaking immigrants in particular, have language advantages, which conceivably would produce better economic outcomes, holding other things equal. Specifications (7) and (8) exclude the English and German ethnic groups from the sample, respectively. The results are largely in line with the above hypothesis. The effect of endogamous marriage of the probability of home ownership strengthens to 0.32 and 0.38 % points (compared to 0.28 using the full sample) when the English and Germans are excluded. The probability of having at least one and two children also gains in size. This offers evidence for the economic advantages of immigrants from countries with a longer migration history to the US and indicates that endogamous marriage could be less crucial to immigrants that are better accustomed to the local cultures.

Specification (9) investigates the effect of endogamous marriage among ethnic groups of European ancestry by excluding the Japanese and Mexican group. Intermarries across Europeans are subject to less cultural difference and so one would expect that the effect of endogamous marriage to be weaker when non-Europeans are excluded. The empirical results partially confirm this hypothesis. A comparison between the estimates in specifications (1) and (9) suggests that inter-marriages that occur among the European ancestry categories produce no significant effect on fertility. European women in endogamy are found to reduce their market labor supply compared to those that marry across ethnic lines (−0.113), but the magnitude is smaller compared with estimates using the full sample (−0.139). The magnitude of the effect is larger for the likelihood to own homes. Taken together, this implies that inter-marriages that possibly involved European and non-European spouses are less complementary in skills that enhance sexual division of labor. This result provides some economic justifications for the strong separation in marriage along the European and non-European dividing line.

7 Concluding remarks

I create instruments based on the exogenous variation of immigrant flows from different countries to the US during the great migration period to estimate the effect of same ethnic marriage on a variety of household outcomes. The results in this paper provide strong evidence that marrying within the same ethnic group alters a number of family outcomes that are favorable to marriage and this relationship is causal.

As proposed by Becker (1973), the mechanism for the linkage between ethnic assortative matching and these favorable family outcomes is that ethnic assortative matching enhances marital surplus, as same ethnicity is one form of complements in marital output. Same-ethnic couples therefore invest more in household public goods such as home ownership and childrearing. Wives who marry endogamously are also found to have lower labor force participation, which is consistent with more intense sexual division of labor within same-ethnic marriages. Such economic effects of ethnic assortative matching have been largely overlooked in the literature. In addition, these effects of endogamy are the most pronounced among the first generation immigrants. This suggests that assimilation makes same-ethnic marriages less crucial to family outcomes, as different ethnic groups in society essentially become more alike.

Conceivably, the findings of this paper can be extended to matching of the likes for traits such as personality, as suggested by Lundberg (2010). After all, the benefits of endogamous marriage mostly come from the fact that these couples share similar cultural backgrounds which facilitate household co-ordination and the complementarity in leisure as well as public goods sharing. Similar effects can be applied to couples that share similar interests and lifestyle.

The results provide important insights for mating of the likes which complement the existing literature on endogamous and inter-ethnic marriage, which has been mostly descriptive. The findings portray a very clear picture that without accounting for the selection problem in mating, the effect of endogamous marriages are heavily underestimated and that the positive effect arising out of matching of the likes are largely present in marriage. Such finding is consistently with the findings in Qian (1997), Chiswick and Houseworth (2011) and Furtado and Theodoropoulos (2011).

There is also evidence that the effects of same-ethnic marriage are stronger among couples that are both first-generation immigrants. This is reasonable since the second-generation immigrants in general are more alike to each other than their first generation counterparts in many attributes such as their ability to speak English and assimilation to the American cultures. Further research with new data possibly from other countries could provide more conclusive evidence as to how same-ethnic marriage affects economic outcomes of the second-generation immigrants and would also add value to our understanding of the process of assimilation. Casey and Dustmann (2010) for instance, found a positive association between ethnic identity and labor market outcomes among second generation male immigrations in Germany.

