INSTITUTO DE ECONOMÍA

DOCUMENTO de TRABAJO

422

2012

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

Jeanne Lafortune.

www.economia.puc.cl · ISSN (edición impresa) 0716-7334 · ISSN (edición electrónica) 0717-7593

Versión impresa ISSN: 0716-7334
Versión electrónica ISSN: 0717-7593

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA

Oficina de Publicaciones
Casilla 76, Correo 17, Santiago
www.economia.puc.cl

MAKING YOURSELF ATTRACTIVE: PRE-MARITAL INVESTMENTS AND THE RETURNS TO EDUCATION IN THE MARRIAGE MARKET

Jeanne Lafortune*

Documento de Trabajo Nº 422

Santiago, Junio 2012

---

* jlafortune@uc.cl

INDEX

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

Jeanne Lafortune†

June 2012

Abstract

This paper explores how a rise in a gender's scarcity may impact educational investments using exogenous variation in the marriage market of second generation Americans in early 20th century. Theoretically, one may expect this to occur through two potential channels: a change in matching possibilities or in post-match bargaining. Empirically, I find that worse marriage market conditions spurs higher pre-marital investments: the effect for males is significant (0.2 years of education for one standard deviation in the sex ratio) while for females, it is only observed in highly endogamous groups. When faced with an exogenously larger number of males per females, males' marriages appear to be less stable and more likely to involve natives and more educated spouses while women are less likely to work and, for those in high endogamous groups, marry more immigrants.

JEL classification: J12, J24, J61
Key words: Pre-marital investments, Sex ratios, Marriage market

*I am particularly indebted to Esther Duflo and David Autor for their comments and advice as well as three anonymous referees and the editor for their valuable suggestions. I also wish to thank Josh Angrist, Abhijit Banerjee, Alan Manning, Tavneet Suri, Tal Gross, Patrick Warren, Suman Basu and José Tessada for their valuable suggestions and support. I also benefited from comments from the participants at seminars at MIT, PUC-Chile, UCLA, UC Berkeley, University of Maryland-College Park, Carleton University, University of Toronto, Université de Montreal, Universitá Bocconi, University of Virginia, University of British Columbia, Georgetown University, Board of Governors of the Federal Reserve Bank, the ZEW "Gender and Labor Markets" workshop and SOLE. I acknowledge financial support from the Social Sciences and Humanities Research Council of Canada. All remaining errors are my own.

†Instituto de Economía, Pontificia Universidad Catolica de Chile. E-mail: jlafortune@uc.cl

1 Introduction

How do the prospects of one's marriage market influence one's investments before the union takes place? While several studies have shown that marriage market conditions (such as divorce laws and sex ratios) affect *post*-marital behavior (Chiappori et al. 2002 and Angrist 2002), there is little empirical work on the impact of these factors on *pre*-marital investments. This is surprising, considering that if individuals are forward-looking, these conditions could then be anticipated and potentially modify pre-marital decisions. For example, if one foresees having a lower share of the post-marital output, one could increase one's pre-marital investment in order to compensate for this loss. These adjustments have been emphasized recently in theoretical models (see for example Iyigun and Walsh 2007 and Chiappori et al. 2009) but have not been explored empirically. Moreover, several studies have demonstrated that educational investments appear to respond to perceived returns in practice (see Foster and Rosenzweig 1996, Jensen 2010 and Nguyen 2008 for returns to education in the labor force and Foster and Rosenzweig 2001 for returns to education in marriage markets). Changes in marriage market conditions may thus have an impact upon agents' behavior before the union is formed through potentially multiple channels. This paper investigates both empirically and theoretically how changes in sex ratios (which will be defined here as the ratio of males to females in a marriage market) modify incentives for pre-marital investments.

This paper first explores the relationship between these two variables in a simple theoretical model where individuals face search constraints. They can match within their ethnic group, which gives them a higher utility, or within an alternative one (the pool of native individuals), which gives them a spouse that is more educated but where the utility of such a match is lower. In this model pre-marital investments are undertaken separately by each individual; individuals are paired and produce an output where investments are complementary and the output is shared according to marriage market conditions.[1]

In this set-up, a change in the sex ratio can be expected to affect pre-marital investments through two channels: its effect on the probability of finding a spouse and the quality of the match (which would hold even in a unitary model of the household) and on the anticipated bargaining power.[2] First, a higher sex ratio may increase males' and decrease females' investments because it influences the probability of matching within one's ethnic group. The higher the probability of matching within one's ethnic group, the lower are one's incentives for investments. One has a higher second-period utility (which decreases the need for trans-

[1]Pre-marital investments are thus undertaken non-cooperatively while consumption decisions once paired are taken cooperatively, as in Konrad and Lommerud (2000).

[2]This distinction was made previously by Grossbard-Shechtman (1984)

ferring resources from the first period to the next one through investments) and also receives a lower return as investments are complementary and members of one's ethnic group are less well educated than natives. This trade-off between the benefits and costs of endogamy was previously discussed in Furtado (2012) and Furtado and Theodoropoulos (2011).[3] The sex ratio may also change how one's investment increases one's attractiveness which would lead males to educate themselves more when faced with a tighter marriage market.[4]

Secondly, sex ratios may also alter incentives for pre-marital investments because they modify the balance of power within a household. The standard framework linking bargaining power and investments draws upon the work of Grossman and Hart (1980), in which agents with linear utility functions are engaged in a contractual arrangement. In that framework, an increase in one's bargaining power always leads one to invest more since the additional bargaining power translates into a larger share of the returns on that investment. However, when the utility function is concave, one must add an income effect to the traditional substitution effect of Grossman and Hart (1980) thus leading individuals to want to invest more to counter-act the fact that the lower bargaining power translates into smaller incomes. The overall effect here is unclear.

Having detailed the two channels through which sex ratios may impact pre-marital investments, this paper then explores the empirical relationship in the context of ethnic marriage markets in the United States for individuals around the turn of the twentieth century. Specifically, I explore whether children of immigrants, here referred to as second-generation Americans, born between 1885 and 1915 modify their human capital when confronted, in their late teens, with a plausibly exogenous shift in the sex ratio of their state-level ethnic marriage market, as measured in Censuses from 1900 to 1970.

I exploit the fact that a large fraction of second-generation Americans tend to marry within their own ethnicity. Therefore, the waves of newly arrived immigrants have a disproportionally higher impact on their ethnicity's marriage markets, as in Angrist (2002). While Angrist looks at national ethnic markets and instruments using flows at entry, this study focuses on state-level, within-ethnic group variation, which allows me to control for many potential confounding factors and better isolate the effect of a change in sex ratios. The variation in sex ratios of immigrants at this level is large but since immigrants may select their location based on labor and marriage market conditions which also affect the second generation, it may be

[3]This model, however, does not capture the possibility that having more education may make natives more attractive as the utility cost of marrying a native is independent of one's education.

[4]The conclusions from this model contrasts with those of Iyigun and Walsh (2007). Their framework suggest that with perfect competition, investments will be made Pareto-optimally (that is everybody will receive the full return to their investment). Investments will thus not be used to attract a more qualified spouse in their context.

endogenous. This paper thus constructs an instrument based on the fact that, during this period, the location choice of immigrants by country within a larger ethnic group are persistent over time, as documented in Lafortune and Tessada (2011). Furthermore, over the course of the early twentieth century, the sex ratio of new immigrants has varied substantially and differently across origin countries. Consequently, one can construct an instrument that allocates flows of immigrants to different states using past distributions, comparable to the strategy used by Card (2001). This variation proves to be highly predictive of both the flow of newly arrived immigrants and their gender composition. In addition, it generates sufficient variation to instrument for the overall sex ratio among the foreign stock (immigrants and second generation individuals combined).

Using this strategy, this paper finds that shifts in sex ratios influence pre-marital human capital investment decisions of individuals, proxied by the completed years of education and the educational characteristics of the occupation they select. In states and ethnic groups where the number of males per female increases due to gender-biased immigration, young adult males acquire more formal education and select occupations where individuals are more educated. A change in the sex ratio equivalent to one standard deviation leads second-generation males to increase their educational investment by 0.2 years. This effect appears to be the same irrespective of the degree of endogamy within one's ethnic group. Females, on the other hand, only decrease their educational investment when they are from an ethnic group with high rates of intermarriage. These results are fairly robust to a variety of specifications and do not appear to be driven by labor market impacts of immigrations as they do not respond to the inflow of immigrants of a slightly older cohort.

Marital outcomes are also altered. Males are not less likely to marry when the sex ratio is high although they are less likely to marry an immigrant of their own ethnic group. This leads them to be more likely matched with a spouse with a high school degree.[5] Females are more likely to marry an immigrant of their ethnic group but again, only when their ethnic group is highly endogamous.

Furthermore, as in previous studies (for example, Angrist 2002, Chiappori et al. 2002, Amuedo-Dorantes and Grossbard 2007 and Oreffice and Bercea 2008), this paper also finds that post-matching labor supply decisions are affected by a change in the sex ratio. Contrary to these previous studies, however, only the gender who did not adjust its pre-marital investment in response to a change in the sex ratio appears to experience a significant change in its labor supply. A rise in the sex ratio from 1 to 1.1 leads to a fall of about 1.5 percent in the prob-

[5]This result differ from previous studies such as Abramitzky and Vasconcelos (2011) which found that lower sex ratios allowed men to marry up in terms of class.

ability of being employed and a decrease of about 0.6 hour worked per week and 0.6 week worked per year for females. Indicative evidence suggests that males would have increased their labor force significantly had it not been from adjustments made pre-maritally.

These results highlight that human capital decisions may be driven by other factors than simply labor market returns and that marriage market conditions can influence that decision. Theoretically, this was previously discussed by Boulier and Rosenzweig (1984), Peña (2006), Chiappori et al. (2009) and Browning et al. (2010) but little empirical evidence has so far been provided.

While the setting is clearly historical in nature and many factors (in particular the average educational attainment of males and females) have evolved since then, the results are still informative in many respects. Maybe surprisingly, this period was characterized by matching patterns that resemble the ones observed today which make the results potentially relevant: males and females shared a similar level of educational attainment, marriages were contracted at an advanced age and the age difference between spouses was very small. However, this setting was also characterized by much lower female labor force participation than in today's market. Nevertheless, if males are responding in a setting where complementarities between males and females investments were potentially more limited than today, due to this lower labor force participation, we may think that the results obtained represent a lower bound on what would be expected in today's context.

