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

Yan Hu[*#]
University of Edinburgh

September 2023

ABSTRACT: This paper provides novel evidence that social interactions of rural-to-urban migrants can enforce traditional norms. Using variation of social pressure to conform to rural norms from migrants from the same hometown in the workplace, I find that the concentration of same-origin co-workers substantially increases the likelihood of early marriage for female rural-to-urban migrants in China, but not male migrants. Consistent with the norm-based explanation for the association between social interactions and early marriage, the gender-differential effect is larger for migrants from regions with more traditional gender norms. The effect is not driven by matching or self-selection into social interactions.

Key words: social interactions, co-workers, marriage, rural-to-urban migrants
JEL classification: D91, J12, J16, R23

[*]31 Buccleuch Place, EH8 9JT, Edinburgh, UK. E-mail: yan.hu@ed.ac.uk. E-mail: yan.hu@ed.ac.uk

[#]I am extremely grateful for the guidance and support of Manuel Arellano and Diego Puga for the paper. I thank Joan Monras, Andreas Stegmann, Ana Tur-Prats and seminar participants at CEMFI, Meeting of the Urban Economics Association, Econometrics Society Asian Meeting, ASREC Europe, Summer School in Urban Economics, and CAGE&IAS Summer School 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. The usual disclaimer applies.

1. Introduction

China's fast-growing cities have attracted a large number of rural migrants. Due to their disadvantaged positions in cities, migrant workers typically form social networks based on their place of origin to share information and resources. Previous literature has mainly focused on the economic consequences of concentration of migrants from the same origin (e.g., Edin, Fredriksson and Åslund, 2003; Munshi, 2003). At the same time, migrant networks can also affect individual behaviour through social interactions and pressure. Given the large discrepancy in economic development as well as norms between rural and urban China, it can have an important consequence for rural-to-urban migrants if social interactions pressure them to adhere to norms from agricultural societies, which may no longer be optimal in cities.

This paper provides novel evidence that social interactions of rural-to-urban migrants can enforce rural norms by examining the marriage decisions of migrant workers in China. 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. Given the strong ties based on kinship and common origin in Chinese society, co-workers from the same rural place can pressure individuals to abide by rural norms even when they now live in an urban setting. Compared to friendship, however, co-workers are less subject to selection bias because individuals have less control over the origin of their co-workers and self-selection into co-workers, if any, is unlikely to be driven primarily by preferences of early marriage.[1] By examining rural-to-urban migrants, the effect of cultural norms can be isolated from that of rural occupations. Because the social norm of early marriage is much more pronounced for females in agricultural societies than for males, I examine the effect separately by gender, which helps to difference out any selection bias that is identical for males and females.

Using a discrete-time hazard model, I find that intensive social interactions (with the majority of co-workers from the same hometown) compel female migrants to abide by norms from their hometown, thereby increasing their probability of getting married before age 22 from 0.21 to 0.43.[2] Conditional on getting married before 40, women with the majority of co-workers from the same hometown are estimated to marry 2.5 years earlier than ones without any same-origin co-workers. For male migrants, however, the association between social interactions and early marriage is much weaker.

To further address selection bias in social interactions, I first use propensity score matching to control for selection on observables. I allow selection to vary by gender and compare individuals who differ only in terms of the composition of their co-workers but are otherwise equal and find that

[1] As expected, the paper finds that same-origin co-workers do not pick up women's preference or tendency to marry early whereas friendships with same-origin individuals are correlated with women's propensity to marry early.

[2] 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 proportion 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 proportion of married female migrants before age 26 from 0.59 to 0.8.

the resulting discrepancy in the effect of social interactions by gender is at least as large.[3] Further, I rule out spurious correlation between early marriage and social interactions (e.g., selection on unobservables) by restricting the sample to individuals who have been married before migration. If same-origin co-workers have no real impact on marriage behaviour but instead capture certain private preferences or tendencies for early marriage, then we would expect a correlation between early marriage and social interactions with same-origin co-workers present even for female migrants who were married prior to entering the workplace. Reassuringly, both the significance level and point estimate drop to zero when we examine females who have been married before migration. This indicates that the association between social interactions and early female marriage is not driven by selection into same-origin co-workers based on preferences or tendencies for early marriage. In addition, I find that there is no correlation between the concentration of same-origin co-workers and female (and male) migrants being more traditional, as proxied by their self-identification with rural origins.

An alternative channel through which social interactions with co-workers of the same origin can affect marriage age is by increasing the likelihood of matching individuals who are similar. However, the findings that the effect of social interactions is much smaller for males and not stronger with a more skewed gender ratio of same-origin migrant workers suggest that matching is not the main mechanism. In addition, I find that female migrants surrounded with same-origin co-workers are more likely to marry someone in the same workplace, regardless of whether or not he is from the same hometown.

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 migrants from regions that hold more traditional views regarding women's role in society. In other words, the same degree of concentration of same-origin co-workers will insert a greater impact on early female marriage if the co-workers are from more traditional regions.

The gender-biased effect of social interactions on marriage echoes the gender-differentiated cultural norms on marriage from agricultural societies. During the prolonged period of agrarian economy in China, early female marriage was practised, thereby facilitating the corresponding norm to form and pass down from generation to generation.[4] According to the 2010 China General Social Survey, compared to urban dwellers, people from rural areas have a greater tendency to believe that the role of women is primarily domestic and that marrying a good husband is of paramount importance for women (Table A1). Such norms translate into behaviour. Individuals tend to marry

[3]Individuals are matched according to an extensive set of characteristics: occupation, educational attainment, age of starting current job, whether it was the first job, whether the job was obtained through referral, the proportion of same-origin migrants in the same city, self-identity, cohort, and place of origin and destination.

[4]Historical records date back to the Western Zhou period (1046 BC–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 BC–188 BC), unmarried women between 15 and 30 years were taxed equivalent to one year's consumption of crops.

earlier in rural areas than in cities, but such a distinction is sharper for females than for males (Figure A1).

As individuals migrate from rural to urban China, taking up jobs in manufacturing and services, the socio-economic environment that allowed for early marriage to operate ceases to exist. However, behaviour may fail to change if individuals continue adhering to old norms, despite the price of early marriage for women being much dearer in the modern economy. I find that female migrants are more likely than male migrants to marry early if they identify more with their rural origin. More importantly, the paper finds that social interactions with co-workers from the same rural origin pressure individuals to comply with old norms, whether or not such norms are internalised.

This paper contributes 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; Kondo and Shoji, 2019). 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 labour market networks (e.g., Cingano and Rosolia, 2012; Hellerstein, Kutzbach and Neumark, 2014) and ethnic enclaves (e.g., Borjas, 1995; Edin, Fredriksson and Åslund, 2003; Munshi, 2003; Zhang and Xie, 2016), which have predominantly been concerned with their economic benefits; and here I explore the social aspects of concentration of same-origin co-workers by showing that their social interactions can enforce rural norms.[5]

In this respect, this paper also contributes to the literature regarding the persistent effect of cultural norms (e.g., Fernández and Fogli, 2006; Fisman and Miguel, 2007) and provides evidence on cultural persistence through social interactions.[6] In particular, I use variations in co-workers from the same rural origin, who can place higher pressure on conformity to rural norms than co-workers from distinct origins, to identify cultural transmission through social interactions. In contrast to Munshi and Myaux (2006) where individuals update their beliefs about the prevailing cultural practice through social interactions, in my case, close social interactions prevent adaptation to changes in the environment by pressuring individuals to comply with the pre-existing norms.

More broadly, the paper contributes to the literature and policy discussion on the gender gap in cities. Urbanisation provides unprecedented job opportunities for female migrants. Previous studies have found that women with labour market opportunities tend to delay marriage and childbearing (Jensen, 2012; Heath and Mobarak, 2015). Nevertheless, social interactions with co-workers from the same hometown can pressure female migrants into early marriage, which may no longer be their optimal choice with enhanced career prospects. Policy interventions aimed at closing the gender

[5]Meng and Xue (2020) also look at migrant workers' social networks in China and examine their effect on migrants' mental health. Their measure of social networks are self-reported numbers of friends and acquaintances living in cities. The strong first-stage result using lagged hometown rainfall suggests that many of the migrant workers' friends come from the same hometown as the migrant workers.

[6]Other mechanisms of cultural persistence have been identified. For example, Giuliano and Nunn (2020) find that cultural persistence and change depends on similarity of the environment across generations. Bau (2021) shows that policies can change cultural persistence.

gap, such as providing equal access to education and jobs, may not be effective under strong social pressure for women to adhere to the traditional norm of early marriage.

The remainder of this paper is organised in the following manner. 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 method, result, and mechanism. Section 5 rules out alternative explanations. Section 6 concludes.

2. Institutional background

In this section, I provide an overview of the causes of the 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 service industries. 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% were concentrated in the Pearl River and Yangtze River Deltas (Figure A2).[7]

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, in turn, 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).[8]

Further, 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, thereby causing them to become socially excluded from the city.[9] Because of their disadvantageous positions in cities, most migrant workers take up jobs that do not appeal to the locals characterised by long hours, poor working conditions, and low and unstable pays (Wang and Zuo, 1999). Moreover, because a strict migration restriction was implemented before 1978, the isolation created large socio-economic barriers to interaction (e.g., language, custom, income) for people from different parts of China.

[7] Source: the Chinese National Bureau of Statistics.

[8] 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, 2011; Lee, 2012).

[9] According to the 2010 survey on migrant workers, over 50% of the migrant workers reported discrimination by local residents and approximately 60% lack of a sense of belongingness in the cities.

Therefore, migrant workers stay closely to their *tongxiang*–the Chinese word for people from the same hometown–to share networks, information, and resources.[10]

The motive to join a *tongxiang* network is usually 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 such interactions can increase the likelihood of matching with a partner. In addition, marriage behaviour can be regulated by social norms and is, therefore, subject to peer pressure. In particular, given the strong ties based on kinship in rural communities of China, individual behaviour conforms highly within the social group of common origin and kinship.[11] This paper finds evidence for the latter channel–that social interactions affect marriage through social pressure to conform to the cultural norms of agricultural societies.