In addition, the findings of this paper provide a possible economic reason why ethnically mixed couples are more vulnerable to marital disruption (Bratter and King 2008; Jones 1996): holding other attributes of mates constant, same-ethnic couples could generate more marital surplus for the same amount of input into marital production and this will lower the likelihood of marital dissolution. Becker et al. (1977) found that discrepancies between intelligence, social background, religion, or race raise the probability of marital dissolution. They also found a very strong positive effect of racially endogamous marriage on fertility.[29] An interesting implication from the results is that the immigration quota acts of 1924 which gave rise to a dramatic

---

[29] In their study, racially mixed couples tend to have one less child than other couples. Note that in their estimates, they did not account for the selection into intermarriage and so it is likely that the true effect of racially endogamous marriage is larger than their estimates.

increase in intermarriage rates might have partially accounted for the increase in marital breakdown in the US. The increasing trend of marital instability that had been observed in the US might have been in part the cost the society paid for cultural assimilation. Such effect has been over-looked in the literature and deserves further investigation.

Appendix A

Table 11 Mean values of selected outcome variables by ethnicity of household heads

Ethnicity	1910			1920			1930		
	Home ownership	Number of children	Wife's labor supply	Home ownership	Number of children	Wife's labor supply	Home ownership	Number of children	Wife's labor supply
Dutch	0.445	2.34	0.023	0.422	2.25	0.034	0.512	2.19	0.068
	(0.498)	(2.04)	(0.150)	(0.495)	(1.99)	(0.181)	(0.500)	(1.86)	(0.253)
English	0.268	2.19	0.038	0.290	2.02	0.048	0.310	1.77	0.091
	(0.443)	(1.88)	(0.191)	(0.454)	(1.73)	(0.214)	(0.463)	(1.63)	(0.289)
Finnish	0.439	2.51	0.032	0.478	2.19	0.059	0.461	1.65	0.098
	(0.497)	(2.01)	(0.177)	(0.500)	(1.85)	(0.235)	(0.500)	(1.48)	(0.298)
Former Austro-Hungarian	0.248	2.19	0.065	0.322	2.56	0.063	0.400	2.24	0.086
	(0.432)	(1.83)	(0.246)	(0.467)	(1.88)	(0.243)	(0.490)	(1.83)	(0.280)
French	0.333	1.85	0.100	0.370	1.91	0.108	0.353	1.48	0.141
	(0.473)	(1.72)	(0.301)	(0.484)	(1.76)	(0.311)	(0.479)	(1.42)	(0.349)
German	0.412	2.18	0.040	0.439	2.05	0.030	0.412	1.74	0.089
	(0.492)	(1.85)	(0.195)	(0.496)	(1.76)	(0.170)	(0.492)	(1.61)	(0.284)
Greek	0.114	1.56	0.139	0.145	1.81	0.096	0.289	2.25	0.072
	(0.323)	(1.40)	(0.351)	(0.353)	(1.54)	(0.295)	(0.454)	(1.80)	(0.260)
Italian	0.157	2.44	0.077	0.250	2.84	0.071	0.364	2.52	0.074
	(0.364)	(1.94)	(0.267)	(0.433)	(2.05)	(0.256)	(0.481)	(2.00)	(0.262)
Japanese	0.060	1.48	0.418	0.063	1.77	0.216	0.138	2.64	0.159
	(0.239)	(1.41)	(0.495)	(0.243)	(1.70)	(0.412)	(0.346)	(1.91)	(0.366)
Mexican	0.223	2.29	0.099	0.139	2.35	0.077	0.178	2.74	0.587
	(0.416)	(1.99)	(0.299)	(0.346)	(2.01)	(0.267)	(0.383)	(2.15)	(0.235)
Polish	0.226	2.41	0.054	0.306	2.70	0.076	0.437	2.67	0.093
	(0.418)	(1.98)	(0.227)	(0.461)	(1.90)	(0.265)	(0.496)	(1.98)	(0.291)
Portuguese	0.264	2.57	0.095	0.288	2.36	0.180	0.303	2.39	0.206
	(0.442)	(2.15)	(0.294)	(0.454)	(1.94)	(0.385)	(0.461)	(2.14)	(0.405)
Romanian	0.094	2.54	0.023	0.234	2.24	0.039	0.313	1.94	0.091
	(0.294)	(2.00)	(0.151)	(0.424)	(1.83)	(0.194)	(0.464)	(1.49)	(0.288)
Russian	0.180	2.55	0.043	0.231	2.37	0.045	0.290	2.03	0.062
	(0.384)	(1.95)	(0.202)	(0.422)	(1.74)	(0.208)	(0.454)	(1.62)	(0.240)

