University of Toronto
Department of Economics

Working Paper 333

Are there Increasing Returns in Marriage Markets?

By Maristella Botticini and Aloysius Siow

September 8, 2008

Are There Increasing Returns in Marriage Markets?*

Maristella Botticini†
Collegio Carlo Alberto, Università di Torino
and Boston University

Aloysius Siow
University of Toronto

Draft of December 2007

Abstract

The returns to scale of marriage markets have important behavioral and welfare consequences. It is quantitatively difficult to estimate the returns to scale because, due to endogenous migration, the marriage market size is endogenous. This paper addresses the endogeneity in two ways. First, it estimates the degree of returns to scale in U.S. marriage markets using the 2000 census. Given that in the United States people move to cities to find marriage partners and, therefore, the size of the marriage market is endogenous, we instrument the current size of a cohort in the marriage market with the size of that cohort twenty years earlier. Second, it estimates city scale effects in two societies—early Renaissance Tuscany and pre-reform China—where there was little internal mobility, and thus, the size of the marriage market can be considered exogenous. The main finding is that in all three societies, there is no evidence of increasing returns to scale in marriage markets, whereas the hypothesis of constant returns to scale cannot be rejected. This is true when looking at marriage odds ratios, total gains to marriage, and the quality of marital match. Given the different characteristics of the three societies in terms of population size, time period, economic structure, and social norms characterizing the marriage market, the similarity and precision of the estimates for returns to scale parameters is remarkable.

JEL Classification: J12, J41
Keywords: marriage markets, matching, thick market externality, returns to scale, spousal search, endogeneity

*We thank Oriana Bandiera, Hoyt Bleakey, Eugene Choo, Andrew Newman, Robert Shimer, and seminar participants at Boston University, European University Institute (Florence), Georgetown, Northwestern, Paris School of Economics, University College London, and University of Zurich (IWE) for helpful comments. This research was supported by the National Science Foundation (grant SES-0339133). The usual disclaimers apply.

†Research Fellow of CEPR and CHILD.

1 Introduction

The returns to scale in marriage markets have important behavioral and welfare consequences. The seminal paper by Mortensen (1988) raised the possibility of a thick market externality or increasing returns to scale (IRS henceforth) in marriage markets. Under IRS, a larger pool of eligible individuals searching for a marriage partner may lead to a higher marriage rate in that market. From a social welfare perspective, there is too little search in marriage markets with IRS.

In the *theoretical* literature, IRS is a popular assumption in models of marital search. For example, Chiappori and Weiss (2000), and Anderberg and Mongrain (2001), adopt it to motivate multiple equilibria in marriage and divorce rates, whereas Gautier, Svarer, and Teulings (2005) use it to motivate their model of two way urban–rural migration and marital search.

Alternatively, in a marriage market with constant returns to scale (CRS henceforth), the marriage rate is independent of the size of the stock of eligibles, as in the static and dynamic marriage market models of Choo and Siow (2006a, 2006b). Models in the Choo Siow class are efficient by assumption. Lastly, in the case of decreasing returns to scale (DRS henceforth), the marriage rate falls as the stock of eligibles increases. Thus it is important to quantitatively estimate the returns to scale in marriage markets.

In the *empirical* literature, researchers have begun to investigate the degree of returns to scale in a marriage market. Following the empirical literature testing for agglomeration effects in labor markets, a first approach is to use a city as the unit of observation and to regress the marriage rate in a city on the size of the marriage market in that city.[1] The coefficient on city size provides an estimate of the degree of returns to scale in marriage markets. For example, using the 1970, 1980 and 1990 U.S. censuses, Gould and Paserman (2003) estimate a probit regression of the probability of marriage of an eligible woman on individual covariates including a measure of men's wage inequality and log population in the Metropolitan Area in which the woman resides. In the pooled regression, they show that there are IRS in the marriage market. When they used fixed effects for the 321 Metropolitan Areas in their sample, the point estimate for log population becomes statistically insignificant.

The econometric problem that stems from using linear probability models when testing for agglomeration effects is well known. If individuals who want to marry, choose to migrate to cities, then the regression suffers from the problem of endogeneity bias.

There is substantial empirical evidence showing that individuals choose to locate in cities in order to engage in marital search or related activities. For example, Costa and Kahn (2000) argue that educated couples are more likely to live in cities because it is easier for both of them to find suitable jobs. Compton and Pollak (2004) suggest that cities are attractive to both married and single educated individuals, and that the higher observed rate of educated couples in cities relative to other types is partly because of the larger stock of educated eligibles in cities. Gautier, Svarer, and Teulings (2005) show that in Denmark,

[1] Petrongolo and Pissarides (2001) present a survey of the empirical literature testing for agglomeration effects in labor markets and conclude that CRS is a reasonable description of the labor market data.

single educated individuals are more likely to migrate to cities and married educated couples are more likely to leave the cities. In Sweden, single women are more likely to migrate to cities (Edlund 2005). Thus the problem of endogeneity bias in OLS or probit regressions when testing for agglomeration effects in marriage markets has to be addressed.

Using the 1980 U.S. census, Drewianka (2003) deals with this problem by regressing the propensity to marry for individuals of a particular gender on individual attributes and the sex ratio (ratio of eligible men to women) of the city in which the individual resides. He concludes that there are IRS in the marriage market. However, if the sex ratio in a city is endogenous as argued above, the endogeneity problem is not solved.

The objective of our paper is to estimate the degree of returns to scale in marriage markets by regressing the marriage rate (or a related variable) in a city on the number of eligibles in that city. We deal with the endogeneity problem in two ways. First, when moving to marry is an important concern, we implement an instrumental variables strategy. Second, we use data from societies in which we can make an *a priori* argument that moving to marry is not an issue.

We use data from three societies that are different in many dimensions (e.g., time period, area and population size, economic structure, and social norms governing the marriage markets).

First, we study the United States as described in the 2000 census. Clearly, the assumption of no in- or out-migration in cities is untenable here. To deal with the endogeneity problem, for each city we use the size of the cohort twenty years earlier as an instrument for the size of the same cohort in the year 2000. For example, when we analyze the marriage rate of 25–29 years old women in the 2000 census, we use the number of 5–9 years old women in the 1980 census in that city as an instrument for the number of 25–29 year old women in 2000. Our assumption is that children (or their parents) do not choose the city in which to reside based on their marital prospects twenty years later.

Next, we consider two societies in which the endogeneity problem is not an issue because people do not migrate for marriage purposes.

The first society is the city of Florence and her dominions consisting of Tuscan towns and hundreds of villages in the countryside. Data on socio-economic characteristics come from the 1427 Florentine *Catasto*, a census and property survey of about 60,000 Tuscan households. The endogeneity problem is not a concern for early Renaissance Tuscany: a sample of more than 7,200 dowry (marriage) contracts, which we collected at the State Archives of Florence, indicate that Tuscan people did not migrate to find marriage partners.

The second society is the People's Republic of China as described in the 1982 census. Until 1978, internal mobility and particularly migration to cities was severely limited.[2] After 1978, farmers began to obtain the right to temporarily migrate to small- and medium-size cities to look for work. Even today, internal migration is restricted in China. Because the census we use was conducted in 1982, we can assume that the population in each city is essentially free from in-migration.

The main finding is that in all three societies, there is no evidence of increasing returns to scale when looking at marriage rate regressions. The hypothesis of constant returns

[2]See Goldstein and Goldstein (1990), Goldstein and Yang (1990), and Yang (1996a, 1996b).

to scale cannot be rejected. More in detail, marriage rate regressions by gender at the city level do not lend support for strong departures from the hypothesis of CRS in early Renaissance Tuscan marriage markets and contemporary U.S. marriage markets. Chinese marriage markets do not display IRS or CRS, but the point estimates suggest a mild DRS.

The same findings hold true when looking at total gains to marriage regressions. The reason for estimating total gains to marriage regressions (in addition to marriage rate regressions) is that the marriage rates of men and women in a marriage market must be related. The marriage gain statistic, which has been proposed by Choo and Siow (2006), is a bivariate extension of the univariate gender-specific marriage rates. The empirical results on marriage rates and total gains to marriage from the three data sets and different empirical methods are consistent with each other: the hypothesis of CRS cannot be rejected.

An important caveat is in order. Even if we find that the *marriage rate* in a city is independent of the number of eligibles in that city, a thick market externality may still exist. The reason is that agents may react to a thick market externality, with faster arrivals of potential partners and/or more disperse match value distribution, by waiting for a better match rather than marrying earlier. In this case, the thick market externality generates higher *marital output* but not necessarily a higher marriage rate. Because women have a finite amount of time to bear children, they may choose to marry in a narrow age window even if the pool of potential partners is poor. If the pool of partners increases, women may continue to marry in that age window but now they may be able to find better matches.

Addressing this issue is similar to what has been done in the labor markets literature. For example, Petrongolo and Pissarides (2005) estimate a structural model of search and unemployment with scale effects. They show that the unemployment hazard is unaffected by scale effects. Scale effects lead to higher reservation wages and higher post-employment wages. This means that it is possible for scale effects to show up in the quality of the match rather than in a higher matching rate.

For the U.S. 2000 sample we study scale effects on the average quality of the marital match using two proxies of average marital output: (i) the fraction of young children who live with two parents, and (ii) the educational attainment of children within marriage. We do not use the divorce rate or marital tenure to proxy for marital quality because one can argue that married individuals in larger cities may also choose to break up more and rematch more and they still have higher average welfare. The marital output evidence does not support the hypothesis that larger marriage markets result in higher average marital output, that is, there is no evidence of a thick market externality even when we consider the quality of the marital match.

The paper is organized as follows. Section 2 describes our methodological approach. Sections 3, 4, and 5 present the datasets and empirical findings for the United States in 2000, Tuscany in 1427, and China in 1982, respectively. Section 6 concludes.