Furthermore, there are circumstances in today's marriage markets where sex ratios are biased, either by sex selective abortion, gender-skewed incarceration rates, immigration or wars.[6] This study highlights the potentially long-lasting effects of these shocks as they may impact educational choices of a full cohort. In some cases, this could even reinforce the preference for a gender which generated the skewed sex ratio to begin with. For example, if, in front of a high sex ratio, males study more than females, parents may be led to believe that their preference for sons over daughters is justified and perpetuate it.

This paper is organized as follows. A model is first presented in Section 2 which inspires the empirical strategy and data presented in the subsequent section. Section 4 presents the empirical results and the last section concludes.

[6]The impact of this imbalance has been studied previously, see for example Porter (2007) and Edlund et al. (2007) for the case of China and Brainerd (2006), Kvasnicka and Bethmann (2012), Angrist (2002) and Abramitzky and Vasconcelos (2011) for historical circumstances. No previous study looked at pre-marital investments.

2 Model

2.1 Set-up

Let us assume a setting where a number N_m of identical men and N_f of identical women from a given ethnic market are endowed with an initial wealth w. Define the sex ratio as the number of males per female in this economy as $z = \frac{N_m}{N_f}$. Individuals have an additive utility function over two periods[7]:

$$u(c_1, c_2) = u(c_1) + E\left(u(c_2)\right).$$

For simplicity, assume that the utility function has constant elasticity of inter-temporal substitution/constant relative risk aversion given by the parameter σ that is $u\left(c_t\right) = \frac{c_t^{1-\sigma}}{1-\sigma}$, $\sigma > 0$, for $t = 1, 2$.[8]

In the first period, an individual can invest in i units of a productive asset at a unit cost of 1. She can also put a fraction of her wealth s in a bank which offers a rate of interest of $\tilde{r} > 1$ which is independent of her marriage market outcome will be and will be consumed privately and solely by her if married. Her consumption in the first period is thus given by $c_1 = w - i - s$.

In the second period, this investment produces an output which depends on one's investment and that of one's spouse. This partner can come from the same ethnic group or from the pool of natives who is assumed to be large, more educated than the individuals in one's marriage market and with marriage possibilities independent of z. Marrying a native, however, incurs a utility cost denoted by γ sufficiently large that marrying within one's ethnic group is the preferred option.

Once paired, individuals bargain and obtain a share of the family output that may depend on marriage market conditions. Denote males' consumption when married within their ethnic group to be given by c_2^{mE} and females' by c_2^{fE} and if married with natives as c_2^{mN} and c_2^{fN} respectively. Assume that these consumption functions are increasing in both partners' investment, are concave in one's own investment and display complementarity in the investment levels. The return on the investment in these functions is assumed to be larger than $\tilde{r}$. If single, however, one simply receives i, which is a lower return to that of savings.

To introduce search frictions, we can think of the second period as being divided into an

[7]It is assumed for simplicity that the discount factor is 1; none of the results derived below depend on this assumption.

[8]None of the results strongly depend on this assumption.

infinite number of shorter intervals and where individuals discount each of these intervals at a rate of r.[9] During each of these intervals, individuals will be either married or single. There are search frictions in the sense that once one becomes separated, they must remain single for at least one period before finding another spouse. This will imply that a relationship-specific surplus will exist and that it will have to be shared between the spouses.[10] While c_2^{kE} may be a function of z, we will assume that the sharing rule with natives will be independent of the sex ratio.

Assume one's probability of matching within one's ethnic group may be related but not perfectly to one's investment level.[11] Let's denote by $p^k(z,i,i')$ the probability that an individual of gender k with investment level i finds a spouse of investment level i' within their ethnic group and by $p^k(z,i)$ the probability that she finds any spouse within their ethnic group where $\sum_{i'} p^k(z,i,i') = p^k(z,i)$. Also assume that couples separate at an exogenous rate given by δ.

Denoting the expected benefit of being single as $S^k(i,s,z)$ and that of being married with a partner of investment level i' within one's ethnic group as as $E^k(i,i',s,z)$ and with a native partner of investment i^N as $N^k(i,i^N,s,z)$ for each gender k gives:

$$S^k(i,s,z) = u(i+\tilde{r}s) + \frac{1}{1+r}\left(\sum_{i'} p^k(z,i,i')(E^k(i,i',s,z)) + (1-p^k(z,i))N^k(i,i^N,s,z)\right)$$

$$E^k(i,i',s,z) = u(c_2^{kE}(i,i',z)+\tilde{r}s) + \frac{1}{1+r}\left(\delta S^k(i,s,z) + (1-\delta)E^k(i,i',s,z)\right)$$

$$N^k(i,i^N,s,z) = u(c_2^{kN}(i,i^N)+\tilde{r}s-\gamma) + \frac{1}{1+r}\left(\delta S^k(i,s,z) + (1-\delta)N^k(i,i^N,s,z)\right)$$

Those can be solved for and then the expected utility of the individual (assuming that the initial probability of finding a spouse in one's ethnic group is given by $p^k(i,z)$) can be shown to be equal to:

$$u\left(w-i-s\right) + \frac{\delta u(i+\tilde{r}s) + (1+r)\sum_{i'} p^k(z,i,i')u(c_2^{kE}+\tilde{r}s) + (1+r)(1-p^k(z,i))u(c^{kN}-\gamma+\tilde{r}s)}{1+r+\delta}$$

[9] We will thus assume that the overall second period is discounted at the opposite rate such that earning a dollar in every of these intervals would be the same as obtaining one in the first period.

[10] Notice that the need for a concave utility function (which is driving some of the results below), complicates the use of the typical Nash Bargaining derivation in this setting. It also renders the setting of Shimer and Smith (2000) much more complex and is thus not explored here.

[11] If only the most educated males could find a spouse, the effect of an increase in the sex ratio would be to decrease the investment of males through what may be called "discouragement effect" and through the fact that one's bargaining in the household also has fallen.

Individuals in the market will simply maximize the above expression by electing the optimal savings and investment levels. As long as $\tilde{r}$ is not too large or too small, individuals will elect to invest part of their wealth in the bank and part of it in pre-marital investment (that is $i, s > 0$). If $\tilde{r}$ is very large, pre-marital investments will never be positive, which would make the empirical exercise that follow useless. Making $\tilde{r}$ very low such that $s = 0$ would not greatly change the results when the native market exist but it would do so if the outside option of individuals was simply to remain single.

2.2 Comparative statics

Appendix A demonstrates that pre-marital investments respond to the sex ratio through three different channels according to the equation:

$$\frac{\partial i}{\partial z} = (1+r)\sum_{i'}\left(-a\frac{\partial p^k(i')}{\partial z} + b\frac{\partial^2 p^k(i')}{\partial i \partial z} - c\frac{\partial c_2^{kE}}{\partial z} + d\frac{\partial^2 c_2^{kE}}{\partial i \partial z}\right) \tag{1}$$

where a, b, c and d represent functions of parameters and a, b and d are all positive.

This setting emphasizes the two distinct roles played by the sex ratio in the investment decision: matching and bargaining effects. The matching effect refers to the fact that a higher sex ratio will lead all males to have a more difficult time finding a spouse within their ethnic group ($\frac{\partial p^k(i')}{\partial z}$) and to the fact that the change in the sex ratio will change how much my probability of finding a spouse of investment level i' will be influenced by my own investment i ($\frac{\partial^2 p^k(i')}{\partial i \partial z}$). The bargaining effect relates to the fact that once paired with someone of their ethnic group, the sex ratio influences how the surplus is shared ($\frac{\partial c_2^{kE}}{\partial z}$) and thus the return to investment in that state of nature ($\frac{\partial^2 c_2^{kE}}{\partial i \partial z}$).

First, when the sex ratio becomes more biased against one's gender, the probability of finding a spouse within one's preferred market falls and that of matching with a native increases, for any investment level. This has two effects, both of which increase the incentives for investment from that gender. One is more likely to be in a situation where they are less happy in the second period which increases their incentives for transferring resources to the second period through pre-marital investment.[12] Furthermore, the expected return to investment increases as natives are more educated than individuals within one's marriage market and given the complementarity within the consumption function, this will translate into a higher return. The sex ratio may also change the return one receives on pre-marital investment in

[12]Part of that can be done through savings but since the expected return on i is still higher than $\tilde{r}$, individuals will want to do some of the adjustments through pre-marital investments.

terms of the quality of the spouse one may obtain within one's ethnic market. If, in a tighter market, investment increases more the probability of finding a high-quality spouse than in a more favorable one, then the incentives in "making oneself more attractive" increases the investment. Thus, the matching channel would mean (if $\frac{\partial^2 p^k(i')}{\partial i \partial z} > 0$) that when faced with a higher sex ratio, males would increase their investment. Furthermore, the change in the sex ratio through the matching channel will lead to a larger response when the "outside option" entails marrying someone who is more educated than the average endogamous partner.

Secondly, the sex ratio may also change the way the household surplus is separated between spouses, which will be called the "bargaining" effect. In this case, the effect of the sex ratio is not so clear because it changes the incentives in two distinct ways. When the sex ratio moves more in one's disadvantage, one's expected second period income will fall, which increases one's incentives for transferring resources from the first to the second period. On the other hand, the return of one's investment decreases with a fall in bargaining power, which makes the investment less attractive. Which one dominates will depend on the size of the parameter σ (when it is 0 as is traditional in this literature, more bargaining power naturally leads to more investments) and the difference in returns between the native and ethnic markets.

While a change in the sex ratio may generate the same response in terms of human capital acquisition through these two paths, it will influence distinctively the consumption in the second period depending on which channel is at play. On the one hand, if individuals simply adjust their pre-marital investments through the matching channel, one should find that the impact of the sex ratio on post-marital consumption would disappear once one controls for the effect of the sex ratio on investment levels. On the other hand, if the sex ratio modifies post-matching bargaining and individuals are responding through their pre-marital investments to counter-act this, we should observe a larger direct impact of the sex ratio on post-marital consumption, once the indirect effect through investment is accounted for.

