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

Jeanne Lafortune†

January 2008

Abstract

While several studies examine the effect of marriage market conditions on post-marital labor supply, few account for the effect of these conditions on pre-marital investment decisions and mate selection. This paper investigates theoretically and empirically how changes in marriage market conditions affect pre-marital investments. I first show how a change in the sex ratio – that is, the ratio of males to females – can alter incentives for investments both in the context of a unitary model of the household and in a non-unitary setting with post-match bargaining. The model predicts that a rise in the sex ratio will lead men to increase pre-matching investments and women to decrease them if agents are sufficiently risk averse. I test this prediction using exogenous variation in the marriage market sex ratio, brought about by immigration, exploiting the preference of second generation Americans for endogamous matches. I find that a worsening of marriage market conditions spurs higher pre-marital investments, measured by years of education, literacy and occupational choice. Specifically, a change in the sex ratio from one to two leads men to increase their educational investment by 0.5 years on average and women to decrease it by (an insignificant) 0.05 years. In addition, the sex ratio significantly affect post-marital labor supply through pre-marital investments suggesting that accounting for these effects when using marriage market conditions as proxies for ex-post bargaining power is important. Overall, the results suggest that there are substantial returns to education in the marriage market, and that both men and women take these returns into account when making education decisions.

---

*I am particularly indebted to my advisors, Esther Duflo and David Autor for their comments and advice. 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 the MIT Labor Seminar, various MIT Labor and Development lunches as well as the PUC-Chile lunch. I acknowledge financial support from the Social Sciences and Humanities Research Council of Canada. All remaining errors are my own.

†Economics Department, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02142. E-mail: jlafor@mit.edu

1 Introduction

Several studies have shown that marriage market conditions (such as divorce laws and sex ratios) affect post-marital behavior (Chiappori, Fortin, and Lacroix 2002 and Angrist 2002). These results are generally interpreted as a rejection of the so called "unitary" model of the household where households, once formed, behave like a single individual. Because marriage market conditions change the outside option of each spouse, they alter bargaining weights and lead to modifications in the way household surplus is shared. However, there is little empirical work on the impact of these factors on pre-marital investments. This is surprising, since if individuals are forward-looking, these conditions should 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. Moreover, several studies have demonstrated that educational investments appear to respond to perceived returns in practice (see Foster and Rosenzweig 1996 and Nguyen 2007 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 impact upon agents' behavior before the union is formed. This paper investigates theoretically and empirically how changes in sex ratios (here defined as the ratio of males to females) modify incentives for pre-marital investments.

Theoretically, a change in the sex ratio can be expected to affect pre-marital investments through two channels: its effect on the probability of matching (which would hold even in a unitary model) and its effect on anticipated bargaining power. To better frame these two channels, I present a simple model. The timing of the model is as follows: first, pre-marital investments are undertaken by each individual, random matching then pairs individuals into couples and finally the output is shared according to rules that may depend on pre-marital investments and external conditions.

Sex ratios affect whether and with whom one can match. The model shows that for any relative risk aversion parameter larger than one, an increase in the sex ratio will lead men to increase and women decrease their pre-marital investments because of the matching effect. If the sex ratio is higher, a man has a higher probability of remaining single. The income effect of having no partner dominates the effect that a partner has on one's returns for high enough risk aversion. Thus, when one's marriage prospects get poorer, one's investment incentives are increased.

Secondly, sex ratios may also alter incentives for pre-marital investments because they modify the balance of power within a household. The model assumes that the division of the marital output occurs such that the bargaining leads to an ex-post Pareto optimal allocation, in the same spirit as in the "collective model" of the household. The bargaining weights may depend on an external determinant of bargaining power, as suggested by Browning and Chiappori (1998) and Chiappori, Fortin, and Lacroix (2002). However, this paper also allows individuals' pre-marital investments to influence the way the output is shared. I restrict the way these investments influence one's bargaining share by assuming that only the ratio of one's own investment to that of one's spouse influences the sharing factor.

The standard framework linking bargaining power and investments draws upon to 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. Since the utility is linear, there is no income effect stemming from obtaining a bigger proportion of the surplus. However, while risk neutrality may be a good approximation in the context of contracting between firms, it may not be appropriate in the context of spouses engaging in marital bargaining where risk aversion is likely to be present.

When the utility function is concave, a rise in the sex ratio decreases the incentives for male investment through a lower bargaining power (and hence return) as emphasized by Grossman and Hart (1980). This corresponds to a substitution effect. However, because the lower bargaining power also translates into smaller incomes, this increases the incentives for investment due to decreasing marginal utility. In the context of the model, this income effect is found to dominate the substitution effect if the elasticity of intertemporal substitution is less than one. Finally, the change in the sex ratio also modifies the incentives for one's spouse and in order for that response not to undo the direct effect of the bargaining power, it suffices to assume that the investments are gross substitutes in consumption.

Most estimates of either the relative risk aversion parameter or the elasticity of intertemporal substitution (which in this case are the inverse of each other) in the literature suggest that the restrictions mentioned above will hold (see for example Hall 1988, Beaudry and van Wincoop 1996 and Vissing-Jorgensen and Attanasio 2003) .

Note that if pre-marital investments modify post-marital outcomes, one would observe that the sex ratio affects post-marital outcomes, even outside a bargaining model. Furthermore, even within a bargaining framework, the estimated effect of marriage market conditions on post-marital outcomes may not properly measure the full impact of changes in bargaining power because part of this shift in bargaining is anticipated and counteracted by a change in pre-marital investments.

The model implies that for realistic values of the elasticity of intertemporal substitution, a rise in the sex ratio leads men to increase their pre-marital investments and (by an analogous argument) women to decrease them. This paper explores whether there is evidence of this pattern in the context of ethnic marriage markets in the United States around the turn of the twentieth century. Did second generation Americans modify their human capital acquisition decision when faced with a plausibly exogenous shift in the sex ratio of their state-level marriage market?

To answer this question, I exploit the fact that a large fraction of the children of immigrants, here referred to as second-generation Americans, tend to marry within their own ethnicity. Therefore, waves of newly arrived immigrants impact on their ethniticy'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 one to control for many potential confounders of the effect of a change in sex ratios. The variation in sex ratios of immigrants at this level is large and influences significantly the marriage market conditions of second-generation Americans. Since immigrants may select their location based on labor and marriage market conditions which also affect the second generation, this shock may be endogenous. To control for endogeneity, this paper constructs an instrument based on the fact that immigrant flows by country within a larger ethnic group are persistent. Each country within a group has located, in the past, to various destinations. Furthermore, over the course of the early twentieth century, the sex ratio of new immigrants has varied substantially and differently across countries. Consequently, one can construct an instrument that allocates shifts in the flows of immigrants to different states using past shares, akin 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.

Using this strategy, this paper finds that shifts in sex ratios influence pre-marital investment decisions of men, whether defined in terms of years of education, literacy or occupational choices. In states and ethnic groups where the number of males per female increases in their state of birth due to gender-biased immigration, young adult males acquire more formal education, are more literate, and pursue higher ranked occupations. A change from a balanced sex ratio among immigrants to one where men are twice as numerous as women leads men to increase their educational investment by 0.5 years and women to decrease it by (an insignificant) 0.05 years. These results are robust to various changes in the start and end dates of the period observed, in the states selected and to variations in the instrument.

As in previous studies (for example, Angrist 2002, Chiappori, Fortin, and Lacroix 2002, Amuedo-Dorantes and Grossbard 2007 and Oreffice and Bercea 2006), this paper also finds that post-matching labor supply decisions are affected by a change in the sex ratio, although the estimated impact is weaker and less significant than previously estimated, possibly due to the difference in empirical strategy. In particular, women's labor force attachment is reduced. A doubling in the sex ratio of newly arrived immigrants from a balanced level lead to a fall of about 4 percent in female labor force participation, of 1.4 hours worked per week, and of 1.3

weeks worked per year. The labor supply response for men appears to be generally positive, but smaller in magnitude and insignificant. The indirect effect of the sex ratio on labor supply through educational attainment, however, appears significant, particularly for males, which suggests that using marriage market conditions as proxies for ex-post bargaining power may lead to inaccurate conclusions regarding the importance of bargaining power.

Finally, this paper combines the model developed and the empirical estimates obtained to calibrate the returns to education in the marriage market. The fact that marriage market conditions influence educational decisions suggests that there are some returns to education in the marriage market, whether stemming from a joint production function or through the effect of pre-marital investments on bargaining weights. Defining returns to schooling on the marriage market as any returns that would not be captured if one was single, I find that these make up around 40 to 60 percent of total returns. These returns are thus important in magnitude and may help to explain why, in this context and many others, women are as educated as men although they have very low rates of labor force participation. It may reflect both the importance of education as an input in household tasks such as child rearing as well as a method to strengthen one's position within the household.

This paper is organized as follows. Section 2 summarizes the existing literature. Section 3 then introduces the model and derives comparative statics, while Section 4 presents the data and explains the empirical strategy. The subsequent section presents the results of the regressions and section 6 then uses both the estimates and the theoretical model to separate the returns to education stemming from the labor market vis-à-vis those from the marriage market. The last section concludes and suggests avenues for future research.

2 Literature Review

This paper is related to the growing theoretical literature linking education and marriage markets in order to address changes in the educational attainment gap between genders (Chiappori, Iyigun, and Weiss 2007 and Pena 2006). In contrast to this literature, the model in this paper assumes symmetry in the production function in order to focus more closely on the effect of the sex ratio. The theoretical work most related to this paper is that of Iyigun and Walsh (2007) who construct a model where the sex ratio can generate gender gaps in educational attainment. They assume a competitive marriage market where consumption levels are independent of spousal investments and this implies that investments are Pareto efficient. This means that their model cannot generate monotone comparative statics for investments with respect to the sex ratio. By contrast, under the specification used in this paper, bargaining may lead to inefficient investment levels and parameters can be selected to ensure monotone comparative statics.

Empirically, this paper relates to a wave of new studies that have explored effects of changes in sex ratios on non-labor outcomes, mostly marital and fertility decisions (Porter 2007, Brainerd 2006, Kvasnicka and Bethmann 2007). They all use large shifts in fertility or mortality rates which altered sex ratios and find that when the sex ratio increases women tend to marry more and to be less likely to have out-of-wedlock births. Porter (2007) also finds that higher sex ratios lead women to marry "better" mates in terms of health, age and height. Angrist (2002) studies the effect of a national shock to the ethnic sex ratio brought about by immigration and uses as an instrument for the gender composition of immigrants the sex ratio at entry. If immigrants leave their home country for reasons that are exogenous to the local marriage and labor market, this instrument identifies the causal effect of changes in sex ratios on post-marital behavior. Using this strategy, he finds that both men and women are more likely to get married and that women's labor supply is reduced while overall incomes are increased when the sex ratio rises. However, no study has yet explored pre-marital investments as a potential outcome.

Finally, this study also contributes to existing work exploring the returns to education in the marriage market. Foster and Rosenzweig (2001) show in a general equilibrium framework that agricultural productivity growth raises the demand for educated wives and confirm it empirically. Behrman, Foster, Rosenzweig, and Vashishtha (1999) suggest that much of this response is due to the capacity of better educated mothers to teach their children. An older branch of this literature have looked at earnings correlations with own and spousal education. Some studies found positive correlation between own earnings and spousal education (Benham 1974 for the United States, Tiefenthaler 1997 for Brazil, Neuman and Ziderman 1992 for Israel) suggesting that in particular in the case of entrepreneurs, one's earnings tends to rise with spousal education. Also, marriage market conditions seem to influence human capital acquisition (Boulier and Rosenzweig 1984 for example). However, no study has yet quantified of how large the return to education on the marriage market are.

3 Model

3.1 General model set-up

Let us assume a setting where each man (m) and woman (f) is endowed with an initial wealth w. Individuals have an additive utility function over two periods[1]:

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

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

For simplicity, assume that the utility function has constant elasticity of intertemporal substitution/constant relative risk aversion given by the parameter σ:

$$u(c_k) = \frac{c_k^{1-\sigma}}{1-\sigma}, \qquad \sigma > 0, \qquad k = 1, 2.$$

In the first period, an individual can invest in a productive asset i at a cost of 1. Her consumption in the first period is thus given by:

$$c_1 = w - i.$$

In the second period, individuals pair and can share resources. The joint output is given by the function h which is assumed to be increasing in both investment levels, twice-continously differentiable and symmetric:

$$\frac{\partial h(i, i')}{\partial i} = \frac{\partial h(i', i)}{\partial i} > 0.$$

In addition, the production function is always positive when one individual has positive investments and offers positive returns even at very low levels of investment:

$$\begin{aligned}
h\left(i^j, 0\right) &> 0, \qquad j = m, f \\
\lim_{i^j \to 0} \left(\frac{\partial h\left(i^j, i^{j'}\right)}{\partial i^j} \right) &> 0, \qquad j = m, f \\
\frac{\partial h\left(i^j, 0\right)}{\partial i^j} &> 0, \qquad j = m, f.
\end{aligned}$$

Notice that these restrictions exclude the Cobb-Douglas production function, for example, but are perfectly compatible with a constant elasticity of substitution function.

Once paired, individuals, through post-matching bargaining, arrive at a Pareto optimal sharing (as in Browning and Chiappori 1998, Chiappori, Fortin, and Lacroix 2002). This implies that the post-matching division of the total output will be given by the maximization of the following program:

$$Max \quad \mu\left(i^m, i^f, z\right) u\left(c_2^m\right) + \left(1 - \mu\left(i^m, i^f, z\right)\right) u\left(c_2^f\right) \quad s.t. \quad c_2^m + c_2^f = h\left(i^m, i^f\right).$$

Pareto weights are allowed to depend on 3 elements: male and female pre-marital investments and potentially the sex ratio denoted by z.