3. Data

The main analysis of this paper is based on survey data from interviews of a cross-sectional 4157 rural-to-urban migrant workers in China in 2010.[12] It took a representative sample of migrant workers working in the Yangtze and Pearl River Deltas in 2010.[13] The quota sampling method was used to correct for representativeness in gender, occupation, and regional distribution of migrants. The data were used by Zhang and Xie (2016) to study the effect of the *tongxiang* network on migrant workers' wages; consequently, in the 2010 wave of the survey, questions regarding the relations with *tongxiang*, or people from the same hometown, were incorporated.

3.1 Measures of social interactions

From the survey data, we obtain the information regarding the percentage of co-workers in the production line from the same province, county, and town, which are respectively first-, third- and fourth-level administrative units in China. In addition, we know whether the best friends of the

[10]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 networks based on their place of origin. Occasionally, this takes the form of *tongxiang* enterprises where hiring decisions are made according to people's place of origin, thereby resulting in a concentration of same origin co-workers within a particular workplace.

[11]Even though co-workers/friends from the same origin are not necessarily related to the individual by blood, they could report individual behaviour to the closely connected rural community back home, which subjects individual behaviour to within-kinship scrutiny.

[12]The survey was part of a project sponsored by the Ministry of Education in China with the aim of studying the status-quo of migrant workers and protecting their rights.

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

migrant worker are *tongxiang*, as identified by the respondents. In the survey, individuals were asked the following question:

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 presents 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 proportion of people have over 30% of co-workers from the same region.

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 presents the levels of the concentration of co-workers (in columns) from the same town, county, and province (in rows).

The concentration of *tongxiang* co-workers is my preferred measure of social interactions because it is a less endogenous choice than friendships. The concentration of co-workers based on the place of origin can be a result of the close proximity between the origin and destination cities, a large out-migration population from the origin, or employers' preferential hiring of *tongxiang*, which is likely to be orthogonal to the decision on the age of marriage.[14]

Moreover, the 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. In addition, 32% of total migrant workers even work 7 days a week. The intensity of the working schedule results in the intensity and closeness of social interactions within the workplace. Migrant workers have numerous opportunities to socialise (Fang, 2012). For example, 36% of total migrant workers live in the dormitory provided by their employers and the percentage increases to 49% for individuals who are single; 62% of workers dine in the canteen of the workplace.

Another potential measure 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 indicate that a considerable proportion of people have

[14]While the availability of same-origin individuals may make a migrant worker more likely to be matched with similar individuals and marry early, Section 5 shows that the gender differential effects of social interactions are not driven by this.

tongxiang as their best friends. The proportion of individuals whose first best friend is *tongxiang* is 0.37. If we consider three best friends, for 13% of the individuals, all their best friends are *tongxiang*, and for 44% of the individuals, at least one of their best friends is *tongxiang*.

However, friendships are subject to a greater extent to selection bias. People choose who their friends are. In particular, individual with more traditional values may be more likely to make friends with *tongxiang*. On the other hand, the concentration of co-workers is more exogenous to the outcome that I am attempting to evaluate. We would expect that the primary function of *tongxiang* enterprises is to advance the economic prospect of migrant workers rather than their marriage. The paper proceeds to demonstrate that self-selection into *tongxiang* co-workers is unlikely to be correlated with a preference for early marriage whereas self-selection into friendship is found to be associated with a preference for early marriage. Therefore, the main analysis is based on the concentration of *tongxiang* co-workers to study the effect of social interactions on marriage age.

3.2 Individual characteristics

The sample consists of 1895 females and 2252 males. After excluding observations that are either widower or divorced, a sample of 4093 individuals remains.

The average migrant worker in the sample has 9 years of education and migrated at an age of 19. The migration is typically temporary. Less than 25% migrant workers express a desire to convert to local *hukou*, not to mention the practical barriers to acquire one. The majority of individuals work in manufacturing and service industries 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 fourth largest employment with a balanced gender representation. The other sectors are relatively small and assimilate approximately 2% of workers (see Figure A3).

For the main analysis in Section 4, I show that the probability of early marriage depends on the prevalence of co-workers from the same rural origin for female migrant workers, but not for male migrants. One concern is that female migrant workers could sort into social interactions with *tongxiang* co-workers and have certain characteristics that make them marry early. Table A2 presents summary statistics of a wide range of individual characteristics by gender as well as by concentration of co-workers from the same town (i.e., over 30% *tongxiang* co-workers and less than 30% *tongxiang* co-workers).[15] 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. In other words, sorting into *tongxiang* enterprises on observables, if any, does not seem to differ considerably between male and female migrants. In

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

Section 5, I use propensity score matching to explicitly correct for selection on observables and show that the gender-asymmetric effect of social interactions is not driven by selection on observables.

3.3 Complementary data

I use the 2010 Chinese General Social Survey which enquires about individual attitudes towards women's role in society to proxy for the differences in marriage norms in different parts of rural China. If social interactions pressure females into early marriage through cultural norms, 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 method and result. I use the discrete-time duration model to examine the effect of social interactions on the distribution of marriage age and quantify the effect by translating the hazard rate to a cumulative distribution function. I then proceed to provide evidence that the association between social interactions with co-workers from the same hometown and early female marriage works through pressure to conform to rural norms.

4.1 Estimation of hazard rate

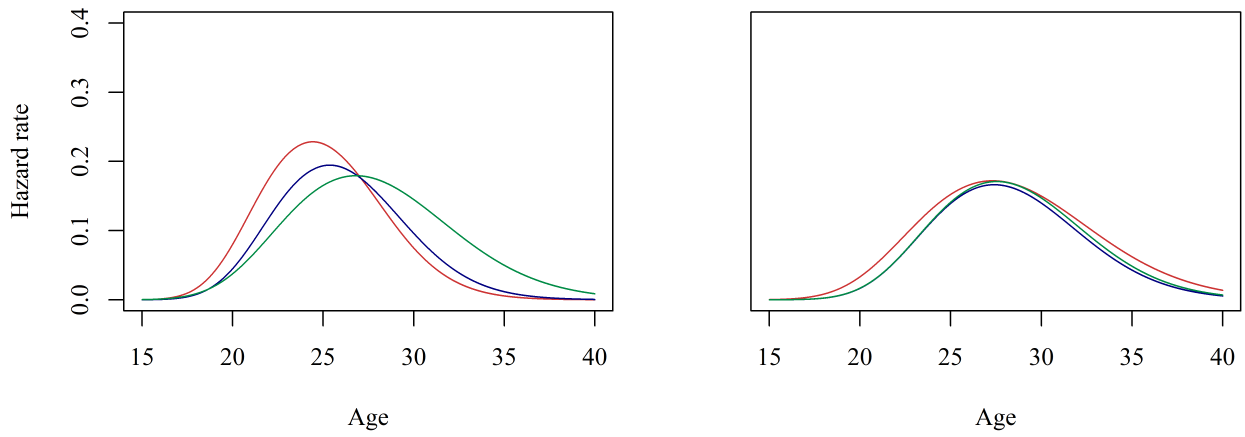
I examine the effect of social interactions on the probability of getting married at different ages. Alternatively, we can estimate the effect of social interactions by comparing average marriage ages, but this comparison would overlook individuals that have not been married in the sample. The fact that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

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

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 that one has not been married before age t. The specification here allows the hazard rate to vary with age t.



(a) Female (b) Male

Figure 1: Estimated hazard rate: pc_town

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

blue: 0−30% of the co-workers come from the same town

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

I examine marriage hazard rates at different ages and also examine how they vary with the 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 town (the fourth-level administrative unit in China). As the proportion of *tongxiang* co-workers increases, we see a gradual shift and intensification of exit (from singlehood) rate at early ages. This indicates that female migrant workers are more likely to marry early when they have a higher proportion of co-workers from the same hometown in the production line. In contrast, the concentration of *tongxiang* co-workers does not seem to affect the marriage age of male migrant workers, as shown in the right panel of Figure 1.

In Section 5.1, I use propensity score matching to control non-parametrically for an extensive set of individual characteristics (i.e., occupation, education attainment, age of starting current job, whether it was the first job, whether the job was obtained through referral, *tongxiang* ratio in the city, self-identity, cohort, and place of origin and destination) and find similar gender-differentiated effects of *tongxiang* co-workers.

Similar gender-differentiated patterns arise by examining co-workers from the same county and the results are presented in Appendix (Figure A4). The effect of co-workers from the same province is less clear (Figure A5). 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 and,

therefore, individuals may stop to identify people from the same province as *tongxiang*.[16]

4.2 Estimation of cumulative distribution function

The estimated hazard rates provide an intuitive overview for the effect of social interactions on marriage age. However, they cannot be used to directly quantify the magnitude of the effect of social interactions. This is because the difference in hazard rates of marriage between the treated and control groups for any given age is a combination of the treatment effect of social interactions and the compositional differences between groups. At each point of time, as more marriageable individuals exit from singlehood and more individuals exit in the group with a higher concentration of *tongxiang* co-workers, the compositions of the remaining population change differently for groups with different proportions of *tongxiang* co-workers.[17] As time goes by, people who remain unmarried in the group with more *tongxiang* co-workers are, on average, less marriageable than people in the group with less *tongxiang* co-workers. Consequently, the difference of hazard rates is likely to underestimate the treatment effect of social interactions, as the treated group with a higher proportion of *tongxiang* co-workers over time has a less marriageable population.

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

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

where $F(t)$ is the proportion 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 over 30% or 50% of the co-workers from the same town with ones with none of the co-workers from the same town. The sample is restricted to individuals who get married after migration. For presentational purposes, the table only shows the probability of getting married before the ages of 20–30.[18] In columns 1 and 4, I show the probability of getting married before a certain 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 hometown. The differences in columns 2, 3 for females and columns 5, 6 for males yield the effect of social interactions, which is the increase relative to the baseline group in the probability of getting married

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

[17]This is evident from the fattening-out of the right-hand tail of the hazard rate for the group of female migrants with high levels of concentration of co-workers from the same rural origin.

