

Rushed to the Altar: The Effect of Social Interactions on Migrant Workers' Marriage Age

Yan Hu^{*#}

CEMFI

17 November 2020

[Latest version available HERE](#)

ABSTRACT: This paper provides causal evidence that social interactions can regulate individual behaviour through norms, by looking at the marriage decisions of rural-to-urban migrants in China. Using variation of social pressure to conform to the rural norm of early marriage from co-workers from the same rural origin in the workplace, I find that the chance of getting married before 23 doubles for female migrants in China if the majority of their co-workers are from the same hometown. Little effect is found on male migrants because early marriage is more valued for females than males in agricultural societies. Consistent with the norm-based explanation for the association between early marriage and social interactions, the gender differential is larger for individuals from regions with more conservative values for women. I show that the gender-biased effect of social interactions is not driven by self-selection into social interactions.

Key words: social interactions, social norms, marriage, migration

JEL classification: D91, J12, J16, R23, Z13

^{*}Centro de Estudios Monetarios y Financieros (CEMFI), Casado del Alisal 5, 28014 Madrid. E-mail: yanhu@cemfi.edu.es.

[#]I am especially grateful for the guidance and support from Manuel Arellano and Diego Puga. I thank seminar participants at CEMFI, Econometrics Society Asian Meeting, Urban Economics Association, Summer School in Urban Economics and ASREC Europe for their helpful comments and suggestions. I thank Linping Liu from Nanjing University for sharing the data and Yanting Zhu for her assistance. Funding from the European Research Council under the European Union's Horizon 2020 Programme (ERC Advanced Grant agreement 695107 - DYNURBAN) and from Spain's Ministerio de Economía y Competitividad (Grants ECO-2016-80411-P, BES-2017-082187 and María de Maeztu Programme for Units of Excellence in R&D, MDM-2016-0684) is gratefully acknowledged.

1. Introduction

This paper provides causal evidence that social interactions can regulate individual behaviour through norms, by looking at the marriage decisions of rural-to-urban migrants in China. Individual behaviour in the society is not isolated. Cultural norms prescribe rules about what people in a social group should and should not do. Social interactions exert pressure on individuals to conform to norms, especially when the interactions are close and frequent. However, it is inherently difficult to identify the effect of social interactions on outcome, since in most cases individuals choose who to associate with in order to maximise their utility.

The empirical strategy developed in this study has several advantages for identifying the effect of social interactions. I use variation in social pressure to conform to the rural norm of early marriage from migrants from the same rural origin in the workplace. Compared to friendship, co-workers are less subject to self-selection bias because individuals have less control over the origin of their co-workers. By looking at rural-to-urban migrants, I isolate the effect of cultural norms from that of rural occupations. Given that the social norm of early marriage is much more pronounced for females in an agricultural society than for males, I can look at the effect separately by gender, which helps to difference out any selection bias that is identical for males and females. The gender differential effect of social interactions on marriage age is interpreted as social pressure to conform to the gender-biased norm of early marriage in an agricultural society.

Using a discrete-time hazard model, I find that intensive social interactions (with the majority of co-workers of the same origin) compels female migrants to abide by norms from their hometown, raising the fraction of females married before age 26 from 0.58 to 0.84.¹ Conditional on getting married before 40, women with a strong presence of co-workers from the same hometown are estimated to marry 2.6 years earlier. Social interactions have little impact on the marriage age of males.

To further address the selection bias in social interactions, I use propensity score matching to control for the selection on observables. I compare individuals who differ only in the composition of their co-workers but are otherwise equal, and find that the discrepancy in the effect of social interactions by gender is as large. I rule out spurious correlation between early marriage and social interactions by restricting the sample to individuals who have been married before migration. If entry into social interactions picks up, for example, a private preference for early marriage, we will also see the correlation for women who arrive at the cities married. However, both the significance level and point estimate drop to zero when we look at females who have been married before migration.

An alternative channel through which social interactions with co-workers of the same origin can affect marriage age is increasing the chance of matching between individuals who are more similar. However, the findings that the effect of social interactions is much less for males and not stronger with a more skewed gender ratio of migrant workers suggest that matching is not the main mechanism.

¹To put the number into perspective, extending the use of contraceptive pills to young unmarried women in the U.S contributed to a reduction in the fraction of married college graduate women before age 26 from 0.7 for the cohort born in 1950 to 0.54 for the cohort born in 1957 (Goldin and Katz, 2000).

In line with the norm-based explanation, I find that gender discrepancy in the effect of social interactions on marriage age is larger for individuals from areas that hold more traditional views on women. The gender-biased effect of social interactions on marriage age is consistent with the gender differential of marriage ages between rural and urban China, both reflecting the gender differentiated cultural norms on marriage.² During the prolonged period of agrarian economy in China, early female marriage was practised from generation to generation, facilitating the corresponding norm to form and be reinforced.³ According to the 2010 China General Social Survey, compared to urban dwellers, people from rural areas have a greater tendency to think that the role of females is primarily domestic and that marrying a good husband is of paramount importance for (Table A1).

As individuals migrate from rural to urban China, taking up jobs in manufacturing and services, the economic and social environment that allowed for early marriage to operate ceased to exist. However, behaviour may fail to adjust if individuals adhere to the old norms under social pressure, despite the price of early marriage for females being much dearer in modern society. I find that female migrants have a greater tendency than male migrants to marry early if they identify with their rural origin.

This paper contributes to the literature on the persistent effect of cultural norms (e.g. [Fernández and Fogli, 2006](#); [Fisman and Miguel, 2007](#)) and highlights how the effect depends on who one closely interacts with. Similar to [Munshi and Myaux \(2006\)](#), this paper focuses on the role social interactions play in coordinating norms, but the mechanism is distinct, with diverging implications on behaviour. In [Munshi and Myaux \(2006\)](#), social interactions serve to disseminate information, which helps individuals in the social group update their beliefs about the prevailing contraceptive practice. Consequently, behaviour (i.e. the use of contraceptive) also evolves. In my case, close social interactions exert pressure on individuals to comply with pre-existing norms. As a result, behaviour (i.e. early female marriage) perpetuates.

By emphasising the role co-workers play in shaping an individual's behaviour, this paper also connects to the literature on peer effects and social pressure (e.g. [Mas and Moretti, 2009](#); [Bandiera, Barankay and Rasul, 2010](#); [Burke and Sass, 2013](#); [Cornelissen, Dustmann and Schönberg, 2017](#)). I focus on a particular type of peers, that is, co-workers with the same place of origin. Hence, this paper relates to studies on ethnic enclaves, which have predominantly concern their pecuniary benefit (e.g. [Borjas, 1995](#); [Edin, Fredriksson and Åslund, 2003](#); [Zhang and Xie, 2013](#)), and explores the social aspects of the concentration of individuals from the same place of origin.

More broadly, the paper contributes to the literature and policy discussion on gender gap by shedding light on another source of gender inequality. Marriage age affects motherhood timing. A later timing of marriage and childbearing entails a substantial increase in lifetime income ([Buckles, 2008](#); [Miller, 2011](#); [Heath and Mobarak, 2015](#)). The social pressure of early marriage distorts the trade-off between career prospect and fertility and can result in sub-optimal (early) marriage age. The policy implication is that simply offering economic incentives for delayed marriage may not be

²Females tend to marry earlier in rural areas than in cities, but such distinction is less sharp for males (Figure A1).

³ Historical records date back to the Western Zhou period (1046–771 BC), wherein the dynasty stipulated the maximum marriage age of 30 for males and 20 for females with obligatory parental consent. During the reign of Emperor Hui (194–188 BC), unmarried women between 15 and 30 years were taxed as much as one year's consumption of crops.

enough to alter behaviour in the presence of strong social pressure. We need to think about ways to change either the social network structure or the way that social interactions regulate behaviour.

The remainder of this paper proceeds as follows. Section 2 introduces the institutional background of the emergence of a large number of migrant workers in China and their concentration based on place of origin. Section 3 describes the data. Section 4 presents the estimation results and the story. Section 5 rules out alternative explanations. Section 6 concludes.

2. Institutional background

The economic reform of China since 1978 set the momentum of a large-scale migration from rural to urban areas that are growing vibrantly with an outburst of employment opportunities in manufacturing and services. The rural reform freed farmers from land, further facilitating migration into cities. The number of migrant workers totalled 252 million in 2011, of which 43.2% concentrated in Pearl River Delta and Yangtze River Delta (Chinese National Bureau of Statistics). Figure A2 shows the migration outflow to Pearl River Delta and Yangtze River Delta from the survey data on migrant workers in 2010, which is the data set that the main analysis of this paper uses.

Hukou system is the institutional cause that gives rise to social interactions of migrant workers based on their place of origin. This housing registration system requires individuals to be registered under their place of origin and to be classified as either rural or urban residents, which is used to link local public welfare programs. The conversion from rural to urban *hukou* remains extremely difficult for rural migrants (Song, 2014). Lacking *hukou* status in the resident city, migrant workers have limited access to the local public goods and therefore have less bargaining power in the labor market (Song, 2014). Studies find that rural migrants face severe labour market discrimination and the estimates range from 28% to 60% income differentials even after controlling for observable characteristics (Quheng, 2007; Gravemeyer, Gries and Xue, 2010; Lee, 2012; Liu, 2005).

Furthermore, due to the pre-existing discrepancy in the level of development between rural and urban China, migrant workers also stand out for their appearance and accent, making them socially excluded from the city. According to the survey on migrant workers in 2010, more than 50% of the migrant workers once are discriminated by local residents and about 60% lack a sense of belongingness in the cities.

Because of their disadvantageous positions in cities, most migrant workers take up jobs that do not appeal to the locals, featuring long hours, poor working conditions and low and unstable pays (Wang and Zuo, 1999). Also because a strict migration restriction was carried out before 1978, the isolation created large socio-economic barriers (i.e. language, customs, income) for people from different parts of China to interact. Therefore migrant workers stay closely with their *tongxiang*, the Chinese word for people from the same home town, to share network, information and resources. Migrants in China resemble the ethnic minorities in the U.S., and similarly respond to discrimination faced in destination cities by forming social network based on their place of origin. Sometimes it takes the form of *tongxiang* enterprises where hiring decisions are made according to people's place of origin, resulting in a concentration of same origin co-workers within the workplace. The reason to join *tongxiang* network can be merely financial, but the effect can go beyond pecuniary

benefits. Frequent interactions with individuals from the same rural origin can affect behaviour that has a social dimension. One such example is marriage, which according to a survey conducted by China Youth Daily is also the main concern of migrant workers. Marriage can be affected by social interactions because they can increase the chances of matching with a partner. Additionally, marriage behavior can be regulated by social norms and subject to peer pressure. This paper focuses on the latter channel through which social interactions affect one of the major decisions one makes in lifetime.