2 Methodology

Marriage Odds Ratio Regressions. Using a city as a unit of observation, let c denote city c, f denote type f women and m denote type m men. μ_f^c is the number of married women of type f in city c. n_f^c is the number of eligible women of type f in city c. μ_m^c is the number of married men of type m in city c. n_m^c is the number of eligible men of type m in city c. Consider the following cross-section log marriage odds ratio regressions:

$$\ln \frac{\mu_g^c}{n_g^c - \mu_g^c} = \pi_g + \alpha_g \ln n_g^c + u_g^c \qquad g = f, m \tag{1}$$

where u_g^c are the error terms of the regressions.

The marriage rate $(\frac{\mu_g^c}{n_g^c})$ is monotonically increasing in the marriage odds ratio $\left(\frac{\mu_g^c}{n_g^c - \mu_g^c}\right)$. The behavioral justification for studying the log odds ratio is that it can be derived from McFadden's (1974) random utility model where the choice is between marriage or otherwise.[3] Under McFadden's interpretation, the log odds ratio, that is the left-hand side of (1), measures the mean difference in utility of type g in city c from marrying versus not marrying.

Estimating the parameters α_f and α_m enables to assess the degree of returns to scale in marriage markets. Under CRS, $\alpha_f = \alpha_m = 0$. Under IRS, $\alpha_f > 0$ and $\alpha_m > 0$. Under DRS, $\alpha_f < 0$ and $\alpha_m < 0$.

To understand the empirical findings in the next sections, it is important to discuss orders of magnitude. In general, for any city, a doubling of its population is a very large change. Suppose that the initial marriage rate in city c, $\frac{\mu_g^c}{n_g^c}$, is 0.8. Then a doubling of the city's population will increase the marriage rate to 0.801 if $\alpha_g = 0.01$, and to 0.81 if $\alpha_g = 0.1$. In contrast, let the initial marriage rate in city c, $\frac{\mu_g^c}{n_g^c}$, be 0.5. Then a doubling of the city's population will increase the marriage rate to 0.5017 if $\alpha_g = 0.01$, and to 0.517 if $\alpha_g = 0.1$. This implies that population growth in a city will generate a more modest increase in the marriage rate if the initial marriage rate in that city is higher.

In general, the difficulty with estimating the above regressions by OLS is the potential endogeneity of n_m^c and n_f^c due to endogenous migration to cities to find marriage partners. In order to obtain consistent OLS estimates of α_f and α_m, one has to make an *a priori* argument that n_m^c and n_f^c are exogenous. Net migration across cities due to differences in labor market conditions and ameneties is fine because migration unrelated to marital behavior generates exogenous variation in n_m^c and n_f^c across marriage markets, which is what one needs for OLS to be consistent. Otherwise, one has to find instruments for n_m^c and n_f^c.

A standard concern in regression inference is when a covariate, n_g^c in our case, appears on both the right- and left-hand sides of the equation. If this covariate is measured with error, then measurement error may cause a correlation between the covariate and the dependent variable causing the OLS estimate of α_g to be inconsistent. Since there is sampling error in observed n_g^c, the estimation strategy may have a potential problem. However, this is not a first order problem in our case because the marriage odds ratio is independent of the sampling error in n_g^c.[4] There is still the conventional problem of measurement error in observed n_g^c. We will deal with this issue in two ways. First, we estimate population weighted

[3] This model is the workhorse model in the empirical discrete choice literature.

[4] Let $n_g^{c*} = n_g^c + u_g^c$ where n_g^{c*} is the measured population, n_g^c is the true population and u_g^c is the

regressions where smaller cities, which should suffer from more sampling error, have less weight. Second, our instrumental variables technique should alleviate the measurement error problem.

Total Gains to Marriage Regressions. Independent of the problem of endogeneity bias, μ_f^c and μ_m^c are not independent of each other. Because each heterosexual marriage consists of a man and a woman, the marriage rate of type m men must be related to the marriage rate of type f women if they marry each other. However, estimating (1) for men and women separately does not impose the restriction that α_f and α_m are related. Type f women can marry type m men, as well as other types of men, and type m men can marry type f women, as well as other types of women. These substitution possibilities make the relationship between α_f and α_m complicated. As Angrist (2004) and Choo and Siow (2006) have shown, separate male and female marriage rate regressions can give conflicting results.

To deal with the above problem, we also estimate total gains to marriage regressions. Let the total gains to an $\{m, f\}$ marriage be π_{mf}:

$$\pi_{mf} = \ln \sqrt{\frac{\mu_{mf}^c}{\mu_{m0}^c} \frac{\mu_{mf}^c}{\mu_{f0}^c}} \tag{2}$$
$$\mu_{m0}^c = n_m^c - \mu_m^c;\ \mu_{f0}^c = n_f^c - \mu_f^c$$

where μ_{mf}^c is the number of $\{m, f\}$ marriages in city c. μ_{m0}^c and μ_{f0}^c are the numbers of type m men and type f women in city c who choose to remain unmarried, respectively.

Following Choo and Siow (2006), who derive total gains to marriage by embedding McFadden's random utility model in a marriage market, we behaviorally interpret total gains to marriage as an average of the mean utilities of the two types of individuals married to each other relative to not marrying. That is, if the log marriage odds ratio $\frac{\mu_{mf}}{\mu_{m0}}$ $(\frac{\mu_{mf}}{\mu_{f0}})$ represents the mean utility that a type m man (type f woman) obtains from marrying a type f woman (type m man) relative to remaining unmarried, in (2) the total gain to a $\{m, f\}$ marriage is the average of the male and female log marriage odds ratios.

To estimate the degree of returns to scale in marriage markets, consider the following total gains to marriage regression:

$$\pi_{mf} = \ln \sqrt{\frac{\mu_{mf}^c}{\mu_{m0}^c} \frac{\mu_{mf}^c}{\mu_{f0}^c}} = \rho_{mf} + \alpha_{mf}(\frac{\ln n_m^c + \ln n_f^c}{2}) + v_{mf}^c \tag{3}$$

ρ_{mf} estimates the common total gains to a type m, f marriage across all cities. α_{mf} measures the degree of returns to scale. If $\alpha_{mf} = 0$, the marriage market is characterized by CRS. If $\alpha_{mf} > 0$, the marriage market displays IRS. v_{mf}^c is an error term, which allows for idiosyncratic deviations of total gains to marriage across different cities.

---

sampling error. As long as the marriage rate, r_g^c, is the same for the true population and sampling error, $\frac{\mu_g^{c*}}{n_g^{c*} - \mu_g^{c*}} = \frac{\mu_g^c}{n_g^c - \mu_g^c} = \frac{1}{r_g^c} - 1$.

In the above specification, the coefficients on male and female populations are restricted to be the same, α_{mf}, which means that doubling the population will increase total gains to marriage by $2\alpha_{mf}$ percent. Given that male and female populations across cities are highly collinear, if we allow for gender-specific population coefficients, multicollinearity results in unstable and implausible point estimates.

Like in (1), n_m^c and n_f^c in (3) are potentially endogenous. If the idiosyncratic gain to marriage in city c, v_{mf}^c, is large, individuals may want to move to city c to find a marriage partner and, therefore, n_m^c and n_f^c may also be large, leading to an upward bias in the OLS estimate of α_{mf}. On the other hand, sampling error in the number of eligibles, n_m^c and n_f^c, may lead to a negative correlation between the covariate and total gains to marriage. This implies that the OLS estimate of α_{mf} will be biased downward. As discussed in the case of log marriage odds ratio regressions, this sampling error is unlikely to be important. We will use OLS and also instrument for $(\ln n_m^c + \ln n_f^c)$ in (3).

Marital Output Regressions. A caveat to our empirical approach is necessary. Even if we find that the marriage rate in a city is independent of the number of eligibles in that city, a thick market externality may still exist. The reason is that agents may react to a thick market externality, with faster arrivals of potential partners and/or a more disperse match value distribution, by waiting for a better match rather than marrying earlier. In this case, the thick market externality generates higher marital output but not necessarily a higher marriage rate. Because women have a finite amount of time to bear children, they may choose to marry in a narrow age window even if their pool of potential partners is poor. If the pool of partners increases, women may continue to marry in that age window but now they may be able to find better matches. In Section 3.3 we will provide some evidence on this alternative hypothesis on how individuals may react to thick market externalities.

We now present the datasets and discuss the results.

3 United States, 2000

For the United States we use the 5 percent random sample of the 2000 census. We use the sample weights to calculate our population counts.

Since there is unrestricted mobility in the United States, the population in each city must be regarded as endogenous in the marriage rate or marriage gains regressions. To address this endogeneity problem, for a particular category of individuals in a city, we use the number of individuals in that group 20 years younger to instrument for the number of the same individuals in the year 2000. For example, we instrument the number of men in the 27–31 age group in a city in the year 2000 with the number of male children in the 7–11 age group in that city in the year 1980. Our assumption is that children (or their parents) do not choose the city in which to reside based on their marital prospects twenty years later.

There are 248 cities in the sample and we use the census definition of a standard metropolitan statistical area, SMSA, for a city. The smallest city in the sample has 105,000 individuals, whereas the largest city has 9,700,000 individuals. Across cities, the average population is 810,787, which is consistent with the well-known fact that most cities in the United States have less than 1 million individuals.

Table 1 displays summary statistics where each city is an observation.

[TABLE 1 HERE]

We consider men in the 27–31 age group and women in the 25–29 age group. The mean marriage rates for men and women are 0.485 and 0.508, respectively. As one can see, there is substantial variation in marriage rates across cities. One standard deviation of the marriage rates exceeds 0.06, which is more than 10% of the mean marriage rates across cities.

