Rushed to the altar:
The effect of social interactions on migrant workers' marriage age

Yan Hu[*#]
CEMFI

13 March 2021

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. Combining placebo test and propensity score matching, 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 Joan Monras, Andreas Stegmann and seminar participants at CEMFI, Econometrics Society Asian Meeting, Meeting of the 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 Competividad (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 novel 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 causal 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) compel female migrants to abide by norms from their hometown, raising their probability of getting married before age 23 from 0.29 to 0.59.[1] Conditional on getting married before 40, women with the majority of co-workers from the same hometown are estimated to marry 2.6 years earlier than ones without any same-origin co-workers. However, social interactions have little impact on the marriage age of male migrants.

To further address selection bias in social interactions, I use propensity score matching to control for selection on observables. I allow selection to vary by gender and compare individuals who differ only in the composition of their co-workers but are otherwise equal, and find that the resulting 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 social interactions have no real impact on marriage behaviour, but instead capture some private preferences or tendencies for early marriage, we would expect correlation between early marriage and social interactions present even for females who are already married before entry into workplace with prevalence of *tongxiang* co-workers. Reassuringly, 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 by increasing the chance of matching between individuals who are

[1] To 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). Intensive social interactions between co-workers of the same rural origin increase the fraction of married female migrants before age 26 from 0.58 to 0.84.

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.

In line with the norm-based explanation, I find that the 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 women (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 ceases to exist. However, behaviour may fail to adjust if individuals adhere to old norms, 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 more 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 provides evidence on a novel channel of cultural persistence through social interactions.[4] 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 the pre-existing norms. As a result, behaviour (i.e. early female marriage) perpetuates. Recent research by Myong, Park and Yi (2020) use a model to carefully quantify the role of social norms in marriage. They estimate a single parameter of social norms (along with a rich set of other behavioural and preference parameters) to rationalise observed patterns of marriage and fertility. The advantage of the unique setting of rural-to-urban migrant workers in China allows me to use individual variations in strength of norms (from concentration of co-workers from the same origin) to identify the effect of social norms on marriage. The finding confirms the model estimates that social norms can explain substantial part of marriage behaviour.

By emphasising the role co-workers play in shaping one's behaviour, this paper also connects to the literature on peer effects and social pressure (e.g. Mas and Moretti, 2009; Bandiera, Barankay

[2]Females 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.

[4]Giuliano and Nunn (2020) find that cultural persistence and change depends on similarity of the environment across generations. Bau (Forthcoming) shows that policies can change cultural persistence.

and Rasul, 2010; Burke and Sass, 2013; Cornelissen, Dustmann and Schönberg, 2017). I focus on a particular type of peers, that is, co-workers from the same place of origin. Hence, this paper relates to studies on ethnic enclaves, which previously have predominantly concern their pecuniary benefit (e.g. Borjas, 1995; Edin, Fredriksson and Åslund, 2003; Zhang and Xie, 2013), and here I explore the social aspects of 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 a new source of gender inequality. Marriage age affects motherhood timing. Late marriage and childbearing is associated with a substantial increase in lifetime income for females (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. Policy interventions that aim at closing gender gap, such as providing equal access to education and jobs, may not be effective in the presence of strong social pressure for females to conform to the rural norm of early marriage.

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 methods, results and the mechanism. Section 5 rules out alternative explanations. Section 6 concludes.

2. Institutional background

In this section, I give an overview of the causes of concentration of rural-to-urban migrant workers based on place of origin in destination cities and its consequences.

The economic reform of China since 1978 set the momentum of a large-scale migration from rural to urban areas, which were 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 the Pearl River and Yangtze River Deltas.

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 classified as either rural or urban residents, which is used to link local public welfare programs. The conversion to local *hukou* remains extremely difficult for rural-to-urban migrants (Song, 2014). Lacking *hukou* status in their cities of residence, migrant workers have limited access to local public goods and therefore have less bargaining power in the labour market (Song, 2014).[5]

Furthermore, due to the pre-existing discrepancy in the levels 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 2010 survey on migrant workers, more than 50% of the migrant workers were once discriminated by local residents and about 60% lacked a sense of belongingness in the cities.

[5]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 (Liu, 2005; Deng, 2007; Gravemeyer, Gries and Xue, 2010; Lee, 2012).

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 implemented before 1978, the isolation created large socio-economic barriers (e.g. language, customs, income) for people from different parts of China to interact. Therefore migrant workers sometimes stay closely with their *tongxiang*, the Chinese word for people from the same hometown, to share network, information and resources.

Rural-to-urban 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 motive to join *tongxiang* network is usually 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 the main concern of migrant workers. Marriage can be affected by social interactions because they can increase the chance of matching with a partner. Additionally, marriage behaviour can be regulated by social norms and therefore subject to peer pressure. This paper finds evidence for the latter channel through which social interactions affect marriage, which is 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-to-urban migrant workers in China in 2010.[6] It took a representative sample of migrant workers working in the Yangtze and Pearl River Deltas in 2010.[7] 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 wave of the survey, questions regarding the relations with *tongxiang*, or people from the same hometown, 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, county, and town, which are respectively 1st, 3rd and 4th level administrative units in China. Additionally, we know whether the three best friends of the migrant worker are *tongxiang*, as identified by the respondents. In the survey, individuals were asked:

[6] The survey was part of a project sponsored by the Ministry of Education in China with the aim to study the status-quo of migrant workers and protect their rights.

[7] According to the Chinese National Bureau of Statistics, the two regions have the highest concentration of migrant workers in China and together assimilated more than 40% of total migrant workers in 2011. Figure A2 shows the migration outflows to the Pearl and the Yangtze River Deltas based on the survey data.

Q1 In the production line that you work, what is the percentage of people from the same town/county/province?
(1) None, (2) Very few (<10%), (3) A few (10-20%), (3) Some (20-30%), (4) Many (30-50%), (5) A lot ($>$ 50%), (6) Do not know.

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

Concentration of *tongxiang* co-workers is my preferred measure of social interactions because it is a less endogenous choice than friendship. Concentration of co-workers based on the place of origin can result from close distance from the origin to the destination city, a large out-migration population from the origin, or employers' preferential hiring of *tongxiang*, which can be orthogonal to marriage age decisions.

Table 1: Proportion of *tongxiang* co-workers

	None	Very few <10%	A few 10%-20%	Some 20%-30%	Many 30%-50%	A lot >50%
Town	0.483	0.265	0.091	0.072	0.040	0.050
County	0.412	0.265	0.117	0.091	0.053	0.061
Province	0.140	0.201	0.137	0.137	0.153	0.233

Notes: Based on the 2010 survey on migrant workers. The table shows the levels of concentration of co-workers (in columns) from the same town, county and province (in rows). The numbers are rounded to three decimal places.

Moreover, concentration of co-workers is a good proxy for social interactions because migrant workers spend a substantial amount of time with their co-workers. According to the same survey, 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 socialise (Fang, 2012). Besides, 36% of total migrant workers live in the dormitory provided by their employers and the ratio increases to 49% for individuals who are 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, respondents were asked to indicate whether three of their best friends are from the same hometown. The data show that a considerable fraction of people have *tongxiang* as their best friends. The fraction of individuals whose first best friend is *tongxiang* is 0.37. If we consider three best friends, for 13% individuals, all their best friends are *tongxiang* and for 44% individuals, at least one of their best friends is *tongxiang*.

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 affairs. Therefore I will base my main analysis on measures of *tongxiang* co-workers to proxy for social pressure to

conform to rural norms and present estimation results 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 migrated at an age of 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 services with an almost equal number of male and female workers. There are significantly more males in construction than females although the total employment is only 308, which also makes up for the difference between the total number of females and males in the sample. Transportation assimilates the 4th largest employment with a balanced gender representation. The other sectors are relatively small and assimilate about 2% of workers (see Figure A3).

