How do Heterogeneous Social Distances affect the Neighborhood Effect in Rural–Urban Migration?

Empirical Evidence and Policy Simulation from China

(This version, 2010-7-4)

Zhao CHEN

China Center for Economic Studies, and School of Economics, Fudan University

220 Handan Road, Shanghai 200433, China

zhaochen@fudan.edu.cn

Shiqing JIANG

Department of Economics, Fudan University

220 Handan Road, Shanghai 200433, China

shiqing.jiang@gmail.com

Ming LU*

School of Economics, Fudan University

220 Handan Road, Shanghai 200433, China

lumingfd@gmail.com

Hiroshi SATO

Graduate School of Economics, Hitotsubashi University

2-1 Naka, Kunitachi, Tokyo, Japan

satohrs@econ.hit-u.ac.jp

satohrs@gol.com

Tel&Fax: +81-(0)42-580-8265

*Corresponding author: Ming Lu

We thank Jack W. Hou, Hubert Jayet, Mark Rosenzweig, Yves Zenou, Yaohui Zhao and seminar/conference participants at K. U. Leuven, Fudan University, University of Lille 1, Xiamen University, and Zhejiang University for their useful comments. We gratefully acknowledge financial support for the CHIP 2002 survey provided mainly by the Ford Foundation, the Swedish International Development Cooperation Agency (SIDA) and partly by the Japan Society for the Promotion of Science (JSPS) and Hitotsubashi University. Financial support from the Chinese Ministry of education, the Shanghai Leading Academic Discipline Project (B101), JSPS Grants-in-Aid for Scientific Research (No. 18203018), the Heiwa Nakajima Foundation, and the Global COE Program "Research Unit for Statistical and Empirical Analysis in Social Sciences," Hitotsubashi University, and the Fukino Project, Hitotsubashi University, is also gratefully acknowledged.

How do Heterogeneous Social Distances affect the Neighborhood Effect in Rural–Urban Migration?

Empirical Evidence and Policy Simulation from China

Abstract: In this paper, we use the "2002 Chinese Household Income Project Survey" (CHIPS2002) data to examine how heterogeneous social distances affect the neighborhood effect in the rural–urban migration decision in China. We find that the neighborhood effect, measured by the village migration ratio, significantly increases the individual probability of outward migration. We also find that the magnitude of the neighborhood effect is heterogeneous, depending on social distance that is modeled as a function of the strength and type of social interactions with other villagers. Interactions in information sharing and modest labor market interaction can increase the magnitude of the neighborhood effect, while frequent interactions in mutual help in farm work in busy seasons, will impede the positive role of the neighborhood effect. The existence of a heterogeneous neighborhood effect and multi-equilibria of migration has rich policy implications. For policy makers to encourage rural–urban migration, it is feasible to directly increase village migration ratio or to have higher neighborhood effect in migration by increasing information sharing or substituting within-village mutual help in farm work by providing more efficient services to rural residents. However, according to the policy simulation with the model of neighborhood effect, only a "big-push" in institutional reform can help China escape the low equilibrium of labor mobility and urbanization.

Keywords: labor migration, neighborhood effect, social distance, social interaction, social multiplier

JEL Classification: J61, O15, R23

1 Introduction

Rural-to-urban migration and hence urbanization are key symbols of economic development. Especially for developing countries, policies promoting migration from the countryside to cities are structural forces for sustainable growth. However, except for well-known migration-facilitating measures such as infrastructure and human capital investment, are there other policies to promote migration? This encourages us to explore further the determination of labor migration. In this paper, we attempt to answer two questions. First, how does the neighborhood effect—interdependence in decision making—affect the migration decisions of rural residents in China? Second, how do heterogeneous social distances affect the neighborhood effect in migration?

Using CHIPS2002 (2002 Chinese Household Income Project Survey) data, we find strong evidence that the neighborhood effect exists in the outward migration decision in rural China. In the presence of the neighborhood effect, other policies, such as increasing the education of rural residents, have larger effects than previous estimates because of the spillover of the neighborhood effect or the so-called social multiplier (Glaeser *et al.*, 2002).