The mean total gains to marriage for men in the 27–31 age group married to women in the 25–29 age group is –0.567. One standard deviation is 0.278, which means that in most cities total gains to marriage are negative. A negative total gain to a specific marriage match is not unusual. It means that a marriage match that is imposed on a randomly chosen $\{m, f\}$ couple will generally be worse than having the couple remaining umarried. Type m and type f individuals who choose to marry each other are not randomly drawn from the $\{m, f\}$ population. These couples, compared with other m and f individuals who do not choose to marry each other, have high idiosyncratic payoffs from marrying each other.

3.1 City Size and Marriage Rates

Table 2 presents log marriage odds ratio weighted regressions at the city level with population weights. The dependent variable is the log of the number of married individuals divided by the number of unmarried individuals by gender and age. The independent variable is the log of the number of individuals in the same gender and age groups.

[TABLE 2 HERE]

Column (1) indicates that the coefficient for men in the 27–31 age group is –0.133 with a standard error of 0.021, which implies slight DRS for men in the marriage market.

Adding (in column 2) the proportions of black and white men in that age group, the proportion of men with college degree, the mean and the variance of income of men in the 27–31 age group as covariates, the point estimate increases to –0.028 with a standard error of 0.024. Now $\alpha_m = 0$ is in the confidence interval and we can no longer reject CRS.

The negative sign of the coefficient of the variance of family income of men indicates that in cities with more family income dispersion, which is consistent with Gould and Passerman's observation that marriage is delayed when there is more heterogeneity in the family incomes that these men are associated with.

To deal with the endogeneity problem, column (3) instruments the log number of men in the 27–31 age group with the log number of men 20 years younger, that is, with the number of 7–11 years old male children in 1980.[5] The point estimate becomes –0.048 with

[5] The first stage IV regression is presented in Table A1 in the Appendix.

a standard error of 0.028. The IV point estimate is slightly more negative than the OLS estimate.[6]

Both the OLS and IV estimates do not reject CRS. Given the standard errors, using a 95% confidence interval, the point estimate for α_m will not exceed 0.01, which means that even in the best case for IRS, the quantitative effect is very modest. From Table 1, the mean marriage rate across cities is 0.485. Using our discussion of order of magnitude in Section 2, a doubling of the population of a city will increase the marriage rate of men by less than half percent. Recall that there is substantial variation in marriage rates across cities (in Table 1 the standard deviation of marriage rates for men across cities is 0.063, which exceeds 10% of the mean marriage rate). However, the findings in Table 2 indicate that variation in city size alone cannot explain the variation in male marriage rates across cities.

Columns (4–6) run the same regressions for women in the 25–29 age group. In column (4) the point estimate is –0.166 with a standard error of 0.030 so we can reject the hypothesis of CRS at the 5% significance level. When we include (in column 5) the proportions of black and white women in that age group, the proportion of women with college degree, the mean and the variance of income of women in the 25–29 age group as covariates, the point estimate increases to –0.010 with a standard error of 0.026. Now we cannot reject the hypothesis of CRS at the 10% significance level.

Also for women, the coefficient of the variance of family income is negative, which is consistent with the delay argument set forth by Gould and Paserman (2003): when the pool of women's family incomes are more dispersed, men also delay their marriage to find a better match.

Column (6) instruments the log number of women in the 25–29 age group with the log number of women 20 years younger. The point estimate drops slightly to –0.021 with a standard error of 0.030. Here we cannot reject the hypothesis of DRS at the 5% significance level. However, the magnitude of the DRS is modest: the 95% confidence interval upper bound is 0.04. From Table 1, the mean marriage rate for women across cities is 0.508. A doubling of the population in a city will increase the marriage rate of women by less than half percent. As explained above, recall that there is substantial variation in marriage rates across cities (in Table 1 the standard deviation of marriage rates for women across cities is 0.069, which exceeds 10% of the mean marriage rate). The findings in Table 2 indicate that variation in city size alone cannot explain the variation in female marriage rates across cities. Both estimates of the returns to scale using male and female marriage odd ratios in Table 2 are consistent with each other.

Two caveats are in order. First, if in large cities individuals are more likely to divorce, using the number of *currently married* individuals may undercount the number of marriages in large cities and explain the absence of increasing returns to scale. To check for this possibility, we also investigate log marriage odds ratio regressions using *ever married* individuals instead of currently married individuals (see Table A2 in the Appendix). The evidence for DRS is even stronger. A possible interpretation is that in larger cities, there may be less

[6] As expected, the more negative IV estimate suggests that measurement error in the covariate is more important than having the number of individuals in both the right- and left-hand sides of the regression.

stigma from getting divorced and not remarrying.

Second, to check the issue of delay, that is, the possibility that in large cities with a large pool of eligibles, individuals may wait and delay marriage, we also run marriage odds ratio regressions for currently married in older age groups: men age 37–41 and women age 35–39. The results (Table A3 in the Appendix) are similar to the ones for the younger individuals: there is no evidence of IRS. There is some slight evidence of DRS and in most regressions, we cannot reject CRS. However, unlike for young women who seem to wait when there is more dispersion in the pool of eligibles (measured by the variance of incomes), older women do not delay their marriages when faced with a more dispersed pool of potential partners.

To sum up the findings from the marriage odds ratio regressions, we can conclude that, unlike Angrist (2004) and Choo and Siow (2006), these regressions tell a consistent story: CRS cannot be rejected as a first approximation in U.S. contemporary marriage markets.[7]

3.2 City Size and Total Gains to Marriage

In order to impose marriage market clearing on our estimation, Table 3 presents total gains to marriage (π_{mf}) weighted regressions at the city level for men in the 27–31 age group and for women in the 25–29 age group. The dependent variable is the log of the number of marriages among men and women in the two age groups minus the log of the geometric average of the number of unmarried individuals in the corresponding age groups. The independent variable of interest is $\frac{1}{2}(\ln n_m + \ln n_f)$, the log of the geometric mean of the number of men and women in the corresponding age groups.

[TABLE 3 HERE]

In Column (1) the point estimate for α_{mf} is -0.047 with a standard error of 0.010. We can reject $\alpha_{mf} = 0$, CRS, at the 5% significance level. When adding in column (2) the proportions of black and white individuals in that age group, the proportion of individuals with college degree, and the mean and variance of family income as covariates, the point estimate increases to 0.010 with a standard error of 0.011, and, therefore, we cannot reject $\alpha_{mf} = 0$, CRS. Columns (3) and (4) instrument $\frac{1}{2}(\ln n_m + \ln n_f)$ with the log of the number of individuals twenty years younger, that is, with the size of the cohort of men who were 7–11 years old and of women who were 5–9 years old in 1980. Without other covariates, the point estimate is -0.045 with a standard error of 0.011. So we can reject $\alpha_{mf} = 0$, CRS, at the 5% significance level. When adding the other covariates, the point estimate increases to 0.011 with a standard error of 0.013, which means we cannot reject CRS.

Although the IV point estimates are slightly larger than the OLS estimates, both sets of estimates tell the same story. Moreover, the standard errors in all the regressions are small and similar. The largest point estimate on $\frac{1}{2}(\ln n_m + \ln n_f)$ from column (4) is 0.011. Given a standard error of 0.013, the upper bound of the 95% confidence interval for α_{mf} is 0.037. Doubling the population will increase total gains by $\alpha_{mf} \ln 2 = 0.025$. From

[7] Note that Angrist (2004) and Choo and Siow (2006) ask different substantive questions unrelated to the issues discussed here.

Table 1, the standard deviation of total gains to marriage across cities is 0.277. Using the 95% confidence interval upper bound estimate and the discussion of orders of magnitude in Section 2, a doubling of the population will explain less than 10% of a standard deviation in total gains. Under the most favorable interpretation, there is a slight evidence in favor of IRS.

Although the size of the marriage market seems largely not to affect total gains to marriage, other factors do affect total gains to marriage in a city. In both columns (2) and (4), as can be anticipated from the increment in R^2, the P-values that all the other variables are different from zero are smaller than 0.001.

Again, to control for the issue of delay, we also estimate total gains to marriage regressions for older individuals (Table A4 in the Appendix): men in the 37–41 age group and women in the 35–39 age group. The findings are the same: we cannot reject CRS. We also estimate total gains to marriage regressions where each observation gets the same weight. The results are not qualitatively different from the weighted regressions.

3.3 City Size and Marital Match Quality

Tables 2 and 3 above indicate that when looking at marriage rates and total gains to marriage, CRS is a reasonable assumption for marriage markets in the United States.

However, despite these findings, a thick market externality may still exist in these marriage markets, in the sense that individuals in larger markets have faster arrival of potential matches and/or a more dispersed match value distribution to draw from. As is well known from the sequential search literature (e.g., Mortensen 1986), both of these effects may lead to delayed marriages in larger markets, thus lowering observed marriage rates and marriage gains. Thus our empirical evidence on marriage rates and marriage gains may still be consistent with IRS. Addressing this issue is similar to what has been done in the labor markets literature. For example, Petrongolo and Pissarides (2005) estimate a structural model of search and unemployment with scale effects. They show that the unemployment hazard is unaffected by scale effects. Scale effects lead to higher reservation wages and higher post-employment wages. This means that scale effects may show up in the quality of the match rather than in a higher matching rate.

We can shed some light on this issue by looking at two measures of marital match quality: (i) the proportion of children age 1–6 living with two parents, and (ii) the educational attainment of 16-year old children of married parents. We do not use the divorce rate or marital tenure to proxy for marital quality because one can argue that married individuals in larger cities may also choose to break up more and rematch more and they still have higher average welfare.

A major reason why individuals marry is to have children. We use the proportion of 1–6 years old children, who live with two parents in a city, as a measure of average marital match quality in that city. Figure 1 shows a plot of log(proportion of children age 1–6 who live with two parents) against log(population) in that city.