One can show that the optimal allocations are given by:

$$
\begin{aligned}
c_2^m &= c_2^m\left(i^m, i^f, z\right) = \frac{\mu^{\frac{1}{\sigma}}}{(1-\mu)^{\frac{1}{\sigma}} + \mu^{\frac{1}{\sigma}}} h\left(i^m, i^f\right) \\
c_2^f &= c_2^f\left(i^m, i^f, z\right) = \frac{(1-\mu)^{\frac{1}{\sigma}}}{(1-\mu)^{\frac{1}{\sigma}} + \mu^{\frac{1}{\sigma}}} h\left(i^m, i^f\right)
\end{aligned}
$$

which imply that one's consumption is a share of the total output and that these shares are determined monotonically by the Pareto weight μ. Further assume that the following conditions for these weights:

$$
\begin{aligned}
\mu\left(0, i^f, z\right) &= 1 - \mu\left(i^m, 0, z\right) = 0 \\
\frac{\partial \mu}{\partial i^m} &> 0 > \frac{\partial \mu}{\partial i^f}.
\end{aligned} \tag{1}
$$

The first condition indicates that when one is paired with a partner who does not invest (that is single), one captures the full output of the pair. The third imposes that one's investment increases one's share of the output.

Finally, weights will always be such that one always consumes at least what they can obtain from being single. If it was not the case, individuals would elect not to marry and thus this would not be an equilibrium. Formally,

$$
c_2^k \geq i^k, \qquad k = f, m.
$$

The main assumption of this model is that the consumption exhibits constant returns to scale in the investment levels, that is

$$
c_2^j\left(\lambda i^m, \lambda i^f, z\right) = \lambda c_2^j\left(i^m, i^f, z\right), \qquad j = m, f.
$$

This will imply that the household production function also has constant returns to scale. More intuitively, this also implies that the Pareto weights (μ) will be homogenous of degree zero in male and female investments: only the ratio of investments influences the share one receives. Thus, in a couple where the male has twice the amount of investment as his wife, this man captures the same share of the total output, no matter what are the absolute levels of investments of both parties. Finally, imposing that consumption must rise at a decreasing rate with own investment

$$
\frac{\partial^2 c_2^j}{\partial i^{j2}} < 0, \qquad j = m, f
$$

implies, because of the assumption of constant returns to scale, that

$$\frac{\partial^2 c_2^j}{\partial i^m \partial i^f} > 0, \qquad j = m, f.$$

In the first period, spouses decide their optimal investment level based on the following maximization problem

$$Max \quad u\left(w - i^k\right) + E\left(u\left(c_2^k\left(i^m, i^f, z\right)\right)\right), \qquad k = m, f$$

taking the sex ratio (z) and the future spouse's investment as given.[2]

Once the investments have been made, individuals match randomly. This can be rationalized by the existence of search frictions that prevent individuals from finding their perfect partner. This excludes the possibility of using investments to capture a better spouse. Because of this, the probability that males and females will stay single is independent of the investment level and given by

$$\begin{aligned} p^m(z) &= \begin{cases} z - 1 & if\ z > 1 \\ 0 & if\ z \leq 1 \end{cases} \\ p^f(z) &= \begin{cases} z & if\ z < 1 \\ 0 & if\ z \geq 1 \end{cases}. \end{aligned}$$

One can also see this $p(z)$ as the fraction of period 2 one will spend being single if there are frictions once matched and one can possibly lose one's partner. The first order conditions to this maximization are given by

$$\left(w - i^k\right)^{-\sigma} = p^k(z)\left(i^k\right)^{-\sigma} + \left(1 - p^k(z)\right)\left(c_2^k\right)^{-\sigma}\frac{\partial c_2^k}{\partial i^k}, \qquad k = m, f \tag{2}$$

Notice that because the matching is random, one cannot invest in order to capture a higher skilled wife which is why there is no term in the first order conditions that relates own investment to that of one's spouse.

Investments would be ex-ante Pareto optimal in this case if one was to be the full residual claimant of the returns since the total output available to the household would be maximized in that case. However, in this model, the investment will never be pareto efficient since $\frac{\partial c_2^j}{\partial i^j} = \frac{\partial h}{\partial i^j}$ cannot hold simultaneously for both spouses, as in Acemoglu (1996). There are two factors that distort the decision away from the optimal one. First, one only receives a share of the

[2] From our assumptions that $c_2^k \geq i^k$ and that $\frac{\partial^2 c_2^k}{\partial i^m \partial i^f} > 0$, we know that the return to investment will always be at least 1 and so that savings will not occur.

total output and thus does not capture the full return to one's investment because part of the benefits of this investment will be captured by one's spouse. This would lead one's investment to be below the Pareto optimal level. On the other hand, because investments can be employed to obtain a larger share of the output, overinvestment may also occur. The investment levels selected are thus not ex-ante Pareto optimal unlike in Peters and Siow (2002) or in Iyigun and Walsh (2007). The bargaining process here does not eliminate the "public good" nature of the pre-marital investment as suggested by Bergstrom, Blume, and Varian (1986).

Lemma 1 *There exists a unique pure strategy Nash Equilibrium to this game.*

Proof. See Appendix A. ■

The existence of a Nash Equilibrium depends on the assumption that the consumption function exhibits constant returns to scale (which bounds the degree of complementarity between inputs in the consumption function) and that single individuals receive a positive return (which prevents the existence of a "no-investment" equilibrium).

3.2 Comparative statics

A change in the sex ratio modifies the incentives for pre-marital investment through three distinct channels: a change in the probability of marriage, in the relative bargaining weight as well as the anticipated investment level of one's spouse.

3.2.1 Effect of a change in the probability of matching

First assume that spousal Pareto weights are independent of the sex ratio. The only effect that the sex ratio has is then through the probability of one matching. Formally,

$$\left.\frac{\partial i^k}{\partial z}\right|_{i^{k\prime}} = \frac{\frac{\partial p^k}{\partial z}\left(\left(i^k\right)^{-\sigma} - \left(c_2^k\right)^{-\sigma}\frac{\partial c_2^k}{\partial i^k}\right)}{\sigma\left(w - i^k\right)^{-\sigma-1} + p^k\left(\sigma\left(i^k\right)^{-\sigma-1}\right) + \left(1-p^k\right)\left(c_2^k\right)^{-\sigma-1}\left(\sigma\frac{\partial c_2^k}{\partial i^k}^2 - c_2^k\frac{\partial^2 c_2^k}{\partial i^{k2}}\right)} \quad k = f, m. \tag{3}$$

Importantly, this will only affect the investment level of an individual who is on the short side of the market. That is

$$\begin{aligned} \left.\frac{\partial i^m}{\partial z}\right|_{i^f} &= 0 \quad if \quad z < 1 \\ \left.\frac{\partial i^f}{\partial z}\right|_{i^m} &= 0 \quad if \quad z > 1. \end{aligned}$$

Proposition 1 *The investment level of the individuals on the short side of the market will increase as the number of potential spouses available to them decreases if* $\sigma > \tilde{\sigma}$*, where* $\tilde{\sigma} < 1$*.*

Proof. See Appendix A. ■

This result can be explained intuitively. When an individual is single, she has a lower income which would entice her to invest more. On the other hand, the return to her investment is lower because she does not have a spouse to increase the value of this investment. If the agent is sufficiently risk averse, her desire to insure her consumption in case she remains single dominates the substitution effect.

3.2.2 Effect of a change in bargaining power

Now, adding in the effect that the sex ratio has on investments through bargaining power, assume that

$$\frac{\partial \mu}{\partial z} < 0.$$

Thus, z is a factor that decreases the consumption of men and increases that of females given investment levels. Furthermore, assume that the sex ratio is limited in the way it can decrease the return to male investment by:

$$\frac{\partial c_2^j}{\partial i^f}\frac{\partial^2 c_2^j}{\partial i^m \partial z} > \frac{\partial c_2^j}{\partial i^m}\frac{\partial^2 c_2^j}{\partial i^f \partial z}, \qquad j = m, f. \tag{4}$$

This implies that the effect of the sex ratio on the return one receives from the spouse's investment must be significant enough in magnitude compared to that on own investment.[3] Thus, it cannot be that the sex ratio penalizes greatly the return that can be obtained from own investment but not from that of the spouse.

In this case, the effect of z on pre-marital investment through the channel of bargaining, conditional on spousal investment, is given by

$$\left.\frac{\partial i^k}{\partial z}\right|_{i^{k'}} = \frac{\left(1-p^k\right)\left(c_2^k\right)^{-\sigma-1}\left(-\sigma\frac{\partial c_2^k}{\partial i^k}\frac{\partial c_2^k}{\partial z}+c_2^k\frac{\partial^2 c_2^k}{\partial i^k \partial z}\right)}{\sigma\left(w-i^k\right)^{-\sigma-1}+p^k\left(\sigma\left(i^k\right)^{-\sigma-1}\right)+\left(1-p^k\right)\left(c_2^k\right)^{-\sigma-1}\left(\sigma\frac{\partial c_2^k}{\partial i^k}^2-c_2^k\frac{\partial^2 c_2^k}{\partial i^{k2}}\right)}, \qquad k = f, m. \tag{5}$$

Proposition 2 *Conditional on spousal investment, an increase in bargaining power will lead an individual to decrease their investment level as long as $\sigma > \bar{\sigma}$, where $\bar{\sigma} < 1$*

Proof. From (5),

$$\left.\frac{\partial i^k}{\partial z}\right|_{i^{k'}} \propto \left(-\sigma\frac{\partial c_2^k}{\partial i^k}\frac{\partial c_2^k}{\partial z} + c_2^k\frac{\partial^2 c_2^k}{\partial i^k \partial z}\right)$$

[3] This can be shown to be equivalent to the conditions $\frac{\partial^2 \alpha}{\partial i^m \partial z} < \min\left\{\frac{-1}{1-\alpha}\frac{\partial \alpha}{\partial i^m}\frac{\partial \alpha}{\partial z}, \frac{1}{\alpha}\frac{\partial \alpha}{\partial i^m}\frac{\partial \alpha}{\partial z}\right\}$ where α is defined as $c_2^m = \alpha(i^m, i^f, z) * h\left(i^m, i^f\right)$. Thus, the effect that the sex ratio has on one's ability to modify their weight in the decision process through education cannot be too large in magnitude.

Appendix A demonstrates that when (4) holds, a sufficient condition for the sex ratio for $\frac{\partial i^m}{\partial z}\Big|_{i^f} > 0$ and vice-versa for females is that $\sigma > 1$. ■

A rise in the sex ratio as a shift in bargaining power towards females influences the investment decision through two channels. Males have lower consumption for any value of investment which increases their incentives for investment through this income effect. On the other hand, this increase in z also reduces the return to investment and through this channel, leads to a lower investment level. For the income effect to dominate and thus for males to increase their investment when the sex ratio rises, σ must be sufficiently large. This is akin to the effect of a productivity shock on investment decisions in a macroeconomic model.

3.2.3 Spousal response

Finally, it must also be that the spousal response will not undo the effect of the bargaining power as presented above. A sufficient, although not necessary, condition for this to occur is that investments be strategic substitutes. That is when one is faced with a spouse who has invested more, the income effect brought about by this is larger than the incentives embedded in the complementarity of investments and this leads one to reduce one's investment. Formally, investments will be strategic substitutes if

$$\begin{aligned}
\frac{\partial i^k}{\partial i^{k'}} &= \frac{\left(1-p^k\right)\left(c_2^k\right)^{-\sigma-1}\left(-\sigma\frac{\partial c_2^k}{\partial i^k}\frac{\partial c_2^k}{\partial i^{k'}}+c_2^k\frac{\partial^2 c_2^k}{\partial i^k \partial i^{k'}}\right)}{\sigma\left(w-i^k\right)^{-\sigma-1}+p^k\left(\sigma\left(i^k\right)^{-\sigma-1}\right)+\left(1-p^k\right)\left(c_2^k\right)^{-\sigma-1}\left(\sigma\frac{\partial c_2^k}{\partial i^k}^2-c_2^k\frac{\partial^2 c_2^k}{\partial i^{k2}}\right)} < 0, \qquad k = f, m \\
&\Leftrightarrow \quad \sigma\frac{\partial c_2^k}{\partial i^k}\frac{\partial c_2^k}{\partial i^{k'}} > c_2^k\frac{\partial^2 c_2^k}{\partial i^k \partial i^{k'}}.
\end{aligned}$$

This implies that investments cannot be so highly complementary that the substitution effect dominates the income effect. This translates into a fairly intuitive condition that is that

$$\sigma > \frac{1}{\rho_c\left(i^m, i^f\right)}$$

where $\rho_c\left(i^m, i^f\right)$ is the elasticity of substitution of investments within the consumption function. If investments are gross substitutes, $\sigma > 1$ is then a sufficient condition. This rules out the possibility that individuals are the full residual claimant of their investment or receive a return above the return they produce on the couple's output (as in Wells and Maher 1998). In both these cases, investments are strategic complements.

Thus, for $\sigma > 1$ and when investments are gross substitutes, an increase in the sex ratio will lead to a decrease in female investment and an increase in male pre-marital investments. Appendix B presents a special case of the model presented above: one where individuals Nash bargain over the surplus.

3.3 Perfect competition

One could remove the assumption of random matching and turn to a model where there is assortative matching. However, in that case, because there are no search frictions, one's consumption will be determined by market forces rather than bargaining. If even one man with a given investment is single, all the other men with the same investment level as his will earn a single man's payoff. If that was not the case, a single man could offer to any woman to pair with them and offer him only ε more than his current pay-off and every woman would accept.