We also develop a model to incorporate heterogeneous neighborhood effect. In our model, the magnitude of the neighborhood effect depends on whether within-neighborhood social interactions consume time during the outward migration decision. Empirically we interact the neighborhood effect with different types of social interaction and obtain some interesting results: higher interaction frequencies in information sharing with other villagers and modest mutual help in the labor market will enhance the magnitude of the neighborhood effect, while higher interaction frequencies in time-consuming local labor exchange activities will reduce the positive role of the neighborhood effect.

The presence of the neighborhood effect implies multiple equilibria in labor migration, either a low-level equilibrium, which may be the current situation in China's rural-urban migration as our model shows, or a high-level one. If social interactions of different types and frequencies affect the neighborhood effect in migration, new social policy tools can be utilized to push ahead rural–urban migration and urbanization. Policy makers seeking to encourage outward migration can either promote information sharing among villagers or substitute within-village mutual help in farm work by providing more efficient services to rural residents. When migration is trapped in the low equilibrium with neighborhood effect, those social multiplier enhancing policy tools, together with direct migration promoting policies, may not be enough to help rural people escape the low equilibrium of migration ratio. Therefore, the policy makers should eliminate urban-rural divide and urban-biased policies and accelerate social integration between rural and urban areas. This is another policy implication in our paper: the institutional “big-push” in China.

The rest of this paper is organized as follows. Section 2 reviews studies on labor migration in China and the emerging studies of network effects in migration decision. Section 3 establishes a simple model to demonstrate how the neighborhood effect is affected by heterogeneous social distances. Section 4 describes the data and Section 5 presents the econometric model and empirical findings. In Section 6 we do some robustness checks. Section 7 contains the policy simulation of how rural-urban migration can be enhances based on our empirical model, and the final section concludes with policy implications.

2 Literature Review

Migration has been understood as both an individual and a household decision, so that the factors affecting the labor migration decision are a group of individual and family characteristics. Many empirical studies have explored migration determination. In migration studies for China, the classical framework is also applicable. Using cross-sectional data in the Sichuan rural areas, Zhao (1999a; 1999b) finds evidence consistent with findings in other countries: male laborers have a higher probability of outward migration, while aging and more household land area will significantly decrease the probability of migration. Zhu (2002) finds that the income gap between farming and nonfarm activities will affect the migration decision, which is consistent with the Harris–Todaro model. Cai *et al.* (2003) discover that although the income gap between west and east China is greater than that between middle and east China, migration is more prevalent from middle to east than from west to east, which seems to contradict the Harris–Todaro prediction but still can be explained by distance effects.

Recent studies add the role of social networks to the analysis of the migration decision. Munshi (2003) finds that networks play a significant role in helping rural Mexican residents migrate to the US. Mckenzie and Rapoport (2007) argue that with the expansion of migration networks, more poor families can engage in migration, thus reducing rural inequality. Using Chinese data, Zhang and Li (2003) find that rural residents have higher probability in nonfarm employment if their family has social ties outside the village. Bao *et al.* (2007) find that province-to-province migration rates rise with the size of the migrant community in the destination province. Zhao (2003) shows that larger numbers of local experienced migrants will significantly increase the migration probability of villagers in the same village, and she argues this is the result of job information sharing among villagers. Although findings about the role of social networks extend our understanding of labor migration, all of these empirical studies only consider homogeneous network effects in migration decision.

Before we go further to focus on the heterogeneity of neighborhood effect, we first explain why we use the term “neighborhood effect”, instead of “network effect”, in this paper. Network can be distinguished as within-community (bonding) and cross-community (bridging) network. Relatives and friends outside the community constitute bridging network, while within-community network is the social network that provides information to reduce migration costs. Focusing on within-village social interaction and network, we prefer to use the terminology, neighborhood effect, while we also control for out-of-village network in the empirical model. However, as Bauer *et al.* (2002) point out, peers in the community also contribute to herd effect, another name of neighborhood effect and within-community interdependence of behavior, even in the presence of migration networks. In fact, in rural China, villagers form strong social and economic ties in their daily lives, so the behavior of a person would be affected by his or her village neighbors. Bauer *et al.* (2002) and Araujo *et al.* (2004) find strong evidence that neighborhood effects exist in labor migration from rural Mexico to urban areas and from Mexico to the USA, respectively. In our study, using data from rural China, we further confirm the existence of the neighborhood effect in labor migration.