For the main analysis in Section 4, I show that the probability of early marriage depends on how connected one remains with her rural community for female migrant workers, but not for male migrants. One concern is that female migrant workers that sort into social interactions with *tongxiang* co-workers have certain characteristics that make them marry early. Therefore, in Table A2 I present summary statistics of a wide range of individual characteristics by gender and concentration of co-workers from the same county (i.e. more than 30% *tongxiang* co-workers and less than 30% *tongxiang* co-workers).[8] The results show that although individuals with more *tongxiang* co-workers are different in some dimensions from ones with less *tongxiang* co-workers on average, they differ in the same fashion for female and male migrants. That is to say, sorting into *tongxiang* enterprises on observables, if any, is not that different between male and female migrants, therefore can not explain the gender asymmetric effects of social interactions on marriage age. In Section 5, I use propensity score matching to correct for the selection on observables, which confirms the patterns found in Table A2.

C *Complementary data*

Additionally, I use the 2010 Chinese General Social Survey which enquires about individual attitudes towards women to proxy for the differences in marriage norms in different parts of rural China. If social norms pressure females into early marriage in the presence of *tongxiang* co-workers, the effect would be amplified by the strength of the norms.

4. Estimation method, result and explanation

In this section, I first present the estimation methods and results. I use discrete-time duration model to examine the effect of social interactions on the distribution of marriage age and quantify the

[8] These include age, age at migration, years since migration, distance to the destination city from hometown, education, working hours and wage.

effect by translating the hazard rate to cumulative distribution function. I then proceed to provide evidence for a potential channel that can explain the association between social interactions and early marriage.

A *Estimation of hazard rate*

The paper examines the effect of social interactions on the probability of getting married at different ages. Alternatively we can estimate the effect of social interactions by comparing average marriage ages, but this comparison would overlook individuals that have not been married in the sample. The fact that one has not been married and at the same time has only a few co-workers of the same rural origin is also informative about the effect of social interactions on marriage age. Using a duration framework, which models probability of getting married conditional on not having been married before, allows us to incorporate this information, as well as to look closely at the effect of social interactions on the whole distribution of the marriage age, in additional to the mean. I follow the specification of the discrete-time duration model by Bover, Arellano and Bentolila (2002) and define the hazard rate of marriage at age t as:

$$h(t) = Pr(T_i = t | T_i \geq t) = G(\gamma_0 + \gamma_1(lnt) + \gamma_2(lnt)^2) \,, \tag{1}$$

where

$$G(x) = \frac{exp(x)}{1 + exp(x)} \,. \tag{2}$$

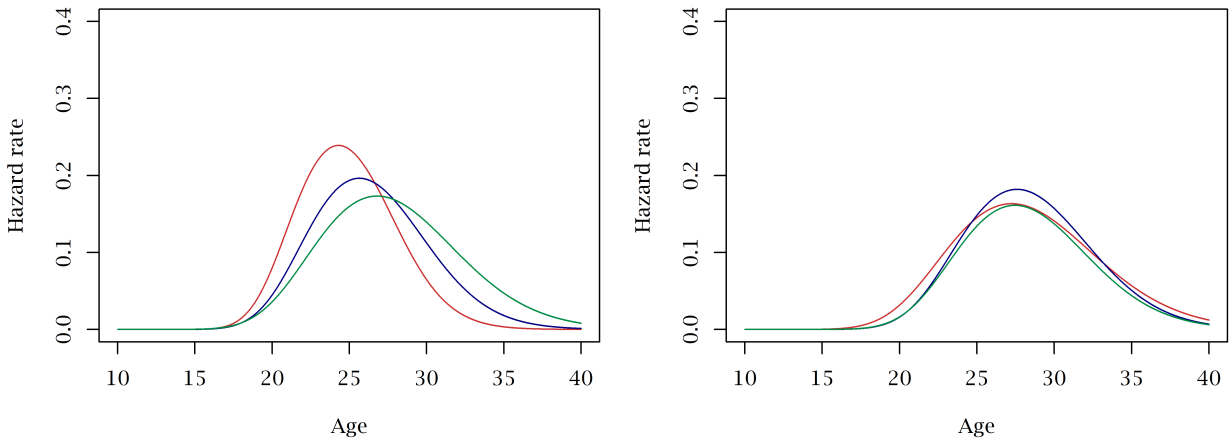
The duration here is years to marriage. The hazard rate $h(t)$ is the conditional probability of getting married at age t given one has not been married before age t. The specification here allows the hazard rate to vary with age t.

I look at marriage hazard rates at different ages and also examine how they vary with concentration of co-workers from the same hometown, separately for male and female migrant workers. Panel a of Figure 1 shows the estimated hazard rates for female migrant workers with different levels of concentration of co-workers from the same county. As the fraction of *tongxiang* co-workers increases, we see a gradual shift and intensification of exit rate at earlier ages. This indicates that female migrant workers are more likely to marry early when they have a higher fraction of co-workers from the same county in the production line. In contrast, concentration of *tongxiang* co-workers does not affect marriage age of male migrant workers, as shown in the right panel of Figure 1.

We obtain very similar gender-differentiated patterns by looking at co-workers from the same town and the results are shown in Appendix (Figure A4). The effect of co-workers from the same province is less clear (Figure A6). Given that 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 province.[9] Individuals stop to identify people from the same province as *tongxiang*.

Next we turn to estimation results using self-perceived friendship as a proxy for social interactions with *tongxiang* (Figure 2). Similar patterns emerge. For females, if her best friend comes

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



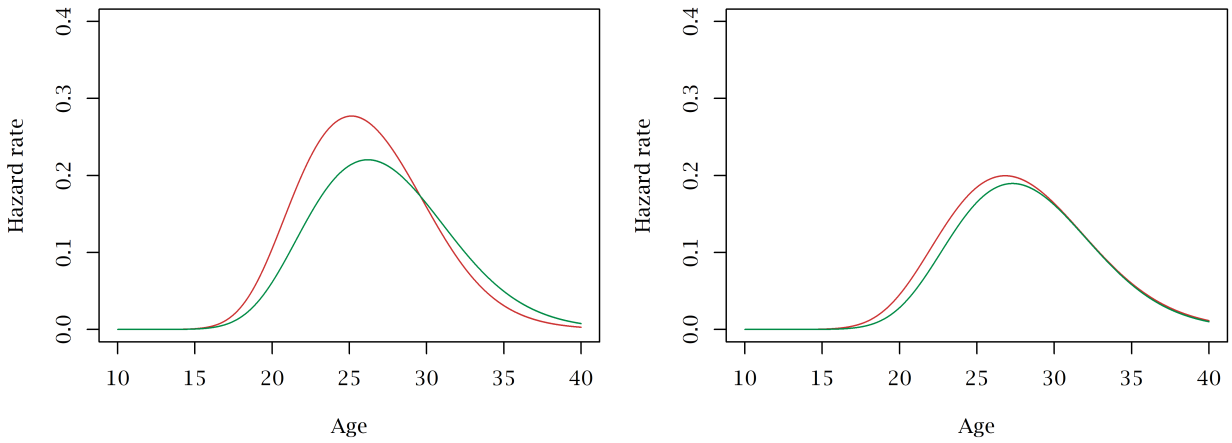
(a) Female (b) Male

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

from the same hometown, 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 A5, A7). Although the friendship measure is more subject to selection bias, the same gender differentiated patterns are consistent with the story that females are pressured into early marriage with social interactions with *tongxiang* to conform to the rural norm of early female marriage.