[18]The probability of marriage can be calculated for all ages starting from 16, which is set to be the minimum marriage age.

Table 2: Marriage age and same-town co-workers—main analysis

	Probability of being married before a certain age					
	Female			Male		
Age	Baseline 0 same origin (1)	Difference >30% (2)	Difference >50% (3)	Baseline 0 same origin (4)	Difference >30% (5)	Difference >50% (6)
20	0.070	0.069*	0.101	0.027	0.035*	0.046
		(0.039)	(0.079)		(0.019)	(0.025)
21	0.128	0.122**	0.167**	0.060	0.054**	0.067*
		(0.048)	(0.081)		(0.026)	(0.032)
22	0.207	0.174***	0.227***	0.114	0.072**	0.084**
		(0.057)	(0.084)		(0.031)	(0.038)
23	0.302	0.208***	0.261***	0.190	0.083**	0.093**
		(0.065)	(0.089)		(0.036)	(0.046)
24	0.404	0.217***	0.265***	0.283	0.086**	0.092*
		(0.071)	(0.093)		(0.041)	(0.054)
25	0.503	0.204***	0.245**	0.385	0.081*	0.082
		(0.077)	(0.095)		(0.046)	(0.063)
26	0.591	0.177**	0.212**	0.483	0.071	0.069
		(0.081)	(0.097)		(0.049)	(0.068)
27	0.664	0.145*	0.174*	0.571	0.060	0.056
		(0.083)	(0.098)		(0.050)	(0.069)
28	0.722	0.114	0.139	0.644	0.050	0.046
		(0.084)	(0.099)		(0.050)	(0.068)
29	0.767	0.087	0.109	0.701	0.042	0.038
		(0.084)	(0.099)		(0.049)	(0.065)
30	0.801	0.064	0.084	0.745	0.037	0.034
		(0.084)	(0.100)		(0.047)	(0.062)

Notes: The estimation is based on the sample of individuals married after migration. The baseline group is migrants with 0 co-workers from the same town. The difference is the increase in probability relative to the baseline group for migrants with over 30% or 50% of the co-workers from the same town. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

for migrant workers with over 30% and 50% of the co-workers from the same hometown. Bootstrap standard errors with 500 replicates are given in parentheses.[19]

A female migrant who does not have any *tongxiang* co-workers is expected to get married by 22 with a probability of 0.21 (column 1). If, instead, the majority of her co-workers in the same production line come from the same hometown, the likelihood of being married by 22 increases by 23 percentage points (column 3). The effects on male workers, however, are much smaller (columns

[19]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 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.

5 and 6).

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 certain private preferences or tendencies for early marriage, we would expect a correlation between early marriage and social interactions with same-origin co-workers present even for females who are already married before entering a workplace with *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 a city. Now, the effect of social interactions disappears for females (columns 2 and 3). Not only does the significance level fall to zero but so do the point estimates. The concentration of *tongxiang* co-workers does not predict past marriage age, thereby indicating that social interactions have a real impact on behaviour instead of picking up some unobservable individual characteristics that simultaneously lead to early marriage.

Notice that for male migrant workers who are married before migration, the estimated association between social interactions and early marriage is slightly positive but very similar to the results for migrants who marry after migration. This suggests positive sorting for male migrants, but almost all positive association disappears once selection on observables is controlled for as shown later in Part B of Section 5.

In Table A5, I present additional results using alternative thresholds of concentration of same-town co-workers. In particular, female migrants with over 10% or 20% of co-workers from the same hometown also have a higher probability of early marriage than females without any *tongxiang* co-workers, and as expected the effects are more modest than the results at higher thresholds (i.e., 30% and 50% same-town co-workers) in Table 2. Consistent with Figure 1, the effect of social interactions increases with the fraction of same-town co-workers.

The estimation results using same-county co-workers (Table A6, Table A7 and Table A8) show similar patterns: (1) female migrant workers with co-workers from the same rural origin are more likely to marry early, (2) the effect is stronger with a higher fraction of same-origin co-workers, (3) the effect is much smaller for male migrants, and (4) the association between same-origin co-workers and early marriage is not present for female migrants who have been married before migration, ruling out spurious correlation.

By comparison, Table A3 and Table A4 show the estimates using the friendship measure. The gender-differentiated effects of social interactions are similar as to those of co-workers. For a female migrant worker who marries after migration, she is more likely to marry early if her best friend is from the same hometown. The effect for male migrants is smaller.[20] Interestingly, if we focus on individuals who have been married before migration, there is still a positive association between early marriage and same-origin best friends for female migrant workers. This indicates, contrary

[20]Note also that the estimated effect of best friends from the same hometown is smaller than the effect of 50% co-workers from the same hometown and comparable to the effect of 10% co-workers from the same hometown.

Table 3: Marriage age and same-town co-workers—placebo analysis

	Probability of being married before a certain age					
	Female			Male		
Age	Baseline 0 same origin (1)	Difference >30% (2)	Difference >50% (3)	Baseline 0 same origin (4)	Difference >30% (5)	Difference >50% (6)
20	0.285	−0.039	−0.063	0.159	0.084*	0.100*
		(0.054)	(0.228)		(0.044)	(0.056)
21	0.437	−0.036	−0.032	0.267	0.096*	0.102
		(0.064)	(0.190)		(0.055)	(0.070)
22	0.594	−0.028	0.011	0.398	0.093	0.087
		(0.070)	(0.144)		(0.063)	(0.081)
23	0.731	−0.021	0.040	0.536	0.079	0.061
		(0.070)	(0.104)		(0.066)	(0.086)
24	0.835	−0.016	0.046	0.662	0.059	0.034
		(0.063)	(0.073)		(0.063)	(0.083)
25	0.904	−0.013	0.038	0.765	0.040	0.011
		(0.051)	(0.047)		(0.055)	(0.076)
26	0.946	−0.011	0.026	0.841	0.025	−0.003
		(0.038)	(0.029)		(0.046)	(0.065)
27	0.971	−0.010	0.016	0.895	0.014	−0.011
		(0.028)	(0.018)		(0.037)	(0.054)
28	0.984	−0.008	0.009	0.930	0.008	−0.014
		(0.020)	(0.011)		(0.029)	(0.044)
29	0.991	−0.007	0.005	0.953	0.004	−0.014
		(0.015)	(0.007)		(0.022)	(0.035)
30	0.995	−0.006	0.003	0.967	0.002	−0.012
		(0.011)	(0.005)		(0.017)	(0.028)

Notes: The estimation is based on the sample of individuals married before migration. The baseline group is migrants with 0 co-workers from the same town. The difference is the increase in probability relative to the baseline group for migrants with over 30% or 50% of the co-workers from the same town. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

to co-workers, friendship is subject to selection on early marriage for female migrant workers. In other words, female migrants who have a tendency for early marriage are more likely to associate themselves with friends from the same hometown.

The results in this section indicate a causal link from concentration of co-workers from the same rural origin to female early marriage. The next section proceeds to provide evidence for the underlying mechanism.

4.3 Effect of social interactions through cultural norms

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.*

Social interactions with co-workers from the same rural origin can exert much stronger social pressure to conform to rural norms of early marriage than interactions with co-workers from distinct origins who form looser relationships. Consistent with the data, while migrants with a diverse social network–whose members could be equally rural–adapt to the new socio-economic environment in cities by marrying late, female migrant workers with the concentration of co-workers from the same rural origin are pressured into early marriage.

I provide two sets of evidence to bolster the idea that the effect of social interactions on marriage age 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 these norms. While the first piece of evidence links norms with marriage age, the second set of evidence links social pressure to conform to norms with marriage age. If the concentration of co-workers from the same rural origin exerts pressure to conform to norms, then the effect will be stronger if the social pressure is higher (i.e., higher concentration of *tongxiang* co-workers) and if the norm is stronger (i.e., more traditional values on women). The previous part showed that the higher concentration of *tongxiang* co-workers, the earlier female migrants marry. This part proceeds to show that, holding constant the degree of concentration of *tongxiang* co-workers, the effect will be larger if the norm is stronger. In particular, I show that the gender-differentiated response to the presence of *tongxiang* co-workers is stronger for those whose hometown has more conservative attitudes towards women's role in society.

4.3.1 Farmer v.s. Worker Identity

Female migrant workers who identify with their rural origin internalise social norms in an agriculture society and may continue to behave according to rural norms when economic conditions change.

I investigate this possibility in this section as supporting evidence that the gender-biased effect of social interactions is attributed to rural norms. 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 the following question:

Q2 Which of the following do you believe to be your identity?
(1) Farmer, (2) Worker, (3) Others, (4) Do not know.

Table 4: Marriage age and identity

	Probability of being married before a certain age			
	Female		Male	
Age	Baseline Worker (1)	Difference Farmer (2)	Baseline Worker (3)	Difference Farmer (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: The 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% levels.

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 get 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 themselves as a farmer for females in column 2 and for males in column 4. Female migrants who recognise themselves as farmers are more likely to get married early (column 2). For males, the association between rural origin and marriage age is much weaker (column 4).

One concern is that people who identify with their rural origin may be more likely to interact with *tongxiang*. Consequently, the correlation between the prevalence of *tongxiang* co-workers and early marriage instead reflects the effect of self-identity on marriage behaviour. In Table A10, I regress the concentration of *tongxiang* co-workers on an indicator for self-identification as a farmer and find little correlation between the two. This also provides additional evidence that individuals do not self-select into same-origin co-workers based on the rural norm of early marriage, proxied by self-identification with rural origin.[21]

Another possibility is that self-identification with rural origin is correlated with occupations, which lead to early marriage. However, in the sample, less than 1% of individuals work in the agriculture sector. Self-perceived identity does not appear 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 who self-identify as workers. For service, the second largest category, the comparison is 22.51% v.s. 23.92 %.