3 Data and empirical strategy

Having described the theoretical link that may exist between pre-marital investments and the sex ratio, I will now explore this relationship empirically. As in Angrist (2002), this paper uses data from second-generation Americans born around the turn of the century (from 1885 to 1915). The advantage of this population is that during this period second-generation Americans tended to marry within their ethnic group (40 percent of second generation males and 45 percent of females among a slightly older cohort marry within their own ethnicity). Thus, for

this population, the "preferred" ethnic marriage market includes new waves of immigration. Since marriage markets are fairly local, I define this market to be within a given ethnic group within a state in a particular cohort. I then look at the impact of changes in local marriage market conditions brought about from waves of immigration occurring at a moment when the sample of second generation individuals was in their late teens, making educational and marriage decisions, which corresponds to the years 1900 to 1930. Over that period, the sex ratio of newly arrived immigrants varied greatly transforming the gender balance within each state's ethnic marriage markets. Given that location choices of immigrants may be endogenous, this paper instruments for both the flow and the sex ratio of new immigrants using the fact that immigrants locate near existing networks (as in Card 2001 and justified by Munshi 2003), thus leading past immigrant stocks in a particular state to predict current immigrant flows.

3.1 Basic specification

The basic regression of this study relates pre-marital investment of second generation Americans to two characteristics of the individual's "preferred" marriage market: its sex ratio and its total size. The first one measures the relative abundance of one's gender. The second variable provides an estimate of the effect of market thickness. In addition, it captures any effect that overall own-ethnic immigration has on local conditions, either through marriage or labor markets.[13]

A key decision in implementing this framework is the appropriate empirical definition that should be used for a "marriage market". In this setting, a marriage market is assumed to be one's given ethnic group within one's state of birth in one's 5-year cohort. This definition of marriage market is quite restrictive but as long as what happens to one's market is more important than what occurs in another group, this approximation will capture relevant variation and this approximation would simply bias against finding a significant effect. Furthermore, it appears to correspond to the marriage market most relevant for this sample as most, although not all, marry someone from this group. Immigrants are matched to a market based on their state of residence, their country of origin and their year of arrival such that they are matched with second generations of a comparable age group who were, at that moment of their arrival, deciding whether or not to remain in school (age 11-19). Naturally, this decision will be taken before most of them officially enter into the marriage market but the assumption is that they can observe the tightness of their market as teenagers and presume

[13]The exclusion of the stock variable from these regressions would not change the results presented below.

that it will remain that way in their early 20s. Appendix B provides more detail regarding the definitions used for the construction of these marriage market measures and the way data were combined to do so.

Given this definition of the marriage market, the estimation equation is given by

$$y_{kjst} = \alpha \frac{N^m_{jst}}{N^f_{jst}} + \beta(N^m_{jst} + N^f_{jst}) + \psi X_{kjst} + \mu_j + \nu_s + \theta_t + \sigma_{js} + \phi_{st} + \gamma_{jt} + \varepsilon_{kjst} \tag{2}$$

where the left hand-side variable is an outcome for an individual k, of ethnic group j, born in state s, of cohort t and N^k_{jst} is the number of individuals of gender k in her marriage market. Individual-level controls X_{kjst} include dummies for age and for the nativity of parents. In order to control for potential confounding factors that affect sampled individuals, the regressions include fixed effects for cohort, state and ethnic group as well as for cohort*state, cohort*ethnic group and state*ethnic group. Conceptually, this regression contrasts the change in outcomes over time among individuals from a given state of two different ethnic groups, or of the same ethnic group but in two different states.

3.2 Instrumentation

The marriage market size and sex ratio may be endogenous as it is mostly influenced by migratory flows. While we may think that the sex ratio of second generations is mostly biologically fixed at around 1, immigrants potentially select their state of residence based on labor and marriage market conditions. Female immigrants may choose to immigrate to a state where women's bargaining power is larger. Males may elect locations where there are good work opportunities. Since these factors influence choices made by second generation Americans, it introduces a bias in the interpretation of $\hat{\alpha}$ as the causal effect of the sex ratio.

To alleviate this problem, I construct an instrument in a similar spirit as that of Card (2001). Since individuals from the same country of origin tend to form networks, they also tend to migrate to similar locations (Munshi 2003). Past location choices are thus a good predictor of future immigration decisions in this cohort as discussed in Lafortune and Tessada (2011). As long as past waves of immigrants did not select the state of migration based on future marriage market conditions for their children, using these shares provide an exogenous source of variation. The first stage will capture how much of the changes in the total foreign stock characteristics can be explained by shocks driven *only* by immigrants, which are likely to be more exogenous to local conditions.

More precisely, two instruments were constructed as follows. All male and female immi-

grants were allocated separately to a given state for each period and country of origin based on the 1900 concentration of that country in that particular state. If 10 percent of all Norwegians were located in Minnesota in 1900, 10 percent of all men and 10 percent of all women immigrants from Norway arriving after 1900 are assigned to Minnesota. This generates a predicted flow of males and females by country of birth. Summing for all countries within an ethnic group, one obtains a measure of the predicted flow of immigrant of each gender for each state, ethnic group, and immigration period cell. The instrument for the flow of immigrants of a given ethnic group is then obtained by adding the predicted flow of males to that of females. The sex ratio instrument is built by dividing the predicted flow of males by that of females. Equations (3) and (4) define formally the instruments for an ethnic group j (which includes many countries of birth l), in state s and in time t:

$$\frac{\hat{N^m_{jst}}}{N^f_{jst}} = \frac{\sum\limits_{l \in j} \left(\frac{I_{ls1900}}{I_l} \right) * I^m_{lt}}{\sum\limits_{l \in j} \left(\frac{I_{ls1900}}{I_l} \right) * I^f_{lt}} \tag{3}$$

$$\hat{N_{jst}} = \sum_{l \in j} \left(\frac{I_{ls1900}}{I_l} \right) * \left(I^m_{lt} + I^f_{lt} \right). \tag{4}$$

In Card (2001), multiple countries of birth were needed to build the instrument for each state-year and for a very similar reason, multiple countries of birth are needed to build the instrument for each state-ethnic group-cohort cell. The issue at hand in this case is that the location choice of immigrants in 1900 is not gender-specific. Thus, if one used only one country of birth or if all countries of birth within one ethnic group selected the same locations, the predicted sex ratio of each state would be *exactly* the sex ratio of immigrants who arrived to the United States for that cohort-ethnic group and would thus be co-linear with the fixed effects for that double interaction. The instrument exploits the interaction between variation in the sex ratio of immigrants by country of birth within an ethnic group and differences in past location shares. In short, this instrumental variable strategy assumes that individuals tend to locate where their fellow countrymen live but marry within the entire ethnic group.[14] The regressions below were performed with slight differences in the allocation of countries

[14]The fact that the instrument clearly predicts changes in the stock and the sex ratio indicates the validity of the first part of this assumption. The second is explored using data from an older cohort in Appendix Table C.1. An individual of a given ethnicity was more likely to marry individuals of their ethnic group (but not of their own country of birth) relative to the prevalence of that ethnicity in the overall population in a majority of groups, in particular in those with large representations. Furthermore, in the cases where it is not, alternative pairings were considered without much better results.

of birth to ethnic group or by excluding the ethnic groups that were the least likely to match across countries of birth with very similar results.

The strategy can best be illustrated by an example using the Scandinavian ethnic group in two key states: Illinois and Wisconsin. In 1900, Illinois had 10.2 percent of the Danes, only 1.3 percent of the Finnish, 8.9 percent of the Norwegians but 17.3 percent of all Swedes. Wisconsin, on the other hand, had a similar fraction of the Danes (10.5 percent), slightly more Finnish (3.5), a much larger share of the Norwegians (18.2) and only 4.6 percent of the Swedes. Because Illinois had a high concentration of Swedes in 1900, the evolution of its predicted sex ratio is highly influenced by the changes in the sex ratio of Sweden immigrants. On the other hand, Wisconsin follows much more closely that of Norwegians and Danes. Had I only used Swedes to construct the instrument, Illinois and Wisconsin would be predicted to have *exactly* the same sex ratio for that group as it would correspond to the Swedes sex ratio at the national level.

This identification strategy relies on one key assumption: that immigrants before 1900 did not select these locations because they anticipated the changes in marriage and labor market conditions for that particular ethnicity after 1900. This assumption will not be violated if immigrants select locations that were more attractive for their ethnic group before 1900 but remained similarly attractive over the next 30 years. It will also not be violated if immigrants anticipated shocks for their ethnicity that were short-lived so that by 1905, no remnants of these shocks were found. Finally, it would not be violated if pre-1900 immigrants selected states in anticipation of better conditions for all ethnic groups (or if their arrival improved conditions for all ethnicities) but not particularly for their particular ethnic group because regressions control for state-time fixed effects. A valid concern, however, arises, as is the case in this literature, if immigrants are so concentrated in one geographical area that local conditions affect the gender composition of immigrants from a given country. The fact that no ethnic group, except for Hispanics, has more than 35 percent of previous immigrants locating in one market should diminish these concerns.

In addition, it must be the case that, once controlling for the total number of immigrants, no other characteristics of the immigrants change at the same time as the sex ratio by location. This would be violated if, when more men than women enter the United States, these men tend to be of lower/better quality. There exists little information on the quality of immigrants to test whether this is violated, except for immigrants' literacy as measured by the Census. No correlation was found between literacy rates for either gender nor the actual or the instrumented sex ratio of immigrants.

3.3 Data

3.3.1 Marriage market measures

Sex ratios and immigrant flows (and foreign stock) for each time period, state and ethnic group, as well as the national flows by gender (I_{jt}^m and I_{jt}^f), were constructed using the IPUMS files for 1910, 1920 and 1930. Location shares were obtained from the 1900 Census tables (United States Census Office 1901).[15]

Immigrants were concentrated in some key states with 72 percent of all immigrants locating in 10 states in 1900–see Appendix Table C.2. The location of immigrants varied importantly by ethnic group and the relative concentration of ethnicities also varied from the most concentrated (Hispanics, with 94 percent living in 10 top states) to the least concentrated (British ancestry, at 75 percent).