This also means that the individual is the full residual claimant on his marginal contribution since:

$$c_2^j = h\left(i^m, i^f\right) - c_2^{j'}$$

and $c_2^{j'}$ is a price outside the control of the agent himself. I will further assume that when $z = 1$, the output is shared equally between spouses (any share $\in [0, 1]$ would be an equilibrium). Assume again that the output function h has constant returns to scale and has positive marginal return when own investments are 0 and that one receives $c_2^k = i^k$ if single.

Proposition 3 *Under perfect competition, when the sex ratio increases, men increase their investment and women decrease their investment as long as $\sigma > 1$.*

Proof. See Appendix A. ■

Thus, the result obtained above also holds in a context where there is assortative matching and perfect competition.

3.4 A different outside option

The previous sections assume that a higher sex ratio will lead males to be more likely to be single. However, in the empirical context that follows, it may be more relevant to think of individuals as being pushed to another marriage market (that of natives). Assume that the second period utility of a member of an ethnic group is given by consumption minus a fixed penalty γ if he marries someone from another ethnic group. An increase in the sex ratio leads men to be more likely to marry native women. This gives them less utility which creates an incentive for higher investment levels. Furthermore, in this particular application, the investment levels in this native pool are higher and this encourages further investment due to complementarity. In this case the size of the preferred marriage pool would be also important since the impact of the sex ratio within your own marriage market may depend on the likelihood of marrying within one's group. Thus, even in this case, tighter marriage market conditions will lead to higher investments, whether or not the sex ratio influences post-matching bargaining.

3.5 Ex-post outcomes and the sex ratio

This analysis also highlights that if the sex ratio affects pre-marital investments, the effect of the sex ratio on post-marital consumption levels will not represent the effect of bargaining power. It is a mixture of the bargaining power effect, the effect of one's investment level on post-marital outcome and the effect of one's spouse's. Formally,

$$\frac{dc_2^k}{dz} = \frac{\partial c_2^k}{\partial z} + \frac{\partial c_2^k}{\partial i^k}\frac{\partial i^k}{\partial z} + \frac{\partial c_2^k}{\partial i^{k\prime}}\frac{\partial i^{k\prime}}{\partial z}.$$

4 Data and Empirical Strategy

Having established a framework where changes in the sex ratio modify individuals' pre-marital investments, I now investigate empirically the link between sex ratios and pre-marital investments.

As in Angrist (2002), this paper uses data from second-generation Americans born around the turn of the century (from 1885 to 1915). This identification strategy is based on the observation that second-generation Americans tend 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 relevant marriage market includes new waves of immigration. Because marriage markets are fairly local, I focus on state-level within ethnic group marriage markets. From 1900 to 1930, the sex ratio of newly arrived immigrants varied greatly transforming the balance of power within each state's ethnic marriage markets. These waves occurred at the moment when the sample of second generation individuals was making educational and marriage decisions. Because 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), which leads past immigrant stocks in a particular state to predict current immigrant flows.

4.1 Basic specification

The basic regression of this study relates pre-marital outcomes of second generation Americans to two characteristics of their marriage market: its sex ratio and its total size. The second variable provides an estimate of the effect of market thickness which may influence decisions as explained in Section 3.4. In addition, it captures any effect that overall own-ethnic immigration has on local conditions, either through the marriage or the labor market.

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. They also include dummies for age, year of birth, year of the Census and for nativity of parents. The estimation equation is given by

$$y_{kjst} = \alpha * sexratio_{jst} + \beta * flow_{jst} + \psi * X_{kjst} + \sigma_{js} + \phi_{st} + \gamma_{jt} + \varepsilon_{kjst} \tag{6}$$

where the left hand-side variable is an outcome for an individual k, of ethnic group j, born in state s, of cohort t. Conceptually, this regression contrasts the change in outcomes over time among individuals from a given state of two different ethnic groups.

The marriage market size and sex ratio may be endogenous. 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 estimation of Equation (6).

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. 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. Various countries of birth are included within each ethnic group and each had previously selected different locations. Since over the period, the sex ratio of immigrants within an ethnic group varied by country of birth, the combination of this variation and differences in past location shares provides the source of variation that the instrument will exploit. In short, this instrumental strategy assumes that individuals tend to locate more where their fellow countrymen live but marry within the entire ethnic group. If all countries of birth within one ethnic group selected the same locations, there would not be no geographical variation to exploit.

More precisely, two instruments were constructed as follows. All male and female immigrants 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 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 and that of females. The sex ratio instrument is built by dividing the predicted flow of males by that of females. Equations (7) and (8) define formally the instruments:

$$pred\_sexratio_{ist} = \frac{\sum_{j \in i} \left(\frac{imm_{js1900}}{imm_j} \right) * males_{jt}}{\sum_{j \in i} \left(\frac{imm_{js1900}}{imm_j} \right) * females_{jt}} \tag{7}$$

$$pred\_flow_{ist} = \sum_{j \in i} \left(\frac{imm_{js1900}}{imm_j} \right) * (males_{jt} + females_{jt}) \,. \tag{8}$$

The strategy can best be illustrated by an example using the Scandinavians 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. Figure 1 presents the evolution of the sex ratio among all four countries over the period studied and the predicted sex ratio of this ethnic group in both states. 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. Figure 2 shows that the same argument holds for flows.

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. It 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 also not violate the exclusion restriction if pre-1900 immigrants selected states in anticipation of better conditions for all ethnic groups but not particularly for their particular ethnic group because regressions control for state-time fixed effects.

In addition, it must also be 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 could be violated if when more men than women enter the United States, these men tend to be of lower/better quality. Little information on immigrants' quality is available to test whether this is violated, except for immigrants' literacy as measured by the Census. No correlation was found between that measure for either gender and the actual or the instrumented sex ratio of immigrants.

4.2 Data

4.2.1 Outcome measures

All outcomes, obtained from IPUMS files between 1900 and 1970, are presented with a detailed description in Appendix Table A1. First, marital outcomes are collected: marital status, measures of marital stability (divorce rates and number of marriages) and country of birth of one's spouse. 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. Because pre-marital investments may be modified because marriage is delayed when the marriage is tight, leaving more time to acquire education, age at first marriage is also measured. To alleviate the problem of sample selection (it 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.

Various measures of human capital investment as proxies for pre-marital investments are considered: literacy, years of schooling and occupational choices. Literacy should be acquired before marriage and could affect post-matching output (see Behrman, Foster, Rosenzweig, and Vashishtha 1999 for an example in India). A more continuous measure of human capital investments is the highest grade a person has attained. While it provides a more detailed categorization of the level of skills acquired, it may also be obtained partially after marriage but there is little evidence of this in my sample. First, 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. Finally, two occupational indices measuring the "quality" of the current occupation are available. These variables could reflect pre-marital investments because the quality of an occupation is correlated with on-the-job training and past human capital accumulation although it could also reflect some labor supply decisions. To alleviate potential problems linked to these measures capturing post-marital investments, they are measured only for those aged 15-25 except in the case of education where education was only available at later ages.

To measure post-matching outcomes, this paper uses labor supply of all individuals aged 25 and above. It does 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 can be obtained.

Table 1 gives the main summary statistics for each outcome. The rate of non-marriage after age 35, around 10 percent, is much above that of natives (about 5 percent).[4] Divorce rates are low but widowhood is not uncommon for women. The age at first marriage is around 23 for women and 27 for men. About 8 percent of second generation women and 3 percent of males are currently married with immigrants from their ethnic group (which accounts for most marriages between second generation and immigrants). This is somewhat low but includes all singles and widows, to avoid selection bias. 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. Labor supply attachment by woman is quite low. Slightly more than 30 percent of women were in the labor force compared to 79 percent for men. While 61 percent of men worked full time, only 18 percent of women did.

4.2.2 Marriage market measures

A key decision in implementing the above framework is the appropriate empirical definition that should be used for a "marriage market". In this setting, a marriage market is assumed to be a given ethnic group within a state in a particular cohort. This definition of marriage market is quite restrictive but as long as what happens to one's market is more relevant than what occurs in another group, this approximation will capture relevant variation.

The marriage market is first defined within an age cohort. The second generation sample born between 1885 and 1915 is divided into 5 year intervals. I maintain the assumption that people marry within their age cohort.[5]

Marriage markets are fairly local with more than 65 percent of sampled individuals married to someone who was born in the same state as them.[6] State is the lowest geographical unit for which place of birth is available in the IPUMS files. Furthermore, mobility is limited: more than 70 percent of individuals in this sample still live in their state of birth and this figure increases to 85 percent among those less than 20.

Marriage markets must also include a definition of ethnic groups. 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.[7] Using

[4] 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).

[5] Around 50 percent of married individuals younger than 40 are matched to someone of the same age group.

[6] 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.

[7] 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

all countries of birth, the sample was divided into 9 ethnic groups, summarized in Appendix Table A2. This division was inspired by that used by Angrist (2002) and based on Pagnini and Morgan (1990), with required modifications.[8] Using marriage patterns of previous immigration waves, these groupings were found to correspond closely to the patterns observed in the data. In almost all cases, the percent of individuals marrying someone within their ethnic group but not from their own country of birth was much higher than the prevalence of those countries in the sample. The regressions below were performed with slight differences in the allocation of ethnicities to ethnic group with very similar results.

Two sets of marriage market conditions were constructed using the IPUMS files for 1910, 1920 and 1930: one for immigrants only, the other incorporating second generation individuals as well to which we will refer to as "foreign stock". The former were classified by their country of birth, year of immigration (grouped into 5-year periods) and state of current residence. The latter were classified by their state of birth, their father's place of birth and grouped with immigrants such that these immigrants arrived while the second generation individuals are in their late teens (age 15-19), an age at which schooling and marriage decisions are made.[9] Only immigrants arriving between the ages of 10 and 25 are included since they are more likely to be part of the marriage pool of the cohort of second generation Americans.[10] For each ethnic group-state-immigration period, the above methodology produces a measure for the number of immigrants and their gender. Measures of total flow of immigrants and total flow of foreign stock are then built by summing all individuals in each state-ethnic group-period cell. Sex ratios were defined as the number of males per female in each cell.[11]

4.2.3 Instrument construction

Equations (7) and (8) employ as national flow measures the sum of all state flows as defined previously. Location shares were obtained from the 1900 Census tables (United States Census

percent of foreign born parents share a common country of birth.

[8] East European Jews are grouped by nationality because it is difficult to identify them after 1930. Also, two countries of birth per ethnic group were required since the instrument relies on differences in 1900 location choices within ethnic groups across countries of birth. Immigrants from Ireland were joined with 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.

[9] To avoid double-counting for the flow indicator, only the 1910 Census is used to compute the flow of immigrants between 1900 and 1909, the 1920 Census for immigrants arriving between 1910 and 1919 and so forth. However, since the sex ratio may suffer more from measurement error because it is a ratio, all three waves of the Census were employed to construct that measure.

[10] Also, a variant using different age distribution by gender such that females are between 10 and 22 but men between 13 and 25 was used to better match spousal age differences and produced very similar results.

[11] 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.

Office 1901).[12] Ideally, the 1890 shares would have been used but a few key countries were only tabulated starting in 1900 and for countries which were similarly identified in both periods, shares are almost identical.

Immigrants were concentrated in some key states with 72 percent of all immigrants locating in 10 states in 1900–see Appendix Table A3. The location of immigrants varied importantly by ethnic group, the traces of which can still be found in the ethnic composition of today's population. 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, variation in location choice across countries of birth and within ethnicity arises. 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 2 summarizes the distribution of the endogenous variables and of the instruments. As can be seen, the major immigrant groups over this period were the Russians, Poles and Romanians, followed by Southern Europeans and Germans. Including second generation individuals, those of Germans and British descent are far more numerous, reflecting the importance of past immigration waves. The sex ratios of Others (mostly Asians) and Italians were among the highest at almost two men per women 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.

5 Results

5.1 First stage

The instruments are very highly predictive of their respective endogenous variable. An increase of one in the predicted flow leads to an increase of about 0.87 in the actual number of immigrants arriving over that period and of about 0.57 for the total foreign stock. Similarly, an increase of one in the predicted sex ratio measure is linked to an increase of about 0.84 among immigrants and about 0.43 among total foreign stock. These are all significant at 0.1 percent level. These results are presented in Table 3.

[12] 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 (2006).

The robustness of the first stage is tested through various specifications presented in Table 4.[13] 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 and finds that this omission does not change the robustness of the first stage. Column (2) restricts itself to the first four periods of immigration and again finds similar results. Although 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. The first stage is also robust to the transformation of the instruments and the endogenous variables in logarithmic form. Finally, removing the key immigration-receiving state over this period (New York) does not change the results. All empty cells where the sex ratio was imputed were dropped without significantly altering the coefficient on the sex ratio. To ensure that this result does not stem from correlated measurement error in the flow of immigrants, another source of data was used to construct the flow of immigrants for the instrument. This measure included all immigrants (irrespective of their year of birth and their age at arrival) by country of birth, arrival period and gender but only for Foreign Born Whites. As can be seen in column (6), the instrument conserves predictive power, although the precision goes down due to the imperfect measure of flows it constitutes.

5.2 Outcomes

5.2.1 Marital outcomes

I first explore the causal effect of marriage market conditions on marital outcomes. The first panel of Table 5 presents regressions where marriage market variables are defined to include only immigrants while the bottom section relates to conditions measured for the total foreign stock. Surprisingly, both men and women are more likely to have ever been married when the sex ratio rises (although this is similar to Angrist 2002). The probability of a man ever having been married rises by 3.3 percent when the sex ratio of immigrants doubles from a balanced level. The next two columns shed some additional light on this result. Men are also more likely to be divorced and more likely to have been married more than once (the opposite being true for women). The rise in the probability of having been married more than once is almost 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.