I find that for individuals who marry before migration, females who identify with their rural origin (responded in the survey after migration and marriage) have also married early but their male counterparts have not (Table A9). This is consistent with the idea that self-identity is stable over time and is associated with a set of values and norms that 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. Note also that the effect of farmer identity is weaker than the effect of same-origin co-workers. This may be due to either (1) farmer identity being an imperfect measure of the social norm of early marriage, or (2) the intensifying and multiplying effects of social pressure on marital behaviour.

4.3.2 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 stronger effect of social pressure of a concentration of *tongxiang* co-workers if the norm is stronger. The strength of the norm can be proxied by views regarding women's role in agricultural societies. I use the response in the 2010 China General Social Survey to the following question:

[21] 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.

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 proportion of rural respondents who agree or strongly agree with the above statement (high proportion 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 the value of marriage compared to career for females, which is closely related to a preference of female early marriage. 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 (s: majority of *tongxiang* co-worker; ns: no *tongxiang* co-workers); gender by M, H (M: female; H: male); and "traditionalness" by c, nc (c: traditional province; nc: non-traditional province); we can write the triple differences of the effect of social interactions on the probability of getting married before age t:

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

As shown in column 1 of Table 5, the gender differential is larger for individuals whose native provinces have more conservative values for females. Females are more likely to get married early when they are surrounded by *tongxiang* co-workers than male migrants (e.g., the differences shown in columns 3 and 6 of Table 2), yet there is heterogeneity in the gender-asymmetric response to the concentration of *tongxiang* co-workers depending on the strength of the norm. In particular, the gender differential of the probability of getting married before 25 is 39 percentage points higher for individuals originally from traditional provinces than non-traditional provinces. In other words, the same degree of concentration of *tongxiang* co-workers will have a stronger effect on female early marriage if the *tongxiang* co-workers are from areas with more traditional norms on women's role in society.

5. Alternative explanations

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

Table 5: Gender-differentiated effect of social interactions

	Triple differences of probability of being married before a certain age	
	Traditional v.s. non-traditional	More males v.s. less males
Age	(1)	(2)
20	0.259	0.087
	(0.428)	(0.352)
21	0.220	0.075
	(0.390)	(0.318)
22	0.212	0.042
	(0.336)	(0.283)
23	0.255	0.000
	(0.269)	(0.259)
24	0.329	−0.038
	(0.220)	(0.246)
25	0.386*	−0.064
	(0.200)	(0.240)
26	0.394**	−0.076
	(0.200)	(0.237)
27	0.364*	−0.076
	(0.202)	(0.235)
28	0.322	−0.067
	(0.203)	(0.233)
29	0.279	−0.053
	(0.202)	(0.231)
30	0.240	−0.038
	(0.199)	(0.230)

Notes: Column 1 compares the gender differential in the probability of getting married before a certain age between individuals from traditional and non-traditional provinces. Traditional provinces are defined as ones that have above median fraction of rural residents who agree that for females it is more useful to have a good husband than a good career (Q2). Column 2 compares the gender differential between individuals in cities with a higher origin-destination specific (*tongxiang*) 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% levels.

5.1 Matching

Migrant workers originating from the same rural area share customs, languages, and similar socio-economic conditions. The concentration of *tongxiang* co-workers can facilitate meeting fellow countrymen who are similar to the individual, thereby increasing the likelihood of meeting a potential partner and a quick transition into marriage.

Although the data do not reveal the identity of the spouse, there are a few 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 same-origin co-workers. Note that the measure of *tongxiang* co-workers for individuals is (only available) for their current job.

Table 6: Spouse workplace and same-town co-workers

	Dep. var.: 𝟙 (Spouse in the same workplace)					
	Married before migration		Married after migration before current job		Married after current job	
Age	Female (1)	Male (2)	Female (3)	Male (4)	Female (5)	Male (6)
Sametown	0.236***	0.197***	0.242**	0.127*	0.456***	0.030
	(0.078)	(0.068)	(0.097)	(0.068)	(0.144)	(0.072)
Constant	0.198***	0.216***	0.129***	0.188***	0.159***	0.159***
	(0.026)	(0.027)	(0.026)	(0.025)	(0.045)	(0.032)
Observations	283	304	190	293	82	169
R-squared	0.042	0.033	0.052	0.014	0.156	0.001

Notes: The estimation utilises the linear probability model to predict whether the spouse is in the same workplace. *Sametown* is a dummy variable that equals to 1 if the individual has over 30% same-town co-workers and 0 if none of the co-workers is from the same hometown. The information about the spouse in this subsample is available for 1,323 respondents, among which 2 lack information regarding gender. Robust standard errors are given in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

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. *Sametown* is a dummy variable that equals one if the individual has over 30% of the co-workers from the same town and zero if none of the co-workers are from the same hometown. Columns 1 and 2 show that for both males and females who are married before migration, their spouse is more likely to be in the same workplace if over 30% of the co-workers (at the current job after migration and marriage) are from the same hometown. On the other hand, for individuals who marry after migration and particularly ones who marry after beginning the current job (where the proportion of *tongxiang* is measured), only husbands of the female migrant workers are more likely to be in the same workplace if their co-workers are from the same hometown (an increase in probability from 0.159 to 0.615 in column 5). In Table A11, I employ an alternative specification that predicts spouse in the same workplace by a categorical variable, *pc_town*, whose value increases with the proportion of *tongxiang* co-workers from the same town. The results show very similar patterns.

These results suggest that (a) individuals who are married before migration (possibly from the same hometown) are more likely to later enter a workplace with a concentration of their *tongxiang*; and (b) unmarried females who enter in a workplace with a higher proportion 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 hometown takes place in the workplace, male migrants who arrive at the current job unmarried would also be more likely to find a spouse from the pool of *tongxiang*

co-workers, but the small and insignificant coefficient on *sametown* in column 6 indicates that this is not the case.[22] In other words, females migrants with same-town co-workers marry someone in the workplace, regardless of whether or not he is from the same hometown.

The results in columns 1 and 2 of Table 6 may appear to be inconsistent with the result from the placebo test in Table 3 at first glance, but they look at distinct issues. Table 6 shows that there is selection into same-town co-workers based on marital status. A married couple, possibly from the same hometown, are more likely to enter a *tongxiang* firm than other firms. Table 3 shows, however, that there is no selection based on a preference for early marriage. Comparing different women, but all of them were married before migration, whether or not one later enters a workplace with *tongxiang* co-workers is not correlated with her age at marriage, even though those who migrated after marriage generally married early.

In addition to results on the probability of early marriage shown in Section 4, examining a different outcome—whether the spouse is in the same workplace—also shows a gender-asymmetric pattern. Both results are consistent with the idea that females are pressured into early marriage with close interactions with co-workers of the same rural origin and marry someone in the workplace, regardless of whether or not he is from the same hometown.

Another perspective to show that matching is not the main mechanism is to examine whether gender ratios change the magnitude of the effect of social interactions. If there are considerably more *tongxiang* male than female co-workers, then some male migrants will not be able to find marriage partners even if they have female *tongxiang* co-workers and prefer them. This can 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 migrants in the construction industry. In addition, I use the gender ratio of *tongxiang* in the city to proxy for the ratio in the workplace and 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).

5.2 Selection

Social interactions with *tongxiang* may correlate with factors that affect the timing of marriage. For example, people who are less educated may be more likely to rely on the *tongxiang* network and, simultaneously, are more likely to marry early. Moreover, people who are more resourceful can join the *tongxiang* network and are also more capable of finding a marriage partner. However, in order 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.

[22] As same-sex marriage has not been legalised in China, the matching story would require the marriage between a male and a female of the same origin. Therefore, we expect a positive association between *tongxiang* co-workers and spouse in the same workplace for male migrants as well, contrary to column 6 of Table 6.

To address these concerns, I use propensity score matching to control for (gender-differential) selection on observables.[23] The idea is to compare individuals who are otherwise similar but only differ in terms of the concentration of co-workers from the same rural area. The matching function allows to control for observable factors that can result in entry into a *tongxiang* concentration in the workplace and early marriage simultaneously. Table A12 shows the matching function, separately by gender, with various factors to predict entry into a workplace with over 30% of the co-workers from the same hometown. People with *tongxiang* co-workers are less educated and it is more likely to be a job acquired through a referral. Further, the ratios of citywide *tongxiang* are positively correlated with entry into a *tongxiang* enterprise, although their effects are not statistically significant.[24] In addition, I use self-identification with farmer and rural origin in the matching function as an attempt to match individuals based on their own preferences for early marriage. Yet as Table A10 already suggests, farmer identity does not predict the concentration of *tongxiang* co-workers. As farmer identity may not fully capture a preference for early marriage, I will also discuss further selection on unobservables at the end of this part.

In Table A13, I recalculate the cumulative distribution function using matched observations. The estimated effects of social interactions remain similar for female migrants, although the significance levels become slightly lower with a smaller sample size. For male migrants, the small positive correlation between early marriage and co-workers from the same town disappears after controlling for selection on observables.

It is also possible that there are unobservable individual characteristics that determine both entry into a *tongxiang* network and early marriage. In particular, there is a concern that individuals with a preference or tendency to marry early may be more likely to enter a workplace with *tongxiang* co-workers. The paper provides the following evidence to show that this is unlikely to be the case. First, there is no association between early marriage and *tongxiang* co-workers among women who married before migration (Table 3). If social interactions with co-workers pick up certain private preferences or tendencies for early marriage, we would expect a correlation between early marriage and same-origin co-workers even among women who already married prior to meeting *tongxiang* co-workers. This contrasts with the results for friends from the same hometown, where friendship choice shows an association with a preference for early marriage (Table A4). Second, migrant workers' identification with their farmer or rural origin predicts early female marriage and is therefore a plausible proxy for the preference for early marriage, yet it is not correlated with the concentration of co-workers from the same town (Table A10).