More importantly for the instrument is the variation in location choice across countries of birth and within ethnicity. For example, among those of British ancestry, English Canadians located mostly in Massachusetts and in Michigan while Australians elected California, the Welsh primarily settled in Pennsylvania and the Irish, in New York and Massachusetts. Even between Poles and Russians, where the same three states are preferred locations (New York, Pennsylvania and Illinois), the Poles were distributed equally across the three states while the Russians were much more concentrated in New York.

Table 1 summarizes the distribution of the endogenous variables and of the instruments. The sex ratios of Others (mostly Asians) and Italians were among the highest while the Francophone and those of British ascendance had close to a balanced sex ratio. The sex ratios of the total foreign stock are more balanced but the same differences across ethnic groups emerge. What is worth noting is that variations in the sex ratios are large, even when looking at the foreign stock. Ethnic groups within a state experienced substantial changes in the sex ratio over the time period studied: the average standard deviation within a group and state was about 0.2.

The last columns of that table present the average prevalence of endogamy within ethnicities that are part of the ethnic group, constructed using data from cohorts born between 1865 and 1884. This is an imperfect measure of the desire of an ethnic group to marry within their group as marriage market conditions in this period may have affected the matching possibilities. In order to attenuate this problem, the measure subtracts from the observed endogamy rate the prevalence of one's ethnic group among all potential mates in 1900. This rate thus

[15]Because these shares are computed using the full population of immigrants and not just a small public-use sample, they are robust to the "small cell bias" as argued by Aydemir and Borjas (2011).

reflects how much one ethnicity marries within one's ethnic group, above and beyond if it randomly matched with a spouse in the general marriage pool. It is defined separately for males and females. In general, males have lower rates of endogamy than females (compared to the relative abundance of potential partners of their ethnic group). There is also ample variation across ethnic groups with those of British Ancestry and Francophone origin having the lowest rates while Southern European women and males of other ethnicities having the largest.

3.3.2 Outcome measures

All outcomes are obtained from IPUMS files between 1900 and 1970 and are presented with a detailed description in Appendix Table C.3.[16]

First, various proxies for pre-marital investments were collected. The preferred measure is the highest grade attained by an individual. This measure is continuous and is often used as a measure of human capital accumulation. However, it is only available after 1940 in the US Census and thus a problem would arise if schooling was mostly acquired in our sample after marriage. This is, however, unlikely to be the case: this sample has an average schooling level below high school completion (9.5 years for females and for males) and the average age at first marriage is 23 for females and 27 for males. Also, while 22 percent of the individuals aged 15-25 attend school, only 1 percent of the married males and 3 percent of the married females report being in school between 1900 and 1930. A dummy indicating whether the individual attended high school will also be used to capture the relevant margin of adjustment. Pre-marital investment could also be obtained through on-the-job training. To try to capture that margin, an occupational scoring measure is employed for the occupation of young adults (15-25), where the score is based on the average educational attainment of individuals holding an occupation in 1950 (similar results were obtained with other occupational scores). Literacy is also included: it should be acquired before marriage and could affect post-matching output (see Behrman et al. 1999 for an example in India). However, given that it is usually acquired at very young ages and that the shock to the marriage market produced by the immigrants occurs when individuals are in their teens and early twenties, it will be more presented as a robustness check. Finally, a delay in marriage could be indicative of a higher level of pre-marital investment and thus the age at first marriage is also measured. To alleviate the problem of

[16]Multiple years of the census were used because some variables were only available in particular years. Because this implies that some cohort were more likely to be answering some questions more often than others, age restrictions were imposed in order to obtain a "representative" sample of each cohort of second generation individuals.

sample selection (age is only measured if one is already married), this variable is restricted to individuals older than 35, for whom most first unions have already been entered into.

I also explore how matching patterns may have been altered by the change in the sex ratio. Marital status, measures of marital stability (divorce rates and number of marriages) and country of birth of one's spouse are all obtained for each individual in the sample. Unfortunately, ethnicity of spouse's parents is not available in either 1940 or 1950 so it is difficult to classify spouses as second generation Americans of a particular group and thus measure this broader definition of endogamy. Finally, spousal education is also included to measure whether the quality of spouses has changed by altering the sex ratio.[17]

To measure post-matching outcomes, this paper uses labor supply of all individuals aged 25 and above. I do not restrict the analysis to married individuals because this would potentially introduce selection bias. For all individuals, a variable indicating labor force participation and employment is available. In addition, measures of weeks worked last year and hours worked last week are included. Obviously, labor supply is a very crude and potentially inadequate measure of ex-post bargaining power but it has been used previously (Angrist 2002 and Chiappori et al. 2002, for example) and is the only proxy available in Census data.

Table 2 gives the main summary statistics for each outcome. The rate of non-marriage is much above that of natives.[18] Divorce rates are low overall. About 7 percent of second generation women and 4 percent of males are currently married to first generation immigrants from their ethnic group (which accounts for most marriages between second generation and immigrants). This is somewhat low but the denominator includes all singles and widows, as discussed above. This is also lower than the total endogamous marriages which include all marriages within second-generation individuals as well. Literacy is very high among second generation Americans (close to 99 percent of them are literate) but varies considerably across ethnicities, with non-European groups having much lower levels. Men and women are both achieving about the same level of schooling (9.5 years) and if anything, women are more educated than men. This is a fact that holds for natives as well. There is substantial variation in the educational attainment of individuals: a standard deviation of about 3 years of education for both males and females. Labor supply attachment by woman is quite low; slightly more than 30 percent of women were in the labor force compared to 88 percent for men.

[17]This variable is only available in 1940 or 1970 and all regressions will limit themselves to 1940 since only long-lasting marriages would still be observed in 1970.

[18]Previous studies have noticed that second generation immigrants have the lowest rate of marriage (Groves and Ogburn 1928, Haines 1996 and Landale and Tolnay 1993). This is surprising since they should be able to participate in two marriage markets (endogamous and native). This may indicate a distaste from natives to marry non-natives.

4 Results

4.1 First stage

The instruments are very highly predictive of their respective endogenous variable, as shown in Table 3. This table represents the first stage performed on the sample used for the highest grade ever attained for males in the top panel and females in the bottom one.[19] Extremely similar results were obtained for other samples used in this paper and are thus omitted to save space. An increase of one in the predicted flow leads to an increase of about 0.96 in the actual number of immigrants arriving over that period. Similarly, an increase of one in the predicted sex ratio measure is linked to an increase of about the same magnitude among actual immigrants. The instrument can very precisely predict shocks to the number and sex ratio of the foreign stock (immigrants and second generation individuals) although the magnitudes are normally much lower. A predicted increase of one in the number of immigrants translates into an actual increase in the size of the foreign stock of about 0.2. An increase of 1 in the predicted sex ratio of immigrants also increases that of the sex ratio of the foreign stock by about 0.2. Since this paper's interest lies in shocks to the overall marriage market, columns (3) and (4), where the endogenous variables are measured for the entire foreign stock, will be the ones used for the rest of the paper. These are all significant at 1 percent level although slightly less precisely estimated in the females' sample. Given that two endogenous variables and two instruments are used here, the partial R-square measure proposed by Shea (1997) indicate whether only one of the instrument is driving the first stages. For the immigrant marriage market measures, I obtained a partial R-square of about 5 percent for the sex ratio instrument and of about 86 percent for the stock instrument. The statistic falls to about 1 to 2 percent for both instruments once the endogenous variables refer to the entire foreign stock, not because the correlation between the two instruments increases but rather because the influence of the instruments on the endogenous variables decreases compared to all the fixed effects included in the regressions. The joint F-tests are, in general, very high for the stock regressions and a little bit less so for the sex ratio measures.[20] Since the model is just identified, the 2SLS estimates are median-unbiased. Furthermore, the reduced form estimates (which are unbiased), not presented here but available from the author, mirror the IV results, thus suggesting little problems stemming from weak instruments.

[19] The difference in samples simply influences the weight given to each ethnic group-state-cohort cell.

[20] Given the fact that the standard errors are clustered, the standard Cragg-Donald statistic is incorrect and Cameron and Trivedi (2010) (p.193) recommend the use of the joint F-test and the Stock-Yogo critical values, which in this case indicate that for we would at most be facing a distortion of 5-7% in our 5% inferences for the sex ratio regressions and much less than 5% for the stock regressions.

The robustness of the first stage is tested through various specifications presented in Table 4. This table includes only those where the sample is that of males but very similar results were found in the female sample as well. To verify that the share of immigrants not only predicts the behavior of immigrants shortly after 1900, Column (1) ignores the first two periods of immigration. Column (2) restricts itself to the first four periods of immigration. Column (3) excludes the largest immigration-receiving state over this period (New York) while column (4) excludes any cell where the sex ratio was imputed given that no males or females could be found. Column (5) presents the same regression but where both the instruments and the endogenous variables are expressed in log terms. The predictive power of the instrument for the sex ratio is very robust across specifications. That of the stock is very sensitive to the exclusion of the state of New York: once that state is excluded, the significance of the instrument falls to just below 5 percent. Finally, although not shown, even if the relationship between the instrument and the actual marriage market measure is stronger for some ethnic group than for others, removing any ethnic group does not alter the significance of the first stage. Overall, the first stage appears to be fairly robust and not driven by a few outliers.

4.2 Pre-marital investments

The main regressions of interest relate pre-marital investments to marriage market conditions. Table 5 first presents correlations obtained from an OLS regression of the marriage market measures and the proxies for pre-marital investments. All the correlations with the sex ratio are close to 0 in magnitude and certainly undistinguishable from zero given the standard errors, except for those related to the educational score of one's occupation which is positive and significant for males.