However, the existing literature of neighborhood effect in labor migration has not told how the magnitude of neighborhood effect and the corresponding social multiplier are determined. In our study, we argue that different people have heterogeneous neighborhood effects due to their type and frequency of social interaction with their village peers. The peer effect, another often-used term for interdependence of behavior, is found in many social and economic behaviors, although the terminology differs according to research contexts (see Durlauf (2004) for an exhaustive literature survey). It is not a new idea that people are interdependent in decision making, but empirically constructing the peer group was once formidable because of the lack of subtle microdata. Therefore, the measurement of the peer effect is always at the core of research. Early research only roughly measured the peer effect as the average outcome in a group. For example, Evans *et al.* (1992) defined the class as the peer group and observed the effect of class average education scores on the probability of becoming an unmarried mother. Recent studies used unique data to identify friend networks and thus peer groups (Ballester *et al.*, 2006; Calvó-Armengol and Zenou, 2005; Patacchini and Zenou, 2008). Some even used the correspondence frequency to measure friendship distance (Marmaros and Sacerdote, 2006). In our research, we assume people in the same village play with all the villagers, but each individual has a heterogeneous social distance from the other villagers. Empirically, we construct the heterogeneous neighborhood effect using interaction terms between the neighborhood effect and social interaction frequencies.

3 A Model of Social Distance and Neighborhood effect

Our model is mainly based on network models such as Ballester *et al.* (2006) and simplifies some of their assumptions. In our model, we explicitly assume that social distance is a function of the type and frequency of social interactions.

There are N individuals in a village. The network N = {1, . . . , n} is a finite set of agents. The n-square matrix G of a network *g* keeps track of the connections in this network. Here, we simply assume each individual is friends with everybody else in network G. Ballester *et al.* (2006), Calvó-Armengol and Zenou (2005), and Patacchini and Zenou (2008) discuss more general cases where individuals face different peer networks; however, we dismiss this idea because of the data limitation. Every person in our model has heterogeneous attitudes toward the behavior of peers, thus is a different social distance from the network. That is to say, in matrix G, $g_{ik} \neq g_{jk}$ if $i \neq j$. It also implies that friendship is not a reciprocal relationship. We also set $g_{ii} = 0$.

Using matrix denotation, we assume:

$$G = \begin{pmatrix} g_{11} & g_{12} \cdots & g_{1n} \\ g_{21} & g_{22} \cdots & g_{2n} \\ \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} \cdots & g_{nn} \end{pmatrix} = \begin{pmatrix} 0 & g_1 \cdots & g_1 \\ g_2 & 0 \ldots & g_2 \\ \vdots & \ddots & \vdots \\ g_n & g_n \cdots & 0 \end{pmatrix} = \begin{pmatrix} g_1 & 0 \ldots & 0 \\ 0 & g_2 \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 \cdots & g_n \end{pmatrix} \begin{pmatrix} 0 & 1 \ldots & 1 \\ 1 & 0 \ldots & 1 \\ \vdots & \ddots & \vdots \\ 1 & 1 \cdots & 0 \end{pmatrix}.$$

Group influence/neighborhood effects are expressed as:

$$\frac{\sum_{i=1} g_{ij} m_j}{\sum_{i=1} g_{ij}} = \frac{g_i}{n-1} \sum_{i=1} m_j = g_i \bar{m} \,. \tag{1}$$

g_i measures the social distance to the network. The standard neighborhood effect model implies g_i equals a constant; thus, g_i cannot be estimated. Here we write $g_i = J + \lambda_i s_i$, with $J, \lambda_i > 0$ to capture the heterogeneous social distances of different individuals to other villagers. J is a constant, while s_i is the individual interaction frequency with other villagers, λ_i is the parameter linking social interaction frequency to social distance. It is natural to assume that if one person is more involved in the local interaction, the social distance is nearer, so that the neighborhood effect will be larger.