Table 11 (continued)

	1910				1920				1930	
Scandinavian	0.453	2.11	0.029	0.473	2.02	0.036	0.394	1.77	0.078
	(0.498)	(1.82)	(0.168)	(0.499)	(1.73)	(0.186)	(0.489)	(1.59)	(0.268)
Spanish	0.154	2.00	0.111	0.195	2.06	0.090	0.283	2.02	0.169
	(0.368)	(2.09)	(0.321)	(0.399)	(1.59)	(0.289)	(0.452)	(1.83)	(0.376)
Turkish	0.153	2.22	0.083	0.222	2.27	0.033	0.328	1.95	0.078
	(0.363)	(1.78)	(0.278)	(0.423)	(1.61)	(0.181)	(0.479)	(1.68)	(0.269)
All	0.299	2.26	0.050	0.317	2.35	0.057	0.360	2.14	0.084
	(0.458)	(1.89)	(0.218)	(0.465)	(1.86)	(0.231)	(0.480)	(1.82)	(0.277)

Standard errors in parentheses. The means are weighted by the household weight; Data source: 1910, 1920, and 1930 Census of Population, Public Use Microdata, 1 % sample

Appendix B

Table 12 Regression results assuming constant marriage market effects across ethnic groups

Panel A			
First Stage	Coefficients of Net	Coefficients of IV(2)	Coefficients of IV(3)
	(1)	(2)	(3)
	0.058***	−0.059***	−0.007***
	(0.005)	(0.004)	(0.002)
F-statistics	163.5	176.7	17.80
Panel B			
Second Stage	NL2S(1)		
Dependent variables	(1)	(2)	(3)
Home ownership	0.325***	0.313***	0.332***
	(0.085)	(0.087)	(0.087)
Wife's labor supply	−0.145***	−0.145***	−0.149***
	(0.049)	(0.049)	(0.051)
Number of children >=1	0.148**	0.146**	0.151**
	(0.070)	(0.070)	(0.072)
Number of children >=2	0.228***	0.231**	0.251**
	(0.113)	(0.113)	(0.117)
N	77,403	77,403	77,403

***variable is statistically significant at 1 % level; **variable is statistically significant at 5 % level; *variable is statistically significant at 10 % level. Robust standard errors clustered at the ethnicity-year level are in parentheses

Appendix C

Grouping of ethnicity

Individuals are grouped by their place of birth or the place of birth of their parents in the Census data. If the individual was native born and both his parents are foreigners, he would be categorized into the ethnic group of his mother as second generation.

The following codes of the places of birth in the IPUMS Census data that are used to merge with data on the immigration flows from the Annual Report of the Commissioner General of Immigration by "race or people".

Categorization of ethnicity in IPUMS Census data

Country code in Census	Country	Ethnicity	Country code in census	Country	Ethnicity
200	Mexico	Mexican	434	Italy	Italian
400	Denmark	Scandinavian	436	Portugal	Portuguese
401	Finland	Finnish	438	Spain	Spanish
402	Iceland	Scandinavian	450	Austria	Former Austro-Hungarian
404	Norway	Scandinavian	452	Czechoslovakia	Former Austro-Hungarian
405	Sweden	Scandinavian	453	Germany	German
410	England	English (Anglophone)	454	Hungary	Former Austro-Hungarian
411	Scotland	English (Anglophone)	455	Poland	Polish
412	Wales	English (Anglophone)	456	Romania	Romanian
413	United Kingdom	English (Anglophone)	462	Lithuania	Russian
412	Ireland	English (Anglophone)	465	Russia	Russian
421	France	French	501	Japan	Japanese
425	Netherlands	Dutch	542	Turkey	Turkish
433	Greece	Greek			

Poland was not identified as a country in the 1910 Census but was again identified in 1920 and onwards after World War I. The Polish stock in the US in 1910 is estimated by population with Polish mother tongue. The same approach is used by Census in estimating the 1910 foreign-born population from Poland (See Gibson 2013). Note that the results are insensitive to exclusion of the Polish ethnic group.