One would expect the OLS results to be biased towards zero if immigrants elect locations where they have more bargaining power and more mating possibilities. Then, locations with a larger number of male immigrants are also those in which second generation men have more benefits and thus less incentive to invest in human capital. There is weak evidence that this is the case, as immigrant men elect states where there are more second generation females and fewer second generation males of their own ethnic group. Also, the correlation between the sex ratio and the probability that a female marries an immigrant of her ethnic group is fairly strong which may be indicative that immigrants select locations where the marriage prospects are good.[21]

Once one uses the instruments, the coefficients are larger in magnitude for males and now

[21] These results are not shown but available upon request.

imply that a higher sex ratio leads males to significantly increase their acquisition of human capital. The results are very significant for the highest grade attained and the probability of attending high school. They suggest that a change in the sex ratio of one's marriage market from a balanced level ($z = 1$) to one where there are 20 percent more males than females ($z = 1.2$), which corresponds to the average standard deviation in the sample, would lead to an increase in the number of years of education obtained by a male of about 0.24 and increase the probability of males attending high school by 4 percent. Little evidence is found here that a higher sex ratio increased male literacy, which is logical since immigrants are arriving when the second generation males are tool old for them to adjust their literacy. Column (5) clearly indicates that this increase in human capital is not acquired through a simple delay of the age at first marriage.

The IV results for females are much less precisely estimated and very little can be said about how the sex ratio influenced the pre-marital investments of females. The point estimates are positive, but we would not be able to reject the possibility that they would be negative given the confidence intervals.

Marriage market size appears to lead males to select much higher paying occupations. This is surprising, but may reflect the fact that immigrants fill low-paid occupations and push second generation Americans to higher-paying ones. It also appears to hurry female's marriages but delay that of males. This may indicate that when one's preferred marriage market expands, it may be easier to select optimally the timing of one's marriage. Thus, if females prefer marrying earlier than men, this could explain these results. Overall, omitting the stock measure usually renders the effect of the sex ratio more significant.

The bottom panel of Table 5 explores whether the estimated coefficients depend on how endogamous the individual's ethnicity was, as described in the last paragraph of Section 3.3.1. The endogamy measure is demeaned to facilitate the interpretation of the coefficients and the interaction term is instrumented by the interaction of the instrument for the sex ratio and the endogamy measure. The results presented in this panel suggest that males increase their educational investment in response to a rise in the sex ratio irrespective of the endogamy prevalence in their group: the interaction terms are small and insignificant in all cases. Females, on the other hand, appear to only decrease their educational score, literacy rate and age at first marriage when their ethnicity has a high endogamy rate. Note, however, that the variance in the endogamy rate is limited by the fact that all groups marry well above what random matching would predict. The difference between the responses of males and females may be due to the fact that males are more likely to seek a partner inside their ethnic group than females (compared to what a random matching process would generate) and that the variance

in the endogamy rate of males between ethnic groups is less for males than for females.

Yet, one may worry that the effect estimated above may only correspond to a shock to the labor market contemporaneous to the immigrant flow predicted by the instrument. In order to check whether this is the case, I introduce two additional endogenous variables to the above regression, measuring the sex ratio and the flow of immigrants who were older at the time of their arrival than the ones used in the above marriage market measure (older than 27 for women and 30 for males). These would have impacted the labor market outcomes of the second generation individuals in the sample but should have not been part of the same marriage market. Those endogenous variables are instrumented using exactly the same instruments as described by equations (3) and (4) but using the older cohorts of immigrants as the national flow. Table 6 explores these results. The introduction of these variables does little to change the results of the previous table. Women's results are still imprecisely estimated and positive. The effect of the sex ratio on male's education remains positive for the highest grade attained. In no case do we see the older generation's sex ratio measures significantly impacting the pre-marital investments of the second generation.

One may wonder if what the results really capture is the discontinuous effect of z around 1 or if smaller variations in z above or below 1 may also generate the patterns presented above. The first four columns of Appendix Table C.4 explore this by restricting the sample only to cells that have a predicted sex ratio larger than 1.1. The results, if anything, suggest that restricting oneself to this sample strengthens the conclusions drawn previously. In particular, in this setting, females appear to decrease their pre-marital investments in response to higher sex ratios, although this effect is still not statistically significant.

The work of Furtado (2012) and the model above seem to indicate that the impact of the sex ratio could depend on how much less educated are one's ethnic group compared to natives. This is explored in Columns (5) to (8) of Appendix Table C.4 but few of the interaction coefficients are significant and, for those that are, wrongly signed.[22] Finally, one may expect that the sex ratio within one's preferred marriage market may have a differential impact depending on how stable are unions formed within that group. The last columns of the table explore this by interacting the impact of the sex ratio with the prevalence of divorce in a previous cohort by country of birth. The results indicate that the sex ratio increases more significantly one's pre-marital investment when one's ethnicity is more prone to divorce.

Although not presented here, variants of the instrument were explored with similar results. The results are almost identical if immigrants are matched to second generation based

[22]The difference in educational attainment is constructed by subtracting the average native educational attainment from the average educational attainment in one's ethnic group for individuals of an older cohort (born between 1865 and 1884).

on their year of birth rather than on the moment of their arrival. Although gender-specific shares by country of birth were not available from the Census tables, overall immigrant sex ratios by state were obtained. If one allocates immigrants based on the interaction between the fraction of all immigrants of a given gender electing a given state and the share of one's ethnicity electing that same state, the results are very similar to the ones presented above. However, this last instrument would be more likely to violate the exclusion restriction and was thus not employed here.[23]

4.3 Matching patterns

Changing marriage market conditions should also affect matching patterns as suggested by the theoretical framework presented above. These outcomes are explored in Table 7. Panels A and C only include the two main regressors of interest for females and males while Panels B and D allow the effect of the sex ratio to differ by the rate of endogamy among one's ethnicity, as defined in the previous section.

Surprisingly, the first column indicates that both men and women are more likely to have ever been married when the sex ratio rises (although this is similar to Angrist 2002) while the model above would have predicted no change in the probability of matching, just in the type of partner one would encounter. The probability of a man ever having been married rises by 0.7 percent when the sex ratio among his ethnic group goes from being balanced to one where there is 10 percent more males. The coefficient is smaller and insignificant for women. The next two columns shed some additional light on this result. Men are also more likely to be divorced (significant at the 10 percent level) and more likely to have been married more than once (the opposite is true for women but the estimates are not significant). The rise in the probability of having been married more than once is comparable in size with the effect of the sex ratio on the rate of marriage. Thus, the effect of the increased sex ratio among this population appears to be more linked to marital stability than to the formation of relationships. Furthermore, if males seek partnerships without marriage and can obtain these when they have more bargaining power, this could also explain these results. The absence of significant results for females is more consistent with the model but could also be related to the fact that one single marriage was slightly more common for females than for males in the sample and that the standard errors for females are larger.

[23]Replications of Angrist (2002) strategy were also conducted with results indicating that both males and females responded positively to a larger sex ratio at the national level within their ethnic group. This puzzling result may be due to the fact that the instrument used in that case may be more likely to capture national factors affecting all individuals of an ethnic group while the strategy mentioned above exploits cross-state variations and nation-wide trends at the ethnic levels are here captured by a fixed effect.

Another reason why this result could arise is if the native marriage market values the investment made by males in their education. Column (4) of Table 7 indicates that men are less likely to marry an immigrant of their own ethnicity when the sex ratio is higher, although the estimate is only significant at the 10 percent level. A change from a balanced sex ratio to one where there are 10 percent more males than females decreases the probability of a man to be married to a female immigrant of his own ethnicity by 0.4 percent. The effect is smaller and imprecisely estimated for females. The fact that males turn to the native market as argued in the model is also visible in the last two columns of the table where the highest grade attained by one's spouse (0 if single) and the probability of being married to a spouse with at least some high school are shown to increase significantly for males when faced with a higher sex ratio [24]. Females, on the other hand, see the educational investment of their partner decreases when faced with a higher sex ratio, although not significantly so.

The effect of the sex ratio is then allowed to differ by the endogamy rate as shown in Panels B and D of Table 7. The effect of the sex ratio on males' and females' propensity to marry is shown to be much larger in ethnicities that are more endogamous. Only females in high endogamy groups are more likely to marry an immigrant of their own ethnic group when the sex ratio of these immigrants increase. Finally, male spousal educational attainment is particularly increased when the sex ratio rises when the endogamy rate is higher, which is consistent with these males being less likely to consider marrying a native unless forced to.

4.4 Labor supply

Finally, this section explores how the sex ratio may affect the bargaining power within the household once matched. As proxies for this, labor supply variables are used as has been done previously in the literature. Those are not, by any means, the best measure of ex-post bargaining power one could measure but they are the only ones available in the Census data. If leisure is a normal good, one would expect that as they acquire more bargaining power within their couple, their labor supply would decrease.

The results of these regressions are presented in Table 8, where Panel A presents the OLS results and Panel B, those using the instrumental variable approach. The OLS coefficients are very small and imprecisely estimated. This could be either an overestimate or an underestimate of the real causal effect. It would be an overestimate if male immigrants select to locate in states where the labor market is booming. On the other hand, if men tend to locate in areas where they have more bargaining power, they would select locations where males are

[24] This result does not stand if one runs these regressions conditional on being married

working less and the OLS would be a lower bound on the magnitude of the causal estimate.

Panel B of Table 8 presents the results of the instrumental variable regressions. The causal effect of the sex ratio appears to lead women to reduce their labor supply while little effect is seen for males except in the case of weeks worked where it is positive but not significant. These results indicate that going from a balanced sex ratio to one where there are 10 percent more males than females lead women to be 1 percent less likely to be in the labor force and 1.5 percent less likely to be employed, in line with the estimates provided by Angrist (2002). A rise in the sex ratio from a balanced level to one where men are 10 percent more numerous than women reduces hours worked per week and the number of weeks per year by about 0.6. The result for hours worked is only significant at the 10 percent level. For males, a similar increase in the sex ratio leads to no effect for either employment, labor force participation, or hours worked per week and only increases the number of weeks worked per year by 0.2 (but this is not significantly different from zero). The OLS results are usually smaller in absolute value than the IV as expected if immigrants locate based on conditions of the marriage market. The difference in the responses of males and females could be due to the fact that men's labor supply was more inelastic than females over this period. However, it could also be related to the fact that males appear to have adjusted their educational attainment more than females, something explored below.

Contrarily to previous works that have employed these types of outcomes but following the above framework, this paper argues that these results should not necessarily indicate that a higher sex ratio leads to a change in relative spousal ex-post bargaining power favoring women if partners have the time to readjust their pre-marital behavior. Labor supply may be particularly influenced by pre-marital investments made by spouses or by changes in matching patterns. As the model suggested above, changes in the sex ratio could translate into changes in pre-marital investments even if they do not affect ex-post bargaining power and thus impact labor supply through that channel only.