3. Data

The main analysis of this paper is based on a survey data that interviewed a cross-sectional 4157 rural migrant workers in China in 2010. The survey was part of a project sponsored by Ministry of Education in China with the aim to study the status-quo of migrant workers and protect their rights. It took a representative sample of migrant workers working in Yangtze River Delta and Pearl River Delta in 2010. The two regions have the highest concentration of migrant workers in China and together assimilated more than 40% of total migrant workers in 2011. Quota sampling method was used to correct for representativeness in gender, occupation and regional distribution of migrants. The data was used by [Zhang and Xie \(2013\)](#) to study the effect of *tongxiang* network on migrant workers' wages and as a result in the 2010 round of the survey, questions regarding the relations with *tongxiang* were incorporated.

A Measures of social interactions

From the survey data, we obtain the information about the percentage of co-workers in the production line from the same province, the same county and the same town, which are respectively 1st, 3rd and 4th level administrative unit in China. Additionally, we know whether the three best friends of the migrant worker are *tongxiang*. In the survey, individuals were asked:

In the production line that you work, how many people come from the same town/county/province?

Table 1 shows the percentage of co-workers by the percentage of *tongxiang* co-workers in the same production line, defined by co-workers from the same town, the same county or the same province respectively. A non-trivial fraction of people have more than 30% of co-workers from the same rural origin.

The concentration of *tongxiang* co-workers is my preferred measure of social interactions because it is a less endogenous choice than friendship. The concentration of co-workers based on place of origin can result from close distance from the cities, a large out-migration population or owners' preferential hiring of *tongxiang*, which can be orthogonal to marriage conditions. Moreover, migrant workers spend a considerable amount of time with co-workers. The average working hours are 9.3 hours per day on an average 6-day working schedule. 32% of migrant workers even work 7 days a week. The intensity of working schedule results in the intensity and closeness of social interactions within the workplace. Migrant workers have plenty opportunities to socialize ([Fang, 2012](#)). Besides,

Table 1: Percentage of *tongxiang* co-workers

| | none | very few <10% | a few 10%-20% | not many 20%-30% | many 30%-50% | a lot >50% |
|----------|--------|------------------|------------------|---------------------|-----------------|---------------|
| Town | 48.29% | 26.51% | 9.07% | 7.17% | 3.96% | 5.01% |
| County | 41.24% | 26.49% | 11.74% | 9.14% | 5.30% | 6.10% |
| Province | 14.02% | 20.12% | 13.66% | 13.68% | 15.25% | 23.27% |

Notes: Based on the 2010 survey on migrant workers. The table shows concentration (in columns) of co-workers from the same town, county and province (in rows).

for the sample we are studying, 36% of total migrant workers live in the dormitory provided by their employers and the ratio increases to 49% for individuals who stay single. 62% of workers dine in the canteen of the workplace.

A second set of measures of social interactions is migrant workers' self-perceived friendship with *tongxiang*. In the survey, they were asked to identify whether three of their best friends are from the same home town:

Is your best friend tongxiang?

Table 2: Percentage of having *tongxiang* as best friends

| | |
|-----------------|-----|
| 1st best friend | 37% |
| 2nd best friend | 33% |
| 3rd best friend | 31% |
| every friend | 13% |
| any friend | 44% |

Notes: Based on the 2010 survey on migrant workers. The table shows fraction of migrant workers whose best friends is *tongxiang*, or a person from the same hometown. *Every friend* is an indicator whether all three best friends are *tongxiang*. *Any friend* is whether any of the three friends is *tongxiang*.

Table 2 shows that a considerable fraction of people have *tongxiang* as their best friends. However, friendship is subject to a greater extent to selection bias. People choose who their friends are. On the other hand, the measure of concentration of co-workers are more exogenous to the outcome we try to evaluate. We would expect that the primary function of *tongxiang* enterprises is to advance the economic prospect of migrant workers rather than their love affair. Therefore I will base my main analysis on the measure of co-workers to proxy for social pressure to conform to rural norms and present estimations using friendship measures as complementary evidence.

B Individual characteristics

The sample consists of 1895 females and 2252 males. After dropping observations that are either widower or divorced, I am left with a sample of 4093 individuals.

The average migrant worker has 9 years of education and migrate at an age around 19. The migration is typically temporary. Less than 25% migrant workers express desires to convert to local

hukou not to mention the practical barriers to acquire one. The majority of individuals work in manufacturing and service with an almost equal number of male and female workers. There are significantly more males in construction than females although the total number of employment is only 308, which makes up for difference in total number of females and males in the sample. Transportation is the 4th largest employment with balanced gender representation. The other sectors are relatively small and assimilate about 2% of workers (Figure A4).

In this paper, we compare marriage age between people who have co-workers of the same rural origin and people who do not, and compare the former difference by gender. Before delving into the estimation, I examine whether systematic difference in observable characteristics exist between individuals who have frequent interactions with *tongxiang* and who do not and whether it varies with gender.

In Table A2, I compare individuals with more than 30% of *tongxiang* co-workers (treated group) with individuals who have less than 30% co-workers from the same county (untreated group), by a wide range of personal traits: age, age at migration, migration length, distance of current city from home, education, working schedule and wage. We can see that although in general there exists some difference between males and females, and between the treated and the untreated individuals, once we take the difference in difference, the gap between the treated and untreated groups are no different between male and female migrants, except the age at migration. A treated male migrant leaves home town on average 1.9 years older than a treated female, but this cannot make it particularly difficult for them to find a partner as the average marriage age gap is 2 years in China and the difference in difference is not significant different from 0 at 10% level. Overall the comparison in the mean statistics suggests that sorting into *tongxiang* enterprises, if there is any, is not that different between male and female migrants.

C Complementary data

I also use 2000 census data that sampled 1% Chinese population to impute the number of *tongxiang* as well as the total number of migrant workers in the destination city⁴. We want to control for citywide presence of *tongxiang* because it may be correlated with *tongxiang* network in the workplace. Although it does not undermine the story of social pressure, but it helps to disentangle whether the influence is from *tongxiang* in the city or *tongxiang* in the workplace.

Additionally, I use Chinese General Social Survey in 2010 that asks about individual attitudes towards women to proxy for the difference in marriage norms in different parts of rural China.

4. Estimation methods, results and explanation

I study the effect of social interactions on the probability of getting married at different ages. If we estimate the effect of social interactions by comparing the average marriage age, we are overlooking individuals that have not been married in the sample. The fact that one is not married and at the

⁴It is ideal to compute the number of *tongxiang* from the same prefecture or county. However, we only know the original prefecture if the individual moves in the past 5 years. Therefore I calculate the number of *tongxiang* as the number of individuals from the same province

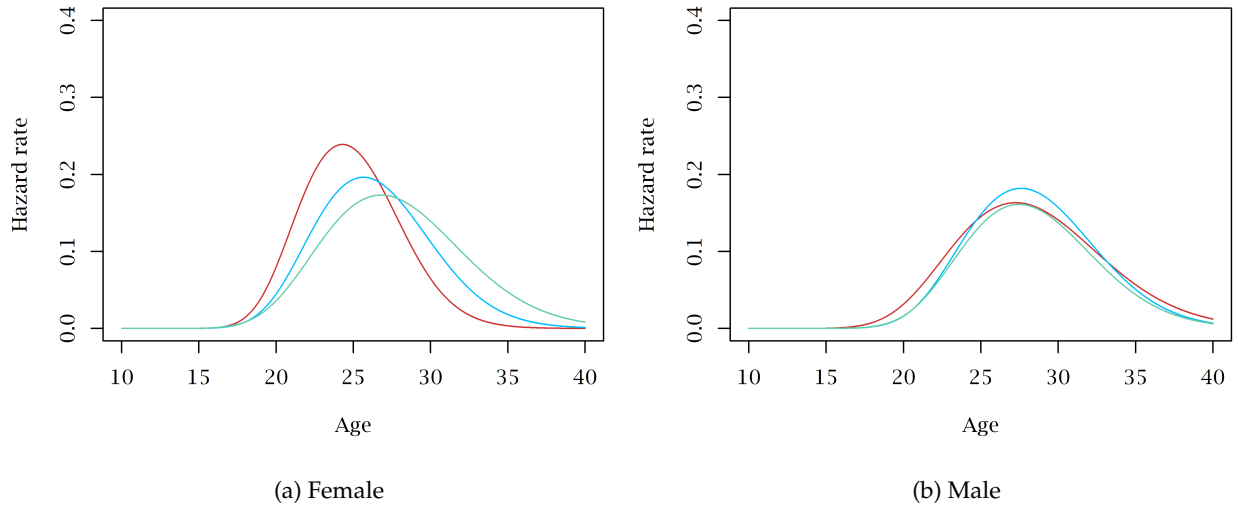


Figure 1: Estimated hazard rate: pc_county
green: none of co-workers come from the same county
blue: 0–30% of co-workers come from the same county
red : >30% of co-workers come from the same county

same has no co-workers of the same origin also provides information about the effect of social interactions on marriage. Using a hazard model allows us to use this information as well as look closely at the effect of social interactions on the whole distribution of marriage age, in addition to a single moment. I follow the specification of discrete time duration model by [Bover, Arellano and Bentolila \(2002\)](#) and define the hazard rate of marriage as:

$$h(t) = Pr(T_i = t | T_i \geq t) = G(\gamma_0 + \gamma_1(\ln t) + \gamma_2(\ln t)^2) \quad (1)$$

Where

$$G(x) = \frac{\exp(x)}{1 + \exp(x)} \quad (2)$$

The duration here is years to marriage. The hazard rate $h(t)$ is the conditional probability of getting married at age t given one has not been married before age t . The specification here allows the hazard rate to vary with age t .

A Estimation: hazard rate

I look at hazard rate of marriage and see how it varies with the percentage of co-workers from the same home town, and by gender.