[23]The estimation procedure of the duration model with propensity score matching follows Austin (2014). The matching function follows Zhang and Xie (2016). In addition, I include indicators for farmer identity, occupational dummies, cohort dummies and dummies for the province of origin and destination. In order to account for differential selection by gender, the matching is done separately by gender. The complete list of variables used for matching and their contribution to the matching function is presented in Table A12.

[24]Controlling for the concentration of *tongxiang* in the city helps to differentiate the effect of social interactions in the workplace from citywide social interactions.

The results indicate that instead of capturing spurious correlations, social interactions with co-workers from the same rural origin have an actual impact on marriage behaviour. The evidence presented in the previous sections reveals that the channel is social pressure and not matching.

6. Conclusion

This paper provides novel evidence that social interactions of rural-to-urban migrants can pressure their behaviour to conform to traditional norms. I focus on the gender-asymmetric norm of early marriage from agricultural societies and use variation in social pressure to conform to the norm from migrants from the same rural origin in the workplace. When women migrate from rural to urban areas, the economic conditions for early marriage cease to exist, but their marriage behaviour can still be subject to old norms when pressured by individuals from the same rural area. 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 the likelihood of early marriage for male migrant workers to a much lesser extent.

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

The findings carry important policy implications: the economic independence of female migrant workers does not necessarily empower them to freely choose when to marry, which can further affect their childbearing decisions and career paths. Policies that aim at closing gender gaps by providing equal access to education and jobs may not be as effective if women 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. https://doi.org/10.1002/sim.5984.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Social incentives in the workplace," *The Review of Economic Studies*, 2010, *77* (2), 417–458. https://doi.org/10.1111/j.1467-937X.2009.00574.x.

Bau, Natalie, "Can policy change culture? Government pension plans and traditional kinship practices," *American Economic Review*, 2021, *111* (6), 1880–1917. https://doi.org/10.1257/aer.20190098.

Borjas, George, "Ethnicity, neighborhoods, and human-capital externalities," *American Economic Review*, 1995, *85* (3), 365–390.

Bover, Olympia, Manuel Arellano, and Samuel Bentolila, "Unemployment duration, benefit duration and the business cycle," *The Economic Journal*, 2002, *112* (479), 223–265. https://doi.org/10.1111/1468-0297.00034.

Burke, Mary A and Tim R Sass, "Classroom peer effects and student achievement," *Journal of Labor Economics*, 2013, *31* (1), 51–82. https://doi.org/10.1086/666653.

Cingano, Federico and Alfonso Rosolia, "People I know: job search and social networks," *Journal of Labor Economics*, 2012, *30* (2), 291–332.

Cornelissen, Thomas, Christian Dustmann, and Uta Schönberg, "Peer effects in the workplace," *American Economic Review*, 2017, *107* (2), 425–456. https://doi.org/10.1257/aer.20141300.

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. https://doi.org/10.1162/00335530360535225.

Fang, I-Chieh, "Keen to get married: Why marriage is so important to "independent" female migrant workers in Shenzhen, China," Working Paper 78, FREIA - Feminist Research Center in Aalborg, Institut for Kultur og Globale Studier 2012. https://doi.org/10.5278/freia.65219031.

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. https://doi.org/10.1162/jeea.2006.4.2-3.552.

Fisman, Raymond and Edward Miguel, "Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets," *Journal of Political Economy*, 2007, *115* (6), 1020–1048. https://doi.org/10.1086/527495.

Giuliano, Paola and Nathan Nunn, "Understanding cultural persistence and change," *The Review of Economic Studies*, 12 2020, *88* (4), 1541–1581. https://doi.org/10.1093/restud/rdaa074.

Goldin, Claudia and Lawrence F Katz, "Career and marriage in the age of the pill," *American Economic Review*, 2000, *90* (2), 461–465. https://doi.org/10.1257/aer.90.2.461.

Gravemeyer, Stefan, Thomas Gries, and Jinjun Xue, "Income determination and income discrimination in Shenzhen," *Urban Studies*, 2011, *48* (7), 1457–1475. https://doi.org/10.1177/0042098010375317.

Heath, Rachel and A Mushfiq Mobarak, "Manufacturing growth and the lives of Bangladeshi women," *Journal of Development Economics*, 2015, *115*, 1–15.

Hellerstein, Judith K., Mark J. Kutzbach, and David Neumark, "Do labor market networks have an important spatial dimension?," *Journal of Urban Economics*, 2014, *79*, 39–58.

Jensen, Robert, "Do labor market opportunities affect young women's work and family decisions? experimental evidence from India," *The Quarterly Journal of Economics*, 03 2012, *127* (2), 753–792.

Kondo, Ayako and Masahiro Shoji, "Peer effects in employment status: Evidence from housing lotteries," *Journal of Urban Economics*, 2019, *113*, 103195.

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. https://doi.org/10.1016/j.chieco.2012.03.004.

Liu, Zhiqiang, "Institution and inequality: the hukou system in China," *Journal of Comparative Economics*, 2005, *33* (1), 133–157. https://doi.org/10.1016/j.jce.2004.11.001.

Mas, Alexandre and Enrico Moretti, "Peers at work," *American Economic Review*, 2009, *99* (1), 112–145. https://doi.org/10.1257/aer.99.1.112.

Meng, Xin and Sen Xue, "Social networks and mental health outcomes: Chinese rural–urban migrant experience," *Journal of Population Economics*, 2020, *33* (1), 155–195.

Munshi, Kaivan, "Networks in the modern economy: Mexican migrants in the US labor market," *The Quarterly Journal of Economics*, 2003, *118* (2), 549–599. https://doi.org/10.1162/003355303321675455.

_ **and Jacques Myaux**, "Social norms and the fertility transition," *Journal of Development Economics*, 2006, *80* (1), 1–38. https://doi.org/10.1016/j.jdeveco.2005.01.002.

Song, Yang, "What should economists know about the current Chinese hukou system?," *China Economic Review*, 2014, *29*, 200–212. https://doi.org/10.1016/j.chieco.2014.04.012.

Wang, Feng and Xuejin Zuo, "Inside China's cities: Institutional barriers and opportunities for urban migrants," *American Economic Review*, 1999, *89* (2), 276–280. https://doi.org/10.1257/aer.89.2.276.

Zhang, Chunni and Yu Xie, "Ethnic enclaves revisited: Effects on earnings of migrant workers in urban China," *Society: Chinese Journal of Sociology/Shehui*, 2016, *2* (2), 214–234. https://doi.org/10.1177/2057150X16633580.

7. Appendix

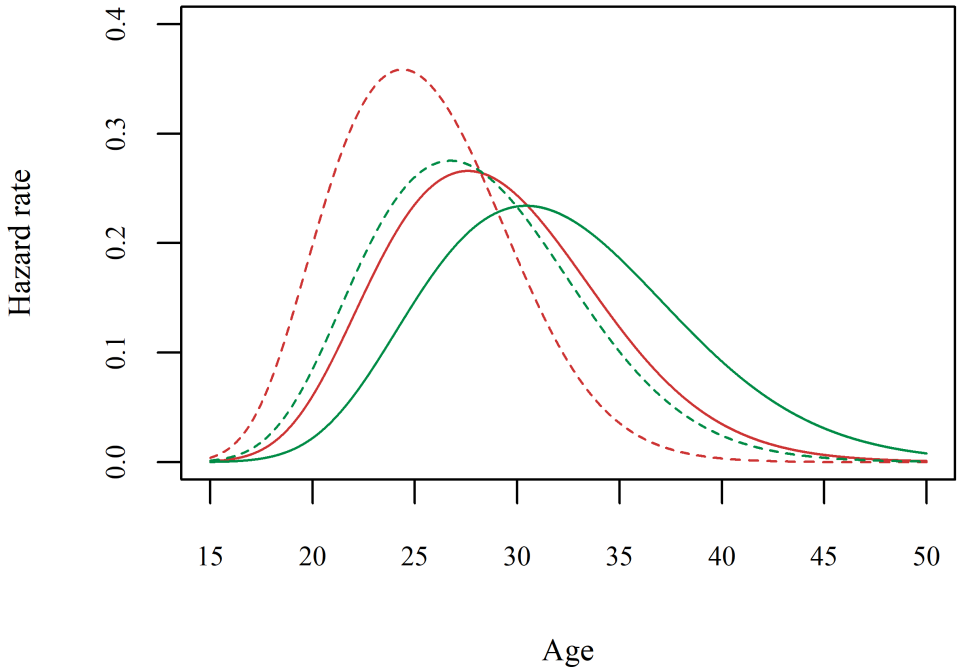


Figure A1: Marriage hazard rate in rural and urban China

— : 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*.