Classification for Ethnicity of Immigrants from the Annual Report of the Commissioner General of Immigration

Race or people	Ethnicity
Bohemian and Moravian (Czech)	Former Austro-Hungarian
Bulgarian, Serbian, and Montenegrin	Former Austro-Hungarian
Dalmatian, Bosnian, Herzegovinian	Former Austro-Hungarian
Dutch and Flemish	Dutch

(continued)

English	English (Anglophone)
Finnish	Finnish
French	French
German	German
Greek	Greek
Irish	English (Anglophone)
Italian North	Italian
Italian South	Italian
Japanese	Japanese
Lithuanian	Russian
Magyar	Former Austro-Hungarian
Mexican	Mexican
Polish	Polish
Portuguese	Portuguese
Rumanian	Romanian
Russian	Russian
Ruthenia	Russian
Scandinavian	Scandinavian
Scotch	English (Anglophone)
Spanish	Spanish
Turkish	Turkish
Welsh	English (Anglophone)

References

Adserà A, Ferrer AM (2014) Immigrants and demography: marriage, divorce, and fertility. IZA Discussion Paper DP7982

Angrist JD (2002) How do sex ratios affect marriage and labor markets? Evidence from America's second generation. Q J Econ 117(3):997–1038

Angrist JD, Pischke J (2009) Mostly harmless econometrics: an empiricist's companion. Princeton University Press, Priceton

Baker M, Benjamin D (1997) The role of family in immigrants' labor-market activity: an evaluation of alternative explanations. Am Econ Rev 87(4):705–727

Becker GS (1973) A theory of marriage: part I. J Polit Econ 81(4):813–846

Becker GS, Landes EM, Michael RT (1977) An economic analysis of marital instability. J Polit Econ 85(6): 1141–1187

Bloch L (1929) Facts about Mexican immigration before and since the quota restriction laws. J Am Stat Assoc 24(165):50–60

Bodnar J (1985) The transplanted: a history of immigrants in urban America. Indiana University Press, Bloomington

Bratter JL, King RB (2008) "But will it last?": marital instability among interracial and same-race couples. Fam Relat 57(2):160–171

Casey T, Dustmann C (2010) Immigrants' identity, economic outcomes and the transmission of identity across generations. Econ J 120(542):F31–F51

Chiappori PA, Fortin B, Lacroix G (2002) Marriage market, divorce legislation, and household labor supply. J Polit Econ 110(1):37–72

Childs EC (2005) Navigating interracial borders: black-white couples and their social worlds. Rutgers University Press, Piscataway

Chiswick CU (2009) The economic determinants of ethnic assimilation. J Popul Econ 22(4):859–880

Chiswick BR, Houseworth C (2011) Ethnic intermarriage among immigrants: human capital and assortative mating. Rev Econ Househ 9(2):149–180

Constant AF, Zimmermann KF (2008) Measuring ethnic identity and its impact on economic behavior. J Eur Econ Assoc 6(2/3):424–433

Dalmage H (2000) Tripping on the color line: black/white multiracial families in a racially divided world. Rutgers University Press, Piscataway

Dribe M, Lundh C (2012) Intermarriage, value context and union dissolution: Sweden 1990–2005. Eur J Popul 28(2):139–158

Duncan B, Trejo S (2005) Ethnic identification, intermarriage, and unmeasured progress by Mexican Americans. IZA Discussion Paper DP1627

England P, Folbre N (1999) The cost of caring. Ann Am Acad Pol Soc Sci 561(1):39–51

Foerster RF (1924) Italian emigration of our times. Harvard University Press, Cambridge

Frimmel W, Halla M, Winter-Ebmer R (2013) Assortative mating and divorce: evidence from Austrian register data. J R Stat Soc A Stat Soc 176(4):907–929

Furtado D (2009) Cross-nativity marriages and human capital levels of children. Res Labor Econ 29:273–296

Furtado D (2012) Human capital and interethnic marriage decisions. Econ Inq 50(1):82–93

Furtado D, Theodoropoulos N (2011) Interethnic marriage decisions: a choice between ethnic and educational similarities. J Popul Econ 24(4):1257–1279