Panel (a) of Figure 1 shows the estimated hazard rate for female migrant workers with different levels of concentration of co-workers from the same county. As the fraction of *tongxiang* co-workers

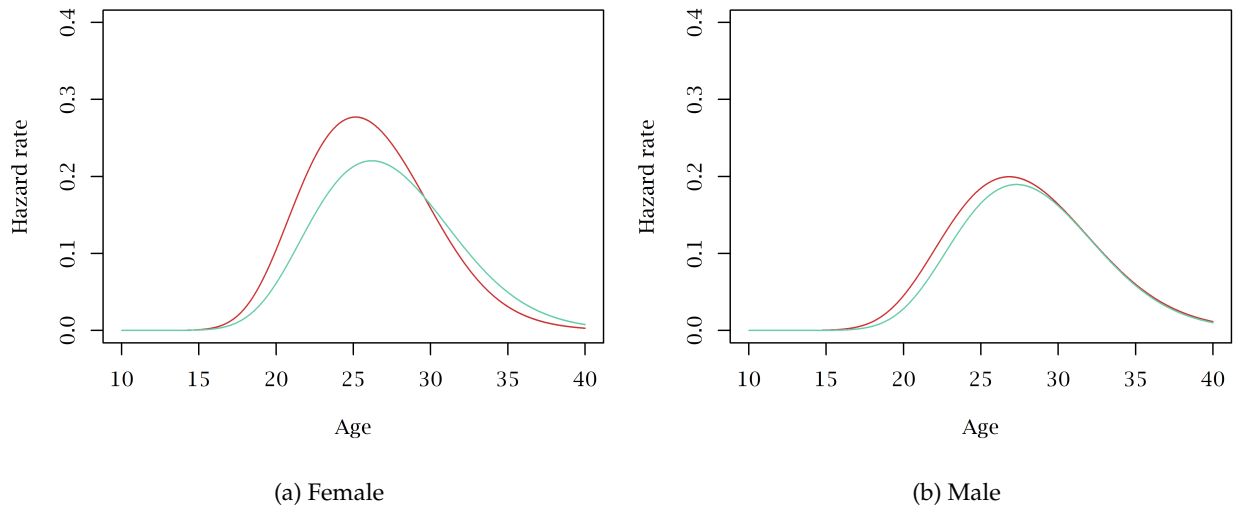


Figure 2: Estimated hazard rate: friend1

red: 1st best friend is *tongxiang*
green: 1st best friend is not *tongxiang*

Notes: *Tongxiang*, or people from the same hometown, self-defined by the survey respondents.

increases, we see a gradual shift and intensification of exit rate at an earlier age. This indicates that female migrant workers marry earlier with more co-workers from the same place in the production line. In contrast, the concentration of *tongxiang* co-workers is not associated with early transition to marriage for male migrant workers.

We have very similar gender-differential patterns if we look at co-workers from the same town and the results are shown in Appendix (Figure A5). The effect of co-workers from the same province is less clear (Figure A7). Given the average area of a province in China is as large as Germany and the average population is that of Spain, there is large heterogeneity within a province⁵ Individuals may stop to associate people from the same province with *tongxiang*.

Next we turn to the effect on marriage age using self perceived friendship as a proxy for social interactions with *tongxiang* (Figure 2). Similarly for females, if her best friend comes from the same home town, she is more likely to marry at an earlier age while for males the effect is much smaller. The results from the 2nd and 3rd best friend are shown in Appendix (Figure A6, A8). Although the friendship measure is subject to selection bias, the same gender differentiated pattern is consistent with the story that females are pressured into early marriage because of social interactions with *tongxiang*.

⁵The average area of a province in mainland China is 352,033 km^2 , excluding provincial level municipalities (Beijing, Shanghai, Tianjin, Chongqing). The average population of a province is 46 million in 2010.

B Estimation: cumulative distribution function

The hazard rate offers qualitative evidence for the effect of social interactions on marriage age. However, it does not tell us quantitatively the magnitude of the effect and its significance level. Importantly, comparing directly the difference in hazard rate for any given age is problematic. At each point of time, as the more marriageable individuals exit, and more individuals exit in the group of people who have intensive social interactions with *tongxiang*, the compositions of population change differently for groups with different levels of social interactions. This can be seen from the fattening out of the right-hand tail of hazard rate for the group of female migrants with high concentration of co-workers from the same county. Therefore as time goes by, people who remain unmarried in the treated group (intensive social interactions) are on average less marriageable than people in the untreated group. The difference in hazard rates is a combination of treatment effect of social interactions and compositional difference between the treated and untreated groups. As the treated group overtime has a less marriageable population, the difference in hazard rates is likely an underestimation of the treatment effect.

In order to quantify the treatment effect of social interactions, I translate the hazard rate into cumulative distribution function, calculated as:

$$F(t) = 1 - \prod_{t=1}^T (1 - h(t)) \quad (3)$$

The cumulative distribution function calculates the fraction of individuals or the probability of an individual gets married before a certain age t . We can interpret the difference in the probability of getting married between the treated and untreated group as the effect of social interactions.

Table 3 compares individuals that have more than 50% of the co-workers from the same county with ones with none of the co-workers from the same county and restricts the sample to individuals who get married after migration. For visibility, the table only shows the probability of getting married before age 20 to 30. The baseline group is migrant workers with none of the co-workers from the same county. The parameter of interest is the difference in the probability of getting married between migrant workers with the majority of co-workers from the same rural area and the baseline group. Bootstrap standard errors with 500 replicates are shown in parenthesis ⁶.

For a female who works in a place without any *tongxiang*, she is expected to get married by 23 with probability 29%. If the majority of her co-workers come from the same county, the chance of being married by 23 doubles, increased by 30 percentage points. The effects on male workers are much smaller and remain insignificant for most ages.

As a placebo test, we can make the same calculation for individuals who have already been married before migration. If we think that social interactions have no real impact on marriage behavior, but instead capture private preference for early marriage, we would expect the correlation

⁶ The estimation of duration model generates $t-15$ (the minimum marriage age is set to be 16) observations for individuals who get married at age t and generates the number of observations that equals to one's age minus 15 if the individuals is not married. The bootstrap error is calculated by clustering at the individual level. Clustered bootstrapping is equivalent to taking a random sample of individuals from the original sample (with replacement) and generates the above-mentioned expanded sample each time.

between early marriage and social interactions present for females who come to the factories married. The result is shown in Table 4. Now the effect of social interactions disappears for females as well. The result is unlikely to be driven by a small sample⁷. Not only does the significance level fall but also the point estimates. The placebo results suggest that the association of *tongxiang* network in the workplace with an earlier marriage age for female migrants is not driven by spurious correlations, ie. selection into social interactions based on factors that also contribute to early marriage. The effect of social interactions on marriage is likely causal. What remains to be known is the mechanism.

C The effect of social interactions through culture

Anecdotal evidence suggests that female migrant workers in the workplace face considerable social pressure to get married. Fang (2012) depicts a vivid picture of social interactions in an electronic factory in Shenzhen, Pearl River Delta:

On the factory floor, women have plenty chances to compare themselves to one another...Married women gossip aggressively about the unmarried, especially the "old" ones, and the gossiping does not take into account any career performance. No matter how well an unmarried woman performs in her job or how high the job position she attains, she will still be singled out for her embarrassing unmarried status. In other words, she has simply 'failed' to marry.

I provide two sets of evidence to support the idea that the effect of social interactions on marriage works through social pressure to conform to rural norms. First, female migrants who identify more with their rural origin are more likely to get married early. This is to show that rural norms make female marry early. Second, the gender differential response to the presence of *tongxiang* co-workers is larger for individuals originated from places that hold more conservative attitudes toward women's role in the society. This is to show that the social pressure to marry early should be stronger if the norm of early marriage is stronger.

a Farmer v.s. worker identity

If we believe cultural norms in agricultural society make females marry early, we expect that females who identify more with their rural origin are more likely to marry early, conditional on the same economic conditions. Although in the sample less than 1% of individuals work in the agriculture sector in cities, a large number of migrants still identify themselves as farmers because of their rural origin and rural *hukou*.

In the survey of migrant workers, people were asked:

Which do you think is your identity? Do you recognize yourself as a farmer or a worker?

Table 5 calculates the probability of getting married before a certain age separately for males and females and only for individuals who are married after migration. The baseline group is migrant

⁷537 females and 565 males marry before migration(the estimation in Table 4 is based on 473(27%) individuals compared to 1360 individuals in Table 3).

workers who identify themselves as workers. Column 3 and Column 5 are the differences in the probability of getting married between an individual who identify herself as a farmer and as a worker. The table shows that female migrants that recognize themselves as farmers are more likely to get married early. For males, there is little effect of identification with farmers on marriage age. Self perceived identity does not seem related to the actual jobs migrants take up. 64.06 % of people who regard themselves as farmers work in manufacturing compared to 62.61% of people as workers. For service, the second largest category the comparison is 22.51% v.s. 23.92 %.

One concern is that people who identify with their rural origin may be more likely to interact with *tongxiang* and the correlation between social interactions and early marriage reflects the effect of self-image on marriage behaviour. Table 7 shows that identities do not correlate with social interactions. Another interpretation of the lack of correlation between the two is that social interactions do not fundamentally modify the values and beliefs of migrant workers. Females rush into early marriage under social pressure without necessarily altering their self identity and preferences.

I find even for individuals who marry before migration, females who identify with their rural origin marry early but males do not (Table 6). Therefore, self identity in this case is likely to reflect a set of values and norms that an individual appreciates and uses to guide behavior. People who identify with their rural origin value early female marriage, which is the norm in agricultural societies, and accordingly marry early.

b Traditional v.s. non-traditional provinces

If the rural norms pressure females into early marriage in the presence of *tongxiang* co-workers, we would expect a stronger effect for individuals from more traditional areas of China and hence a larger gender differential. I define traditional provinces as ones that hold more conservative views regarding the role of females in the society, using the response in China General Social Survey in 2010 to the question:

Do you agree that for females, it is more useful to have a good husband than a good career?

I divide provinces by "traditionalness" by the median value of the proportion of rural respondents agree to the above statement (Figure A9). I compare the difference in the gender-differential of the effect of social interactions on marriage age between individuals from the more traditional and the less traditional provinces. If we index social interactions s, ns , females and men F, M , and "traditionalness" t, nt , the triple difference of the effect of social interactions on the probability of getting married ΔF is:

$$\begin{aligned} \Delta F = & F_{s,F,t} - F_{ns,F,t} && \text{effect on females from traditional provinces} \\ & - (F_{s,M,t} - F_{ns,M,t}) && \text{effect on males from traditional provinces} \\ & - (F_{s,F,nt} - F_{ns,F,nt}) && \text{effect on females from non-traditional provinces} \\ & + (F_{s,M,nt} - F_{ns,M,nt}) && \text{effect on males from non-traditional provinces} \end{aligned} \quad (4)$$

As shown in column (1) of Table 8, the gender differential is larger for individuals from provinces with more conservative values on females. The gender differential of the probability of getting married before 24 is 39 percentage points higher in the more traditional provinces than in the less traditional provinces. Females from areas with more traditional norms face a higher social pressure of getting married early.