[FIGURE 1 HERE]

As can be seen in the scatter plot and the non-parametric regression line, there is little evidence that average match quality is increasing in city size. There is an increase for the largest cities but there are few observations there.

Table 4 presents population weighted regressions of the ln(proportion of children 1–6 who live with two parents) on ln(population) and covariates.

[TABLE 4 HERE]

The point estimate in column (1), -0.016, shows no relationship between the proportion of children age 1–6, who live with two parents, and log population in a city. When controlling in column (2) for the proportions of black and white men and women in the 27–31 and 25–29 age groups, respectively, and for the proportions of those individuals with college degrees, the estimated coefficient, -0.015, is still zero at the 5% significance level. Column (3) adds regional dummies. This specification does show a statistically significant negative estimated coefficient, -0.015, at the 5% significance level, which is consistent with the negative slope for the largest cities in Figure 1. Weighting the observations by population size gives increasing weight to larger cities, and therefore, this negative slope is driven by very few observations. When we rerun the regressions treating all observations equally, the estimated coefficients on log population are not significantly different from zero at the 5% significance level in all specifications. Thus, the findings in Table 4 indicate that delay in marriage does not seem to increase average marital match quality as proxied by the proportion of children age 1–6, who live with two parents. If anything, there is some evidence that the largest cities reduce marital match quality.

A possible objection to using the proportions of children age 1–6 living with two parents as a proxy for average marital match quality is another endogeneity problem: if married individuals leave a city (for example, when they have children, they decide to move to the suburbs), then cities may be left with unmarried individuals who are still hoping to marry. If there are IRS, larger cities with their higher *actual* marriage rates, may not have higher *observed* marriage rates if married individuals leave the cities at a higher rate than unmarried individuals. Since we use the metropolitan definition of a city (SMSA), which includes the nearby suburbs, this endogeneity problem of leaving the city after marriage should not be a first order problem for the regressions in Table 4.

The second proxy for the quality of marital match is the educational attainment of 16-year old children of married parents. Age 16 is chosen because most of these children are still living with their parents and a non-trivial proportion of them is behind their birth cohort in terms of attained years of schooling. This proxy does not suffer of the endogeneity problem described above: by looking at the educational attainment of 16-year old children, we are basically comparing the average marital match quality of couples in small cities, who did not leave the city after marriage, with the average marital match quality of couples in large cities, who did not leave the city after marriage. Table 5 presents the results of two regressions.

[TABLE 5 HERE]

In column (1), the number of years of education of 16-year old children with married parents in a city is regressed on the log population of that city, father's age, mother's age, father's years of schooling, mother's years of schooling, and dummy variables for father's and mother's races. The point estimate indicates that, controlling for parents' education, a 1% increase in the population size of a city decreases the years of schooling of a 16-year old by 0.004 years. However, the standard error, which is clustered by city, shows that the point estimate is not statistically different from zero at the 10% significance level.

Column (2) reports the results of a linear probability model where the dependent variable is equal to 1 if the 16 year-old child of a couple attained the median years of education of 16-year old children in his or her state, and 0 otherwise. The estimated coefficient on log population is 0.012 but again the standard error, which is clustered by city, shows that the point estimate is not statistically different from zero at the 10% significance level.

The two regressions lead to similar conclusions. Marital output, as proxied by the educational attainment of 16-year old children, does not rise with city size. Also, regardless of which proxy for mariatl match quality, this evidence on marital output is unlikely to be affected by the out-migration of married couples from the city.

In both regressions, the estimated coefficients on the interaction between the parents' years of schooling are positive and significantly different from zero at the 5% level. So there is evidence of gains in marital quality, as proxied by educational attainment of the children, by matching by education of the parents. It should be clear that gains from marital matching is separate from the question of whether the degree of returns to scale in the marriage market.

To sum up the findings for U.S. marriage markets in the year 2000, we conclude that there is no evidence of IRS, whereas the hypothesis of CRS cannot be rejected. A thick market externality is rejected when one looks at marriage rates and totals gains to marriage (Tables 2 and 3), as well as at marital match quality (Tables 4 and 5).

Given the potential endogeneity of marriage market size in the United States, the credibility of our findings is predicated upon the exogeneity of our instruments for population vectors. While we have made apriori arguments for the exogeneity of these instrumental variables, we cannot externally validate the exogeneity assumption.

We now turn our attention to two societies in which the potential endogeneity problem that characterizes U.S. marriage markets does not exist. Both in early Renaissance Tuscany and in China in 1982 individuals did not move to cities to search for marriage partners. Therefore, we can assume that in these two societies the size of the marriage market is exogenous.

4 Tuscany, 1427

Late medieval and early Renaissance Tuscany was one of the most urban and commercial economies in Europe, with Florence being one of the most important trade and banking centers. The other Tuscan towns under Florentine rule (e.g., Pisa, Prato, Arezzo, Pistoia, San Gimignano) were also commercial economies, though on a smaller scale. The rest of Tuscany consisted of hundreds of rural villages, with a mix of farmers, who owned and worked on their farms, and sharecroppers and fixed-rent tenants, who worked on the farms owned by town- and village-dwellers.

Two features of Tuscan marriage markets are relevant for our study. First, early Renaissance Tuscany was a virilocal society in which daughters left their natal households upon marriage and moved into their in-laws' households, whereas most married sons continued to live with, and work for, their parents (Botticini 1999; and Botticini and Siow 2003). Second, Tuscan people rarely moved to other towns or villages for the purpose of marriage as shown in Table 6, which illustrates the place of residence of the bride's and groom's families for a sample of more than 7000 dowry and marriage contracts we collected at the State Archives of Florence.

[TABLE 6 HERE]

From the thirteenth to the fifteenth century, the vast majority of urban grooms married urban brides, and vice versa, rural grooms mainly married rural brides. There were very few instances of an urban groom marrying a rural bride, or vice versa. This implies that the size of the marriage market in Tuscan cities and villages can be taken as exogenous.

The data on marriage matches come from the machine readable data file *Census and Property Survey of Florentine Dominions in the Province of Tuscany, 1427-1480*, prepared by David Herlihy and Christiane Klapisch-Zuber (1978).[8] Faced with a fiscal crisis because of the protracted warfare against the duchy of Milan, in 1427 the Priors of the Florentine Republic instituted a new tax survey for all citizens of the city of Florence, the Tuscan towns, and hundreds of villages in her territories. Government officials and their staffs interviewed every head of household in Tuscany, for a total of more than 60,000 households (about 260,000 people). The survey, completed within a few months, included information on all real property (houses in the city, farms and land holdings), business investments, loans, shares of the public debt, debts, occupations, and household demographics (number, gender, and age of children and other household members).

The machine readable data file divides the Tuscan population into 29 distinct series (Table A5 in the Appendix). For the purpose of our study, each series is labeled as one distinct "district, " where each district corresponds to either a city (town) or to the rural area surrounding a city (town). There was clearly a large variance in the size of the population, and hence, of marriage markets, in each of the 29 districts. For example, district no. 1 corresponds to the city of Florence, which had 9780 households, whereas district no. 18 consists of the villages in the Garfagnana rural area, which hosted a tiny population of only 175 households.

[8] The tax survey on which the Catasto data are based, and the documentary sources for the Catasto, are fully described in David Herlihy and Christiane Klapisch-Zuber (1978), *Les Toscans et leurs familles: Un étude du catasto Florentin de 1427* (English abridged edition: *Tuscans and their Families: A Study of the Florentine Catasto of 1427.* New Haven: Yale University Press, 1985).

4.1 City Size and Marriage Rates

Figure 2 describes the marriage behavior of individuals surveyed in the 1427 census. Women married in a very narrow age window, 15–25, whereas men married much later.

[FIGURE 2 HERE]

We consider men in the 22–32 and women in the 18–27 age groups, respectively. Table 7 indicates that women's marriage rate was very high: 73% of 18–27 years old women were married versus 52% of 22–32 years old men.

[TABLE 7 HERE]

Because of the lack of mobility for marital reasons in early Renaissance Tuscany, reverse causality is not a concern and, therefore, we can investigate the returns to scale hypothesis using OLS.

Table 8 presents log marriage odds ratio population weighted regressions at the district level. The dependent variable is the log of the *number of married individuals* divided by the *number of unmarried individuals* by gender and age. The independent variable is the log of the *number of individuals* in the same gender and age group. As these are log linear regressions, the point estimates are interpretable as elasticities.

[TABLE 8 HERE]

Column (1) indicates that a 1% increase in the number of women in the 18–27 age group results in a 0.038% decrease in the marital odds ratio for women in that age group. Since the standard error is 0.107, we cannot reject the null hypothesis of CRS ($\alpha_f = 0$) at any reasonable significance level.

Column (2) includes the average log assets per adult in the district, that is, the average wealth per adult in the district. Again, the size of the marriage market in a given age group does not affect the probability of being married in that age group. Also wealth per adult does not affect the odds of being married.

In columns (3) and (4), we instrument the log of the number of women in the 18–27 age group with the log total population of the district. Again, there is no effect of the size of marriage market on the number of marriages in that age group.

So for women in the 18–27 age groups, we cannot reject the hypothesis of CRS ($\alpha_f = 0$) in marriage markets at the district level. There is also no discernible effect of average wealth in a given district on the marriage rate of women.

Columns (5)-(8) present the results for 22–32 years old men. The OLS point estimate for α_m in column (5) is -0.035 with a standard error of 0.153. We cannot reject the hypothesis that $\alpha_m = 0$ (CRS) at the 10% significance level. When adding in column (6) the average log assets per adult in the district for that male age group, the OLS point estimate for α_m is

-0.037 with a standard error of 0.056. Again, we cannot reject the hypothesis that $\alpha_m = 0$ (CRS) at the 10% significance level. The point estimate for average log assets per adult in the district is -0.353 with a standard error of 0.089. Thus, unlike women whose marriage rates did not seem to be affected by the average wealth in the district, 22–32 years old men in richer districts were more likely to remain unmarried.