Ideally, one would like to add education and any other form of pre-marital investments that can influence labor supply in the regressions presented in Table 8. However, education would clearly be endogenous in this context and given that the previous model was just identified, no instruments are available. Alternatively, I here attempt to directly estimate the impact of education on labor supply and then use that estimate to compute how the impact of the sex ratio on labor supply would change when controlling for education. Note that economic theory does not predict whether education increases or decreases labor supply. The income effect decreases labor supply. On the other hand, the substitution effect increases the number of hours spent working.

To isolate the effect of education on labor supply in this population, I use compulsory schooling laws as tabulated by Lleras-Muney (2002) as instrument for education in a sample of individuals born between 1900 and 1924, a slightly younger cohort than the one studied above.[25] Labor supply and education are measured in the 1940-1970 IPUMS files and the sample is restricted to native-born males and females older than 25. The results presented in Appendix Table C.5 use as instruments a set of dummy variables for each minimum number of years of schooling required by the state.[26] The first stage suggests that each additional year of compulsory schooling leads men to increase their level of schooling by about 0.04 years and women to do so by about 0.06 years.[27] The OLS estimates are, as expected, positive and large suggesting that individuals who acquire more education are also more likely to be working and to do so for longer hours. The IV estimates, on the other hand, suggest that education decreases labor supply of males, whether measured in terms of labor force participation rates or hours worked. The estimates are fairly large suggesting that one more year of education reduces hours worked per week by 1.5 hours for males. Among females, the results are imprecisely estimated.

The next set of regressions attempts to measure the overall effect of both spouses' education. It is restricted to married individuals for that reason. The instrument is based on the compulsory schooling that affected each spouse in his or her state of birth. Two caveats must be mentioned. First, the first stages are much weaker in this context than before, simply because there are a few spouses who were subject to different compulsory schooling laws (since individuals tend to marry within their state and within a relatively close age cohort). The compulsory schooling laws affecting females tend to be a better predictor of the education of both spouses. Second, even if both educational levels are instrumented, this regression does not control for the potential endogeneity of the match. Nevertheless, these results are presented as a robustness check on the previous estimates.[28] They suggest that for both genders, one's own education decreases labor supply while that of one's spouse tends to attenuate this effect.

Combining these estimates with the ones from section 4.2, an increase of 10 percent of the sex ratio from a balanced level (through the educational channel alone), decreases the number

[25]One could worry that this younger cohort differs substantially from the one used in the analysis above. Even if this is the case, the difference between the cohorts should be orthogonal to the variation in the sex ratio employed in the main analysis since all above regressions used state-cohort fixed effects that should be capturing all of this variation.

[26]Similar results, although slightly less precise, were obtained by using a continuous measure of the minimum number of years of schooling.

[27]This is very similar to the first stage presented jointly for both genders by Lleras-Muney (2002).

[28]This is very similar to the argument made by Boulier and Rosenzweig (1984) that unobservables involved in the matching process cannot be assumed to be orthogonal to the educational level of spouses.

of hours worked by males by about 0.15 hours per week.[29] This suggests that the effect of the sex ratio on labor supply obtained in Table 8 is underestimating the true impact of the sex ratio on post-matching outcomes for males.

5 Conclusions

This paper first presented a simple model to understand the reasons why changes in sex ratios could translate into changes in pre-marital investments, highlighting two main channels: matching and bargaining. It then documents that males increased their human capital investment when faced with a plausibly exogenous increase in the number of males per females in their state-ethnic-cohort marriage market. While the magnitude of the estimated effects are not extremely large, the magnitudes are still economically relevant: within a given ethnic group and state, an increase in the sex ratio by an average standard deviation would lead males to increase their years of education by about 0.2 and their likelihood of attending high school by 4 percent. Furthermore, the type of investment that is considered here is very costly and requires a fairly long planning horizon. One could thus think that for simpler forms of pre-marital investments which were unobserved in this context (such as improving one's looks, learning how to cook or dance, etc), the effect may even be more relevant.

Thus, this suggests that when undertaking educational decisions, labor market prospects are not the only element taken into account by individuals. The importance of incentives linked to returns received once matched may partially explain why the educational gap by gender is not always correlated with either differences in labor force attachment or differences in wages between men and women. This is particularly relevant in today's world where, despite the fact that women still face lower wages and higher probability of child-related work interruptions, they are now acquiring often more education than men. The demographic changes that have been seen recently in many countries (increase in cohabitation, decreases in marriage rates, increase in non-marital births, increase in marriage age, etc) could thus heavily influence educational investments of current cohorts and not just be driven by them as is often argued.

In addition, the magnitude of the educational shifts found in this study combined with estimates of the effect of education on labor supply suggest that the interpretation of the effect of marriage market conditions on post-marital outcomes is difficult and cautions against the use of such methods as a test of the unitary framework.

[29] A similar range of values would be given for males if using the effect of both own and spousal education on labor supply decisions.

Finally, the conclusions of this paper also suggest that our understanding of the household would be enhanced by a more careful analysis of how marriage market conditions may affect both the process of household formation and pre-marital decisions as well as post-marital outcomes. For example, while divorce laws have been previously envisaged as strong determinants of ex-post bargaining power within the household, little is known about how these may modify matching patterns and other decisions undertaken before the union is formed. More research is warranted.

6 References

Abramitzky, Ran, D. A. and L. Vasconcelos (2011). Marrying Up: The Role of Sex Ratio in Assortative Matching. *American Economic Journal: Applied Economics 3*(3), 124–57.

Amuedo-Dorantes, C. and S. Grossbard (2007). Cohort-level sex ratio effects on women's labor force participation. *Review of Economics of the Household 5*(3), 249–278.

Angrist, J. (2002). How do sex ratios affect marriage and labor markets? Evidence from America's second generation. *The Quarterly Journal of Economics 117*(3), 997–1038.

Aydemir, A. and G. Borjas (2011). Attenuation bias in measuring the wage impact of immigration. *Journal of Labor Economics 21*(1), 69–113.

Behrman, J. R., A. D. Foster, M. R. Rosenzweig, and P. Vashishtha (1999). Women's schooling, home teaching, and economic growth. *The Journal of Political Economy 107*(4), 682–714.

Boulier, B. L. and M. R. Rosenzweig (1984). Schooling, search, and spouse selection: Testing economic theories of marriage and household behavior. *The Journal of Political Economy 92*(4), 712–732.

Brainerd, E. (2006). Uncounted costs of World War II: The effect of changing sex ratios on marriage and fertility of Russian women. http://urban.hunter.cuny.edu/RePEc/seminar/old/rfwomen.pdf.

Browning, M., P.-A. Chiappori, and Y. Weiss (2010). *Family Economics*. in progress.

Cameron, A. C. and P. K. Trivedi (2010). *Microeconometrics Using Stata, Revised Edition*. StataCorp LP.

Card, D. (2001). Immigrant inflows, native outflows, and the local market impacts of higher immigration. *Journal of Labor Economics 19*(1), 22–64.

Chiappori, P.-A., B. Fortin, and G. Lacroix (2002). Marriage market, divorce legislation, and household labor supply. *The Journal of Political Economy 110*(1), 37–72.

Chiappori, P.-A., M. Iyigun, and Y. Weiss (2009). Investment in schooling and the marriage market. *American Economic Review 99*(5), 1689–1713.

Edlund, L., H. Li, J. Yi, and J. Zhang (2007). Sex ratios and crime: Evidence from china. IZA Discussion paper No. 3214.

Foster, A. and M. Rosenzweig (2001). Missing women, the marriage market and economic growth. http://adfdell.pstc.brown.edu/papers/sex.pdf.

Foster, A. D. and M. R. Rosenzweig (1996). Technical change and human-capital returns and investments: Evidence from the green revolution. *American Economic Review 86*(4), 931–53.

Furtado, D. (2012). Human capital and interethnic marriage decisions. *Economic Inquiry 50*(1), 82–93.

Furtado, D. and N. Theodoropoulos (2011). Interethnic marriage: A choice between ethnic and educational similarities. *Journal of Population Economics 24*, 1257–1279.

Grossbard-Shechtman, A. (1984). A theory of allocation of time in markets for labour and marriage. *The Economic Journal 94*(376), 863–882.

Grossman, S. J. and O. D. Hart (1980). Takeover bids, the free-rider problem, and the theory of the corporation. *The Bell Journal of Economics 11*(1), 42–64.

Groves, E. R. and W. F. Ogburn (1928). *American Marriage and Family Relationships*. Henry Holt & Co.

Haines, M. R. (1996). Long term marriage patterns in the United States from colonial times to the present. NBER Historical Working Papers 0080.

Iyigun, M. and R. P. Walsh (2007). Building the family nest: Premarital investments, marriage markets, and spousal allocations. *Review of Economic Studies 74*(2), 507–535.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics 125*(2), 515–548.

Konrad, K. A. and K. E. Lommerud (2000). The bargaining family revisited. *Canadian Journal of Economics 33*(2), 471–487.

Kvasnicka, M. and D. Bethmann (2012). World war II, missing men, and out-of-wedlock childbearing. *The Economic Journal* (accepted).

Lafortune, J. and J. Tessada (2011). Smooth(er) landing? the dynamic role of networks in the location and occupational choice of immigrants. http://sites.google.com/site/jeannelafortune/papers/Lafortune-Tessada-networksimmigration.pdf.

Landale, N. S. and S. E. Tolnay (1993). Generation, ethnicity, and marriage: Historical patterns in the Northern United States. *Demography 30*(1), 103–126.

Lleras-Muney, A. (2002). Were compulsory attendance and child labor laws effective? An analysis from 1915 to 1939. *Journal of Law & Economics 45*(2), 401–35.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the U.S. labor market. *The Quarterly Journal of Economics 118*(2), 549–599.

Nguyen, T. (2008). Information, role models and perceived returns to education: Experimental evidence from Madagascar. http://www.povertyactionlab.org/sites/default/files/documents/Nguyenpdf.

Oreffice, S. and B. Bercea (2008). Quality of available mates, education and household labor supply. *Economic Inquiry*, 1–17.