5. Alternative explanations

In this part I try to rule out alternative mechanisms that can generate the gender differential effect of social interactions.

A Matching

Migrant workers originated from the same rural area share customs, languages and similar socio-economic conditions. If *tongxiang* network facilitates meeting fellow countrymen that are more similar to the individual, then migrant workers will have a better chance of matching with a potential partner and a quicker transition into marriage. Although I do not know in the sample the identity of the spouse, there are some testable implications of the matching story. In Table 9, I examine whether the spouse of the migrant worker in a *tongxiang* enterprise is more likely to be in the same enterprise. The dependent variable is an indicator for whether the spouse works in the same enterprise. The results show that the spouse is more likely to be in the same workplace if one works with more co-workers from the same county. However, this is true even for people who marry before marriage. Therefore instead of the concentration of co-workers resulting in marrying a co-worker and having the spouse in the same workplace, the causality is likely to reverse: people who are married (possibly come from the same place) are more likely to enter *tongxiang* network. For people who marry after migration, only the spouse of a female who has a lot of *tongxiang* co-workers is more likely to be in the same workplace. This is compatible with the culture story. Females are pressured into early marriage with close interactions with co-workers of the same origin and marry someone in the workplace, regardless of whether he is a *tongxiang* or not.

The gender gap in the effect of social interactions can also be reconciled with the matching story if there are more male migrants than females. Still if matching is the main operating mechanism, we would expect that the gender differential to be larger in areas with a lower female to male ratio. Column (2) of Table 8 shows that the gender gap does not increase with a more skewed gender ratio.

B Spurious correlation

Social interactions with *tongxiang* may correlate with other factors that also make people follow rural norms. For instance, people who enter into *tongxiang* network may be more likely to return to their home towns and therefore abide by the rural norms. However I find that for both females and males in the sample, people who are surrounded with *tongxiang* co-workers stay longer in cities (Table A10).

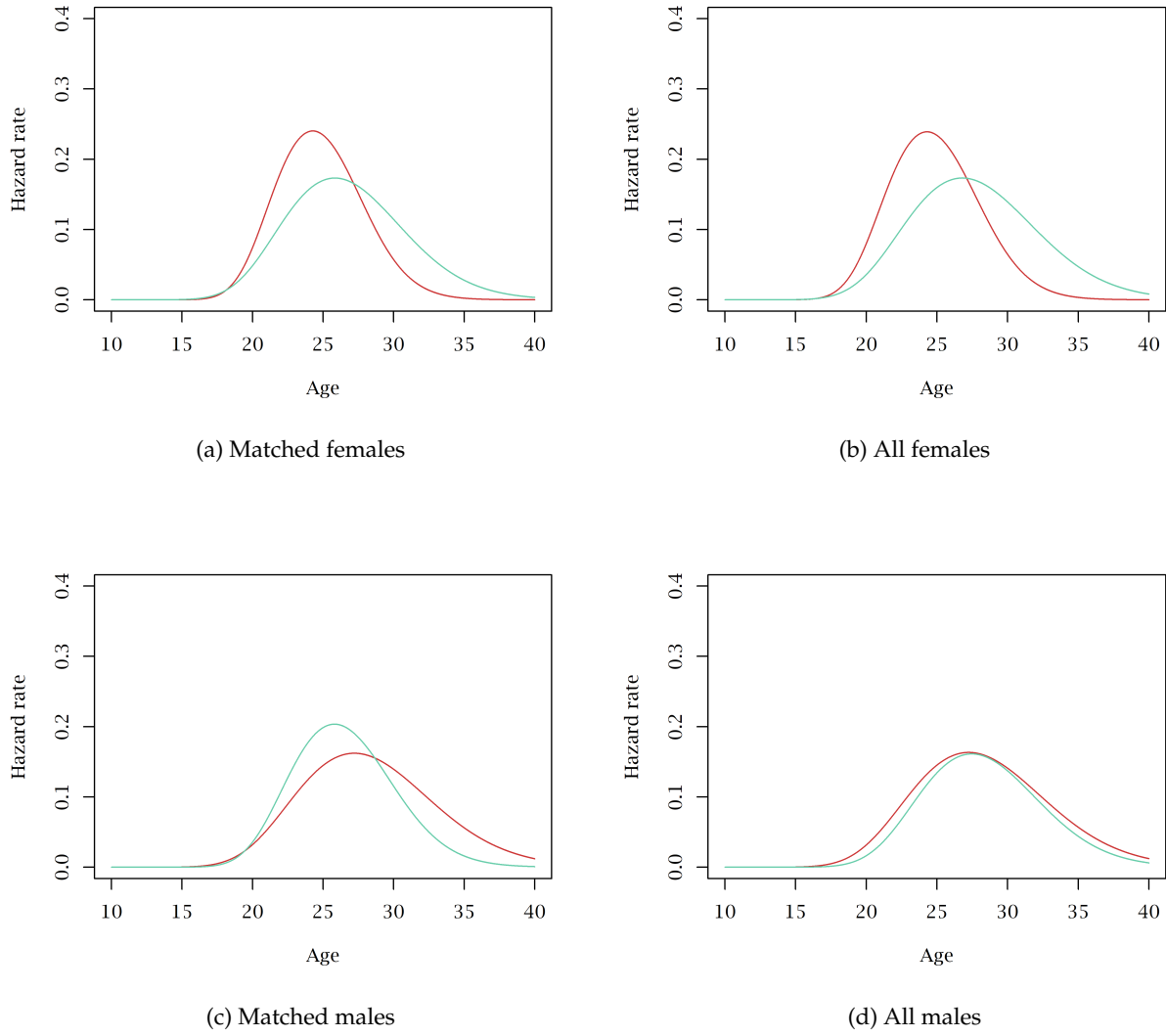


Figure 3: Estimated hazard rate for matching and all Samples

green: none of co-workers come from the same county
red: >30% of co-workers come from the same county

Notes: The right panels are the same as Figure 1 without the intermediate group. The left panels match the control group (in green) with the treated group (in red) by a set of individual attributes: occupation, education, age starting to work in the current job, whether it is a first job, whether it is a job by referral, the number of *tongxiang* in the same city, the number of total migrant workers in the same city, and origin and destination provinces.

More generally, social interactions can correlate with factors that affect the timing of marriage. For example, people who are less educated may be more likely to rely on *tongxiang* network and at the same time more likely to get married early. Also people who are more resourceful get to join *tongxiang* network and meanwhile have a better chance of finding a partner. Additionally, the concentration of *tongxiang* co-workers can be correlated with the concentration of *tongxiang* in the city and the results capture the effect of social interactions citywide instead of in the workplace. Although to produce the gender differential effect, we need these spurious correlations to differ

between men and women.

To address these concerns, I use propensity score matching separately by gender⁸. The idea is to compare individuals that are otherwise similar, but only differ in composition of co-workers from the same rural area. We can also control for individual characteristics by including them as additive terms in Equation 1. The general picture is similar.

Table 10 shows various factors that affect one's chances of entering into a *tongxiang* network with more than 30% of co-workers from the same county. These factors are used to calculate propensity scores to match individuals. People with *tongxiang* co-workers are less educated and it is more likely to be their first job acquired through a referral. The ratio of *tongxiang* citywide is positively correlated with entering into a *tongxiang* enterprise.

The hazard rate of marriage using the matched sample is shown in Figure 3. The results do suggest positive sorting into *tongxiang* network as the matched control groups shift to the left. People who enters into *tongxiang* network have traits that make them marry early. However, the sorting happens in the same direction for male and female migrant workers, and is perhaps stronger for males, which makes the gender differential effect of social interactions even larger.

6. Conclusion

This paper investigates one hypothesis put forward in economics and sociology literature, that is, the broader social context can influence individual behaviour by looking at marriage, which is a major choice in lifetime. I study migrant workers in China who set foot in urbanized coastal areas to seize the economic opportunities brought by the reform since 1978. The institutional set-up is appealing because of the large discrepancy between rural and urban China both in the levels of development as well as in culture and norms.

When an individual migrates from rural to urban China, the economic conditions drastically change, but their behaviour may fail to adjust in the face of social pressure to conform to rural norms. Many females were empowered with economic autonomy for the first time in history, which fundamentally changed the trade-off between career and family. Meanwhile *tongxiang* network provided migrant workers a temporary shelter from the outside discrimination faced in cities. Given strong collectivistic tendencies in China, the concentration of co-workers from the same rural origin can pressure individuals to conform to norms from their hometown. In rural China, the norm of early female marriage formed in the course of its 3000 years' history of agricultural economy. The paper shows that female migrants surrounded by co-workers from the same rural origin are more likely to get married at a younger age whereas such effect is not found for male migrants.

Consistent with the explanation that social interactions pressure females into early marriage to conform to rural norms, I find that the gender differential to be larger for individuals from areas that hold more conservative values for women. The effect is unlikely to be driven by selection bias as I find no effect for individuals who marry before migration. Close interactions can also facilitate matching between individuals from the same region who share customs and languages. However, I

⁸The estimation procedure of the duration model with propensity score matching follows Austin (2014). The matching function follows Zhang and Xie (2013). In addition, I include occupational dummies and dummies for the origin province.

do not find a similar effect for males and the gender gap is not larger with a more skewed gender ratio.

The findings carry important policy implications: the economic independence of female migrant workers does not empower them to freely choose when to marry. We can analyse the welfare change of early marriage due to social pressure for the individual as well as for the society as a whole, which can be an interesting topic for future research.