B *Estimation of cumulative distribution function*

The estimated hazard rates provide intuitive overview for the effect of social interactions on marriage age. However, they do not give us quantitatively the magnitude of the effect and its significance level. Furthermore, comparing directly the difference in hazard rates of marriage for any given age and interpret it as the treatment effect of social interaction can be problematic. At each point of time, as more marriageable individuals exit (i.e. get married), and more individuals exit in the group with higher concentration of *tongxiang* co-workers, the compositions of the remaining population change differently for groups with different fractions of *tongxiang* co-workers. This can be seen from the fattening out of the right-hand tail of the hazard rate for the group of female migrants with a high level of concentration of co-workers from the same county. As time goes by, people who remained unmarried in the group with more *tongxiang* co-workers are on average less marriageable than people in the other groups. The difference in hazard rates is a combination of the treatment effect of social interactions and compositional differences between different groups, and is likely to underestimate the treatment effect of social interactions as the treated group (with a higher fraction of *tongxiang* co-workers) overtime has a less marriageable population.



(a) Female (b) Male

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-identified by the survey respondents.

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

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

$F(t)$ is the fraction of individuals that get married or the probability that an individual gets married before age t. We can interpret the difference in this probability between individuals with more or less *tongxiang* co-workers as the treatment effect of social interactions.

Table 2 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, restricting the sample to individuals who get married after migration. For visibility, the table only presents the probability of getting married before age 20 to 30.[10] In column 1 and 3, I show the probability of getting married before a given age for the baseline group for female and male migrants respectively. The baseline group is migrant workers with none of the co-workers from the same county. The differences in column 2 and 4 give us the effect of social interactions, which is the increase relative to the baseline group in the probability of getting married for migrant workers with the majority of their co-workers from the same county. Bootstrap standard errors with 500 replicates are shown in parenthesis.[11]

[10]The estimation can be done for all ages starting from 16, which is set to be the minimum marriage age.

[11]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.

For a female migrant who does not have any *tongxiang* co-workers, she is expected to get married by 23 with a probability of 29% (column 1). If instead the majority of her co-workers in the production line come from the same county, the chance of being married by 23 more than doubles, increased by 30 percentage points (column 2). On the contrary, the effects on male workers are much smaller and remain insignificant for most ages (column 4).

Table 2: Differences in marriage between 0% and 50%+ *tongxiang* co-workers for individuals married after migration

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

Notes: Estimation is based on the sample of individuals married after migration. The baseline group is migrants with 0 *tongxiang* co-workers, or co-workers from the same county. The difference is the increase in probability relative to the baseline group for migrants whose majority of co-workers are *tongxiang*. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

As a placebo test, we can make the same calculation for individuals who have already been married before migration. If social interactions have no real impact on marriage behaviour, but instead capture some private preferences or tendencies for early marriage, we would expect correlation between early marriage and social interactions present for females who are already married before entry into workplace with prevalence of *tongxiang* co-workers. In Table 3, I estimate the treatment effect of social interactions with current *tongxiang* co-workers but restrict the sample to individuals who are already married before migration into city. Now the effect of social interactions disappears for females as well (column 2). Not only does the significance level fall to zero but

also the point estimates. This means future co-worker composition does not predict past marriage, indicating that social interactions do not pick up some unobservable individual characteristics that at the same time lead to early marriage.

The results in Table 2 and placebo test in Table 3 demonstrate that concentration of co-workers from the same rural origin is not only correlated but causes female early marriage. The next part investigates the underlying mechanism.

Table 3: Differences in marriage between 0% and 50%+ *tongxiang* co-workers for individuals married before migration (placebo test)

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

Notes: Estimation is based on the sample of individuals married before migration. The baseline group is migrants with 0 *tongxiang* co-workers, or co-workers from the same county. The difference is the increase in probability relative to the baseline group for migrants whose majority of co-workers are *tongxiang*. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

C *The effect of social interactions through cultural norm*

Anecdotal evidence suggests that female migrant workers in the workplace face considerable social pressure to get married from social interactions with their co-workers. 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 bolster the idea that the effect of social interactions on marriage works through social pressure to conform to the rural norm of early female marriage. First, I show a gender-differentiated reaction to rural norms in marriage behaviour. Female migrants who identify more with their rural origin are more likely to get married early. For male migrants, the association between rural identity and early marriage is much weaker. This indicates that the rural norm of early marriage, proxied by self-identification with rural origin, is more pronounced for females than for males.

While individuals who identify with their rural origin internalise rural norms and behave accordingly, social interactions can place pressure to conform to rural norms even if the individual does not necessarily identify with those norms. In addition to linking norms with marriage, the second set of evidence links social pressure to conform to norms with marriage. If concentration of co-workers exert pressure to conform to norms, then the effect will be larger if the social pressure is larger (i.e. higher concentration of *tongxiang* co-workers) and if the norm is stronger (i.e. more traditional values on women). I show that the gender differentiated response to the presence of *tongxiang* co-workers is stronger for individuals originated from places that hold more conservative attitudes toward women's role in the society.

a Farmer v.s. worker identity

If we believe cultural norms in agricultural societies make females marry early, we would expect that females who identify more with their rural origin are more likely to abide by the rural norms and therefore marry early as they could internalise rural norms, 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 same 2010 survey of migrant workers, individuals were asked:

Q2 Which do you think is your identity?
(1) Farmer, (2) Worker, (3) Others, (4) Do not know.

Table 4 calculates the probability of getting married before a certain age by self-identity, separately for males and females, and only for individuals who are married after migration. The baseline groups are migrant workers who identify themselves as workers for females in column 1 and for males in column 3. I show the increase relative to the baseline in the probability of getting married for individuals who identify herself as a farmer for females in column 2 and for males in column 4 . Female migrants who recognise herself as a farmer are more likely to get married early (column 2). For males, the association between rural origin and marriage age is much weaker (column 4). Self perceived identity does not seem to be 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* co-workers. As a result, the correlation between prevalence of *tongxiang* co-workers and early marriage would reflect instead the effect of self-identity on marriage behaviour. In Table A4, I regress concentration of *tongxiang* co-workers on an indicator for self-identification with farmer and find no correlation between the two. Another interpretation of the result 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.

Table 4: Differences in marriage between worker and farmer identity for individuals married after migration

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

Notes: Estimation is based on the sample of individuals married after migration. The baseline group is individuals who identify themselves as workers. The difference is the increase in probability relative to the baseline group 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.

I find that for individuals who marry before migration, females who identify with their rural origin also marry early but males do not (Table A3). This is consistent with the idea that self-identity is stable over time and is associated with a set of values and norms individuals internalise to guide

their behaviour. Female migrants who identify with their rural origin value early marriage, which is the norm in agricultural societies, and accordingly marry early. For males, as the norm of early marriage is not as strong, self-identification with rural origin does not translate into early marriage.

b Traditional v.s. non-traditional provinces

If rural norms pressure females into early marriage in the presence of *tongxiang* co-workers, we would expect a larger effect of social pressure if the norm is stronger. The strength of the norm can be proxied by views regarding the role of females in the agricultural society. I use the response in the 2010 China General Social Survey to the question:

Q3 Do you agree that for females, it is more useful to have a good husband than a good career? (1) Completely disagree, (2) Disagree, (3) Indifferent, (4) Agree, (5) Strongly agree, (6) Do not know.

I divide provinces into two equal groups and code them as traditional and non-traditional provinces depending on the fraction of rural respondents who agree or strongly agree with the above statement (high fraction defined as traditional province, see Figure A8 for the province classification). I classify traditional provinces along the dimension of this question in the general social survey because it solicits value of marriage for females, which is exactly the outcome the paper studies. I compare the difference in the gender-differential of the effect of social interactions on marriage age between individuals from the traditional and non-traditional provinces. If we index social interactions by s, ns, gender by F, M, and "traditionalness" by t, nt, we can write the triple difference of the effect of social interactions on the probability of getting married before age t:

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

As shown in column 1 of Table 5, the gender differential is larger for individuals from provinces with more conservative values for females. Females are more likely to get married early with concentration of *tongxiang* co-workers than male migrants (the differences are shown in column 2 and 4 of Table 2), yet there is heterogeneity in the gender asymmetric response to concentration of *tongxiang* co-workers depending on the strength of the norm. In particular, the gender differential of the probability of getting married before 24 is 39 percentage points higher in traditional provinces than in non-traditional provinces. The result is consistent with the idea that females from areas with more traditional norms on women's role in the society face a higher social pressure of getting married early with concentration of *tongxiang* co-workers than their counterparts from less traditional provinces.