In columns (7) and (8), we instrument the log of the number of men in the 22–32 age group with the log total population of the district. Again, there is no effect of the size of the marriage market on the number of marriages of that male age group. The point estimate for average log assets per adult in the district is -0.353 with a standard error of 0.090. Consistent with the OLS results, 22–32 years old men in richer districts were more likely to be unmarried. In general, the OLS and corresponding IV point estimates and standard errors are very similar. Thus, there is no evidence of endogenous mobility in a district for marital reasons, which is consistent with our prior.

We also run unweighted regressions and the results are quantitatively similar to those in Table 8. Thus when looking at both OLS and IV, weighted or unweighted, male or female marriage odds ratio specifications, we cannot reject the hypothesis that $\alpha_m = \alpha_f = 0$, that is, the hypothesis of CRS. However, because of the small number of districts (observations) and the relatively large standard errors, it is premature to reject modest IRS. For example, we often cannot also reject the hypothesis that $\alpha_m = \alpha_f = 0.1$. The marriage rates for 22-32 year old men and 18-27 year old women were 0.518 and 0.838 respectively. Assuming the upper bound estimates $\alpha_m = \alpha_f = 0.1$ and using the orders of magnitude calculations in Section 2, a doubling of the population will increase the marriage rate by less than 4% and 2% for men and women respectively.

4.2 City Size and Total Gains to Marriage

Table 9 presents unweighted total gains to marriage regressions at the district level for men in the 22–32 age group and for women in the 18–27 age group. The dependent variable is the log of the number of marriages among men and women in the two age groups divided by the geometric mean of the number of unmarried men and women in the two age groups. The independent variable is the log of the geometric mean of the numbers of individuals in the two age groups .

[TABLE 9 HERE]

Column (1) indicates that the OLS point estimate of α_{mf} is -0.012 with a standard error of 0.131. Column (2) includes the average log assets per adult in the district. The OLS point estimate for α_{mf} is -0.027 with a standard error of 0.092. Neither estimate can reject the hypothesis that $\alpha_{mf} = 0$ (CRS) at standard confidence intervals. The point estimate for log assets per adult in the district is -0.206 with a standard error of 0.149. So there is some evidence that wealthier districts lower total gains to marriage between 18–27 years old women and 22–32 years old men.

In columns (3) and (4), we instrument the log of the geometric mean of the numbers of individuals in the two age groups with the log population in the district. In column (3), the

IV estimate is 0.021 with a standard error of 0.132. When including the average log assets per adult in the district, the IV point estimate in column (4) is 0.009 with a standard error of 0.100. Also, the IV point estimate for average log assets shows that wealthier districts have a negative impact on total gains to marriage between 18–27 years old women and 22–32 men years old men suggesting that wealthier individuals delay marriage.

We also run unweighted total gains to marriage regressions and obtain similar results. Whatever specification we run, we cannot reject the hypothesis that $\alpha_{mf} = 0$, that is, the hypothesis of CRS. However, because of the small number of observations and the relatively large standard errors, it is premature to reject modest IRS. For example on a few occasions, we cannot also reject the hypothesis that $\alpha_{mf} = 0.2$ at the 5% confidence interval. If $\alpha_{mf} = 0.2$, a doubling of the population will increase total gains to marriage by $0.2 * \ln 2 = 0.139$, which is 27% of the standard deviation of total gains to marriage across districts. These estimates are the largest returns to scale estimates in this paper.

5 China, 1982

The data comes from the 1982 census, which is a 1 percent random sample of the population. We restrict our analysis to people living in the 245 cities identified in the census.

Until 1978, internal migration in China was strictly restricted and residents of cities had to live in the cities where they were born. After 1978, at the same time when economic reforms in the countryside started being implemented, migration restrictions from rural to small– and medium–size cities began to be relaxed. Even though the pace of social and economic changes has been enormous since 1978, internal migration is still restricted in China. At the time of the 1982 census, most residents had no choice about where they could live, and even if individuals wanted to migrate across cities for marital purpose, they could not do so (Goldstein and Goldstein 1990).

Only under two circumstances internal migration was relatively free. First, individuals who got into institutions of higher education were allowed to move there. In 1982, though, few Chinese people had higher education. Second, individuals could move to another city if a firm was willing to hire them. However, because there was no free market for labor, very few people moved for this reason.

This restricted and limited internal migration makes the endogeneity problem not a issue when we use the 1982 Chinese census to study the effect of city size on marriage rates. Au and Henderson (forthcoming) also exploited labor immobility across and into Chinese cities to study other agglomeration effects of cities.

Since there was little divorce in China in that time period, the marriage rate in the 1982 census refers to permanent marriages.

5.1 City Size and Marriage Rates

Table 10 shows summary statistics for the Chinese data. There were 244 (OR 245???????) cities. We study the marital status of men in the 24–28 and 29–33 age groups, and women in the 21–25 and 26–30 age groups. Each age group has over 70,000 observations.

[TABLE 10 HERE]

By age 30, marriage was essentially universal for women: 46 percent of women in the 21–25 age group was married, whereas 97 percent of women in the 26–30 age group was married. For young and old men, the percentages of married were 58 and 90 percent, respectively. Ninety six percent of men and ninety two percent of women in the two age cohorts had completed primary school.

Table 11 presents marriage odds ratio OLS regressions at the city level. Each observation is weighted by the population of the city. The dependent variable is, by gender and age, the log of the *number of married individuals* divided by the *number of unmarried individuals*. The independent variable is the log of the *number of individuals* in the same gender and age group.

[TABLE 11 HERE]

Column (1) indicates that a 1% increase in the number of women in the 21–25 age group results in a 0.35% decrease in the number of marriages in that age group. This estimate is statistically significant at the 5% significance level.

Columns (2) adds the log odds ratio of women in the 21–25 age group, who have completed primary education in a given city. The point estimate implies a population elasticity of -0.17. It continues to be statistically significant at the 5 percent level. The estimated coefficient on log odds ratio of education is negative, consistent with the hypothesis that educated women delay marriage. The negative estimated population elasticities suggests that there may be some delay in marriage in larger cities.

Column (3) runs the same log marriage odds ratio regression for women in the 31–35 age group. The estimated population elasticity is -0.24, which is statistically significant at the 10% level. Column (4) adds the log odds of primary education. The estimated population elasticity becomes -0.02 and not statistically different from zero. Although the standard error is reasonably small, here we cannot reject the hypothesis that $\alpha_f = 0.1$ at the 95% confidence interval. If $\alpha_f = 0.1$, a doubling of the population will increase the marriage rate by 2%. The point estimate for α_f suggests a much smaller increase, if any. The results in columns (3) and (4) do not suggest any strong departure from CRS in long run female marriage rates.

Columns (5)–(8) run similar regressions for men in the 24–28 age group. Column (5) indicates that a 1% increase in the number of men in the 24–28 age group results in a 0.23% decrease in the number of marriages in that age group. This estimate is statistically significant at the 10% level.

Columns (6) adds the log odds of men in the same age group, who have completed primary education in a given city. The point estimate implies a population elasticity of -0.17, which is almost the same as the women's point estimate in column (2), and it is statistically significant at the 5 percent level. The estimated coefficient on log odds of education is negative, albeit not statistically different from zero. The negative estimated population elasticities suggests that there may be some delay in marriage in larger cities.

Column (7) runs the same log marriage odds ratio regression for men in the 31–35 age group. The estimated population elasticity is -0.10, which is statistically not different from zero. Column (8) adds the log odds ratio of primary education. The estimated population elasticity becomes -0.18 and statistically different from zero at the 5% level. The negative estimated education elasticity is again not different from zero suggesting that the long run marriage rates of men did not differ by education. The results in columns (7) and (8) do not suggest any strong departure from CRS in long run male marriage rates.

[TABLE 12 HERE]

Table 12 presents population weighted IV regression results. The instrument for log number of individuals by gender and age is the log population of the city. The columns correspond to the similarly numbered columns in Table 11. By comparing the results for the same numbered columns in Tables 11 and 12, one can see that the results are essentially the same between the two tables. Thus, instrumenting the number of individuals by gender and age does not change the conclusion that there is minimal evidence of IRS in both short and long run marriage rates in China. A first approximation is that Chinese marriage markets are characterized by CRS.[9]

5.2 City Size and Total Gains to Marriage

Table 13 presents population weighted total gains to marriage regressions.

In column (1), the dependent variable is the total gains to marriage between 21–25 years old women and 24–28 years old men for each city. The independent variable is the log of the geometric mean of the populations of those women and men for each city. The OLS estimated population elasticity is -0.270 and it is statistically significant at the 5% level.

Column (2) adds the log odds ratio of educated men and women as additional regressors. The OLS estimated population elasticity is -0.143 and it is statistically significant at the 5% level.

Columns (3) and (4) estimate the total gains for women in the 31–35 age group and men in the 34–38 age group. In column (3), the estimated population elasticity is -0.198 and it is statistically significant. Adding educational regressors in column (4) changes the estimated population elasticity to -0.123. It is now only significant at the 5% level. Thus, the total gains estimates suggest that the long run total gains to marriage in China in 1982 is best approximated by CRS.

[TABLE 13 HERE]

Columns (5) to (8) re-estimate columns (1) to (4) using log of the total population as an instrument for $\frac{1}{2}\ln n_f n_m$. There is little difference between the OLS and the IV estimates for the younger age match (compare columns [5] and [6] with [1] and [2], respectively). Here the evidence suggests that larger cities have lower total gains for these young marriages.