References

- Austin, Peter C**, "The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments," *Statistics in medicine*, 2014, 33 (7), 1242–1258.
- Bandiera, Oriana, Iwan Barankay, and Imran Rasul**, "Social incentives in the workplace," *The Review of Economic Studies*, 2010, 77 (2), 417–458.
- Borjas, George**, "Ethnicity, Neighborhoods, and Human-Capital Externalities," *American Economic Review*, 1995, 85 (3), 365–90.
- Bover, Olympia, Manuel Arellano, and Samuel Bentolila**, "Unemployment duration, benefit duration and the business cycle," *The Economic Journal*, 2002, 112 (479), 223–265.
- Buckles, Kasey**, "Understanding the returns to delayed childbearing for working women," *American Economic Review*, 2008, 98 (2), 403–07.
- Burke, Mary A and Tim R Sass**, "Classroom peer effects and student achievement," *Journal of Labor Economics*, 2013, 31 (1), 51–82.
- Cornelissen, Thomas, Christian Dustmann, and Uta Schönberg**, "Peer effects in the workplace," *American Economic Review*, 2017, 107 (2), 425–56.
- Edin, Per-Anders, Peter Fredriksson, and Olof Åslund**, "Ethnic enclaves and the economic success of immigrants—Evidence from a natural experiment," *The quarterly journal of economics*, 2003, 118 (1), 329–357.
- Fang, I-Chieh**, *Keen to get married: Why marriage is so important to independent female migrant workers in Shenzhen, China*, Feminist Research Centre in Aalborg (FREIA), Institute for History, International and Social Studies, Aalborg University, 2012.
- Fernández, Raquel and Alessandra Fogli**, "Fertility: The role of culture and family experience," *Journal of the European Economic Association*, 2006, 4 (2-3), 552–561.
- Fisman, Raymond and Edward Miguel**, "Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets," *Journal of Political economy*, 2007, 115 (6), 1020–1048.
- Goldin, Claudia and Lawrence F Katz**, "Career and Marriage in the Age of the Pill," *American Economic Review*, 2000, 90 (2), 461–465.
- Gravemeyer, Stefan, Thomas Gries, and Jinjun Xue**, "Income determination and income discrimination in Shenzhen," *Urban Studies*, 2010, p. 0042098010375317.
- Heath, Rachel and A Mushfiq Mobarak**, "Manufacturing growth and the lives of Bangladeshi women," *Journal of Development Economics*, 2015, 115, 1–15.

- Lee, Leng**, "Decomposing wage differentials between migrant workers and urban workers in urban China's labor markets," *China Economic Review*, 2012, 23 (2), 461–470.
- Liu, Zhiqiang**, "Institution and inequality: the hukou system in China," *Journal of comparative economics*, 2005, 33 (1), 133–157.
- Mas, Alexandre and Enrico Moretti**, "Peers at work," *American Economic Review*, 2009, 99 (1), 112–45.
- Miller, Amalia R**, "The effects of motherhood timing on career path," *Journal of Population Economics*, 2011, 24 (3), 1071–1100.
- Munshi, Kaivan and Jacques Myaux**, "Social norms and the fertility transition," *Journal of development Economics*, 2006, 80 (1), 1–38.
- Quheng, Deng**, "Earnings Differential between Urban Residents and Rural Migrants: Evidence from Oaxaca-Blinder and Quantile Regression Decompositions [J]," *Chinese Journal of Population Science*, 2007, 2, 8–16.
- Song, Yang**, "What should economists know about the current Chinese hukou system?," *China Economic Review*, 2014, 29, 200–212.
- Wang, Feng and Xuejin Zuo**, "Inside China's cities: Institutional barriers and opportunities for urban migrants," *The American Economic Review*, 1999, 89 (2), 276–280.
- Zhang, Chunni and Yu Xie**, "Ethnic Enclaves Revisited: Effects on Earnings of Migrant Workers in Urban China.," *Society: Chinese Journal of Sociology/Shehui*, 2013, 33 (1).

7. Table

Table 3: Differences in marriage between 0%+ and 50% *tongxiang* co-workers for individuals married after migration

| Age | Probability of married before certain age | | | |
|-----|---|---------------------|-------------------------------------|-------------------|
| | Female | | Male | |
| | Baseline: 0 <i>tongxiang</i> (1) | Difference (2) | Baseline: 0 <i>tongxiang</i> (3) | Difference (4) |
| 20 | 0.067 | 0.106 (0.117) | 0.027 | 0.029 (0.020) |
| 21 | 0.122 | 0.180* (0.109) | 0.058 | 0.044* (0.026) |
| 22 | 0.199 | 0.252** (0.101) | 0.110 | 0.057* (0.032) |
| 23 | 0.292 | 0.298*** (0.095) | 0.184 | 0.065* (0.038) |
| 24 | 0.392 | 0.310*** (0.091) | 0.273 | 0.067 (0.046) |
| 25 | 0.489 | 0.295*** (0.088) | 0.371 | 0.064 (0.053) |
| 26 | 0.576 | 0.263*** (0.085) | 0.466 | 0.058 (0.058) |
| 27 | 0.650 | 0.225*** (0.082) | 0.552 | 0.051 (0.060) |
| 28 | 0.708 | 0.189** (0.080) | 0.624 | 0.046 (0.059) |
| 29 | 0.754 | 0.158** (0.079) | 0.681 | 0.043 (0.057) |
| 30 | 0.788 | 0.132* (0.079) | 0.724 | 0.041 (0.055) |

Notes: Marriage hazard estimated based on the sample of individuals married after migration. Calculation of the probability of getting married for age 20 to 30. The baseline group is migrants with none of their co-workers from the same county. The difference is the probability of getting married for migrants with the majority of their co-workers from the same county relative to the baseline group. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

Table 4: Differences in marriage between 0%+ and 50% *tongxiang* co-workers for individuals married before migration

| Age | Probability of married before certain age | | | |
|-----|---|-------------------|-------------------------------------|-------------------|
| | Female | | Male | |
| | Baseline: 0 <i>tongxiang</i> (1) | Difference (2) | Baseline: 0 <i>tongxiang</i> (3) | Difference (4) |
| 20 | 0.314 | -0.111 (0.113) | 0.159 | 0.071 (0.051) |
| 21 | 0.450 | -0.085 (0.108) | 0.268 | 0.087 (0.063) |
| 22 | 0.587 | -0.039 (0.097) | 0.400 | 0.092 (0.073) |
| 23 | 0.708 | 0.003 (0.084) | 0.537 | 0.086 (0.077) |
| 24 | 0.805 | 0.026 (0.067) | 0.662 | 0.072 (0.074) |
| 25 | 0.874 | 0.031 (0.049) | 0.764 | 0.056 (0.065) |
| 26 | 0.921 | 0.027 (0.034) | 0.840 | 0.041 (0.054) |
| 27 | 0.951 | 0.020 (0.023) | 0.892 | 0.029 (0.043) |
| 28 | 0.970 | 0.013 (0.015) | 0.927 | 0.020 (0.034) |
| 29 | 0.981 | 0.008 (0.010) | 0.950 | 0.015 (0.026) |
| 30 | 0.988 | 0.005 (0.007) | 0.964 | 0.011 (0.020) |

Notes: Marriage hazard estimated based on the sample of individuals married after migration. Calculation of the probability of getting married for age 20 to 30. The baseline group is migrants with none of their co-workers from the same county. The difference is the probability of getting married for migrants with the majority of their co-workers from the same county relative to the baseline group. Bootstrap standard errors with 500 replicates are reported in parentheses.

***, **, and * indicate significance at the 1, 5, and 10 percent levels.

Table 5: Differences in marriage between worker and farmer identity
for individuals married after migration

| Age | Probability of married before certain age | | | |
|-----|---|--------------------|-------------------------|-------------------|
| | Female | | Male | |
| | Baseline: worker (1) | Difference (2) | Baseline: worker (3) | Difference (4) |
| 20 | 0.076 | 0.035 (0.029) | 0.035 | 0.003 (0.011) |
| 21 | 0.135 | 0.060 (0.037) | 0.070 | 0.009 (0.018) |
| 22 | 0.214 | 0.086* (0.045) | 0.124 | 0.020 (0.026) |
| 23 | 0.307 | 0.105** (0.053) | 0.197 | 0.034 (0.034) |
| 24 | 0.406 | 0.113* (0.058) | 0.283 | 0.048 (0.041) |
| 25 | 0.502 | 0.108* (0.062) | 0.376 | 0.059 (0.045) |
| 26 | 0.588 | 0.095 (0.063) | 0.466 | 0.065 (0.047) |
| 27 | 0.660 | 0.077 (0.063) | 0.547 | 0.067 (0.047) |
| 28 | 0.718 | 0.058 (0.064) | 0.616 | 0.064 (0.045) |
| 29 | 0.763 | 0.041 (0.065) | 0.672 | 0.059 (0.044) |
| 30 | 0.798 | 0.026 (0.067) | 0.716 | 0.053 (0.044) |

Notes: Marriage hazard estimated based on the sample of individuals married after migration. Calculation of the probability of getting married for age 20 to 30. The baseline group is individuals who identify themselves as workers. The difference is the probability of getting married for migrants who identify themselves as farmers. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

Table 6: Differences in marriage between worker and farmer identity
for individuals married before migration

| Age | Probability of married before certain age | | | |
|-----|---|--------------------|-------------------------|-------------------|
| | Female | | Male | |
| | Baseline: worker (1) | Difference (2) | Baseline: worker (3) | Difference (4) |
| 20 | 0.303 | -0.009 (0.059) | 0.177 | -0.012 (0.038) |
| 21 | 0.423 | 0.049 (0.061) | 0.293 | -0.021 (0.047) |
| 22 | 0.546 | 0.102* (0.060) | 0.430 | -0.030 (0.053) |
| 23 | 0.660 | 0.129** (0.056) | 0.571 | -0.037 (0.058) |
| 24 | 0.756 | 0.126** (0.050) | 0.695 | -0.041 (0.058) |
| 25 | 0.831 | 0.106** (0.043) | 0.794 | -0.041 (0.054) |
| 26 | 0.886 | 0.081** (0.037) | 0.865 | -0.038 (0.046) |
| 27 | 0.925 | 0.058* (0.030) | 0.914 | -0.033 (0.038) |
| 28 | 0.950 | 0.040* (0.024) | 0.945 | -0.028 (0.030) |
| 29 | 0.967 | 0.027 (0.019) | 0.964 | -0.024 (0.024) |
| 30 | 0.978 | 0.018 (0.015) | 0.976 | -0.020 (0.019) |

Notes: Marriage hazard estimated based on the sample of individuals married before migration. Calculation of the probability of getting married for age 20 to 30. The baseline group is individuals who identify themselves as workers. The difference is the probability of getting married for migrants who identify themselves as farmers. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

Table 7: Correlation between identity and *tongxiang* co-workers

| | Dep. var.: Concentration of co-workers from the same county |
|----------------------|---|
| | (1) |
| Male | 0.132* (0.070) |
| Farmer | 0.033 (0.080) |
| Farmer \times male | -0.013 (0.114) |
| Constant | 1.236*** (0.050) |
| Observations | 2,928 |
| R-squared | 0.002 |

Notes: The dependent variable is a categorical variable that increases with the level of concentration of co-workers from the same county (0: none; 1: 0%-10%; 2: 10%-20%; 3: 20%-30%; 4: 30%-50%; 5: more than 50%). Robust standard errors in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

