Making Yourself Attractive: Pre-Marital Investments and the Returns to Education in the Marriage Market Online Appendix

By Jeanne Lafortune

A. Theoretical derivations

Given the utility function of each individual and the fact that they are too small to alter the behavior of other agents in the economy, the first order conditions for i and s (denoting them by $Y(i,s,z)=0$ and $Z(i,s,z)=0$ respectively) are given by

(A1)

$$\begin{aligned}(1+r+\delta)(w-i-s)^{-\sigma} = {} & \delta(i+\tilde{r}s)^{-\sigma} + (1+r)\sum_{i'} p^k(i')(c_2^{kE}+\tilde{r}s)^{-\sigma}\frac{\partial c_2^{kE}}{\partial i} \\ & + (1+r)(1-p^k)(c_2^{kN}+\tilde{r}s-\gamma)^{-\sigma}\frac{\partial c_2^{kN}}{\partial i} \\ & + \frac{1+r}{1-\sigma}\sum_{i'}\left((c_2^{kE}+\tilde{r}s)^{1-\sigma}-(c_2^{kN}+\tilde{r}s-\gamma)^{1-\sigma}\right)\frac{\partial p^k(i')}{\partial i}\end{aligned}$$

(A2)

$$\frac{(1+r+\delta)}{\tilde{r}}(w-i-s)^{-\sigma} = \delta(i+\tilde{r}s)^{-\sigma}+(1+r)\sum_{i'} p^k(i')(c_2^{kE}+\tilde{r}s)^{-\sigma}+(1+r)(1-p^k)(c_2^{kN}+\tilde{r}s-\gamma)^{-\sigma}$$

As long as $\tilde{r}$ is not too large or too small, the two above conditions will be satisfied with equality and an individual will invest a positive amount in both types of investments.

Using these first order conditions, we can derive the effect of a change in z, remembering that z enters into two of the elements of the utility of the agent, namely $p^k(i,i',z)$ and $c^{kE}(i,i',z)$. Solving for the effect on investment of a change in z, we obtain:

(A3)

$$\frac{\partial i}{\partial z} \propto (1+r)\sum_{i'}\left(\left((c_2^{kE}+\tilde{r}s)^{-\sigma}\left(\frac{\partial c_2^{kE}}{\partial i}-X\right)-(c_2^{kN}+\tilde{r}s-\gamma)^{-\sigma}\left(\frac{\partial c_2^{kN}}{\partial i}-X\right)\right)\frac{\partial p^k(i')}{\partial z}\right.$$
$$+\frac{1}{1-\sigma}\left((c_2^{kE}+\tilde{r}s)^{1-\sigma}-(c_2^{kN}+\tilde{r}s-\gamma)^{1-\sigma}\right)\frac{\partial^2 p^k(i')}{\partial i\partial z}$$
$$\left.+(c_2^{kE}+\tilde{r}s)^{-\sigma-1}\left(-\sigma p^k(i')\left(\frac{\partial c_2^{kE}}{\partial i}-X\right)\frac{\partial c_2^{kE}}{\partial z}+(c_2^{kE}+\tilde{r}s)\left(\frac{\partial p_k}{\partial i}\frac{\partial c_2^{kE}}{\partial z}+p^k(i')\frac{\partial^2 c_2^{kE}}{\partial i\partial z}\right)\right)\right)$$

where $X=\frac{\frac{\partial Z}{\partial s}}{\frac{\partial Y}{\partial s}}$.

The right-hand side can be decomposed into four elements. The first two, called the "matching effect", highlights the impact that z has on the probability of matching with a native. Given that the marginal utility of matching with a native is higher than that of matching with someone from one's ethnic group (since one is less happy), that $\frac{\partial c_2^{kE}}{\partial i} < \frac{\partial c_2^{kN}}{\partial i}$ and that $\frac{\partial c_2^{kN}}{\partial i} > X$, it is clear that the matching effect will be in the opposite direction as $\frac{\partial p^k(i')}{\partial z}$. The second term of the "matching effect" corresponds to the effect that z has on the return to one's investment in terms of probability of matching within one's preferred market. Since utility when matched within one's ethnic group is higher than match outside, this term will be of the same sign as $\frac{\partial^2 p^k(i')}{\partial i\partial z}$ that is, it depends on whether the sex ratio increases or decreases the attractiveness of the investment for spouses of distinctive investment levels.