[9] We have also run the equivalent unweighted regressions for Tables 11 and 12 and the results largely do not change.

The estimates of α_{mf} in columns (7) and (8) are different from those in columns (3) and (4). Although statistically not different from zero, the point estimates in (7) and (8) are positive, unlike the estimated coefficients for younger marriages. At the 95% confidence interval, the upper bound estimate for α_{mf} is 0.14 from column (7). This estimate implies that a doubling of the population will increase total gains to marriage by 0.097, which is less than 20% of the standard deviation of the total gains to marriage across cities. So while we cannot reject the hypothesis of CRS for the older matches, we can also not reject mild IRS.

6 Concluding Remarks

The empirical results on marriage rates and marriage gains from our three data sets and different empirical methods are consistent with each other. Given the widely different circumstances of the three societies in terms of geography, time periods, social norms, and population size differences, the similar estimates for returns to scale parameters in all three data sets are remarkable.

CRS is a reasonable approximation for marriage markets. Using a 95% confidence interval upper bound estimate for α_{mf}, there is mild evidence for IRS in some specifications. Even in these most favorable cases, the quantitative effects are small. A doubling of the population cannot increase total gains by more than 30% of a standard deviation of total gains across observations. Also, there is more evidence in favor of DRS in Chinese marriage markets.

In order to rule out IRS, which does not show up in marriage rates or marriage gains, we also investigated the relationship between a proxy for average marital quality, the proportion of children in the 1–6 age group who live with two parents, and population in a city. There is little evidence of a such positive relationship in 2000 US census. We also investigated the relationship between city size and educational attainment of 16 year olds living with two married parents. The evidence suggest that, controlling for parental characteristics, education attainment fell with city size.

These findings suggest that economists should focus on models of the marriage market that deliver CRS. One rationale for the CRS finding is that, given their individual characteristics, marriage within a narrow age window is a compelling experience for most individuals. This customary narrow age window of marriage adjusts due to income equality (as also found by Gould and Passerman), educational attainment and other factors. Exogenous population variation is not one of these factors.

TABLE 1—Summary Statistics, United States 2000

	Mean	Standard deviation	Observations (no. of cities)
City population	810787.1	1278090	248
	Men (age 27–31), women (age 25–29)		
Male marriage rate	0.484	0.063	248
Female marriage rate	0.507	0.069	248
Log male marriage odds ratio	–0.061	0.258	248
Log female marriage odds ratio	0.032	0.283	248
Total gains to marriage	–0.567	0.277	248
	Men (age 27–31)		
Married men	13501.24	20702.69	248
Total number of men	29899.75	50690.39	248
Fraction of white non-Hispanic men	0.695	0.169	248
Fraction of black non-Hispanic men	0.110	0.094	247
Fraction of Hispanic men	0.137	0.156	245
Fraction of men with college education	0.250	0.092	248
Mean family income of married men	10703	0.175	248
Variance, family income of married men	0.923	0.556	248
	Women (age 25–29)		
Married women	13272.79	20430.38	248
Total number of women	28417.72	48995.97	248
Fraction of white non-Hispanic women	0.686	0.176	248
Fraction of black non-Hispanic women	0.125	0.113	241
Fraction of Hispanic women	0.130	0.157	245
Fraction of women with college education	0.282	0.097	248
Mean family income of married women	10682	0.168	248
Variance, family income of married women	0.971	0.602	248

Source: U.S. 2000 Census (5% random sample of the population).

TABLE 2—Marriage Odds Ratio Regressions, United States 2000

	Dependent Variable = $\ln \frac{\mu}{n-\mu}$					
	Men 27–31	Men 27–31	Men 27–31	Women 25–29	Women 25–29	Women 25–29
	OLS	OLS	IV	OLS	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
ln(men 27–31)	–0.133	–0.028	–0.048			
	(0.021)	(0.024)	(0.028)			
$\frac{\text{White men (27–31)}}{\text{Men (27–31)}}$		0.511	0.472			
		(0.189)	(0.198)			
$\frac{\text{Black men (27–31)}}{\text{Men (27–31)}}$		0.079	0.100			
		(0.216)	(0.220)			
$\frac{\text{Men (27–31), college}}{\text{Men (27–31)}}$		–2.071	–1.990			
		(0.338)	(0.359)			
Mean family income		–0.127	–0.113			
		(0.114)	(0.112)			
Variance family income		–0.062	–0.062			
		(0.027)	(0.027)			
ln(women 25–29)				–0.166	0.010	–0.021
				(0.030)	(0.026)	(0.030)
$\frac{\text{White women (25–29)}}{\text{Women (25–29)}}$					0.671	0.618
					(0.129)	(0.139)
$\frac{\text{Black women (25–29)}}{\text{Women (25–29)}}$					–0.514	–0.481
					(0.206)	(0.209)
$\frac{\text{Women (25–29), college}}{\text{Women (25–29)}}$					–2.166	–2.129
					(0.298)	(0.310)
Mean family income					–0.622	–0.533
					(0.171)	(0.173)
Variance family income					–0.123	–0.111
					(0.032)	(0.031)
Instruments			$\ln(\text{men}^{1980}_{7-11})$			$\ln(\text{women}^{1980}_{5-9})$
R-squared	0.19	0.53	0.53	0.53	0.66	0.65
Number of cities	248	247	247	248	241	241

Note: Regressions weighted by city population with robust standard errors in parentheses.
μ = number of currently married individuals. n = total number of inviduals.
Men 27–31 = number of men age 27–31. Women 25–29 = number of women age 25–29.
Men (27–31), college = number of men age 27–31 with college degree. Women (25–29), college = number of women age 25–29 with college degree.
Men^{1980}_{7-11} = number of men who were 7–11 years old in 1980. $\text{Women}^{1980}_{5-9}$ = number of women who were 5–9 in 1980.

TABLE 3–TOTAL GAINS TO MARRIAGE REGRESSIONS, UNITED STATES 2000

	Dependent Variable = gains to marriage $\pi_{mf} = \ln \mu_{mf} - \frac{1}{2}(\ln \mu_{m0} + \ln \mu_{f0})$			
	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)
$\frac{1}{2}(\ln\ n_m + \ln\ n_f)$	–0.047	0.010	–0.045	0.011
	(0.010)	(0.011)	(0.011)	(0.013)
$\frac{\text{White men (27–31)}}{\text{Men (27–31)}}$		0.463		0.482
		(1.027)		(1.029)
$\frac{\text{Black men (27–31)}}{\text{Men (27–31)}}$		0.101		0.125
		(1.289)		(1.305)
$\frac{\text{Men (27–31), college}}{\text{Men (27–31)}}$		0.974		0.976
		(1.061)		(1.057)
Mean income, married men 27–31		–0.467		–0.469
		(0.522)		(0.519)
Variance income, married men 27–31		–0.036		–0.036
		(0.091)		(0.091)
$\frac{\text{White women (25–29)}}{\text{Women (25–29)}}$		0.240		0.227
		(0.872)		(0.871)
$\frac{\text{Black women (25–29)}}{\text{Women (25–29)}}$		–0.433		–0.454
		(1.092)		(1.106)
$\frac{\text{Women (25–29), college}}{\text{Women (25–29)}}$		–1.551		–1.554
		(0.928)		(0.924)
Mean income, married women 25–29		–0.061		–0.064
		(0.544)		(0.546)
Variance income, married women 25–29		–0.145		–0.145
		(0.101)		(0.101)
Instruments			$\ln(\text{men}_{7-11}^{1980})$, $\ln(\text{women}_{5-9}^{1980})$	$\ln(\text{men}_{7-11}^{1980})$, $\ln(\text{women}_{5-9}^{1980})$
R-squared	0.126	0.539	0.126	0.539
Number of cities	248	241	248	241

Source: U.S. 2000 Census (5% random sample of the population).
Note: regressions weighted by city population with robust standard errors in parentheses.
μ_{mf} = number of marriages between men age 27–31 and women age 25–29.
μ_{m0} = number of unmarried men in the 27–31 age group.
μ_{f0} = number of unmarried women in the 25–29 age group.
n_m = total number of men in the 27–31 age group.
n_f = total number of women in the 25–29 age group.
Men_{7-11}^{1980} = number of men who were 7–11 years old in 1980. $\text{Women}_{5-9}^{1980}$ = number of women who were 5–9 in 1980.

TABLE 4–MARITAL OUTPUT REGRESSIONS, UNITED STATES 2000

	DEPENDENT VARIABLE = $\ln\left(\frac{\text{children (age 1-6) living with two parents}}{\text{children (age 1-6)}}\right)$		
	(1)	(2)	(3)
ln(population)	–0.016	–0.015	–0.015
	(0.010)	(0.009)	(0.006)
Race and education controls	No	Yes	Yes
Region dummies	No	No	Yes
R-squared	0.04	0.45	0.55
Number of cities	248	241	241

Source: U.S. 2000 Census (5% random sample of the population).
Note: robust standard errors in parentheses.
Children (1–6) = number of children in the 1–6 age group.

TABLE 5—Marital Output Regressions, United States 2000

	Dependent Variable	
	Years of education of 16-year old children of married couples	Dummy = 1 if 16-year old children attained median education
	(1)	(2)
ln(population)	–0.004	0.012
	(0.014)	(0.016)
Boy dummy	–0.144	–0.053
	(0.011)	(0.005)
Father's age	0.002	0.000
	(0.002)	(0.001)
Mother's age	0.010	0.002
	(0.002)	(0.001)
Father's education	0.018	0.006
	(0.007)	(0.003)
Mother's education	0.021	0.006
	(0.008)	(0.002)
F x M education	–0.119	–0.042
	(0.047)	(0.020)
Race dummies	yes	yes
	(0.030)	(0.012)
R-squared	0.01	0.01
Observations	81772	82021

Source: U.S. 2000 Census (5% random sample of the population).