As presented in the last section of the model, marriage rates may also be unaffected if the outside option is not to remain single but rather to elect a less desirable marriage market.

[13] It only presents the results for the marriage market measures based on immigrants but very similar results were obtained for the ones based on total foreign stock.

Column (4) of Table 5 indicates that men are significantly less likely to marry an immigrant of their own ethnicity when the sex ratio is higher. A change from a balanced sex ratio to one where there are twice as many male as female immigrants decreases the probability of a man to be married to a female immigrant of his own ethnicity by 1.7 percent. The effect is smaller and imprecisely estimated for females. Marriage market size strongly increases the probability of marrying an immigrant as predicted from the model. Also, age at first marriage does not appear to be modified suggesting that any effect of the sex ratio on pre-marital investments will not mechanically stem from a delay in marriage timing. Marriage market sizes appear to hurry the timing of marriage of women but delay that of men. When one's preferred marriage market expands, it may be easier to select optimally the timing of one's marriage. If females prefer marrying earlier than men, this could explain these results.

5.2.2 Pre-marital investments

The main outcomes of interest relate pre-marital investments to marriage market conditions. Table 6 first presents correlations obtained from an OLS regression. The effects observed here are in the predicted direction and significant in some cases for females but the coefficients are extremely small. The OLS results should 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 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 while the IV result presented above is small and insignificant, which may be indicative that immigrants select locations where the marriage prospects are good.

Once one purges endogeneity using instruments, the coefficients are larger in magnitude for both genders (except for literacy). The results now indicate that a marriage market favoring women leads men to be more literate, to have more years of completed education and select more highly paid occupations. All these are significant at least at 10% significance level. The coefficients for females are negative except for literacy but are neither very large nor significant.

This suggests that a shift from a sex ratio among immigrants of 1 to 2 (more than two standard deviations in the sex ratio) leads to a 1.7 percent point increase in the probability that a male is literate and, on average, to about half a year more of education. Furthermore, young men were selecting much more highly ranked occupations when faced with higher sex ratios. Women's responses are smaller in magnitude, except for occupational choices. Marriage market size appears to lead both genders to select much more highly paying occupations, in particular for women. This is surprising but may reflect the fact that immigrants fill low-paid occupations and push second generation Americans to higher-paying ones. Overall, omitting the flow measure usually renders the effect of the sex ratio more significant.

Table 7 tests in various ways the robustness of these results.[14] The first column removes the first immigration period and finds very similar results, indicating that the result is not driven by the early years of the period in question. Dropping the last period does not modify the point estimate for education by much but does increases the standard errors while it greatly increases the size of the effect of the sex ratio on the occupational ranking variable. Removing the major immigrant-receiving state over this period (New York), if anything, strengthens the relationship. Adding dummies for each country of origin (rather than ethnic group dummies as in the base specification) does not weaken the pattern observed. Restricting attention to older or younger respondents leaves the results unchanged. Similarly, ignoring a particular Census year does not affect the results.[15] Although not presented here, variants of the instrument were explored with similar results. For example, 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 of a state's attractiveness for a particular gender and its attractiveness for a particular ethnicity, the results are very similar to the ones presented above.[16]

5.2.3 Labor supply

The model presented above argues that the change in pre-marital investments due to altered sex ratios stems from a desire to partially offset the expected effect of the sex ratio on post-marital outcomes. Having found a significant effect of marriage market conditions on human capital decisions, this paper now turns to proxies of post-marital outcomes. The OLS regressions presented in Table 8 suggest that higher sex ratios among immigrants are correlated with higher labor force participation of both men and women. 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 working less and the OLS would be a lower bound on the magnitude of the causal estimate.

The right-hand side panel of Table 8 presents the results of the instrument variable regres-

[14] It focuses on males and on two specific measures, years of education and Duncan Index of occupational choices, although the results are similar for females and other outcome measures.

[15] These last two variants are only presented for the educational variable because the occupational score was restricted to individuals between the ages of 15 and 25.

[16] An interesting data set including the intended state of residence of immigrants at the port of entry was also collected. Unfortunately, the first stage using this data proved to be too weak to be of use for this paper.

sions. The causal effect of the sex ratio appears to lead women to reduce their labor inputs while men increase theirs. These results indicate that a doubling of the sex ratio (from 1 to 2) lead women to be 4 percent less likely to be in the labor force and 3 percent less likely to be employed. This is smaller than the 9 percent found by Angrist (2002) which included females aged 16-33, an age at which labor supply is much more variable and potentially influenced by pre-marital decisions. A rise in the sex ratio of immigrants from a balanced level to one where immigrant men are twice as numerous reduces hours worked per week and the number of weeks per year by about 1.3. These results are significant only at 10 percent significance. For males, a change from a balanced sex ratio to one where men are twice as numerous leads to no effect for either employment or labor force participation and raises hours worked per week and weeks per year by about 0.5, although these are very imprecisely estimated and insignificant. The OLS results are usually lower, although not significantly so, than the IV for males and higher for females as expected if immigrants locate based on the bargaining conditions of the marriage market.[17]

5.3 Labor supply, pre-marital investment and mate selection

The previous section found that the sex ratio had modest albeit imprecisely measured effects on labor supply. One could conclude that this implies little evidence of ex-post bargaining. However, one must also take into account that the effect of the sex ratio measured by the above regression includes not only the ex-post bargaining effect but also any effect that the sex ratio may have had on post-marital outcomes through its effect on education. 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. Labor supply and education are measured in the 1940-1970 IPUMS files. Two sets of results are obtained: one for the full sample and another for second generation Americans. The results presented in Table 9 use as instruments a set of dummy variables for each minimum number of years of schooling required by the state.[18] The first stage suggests that each additional year of compulsory schooling leads men to increase their level of schooling by about 0.05 years and women to do so by about 0.8 years.[19] The IV estimates suggest that education decreases labor

[17] Selecting only married females would show a much clearer pattern where females greatly reduce their labor supply. However, it is unclear whether this would stem from selection or ex-post bargaining.

[18] Similar results were obtained by using a continuous measure of the minimum number of years of schooling.

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

supply, 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 about 0.5 hours for females and 1.5 hours for males, but are only significant for males. Among females, the results are stronger when the sample is restricted to second generation individuals, increasing the magnitude and the significance of the effect to about 1.5 hours. The first stage is much weaker in this subsample for males and the effect for hours worked falls to about 0.5. Results for labor force participation are much smaller and weaker, in particular for males.

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. 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 the above section, a doubling of the sex ratio, through the educational channel itself, decreases the number of hours worked by males by about 0.5-0.75 hours per week.[20] The effect for females is in the same direction albeit much smaller. This suggests that the effect of the sex ratio on labor supply obtained in the previous section is underestimating the true effect of the sex ratio on post-matching outcomes, as predicted by the model presented above. The effect of the sex ratio on post-matching outcomes, once purged of the effect it has through changes in pre-marital investments and matching patterns, then provides an estimate of the effect of an external shifter in bargaining power on post-marital labor supply.

Furthermore, in a case where only matching influenced the choice of pre-marital investment, the effect of the sex ratio on labor supply, for example, is entirely driven by its effect on one's own and spouse's education because the sex ratio does not alter ex-post decisions. These results are thus not in accordance with a hypothesis where only matching is at play.[21]

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

[21] Other evidence that matching is not solely driving the results was obtained. First, while the above matching model suggests that the effect of the sex ratio is largest when one is on the short side of the market, no evidence of this was found. Second, the effect of the sex ratio appears to be larger in larger communities, which is inconsistent with a matching model since a similar change in the sex ratio implies many fewer potential mates in a small than in a large community.

This exercise is meant as an illustration of the importance of considering the link between pre-marital behavior and post-marital decisions. It suggests that using changes in sex ratios as proxies for ex-post bargaining power without taking into effect the potential link that marriage market conditions have on pre-marital behavior and matching patterns may lead to misleading inference. It would have been best to include education in the above labor supply regressions and use another source of exogenous variation to instrument for it. Unfortunately, the sample of second generation Americans employed in this study was too small to use a measure of compulsory schooling as an additional instrument for the educational attainment of an individual in the labor supply regression.

6 Returns to education in the marriage market

The results presented above suggest that marriage market conditions influence pre-marital investments. Assuming this is due to a reaction to changes in the incentives imbedded within the marriage market, these results can be used to infer returns to education in the marriage market.

6.1 General framework

Let us define the returns to education in the marriage market as any additional benefit that is given by one's human capital investment that would not be observed if one were single. First, there could be additional benefits captured once married simply because the educational investments of each spouse are complementary in the household production function (from utility derived from conversations, from the role of parental education in child-rearing or even because of learning spillovers). Because those benefits are shared between spouses, the "public good" aspect of this return may lead individuals to underinvest compared to the optimal level. Secondly, marriage market returns could arise if one's bargaining weight depends on one's educational level. Thus, if single, one's education simply affects the output produced but if married, it affects both the output and the share of it one can capture. In this setting, the spouses simply play a zero-sum game where education does not have any additional productive element but serves as a negotiation tool. This would lead individuals to overinvest.[22]

Disentangling the various sources of incentives for human capital investment is not easy.[23]

[22] Finally, marriage market may also stimulate investments through competition between individuals of the same gender if the matching is not random. It can be shown, however, that in this case, a rise in the sex ratio would lead to a fall in males' investment because as the sex ratio rises, the value of the benefit of more education (i.e. a spouse) falls because fewer females are available. Nevertheless, in this case, the gender who is on the short side of the market may over invest in education simply to compete with one another.

[23] Studying the effects of a policy that increases education of only one gender on the other gender's investment decision would be a key input in this analysis.

Nevertheless, as an illustration, this section attempts to derive some estimates of the importance of marriage market-related returns to education by combining the empirical estimates found above with the model developed in Section 3. I fit the parameters of the model to be the most consistent with the observed educational choices of males, females and their spouse and the measured effect of the sex ratio on education. The estimated parameters are then used to compute the fraction of returns to education in the marriage market. Formally, let me define total returns to education as

$$\frac{\partial \log c_2^k\left(i^m, i^f\right)}{\partial i^k} = \frac{1}{c_2^k\left(i^m, i^f\right)} \frac{\partial c_2^k}{\partial i^k}.$$

To separate marriage and labor market returns, I take the marriage market returns to correspond to returns not captured by a single individual, that is

$$\frac{1}{c_2^k\left(i^m, i^f\right)} \left(\frac{\partial c_2^k\left(i^m, i^f\right)}{\partial i^k} - \frac{\partial c_2^k\left(i^k, 0\right)}{\partial i^k} \right).$$

Notice that in the models presented below, $c_2^k\left(i^k, 0\right) = i^k$ and thus the labor market returns will be given by $1/c_2^k$. I further parameterize the model by assuming that the household production is given by

$$h\left(i^m, i^f\right) = \left(i^{m\alpha} + i^{f\alpha}\right)^{\frac{1}{\alpha}},$$

a constant elasticity of substitution function (CES) with an elasticity given by $\frac{1}{1-\alpha}$.

6.2 Spouse selection model

Let me first consider a model where the sex ratio affects the matching patterns but not the way the household surplus is shared. Formally, male and female consumption are given by

$$\begin{aligned} c_2^m &= i^m + \lambda \left(\left(i^{m\alpha} + i^{f\alpha}\right)^{\frac{1}{\alpha}} - i^m - i^f \right) \\ c_2^f &= i^f + (1-\lambda) \left(\left(i^{m\alpha} + i^{f\alpha}\right)^{\frac{1}{\alpha}} - i^m - i^f \right) \end{aligned} \tag{9}$$

In addition, if one matches outside one's preferred marriage market, one receives a penalty of γ in utility terms. Notice that in this case, the first order condition is given by

$$\left(w^m - i^m\right)^{-\sigma} = p\left(z\right)\left(c_2^{mN} - \gamma\right)^{-\sigma} \frac{\partial c_2^{mN}}{\partial i^m} + \left(1 - p\left(z\right)\right)\left(c_2^{mO}\right)^{-\sigma} \frac{\partial c_2^{mO}}{\partial i^m} \tag{10}$$

where $p\left(z\right)$ is the probability of marrying a native, c_2^{mN} corresponds to the consumption when married to a native female, c_2^{mO} to the consumption level when married to a female of one's own ethnicity and γ the utility cost of marrying out of one's ethnic group. In this case, one can also find the effect of a change in z which is given by

$$\frac{\partial i^m}{\partial z} = \frac{\frac{\partial p}{\partial z}\left(\left(c_2^{mN}-\gamma\right)^{-\sigma}\frac{\partial c_2^{mN}}{\partial i^m}-c_2^{mO-\sigma}\frac{\partial c_2^{mO}}{\partial i^m}\right)+(1-p)\left(c_2^{mO}\right)^{-\sigma-1}\left(-\sigma\frac{\partial c_2^{mO}}{\partial i^m}\frac{\partial c_2^{mO}}{\partial i^f}+c_2^{mO}\frac{\partial^2 c_2^{mO}}{\partial i^m \partial i^f}\right)\frac{\partial i^{fO}}{\partial z}}{-SOC^m} \tag{11}$$

where the denominator simply corresponds to the second order condition. Using Equations (10) and (11), the mean value of i^m, i^{fN}, i^{fO} from the data and the computed estimate of $\frac{\partial i^m}{\partial z}$ and $\frac{\partial i^{fO}}{\partial z}$ from above, I find the set of parameters α and σ that offer the best fit.[24] An assumption must also be made about the average initial wealth of each individual (w). Since educational investment ranges from 0 to 18 with an average slightly above 9, the results below will use variations in the average wealth ranging from 22 to 30 since one should never want to invest more than half of one's wealth based on the model presented above. Three values of bargaining weights (λ) are also evaluated. Finally, two more parameters are required in this case. I must calibrate the cost of marrying outside one's ethnic group and impose that this cost be 9, slightly less than the consumption level one would receive if there were no complementarity between investment levels. Similar results were obtained with other values. Finally, an estimate of $\frac{\partial p}{\partial z}$ was computed to correspond to the estimated effect of the sex ratio on spousal education level since the effect of the sex ratio on the probability of marrying a member of one's own ethnic group (immigrant and second generation Americans) could not be estimated.