Table 8: Gender differentiated effect of social interactions on marriage

| Age | Triple differences of probability of married before certain age | |
|-----|---|-----------------------------------|
| | More traditional v.s. less traditional (1) | More males v.s. less males (2) |
| 20 | 0.313 (0.408) | 0.122 (0.359) |
| 21 | 0.287 (0.361) | 0.141 (0.330) |
| 22 | 0.286 (0.298) | 0.130 (0.287) |
| 23 | 0.327 (0.228) | 0.086 (0.247) |
| 24 | 0.389** (0.181) | 0.018 (0.222) |
| 25 | 0.427** (0.167) | -0.061 (0.212) |
| 26 | 0.423** (0.169) | -0.135 (0.208) |
| 27 | 0.397** (0.172) | -0.191 (0.203) |
| 28 | 0.367** (0.174) | -0.227 (0.199) |
| 29 | 0.341* (0.175) | -0.241 (0.194) |
| 30 | 0.318* (0.175) | -0.239 (0.190) |

Notes: Column (1) compares the gender differential in the probability of getting married before a given age between individuals from more traditional and less traditional provinces. More traditional provinces are defined as ones that have a larger fraction of rural residents agreeing that for females it is more useful to have a good husband than a good career. Column (2) compares the gender differential between individuals in cities with a higher origin-destination specific male to female migrants ratio and in ones with a lower male to female migrants ratio, imputed from the population census in 2000. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

Table 9: Correlation between *tongxiang* co-workers and spouse workplace

| | Dep. var.: 1(Spouse in the same workplace) | | |
|----------------|--|---|--------------------------------------|
| | Married before migration (1) | Married after migration before the current job (2) | Married after the current job (3) |
| Male | 0.017 (0.034) | 0.051 (0.036) | 0.058 (0.051) |
| Pc_county | 0.053*** (0.014) | 0.045*** (0.016) | 0.094*** (0.028) |
| MaleXPc_county | -0.016 (0.018) | -0.027 (0.020) | -0.086*** (0.031) |
| Constant | 0.180*** (0.024) | 0.146*** (0.027) | 0.147*** (0.041) |
| Observations | 1,050 | 790 | 444 |
| R-squared | 0.026 | 0.015 | 0.031 |

Notes: *Pc_county* is a categorical variable whose value increases with the fraction of co-workers from the same county (0: none; 1: 0%-10%; 2: 10%-20%; 3: 20%-30%; 4: 30%-50%; 5: more than 50%). Robust standard errors in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

Table 10: Determinants of concentration of *tongxiang* co-workers

| | Dep. var.: 1 (more than 30% <i>tongxiang</i> co-workers) | |
|------------------------------------|--|----------------------|
| | Female (1) | Male (2) |
| Education | -0.502*** (0.71) | -0.599*** (0.132) |
| Age_work | 0.037** (0.018) | -0.003 (0.017) |
| First job | 0.485* (0.289) | 0.681*** (0.244) |
| Referral | 0.835*** (0.256) | 0.779*** (0.201) |
| Log (<i>tongxiang</i>) | 0.439* (0.234) | 0.195 (0.156) |
| Log (migrant) | -0.707** (0.275) | -0.277 (0.197) |
| Occupation dummies | Yes | Yes |
| Origin province fixed effects | Yes | Yes |
| Destination province fixed effects | Yes | Yes |
| Observations | 614 | 820 |

Notes: The dependent variable is an indicator for having more than 30% of the co-workers from the same county. *Education* is a categorical variable that increases with the level of education (0: less than or equal to primary school; 1: junior high school; 2: senior high school; 3: more than high school). *Age_work* is the age that the individual starts to work for the current job. *First job* is an indicator for whether the job is the first job. *Referral* indicates whether the job is obtained through referral. *Log (tongxiang)* is natural logarithm of the total number of rural migrants from the same county in the same city. *Log (migrant)* is natural logarithm of the total number of migrant workers in the same city. Robust standard errors in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

8. Appendix

Table A1: Attitudes towards females

| | Dependent variable | |
|------------|------------------------|------------------------|
| | Family first (1) | Marriage first (2) |
| Rural | 0.292*** (0.0164) | 0.127*** (0.0182) |
| Male | 0.0578*** (0.0186) | -0.0143 (0.0185) |
| RuralXmale | -0.0743*** (0.0242) | -0.0619*** (0.0270) |
| Constant | 0.519*** (0.0131) | 0.426*** (0.0130) |

Notes: *Family first* is an indicator for whether the respondent agrees that men should prioritise career and women family. *Marriage first* is an indicator whether the respondent agrees that for females marrying a good husband is more useful than having a good job. *Rural* is an indicator for being born and having stayed in the rural region. Robust standard errors in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

Table A2: Summary statistics of individual characteristics: to continue
by *tongxiang* concentration and gender

| | Non-treated average <30% <i>tongxiang</i> | Treated average >30% <i>tongxiang</i> | Difference |
|---|--|--|-------------------|
| <i>Panel A: Age</i> | | | |
| Female | 28.45 (8.70) | 30.69 (10.38) | 2.25 (0.83) |
| Male | 31.35 (9.41) | 34.36 (10.98) | 3.02 (0.70) |
| Difference | 2.90 (0.31) | 3.67 (1.04) | 0.77 (1.08) |
| <i>Panel B: Migration age</i> | | | |
| Female | 20.70 (6.11) | 19.65 (6.57) | -1.05 (0.70) |
| Male | 21.12 (6.93) | 21.50 (7.30) | 0.38 (0.60) |
| Difference | 0.42 (0.27) | 1.85 (0.89) | 1.43 (0.92) |
| <i>Panel C: Years since migration</i> | | | |
| Female | 7.65 (5.68) | 9.19 (8.00) | 1.54 (0.84) |
| Male | 9.97 (7.31) | 12.16 (7.80) | 2.20 (0.64) |
| Difference | 2.32 (0.27) | 2.97 (1.03) | 0.65 (1.06) |
| <i>Panel D: Distance from home (kilometres)</i> | | | |
| Female | 602.68 (443.28) | 518.91 (438.45) | -83.77 (36.20) |
| Male | 637.38 (449.88) | 544.20 (413.72) | -93.18 (27.95) |
| Difference | 34.70 (15.70) | 25.30 (43.03) | -9.41 (45.73) |
| <i>Panel E: Education</i> | | | |
| Female | 1.30 (0.90) | 1.05 (0.85) | -0.26 (0.07) |
| Male | 1.45 (0.85) | 1.12 (0.77) | -0.33 (0.05) |
| Difference | 0.15 (0.03) | 0.08 (0.08) | -0.07 (0.09) |

Table A2: Summary statistics of individual characteristics: continued
by *tongxiang* concentration and gender

| | Non-treated average <30% <i>tongxiang</i> | Treated average >30% <i>tongxiang</i> | Difference |
|-------------------------------------|--|--|-----------------|
| <i>Panel F: Weekly working time</i> | | | |
| Female | 55.34 (15.01) | 59.42 (16.68) | 4.08 (1.36) |
| Male | 56.29 (14.70) | 59.92 (17.55) | 3.63 (1.13) |
| Difference | 0.95 (0.51) | 0.50 (1.70) | -0.45 (1.77) |
| <i>Panel G: Log hourly wage</i> | | | |
| Female | 2.00 (0.45) | 1.90 (0.62) | -0.10 (0.05) |
| Male | 2.23 (0.52) | 2.21 (0.52) | -0.02 (0.03) |
| Difference | 0.23 (0.02) | 0.31 (0.06) | 0.08 (0.06) |

Notes: Treated individuals have more than 30% of their co-workers from the same county while non-treated individuals have less than 30% . In *Panel E*, *Education* is a categorical variable that increases with the level of education (0: less than or equal to primary school; 1: junior high school; 2: senior high school; 3: more than high school). Standard deviations in parentheses.

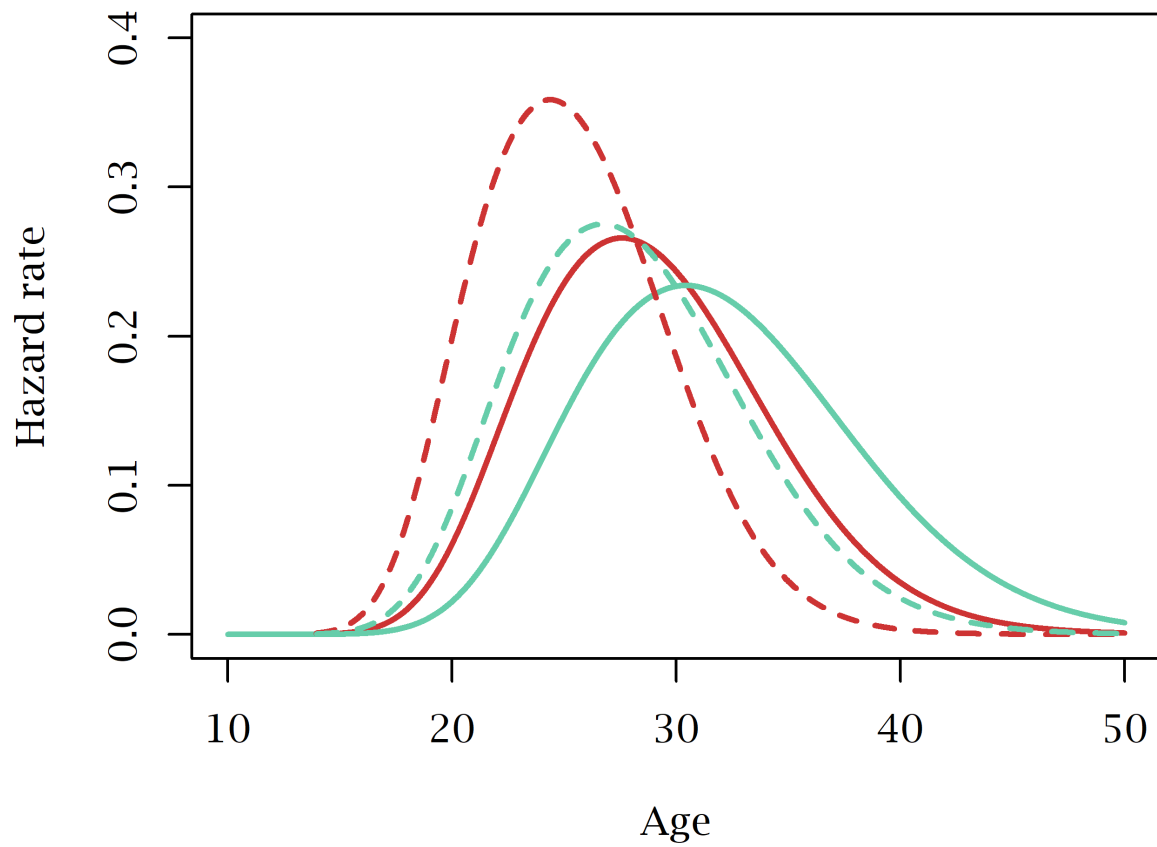


Figure A1: Marriage hazard rate in rural and urban China

— : urban female - - - : rural female — : urban male - - - : rural male

Notes: Based on the 2010 Chinese General Social Survey. *rural* is defined as individuals who are born in rural regions and have never left. *urban* is defined as individuals who are born with urban *hukou*.