Note: standard errors clustered by city in parentheses.

Median education refers to the median years of education of 16-year old children in the state where the city is located.

ln(population) is the log of population size in each city.

Father's education = father's years of schooling.

Mother's education = mother's years of schooling.

F x M education = $\frac{\text{father's years of schooling x mother's years of schooling}}{100}$.

Race dummies are eight dummy variables that control for the race of the married couple. The categories are (father white, mother black), (father white, mother Hispanic), (father black, mother white), (father black, mother black), (father black, mother Hispanic), (father Hispanic, mother white), (father Hispanic, mother black), (father Hispanic, mother Hispanic). The left out category is (father white, mother white).

TABLE 6

MARRIAGES BY TYPE OF MATCH, MEDIEVAL TUSCANY

	Years		
	1260–1299	1340–1360	1420–1435
Urban groom – urban bride	17.9	31.7	44.7
Rural groom – rural bride	74.1	61.0	49.7
Urban groom – rural bride	5.7	2.7	3.6
Rural groom – urban bride	2.3	4.7	0.1
Number of marriage contracts	475	2955	3721

Source: State Archives of Florence, Notarile Antecosimiano, 794 volumes of notarial deeds (see Botticini 2006 for details).
Note: The numbers are the percentages of marriages by type of match.

TABLE 7

MARRIAGE PATTERNS BY GENDER, TUSCANY 1427

	Mean	Standard deviation	Number of observations
Male marriage rate (men age 22–32)	0.518	—	22179
Male marriage rate (men age 30–39)	0.749	—	15427
Female marriage rate (age 18–27)	0.838	—	18961
Female marriage rate (age 25–34)	0.926	—	16135
Assets per household[a, b]	263.9	1534.0	61328
Assets per adult individual[c]	119.1	627.5	58225

Source: Herlihy and Klapisch-Zuber, 1427 Florentine Catasto machine readable data file.
[a] Assets per household refer to the household's total wealth (in gold florins) as given in the 1427 Florentine census. This was equal to the present discounted value (at 7 percent interest rate) of the income from houses, land holdings, commercial partnerships, plus loans to private individuals and shares of the public debt, minus household's total debt.
[b] For men, household wealth refers to either their own wealth (if their parents were no longer alive), or to the wealth of the natal household in which they kept living even after marriage. In contrast, as brides moved into their grooms' households upon marriage, in the 1427 census their wealth is identified with the wealth of their husbands' households. Given the high degree of positive assortative matching in marriage, using the husband's wealth instead of the natal household's wealth for married women is not a major concern.
[c] Assets per adult individual $= \frac{\text{household's total wealth}}{\text{number of adult individuals in the household}}$ where an adult individual is one who is 19 or older.

TABLE 8

MARRIAGE ODDS RATIO REGRESSIONS, TUSCANY 1427

	DEPENDENT VARIABLE = $\ln \frac{\mu}{n-\mu}$			
	Men 22–32	Men 22–32	Men 22–32	Men 22–32
	(1) OLS	(2) OLS	(3) IV	(4) IV
ln(men 22–32)	–0.035	–0.037	0.007	–0.015
	(0.153)	(0.056)	(0.141)	(0.061)
ln(assets)		–0.353		–0.353
		(0.089)		(0.090)
Instrument			ln(population)	ln(population)
R-squared	0.00	0.55	—	—
Number of observations	29	29	29	29
	Women 18–27	Women 18–27	Women 18–27	Women 18–27
	OLS (5)	OLS (6)	IV (7)	IV (8)
ln(women 18–27)	–0.038	–0.023	0.024	0.030
	(0.107)	(0.120)	(0.118)	(0.132)
ln(assets)		0.102		0.102
		(0.183)		(0.189)
Instrument			ln(population)	ln(population)
R-squared	0.00	0.02	—	—
Number of districts	29	29	29	29

Source: Herlihy and Klapisch-Zuber, 1427 Florentine Catasto machine readable data file.
Note: regressions weighted by district population with robust standard errors in parentheses.
μ = number of currently married individuals. n = total number of inviduals.
Men 22–32 = number of men in the 22–32 age group.
Women 18–27 = number of women in the 18–27 age group.
ln (assets) = log (household's total wealth). See Table 7's footnote for the definition of total wealth.
ln (assets per adult) = log $\left(\frac{\text{household's total wealth}}{\text{number of adult individuals in the household}}\right)$.

TABLE 9

Total Gains to Marriage Regressions, Tuscany 1427

	Dependent Variable = $\pi_{mf} = \ln \mu_{mf} - \frac{1}{2}(\ln\ \mu_{m0} + \ln\ \mu_{f0})$			
	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)
$\frac{1}{2}(\ln\ n_m + \ln\ n_f)$	–0.012	–0.027	0.021	0.009
	(0.131)	(0.092)	(0.132)	(0.100)
ln (assets per adult)		–0.206		–0.205
		(0.149)		(0.153)
Instrument			ln (pop)	ln (pop)
R-squared	0.00	0.13		
Number of districts	29	29	29	29

Source: see Herlihy and Klapisch-Zuber, 1427 Florentine Catasto machine readable data file.
Note: regressions weighted by district population with robust standard errors in parentheses.
μ_{mf} = number of marriages between men in the 22–32 age group and women in the 18–27 age group.
μ_{m0} = number of unmarried men in the 22–32 age group.
μ_{f0} = number of unmarried women in the 18–27 age group.
n_m = total number of men in the 22–32 age group.
n_f = total number of women in the 18–27 age group.
ln (pop) = log (total population in a given district).
ln (assets per adult) = log $\left(\frac{\text{household's total wealth}}{\text{number of adult individuals in the household}}\right)$.

TABLE 10

Marriage Patterns by Gender, China 1982

	Mean	Number of observations
Proportion of married men (age 24–28)	0.582	87696
Proportion of married men (age 29–33)	0.903	66938
Proportion of married women (age 21–25)	0.456	72143
Proportion of married women (age 26–30)	0.965	75382
Proportion of men who completed primary school	0.968	154634
Proportion of women who completed primary school	0.917	147525
Number of cities		245

Source: 1982 Census of China (1% random sample of the population).

TABLE 11

MARRIAGE ODDS RATIO REGRESSIONS (OLS), CHINA 1982

	DEPENDENT VARIABLE = $\ln \frac{\mu}{n-\mu}$			
	Men 24–28	Men 24–28	Men 34–38	Men 34–38
	(1)	(2)	(3)	(4)
ln (men 24–28)	–0.226	–0.165		
	(0.092)	(0.057)		
ln (men 34–38)			–0.103	–0.181
			(0.060)	(0.059)
P_m > primary education		–0.115		0.205
		(0.106)		(0.112)
R-squared	0.160	0.184	0.028	0.087
Number of cities	243	243	217	217
	Women 21–25	Women 21–25	Women 31–35	Women 31–35
	(5)	(6)	(7)	(8)
ln (women 21–25)	–0.351	–0.170		
	(0.067)	(0.041)		
ln (women 31–35)			–0.243	–0.020
			(0.099)	(0.066)
P_f > primary education		–0.320		–0.474
		(0.046)		(0.065)
R-squared	0.378	0.600	0.108	0.423
Number of cities	243	240	203	203

Source: 1982 Census of China (1% random sample of the population).
Note: regressions weighted by city population with robust standard errors in parentheses.
ln (men 24–28) = log (number of men in the 24–28 age group).
ln (men 34–38) = log (number of men in the 34–38 age group).
ln (women 21–25) = log (number of women in the 21–25 age group).
ln (women 31–35) = log (number of women in the 31–35 age group).
P_m = log odds ratio of men age 24–28 (in column 2) and age 34–38 (in column 4), who have completed primary education in a given city.
P_f = log odds ratio of women age 21–25 (in column 6) and age 31–35 (in column 8), who have completed primary education in a given city.

TABLE 12
MARRIAGE ODDS RATIO REGRESSIONS (IV), CHINA 1982

	Dependent Variable = $\ln \frac{\mu}{n-\mu}$			
	Men 24–28	Men 24–28	Men 34–38	Men 34–38
	(1)	(2)	(3)	(4)
ln (men 24–28)	–0.224	–0.166		
	(0.091)	(0.060)		
ln (men 34–38)			–0.100	–0.191
			(0.061)	(0.056)
P_m > primary education		–0.114		0.210
		(0.106)		(0.112)
Instrument	ln pop	ln pop	ln pop	ln pop
R-squared	0.160	0.184	0.028	0.087
Number of cities	243	243	217	217
	Women 21–25	Women 21–25	Women 31–35	Women 31–35
	(5)	(6)	(7)	(8)
ln (women 21–25)	–0.314	–0.145		
	(0.072)	(0.042)		
ln (women 31–35)			–0.266	–0.037
			(0.091)	(0.061)
P_f > primary education		–0.333		–0.467
		(0.046)		(0.063)
Instrument	ln pop	ln pop	ln pop	ln pop
R-squared	0.373	0.599	0.107	0.423
Number of cities	243	240	203	203

Source: 1982 Census of China (1% random sample of the population).
Note: regressions weighted by city population with robust standard errors in parentheses.
μ = number of currently married individuals. N = total number of inviduals.
ln (men 24–28) = ln (number of men in the 24–28 age group).
ln (men 34–38) = ln (number of men in the 34–38 age group).
ln (women 21–25) = ln (number of women in the 21–25 age group).
ln (women 31–35) = ln (number of women in the 31–35 age group).
P_m = log odds ratio of men age 24–28 (in column 2) and age 34–38 (in column 4), who have completed primary education in a given city.
P_f = log odds ratio of women age 21–25 (in column 6) and age 31–35 (in column 8), who have completed primary education in a given city.
ln (pop) = ln (total population in a given city).