Pagnini, D. L. and S. P. Morgan (1990). Intermarriage and social distance among U.S. immigrants at the turn of the century. *The American Journal of Sociology 96*(2), 405–432.

Peña, X. (2006). Assortative matching and the education gap. http://sites.google.com/site/ximenapena/matching.pdf.

Porter, M. (2007). Imbalance in China's marriage market and its effect on household bargaining. Unpublished.

Ruggles, S., M. Sobek, T. Alexander, C. A. Fitch, R. Goeken, P. K. Hall, M. King, and C. Ronnander (2009). *Integrated Public Use Microdata Series: Version 4.0 [Machine-readable database]*. Minneapolis, MN: Minnesota Population Center.

Shea, J. (1997). Instrument relevance in multivariate linear models: A simple measure. *The Review of Economics and Statistics 79*(2), 348–352.

Shimer, R. and L. Smith (2000, March). Assortative matching and search. *Econometrica 68*(2), 343–370.

United States Census Office (1901). *Census Reports, Twelfth Census of the United States 1900*, Volume 1. United States Census Office.

7 Tables and figures

Table 1: **Summary statistics-Marriage market conditions and instrument**

Ethnic group	**Immigrants**		**Foreign stock**		**Endogamy**	
	Number (000s)	Sex ratio	Number (000s)	Sex ratio	Males	Females
British Ancestry	0.162	1.012	0.686	0.955	0.249	0.305
	[0.159]	[0.536]	[0.514]	[0.146]		
Francophone	0.033	0.985	0.136	0.956	0.287	0.304
	[0.033]	[0.711]	[0.116]	[0.354]		
South Europeans	0.256	1.591	0.685	1.371	0.362	0.615
	[0.277]	[1.061]	[0.630]	[0.796]		
Hispanics	0.152	1.352	0.302	1.266	0.567	0.592
	[0.117]	[1.079]	[0.242]	[0.851]		
Scandinavian	0.056	1.994	0.296	1.135	0.465	0.507
	[0.054]	[1.748]	[0.240]	[0.388]		
Germanic	0.184	1.432	0.746	1.008	0.360	0.430
	[0.215]	[0.765]	[0.563]	[0.148]		
Russians and others	0.394	1.105	0.986	1.084	0.532	0.602
	[0.483]	[0.526]	[0.853]	[0.235]		
Other Europe	0.095	1.405	0.249	1.297	0.392	0.498
	[0.095]	[1.213]	[0.200]	[0.803]		
Other countries	0.047	3.120	0.068	2.413	0.666	0.207
	[0.043]	[3.123]	[0.062]	[1.883]		

Standard deviations in brackets. All summary statistics are weighted by the size of the foreign stock in each cell.

Table 2: **Summary statistics-Outcomes**

	Males			Females		
	N. Obs.	Mean	Sd	N. Obs.	Mean	Sd
General characteristics						
Age	231980	39.84	16.18	246880	40.56	16.54
Mother foreign	231980	0.79	0.41	246880	0.78	0.41
Father foreign	231980	0.88	0.32	246880	0.88	0.33
Pre-marital investments						
Years of education	87515	9.32	3.40	97885	9.35	3.13
Some high school	87515	0.46	0.50	97885	0.49	0.50
Educational score	73010	74.43	122.11	75582	77.94	166.01
Literate	73010	0.99	0.10	75582	0.99	0.09
Marital status (after age 35)						
Never married	115754	0.12	0.32	126683	0.11	0.32
Divorced	115754	0.03	0.18	126683	0.03	0.17
Married more than once	62813	0.12	0.32	75423	0.11	0.32
Married to an immigrant of own ethnic group	90789	0.04	0.20	101033	0.07	0.26
Age at first marriage	54527	26.97	6.38	65838	23.32	5.80
Labor supply (after 25)						
Employed	142236	0.82	0.38	153590	0.29	0.46
In the labor force	158970	0.88	0.33	171298	0.30	0.46
Hours last week	102922	35.30	20.59	114628	11.05	18.23
Weeks last year	104519	40.25	18.01	114665	14.72	21.57

All summary statistics are weighted by Census sample-line weights

Table 3: **First stage**

	Immigrants		Foreign stock	
	Sex ratio	Flow	Sex ratio	Stock
	(1)	(2)	(3)	(4)
		Males		
Predicted flow of immigrants	0.043	0.954***	0.001	0.233***
	(0.150)	(0.017)	(0.040)	(0.034)
Predicted sex ratio of immigrants	0.893*	-0.005	0.170***	0.020
	(0.334)	(0.005)	(0.046)	(0.023)
N. Obs	96707	96707	96707	96707
R-squared	0.579	0.984	0.570	0.977
Joint F-test	7.02**	1560.96***	6.21**	30.82***
		Females		
Predicted flow of immigrants	0.021	0.960***	-0.037	0.209***
	(0.152)	(0.018)	(0.060)	(0.033)
Predicted sex ratio of immigrants	1.129**	-0.003	0.211**	0.018
	(0.331)	(0.004)	(0.061)	(0.021)
N. Obs	86683	86683	86683	86683
R-squared	0.572	0.984	0.633	0.976
Joint F-test	4.58*	1745.17***	7.42**	34.74***

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. All regressions are weighted by the Census sample-line weight.
* significant at 5%; ** significant at 1%; *** significant at 0.1%

Table 4: **Robustness checks of the first stage-male sample**

	1910-1929 (1)	1900-1919 (2)	Without NY (3)	No missing (4)	In logs (5)
			Panel A: Sex ratio		
Predicted sex ratio	0.240***	0.452***	0.241***	0.200**	0.276**
	(0.064)	(0.125)	(0.067)	(0.067)	(0.076)
Predicted flow	-0.034	-0.146*	-0.164	0.013	-0.039
	(0.071)	(0.059)	(0.134)	(0.039)	(0.031)
R-squared	0.631	0.805	0.629	0.653	0.624
Joint F-test	7.00**	8.23***	6.79 **	6.24**	6.67**
			Panel B: Stock		
Predicted sex ratio	0.022	-0.062	-0.002	0.019	-0.004
	(0.013)	(0.049)	(0.012)	(0.022)	(0.128)
Predicted flow	0.173**	0.659***	0.360	0.208***	0.307***
	(0.052)	(0.056)	(0.184)	(0.033)	(0.067)
R-squared	0.985	0.987	0.965	0.976	0.991
Joint F-test	5.57**	86.95 ***	3.29*	29.60***	11.49***
N. Obs	70407	46055	71150	86317	86501

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. All regressions are weighted by the Census sample-line weight.
* significant at 5%; ** significant at 1%; *** significant at 0.1%

Table 5: **Pre-marital investments**

	Males					Females				
	Highest grade attained	Some high school	Educ. score	Liter-acy	Age at first marr.	Highest grade attained	Some high school	Educ. score	Liter-acy	Age at first marr.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: OLS										
Sex ratio	-0.059	-0.006	5.718*	0.001	0.059	-0.058	-0.005	1.725	-0.001	0.192
	(0.071)	(0.007)	(2.328)	(0.003)	(0.148)	(0.063)	(0.010)	(4.969)	(0.003)	(0.170)
Stock	-0.249*	-0.046**	11.519*	-0.011	-0.005	-0.189	-0.018	-6.914	0.002	0.750**
	(0.114)	(0.016)	(4.888)	(0.007)	(0.166)	(0.130)	(0.018)	(5.189)	(0.003)	(0.216)
Panel B: IV										
Sex ratio	1.177**	0.195*	9.788	0.031	-1.078	0.567	0.099	10.911	0.072	-1.014
	(0.454)	(0.097)	(22.267)	(0.027)	(1.424)	(0.561)	(0.077)	(23.195)	(0.068)	(1.185)
Stock	0.781	-0.056	130.114*	-0.014	6.558**	-0.233	-0.041	65.657	-0.032	-1.484*
	(0.926)	(0.161)	(51.948)	(0.027)	(2.507)	(0.882)	(0.136)	(34.000)	(0.025)	(0.741)
Panel C: IV with interactions										
Sex ratio	1.195**	0.206*	12.362	0.032	-0.686	0.705	0.131	4.060	0.067	-0.763
	(0.440)	(0.097)	(22.988)	(0.027)	(1.429)	(0.655)	(0.096)	(32.201)	(0.044)	(1.395)
Stock	0.776	-0.058	128.825*	-0.014	6.292**	-0.358	-0.056	52.031	-0.046*	-1.733*
	(0.927)	(0.161)	(52.091)	(0.027)	(2.402)	(0.908)	(0.140)	(33.986)	(0.023)	(0.769)
Sex ratio*	-0.179	-0.062	-18.441	-0.001	-4.277	-1.664	-0.201	-184.245***	-0.197**	-3.805**
endogamy	(0.780)	(0.114)	(30.521)	(0.023)	(3.833)	(1.250)	(0.150)	(45.127)	(0.074)	(1.309)
N	86683	86683	66565	66565	54340	96707	96707	68916	68916	65573

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 ethnicity's endogamy rate for Panel C. All regressions are weighted by the Census sample-line weight.
* significant at 5%; ** significant at 1%; *** significant at 0.1%

Table 6: **Pre-marital investments-Additional checks**

	Highest grade attained	Some high school	Education score	Literacy	Age at first marriage
	(1)	(2)	(3)	(4)	(5)
			Panel A: Males		
Sex ratio	1.602**	0.202	21.741	0.025	-0.837
	(0.533)	(0.113)	(15.259)	(0.026)	(1.903)
Stock	-2.320	-0.458	42.270	0.024	-3.559
	(3.213)	(0.449)	(68.654)	(0.039)	(5.492)
Sex ratio (older)	-0.309	0.061	-2.579	0.004	1.315
	(0.511)	(0.094)	(8.765)	(0.017)	(1.573)
Stock (older)	3.080	0.247	66.061	-0.031	4.412*
	(3.174)	(0.415)	(75.412)	(0.034)	(1.787)
N	86683	86683	66565	66565	54340
			Panel B: Females		
Sex ratio	0.763	0.103	21.515	0.067	-0.817
	(0.670)	(0.087)	(17.054)	(0.066)	(2.935)
Stock	-1.205	-0.256	56.306	0.031	-10.513
	(1.802)	(0.289)	(102.982)	(0.065)	(7.669)
Sex ratio (older)	-0.137	0.037	-25.237	-0.009	1.133
	(0.567)	(0.077)	(14.823)	(0.018)	(2.281)
Stock (older)	1.046	0.163	33.254	-0.039	8.620
	(2.171)	(0.326)	(85.940)	(0.049)	(6.132)
N	96707	96707	68916	68916	65573