5. Alternative explanations

In this section, I rule out alternative mechanisms that can generate the gender differentiated effect of social interactions.

A *Matching*

Migrant workers originated from the same rural area share customs, languages and similar socioeconomic conditions. Concentration of *tongxiang* co-workers can facilitate meeting fellow countrymen that are more similar to the individual, therefore increasing the chance of meeting a potential partner and a quick transition into marriage.

Table 5: Gender differentiated effects of social interactions on marriage

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

Notes: Column 1 compares the gender differential in the probability of getting married before a given age between individuals from traditional and non-traditional provinces. Traditional provinces are defined as ones that have a large fraction of rural residents who agree that for females it is more useful to have a good husband than a good career (Q3). Column 2 compares the gender differential between individuals in cities with a higher origin-destination specific male to female migrant ratio and a lower ratio, imputed from the 2000 population census. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

Although we do not know in the data the identity of the spouse, there are some testable implications of the matching story. In Table 6, I examine whether the spouse of the migrant worker is more likely to be in the same workplace if the workplace has a higher fraction of *tongxiang*

co-workers. Notice that the measure of *tongxiang* co-workers for individuals is (only available) for their current job. Therefore I divide the sample into individuals who are (1) married before migration, (2) married after migration but before the current job, and (3) married after the current job. I predict whether the spouse is in the same workplace using a linear probability model. *Pc_county* is a categorical variable whose value increases with the fraction of co-workers from the same county. Column 1 and 2 show that for both females and males who are married before migration, their spouse is more likely to be in the same workplace if the fraction of tongxiang co-workers is high. For individuals who marry after migration and especially ones who marry after starting the current job, only husbands of the female migrant workers are more likely to be in the same workplace with higher fraction of co-workers from the same county (column 3 and column 5).

This results suggest that (a) individuals who are married (possibly from the same county) are more likely to later enter a workplace with concentration of their *tongxiang*; and (b) Unmarried females who enter in a workplace with a higher fraction of *tongxiang* co-workers, are later more likely to get married with someone in the workplace. If matchmaking of individuals who come from the same county takes place in the workplace, male migrants who are unmarried before the current job are also more likely to find his spouse in the current workplace from the pool of *tongxiang* co-workers, but the small and insignificant coefficient on pc_county in column 6 indicates that this is not the case. The gender asymmetric patterns here are consistent with the idea that 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 *tongxiang* or not.

Table 6: Correlation between *tongxiang* co-workers and spouse workplace

	Dep. var.: 𝟙(Spouse in the same workplace)					
	Married before migration		Married after migration before current job		Married after current job	
	Female (1)	Male (2)	Female (3)	Male (4)	Female (5)	Male (6)
Pc_county	0.0526***	0.0371***	0.0453***	0.0181	0.0935***	0.00715
	(0.0136)	(0.0122)	(0.0160)	(0.0124)	(0.0279)	(0.0135)
Constant	0.180***	0.197***	0.146***	0.197***	0.147***	0.205***
	(0.0238)	(0.0246)	(0.0266)	(0.0244)	(0.0408)	(0.0306)
Observations	517	533	329	461	147	297
R-squared	0.033	0.020	0.028	0.005	0.083	0.001

Notes: The estimation uses linear probability model to predict whether the spouse is in the same workplace. *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%). The information about the spouse is available for 2381 respondents, among which 95 lack information about *pc_county* and 2 lack gender. Robust standard errors in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

Another perspective to show that matching is not the main mechanism is to look at whether gender ratios change the magnitude of the effect of social interactions. If there are more *tongxiang* male than female co-workers, then some male migrants will not be able to find marriage partners

even if there is concentration of similar *tongxiang* co-workers, which could also generate the gender asymmetric effect of social interactions. Although we do not observe the gender ratio in the workplace or the gender composition of *tongxiang* co-workers, the data exhibit balanced gender representations in all industries except for construction (Figure A3) and the results are robust to excluding observations in the construction industries. Alternatively, I use *tongxiang* gender ratio in the city to proxy for the ratio in the workplace, and I find that the gender differential in the effect of social interactions is not stronger with a higher relative ratio of *tongxiang* male migrants (column 2 of Table 5).

B *Selection*

Social interactions with *tongxiang* may correlate with other factors that at the same time make people follow rural norms. For instance, people who enter into *tongxiang* network may be more likely to return to their rural hometowns, and therefore have higher incentives to abide by the rural norms. However I find that for both female and male migrants in the sample, those who are surrounded with *tongxiang* co-workers stay longer in cities (Table A9).

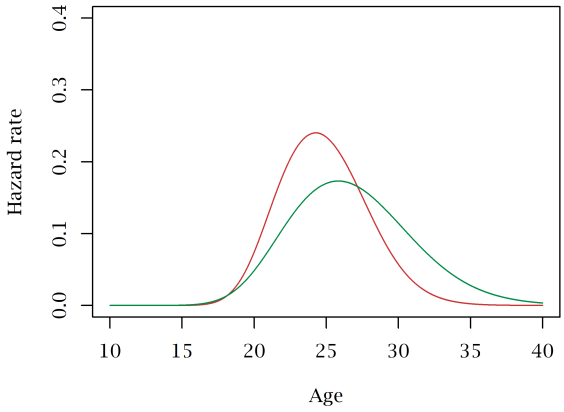
Additionally, social interactions can correlate with factors that affect the timing of marriage. For example, people who are less educated may be more likely to rely on *tongxiang* network and at the same time are more likely to marry early. Also people who are more resourceful join *tongxiang* network and meanwhile are more able to find a partner. Besides, the concentration of *tongxiang* co-workers can be correlated with the concentration of *tongxiang* in the city, therefore the results capture the effect of citywide social interactions instead of workplace. Although to produce the gender differentiated association between social interactions and marriage age, we need the selection on these dimensions to differ between men and women.

To address these concerns, I use propensity score matching and do the matching separately by gender.[12] The idea is to compare individuals otherwise similar, but differ in the concentration of co-workers from the same rural area. The matching function allows to control for observable factors that can predict entry into *tongxiang* concentration in the workplace and marriage at the same time. Table A5 shows the matching function with various factors to predict entry into a workplace with more than 30% of co-workers from the same county. People with *tongxiang* co-workers are less educated and it is more likely to be their first job and to be a job acquired through a referral. The ratio of citywide *tongxiang* is positively correlated with entry into a *tongxiang* enterprise.

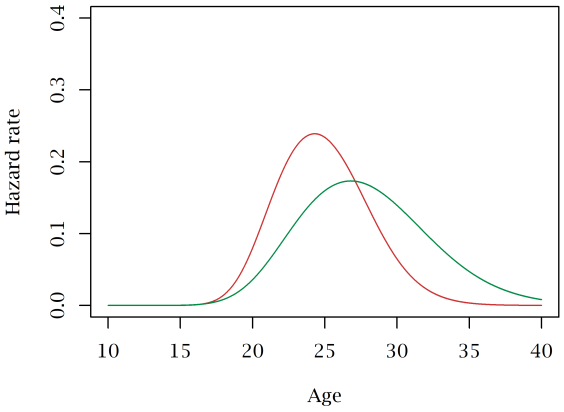
The estimated marriage hazard rates using the matched samples are shown in the left panel of Figure 3 and the estimation using the original sample in the right panel in comparison. The results do suggest positive sorting into *tongxiang* network as the hazard rates for the matched baseline groups (zero *tongxiang* co-workers in red) shift to the left, indicating that migrants who enters into *tongxiang* network have observable 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

[12]The 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 province of origin. The complete list of variables used for matching is shown in Table A5. Alternatively, we can also control for individual characteristics by including them as additive terms in Equation 1. The results are similar.