Finally, the sex ratio also changes the way the surplus is allocated between spouses, which is captured by the third line of equation (A3) and is referred to as the "bargaining effect". Through the fact that it influences second period sharing between spouses, the sex ratio modifies both the return on one's investment and the level of consumption in that second period. As the marriage market makes one's second-period return less valuable (both because the return in that second period falls and because the impact of increasing one's probability of matching within one's ethnic group becomes less attractive), one's incentives for investing fall. On the other hand, one is now poorer, which raises the incentives for investment although some of it may be done through an increase in s.

B. Construction of marriage market variables

Marriage markets are first defined at the state level as it is the lowest geographical unit for which place of birth is available in the IPUMS files, which is used to alleviate concerns of endogenous mobility. However, more than 65 percent of

sampled individuals are married to someone born in the same state as them.[1] Immigrants are classified based on their current state of residence as it is the only information available.

Furthermore, marriage markets are defined within an ethnic group, which is composed of a number of ethnicities. From 1900 to 1970, the IPUMS files include information on parents' country of origin. Using this variable, each second generation individual is associated with a particular ethnicity based on father's ethnicity.[2] Immigrants are classified according to their own country of birth. Using all countries of birth, the sample was divided into 9 ethnic groups, summarized in Appendix Table C1. This division was inspired by that used by Angrist (2002) and based on Pagnini (1990), with required modifications.[3]

Finally, the marriage market is defined within an age cohort of 5 years. While this is restrictive, around 50 percent of all married individuals younger than 40 are matched within that age group. Since the interest of the paper lies in capturing the marriage market as perceived by individuals when taking their educational decision, the marriage market here will include second-generation individuals, immigrants arriving as children (before age 8 if females, before age 10 if males) and newly arrived immigrants. Second generation individuals and child immigrants are restricted to those born between 1885 and 1915 and are divided into 5 year-of-birth intervals to form a marriage market. To that number, immigrants who arrive while second generations are deciding whether or not to remain in school (age 11-19) are added. Immigrants who arrive between 1900 and 1904, for example, are matched with individuals born between 1885 and 1889, those arriving between 1905 and 1909, to those born between 1890 and 1895, etc. Only immigrants who are in the appropriate age groups (10-25 for males, 8-23 for females) are included since they are most likely to be part of the marriage pool.[4]

The number of immigrants or foreign stock in a marriage market is simply the number of individuals in a given cell; the sex ratio is the number of males per females in each cell.[5] To avoid double-counting for the flow/stock indicator, only

[1]This is almost as large as the proportion of individuals still living in their state of birth. One finds very small proportion of "out-of-state" marriages for individuals who are still living in their state of birth.

[2]While Angrist (2002) uses mother's ethnicity, I employ father's ethnicity because in 1960 and 1970, only father's ethnicity is reported when the father is foreign born. This is of little importance, however, because 95 percent of foreign born parents share a common country of birth.

[3]East European Jews are grouped by nationality because it is difficult to identify them after 1930. Also, two countries of birth per ethnic group are required since the instrument relies on differences in 1900 location choices within ethnic groups across countries of birth. Immigrants from Ireland were joined with those from other British Isles. Italians were grouped with other Catholic Southern European countries: Spain and Portugal. Finally, Mexicans were included with other immigrants from the Caribbean, Central and South America.

[4]Similar results were obtained using the same age for both males and females. A variant also built this measure matching all individuals based on their cohort of birth rather than time of arrival (restricting it to individuals arriving before schooling decisions are made) and the results were very similar to the ones presented here.

[5]If the cell is empty, the sex ratio is set to 1. If there are only men, the sex ratio is equal to 1.5 times the number of males. Neither adjustment is crucial; similar results were obtained with various modifications.

the 1910 Census is used to compute the flow of immigrants arriving between 1900 and 1909 and the stock of second generation Americans born between 1885 and 1895, the 1920 Census for immigrants arriving between 1910 and 1919 and so forth. However, since the sex ratio may suffer more from measurement error in small cells because it is a ratio, all three waves of the Census were employed to construct that measure for immigrants.