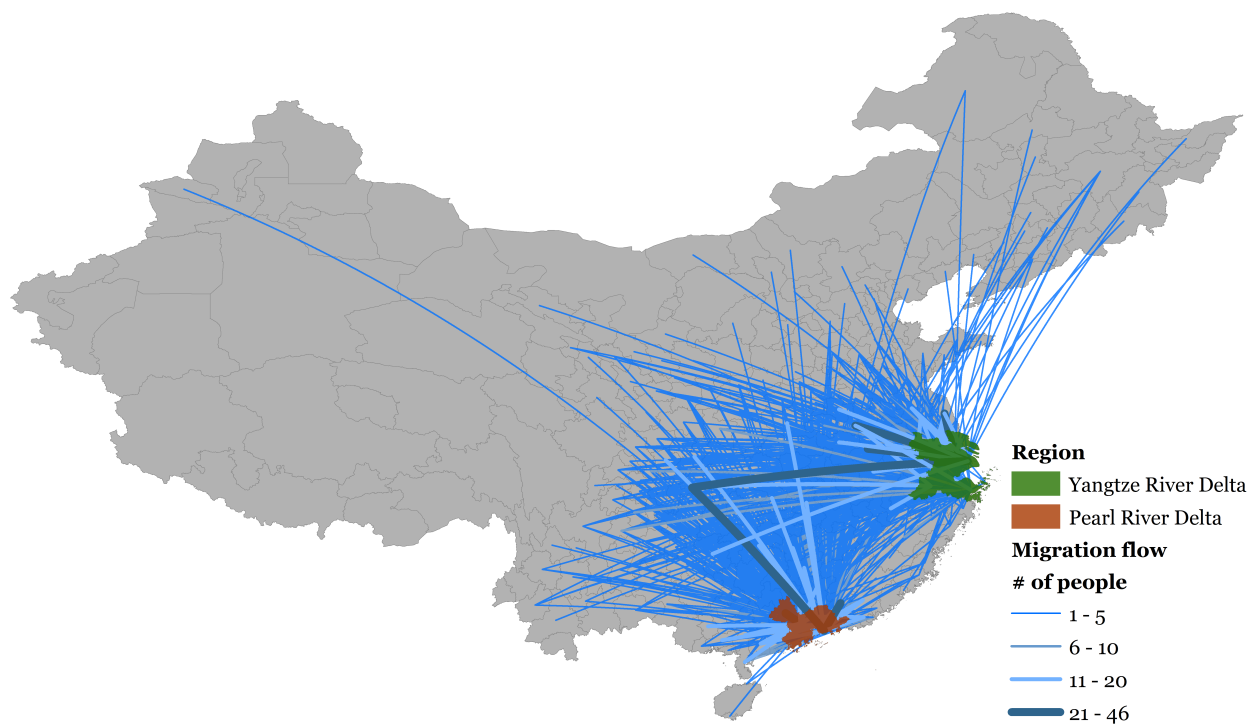


Figure A2: Migration flows to Yangtze and Pearl River Deltas

Notes: Migration flows to the Yangtze and Pearl River Deltas based on the 2010 survey data on migrant workers.

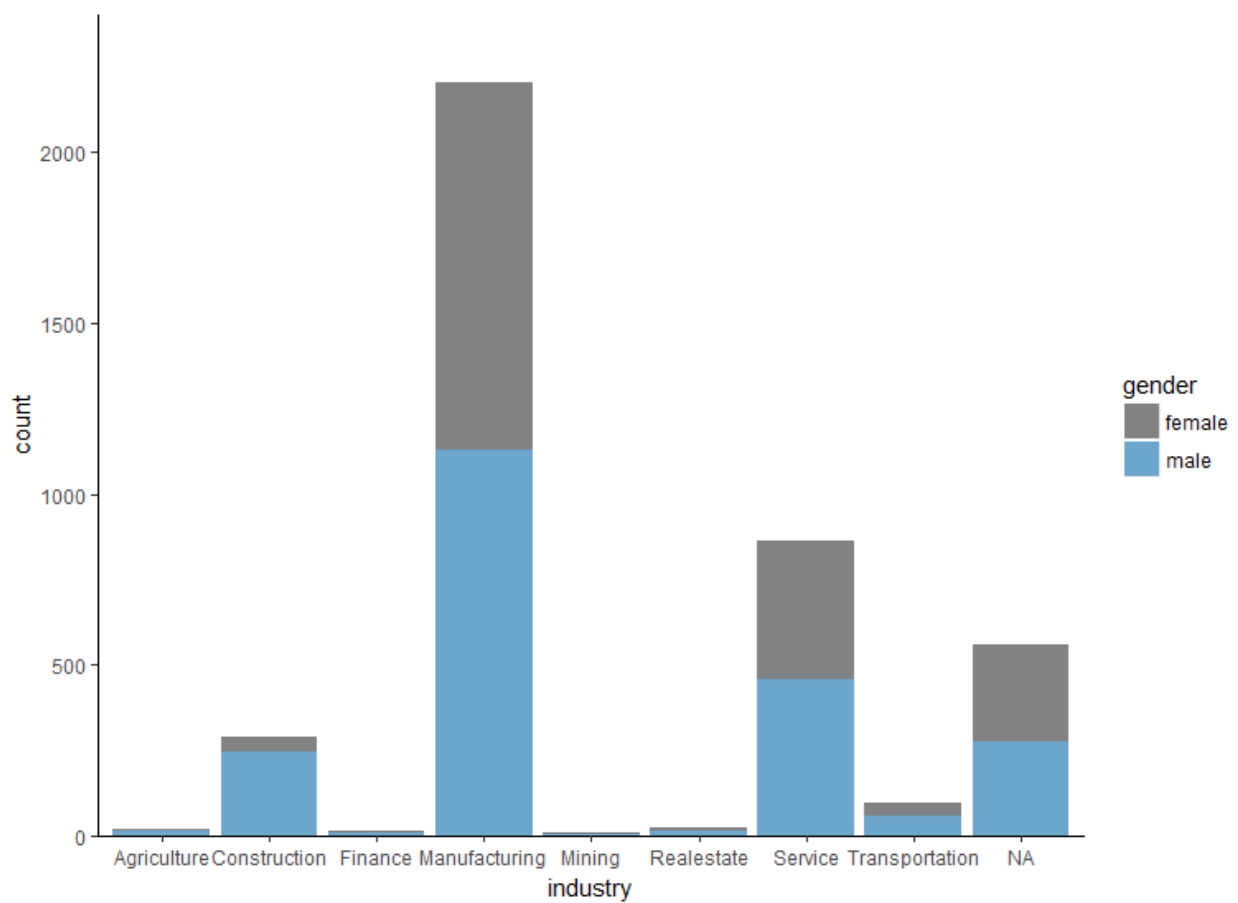
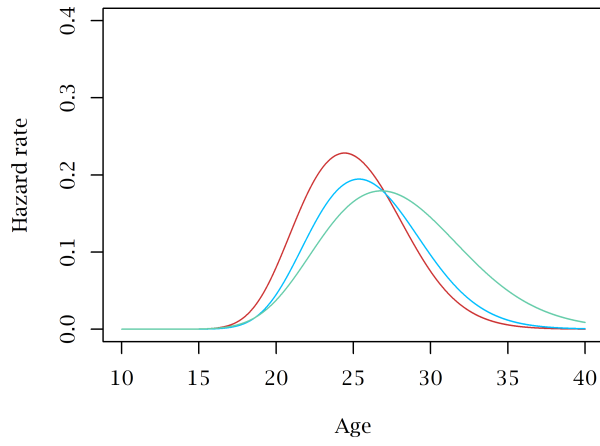
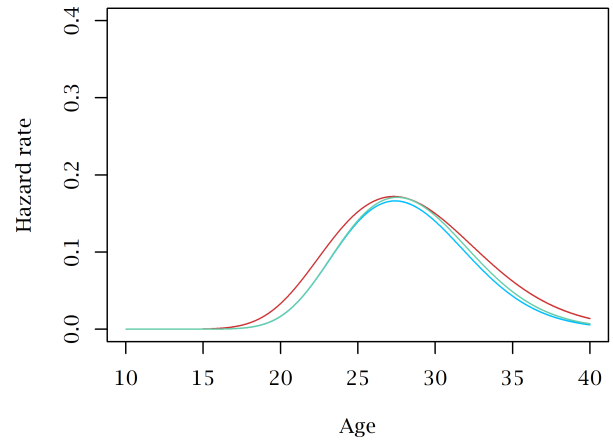


Figure A4: Industrial employment by gender

Notes: Based on the 2010 survey data on migrant workers.