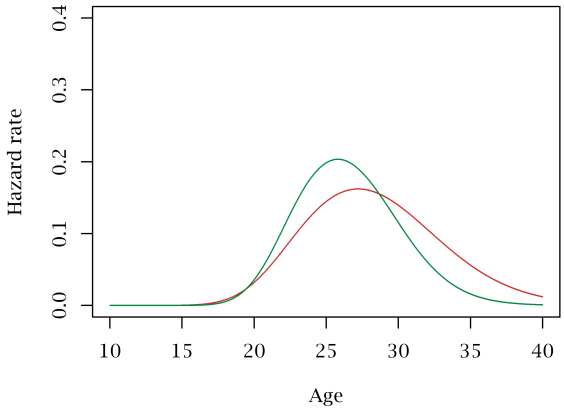
males. The gender differentiated effects of social interactions remain and become even larger when controlling for selection on observables.



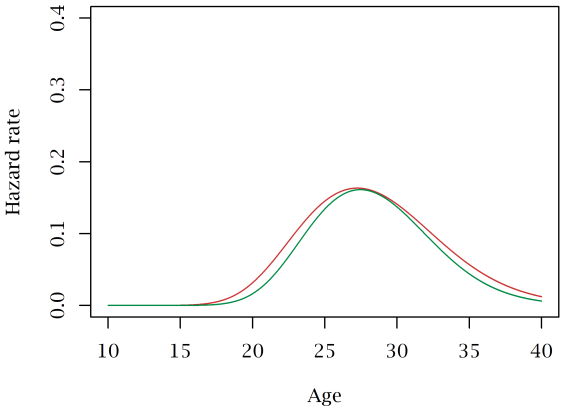
(a) Female, matched



(b) Female, original



(c) Male, matched



(d) Male, original

Figure 3: Estimated hazard rate for matched 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 panel is the same as Figure 1 without the intermediate group, which shows the estimated hazard rates using the original sample. In the left panel, the baseline group (0 *tongxiang* co-workers in green) is a constructed sample matched with the treated group (more than 30% *tongxiang* co-workers 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.

It is also likely that there is unobservable individual characteristics that determine entry into *tongxiang* network and early marriage simultaneously. For instance, social interactions with *tongxiang* co-workers can pick up a private preference for early marriage. If this is the case, we would see correlation between concentration of *tongxiang* co-workers and early marriage present

even for female migrants who are already married before entry into workplace with prevalence of *tongxiang* co-workers. The placebo analysis in Table 3 using the sample of females who marry before migration shows that future interactions with *tongxiang* co-workers do not predict past marriage age. This indicates that instead of capturing spurious correlations, social interactions with co-workers from the same rural origin have a real impact on marriage behaviour. The evidence presented in the previous parts shows that the channel is through social pressure rather than matching.

6. Conclusion

This paper provides novel evidence that social interactions can regulate behaviour by pressuring individuals to conform to cultural norms. I focus on rural-to-urban migrant workers in China and look at the gender asymmetric norm of early marriage from agricultural societies. When females migrate from rural to urban areas, the economic conditions for early marriage to operate cease to exist, but their marriage behaviour can still be subject to old norms when surrounded by individuals from the same rural place. Using variation in social pressure to conform to the rural norm from migrants from the same rural origin in the workplace, I find that social interactions with co-workers from the same hometown substantially increase the probability of early marriage for female migrants. In contrast, social interactions affect to a much less extent the chance of early marriage for male migrant workers.

Consistent with the explanation that social interactions pressure females into early marriage to conform to the rural norms, I find that the gender asymmetric effect is larger for individuals from areas that hold more conservative values for women. Combing propensity score matching and placebo test, I show that the effect is not driven by spurious correlations and selection into social interactions. While close interactions can also facilitate matching between individuals from the same region who share customs and languages, matching story requires either (1) symmetric effect on males, or (2) asymmetric effect on males with a skewed gender ratio, which are not consistent with the data.

The findings carry important policy implications: the economic independence of female migrant workers does not empower them to freely choose when to marry, which can further affect their childbearing decisions and career path. Policies that aim at closing gender gaps by providing equal access to education and jobs may not be as effective if females are expected to abide by traditional norms.

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.

Bau, Natalie, "Can policy change culture? Government pension plans and traditional kinship practices," *American Economic Review*, Forthcoming.

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.

Deng, Quheng, "Earnings differential between urban residents and rural migrants: Evidence from Oaxaca-Blinder and quantile regression decompositions," *Chinese Journal of Population Science*, 2007, *2*, 8–16.

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.

Giuliano, Paola and Nathan Nunn, "Understanding cultural persistence and change," *The Review of Economic Studies*, 12 2020. rdaa074.

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.

Myong, Sunha, JungJae Park, and Junjian Yi, "Social norms and fertility," *Journal of the European Economic Association*, 2020.

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. Appendix

Table A1: Attitudes towards females

	Family first (1)	Marriage first (2)
Rural	0.292***	0.127***
	(0.0164)	(0.0182)
Male	0.0578***	-0.0143
	(0.0186)	(0.0185)
Rural×male	-0.0743***	-0.0619***
	(0.0242)	(0.0270)
Constant	0.519***	0.426***
	(0.0131)	(0.0130)
Observations	6,408	6,390
R-squared	0.077	0.013

Notes: Based on the 2010 Chinese General Social Survey. *Family first* is an indicator for whether the respondent agrees that men should prioritise career and women should prioritise family. *Marriage first* is an indicator for whether the respondent agrees that for females marrying a good husband is more useful than having a good job. The estimation uses linear probability model to predict *famile first* and *marriage first*. *Rural* is an indicator variable that equals one if the respondent was born and have stayed in the rural region and zero if the respondent was born and have stayed in the urban region. Negative and statistically significant interactions between *rural* and *male* indicate gender differentiated rural norms. Robust standard errors in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

Table A2: Summary statistics: to continue

	(1) <30% *tongxiang*	(2) >30% *tongxiang*	(3) Difference
	Panel A: Age		
Female	28.45	30.69	2.25
	(8.70)	(10.38)	(0.83)
Male	31.35	34.36	3.02
	(9.41)	(10.98)	(0.70)
Difference	2.90	3.67	0.77
	(0.31)	(1.04)	(1.08)
	Panel B: Migration age		
Female	20.70	19.65	-1.05
	(6.11)	(6.57)	(0.70)
Male	21.12	21.50	0.38
	(6.93)	(7.30)	(0.60)
Difference	0.42	1.85	1.43
	(0.27)	(0.89)	(0.92)
	Panel C: Years since migration		
Female	7.65	9.19	1.54
	(5.68)	(8.00)	(0.84)
Male	9.97	12.16	2.20
	(7.31)	(7.80)	(0.64)
Difference	2.32	2.97	0.65
	(0.27)	(1.03)	(1.06)
	Panel D: Distance from home (kilometres)		
Female	602.68	518.91	-83.77
	(443.28)	(438.45)	(36.20)
Male	637.38	544.20	-93.18
	(449.88)	(413.72)	(27.95)
Difference	34.70	25.30	-9.41
	(15.70)	(43.03)	(45.73)
	Panel E: Education		
Female	1.30	1.05	-0.26
	(0.90)	(0.85)	(0.07)
Male	1.45	1.12	-0.33
	(0.85)	(0.77)	(0.05)
Difference	0.15	0.08	-0.07
	(0.03)	(0.08)	(0.09)

Table A2: Summary statistics: continued

	(1) <30% *tongxiang*	(2) >30% *tongxiang*	(3) Difference
	Panel F: Weekly working hours		
Female	55.34	59.42	4.08
	(15.01)	(16.68)	(1.36)
Male	56.29	59.92	3.63
	(14.70)	(17.55)	(1.13)
Difference	0.95	0.50	-0.45
	(0.51)	(1.70)	(1.77)
	Panel G: Log hourly wage		
Female	2.00	1.90	-0.10
	(0.45)	(0.62)	(0.05)
Male	2.23	2.21	-0.02
	(0.52)	(0.52)	(0.03)
Difference	0.23	0.31	0.08
	(0.02)	(0.06)	(0.06)

Notes: The table shows the mean and standard deviation (in parentheses) of individual characteristics by gender and concentration of *tongxiang* co-workers ($<$ 30% co-workers from the same county in column 1 and $>$ 30% co-workers from the same county in column 2). The difference by gender is shown in the last row of each panel and the difference by *tongxiang* concentration is shown in column 3. 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).