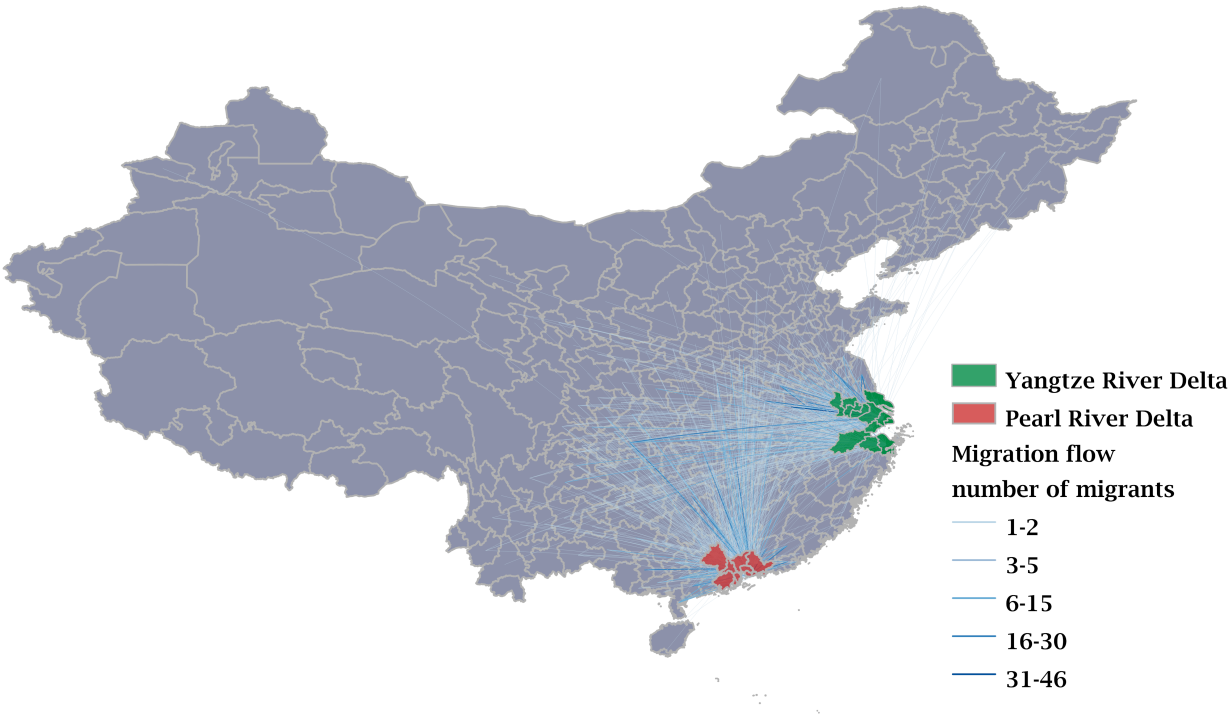


Figure A2: Migration flows to the Yangtze and Pearl River Deltas

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

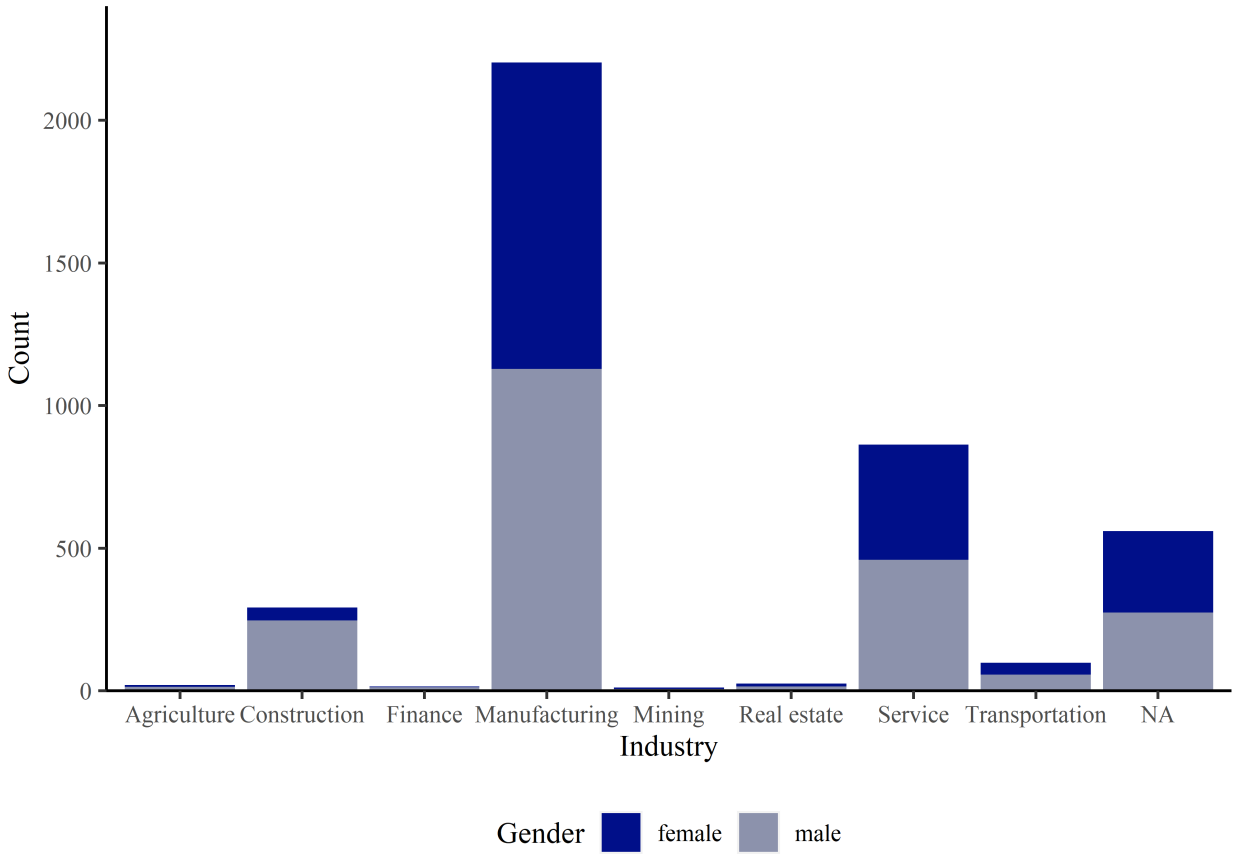
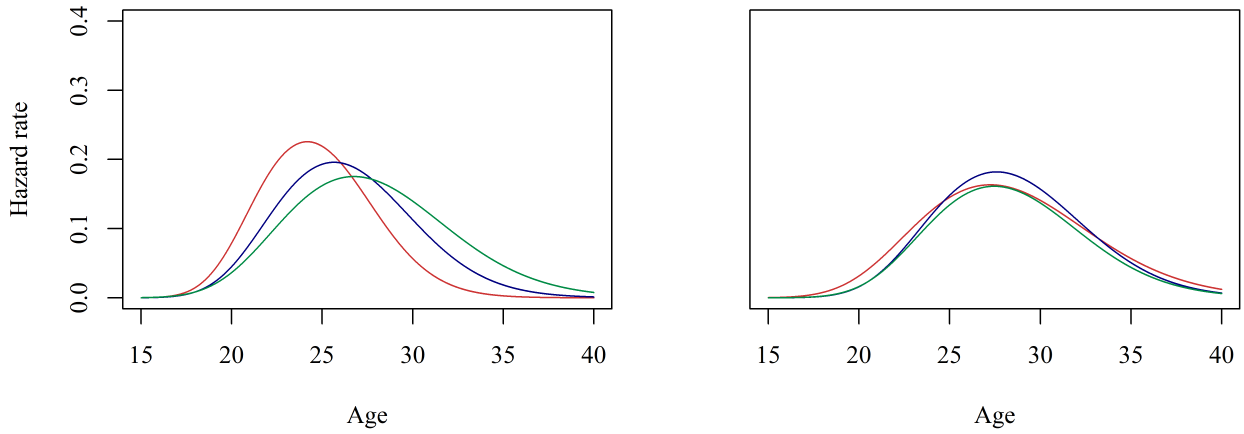


Figure A3: Industrial employment by gender

Source: the 2010 survey data on migrant workers.