The top panel of Table 10 presents these results. Both for males and females, estimates of the parameter α are very comparable and vary between 0.26 and 0.49. This is somewhat surprising because females modified their education by a much smaller fraction in response to a change in the sex ratio in the estimates presented above. The reason for this result is that despite the fact that they have barely modified their behavior, they are now facing men who have changed their behavior substantially ($\frac{\partial i^{mO}}{\partial z}$is large and positive). In response to this females would want to decrease their investment decision by a large fraction. To match the small decrease observed in the data, men and women must be fairly high complements in the production function. Men's effect operates mostly through a change in probability of marrying a native and this leads to similar estimates.

These parameter estimates then imply fairly substantial returns to education when married (between 2 and 5 percent). These estimates suggest that about 40-60 percent of all returns are obtained because of the role education plays within the household production function.

[24] Formally, the parameters selected minimize the sum of the squared errors in the two equations. Imposing that the first order condition holds with equality and then finding the set of parameters which offers the best fit for the comparative static equation offers very similar estimates of α although lower estimates of σ.

6.3 Bargaining power model

For purpose of comparison, let me now assume that the sex ratio only affects one's bargaining power within the household and that the sex ratio has no influence on marital patterns. Consumption levels are now given by

$$\begin{aligned} c_2^m &= i^m + \beta(z)\left(\left(i^{m\alpha} + i^{f\alpha}\right)^{\frac{1}{\alpha}} - i^m - i^f\right) \\ c_2^f &= i^f + (1 - \beta(z))\left(\left(i^{m\alpha} + i^{f\alpha}\right)^{\frac{1}{\alpha}} - i^m - i^f\right). \end{aligned} \tag{12}$$

Further assume that the sharing factor is $\beta(z) = \exp((\ln\lambda) * z)\,, \quad \lambda < 1$. This is a suitable parameterization since it implies that $\beta(0) = 1$ and $\beta(\infty) = 0$.

Using this framework, the first order conditions for each gender is given by

$$(w - i^m)^{-\sigma} = (c_2^m)^{-\sigma}\left(\frac{\partial c_2^m}{\partial i^m}\right) \tag{13}$$

and the effect of the sex ratio on investments given by

$$\frac{\partial i^m}{\partial z} = \frac{(c_2^m)^{-\sigma-1}\left[\left(-\sigma\frac{\partial c_2^m}{\partial i^m}\frac{\partial c_2^m}{\partial z} + c_2^m\frac{\partial^2 c_2^m}{\partial i^m \partial z}\right) + \left(-\sigma\frac{\partial c_2^m}{\partial i^m}\frac{\partial c_2^m}{\partial i^f} + c_2^m\frac{\partial^2 c_2^m}{\partial i^m \partial i^f}\right)\frac{\partial i^f}{\partial z}\right]}{SOC^m}. \tag{14}$$

Equations (13) and (14) are then used as above to find the parameter values of σ and α most consistent with the empirical results obtained.

Although not shown here, the sex ratio increases the investment levels of the spouses of both second generation males and females. Because of this, the direct effect of the sex ratio on male investment in this setting is partially counterbalanced by the fact that they are now paired with higher investment females. Females, on the other hand, experience both the direct and the spousal effect in the same direction. Because of this, the algorithm must estimate men and women's investment to be fairly complementary similar to the case where only matching was at play.

Estimates of the key parameters are shown at the bottom of Table 10. The first two panels include the estimates obtained from individual calibrations for both men and women. Those results suggest that the parameter of the CES function is fairly close to 0.4 which implies fairly high rates of return to education in the marriage market. Although obtained from different sets of estimates, the results for men and women are surprisingly consistent with each other. Panel C thus solves simultaneously for the four sets of restrictions and finds similar results, except for the case where men have more bargaining power where the returns to education in the marriage market are estimated to be much larger. On average, results from Table 10 imply that marriage market returns to education are of the order of about 2-5 percent and correspond to about 50 percent of total returns with slightly larger shares for women than for men. Thus, this model, while making very different assumptions, generates very similar results to the ones presented above where no bargaining power was assumed.

Furthermore, one can, in this case, compare the level of investment observed (about 9.5 years of schooling for both men and women) to the one that would maximize the sum of male and female utility if education decisions could be made jointly. The optimal estimated investment levels, in this case, are well above the observed level. This implies that individuals invest less than would be optimal because of the "public good" nature of the household production. Also, the use of education as a bargaining tool which would lead to over-investment is dominated by this effect. However, this is driven by the assumption of Nash Bargaining, which imposes that returns to investments are lower than optimal.

Despite the fact that these estimates stem from calibrations and would benefit greatly from being refined using more empirically-based techniques, they nevertheless point to substantial returns to education related to the marriage market. Furthermore, they emphasize that household production may produce human capital externalities, a channel that is yet to be explored.

7 Conclusions

Overall, the results of this paper suggest that substantial returns to education derived from the marriage market, whether it be through household production or bargaining power effects. Furthermore, men and women appear to understand, and respond to, the incentives for human capital accumulation embedded in marriage market conditions.

I first derive a framework to explore the relationship between marriage market conditions and pre-marital investments. The model shows that the usual effect of bargaining power on investment within the incomplete contracts framework does not hold for fairly reasonable values of the elasticity of intertemporal substitution. In addition, because sex ratios may affect both bargaining power and matching patterns, a change in the sex ratio would have a similar effect on pre-marital investments under both a unitary framework and a bargaining model. Finally, the model highlights that using marriage market conditions as proxies for ex-post bargaining may not be appropriate as post-marital outcomes will be influenced by marriage market conditions even without any effect through a change in bargaining power.

Empirical support for these conclusions was found in the data. Using shocks to one's marriage market coming from immigration, this paper shows that an increase in the sex ratio increases pre-marital investments for males and lowers them for females, although only significantly so for males. This is confirmed by a variety of outcome measures and is robust to changes in specifications. It appears to stem, at least partially, from changes in bargaining power. In addition, the magnitude of these shifts 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, in particular for the case of males. Finally, these empirical estimates merged with the structure of the model suggest that returns to education in the marriage market are substantial.

These results provide interesting insights into the determinants of educational decisions. The importance of incentives linked to returns received once married may partially explain why the educational gap by gender is not always correlated with either difference in labor force attachment or differences in wages between men and women. Furthermore, while conventional wisdom maintains that women's educational attainment will increase as their bargaining power in developing countries, the results of this paper indicate that this may not be the case.

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. Furthermore, the fact that marriage market conditions may affect post-marital outcomes through modifications in pre-marital conditions even when no post-matching bargaining occurs cautions the use of such measures as tests of the unitary framework.

Finally, these findings may also shed light on the persistence of skewed sex ratios. Willis (1999), for example, suggests that out-of-wedlock births may be more likely when the sex ratio is lower and when men have fewer economic opportunities. This has been used to explain the high rates of single motherhood among inner city African-Americans where the bias in the marriage market sex ratio in favor of males is due to high male incarceration rates and fast population growth (Guttentag and Secord 1983). This paper suggests that this low sex ratio would lead males to invest less in their human capital and thus offer them even worse economic outcomes. This would reinforce the existing gap between male and female economic opportunities and thus generate even worse marriage pools for African-American females. Similarly, markedly high sex ratios in the context of Asian countries would be predicted to induce lower educational attainment by females. Because of this, parents may be less likely to rely on their girls for future economic support and this could reinforce a pre-existing cultural bias for boys. These are fruitful topics left for further research.

References

Acemoglu, D. (1996). A microfoundation for social increasing returns in human capital accumulation. *The Quarterly Journal of Economics 111*(3), 779–804.

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 (2006). Attenuation bias in measuring the wage impact of immigration. Sabanci University working paper series.

Beaudry, P. and E. van Wincoop (1996). The intertemporal elasticity of substitution: An exploration using a US panel of state data. *Economica 63*(251), 495–512.

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.

Benham, L. (1974). Benefits of women's education within marriage. *The Journal of Political Economy 82*(2), S57–S71.

Bergstrom, T., L. Blume, and H. Varian (1986). On the private provision of public goods. *Journal of Public Economics 29*(1), 25–49.

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. unpublished mimeo, available at http://urban.hunter.cuny.edu/RePEc/seminar/old/rfwomen.pdf.

Browning, M. and P.-A. Chiappori (1998). Efficient intra-household allocations: A general characterization and empirical tests. *Econometrica 66*(6), 1241–1278.

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. F. Iyigun, and Y. Weiss (2007). Public goods, transferable utility and divorce laws. unpublished mimeo, available at ftp://repec.iza.org/RePEc/Discussionpaper/dp2646.pdf.

Foster, A. and M. Rosenzweig (2001). Missing women, the marriage market and economic growth. Unpublished manuscript.

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.

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.

Guttentag, M. and P. F. Secord (1983). *Too Many Women? The Sex Ratio Question*. Newbury Park, California: Sage Publications.

Haines, M. R. (1996). Long term marriage patterns in the United States from colonial times to the present. available at http://www.nber.org/papers/h0080.pdf.

Hall, R. E. (1988). Intertemporal substitution in consumption. *The Journal of Political Economy 96*(2), 339–357.

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.

Kvasnicka, M. and D. Bethmann (2007). World war II, missing men, and out-of-wedlock childbearing. SFB 649 Discussion Paper 2007-053.

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.

Neuman, S. and A. Ziderman (1992). Benefits of women's education within marriage: Results for Israel in a dual labor market context. *Economic Development and Cultural Change 40*(2), 413–424.

Nguyen, T. (2007). Information, role models and perceived returns to education: Experimental evidence from Madagascar. available at http://econ-www.mit.edu/grad/trang/research.

Oreffice, S. and B. Bercea (2006). Quality of available mates, education and intra-household bargaining power. Fondazione Eni Enrico Mattei Working Paper, available at http://www.feem.it/NR/rdonlyres/23FB2DA0-9D56-4623-919E-EF744D3B8D34/2157/13308.pdf.

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.

Pena, X. (2006). Assortative matching and the education gap. University of Georgetown mimeo, available at http://www12.georgetown.edu/students/xp/matching.pdf.

Peters, M. and A. Siow (2002). Competing premarital investments. *The Journal of Political Economy 110*(3), 592–608.

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

Tiefenthaler, J. (1997). The productivity gains of marriage: Effects of spousal education on own productivity across market sectors in Brazil. *Economic Development and Cultural Change 45*(3), 633–650.

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

Vissing-Jorgensen, A. and O. P. Attanasio (2003). Stock-market participation, intertemporal substitution, and risk-aversion. *American Economic Review 93*(2), 383–391.

Wells, R. and M. Maher (1998). Time and surplus allocation within marriage. Massachusetts Institute of Technology Working Paper.

Willis, R. J. (1999). A theory of out-of-wedlock childbearing. *The Journal of Political Economy 107*(6), S33–S64.

A Omitted Proofs of Results

Proof of Lemma 1. There will be no strategic behavior between men or between women since the matching is random: thus men do not compete against other men to capture the best females. Assume without loss of generality that the sex ratio is above 1. Thus, $p^m(z) > 0$ and $p^f(z) = 0$. Let $p = p^m(z)$. The first order conditions (2) define best response functions $i^f(i^m)$ and $i^m\left(i^f\right)$ for both husband and wife.