Table A3: Differences in marriage between worker and farmer identity for individuals married before migration

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

Notes: Estimation is based on the sample of individuals married before migration. The baseline group is individuals who identify themselves as workers. The difference is the increase in probability relative to the baseline group 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 A4: 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 × 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%). *Male* is an indicator for male migrants. *Farmer* is an indicator for self-identification with farmer or rural origin. Robust standard errors in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

Table A5: Determinants of concentration of *tongxiang* co-workers

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

Notes: The table estimates propensity scores using logistic regression. The dependent variable is an indicator for having more than 30% of 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 province in the same city and *log (migrant)* is natural logarithm of the total number of migrant workers in the same city, calculated from the 2000 population census due to data availability. Estimation is based on observations with none of co-workers from the same county or more than 30% co-workers from the same county. Robust standard errors in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent levels.

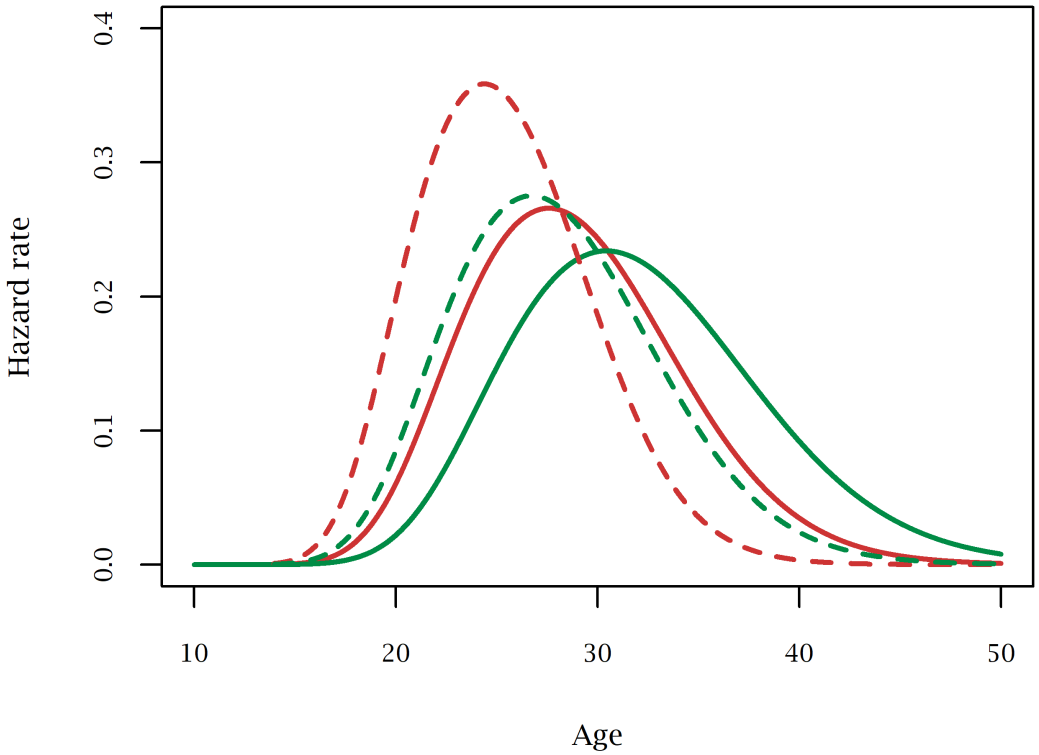


Figure A1: Marriage hazard rate in rural and urban China

— : 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* and currently hold urban *hukou*.

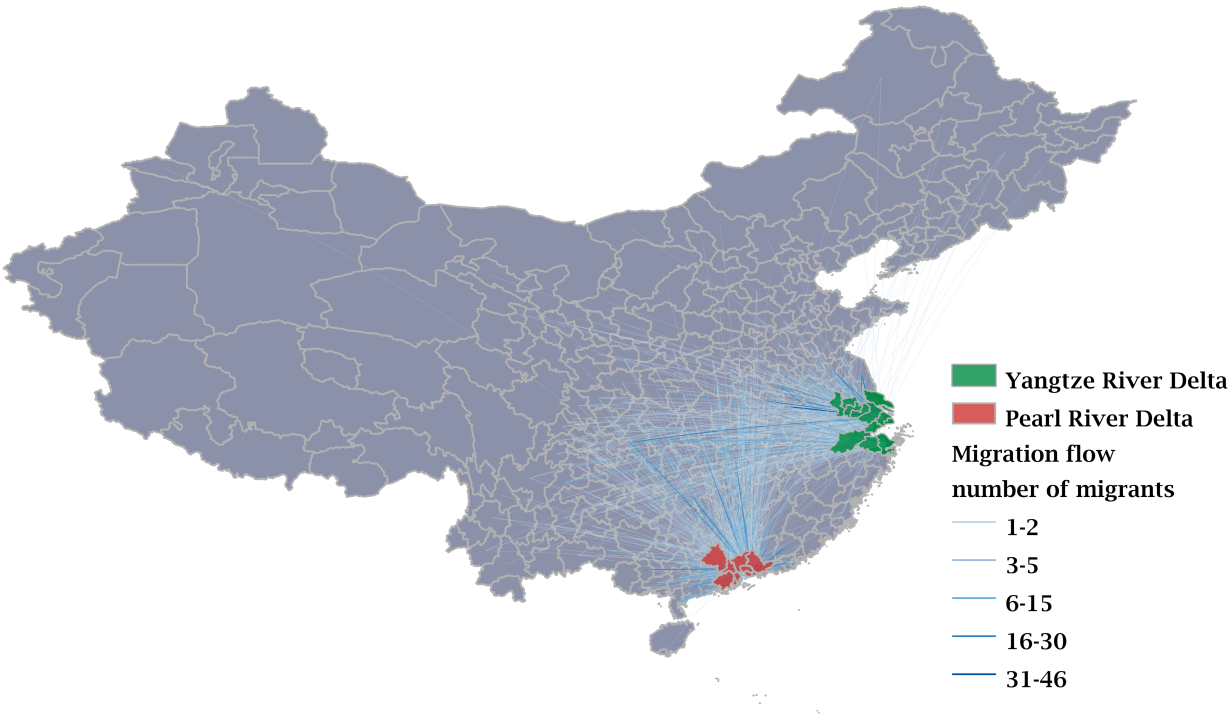


Figure A2: Migration flows to Yangtze and Pearl River Deltas

Notes: Number of rural-to-urban migrant workers (prefecture to prefecture migration) are calculated from the 2010 survey data on migrant workers. The data is representative of regional distribution of migrants.