TABLE 13

Total Gains to Marriage Regressions, China 1982

	Dependent Variable = $\pi_{mf} = \ln \mu_{mf} - \frac{1}{2}(\ln \ \mu_{m0} + \ln \ \mu_{f0})$			
	Men 24–28 Women 21–25	Men 24–28 Women 21–25	Men 24–28 Women 21–25	Men 24–28 Women 21–25
	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)
$\frac{1}{2}\left(\ln \ n_m^{24-28} + \ln \ n_f^{21-25}\right)$	–0.270	–0.143	–0.139	–0.127
	(0.080)	(0.045)	(0.033)	(0.032)
P_m^{24-28} > primary education		–0.182		–0.017
		(0.109)		(0.072)
P_f^{21-25} > primary education		–0.071		–0.098
		(0.079)		(0.061)
Instrument			ln pop	ln pop
R-squared	0.300	0.429	0.104	0.139
Number of cities	243	240	243	240
	Men 34–38 Women 31–35	Men 34–38 Women 31–35	Men 34–38 Women 31–35	Men 34–38 Women 31–35
	OLS	OLS	IV	IV
	(5)	(6)	(7)	(8)
$\frac{1}{2}\left(\ln \ n_m^{34-38} + \ln \ n_f^{31-35}\right)$	–0.198	–0.123	0.045	0.045
	(0.072)	(0.048)	(0.047)	(0.046)
P_m^{34-38} > primary education		–0.211		–0.037
		(0.133)		(0.105)
P_f^{31-35} > primary education		0.036		0.032
		(0.115)		(0.088)
Instrument			ln pop	ln pop
R-squared	0.141	0.198	0.011	0.012
Number of cities	189	189	189	189

Source: 1982 Census of China (1% random sample of the population).
Note: regressions weighted by city population with robust standard errors in parentheses.
μ_{mf} = number of marriages between men in a given age group and women in a given age group.
μ_{m0} = number of unmarried men in a given age group.
μ_{f0} = number of unmarried women in a given age group.
n_m = total number of men in a given age group.
n_f = total number of women in a given age group.
P_m = log odds ratio of men age 24–28 (in column 2) and age 34–38 (in column 4), who have completed primary education in a given city.
P_f = log odds ratio of women age 21–25 (in column 6) and age 31–35 (in column 8), who have completed primary education in a given city.
log (pop) = log (total population in a given city).

TABLE A.1

MARRIAGE REGRESSIONS (1ST STAGE IV), UNITED STATES 2000

	Dependent Variable	
	ln (men 27–31)	ln (women 25–29)
	(1)	(2)
ln $(\text{men}^{1980}_{7-11})$	0.930	
	(0.041)	
$\frac{\text{White men (27–31)}}{\text{Men (27–31)}}$	–1.056	
	(0.233)	
$\frac{\text{Black men (27–31)}}{\text{Men (27–31)}}$	–0.621	
	(0.352)	
fraction of 27-31 males – bachelor +	1.204	
	(0.486)	
Mean income, married men 27–31	0.218	
	(0.350)	
Variance family income	–0.139	
	(0.087)	
ln $(\text{women}^{1980}_{5-9})$		0.968
		(0.037)
$\frac{\text{White women (25–29)}}{\text{Women (25–29)}}$		–0.806
		(0.286)
$\frac{\text{Black women (25–29)}}{\text{Women (25–29)}}$		–0.390
		(0.320)
fraction of 25-29 females – bachelor +		0.945
		(0.454)
Mean income, married women 25–29		0.134
		(0.292)
Variance income, married women 25–29		–0.066
		(0.093)
R-squared	0.90	0.91
Number of cities	247	241

Source: U.S. 2000 census.

Note: regressions weighted by city population with robust standard errors in parentheses.

ln (men 27–31) = log (number of men in the 27–31 age group).

ln (women 25–29) = log (number of women in the 25–29 age group).

ln $(\text{men}^{1980}_{7-11})$ = log (number of men who were 7–11 years old in 1980).

ln $(\text{women}^{1980}_{5-9})$ = log (number of women who were 5–9 years old in 1980).

TABLE A2

1427 Florentine Catasto: the 29 Districts

District no.	Name of city or rural area	Population size (number of households)
1	Florence	9780
2	Rural quarter of S. Spirito (countryside of Florence)	7530
3	Rural quarter of S. Croce (countryside of Florence)	4881
4	Rural quarter of S. Maria Novella (countryside of Florence)	7369
5	Rural quarter of S. Giovanni (countryside of Florence)	7147
6	Cortona	898
7	Countryside of Cortona	1121
8	Castiglion Fiorentino (city and countryside)	559
9	Montepulciano (city and countryside)	733
10	Arezzo	1194
11	Countryside of Arezzo	3694
12	Mountain villages near Arezzo	1194
13	Pisa	1740
14	Countryside of Pisa	3966
15	Pistoia	1247
16	Countryside of Pistoia	2043
17	Mountain villages near Pistoaia	492
18	Rural villages in the Garfagnana area	175
29	Towns and rural villages in Val di Nievole	1262
20	Towns and rural villages in Val d'Arno di Sotto	642
21	San Gimignano (city and countryside)	576
22	Colle	571
23	Volterra	797
24	Coutryside of Volterra	751
25	Sillano and Montecastelli	119
26	??	
27	??	
28	??	
29	??	

References

[1] Anderberg, Dan, and Steeve Mongrain. "Marriage, Divorce and Earnings-Risk." Manuscript, University of Stirling and Simon Fraser University, 2001.

[2] Angrist, Joshua. "How Do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation." *Quarterly Journal of Economics* CXVII, no. 3 (2002): 997–1038.

[3] Au, Chun-Chung, and J. Vernon Henderson. "Are Chinese Cities too Small?" *Review of Economic Studies,* forthcoming.

[4] Chiappori, Pierre-Andre, and Yoram Weiss. "Marriage Contracts and Divorce: An Equilibrium Analysis." Manuscript, University of Chicago, 2000.

[5] Botticini, Maristella. "A Loveless Economy? Intergenerational Altruism and the Marriage Market in a Tuscan Town, 1415-1436." *Journal of Economic History* 59 (March 1999): 104–21.

[6] Botticini, Maristella. "Social Norms and Demographic Shocks: Dowry Contracts in Florence, 1250–1450." Manuscript, September 2006.

[7] Botticini, Maristella, and Aloysius Siow. "Why Dowries?" *American Economic Review* 93, no. 4 (September 2003): 1385–98.

[8] Choo, Eugene, and Aloysius Siow. "Who Marries Whom and Why." *Journal of Political Economy* 114, no. 1 (2006): 175–201.

[9] Compton, Janice, and Robert A. Pollak. "Why Are Power Couples Increasingly Concentrated in Large Metropolitan Areas?" NBER Working Paper no. 10918, 2004.

[10] Costa, Dora L., and Matthew E. Kahn. "Power Couples: Changes in the Location Choice of the College Educated." *Quarterly Journal of Economics* CXV, no. 4 (2000): 1287–1315.

[11] Drewianka, Scott. "Estimating Social Effects in Matching Markets: Externalities in Spousal Search." *Review of Economics and Statistics* 85, no. 2 (2003): 409–23

[12] Edlund, Lena. "Sex and the City." *Scandinavian Journal of Economics* 107, no. 1 (2005): 25–44.

[13] Gautier, Pieter, Michael Svarer, and Coen Teulings. "Marriage and the City." CESIFO Working Paper no. 1422, 2005.

[14] Goldstein, Sidney, and Alice Goldstein. "China." In Charles Nam, William Serow, and David Sly (eds.), *International Handbook on Internal Migration.* New York: Greenwood Press, 1990.

[15] Goldstein, Sidney, and Xiushi Yang. "Population Movement in Zhejiang Province: The Impact of Government Policies." *International Migration Review* 24, no. 3 (1990): 509–33.

[16] Gould, Eric G., and Daniele Paserman. "Waiting for Mr. Right: Rising Inequality and Declining Marriage Rates." *Journal of Urban Economics* 53, no. 2 (2003): 257–81.

[17] Herlihy, David, and Christiane Klapisch-Zuber. *Census and Property Survey of Florentine Dominions in the Province of Tuscany, 1427-1480.* Machine readable data file available at the University of Wisconsin Data and Program Library Service.

[18] Herlihy, David, and Christiane Klapisch-Zuber. *Les Toscans et leurs familles: Un étude du catasto Florentin de 1427. Paris:* Presses de la Fondation Nationale des Sciences Politiques, 1978 (English abridged edition: *Tuscans and their Families: A Study of the Florentine Catasto of 1427.* New Haven: Yale University Press, 1985).

[19] Mortensen, Dale T. "Matching: Finding a Partner for Life or Otherwise." *American Journal of Sociology* 94, supplement (1988): S215–40.

[20] Mortensen, Dale T. "Job Search and Labor Market Analysis." In Orley Ashenfelter and Richard Layard (eds.), *Handbook of Labor Economics*, vol. 2, pp. 849–920. Amsterdam: North Holland, 1986.

[21] Petrongolo, Barbara, and Christopher A. Pissarides. "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature* 39, no. 2 (2001): 390–431.

[22] Petrongolo, Barbara, and Christopher A. Pissarides. "Scale Effects in Markets with Search." *Economic Journal* 116 (2005): 21–44.

[23] Yang, Xiushi. "Patterns of Economic Development and Patterns of Rural-Urban Migration in China." *European Journal of Population* 12, no. 3 (1996a): 195–218.

[24] Yang, Xiushi. "Economic Reforms and Spatial Mobility." In *China: The Many Facets of Demographic Change*, edited by Alice Goldstein and Wang Feng, pp. 167–85. Boulder: Westview Press, 1996b.