Table C1—Ethnic group composition

Country	Same country		Same group (excl. own country)		Other groups (excl. natives)	
	Males	Females	Males	Females	Males	Females
1. British Ancestry						
Australia	**0.08**	**0.08**	**0.13**	**0.13**	**-0.28**	**-0.28**
English Canada	**0.18**	**0.18**	**0.05**	**0.05**	**-0.11**	**-0.11**
England	**0.13**	**0.13**	**0.06**	**0.06**	**-0.13**	**-0.13**
Ireland	**0.35**	**0.35**	**0.05**	**0.05**	**-0.13**	**-0.13**
Scotland	**0.08**	**0.08**	**0.13**	**0.13**	**-0.13**	**-0.13**
Wales	**0.22**	**0.22**	**0.09**	**0.09**	**-0.17**	**-0.17**
2. Francophone						
Belgium	**0.42**	**0.42**	**0.01**	**0.01**	**-0.07**	**-0.07**
French Canada	**0.42**	**0.42**	-0.03	-0.03	**-0.14**	**-0.14**
France	**0.08**	**0.08**	-0.02	-0.02	0.10	0.10
3. South Europeans						
Italy	**0.63**	**0.63**	-0.01	-0.01	**-0.18**	**-0.18**
Spain	**0.56**	**0.56**	-0.03	-0.03	**-0.15**	**-0.15**
Portugal	**0.08**	**0.08**	**0.01**	**0.01**	0.07	0.07
4. Hispanics						
Central America*	**0.00**	**0.00**	-0.01	-0.01	**-0.39**	**-0.39**
Cuba	**0.25**	**0.25**	-0.01	-0.01	**-0.04**	**-0.04**
Mexico	**0.69**	**0.69**	**0.00**	**0.00**	**-0.32**	**-0.32**
South America	**0.00**	0.00	-0.01	-0.01	0.04	0.04
Other West Indies	**0.30**	**0.30**	-0.01	-0.01	**-0.24**	**-0.24**
5. Scandinavians						
Denmark	**0.26**	**0.26**	**0.04**	**0.04**	**-0.08**	**-0.08**
Finland	**0.40**	**0.40**	**0.09**	**0.09**	**-0.28**	**-0.28**
Norway	**0.58**	**0.58**	**0.06**	**0.06**	**-0.21**	**-0.21**
Sweden	**0.44**	**0.44**	**0.04**	**0.04**	**-0.17**	**-0.17**
6. Germanic						
Austria	**0.45**	**0.45**	**0.10**	**0.10**	**-0.12**	**-0.12**
Germany	**0.42**	**0.42**	**0.00**	**0.00**	**-0.13**	**-0.13**
Luxembourg	**0.27**	**0.27**	**0.28**	**0.28**	**-0.11**	**-0.11**
Netherlands	**0.43**	**0.43**	**0.04**	**0.04**	**-0.15**	**-0.15**
Switzerland	**0.12**	**0.12**	**0.20**	**0.20**	**-0.12**	**-0.12**
7. Russians and others						
Poland	**0.58**	**0.58**	**0.02**	**0.02**	**-0.12**	**-0.12**
Romania	**0.22**	**0.22**	**0.20**	**0.20**	0.07	0.07
Russia	**0.60**	**0.60**	**0.04**	**0.04**	**-0.17**	**-0.17**
8. Other Europe						
Bohemia	**0.60**	**0.60**	**0.00**	**0.00**	**-0.13**	**-0.13**
Greece	**0.33**	**0.33**	-0.02	-0.02	**-0.05**	**-0.05**
Hungary	**0.46**	**0.46**	**0.00**	**0.00**	0.07	0.07
Other Europe	**0.29**	**0.29**	-0.02	-0.02	**-0.11**	**-0.11**
9. Other countries						
Africa	**0.12**	**0.12**	-0.02	-0.01	**-0.25**	**-0.27**
Atlantic Islands*	**0.00**	**0.00**	-0.02	-0.01	0.62	0.61
China*	**0.73**	**0.73**	-0.02	-0.01	**-0.22**	**-0.24**
India*	**0.00**	**0.00**	-0.02	-0.01	**-0.38**	**-0.39**
Japan*	**0.00**	**0.00**	-0.02	**0.00**	**-0.38**	**-0.39**
Pacific Islands*	**0.00**	**0.00**	-0.02	-0.01	**-0.04**	**-0.06**
Turkey*	**0.00**	**0.00**	-0.02	-0.07	**-0.38**	**-0.33**
Other Asia*	**0.00**	**0.00**	-0.02	-0.01	**-0.13**	**-0.14**
Other countries*	**0.33**	**0.33**	-0.02	-0.01	**-0.04**	**-0.06**