Because more local interactions s_i will squeeze out the time that can be allocated to outward migration, we standardize s_i and m_i to be continuous and $m_i, s_i \in [0,1]$, and simply assume:

$$s_i + m_i \leq 1, \tag{2}$$

where the total time available is standardized as 1. Social interactions with neighbors make equation (2) binding or loose depending of the type and frequency of interactions. Usually, when people exchange labor market information, it's not time consuming. The utility that individual i obtains from outward migration is:

$$U(m_i) = a + b_i m_i - c m_i^2 + g_i \cdot m_i \cdot \overline{m} \,, \tag{3}$$

where a>0 is a constant, $b_i > 0$ is the linear marginal benefit of migration. The cost of marginal is simplified as a squared term of migration time, with c>0. The last term on the righ-hand-side is the social utility that depends on the average migration time of the peer villagers. g_i captures the heterogeneous social distance of villager i.

If equation (2) is not binding, we may insert equation $g_i = J + \lambda_i s_i$ into (3) and we obtain:

$$U(m_i) = a + b_i m_i - cm_i^2 + [J + \lambda_i s_i] \cdot m_i \cdot \overline{m} \tag{4}$$

An individual optimizes the time allocated to outward migration work (First order condition):

$$dU(m_i)/dm_i = b_i - 2cm_i + \overline{m}(J + \lambda_i s_i) = 0 \tag{5}$$

So,

$$m_i = \frac{b_i + \overline{m}(J + \lambda_i s_i)}{2c} \tag{6}$$

where m_i is positively correlated with $\overline{m}$ and s_i amplifies the interdependence between m_i and $\overline{m}$.

However, for some people who mutually help in the busy season, social interactions are time-consuming, so equation (2) is binding. Thus, we may insert equations $g_i = J + \lambda_i s_i$ and m_i=1-s_i into (3), and we obtain:

$$U(m_i) = a + b_i m_i - cm_i^2 + [J + \lambda_i (1 - m_i)] \cdot m_i \cdot \overline{m} \tag{7}$$

An individual optimizes the time allocated to outward migration work (First order condition):

$$dU(m_i)/dm_i = b_i - 2cm_i + \overline{m}[J + \lambda_i - 2\lambda_i m_i] = 0 = G(m_i, \overline{m}) \tag{8}$$

The following second-order condition guarantees an interior solution:

$$\partial G(m_i, \overline{m})/\partial m_i = -2c - 2\lambda_i \overline{m} < 0, \tag{9}$$

$$\partial G(m_i, \overline{m})/\partial \overline{m} = J + \lambda_i - 2\lambda_i m_i \tag{10}$$

From the derivation calculus of implicit functions, we obtain:

$$\frac{dm_i}{d\overline{m}} = -\frac{G_{\overline{m}}'}{G_{m_i}'} = \frac{J}{-G_{m_i}'} + \frac{\lambda_i}{-G_{m_i}'} + \frac{2\lambda_i m_i}{G_{m_i}'} \tag{11}$$

$\frac{dm_i}{d\overline{m}}$ is the core concept in our paper: the neighborhood effect. Here, it can be decomposed into three parts. $\frac{dJ}{-G_{m_i}'} > 0$, where $G_{m_i}' < 0$ is guaranteed by (9), corresponds to the standard linear neighborhood effect, and we can see that when the village migration ratio increases, the individual allocates more time to migration work. $\frac{\lambda_i}{-G_{m_i}'} > 0$ represents the positive effect of social interaction on the neighborhood effect. When an individual increases his or her social interaction strength λ_i, the social distance is shortened and the neighborhood effect rises. $\frac{2\lambda_i m_i}{G_{m_i}'} < 0$ is the third part and it shows that when individual i spends more time in outward migration, the time for social interaction will be squeezed out due to time constraint, and thus the effect of peer behavior decreases. In summary, equation (11) can lead to two hypotheses: (1) individual migration time is positively related to group mean migration time; and (2) combining terms 2 and 3, social interaction can have either positive or negative effects on the neighborhood effect, depending on whether social interaction greatly reduces outward migration time.