Furtado D, Trejo SJ (2013) 15 Interethnic marriages and their economic effects. In: Constant AF, Zimmermann KF (eds) International handbook on the economics of migration. Edward Elgar Publishing, Northampton, pp 276–292

Gibson, CJ, (2013) Country of Birth of the Foreign-Born Population. In Gibson CJ (ed) *American Demographic History Chartbook: 1790 to 2010.* http://www.demographicchartbook.com

Gilbertson GA, Fitzpatrick JP, Lijun Y (1996) Hispanic intermarriage in New York City: new evidence from 1991. Int Migr Rev 30(2):445–459

Gray J (1998) Divorce law changes, household bargaining, and married women's labor supply. Am Econ Rev 88(3):628–642

Huang W, Zhou Y (2015) Impact of one-child policy on inter-ethnic marriage in China. Working Paper

Hutchinson EP (1956) Immigrants and their Children, 1850–1950. Wiley, New York

Imbens GW, Angrist JD (1994) Identification and estimation of local average treatment effects. Econometrica 62(2):467–475

Johnson WR, Skinner J (1986) Labor supply and marital separation. Am Econ Rev 76(3):455–469

Jones FL (1996) Convergence and divergence in ethnic divorce patterns: a research note. J Marriage Fam 58(1):213–218

Kalmijn M (1998) Intermarriage and homogamy: causes, patterns, trends. Annu Rev Sociol 24:395–421

Kalmijn M, de Graff PM, Jacques PG (2005) Intermarriage and the risk of divorce in the Netherlands: the effects of differences in religion and in nationality, 1974–94. Popul Stud 59(1):71–85

Kulu H, Milewski N (2008) Family change and migration in the life course: an introduction. Demogr Res 17: 567–590

Lafortune J (2013) Making yourself attractive: pre-marital investments and the returns to education in the marriage market. Am Econ J Appl Econ 5(2):158–178

Lam D (1988) Marriage markets and assortative mating with household public goods: theoretical results and empirical implications. J Hum Resour 23(4):462–487

Lundberg S (2010) Personality and marital surplus. IZA Discussion Paper DP7982

Lundberg S, Pollak AP (1993) Separate spheres bargaining and the marriage market. J Polit Econ 101(6):988–1010

Massey DS, Durand J, Malone NJ (2002) Beyond smoke and mirror: Mexican immigration in an era of economic integration. Russell Sage, New York

Meng X, Gregory RG (2005) Intermarriage and the economic assimilation of immigrants. J Labor Econ 23(1): 135–175

Newey WK (1990) Efficient instrumental variables estimation of nonlinear models. Econometrica 58(4):809–837

Pagnini DL, Morgan SP (1990) Intermarriage and social distance among U.S. immigrants at the turn of the century. Am J Sociol 26(2):405–432

Parkman AM (1992) Unilateral divorce and the labor force participation rate of married women, revisited. Am Econ Rev 82(3):671–678

Qian Z (1997) Breaking the racial barriers: variations in interracial marriage between 1980 and 1990. Demography 34(2):263–276

Qian Z, Lichter DT (2007) Social boundaries and marital assimilation: interpreting trends in racial and ethnic intermarriage. Am Sociol Rev 72(1):68–94

Qian Z, Blair SL, Ruf SD (2001) Asian American interracial and interethnic marriages: differences by education and nativity. Int Migr Rev 35(2):557–586

Rosenblum MR, Kandel WA, Seelke CR, Wasem RE (2012) Mexican Migration to the United States: Policy and Trends. CRS Report for Congress R42560.

Stevenson B (2007) The impact of divorce laws on marriage-specific capital. J Labor Econ 25(1):75–93

Stevenson B (2008) Divorce law and women's labor supply. J Empir Leg Stud 5(4):853–873

Wildsmith E, Myron PG, Gratton B (2003) Assimilation and intermarriage for U.S. immigrant Groups, 1880–1990. Hist Fam 8(4):563–584

Wooldridge JM (2001) Econometric analysis of cross section and panel data. MIT Press, Cambridge, p 623

Zimmermann L, Zimmermann KF, Constant A (2007) Ethnic self-identification of first-generation immigrants. Int Migr Rev 41(3):769–781