Note: Each entry represents the difference between the proportion of second generation individuals born between 1865 and 1884 married with each type of partner and the proportion of individuals of that type among all individuals of the other gender in the sample. For example, second-generation Australians married second-generation or immigrant Australians 8 percent more than their relative abundance within that cohort and married other individuals of British Ancestry (but not from Australia) 13 percent more than would have been expected by pure random matching. An asterisk indicates that the sample included fewer than 5 second-generation males and females. Bold entries correspond to the ones that are supportive of the grouping that is where individuals of a group were more likely to marry within their country of birth and their ethnic group and less likely to do in other ethnic groups than would simply be generated by random matching.

TABLE C2—SPATIAL DISTRIBUTION OF IMMIGRANTS BY ETHNIC GROUP, 1900

Ethnic group	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	TOP 10
British	NY	MA	PA	IL	MI	NJ	OH	CA	CT	MN	75.4
ancestry	(19.4)	(14.5)	(11.2)	(6.9)	(6.6)	(4.5)	(3.9)	(3.4)	(3.0)	(2.1)	
French	MA	MI	NY	NH	IL	RI	ME	WI	CT	NJ	75.6
	(21.9)	(10.7)	(9.2)	(7.0)	(6.8)	(5.1)	(4.9)	(3.6)	(3.4)	(2.8)	
South	NY	PA	MA	NJ	CA	IL	CT	LA	RI	OH	86.3
Europeans	(34.8)	(12.7)	(8.0)	(8.0)	(6.8)	(4.5)	(3.7)	(3.4)	(2.2)	(2.2)	
Hispanics	TX	AZ	FL	CA	NY	NM	PA	MA	LA	NJ	93.9
	(51.9)	(10.3)	(8.6)	(7.4)	(6.5)	(4.9)	(1.2)	(1.2)	(0.9)	(0.9)	
Germanic	NY	IL	PA	WI	OH	NJ	MI	IA	MN	MO	77.1
	(18.7)	(11.8)	(9.4)	(8.4)	(7.4)	(4.6)	(4.4)	(4.2)	(4.2)	(3.9)	
Scandinavians	MN	IL	WI	IA	NY	MI	ND	MA	NE	SD	76.3
	(21.9)	(12.9)	(9.4)	(6.4)	(6.0)	(5.3)	(3.8)	(3.8)	(3.6)	(3.1)	
Russians and	NY	PA	IL	MA	WI	NJ	MI	OH	CT	MN	83.6
others	(29.1)	(15.8)	(12.0)	(6.0)	(4.5)	(4.2)	(4.0)	(3.1)	(2.7)	(2.1)	
Other	NY	PA	IL	OH	NE	NJ	WI	MN	IA	TX	85.9
Europeans	(20.1)	(16.1)	(14.4)	(9.7)	(5.1)	(5.1)	(4.7)	(4.2)	(3.5)	(3.1)	
Other	HI	CA	MA	NY	OR	WA	MT	PA	AK	IL	86.4
countries	(34.6)	(24.5)	(5.6)	(5.5)	(5.3)	(0.42)	(1.9)	(1.9)	(1.5)	(1.4)	

Note: For each ethnic group, the first row represents the states with the highest concentration and the second, the actual concentration in each state. The last column measures the share of all immigrants from that ethnic group located in the ten most popular states for that ethnic group.

TABLE C3—DATA DESCRIPTION

Variables	**Census years**	**Age sampled**	**Details**
Pre-marital investments			
Literacy	1900-30	16-25	Literacy in any language
Highest grade achieved	1940-70	46-75	Only available from 1940
Some high school	1940-1970	46-75	Only available from 1940
Educational score	1900-30	16-25	Based on average education of workers in the current occupation in 1950
Marital outcomes			
Ever married	1900-70	36-75	
Currently divorced	1900-70	36-75	
Married more than once	1910 1940-60	36-75	
Currently married to same ethnic immigrant	1900-70	36-75	
Age at first marriage	1930-40 1960-70	36-75	
Post-marital labor supply			
In the labor force	1910-70	26-75	
Employed	1910 1930-70	26-75	
Hours worked per week	1940-70	26-75	Transformed from intervals
Weeks worked per year	1940-70	26-75	to continuous variables by selecting the mid-point of the interval