4 Data Description

The data used in our research are from the 2002 Chinese Household Income Project Survey (CHIP 2002) collected by the Chinese Academy of Social Science. Survey data are from 121 counties, 961 administrative villages, 9200 households and 37,969 individuals. The sampling frame for the survey is a subsample of the official rural household survey conducted by the National Bureau of Statistics (NBS).[1] The questionnaires were collected in February 2003, the Chinese Lunar New Year when almost all the Chinese including rural migrants return home and celebrated the spring festival together. Therefore, the survey captures information of all members of rural households including outward migrants. The data contain

[1] The stratified sampling of the NBS rural household survey followed two steps. First, sample administrative villages were directly selected in each province according to income level, and second, sample households (generally 10) were chosen from each sample village. For details of the sampling framework and sampling method of the CHIP 2002 survey, see Gustafsson, Li, and Sicular (2008).

individual information, such as sex, age, education, job status, family information such as family structure, family economic condition and village geography, village population and economic conditions. More importantly, it also includes information on family social interactions with other villagers.

The explained variable "migrant or not" is a 0–1 dummy variable. Defining the migration variable as discrete makes identification possible in the presence of the reflection problem. The reflection problem coined by Manski (1993) is a difficulty in estimating the neighborhood effect. Simply speaking, in a linear model, individual characteristics affect one's decision linearly. The average characteristics and average choice (measurement of the neighborhood effect) are perfectly collinear so that parameters cannot be identified if we control them simultaneously in the regression model. However, Brock and Durlauf (2001) prove that the reflection problem can be avoided in the nonlinear model. Personal characteristics influence the choice nonlinearly in a nonlinear model such as probit or logit, so that they are not linearly correlated if we put them together in the regression model.

In CHIPS2002, individuals reported the days away from their family in a year. Because of the data limitation, we consider the urban areas in China as the only migration destination in our paper and do not differentiate between migration locations. We follow Zhao (2003) and define an individual as a migrant if he or she lives away from home more than 180 days in a year. Obviously, leaving family for more than six months in a year does not necessarily mean a person is a migrant. Therefore, with personal job status information in the questionnaire, we dropped all long-term out-of-village students, as well as the nonfarm employees who work in the township enterprise outside the village. The largest change in our sample is that we only include the working-age population, i.e., observations of male individuals aged 16–60 and females aged 16–55 according to the official definition in China. We also drop observations whose important variables are missing. Finally, 16,401 observations remain for the analysis.

From the CHIPS2002 questionnaire data, we obtain information on the village population and village migrant numbers. We calculate the village migration ratio using the following equation:

$$village\ migration\ ratio = \frac{no.of\ \ village\ migrants\ \text{-}\ no.of\ \ family\ migrants}{village\ \ population\ \text{-}\ \ family\ \ population}\ . \tag{12}$$

This is the measure of village peer behavior in our paper. For a certain household, the village migration ratio is calculated as the ratio for other villagers to exclude the effects from one's own family. Another advantage of this is that we can have variances in “village migration ratio” among different households. This definition is close to the one used in Zhao (2003), who used the absolute number of migrants in a village to measure the network effects of migration. While using the migration ratio of village peers to capture neighborhood effects, we also control the number of friends and relatives outside the village as the household's bridging network.

Social interactions between one's family and other villagers are assumed to affect social distance and then the magnitude of neighborhood effects in our study. Here, we categorize the social interactions into two types: interactions in information sharing, and in labor markets. In Chinese rural areas, the widening rural–urban income gap makes outward migration an effective way of income earning. Since urban and rural labor markets are segmented, and many migrants seek urban jobs across provinces, rural residents exchange information for better job destination and higher income. The market for labor services is still so unfledged that rural residents cooperate a lot in labor-sharing activities. In the CHIPS2002 data, a series of questions record the social interaction strengths of a family with their relatives and neighbors in “exchange information on employment”, and “mutual help during busy seasons”. The answers to these questions are discrete: (1) very frequently, (2) often, (3) just so-so, (4) sometimes, and (5) none/few. We construct two variables with cardinal order of information sharing and mutual help strength. On that basis, we interact the two social interaction variables with the village migration ratio (neighborhood effect), and use the interaction terms to capture the heterogeneous neighborhood effect. We use the continuous measurement of social interaction as our baseline model. The discrete measurements of social interactions will be used for the robust checks.