Define a Nash Equilibrium as

$$i^{f*} = i^f\left(i^m\left(i^{f*}\right)\right)$$

The function

$$\bar{\imath}^f - i^f\left(i^m\left(\bar{\imath}^f\right)\right)$$

is strictly increasing iff

$$1 > \frac{\partial i^f}{\partial i^m} \frac{\partial i^m}{\partial i^f}$$

which implies

$$\begin{aligned}
1 \quad &> \quad \left(\frac{(1-p) \left(-\sigma \frac{\partial c_2^m}{\partial i^m} \frac{\partial c_2^m}{\partial i^f} + c_2^m \frac{\partial^2 c_2^m}{\partial i^m \partial i^f} \right)}{\sigma \left(\left(\frac{w - i^m}{c_2^m} \right)^{-\sigma-1} + p \left(\frac{i^m}{c_2^m} \right)^{-\sigma-1} \right) + (1-p) \left(\sigma \frac{\partial c_2^m}{\partial i^m}^2 - c_2^m \frac{\partial^2 c_2^m}{\partial i^{m2}} \right)} \right) \left(\frac{\left(-\sigma \frac{\partial c_2^f}{\partial i^m} \frac{\partial c_2^f}{\partial i^f} + c_2^f \frac{\partial^2 c_2^f}{\partial i^m \partial i^f} \right)}{\sigma \left(\frac{w - i^f}{c_2^f} \right)^{-\sigma-1} + \left(\sigma \frac{\partial c_2^f}{\partial i^f}^2 - c_2^f \frac{\partial^2 c_2^f}{\partial i^{f2}} \right)} \right) \\
1 \quad &> \quad \frac{(1-p) \left(\sigma^2 \frac{\partial c_2^m}{\partial i^m} \frac{\partial c_2^f}{\partial i^f} \frac{\partial c_2^m}{\partial i^f} \frac{\partial c_2^f}{\partial i^m} - \sigma \left(c_2^f \frac{\partial c_2^m}{\partial i^m} \frac{\partial c_2^m}{\partial i^f} \frac{\partial^2 c_2^f}{\partial i^m \partial i^f} + c_2^m \frac{\partial c_2^f}{\partial i^m} \frac{\partial c_2^f}{\partial i^f} \frac{\partial^2 c_2^m}{\partial i^m \partial i^f} \right) + c_2^m c_2^f \frac{\partial^2 c_2^m}{\partial i^m \partial i^f} \frac{\partial^2 c_2^f}{\partial i^m \partial i^f} \right)}{A + (1-p) \left(-\sigma \left(c_2^f \frac{\partial c_2^m}{\partial i^m}^2 \frac{\partial^2 c_2^f}{\partial i^{f2}} + c_2^m \frac{\partial c_2^f}{\partial i^f}^2 \frac{\partial^2 c_2^m}{\partial i^{m2}} \right) + \sigma^2 \frac{\partial c_2^m}{\partial i^m}^2 \frac{\partial c_2^f}{\partial i^f}^2 + c_2^m c_2^f \frac{\partial^2 c_2^m}{\partial i^{m2}} \frac{\partial^2 c_2^f}{\partial i^{f2}} \right)} \qquad where \quad A > 0
\end{aligned}$$

This will hold because male and female consumption exhibits constant returns to scale which imply

$$\frac{\partial^2 c_2^m}{\partial i^{m2}} \frac{\partial^2 c_2^f}{\partial i^{f2}} = \left(-\frac{i^f}{i^m} \frac{\partial^2 c_2^m}{\partial i^m \partial i^f} \right) \left(-\frac{i^m}{i^f} \frac{\partial^2 c_2^f}{\partial i^m \partial i^f} \right) = \frac{\partial^2 c_2^m}{\partial i^m \partial i^f} \frac{\partial^2 c_2^f}{\partial i^m \partial i^f}$$

and also

$$\begin{aligned}
c_2^f \frac{\partial c_2^m}{\partial i^m}^2 \frac{\partial^2 c_2^f}{\partial i^{f2}} + c_2^m \frac{\partial c_2^f}{\partial i^f}^2 \frac{\partial^2 c_2^m}{\partial i^{m2}} \quad &= \quad -\left(\frac{i^m}{i^f} c_2^f \frac{\partial c_2^m}{\partial i^m}^2 \frac{\partial^2 c_2^f}{\partial i^m \partial i^f} + \frac{i^f}{i^m} c_2^m \frac{\partial c_2^f}{\partial i^f}^2 \frac{\partial^2 c_2^m}{\partial i^f \partial i^m} \right) \\
&= \quad -\left(c_2^f \frac{\partial c_2^m}{\partial i^m} \frac{\partial^2 c_2^f}{\partial i^m \partial i^f} \left(\frac{c_2^m}{i^f} - \frac{\partial c_2^m}{\partial i^f} \right) + c_2^m \frac{\partial c_2^f}{\partial i^f} \frac{\partial^2 c_2^m}{\partial i^f \partial i^m} \left(\frac{c_2^f}{i^m} - \frac{\partial c_2^f}{\partial i^m} \right) \right) \\
&< \quad c_2^f \frac{\partial c_2^m}{\partial i^m} \frac{\partial c_2^m}{\partial i^f} \frac{\partial^2 c_2^f}{\partial i^m \partial i^f} + c_2^m \frac{\partial c_2^f}{\partial i^m} \frac{\partial c_2^f}{\partial i^f} \frac{\partial^2 c_2^m}{\partial i^m \partial i^f}
\end{aligned}$$

and by our assumption that

$$\frac{\partial \mu}{\partial i^m} > 0 > \frac{\partial \mu}{\partial i^f}$$

this ensures that

$$\frac{\partial c_2^m}{\partial i^m} \frac{\partial c_2^f}{\partial i^f} > \frac{\partial c_2^m}{\partial i^f} \frac{\partial c_2^f}{\partial i^m}$$

There will be a unique Nash Equilibrium if in addition

$$\begin{aligned}
0 \quad &< \quad i^f \left(i^m \left(0 \right) \right) \\
w \quad &> \quad i^f \left(i^m \left(w \right) \right)
\end{aligned}$$

which holds since by the concavity of the utility function and the fact that the return to investment is strictly positive, one always wants to invest a strictly positive amount and consume a strictly positive amount in the first period. ■

Proof of Proposition 1. From Equation (3) it follows that

$$\frac{\partial i^k}{\partial z} \propto \frac{\partial p(z)}{\partial z} \qquad if \quad i^k < c_2^k \left(\frac{\partial c_2^k}{\partial i^k} \right)^{\frac{-1}{\sigma}}$$

which can be rewritten as

$$\begin{aligned} \left(\frac{\partial c_2^k}{\partial i^k} \right)^{\frac{1}{\sigma}} &< \frac{c_2^k}{i^k} \\ \frac{1}{\sigma} \ln \left(\frac{\partial c_2^k}{\partial i^k} \right) &< \ln \left(\frac{c_2^k}{i^k} \right) \\ \frac{\ln \left(\frac{\partial c_2^k}{\partial i^k} \right)}{\ln \left(\frac{c_2^k}{i^k} \right)} &= \tilde{\sigma} < \sigma \end{aligned}$$

Using the fact that c_2^k has constant returns to scale

$$\begin{aligned} \tilde{\sigma} &= \frac{\ln \left(\frac{\partial c_2^k}{\partial i^k} \right)}{\ln \left(\frac{i^k \frac{\partial c_2^k}{\partial i^k} + i^{k\prime} \frac{\partial c_2^k}{\partial i^{k\prime}}}{i^k} \right)} \\ &= \frac{\ln \left(\frac{\partial c_2^k}{\partial i^k} \right)}{\ln \left(\frac{\partial c_2^k}{\partial i^k} + \frac{i^{k\prime}}{i^k} \frac{\partial c_2^k}{\partial i^{k\prime}} \right)} \\ &< 1 \end{aligned}$$

As long as $\frac{\partial c_2^k}{\partial i^{k\prime}} > 0$ since $\frac{\partial c_2^k}{\partial i^k} > 1$. ■

Proof of Proposition 2. To show that $\sigma > 1$ is a sufficient condition for

$$\sigma > \bar{\sigma} = \frac{c_2^m \frac{\partial^2 c_2^m}{\partial i^m \partial z}}{\frac{\partial c_2^m}{\partial i^m} \frac{\partial c_2^m}{\partial z}} \tag{15}$$

Let me use the fact that the consumption functions exhibit constant returns to scale and thus this is equal to

$$\bar{\sigma} = \frac{\left(i^m \frac{\partial c_2^m}{\partial i^m} + i^f \frac{\partial c_2^m}{\partial i^f} \right) \frac{\partial^2 c_2^m}{\partial i^m \partial z}}{\frac{\partial c_2^m}{\partial i^m} \left(i^m \frac{\partial^2 c_2^m}{\partial i^m \partial z} + i^f \frac{\partial^2 c_2^m}{\partial i^f \partial z} \right)}$$

$$
\begin{aligned}
\bar{\sigma} &= 1+\frac{i^{f}\left(\frac{\partial c_{2}^{m}}{\partial i^{f}}\frac{\partial^{2}c_{2}^{m}}{\partial i^{m}\partial z}-\frac{\partial c_{2}^{m}}{\partial i^{m}}\frac{\partial^{2}c_{2}^{m}}{\partial i^{f}\partial z}\right)}{\frac{\partial c_{2}^{m}}{\partial i^{m}}\left(i^{m}\frac{\partial^{2}c_{2}^{m}}{\partial i^{m}\partial z}+i^{f}\frac{\partial^{2}c_{2}^{m}}{\partial i^{f}\partial z}\right)} \\
\bar{\sigma} &< 1\ if \quad \frac{\partial c_{2}^{m}}{\partial i^{f}}\frac{\partial^{2}c_{2}^{m}}{\partial i^{m}\partial z}>\frac{\partial c_{2}^{m}}{\partial i^{m}}\frac{\partial^{2}c_{2}^{m}}{\partial i^{f}\partial z}
\end{aligned}
$$

A similar derivation would lead us to conclude that $\sigma > 1$ is a sufficient condition for $\sigma\frac{\partial c_{2}^{f}}{\partial i^{f}}\frac{\partial c_{2}^{f}}{\partial z} < c_{2}^{f}\frac{\partial^{2}c_{2}^{f}}{\partial i^{f}\partial z}$. ■

Proof of Proposition 3. When an individual is single, he will invest

$$i=\frac{w}{2}$$

When the sex ratio is 1 and each individual is matched with someone identical to them and the surplus is shared equally. The Nash Equilibrium exists and is unique and, given the fact that the production function is symmetric, is equal to

$$
\begin{aligned}
-(w-i)^{-\sigma}+\left(\frac{1}{2}h(i,i)\right)^{-\sigma}\frac{\partial h(i,i)}{\partial i} &= 0 \qquad (16)\\
(w-i)^{-\sigma} &= \left(i\frac{\partial h(i,i)}{\partial i}\right)^{-\sigma}\frac{\partial h(i,i)}{\partial i} \\
i &= \frac{w}{1+\left(\frac{\partial h(1,1)}{\partial i}\right)^{\frac{\sigma-1}{\sigma}}}
\end{aligned}
$$

This will be less than $\frac{w}{2}$ if $\sigma > 1$ since

$$
\begin{aligned}
1+\left(\frac{\partial h(1,1)}{\partial i}\right)^{\frac{\sigma-1}{\sigma}} &> 2 \\
\left(\frac{\partial h(1,1)}{\partial i}\right)^{\frac{\sigma-1}{\sigma}} &> 1
\end{aligned}
$$

If the sex ratio is in one's favor ($z > 1$ for women, $z < 1$ for men), one is match with a partner who invests like a single individual (and thus more than when $z = 1$) and can be offered the single individual pay-off. Their first-order condition is given by

$$-(w-i)^{-\sigma}+\left(h\left(i,i'\right)-i'\right)^{-\sigma}\frac{\partial h\left(i,i'\right)}{\partial i}=0 \qquad (17)$$

Because the second order condition is negative, the solution to (17) will be a lower investment level than that to (16) if

$$-\left(w-i\right)^{-\sigma}+\left(\frac{1}{2}h\left(i,i\right)\right)^{-\sigma}\frac{\partial h\left(i,i\right)}{\partial i}>-\left(w-i\right)^{-\sigma}+\left(h\left(i,i'\right)-i'\right)^{-\sigma}\frac{\partial h\left(i,i'\right)}{\partial i}$$

using the fact that $h\left(i^{m},i^{f}\right)$ has constant returns to scale this implies

$$\left(i\frac{\partial h\left(i,i\right)}{\partial i}\right)^{-\sigma}\frac{\partial h\left(i,i\right)}{\partial i}>\left(i\frac{\partial h\left(i,i'\right)}{\partial i}+i'\left(\frac{\partial h\left(i,i'\right)}{\partial i'}-1\right)\right)^{-\sigma}\frac{\partial h\left(i,i'\right)}{\partial i}$$

and because $\frac{\partial h(i,i')}{\partial i'}>1$ this will hold if

$$i^{-\sigma}\left(\frac{\partial h\left(i,i\right)}{\partial i}\right)^{1-\sigma}>i^{-\sigma}\left(\frac{\partial h\left(i,i'\right)}{\partial i}\right)^{1-\sigma}$$

which will hold if $\sigma>1$ since $i'>i$ and thus $\frac{\partial h(i,i')}{\partial i}>\frac{\partial h(i,i)}{\partial i}$. ■

B Nash Bargaining

A special and traditional case of the model presented above is one where the sharing decision is made through Nash bargaining and thus where consumption levels will be given by

$$\begin{aligned}c_2^m &= i^m+\beta\left(z\right)\left(h\left(i^m,i^f\right)-i^m-i^f\right)\\ c_2^f &= i^f+\left(1-\beta\left(z\right)\right)\left(h\left(i^m,i^f\right)-i^m-i^f\right)\end{aligned}$$

We will assume that the sharing parameter is influenced by bargaining power. It is easy to show that in this case, returns to investments when single are equal to 1 and thus positive.

Define the surplus to be shared as

$$g(i^m,i^f)=h\left(i^m,i^f\right)-i^m-i^f$$

Assume $h\left(\cdot\right)$ displays constant returns to scale and thus $g\left(\cdot\right)$ also does.