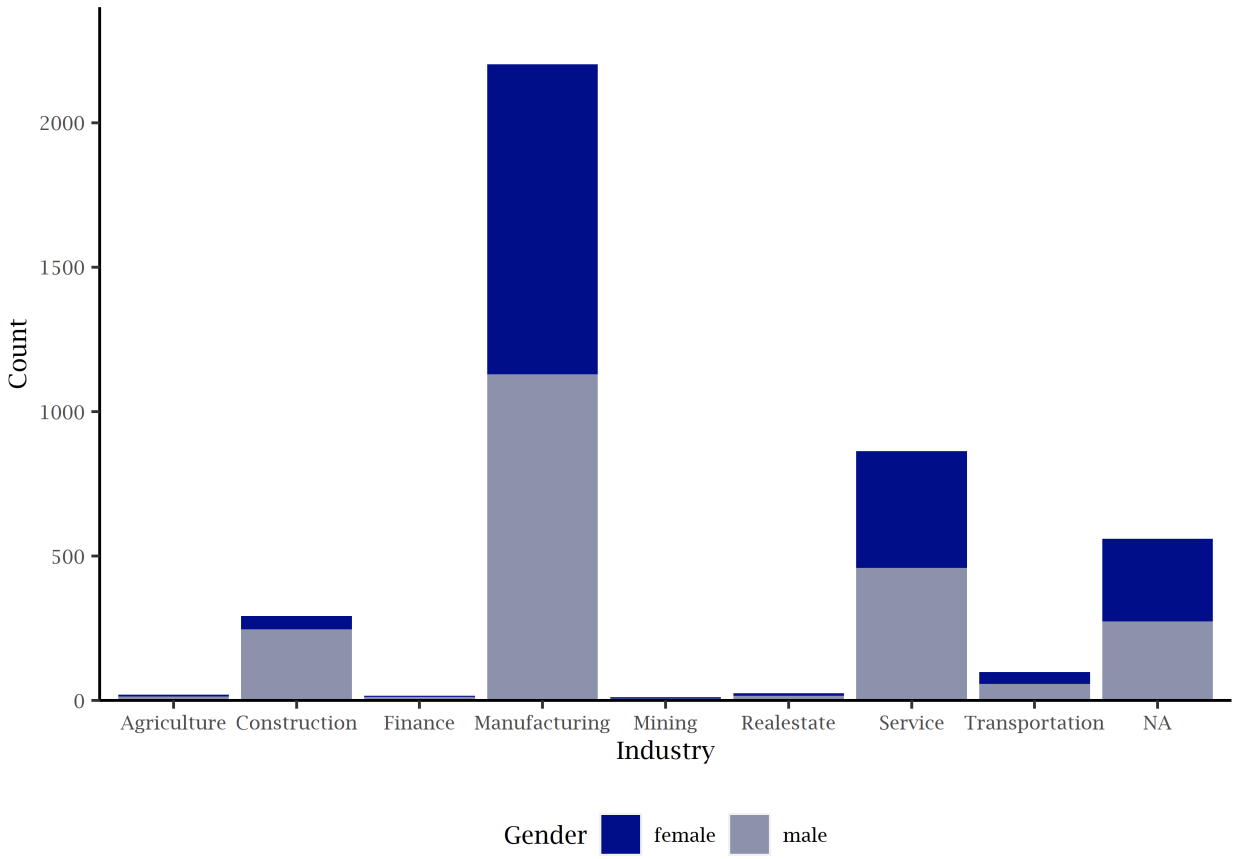
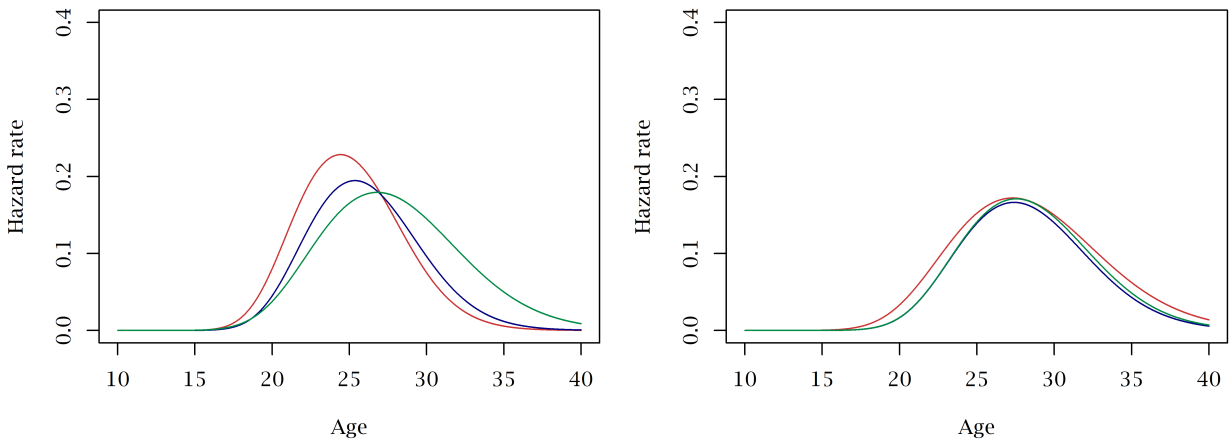


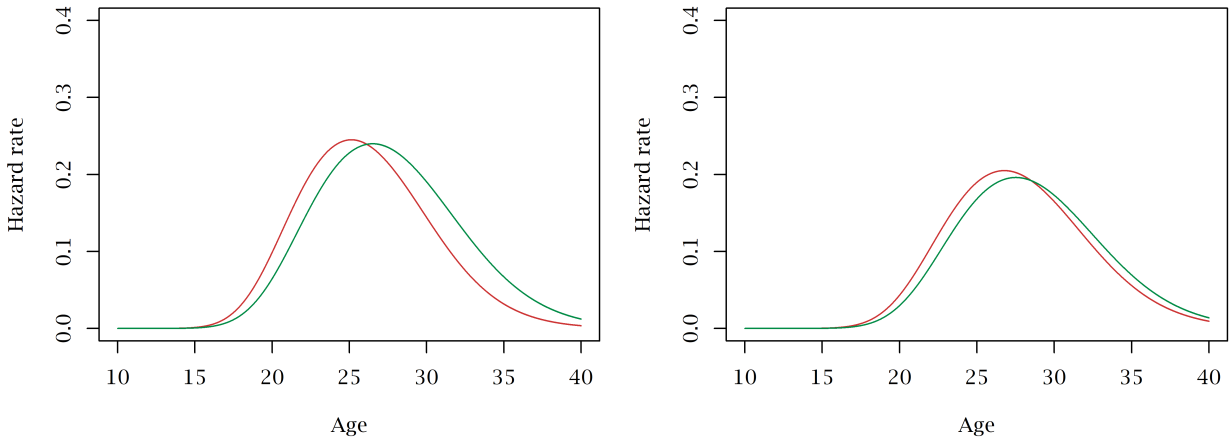
Figure A3: Industrial employment by gender

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



(a) Female (b) Male

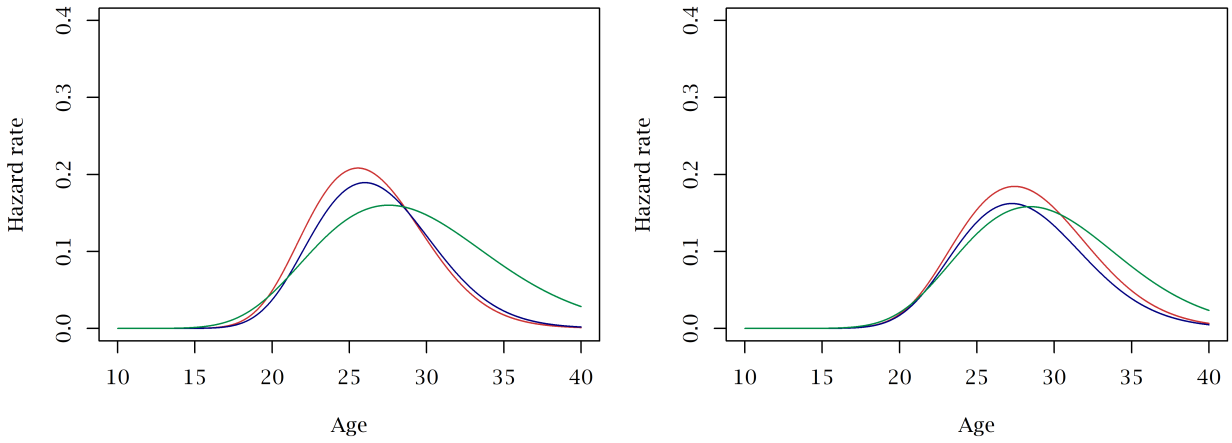
Figure A4: Estimated hazard rate: pc_town
green: none of co-workers come from the same town
blue: 0−30% of co-workers come from the same town
red : >30% of co-workers come from the same town