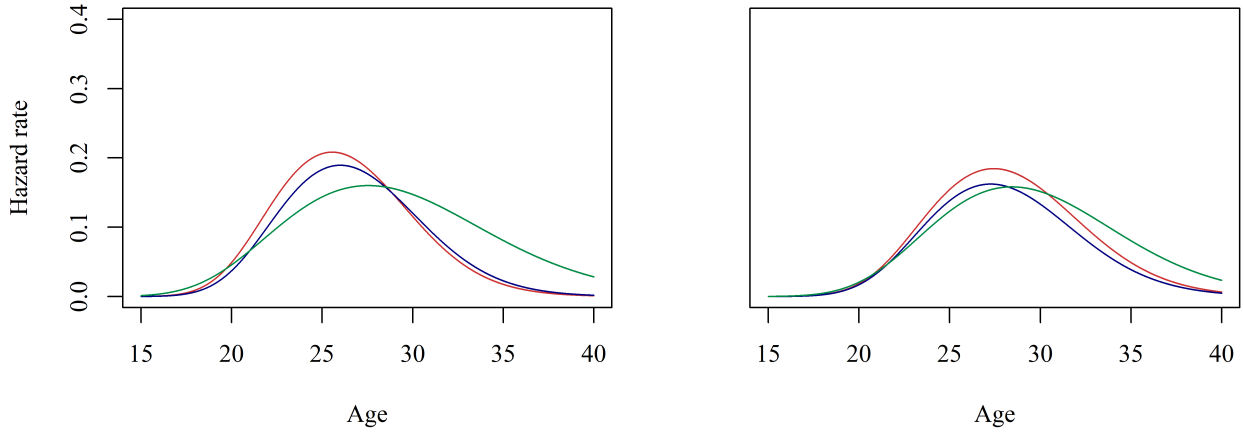
(a) Female (b) Male

Figure A4: Estimated hazard rate: pc_county

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

blue: 0−30% of the co-workers come from the same county

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



(a) Female (b) Male

Figure A5: Estimated hazard rate: pc_province

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

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

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

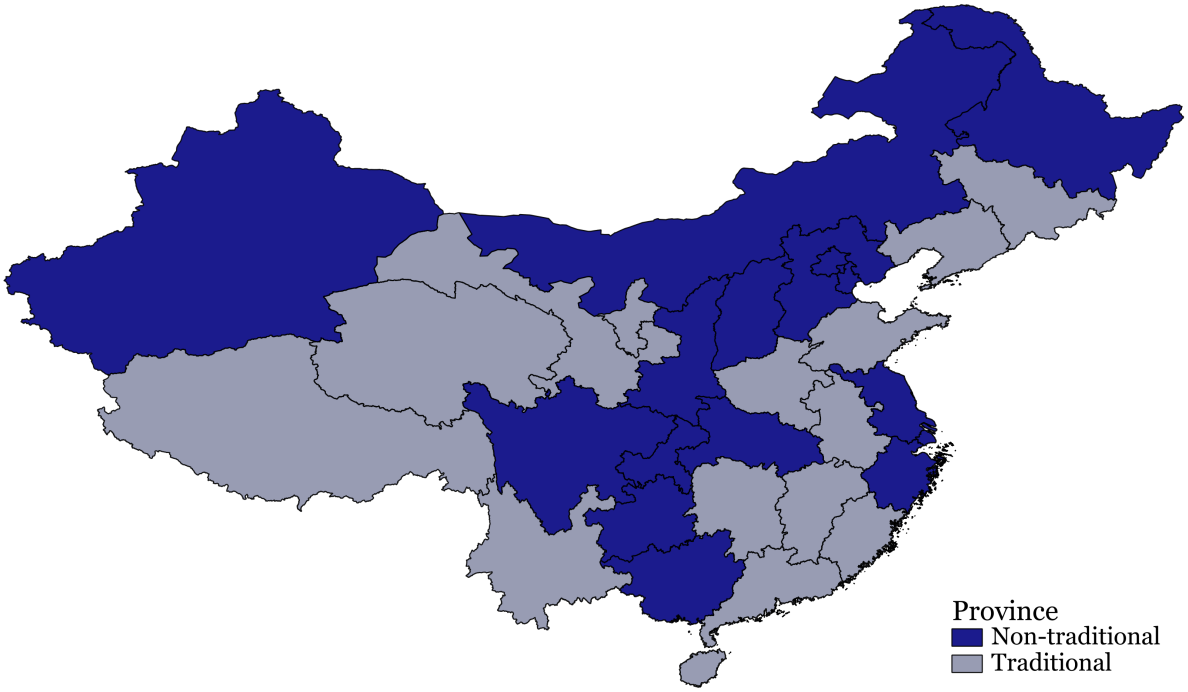


Figure A8: Provinces in China by 'Traditionalness'

Notes: 'Traditionalness' is defined by attitudes towards the role of female in 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 who agree to the following statement that 'marrying a good husband is more important than having a good job for females.' Provinces above the median are defined as traditional provinces and provinces below the median as non-traditional provinces.

Table A1: Attitudes towards females

	Family first (1)	Marriage first (2)
Rural	0.292^{***}	0.127^{***}
	(0.016)	(0.018)
Male	0.058^{***}	-0.014
	(0.019)	(0.019)
Rural$\times$male	-0.074^{***}	-0.062^{***}
	(0.024)	(0.027)
Constant	0.519^{***}	0.426^{***}
	(0.013)	(0.013)
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 the linear probability model to predict *famile first* and *marriage first*. *Rural* is an indicator variable that equals one if the respondent was born and has stayed in the rural region and zero if the respondent was born and has stayed in the urban region. Negative and statistically significant interactions between *rural* and *male* indicate gender-differentiated rural norms. Robust standard errors are given in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table A2: Summary statistics: to continue

	(1) <30% *tongxiang*	(2) >30% *tongxiang*	(3) Difference
	Panel A: Age		
Female	28.53	30.15	1.62
	(8.80)	(9.97)	(0.88)
Male	31.47	33.85	2.37
	(9.51)	(10.81)	(0.77)
Difference	2.94	3.69	0.75
	(0.31)	(1.13)	(1.17)
	Panel B: Migration age		
Female	20.68	19.55	-1.13
	(6.12)	(6.57)	(0.76)
Male	21.13	21.48	0.35
	(6.87)	(7.71)	(0.70)
Difference	0.45	1.93	1.48
	(0.27)	(1.00)	(1.43)
	Panel C: Years since migration		
Female	7.70	8.76	1.06
	(5.75)	(7.80)	(0.90)
Male	10.00	12.34	2.35
	(7.30)	(8.09)	(0.73)
Difference	2.30	3.59	1.29
	(0.27)	(1.13)	(1.16)
	Panel D: Distance from home (kilometres)		
Female	598.90	578.97	-19.94
	(443.49)	(453.38)	(41.35)
Male	630.98	578.46	-52.52
	(448.12)	(420.56)	(31.50)
Difference	32.08	-0.50	-32.58
	(15.45)	(49.76)	(51.99)
	Panel E: Education		
Female	1.30	0.99	-0.32
	(0.90)	(0.85)	(0.08)
Male	1.44	1.13	-0.31
	(0.85)	(0.79)	(0.06)
Difference	0.14	0.15	0.01
	(0.03)	(0.09)	(0.10)

Table A2: Summary statistics: continued

	(1) <30% *tongxiang*	(2) >30% *tongxiang*	(3) Difference
	Panel F: Weekly working hours		
Female	55.24	61.02	5.78
	(15.03)	(16.52)	(1.48)
Male	56.46	59.77	3.31
	(14.96)	(16.46)	(1.20)
Difference	1.22	-1.26	-2.48
	(0.51)	(1.84)	(1.90)
	Panel G: Log hourly wage		
Female	2.00	1.89	-0.12
	(0.45)	(0.66)	(0.06)
Male	2.23	2.22	-0.01
	(0.52)	(0.51)	(0.04)
Difference	0.23	0.33	0.10
	(0.02)	(0.07)	(0.07)

Notes: The table presents the mean and standard deviation (in parentheses) of individual characteristics by gender and concentration of *tongxiang* co-workers (i.e., $<$ 30% co-workers from the same town in column 1 and $>$ 30% co-workers from the same town in column 2). The difference by gender is given 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: Marriage age and friendship—married after migration

	Probability of being married before a certain age			
	Female		Male	
	Baseline	Difference *Tongxiang* friend	Baseline	Difference *Tongxiang* friend
Age	(1)	(2)	(3)	(4)
20	0.058	0.049***	0.023	0.014**
		(0.015)		(0.007)
21	0.114	0.080***	0.053	0.022**
		(0.021)		(0.011)
22	0.194	0.111***	0.105	0.029*
		(0.028)		(0.015)
23	0.293	0.132***	0.180	0.033*
		(0.034)		(0.019)
24	0.399	0.140***	0.274	0.032
		(0.038)		(0.024)
25	0.500	0.136***	0.377	0.028
		(0.040)		(0.027)
26	0.587	0.125***	0.477	0.021
		(0.040)		(0.029)
27	0.658	0.110***	0.566	0.015
		(0.040)		(0.030)
28	0.712	0.096**	0.640	0.010
		(0.040)		(0.030)
29	0.752	0.084**	0.698	0.007
		(0.041)		(0.030)
30	0.781	0.074*	0.741	0.006
		(0.042)		(0.029)

Notes: The estimation is based on the sample of individuals married after migration. The baseline group is individuals whose best friend is not from the same hometown. The difference is the increase in probability relative to the baseline group for migrants whose best friend is from the same hometown. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table A4: Marriage age and friendship—married before migration

	Probability of being married before a certain age			
	Female		Male	
	Baseline	Difference *Tongxiang* friend	Baseline	Difference *Tongxiang* friend
Age	(1)	(2)	(3)	(4)
20	0.288	0.059	0.158	0.041
		(0.037)		(0.029)
21	0.438	0.082**	0.267	0.059*
		(0.039)		(0.035)
22	0.594	0.088**	0.400	0.073*
		(0.040)		(0.039)
23	0.731	0.075**	0.538	0.078*
		(0.038)		(0.040)
24	0.834	0.053*	0.664	0.074**
		(0.032)		(0.037)
25	0.904	0.032	0.767	0.064**
		(0.025)		(0.032)
26	0.947	0.017	0.844	0.052**
		(0.018)		(0.026)
27	0.971	0.008	0.897	0.040**
		(0.012)		(0.020)
28	0.984	0.003	0.931	0.030*
		(0.008)		(0.016)
29	0.991	0.000	0.954	0.022*
		(0.006)		(0.012)
30	0.995	−0.001	0.968	0.017*
		(0.004)		(0.010)

Notes: The estimation is based on the sample of individuals married before migration. The baseline group is individuals whose best friend is not from the same hometown. The difference is the increase in probability relative to the baseline group for migrants whose best friend is from the same hometown. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table A5: Marriage age and same-town co-workers—additional results

	Probability of being married before a certain age					
	Female			Male		
Age	Baseline 0 same origin (1)	Difference >10% (2)	Difference >20% (3)	Baseline 0 same origin (4)	Difference >10% (5)	Difference >20% (6)
20	0.070	0.027	0.050**	0.027	0.012	0.028**
		(0.020)	(0.025)		(0.010)	(0.014)
21	0.128	0.059*	0.091***	0.060	0.020	0.043**
		(0.026)	(0.032)		(0.014)	(0.019)
22	0.207	0.100***	0.135***	0.114	0.029	0.057**
		(0.032)	(0.040)		(0.019)	(0.024)
23	0.302	0.137***	0.168***	0.190	0.037	0.066**
		(0.039)	(0.047)		(0.024)	(0.028)
24	0.404	0.161***	0.181***	0.283	0.042	0.069**
		(0.045)	(0.052)		(0.027)	(0.032)
25	0.503	0.167***	0.176***	0.385	0.044	0.065*
		(0.047)	(0.055)		(0.030)	(0.036)
26	0.591	0.159***	0.158***	0.483	0.043	0.058
		(0.047)	(0.056)		(0.032)	(0.038)
27	0.664	0.143***	0.134**	0.571	0.041	0.051
		(0.047)	(0.056)		(0.033)	(0.039)
28	0.722	0.124***	0.109*	0.644	0.038	0.044
		(0.046)	(0.057)		(0.033)	(0.039)
29	0.767	0.105**	0.086	0.701	0.036	0.040
		(0.046)	(0.058)		(0.032)	(0.038)
30	0.801	0.089*	0.067	0.745	0.034	0.037
		(0.046)	(0.060)		(0.032)	(0.037)

Notes: The estimation is based on the sample of individuals married after migration. The baseline group is migrants with 0 co-workers from the same town. The difference is the increase in probability relative to the baseline group for migrants with over 10% or 20% of the co-workers from the same town. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table A6: Marriage age and same-county co-workers—married after migration