Proposition 4 *There exists a unique pure strategies Nash Equilibrium in this setting.*

Proof. Since g exhibits constant returns to scale in investments, so will $c_2^m$and c_2^f. It is easy to show that conditions (1) are satisfied since

$$\begin{aligned} c_2^m\left(i^m,0,z\right) &= i^m \Longrightarrow \mu\left(i^m,0,z\right)=1 \\ c_2^f\left(0,i^f,z\right) &= i^f \Longrightarrow \mu\left(0,i^f,z\right)=0 \end{aligned}$$

and defining

$$c_2^m\left(i^m,i^f,z\right)=\alpha\left(i^m,i^f,z\right)h\left(i^m,i^f\right)$$

where

$$\alpha\left(i^m,i^f,z\right)=\beta(z)+\frac{i^m\left(1-\beta(z)\right)-\beta(z)\,i^f}{h\left(i^m,i^f\right)}$$

$$\begin{aligned} \frac{\partial\mu}{\partial i^m} &= \frac{\partial\mu}{\partial\alpha}\left(\frac{\left(1-\beta(z)\right)\left(h\left(i^m,i^f\right)-i^m\frac{\partial h}{\partial i^m}\right)+\beta(z)\,i^f\frac{\partial h}{\partial i^m}}{h\left(i^m,i^f\right)^2}\right)>0 \\ \frac{\partial\mu}{\partial i^f} &= \frac{\partial\mu}{\partial\alpha}\left(\frac{-\beta(z)\left(h\left(i^m,i^f\right)-i^f\frac{\partial h}{\partial i^f}\right)-i^m\left(1-\beta(z)\right)\frac{\partial h}{\partial i^f}}{h\left(i^m,i^f\right)^2}\right)<0 \end{aligned}$$

■

A rise in the sex ratio will lead to an increase in men's investments if (15) holds which in this case implies

$$\sigma>1>\min\left\{\frac{\left(i^m+\beta g\right)\frac{\partial g}{\partial i^m}}{\left(1+\beta\frac{\partial g}{\partial i^m}\right)g},\frac{\left(i^f+(1-\beta)\,g\right)\frac{\partial g}{\partial i^f}}{\left(1+(1-\beta)\frac{\partial g}{\partial i^f}\right)g}\right\}$$

And this is satisfied since $g\left(i^m,i^f\right)$ exhibits constant returns to scale. Investments will be strategic substitutes if

$$\sigma>\min\left\{\frac{\left(i^m+\beta g\right)\frac{\partial^2 g}{\partial i^m\partial i^f}}{\left(1+\beta\frac{\partial g}{\partial i^m}\right)\frac{\partial g}{\partial i^f}},\frac{\left(i^f+(1-\beta)\,g\right)\frac{\partial^2 g}{\partial i^m\partial i^f}}{\left(1+(1-\beta)\frac{\partial g}{\partial i^f}\right)\frac{\partial g}{\partial i^m}}\right\}$$

for which a sufficient condition is

$$\sigma>\frac{1}{\rho_g}$$

where ρ_g is the elasticity of substitution between the two inputs in the surplus function $g\left(i^m,i^f\right)$.

Table 1: Summary statistics-Outcomes

	Males			**Females**		
	N. Obs.	Mean	Sd	N. Obs.	Mean	Sd
General characteristics						
Age	203954	40.31	16.66	219800	41.26	17.22
Mother foreign	203954	0.79	0.41	219800	0.78	0.41
Father foreign	203954	0.88	0.32	219800	0.88	0.32
Pre-marital investments						
Literate	59946	0.99	0.10	62272	0.99	0.09
Duncan Index	59946	19.99	19.64	62272	15.82	22.48
Wage Index	59946	16.71	12.37	62272	9.14	11.15
Years of education	109252	9.44	3.34	121352	9.45	3.08
Labor supply						
Employed	127283	0.81	0.39	157528	0.30	0.46
In the labor force	144008	0.86	0.34	139844	0.29	0.45
Hours last week	105985	34.61	20.95	119725	10.73	18.06
Weeks last year	107553	39.53	18.61	119752	14.31	21.41
Marital status						
Never married (after 35)	113709	0.12	0.32	123310	0.11	0.32
Divorced	203954	0.02	0.14	219800	0.02	0.14
Widowed	203954	0.03	0.16	219800	0.11	0.31
Married more than once	87070	0.09	0.29	104453	0.09	0.29
Age at first marriage (older than 35)	48712	26.92	6.33	59017	23.33	5.80
Married to an immigrant of own ethnic group	203679	0.03	0.18	219564	0.08	0.26

All summary statistics are weighted by Census sample-line weights

Table 2: Summary statistics-Marriage market conditions and instrument

Ethnic group	**Immigrants**		**Foreign stock**		**Instrumented**	
	Flow (000s)	Sex ratio	Flow (000s)	Sex ratio	Flow (000s)	Sex ratio
British Ancestry	0.138	0.941	0.605	0.918	0.124	0.851
	[0.136]	[0.546]	[0.483]	[0.158]	[0.108]	[0.067]
Francophone	0.029	1.012	0.112	0.948	0.027	0.882
	[0.033]	[0.734]	[0.102]	[0.424]	[0.030]	[0.174]
South Europeans	0.252	1.536	0.451	1.347	0.256	1.427
	[0.294]	[1.118]	[0.396]	[1.010]	[0.317]	[0.365]
Hispanics	0.129	1.306	0.195	1.174	0.162	1.375
	[0.113]	[0.783]	[0.163]	[0.594]	[0.180]	[0.185]
Scandinavian	0.059	1.724	0.274	1.142	0.069	1.330
	[0.056]	[1.425]	[0.239]	[0.347]	[0.075]	[1.127]
Germanic	0.183	1.404	0.704	1.038	0.177	1.262
	[0.220]	[0.697]	[0.555]	[0.260]	[0.225]	[0.104]
Russians and others	0.407	1.068	0.707	1.065	0.399	1.044
	[0.495]	[0.429]	[0.635]	[0.314]	[0.520]	[0.271]
Other Europe	0.095	1.410	0.175	1.329	0.093	1.252
	[0.101]	[1.337]	[0.131]	[1.215]	[0.112]	[0.643]
Other countries	0.043	2.609	0.058	2.191	0.036	2.257
	[0.039]	[2.249]	[0.055]	[1.800]	[0.040]	[1.323]

Standard deviations in brakets

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

Table 3: First stage

	Sex ratio of immigrants	Sex ratio of foreign stock	Flow of immigrants	Flow of foreign stock
	(1)	(2)	(3)	(4)
Predicted sex ratio of immigrants	0.835***	0.434***	-0.011	-0.007
	(0.179)	(0.084)	(0.007)	(0.023)
Predicted flow of immigrants	0.165	0.051	0.868***	0.573***
	(0.199)	(0.069)	(0.041)	(0.059)
N. Obs	2343	2343	2343	2343
R-squared	0.379	0.362	0.986	0.967

Standards errors clustered at the state level in parantheses
* significant at 5%; ** significant at 1%; *** significant at 0.1%
All regressions include state, ethnic groups, immigration period fixed effects and all double interactions.
All regressions are weighted by the size of the total foreign stock in each cell.

Table 4: First stage-Robustness checks for immigrant measures

	1910-1929	1900-1919	In logs	Without NY	No missing	Census table
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Sex ratio						
Predicted sex ratio	0.916***	0.798***	0.693***	0.867***	0.867***	0.571**
	(0.246)	(0.197)	(0.126)	(0.187)	(0.159)	(0.201)
Predicted flow	0.130	0.098	-0.046	0.619	0.170	-0.100
	(0.212)	(0.226)	(0.078)	(0.403)	(0.241)	(0.142)
R-squared	0.273	0.442	0.529	0.360	0.366	0.341
Panel B: Flow						
Predicted sex ratio	-0.010	-0.013	0.013	-0.013*	-0.012	-0.009
	(0.006)	(0.012)	(0.194)	(0.006)	(0.009)	(0.023)
Predicted flow	0.907***	0.846***	0.466***	0.801***	0.861***	0.648***
	(0.034)	(0.065)	(0.104)	(0.138)	(0.045)	(0.035)
R-squared	0.989	0.986	0.966	0.951	0.984	0.967
N. Obs	1556	1606	1748	2289	1748	1909

Standards errors clustered at the state level in parantheses
* significant at 5%; ** significant at 1%; *** significant at 0.1%
All regressions include state, ethnic groups, immigration period fixed effects and all double interactions.
All regressions are weighted by the size of the total foreign stock in each cell.

Table 5: Marriage market outcomes

	Ever married	Divorced	Ever married twice	Married to own ethnic immigrant	Age at first marriage
	(1)	(2)	(3)	(4)	(5)
			Immigrants		
Panel A: Males					
Sex ratio of immigrants	0.033*	0.005	0.026*	-0.017	-0.027
	(0.014)	(0.003)	(0.011)	(0.009)	(0.371)
Flow of immigrants	0.002	-0.003	0.004	0.020*	0.881**
	(0.013)	(0.003)	(0.012)	(0.009)	(0.296)
N. Obs	203954	203954	87070	203679	48712
Panel B: Females					
Sex ratio of immigrants	0.001	-0.002	-0.004	-0.002	-0.009
	(0.009)	(0.005)	(0.010)	(0.015)	(0.301)
Flow of immigrants	0.008	-0.001	-0.010	0.026*	-0.463*
	(0.011)	(0.004)	(0.016)	(0.011)	(0.210)
N. Obs	219800	219800	104453	219564	60276
			Foreign Stock		
Panel A: Males					
Sex ratio of foreign stock	0.081*	0.011	0.145	-0.050*	-0.036
	(0.034)	(0.010)	(0.094)	(0.024)	(1.091)
Flow of foreign stock	0.020	-0.003	0.011	0.035	1.700**
	(0.034)	(0.006)	(0.028)	(0.018)	(0.590)
N. Obs	203954	203954	87070	203679	48712
Panel B: Females					
Sex ratio of foreign stock	0.003	-0.007	-0.012	-0.006	-0.063
	(0.029)	(0.015)	(0.032)	(0.051)	(1.023)
Flow of foreign stock	0.016	-0.003	-0.021	0.048*	-0.817*
	(0.021)	(0.007)	(0.031)	(0.022)	(0.369)
N. Obs	219800	219800	104453	219564	60276

Standard errors clustered at the state level in parantheses.
* significant at 5%; ** significant at 1%; *** significant at 0.1%
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 parents' nativity.
All regressions are weighted by the Census sample-line weight.

Table 6: Pre-marital investments

	Literacy	Duncan Index	Wage index	Highest grade	Literacy	Duncan Index	Wage index	Highest grade
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Immigrants				
Panel A: Males		**OLS**				**IV**		
Sex ratio of immigrants	0.000	-0.015	0.085	0.026	0.017*	2.351*	0.788	0.443*
	(0.000)	(0.130)	(0.069)	(0.017)	(0.008)	(0.971)	(0.398)	(0.173)
Flow of immigrants	0.003	2.761*	1.382**	0.049	0.001	1.887	0.960	-0.039
	(0.005)	(1.224)	(0.445)	(0.145)	(0.007)	(1.684)	(0.614)	(0.208)
N. Obs	59946	59946	59946	109252	59946	59946	59946	109252
Panel B: Females								
Sex ratio of immigrants	-0.001*	-0.116	-0.106*	-0.001	0.008	-0.596	-0.696	-0.052
	(0.000)	(0.082)	(0.049)	(0.010)	(0.010)	(0.697)	(0.490)	(0.085)
Flow of immigrants	-0.003	3.700**	1.821*	-0.037	-0.002	3.562*	2.077**	-0.101
	(0.005)	(1.529)	(0.749)	(0.157)	(0.007)	(1.401)	(0.758)	(0.207)
N. Obs	62272	62272	62272	121352	62272	62272	62272	121352
				Foreign Stock				
Panel A: Males		**OLS**				**IV**		
Sex ratio of foreign stock	0.000	0.387	0.226	-0.037	0.039	5.542	1.862	1.789*
	(0.001)	(0.200)	(0.118)	(0.032)	(0.021)	(2.968)	(1.147)	(0.874)
Flow of foreign stock	-0.009	0.785	0.682	-0.284*	0.004	3.793	1.895	-0.022
	(0.006)	(1.121)	(0.662)	(0.133)	(0.011)	(2.490)	(0.980)	(0.578)
N. Obs	59946	59946	59946	109252	59946	59946	59946	109252
Panel B: Females								
Sex ratio of foreign stock	-0.001	0.412	0.0446	-0.007	0.051	-3.484	-4.333	-0.164
	(0.003)	(0.556)	(0.250)	(0.027)	(0.062)	(4.454)	(3.345)	(0.282)
Flow of foreign stock	0.001	-0.677	0.508	-0.291	-0.002	6.682*	3.817*	-0.212
	(0.004)	(1.027)	(0.597)	(0.168)	(0.010)	(3.145)	(1.701)	(0.398)
N. Obs	62272	62272	62272	121352	62272	62272	62272	121352

Standard errors clustered at the state level in parantheses.
* significant at 5%; ** significant at 1%; *** significant at 0.1%
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 parents' nativity.
All regressions are weighted by the Census sample-line weight.

Table 7: Pre-marital investments-Robustness checks for males

	All	Excluding oldest cohort	Excluding youngest cohort	Without NY	With ethnicities dummies	Younger than 65	Excluding 1940
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Highest grade attained			
Sex ratio of immigrants	0.443*	0.409*	0.538	0.470**	0.510*	0.431*	0.465**
	(0.173)	(0.170)	(0.525)	(0.165)	(0.224)	(0.177)	(0.164)
Flow of immigrants	-0.039	0.036	-0.336	0.584	-0.028	-0.130	0.182
	(0.208)	(0.220)	(0.307)	(0.422)	(0.186)	(0.208)	(0.221)
N. Obs	109252	101195	79446	89533	109252	86807	92703
				Duncan Index			
Sex ratio of immigrants	2.351*	2.390*	5.576*	1.809	2.768*		
	(0.971)	(0.935)	(2.647)	(0.949)	(1.114)		
Flow of immigrants	1.887	1.725	-1.254	-2.111	1.755		
	(1.684)	(1.661)	(2.588)	(2.861)	(1.715)		
N. Obs	59946	49068	53005	49818	59946		

Standard errors clustered at the state level in parantheses.
* significant at 5%; ** significant at 1%; *** significant at 0.1%
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 parents' nativity.
All regressions are weighted by the Census sample-line weight.