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. All regressions are weighted by the Census sample-line weight.
* significant at 5%; ** significant at 1%; *** significant at 0.1%

Table 7: **Marriage market outcomes**

	Ever married	Currently divorced	Married twice	Married own ethnic immigrant	Highest grade of spouse	Spouse has some high school
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Females			
Sex ratio	0.022	-0.005	-0.047	0.027	-0.615	-0.106
	(0.052)	(0.026)	(0.070)	(0.068)	(1.592)	(0.127)
Stock	0.010	-0.002	0.029	0.162*	0.022	-0.250
	(0.051)	(0.024)	(0.072)	(0.068)	(1.939)	(0.171)
N. Obs.	125487	125487	75140	99896	16029	16029
			Panel B: Females-with interactions			
Sex ratio	-0.011	0.011	-0.076	-0.088	1.528	0.048
	(0.084)	(0.031)	(0.080)	(0.083)	(1.984)	(0.193)
Stock	0.056	-0.004	0.053	0.207**	0.616	-0.277
	(0.059)	(0.026)	(0.075)	(0.077)	(1.994)	(0.153)
Sex ratio*	0.563***	-0.038	0.393**	0.212*	7.156	-0.330
endogamy rate	(0.159)	(0.058)	(0.145)	(0.083)	(6.253)	(0.510)
N. Obs.	125194	125194	74976	99620	16015	16015
			Panel C: Males			
Sex ratio	0.072*	0.020	0.081	-0.040	2.227*	0.237**
	(0.028)	(0.010)	(0.065)	(0.025)	(1.131)	(0.091)
Stock	-0.055	-0.035	0.161*	0.087	-0.058	-0.005
	(0.075)	(0.028)	(0.073)	(0.070)	(2.211)	(0.185)
N. Obs.	114909	114909	62612	90014	16549	16549
			Panel D: Males-with interactions			
Sex ratio	0.050	0.038**	0.067	-0.038	1.794	0.203*
	(0.026)	(0.013)	(0.064)	(0.025)	(1.052)	(0.084)
Stock	-0.048	-0.036	0.167*	0.086	0.520	0.042
	(0.071)	(0.027)	(0.072)	(0.071)	(2.063)	(0.178)
Sex ratio*	0.203**	-0.086	0.082	0.000	7.605**	0.605
endogamy rate	(0.073)	(0.067)	(0.132)	(0.059)	(2.946)	(0.315)
N. Obs.	114905	114905	62609	90012	16549	16549

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 ethnicity's endogamy rate for Panels B and D. All regressions are weighted by the Census sample-line weight.

* significant at 5%; ** significant at 1%; *** significant at 0.1%

Table 8: **Labor supply**

	Males				**Females**			
	In LF	Employed	Hours	Weeks	In LF	Employed	Hours	Weeks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: OLS				
Sex ratio	0.000	0.000	-0.125	0.069	-0.003	0.000	-0.154	0.073
	(0.004)	(0.003)	(0.291)	(0.274)	(0.007)	(0.007)	(0.352)	(0.412)
Stock	-0.002	-0.008	-0.180	-0.756	0.002	0.003	0.226	-0.202
	(0.009)	(0.008)	(0.818)	(0.726)	(0.014)	(0.014)	(0.612)	(0.635)
				Panel B: IV				
Sex ratio	-0.010	0.003	0.007	2.100	-0.106*	-0.153*	-6.051	-6.117*
	(0.027)	(0.018)	(2.478)	(1.659)	(0.054)	(0.064)	(3.192)	(2.741)
Stock	0.029	0.031	10.517	3.536	-0.067	-0.060	1.593	0.610
	(0.048)	(0.049)	(5.853)	(5.656)	(0.041)	(0.040)	(3.370)	(4.751)
N. Obs	141367	158083	102109	103684	152373	170061	113454	113495

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, year of birth and year of Census fixed effects and dummies for whether both or only one parent is foreign-born. All regressions are weighted by the Census sample-line weight.
* significant at 5%; ** significant at 1%; *** significant at 0.1%

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

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

$$\frac{(1+r+\delta)}{\tilde{r}}(w-i-s)^{-\sigma} = \delta(i+\tilde{r}s)^{-\sigma} + (1+r)\sum_{i'} p^k(i')(c_2^{kE}+\tilde{r}s)^{-\sigma} + (1+r)(1-p^k)(c_2^{kN}+\tilde{r}s-\gamma)^{-\sigma} \tag{A.2}$$

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:

$$\begin{aligned}\frac{\partial i}{\partial z} \propto (1+r)\sum_{i'}\Bigg(\left((c_2^{kE}+\tilde{r}s)^{-\sigma}\left(\frac{\partial c_2^{kE}}{\partial i}-X\right) - (c_2^{kN}+\tilde{r}s-\gamma)^{-\sigma}\left(\frac{\partial c_2^{kN}}{\partial i}-X\right)\right)\frac{\partial p^k(i')}{\partial z} \\ + \frac{1}{1-\sigma}\left((c_2^{kE}+\tilde{r}s)^{1-\sigma} - (c_2^{kN}+\tilde{r}s-\gamma)^{1-\sigma}\right)\frac{\partial^2 p^k(i')}{\partial i \partial z} \\ + (c_2^{kE}+\tilde{r}s)^{-\sigma-1}\left(-\sigma p^k(i')\left(\frac{\partial c_2^{kE}}{\partial i}-X\right)\frac{\partial c_2^{kE}}{\partial z} + (c_2^{kE}+\tilde{r}s)\left(\frac{\partial p_k}{\partial i}\frac{\partial c_2^{kE}}{\partial z} + p^k(i')\frac{\partial^2 c_2^{kE}}{\partial i \partial z}\right)\right)\Bigg)\end{aligned} \tag{A.3}$$

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 match-

ing 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 (A.3) 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 (Ruggles et al. 2009) 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.[30] 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.[31] 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 C.1. This division was inspired by that used by

[30] 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.

[31] 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.

Angrist (2002) and based on Pagnini and Morgan (1990), with required modifications.[32]

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.[33]

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.[34] To avoid double-counting for the flow/stock indicator, only 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.

[32]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.

[33]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.

[34]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.

Table C.1: Ethnic group composition

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

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 C.2: **Spatial distribution of immigrants by ethnic group, 1900**

Ethnic group	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	TOP 10
British ancestry	NY	MA	PA	IL	MI	NJ	OH	CA	CT	MN	75.4
	(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 Europeans	NY	PA	MA	NJ	CA	IL	CT	LA	RI	OH	86.3
	(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 Europeans	NY	PA	IL	OH	NE	NJ	WI	MN	IA	TX	85.9
	(20.1)	(16.1)	(14.4)	(9.7)	(5.1)	(5.1)	(4.7)	(4.2)	(3.5)	(3.1)	
Other countries	HI	CA	MA	NY	OR	WA	MT	PA	AK	IL	86.4
	(34.6)	(24.5)	(5.6)	(5.5)	(5.3)	(0.42)	(1.9)	(1.9)	(1.5)	(1.4)	

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 C.3: **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 to a continuous variable by selecting the mid-point of the interval
Weeks worked per year	1940-70	26-75	

Table C.4: 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	Liter-acy	Highest grade attained	Some high school	Educ. score	Liter-acy	Highest grade attained	Some high school	Educ. score	Liter-acy
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Males												
Sex ratio	2.155***	0.255*	25.529	0.063	1.324*	0.236**	19.118	0.062*	0.986	0.173	6.644	0.027
	(0.586)	(0.110)	(17.909)	(0.044)	(0.532)	(0.086)	(34.586)	(0.030)	(0.623)	(0.124)	(22.292)	(0.028)
Stock	-0.404	-0.109	26.393**	-0.005	0.674	-0.082	123.824*	-0.037	0.768	-0.058	130.244*	-0.015
	(0.600)	(0.085)	(9.653)	(0.018)	(0.839)	(0.153)	(55.871)	(0.030)	(0.924)	(0.161)	(51.347)	(0.027)
Interaction					0.057	0.015	3.423	0.011**	0.199	0.024	4.461***	0.006***
					(0.116)	(0.022)	(5.965)	(0.004)	(0.233)	(0.038)	(1.061)	(0.001)
N	45152	45152	34177	34177	86683	86683	66565	66565	86683	86683	66565	66565
Panel B: Females												
Sex ratio	-2.657	-0.468	-533.798	-0.196	-0.254	-0.071	8.005	0.102	0.424	0.064	-72.441	0.026
	(4.624)	(0.743)	(1308.126)	(0.408)	(0.750)	(0.096)	(32.599)	(0.078)	(0.567)	(0.088)	(44.862)	(0.047)
Stock	4.989	0.586	506.203	0.161	0.246	0.072	66.346	-0.069	-0.253	-0.045	52.820	-0.039
	(5.677)	(0.966)	(1548.995)	(0.501)	(0.851)	(0.114)	(37.103)	(0.044)	(0.888)	(0.138)	(33.824)	(0.027)
Interaction					-0.472	-0.099**	-2.043	0.021	0.137	0.034	90.573***	0.052
					(0.248)	(0.035)	(13.911)	(0.013)	(0.222)	(0.036)	(25.006)	(0.034)
N	46484	46484	33745	33745	96707	96707	68916	68916	96707	96707	68916	68916

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 (5)-(8) and the divorce rate for each ethnicity in columns (9)-(12). In columns (1)-(4), 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.
* significant at 5%; ** significant at 1%; *** significant at 0.1%

Table C.5: **Effect of education on labor supply**

	In LF	Hours	In LF	Hours	In LF	Hours	In LF	Hours
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(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***	
F-test (instruments): 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***	
F-test (instruments): spouse							6.91***	
N. Obs	816934	804607	816934	804607	451970	445720	451970	445720

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.
* significant at 5%; ** significant at 1%; *** significant at 0.1%