	Probability of being married before a certain age					
	Female			Male		
Age	Baseline 0 same origin (1)	Difference >30% (2)	Difference >50% (3)	Baseline 0 same origin (4)	Difference >30% (5)	Difference >50% (6)
20	0.067	0.067*	0.106	0.027	0.033*	0.029
		(0.035)	(0.117)		(0.017)	(0.020)
21	0.122	0.126***	0.180*	0.058	0.050**	0.044*
		(0.046)	(0.109)		(0.023)	(0.026)
22	0.199	0.187***	0.252**	0.110	0.067**	0.057*
		(0.056)	(0.101)		(0.028)	(0.032)
23	0.292	0.230***	0.298***	0.184	0.078**	0.065*
		(0.064)	(0.095)		(0.033)	(0.038)
24	0.392	0.244***	0.310***	0.273	0.081**	0.067
		(0.068)	(0.091)		(0.037)	(0.046)
25	0.489	0.232***	0.295***	0.371	0.078*	0.064
		(0.070)	(0.088)		(0.041)	(0.053)
26	0.576	0.204***	0.263***	0.466	0.069	0.058
		(0.071)	(0.085)		(0.044)	(0.058)
27	0.650	0.169**	0.225***	0.552	0.059	0.051
		(0.072)	(0.082)		(0.045)	(0.060)
28	0.708	0.135*	0.189**	0.624	0.050	0.046
		(0.073)	(0.080)		(0.045)	(0.059)
29	0.754	0.105	0.158**	0.681	0.043	0.043
		(0.074)	(0.079)		(0.044)	(0.057)
30	0.788	0.080	0.132*	0.724	0.039	0.041
		(0.075)	(0.079)		(0.043)	(0.055)

Notes: The estimation is based on the sample of individuals married after migration. The baseline group is migrants without any co-workers from the same county. The difference is the increase in probability relative to the baseline group for migrants with over 30% or 50% of the co-workers from the same county. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table A7: Marriage age and same-county co-workers—married before migration

	Probability of being married before a certain age					
	Female			Male		
Age	Baseline 0 same origin (1)	Difference >30% (2)	Difference >50% (3)	Baseline 0 same origin (4)	Difference >30% (5)	Difference >50% (6)
20	0.314	−0.064	−0.111	0.159	0.031	0.071
		(0.066)	(0.113)		(0.038)	(0.051)
21	0.450	−0.025	−0.085	0.268	0.055	0.087
		(0.069)	(0.108)		(0.047)	(0.063)
22	0.587	0.023	−0.039	0.400	0.077	0.092
		(0.066)	(0.097)		(0.052)	(0.073)
23	0.708	0.055	0.003	0.537	0.091*	0.086
		(0.056)	(0.084)		(0.053)	(0.077)
24	0.805	0.064	0.026	0.662	0.092*	0.072
		(0.044)	(0.067)		(0.050)	(0.074)
25	0.874	0.057*	0.031	0.764	0.082*	0.056
		(0.032)	(0.049)		(0.043)	(0.065)
26	0.921	0.043*	0.027	0.840	0.068*	0.041
		(0.023)	(0.034)		(0.035)	(0.054)
27	0.951	0.030*	0.020	0.892	0.053**	0.029
		(0.016)	(0.023)		(0.027)	(0.043)
28	0.970	0.020*	0.013	0.927	0.040*	0.020
		(0.011)	(0.015)		(0.021)	(0.034)
29	0.981	0.013	0.008	0.950	0.030*	0.015
		(0.008)	(0.010)		(0.016)	(0.026)
30	0.988	0.008	0.005	0.964	0.023*	0.011
		(0.006)	(0.007)		(0.012)	(0.020)

Notes: The estimation is based on the sample of individuals married before migration. The baseline group is migrants without any co-workers from the same county. The difference is the increase in probability relative to the baseline group for migrants with over 30% or 50% of the co-workers from the same county. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table A8: Marriage age and same-county co-workers—additional results

	Probability of being married before a certain age					
	Female			Male		
Age	Baseline 0 same origin (1)	Difference >10% (2)	Difference >20% (3)	Baseline 0 same origin (4)	Difference >10% (5)	Difference >20% (6)
20	0.067	0.014	0.033	0.027	0.011	0.019*
		(0.017)	(0.022)		(0.009)	(0.012)
21	0.122	0.039	0.062**	0.058	0.020	0.032*
		(0.024)	(0.030)		(0.013)	(0.017)
22	0.199	0.072**	0.096**	0.110	0.031*	0.046**
		(0.031)	(0.038)		(0.018)	(0.022)
23	0.292	0.106***	0.125***	0.184	0.042*	0.058**
		(0.037)	(0.046)		(0.022)	(0.026)
24	0.392	0.131***	0.142***	0.273	0.051**	0.066**
		(0.042)	(0.052)		(0.026)	(0.030)
25	0.489	0.141***	0.145***	0.371	0.055*	0.069**
		(0.044)	(0.055)		(0.029)	(0.032)
26	0.576	0.139***	0.136**	0.466	0.056*	0.068**
		(0.045)	(0.056)		(0.031)	(0.034)
27	0.650	0.128***	0.122**	0.552	0.053*	0.065*
		(0.045)	(0.055)		(0.032)	(0.035)
28	0.708	0.113**	0.105*	0.624	0.049	0.061*
		(0.044)	(0.054)		(0.032)	(0.035)
29	0.754	0.097**	0.088	0.681	0.045	0.057*
		(0.045)	(0.054)		(0.031)	(0.034)
30	0.788	0.083*	0.074	0.724	0.041	0.054
		(0.046)	(0.054)		(0.031)	(0.034)

Notes: The estimation is based on the sample of individuals married after migration. The baseline group is migrants without any co-workers from the same county. The difference is the increase in probability relative to the baseline group for migrants with more than 10% or 20% of the co-workers from the same county. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table A9: Marriage age and identity—married before migration

	Probability of being married before a certain age			
	Female		Male	
Age	Baseline Worker (1)	Difference Farmer (2)	Baseline Worker (3)	Difference Farmer (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: The 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% levels.

Table A10: Correlation between identity and *tongxiang* co-workers

	Dep. var.: Concentration of co-workers from the same town
	(1)
Male	0.115*
	(0.065)
Farmer	0.093
	(0.077)
Farmer × male	−0.012
	(0.109)
Constant	1.000***
	(0.047)
Observations	2,938
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 town (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 are given in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table A11: Spouse workplace and same-town co-workers—additional results

	Dep. var.: 𝟙 (Spouse in the same workplace)					
	Married before migration		Married after migration before current job		Married after current job	
	Female	Male	Female	Male	Female	Male
Age	(1)	(2)	(3)	(4)	(5)	(6)
Pc_town	0.054***	0.040***	0.056***	0.025*	0.088***	0.009
	(0.015)	(0.013)	(0.017)	(0.014)	(0.029)	(0.014)
Constant	0.195***	0.205***	0.151***	0.196***	0.154***	0.208***
	(0.023)	(0.023)	(0.025)	(0.023)	(0.040)	(0.029)
Observations	512	534	333	465	145	296
R-squared	0.029	0.020	0.037	0.008	0.080	0.001

Notes: The estimation utilises the linear probability model to predict whether the spouse is in the same workplace. *Pc_town* is a categorical variable whose value increases with the fraction of co-workers from the same town (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 2,381 respondents, among which 94 lack information regarding *pc_town* and 2 lack gender. Robust standard errors are given in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table A12: Determinants of *tongxiang* co-workers

	Dep. var.: 𝟙 (over 30% *tongxiang*)	
	Female	Male
	(1)	(2)
Education	-0.678^{***}	-0.545^{***}
	(0.237)	(0.191)
Age work	-0.015	-0.021
	(0.052)	(0.034)
First job	0.163	0.609^{*}
	(0.412)	(0.350)
Referral	0.933^{***}	0.914^{***}
	(0.342)	(0.289)
Tongxiang ratio	5.394	6.859
	(5.461)	(4.940)
Farmer	-0.183	0.301
	(0.387)	(0.314)
Occupation dummies	Yes	Yes
Cohort fixed effects	Yes	Yes
Origin province fixed effects	Yes	Yes
Destination city fixed effects	Yes	Yes
Observations	466	587

Notes: The table estimates propensity scores using logistic regression. The dependent variable is an indicator for having over 30% co-workers from the same town. *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 began working for the current job. *First job* is an indicator for whether the job is the first job of the individual. *Referral* indicates whether the job is obtained through referral. *City (tongxiang)* is the ratio of same origin migrant workers over the total number of migrant workers in the same destination city, proxied from the 2000 population census. *Farmer* is an indicator for self-identification with farmer or rural origin. The estimation is based on observations with none of the co-workers from the same town or over 30% co-workers from the same town. Robust standard errors are given in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table A13: Marriage age and same-town co-workers—matched sample

	Probability of being married before a certain age			
	Female		Male	
Age	Baseline 0 same origin (1)	Difference >30% (2)	Baseline 0 same origin (3)	Difference >30% (4)
20	0.086	0.070	0.075	0.012
		(0.089)		(0.039)
21	0.148	0.123	0.152	−0.006
		(0.095)		(0.054)
22	0.229	0.176*	0.259	−0.037
		(0.103)		(0.068)
23	0.322	0.213*	0.381	−0.070
		(0.110)		(0.080)
24	0.420	0.227**	0.497	−0.092
		(0.114)		(0.087)
25	0.514	0.219*	0.594	−0.098
		(0.114)		(0.090)
26	0.597	0.197*	0.669	−0.089
		(0.114)		(0.090)
27	0.667	0.169	0.722	−0.071
		(0.115)		(0.088)
28	0.723	0.141	0.759	−0.048
		(0.117)		(0.087)
29	0.768	0.115	0.783	−0.025
		(0.120)		(0.087)
30	0.801	0.093	0.798	−0.003
		(0.123)		(0.087)

Notes: The estimation is based on the sample of matched individuals married after migration. The baseline group is migrants with 0 co-workers from the same town. The difference is the increase in probability relative to the baseline group for a subsample of migrants with over 30% of the co-workers from the same town. Each individual in the subsample is matched with an individual in the baseline group using propensity score matching. Bootstrap standard errors with 500 replicates are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.