Table 8: Labor supply

	In LF (1)	Employed (2)	Hours (3)	Weeks (4)	In LF (5)	Employed (6)	Hours (7)	Weeks (8)
				Immigrants				
Panel A: Males		**OLS**				**IV**		
Sex ratio of immigrants	0.002	0.002	0.145*	0.145*	-0.003	-0.004	0.534	0.523
	(0.001)	(0.001)	(0.065)	(0.065)	-0.007	(0.009)	(0.690)	(0.359)
Flow of immigrants	0.012	0.009	1.504*	1.504*	0.004	0.003	1.412	0.494
	(0.014)	(0.012)	(0.616)	(0.616)	(0.012)	(0.012)	(0.735)	(0.773)
N. Obs	144008	127283	105985	107553	144008	127283	105985	107553
Panel B: Females								
Sex ratio of immigrants	0.001	0.001	0.038	0.004**	-0.038	-0.028	-1.387	-1.308
	(0.002)	(0.002)	(0.062)	(0.001)	(0.019)	(0.018)	(0.827)	(0.771)
Flow of immigrants	-0.003	-0.007	0.579	0.001	-0.003	-0.005	0.520	0.367
	(0.010)	(0.011)	(0.710)	(0.023)	(0.015)	(0.015)	(0.846)	(1.142)
N. Obs	157528	139844	119725	119752	157528	139844	119725	119752
				Foreign Stock				
Panel A: Males		**OLS**				**IV**		
Sex ratio of foreign stock	0.002	0.001	-0.158	-0.000	-0.009	-0.015	2.504	2.242
	(0.002)	(0.003)	(0.225)	(0.004)	(0.021)	(0.027)	(3.755)	(1.943)
Flow of foreign stock	-0.014	-0.012	-0.433	-0.020	0.007	0.005	2.993	1.053
	(0.011)	(0.011)	(0.822)	(0.024)	(0.022)	(0.021)	(1.758)	(1.713)
N. Obs	144008	127283	105985	107553	144008	127283	105985	107553
Panel B: Females								
Sex ratio of foreign stock	-0.001	-0.001	-0.272	-0.003	-0.109*	-0.084	-4.398	-4.113
	(0.004)	(0.004)	(0.181)	(0.006)	(0.051)	(0.046)	(2.323)	(2.149)
Flow of foreign stock	0.008	0.008	1.052	0.000	-0.018	-0.018	0.775	0.481
	(0.013)	(0.014)	(0.679)	(0.015)	(0.019)	(0.020)	(1.243)	(1.806)
N. Obs	157528	139844	119725	119752	157528	139844	119725	119752

Standard errors clustered at the state level in parantheses.
* significant at 5%; ** significant at 1%; *** significant at 0.1%
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 parents' nativity.
All regressions are weighted by the Census sample-line weight.

Table 9: 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)
				Full sample				
Panel A: Males	**OLS**		**IV**		**OLS**		**IV**	
Own education	0.007***	0.688***	-0.014	-1.439**	0.006***	0.531***	-0.026	-3.665**
	(0.000)	(0.032)	(0.013)	(0.521)	(0.000)	(0.023)	(0.029)	(1.352)
Spouse's education					0.003***	0.370***	0.024	2.833**
					(0.000)	(0.019)	(0.018)	(1.007)
F-test (instruments): own			11.79***				10.20***	
F-test (instruments): spouse							28.29***	
N. Obs	682362	658532	682362	658532	510355	491396	510355	491396
Panel B: Females								
Own education	0.016***	0.669***	-0.018	-0.511	0.024***	0.966***	-0.030	-1.926
	(0.001)	(0.038)	(0.015)	(0.659)	(0.001)	(0.029)	(0.044)	(1.849)
Spouse's education					-0.010***	-0.345***	0.034	2.546
					(0.001)	(0.025)	(0.051)	(2.159)
F-test (instruments): own			9.40***				16.48***	
F-test (instruments): spouse							7.54***	
N. Obs	655981	647750	655981	647750	499269	492756	499269	492756
				Only second generation Americans				
Panel A: Males	**OLS**		**IV**		**OLS**		**IV**	
Own education	0.003***	0.517***	0.008	-0.575	0.004***	0.473***	-0.005	-3.236
	(0.000)	(0.044)	(0.027)	(1.349)	(0.000)	(0.042)	(0.034)	(2.181)
Spouse's education					0.001***	0.311***	0.017	3.087
					(0.000)	(0.026)	(0.027)	(1.874)
F-test (instruments): own			5.58***				11.31***	
F-test (instruments): spouse							20.44***	
N. Obs	104819	102015	104819	102015	76478	74340	76478	74340
Panel B: Females								
Own education	0.010***	0.380***	-0.045**	-1.695**	0.019***	0.695***	-0.087	-2.664
	(0.002)	(0.058)	(0.014)	(0.535)	(0.002)	(0.057)	(0.044)	(1.553)
Spouse's education					-0.009***	-0.323***	0.091	2.243
					(0.001)	(0.040)	(0.060)	(2.080)
F-test (instruments): own			14.03***				6.05***	
F-test (instruments): spouse							4.75***	
N. Obs	99519	98201	99519	98201	74828	73789	74828	73789

Standard errors clustered at the state level in parantheses.
* significant at 5%; ** significant at 1%; *** significant at 0.1%
All regressions include state, year of birth and Census year fixed effets as well as age and age squared.
Columns (5) and (6) also include spouse's state of birth and year of birth fixed effects.
All regressions are weighted by the Census sample-line weight.

Table 10: Calibration exercise

	λ=1/3			λ=1/2			λ=2/3		
	w=22	w=26	w=30	w=22	w=26	w=30	w=22	w=26	w=30
					Spouse selection model				
Panel A: Males									
α	0.36	0.32	0.29	0.42	0.37	0.34	0.46	0.41	0.37
σ	4.88	4.96	4.75	4.93	4.83	4.86	4.82	4.98	4.59
Returns in marriage	0.04	0.04	0.05	0.04	0.04	0.05	0.04	0.04	0.05
Panel B: Females									
α	0.34	0.29	0.26	0.40	0.34	0.30	0.45	0.38	0.33
σ	5.00	4.98	4.77	4.95	4.93	4.99	5.00	5.00	4.99
Returns in marriage	0.04	0.05	0.05	0.04	0.05	0.05	0.04	0.05	0.05
					Bargaining model				
Panel A: Males									
α	0.40	0.30	0.27	0.51	0.39	0.33	0.63	0.47	0.39
σ	8.44	7.81	8.32	8.03	8.41	8.03	7.06	8.31	8.40
Returns in marriage	0.04	0.04	0.05	0.02	0.03	0.04	0.02	0.02	0.03
Panel B: Females									
α	0.45	0.36	0.31	0.42	0.34	0.29	0.39	0.31	0.27
σ	8.48	8.08	8.17	8.33	8.29	7.47	8.50	8.17	8.14
Returns in marriage	0.03	0.04	0.04	0.03	0.04	0.04	0.03	0.04	0.04
Panel C: Joint									
α	0.30	0.23	0.20	0.51	0.39	0.33	0.18	0.13	0.11
σ	4.46	3.83	3.67	8.05	8.50	8.50	2.56	2.22	2.00
Male returns in marriage	0.03	0.04	0.05	0.02	0.03	0.04	0.05	0.05	0.05
Female returns in marriage	0.04	0.05	0.05	0.02	0.03	0.04	0.05	0.05	0.05

Figure 1: Sex ratios of Scandinavians by country and state

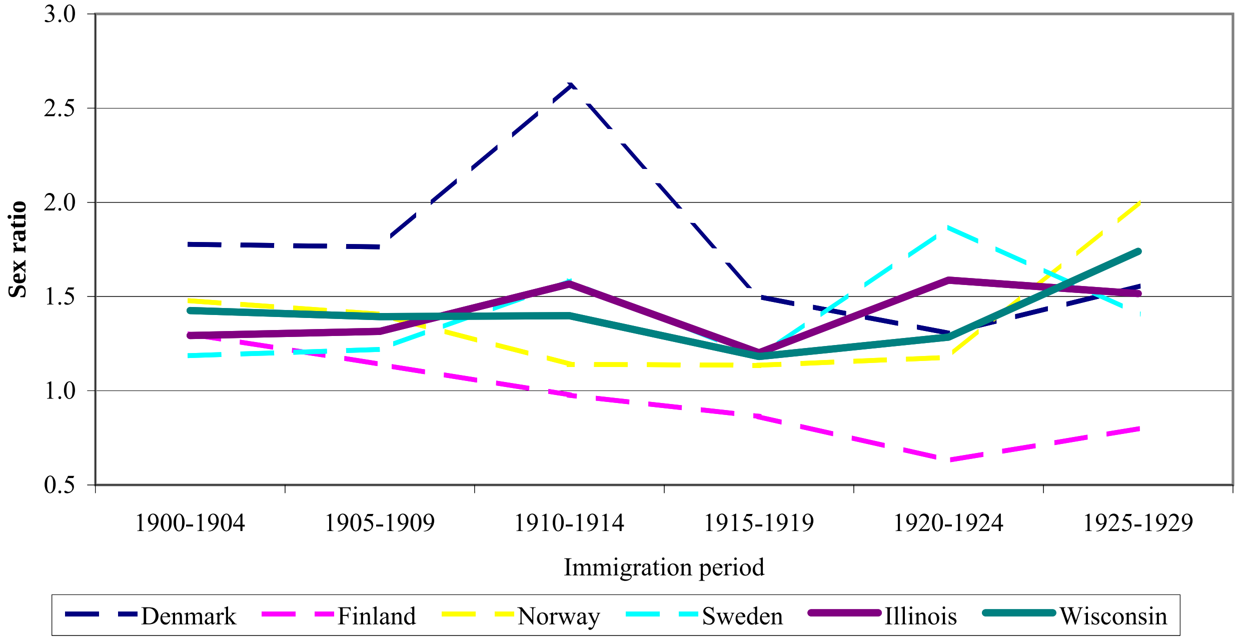
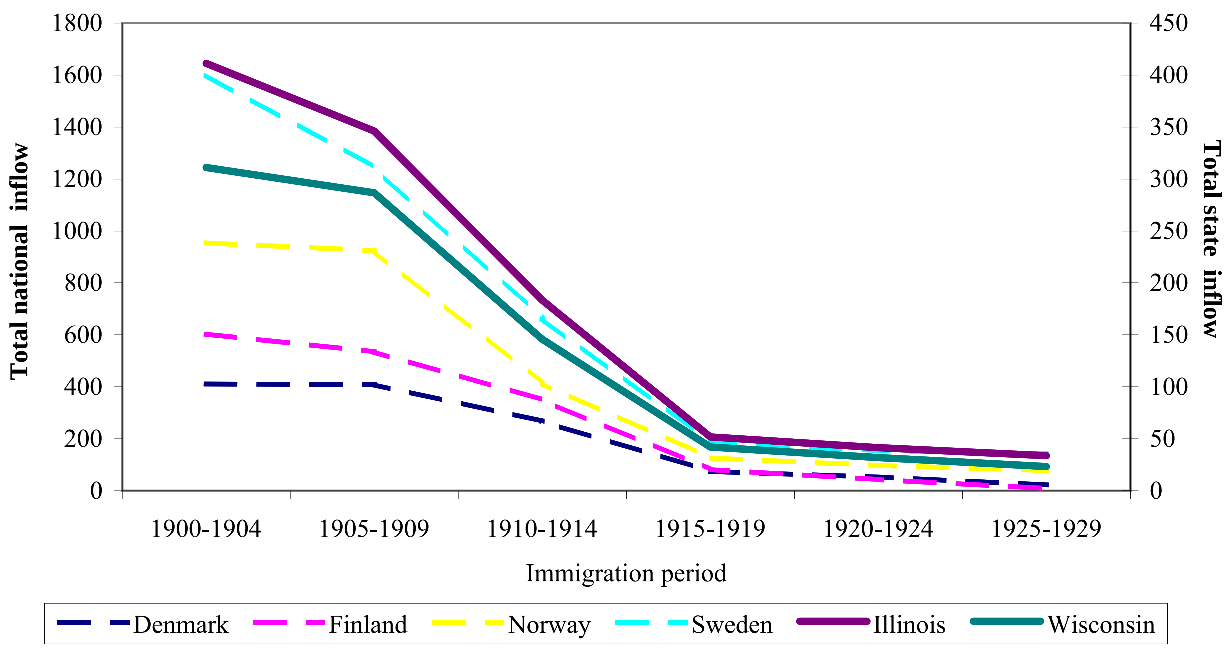


Figure 2: Flows of Scandinavian immigrants by country and state



Appendix Table A1: Data description

Variables	Census years	Age sampled	Details
Marital outcomes			
Ever married	1900-70	15+	
Currently married to a same ethnic immigrant	1900-70	15+	
Currently divorced	1900-70	15+	
Number of marriages	1910 1940-60	15+	
Age at first marriage	1930-40 1960-70	35+	
Pre-marital investments			
Literacy	1900-30	15-25	Literacy in any language
Highest grade achieved	1940-70	25+	Only available from 1940, when the youngest cohort is 25
Duncan index	1900-30	15-25	Based on a measure of prestige linked to wage and education
Wage index	1900-30	15-25	Based on 1950 wages
Post-marital labor supply			
In the labor force	1910-70	25+	
Employed	1910-10 1930-70	25+	
Hours worked per week	1940-70	25+	Transformed from intervals to a continuous variable by selecting the mid-point of the interval
Weeks worked per year	1940-70	25+	

Appendix Table A2: Ethnic group composition

Ethnic group	Countries of Birth
British Ascendance	Australia, English Canada, English, Ireland, Scotland and Wales
Francophone	Belgium, French Canada and France
South Europeans	Italy, Spain and Portugal
Hispanics	Mexico, Cuba, Other West Indies, Central America and South America
Scandinavian	Denmark, Finland, Norway and Sweden
Germanic	Austria, Germany, Luxemburg, Netherlands and Switzerland
Russians and others	Russia, Poland and Romania
Other Europe	Bohemia (Czechoslovakia), Greece, Hungary and Other Europe
Other Countries	Africa, Atlantic Islands, China, India, Japan, Other Asia, Pacific Islands, Turkey and Other countries

Appendix Table A3: Spatial distribution of immigrants by ethnic group

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