(a) Female (b) Male

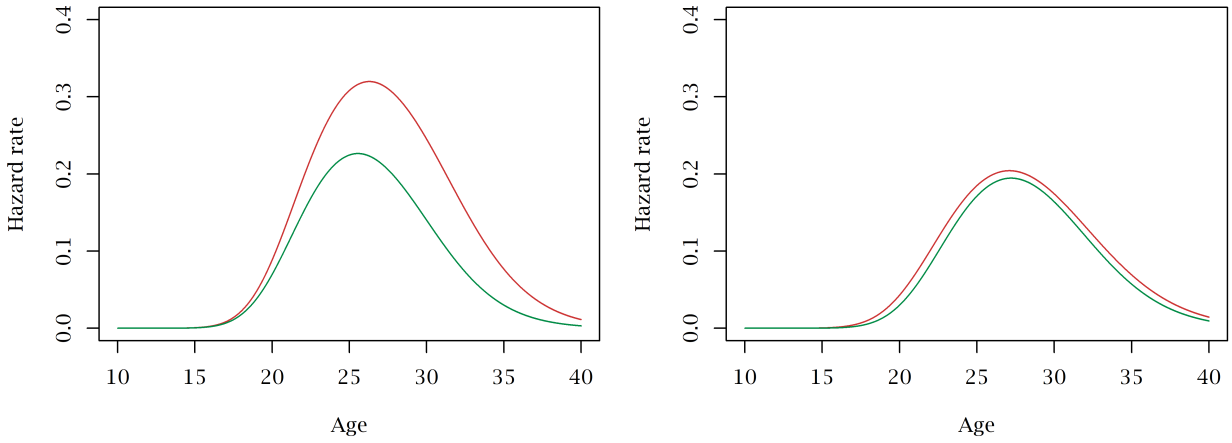
Figure A5: Estimated hazard rate: friend2
red: 2nd best friend is *tongxiang*
green: 2nd best friend is not *tongxiang*

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



(a) Female (b) Male

Figure A6: 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 A7: 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-identified by the survey respondents.

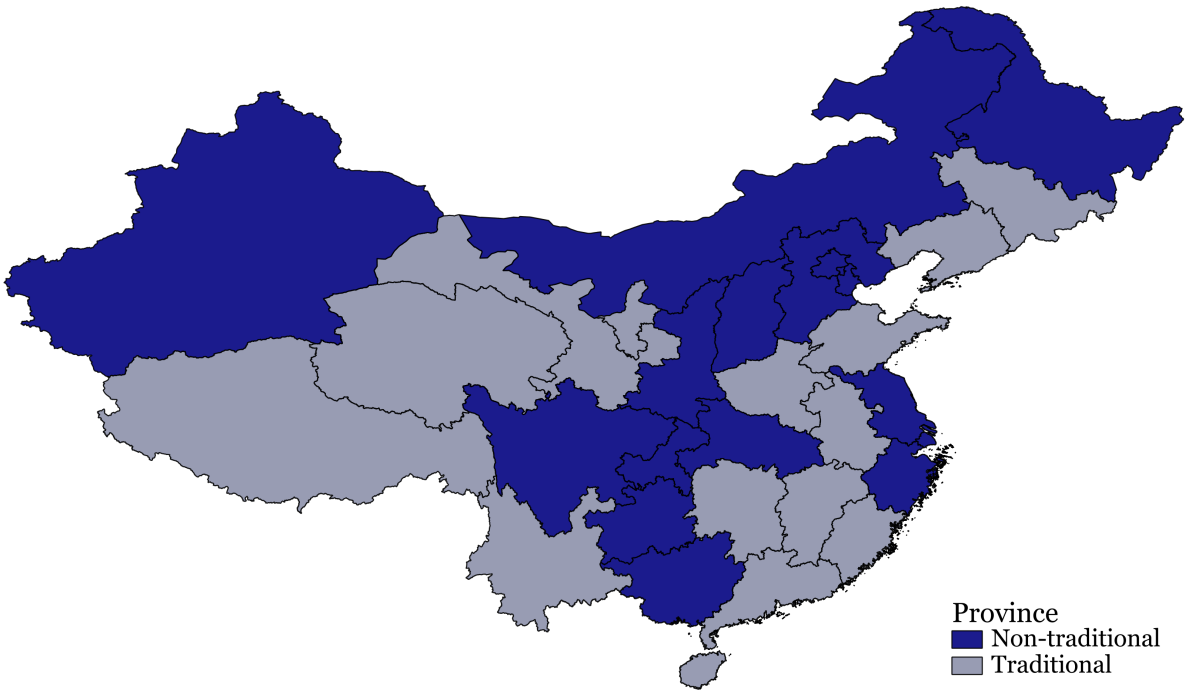
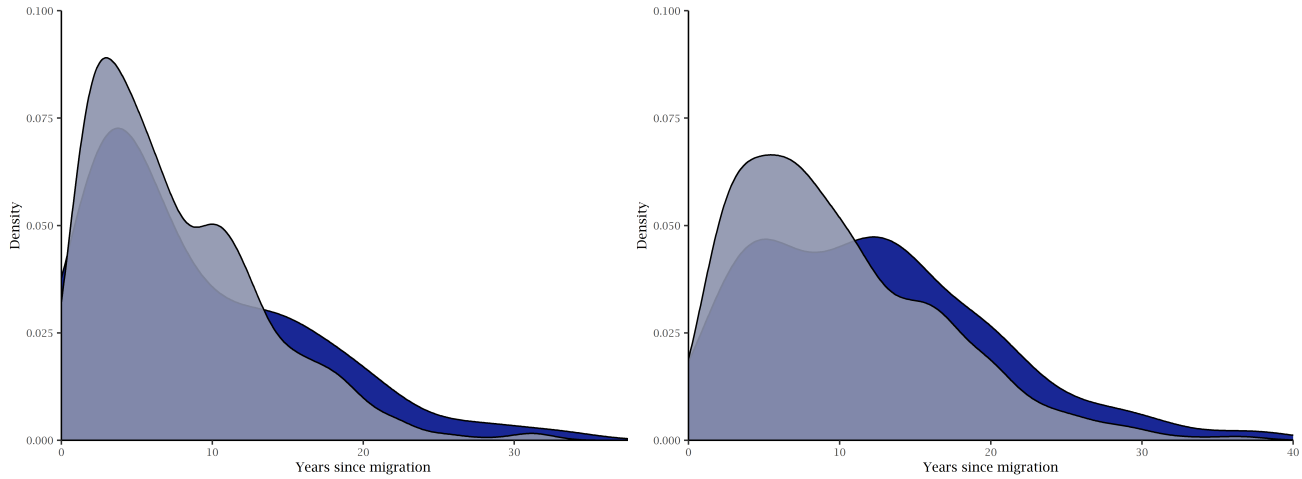


Figure A8: Provinces in China by "Traditionalness"

Notes: "Traditionalness" is defined by attitudes towards the role of female in the society, calculated from the responses to the 2010 Chinese General Social Survey. I compute 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". Provinces above the median is defined as traditional provinces and provinces below the median as non-traditional provinces.



(a) Female (b) Male

Figure A9: Years since migration: pc_county
gray: <30% of co-workers come from the same county
blue: >30% of co-workers come from the same county

Notes: Based on the 2010 survey on migrant workers.