TABLE C4—EFFECT OF SEX RATIO ON PRE-MARITAL INVESTMENTS-ADDITIONAL RESULTS

	Only high predicted sex ratio cells			**Interacted with educ. difference**			**Interacted with divorce rate**		
	Highest grade attained	Some high school	Educ. score	Highest grade attained	Some high school	Educ. score	Highest grade attained	Some high school	Educ. score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				Panel A: Males					
Sex ratio	2.155***	0.255**	25.529	1.324**	0.236***	19.118	0.986	0.173	6.644
	(0.586)	(0.110)	(17.909)	(0.532)	(0.086)	(34.586)	(0.623)	(0.124)	(22.292)
Stock	-0.404	-0.109	26.393***	0.674	-0.082	123.824**	0.768	-0.058	130.244**
	(0.600)	(0.085)	(9.653)	(0.839)	(0.153)	(55.871)	(0.924)	(0.161)	(51.347)
Interaction				0.057	0.015	3.423	0.199	0.024	4.461***
				(0.116)	(0.022)	(5.965)	(0.233)	(0.038)	(1.061)
N	45152	45152	34177	86683	86683	66565	86683	86683	66565
				Panel B: Females					
Sex ratio	0.829	0.091	7.882	-0.254	-0.071	8.005	0.424	0.064	-72.441
	(0.927)	(0.082)	(27.835)	(0.750)	(0.096)	(32.599)	(0.567)	(0.088)	(44.862)
Stock	-0.192	-0.026	2.744	0.246	0.072	66.346*	-0.253	-0.045	52.820
	(0.333)	(0.053)	(17.400)	(0.851)	(0.114)	(37.103)	(0.888)	(0.138)	(33.824)
Interaction				-0.472*	-0.099***	-2.043	0.137	0.034	90.573***
				(0.248)	(0.035)	(13.911)	(0.222)	(0.036)	(25.006)
N	50223	50223	35171	96707	96707	68916	96707	96707	68916

Note: Standard errors clustered at the state level in parentheses. All regressions include state, ethnic groups, immigration period fixed effects and all double interactions. Also includes age fixed effects and dummies for whether both or only one parent is foreign-born and the educational difference between the ethnicity and natives in columns (4)-(6) and the divorce rate for each ethnicity in columns (7)-(9). In columns (1)-(3), the sample is restricted to cells where the predicted sex ratio is larger than 1.1.All regressions are weighted by the Census sample-line weight.

TABLE C5—EFFECT OF EDUCATION ON LABOR SUPPLY

	In LF (1)	Hours (2)	In LF (3)	Hours (4)	In LF (5)	Hours (6)	In LF (7)	Hours (8)
					Panel A: Males			
	OLS		**IV**		**OLS**		**IV**	
Own education	0.010***	0.814***	-0.026	-1.545**	0.006***	0.554***	-0.022	-4.122***
	(0.000)	(0.027)	(0.017)	(0.702)	(0.000)	(0.025)	(0.022)	(1.373)
Spouse's education					0.003***	0.361***	0.019	3.724***
					(0.000)	(0.022)	(0.018)	(1.160)
F-test (instruments)								
Own			11.10***				14.88***	
Spouse							11.78***	
N. Obs	755241	731843	755241	731843	452846	438832	452846	438832
					Panel B: Females			
Own education	0.020***	0.866***	0.014	0.712	0.025***	0.988***	-0.084	-2.821
	(0.001)	(0.032)	(0.022)	(0.785)	(0.001)	(0.033)	(0.065)	(2.840)
Spouse's education					-0.010***	-0.357***	0.106	3.913
					(0.001)	(0.025)	(0.068)	(2.819)
F-test (instruments)								
Own			14.77***				13.42***	
Spouse							6.91***	
N. Obs	816934	804607	816934	804607	451970	445720	451970	445720

Note: Standard errors clustered at the state level in parentheses. All regressions include state, year of birth and Census year fixed effects as well as age and age squared. All regressions are weighted by the Census sample-line weight.