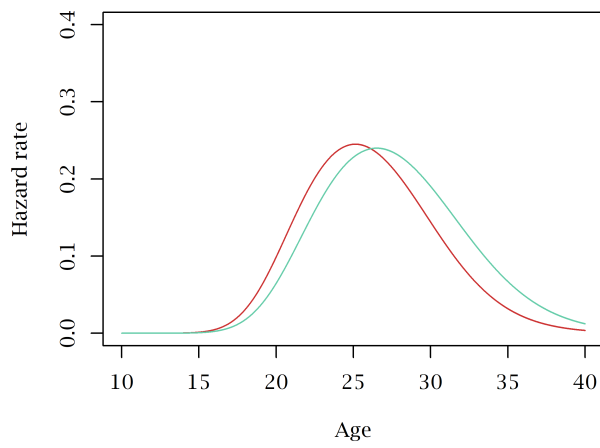
(a) Female



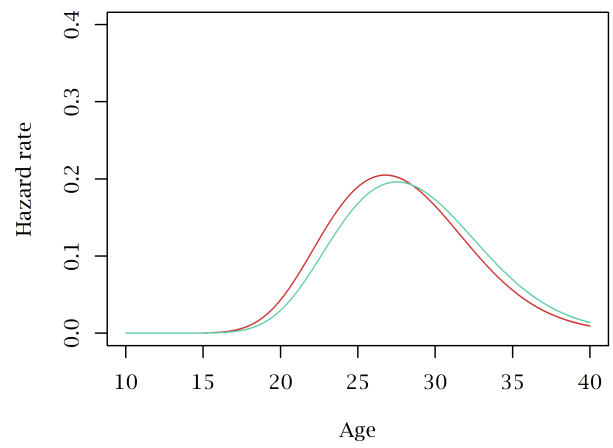
(b) Male

Figure A5: Estimated hazard rate: pc_town

green: none of co-workers come from the same town
blue: 0–30% of co-workers come from the same town
red : >30% of co-workers come from the same town



(a) Female

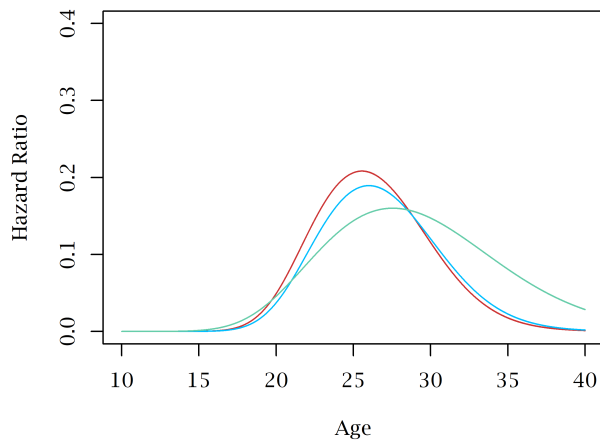


(b) Male

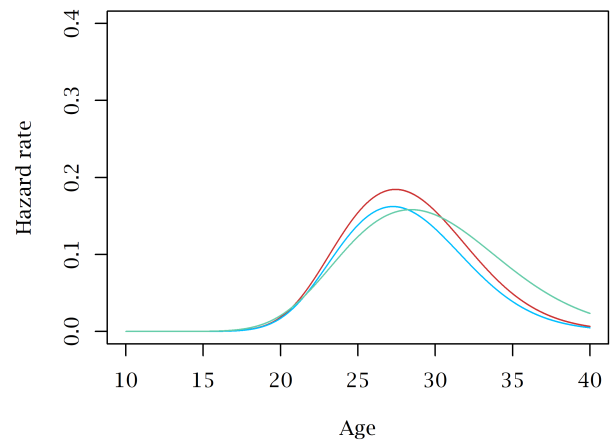
Figure A6: Estimated hazard rate: friend2

red: 2nd best friend is *tongxiang*
green: 2nd best friend is not *tongxiang*

Notes: *Tongxiang*, or people from the same hometown, self-defined by the survey respondents.



(a) Female



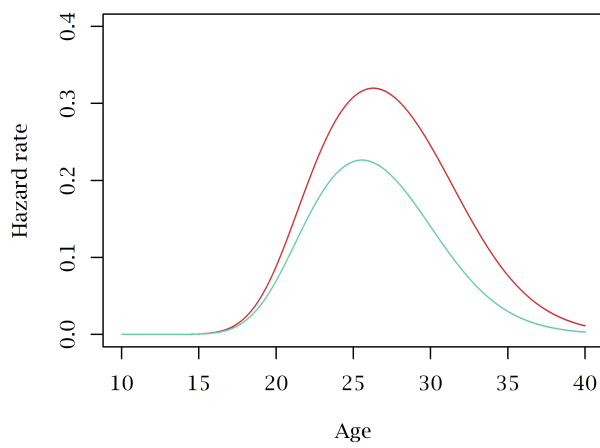
(b) Male

Figure A7: Estimated hazard rate: pc_province

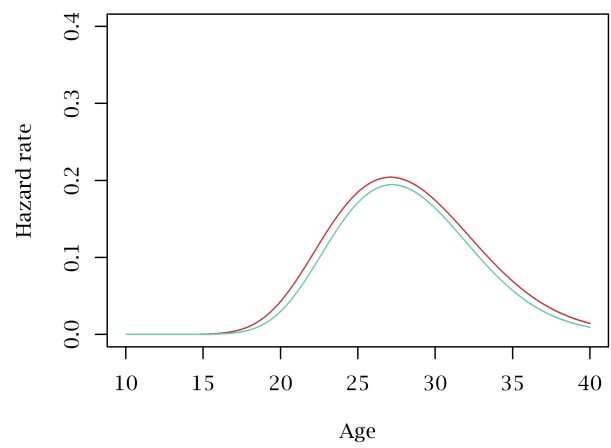
green: none of co-workers come from the same province

blue: 0–30% of co-workers come from the same province

red : >30% of co-workers come from the same province



(a) Female



(b) Male

Figure A8: Estimated hazard rate: friend3

red: 3rd best friend is *tongxiang*

green: 3rd best friend is not *tongxiang*

Notes: *Tongxiang*, or people from the same hometown, self-defined by the survey respondents.

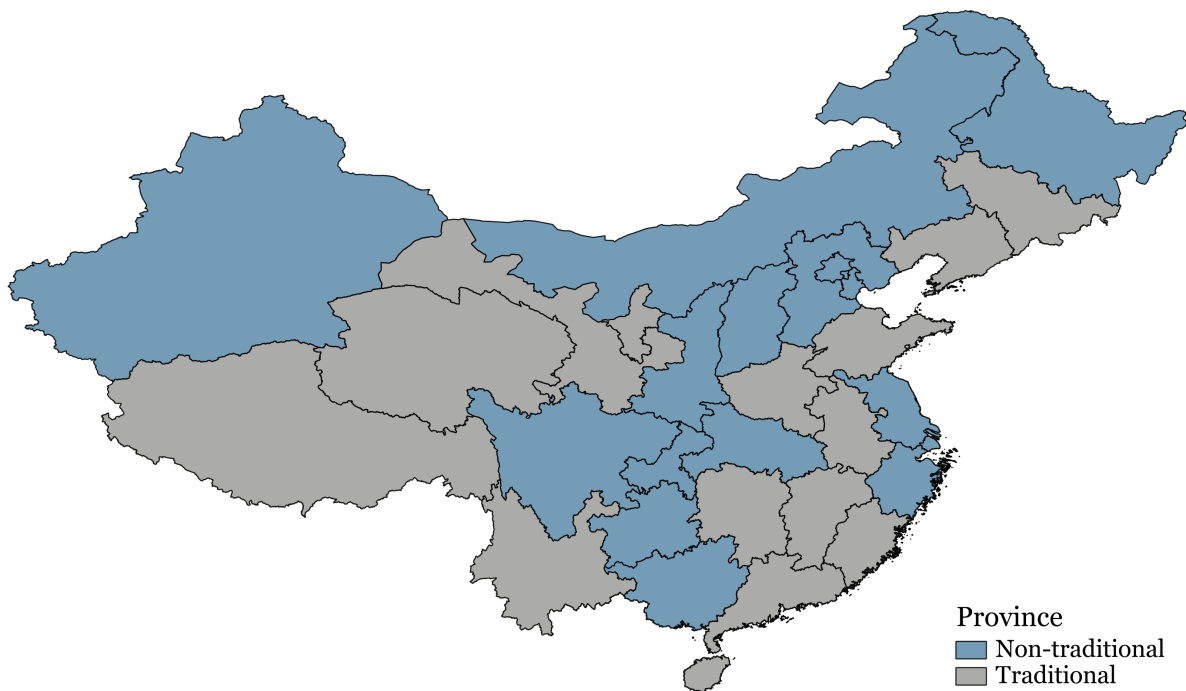


Figure A9: Provinces in China by "Traditionalness"

Notes: "Traditionalness" is defined by attitudes towards the role of female in the society, calculated from the responses to the 2010 China General Social Survey. First the median value of the fraction of rural individuals in each province that agree to "marrying a good husband is more important than having a good job for females" is computed. Provinces above the median is defined as traditional provinces and provinces below the median non-traditional provinces.

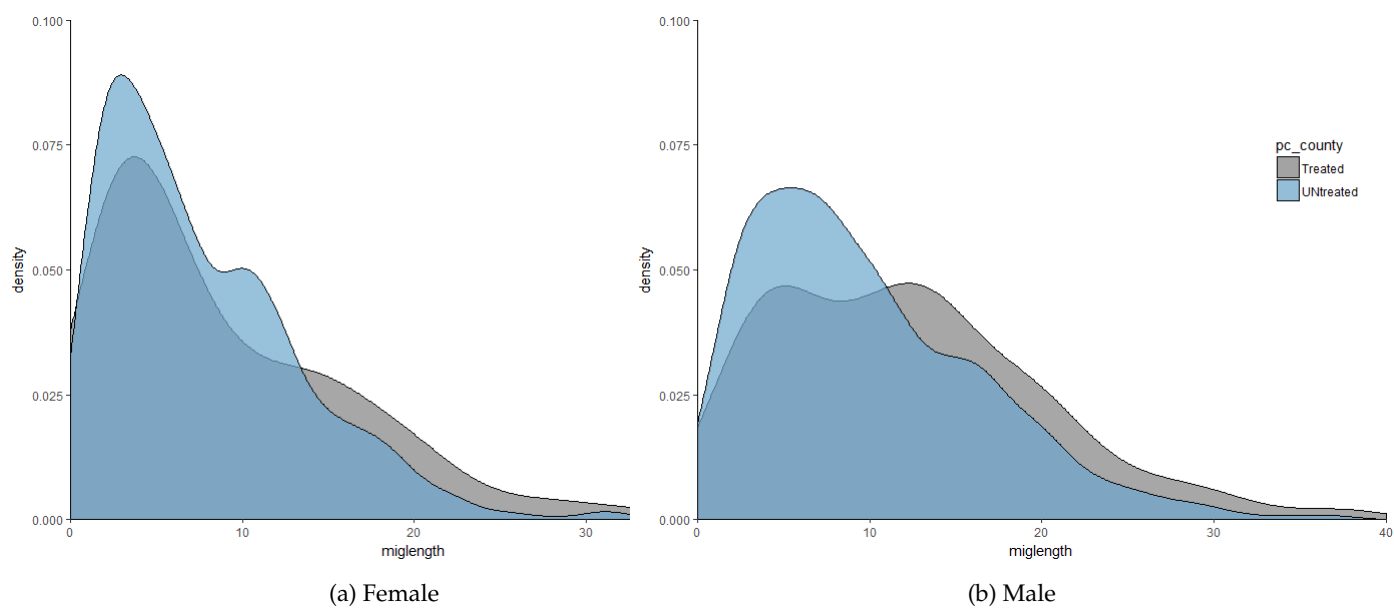


Figure A10: Years since migration by social interactions

Notes: Based on the 2010 survey on migrant workers. Treated individuals are ones with more than 30% co-workers from the same county. The untreated have less than 30% co-workers from the same county.