All the explanatory variables are listed in Table 1, and the basic statistical descriptions are in Table 2. We can see from Table 2 that among the 16,401 rural laborers, 2675 individuals participated in outward migration in 2002, which indicates an overall migration ratio of 16.31%. Even in the basic statistics, we can see some differences between migrants and nonmigrants. Fewer women are employed in outward migration, and in the migrants sample, 50.24% individuals are unmarried, compared with 25.21% in the nonmigrants sample. The outward migrants are much younger with an average age of 27.1, lower than the 36.1 years in the nonmigrants' sample. All these explanatory variables are controlled for in our regression model. However, in the regression analysis, we focus on the magnitude and direction of the neighborhood effect and the interaction term between social interaction strength and the neighborhood effect.

<Insert Table 1 and Table 2 here.>

5 Regression Model and Result

Based on the theoretical model, we define a latent variable Y^*, thus the latent utility function is:

$$Y_i^* = X_i\beta + g_i\bar{M}_i + \varepsilon_i \,, \tag{13}$$

with:

$$\begin{aligned} M_i &= 1 \quad if \quad Y_i^* \geq \alpha_i \\ &= 0 \quad otherwise \end{aligned} \tag{14}$$

Here M_i represents the migration decision of household i, and equals one if the utility from migration is greater than some subjective threshold, which we denote as α_i. Thus we can write a probit model as follows:

$$\begin{aligned}\Pr(M_i = 1) &= \Pr(Y_i^* \geq \alpha_i) = \Pr(X_i\beta + g_i\bar{M}_i + \varepsilon_i \geq \alpha_i) \\ &= \Pr(\varepsilon_i \geq \alpha_i - X_i\beta - g_i\bar{M}_i) = \Phi(-\alpha_i + X_i\beta + g_i\bar{M}_i)\end{aligned} \tag{15}$$

The marginal effect of the neighborhood effect is $\partial \Pr(M_i = 1) \big/ \partial \bar{M}_i = \Phi' \cdot g_i$, where Φ is the cumulative distribution function (CDF) of a standard normal distribution, since ε_i is assumed to follow a standard normal distribution, and $g_i = J + \lambda_i s_i$ as explained before.

To fit our data to the model, we establish the following probit model to explore the determinants of outward migration:

$$P(Y_{ijk} = 1) = \Phi(X_{ijk}\beta + J\bar{M}_{jk} + \bar{M}_{jk} \times \sum \lambda_s s_{jks}) . \tag{16}$$

Equation (16) is the determination function of outward migration probability. i, j and k represent the individual, family and village, respectively. X_{ijk} is a vector of individual, family and village characteristics variables. $\bar{M}_{jk}$ is the village migration ratio and it is the measurement of the neighborhood effect that we are mostly concerned with in our paper. Notice that each household j within the same village k may have a different "village migration ratio" since one's own family are excluded from his household's "village migration ratio". s_{jks} are the social interactions of a family with their relatives and neighbors, where subscript s denotes either "exchange information on employment" or "mutual help during busy seasons" interactions.

<Insert Table 3 here.>

The regression results are reported in Table 3. Equation (1) is the baseline regression, where only individual and household variables are included as explanatory variables. Equation 2 adds the village migration ratio (neighborhood effect) and its interaction terms with social interactions in information exchange and labor market. R-squared increases from 0.1828 to 0.2136, indicating an unneglectable neighborhood effect in migration. Worrying about the potential missing-variable-bias in the estimation of the neighborhood effect, we put the village migration ratio in 1998, a variable based on retrospective information in our survey, in equation 3. Equation 4 controls for village-level characteristics. Equation 5 adds the county dummies to avoid remaining missing-variable-bias in the estimation.

In equations (2) to (5), 3 results have positive signs for the coefficients of current neighborhood effects. Compared with equation (2), the coefficients of current neighborhood effects drop substantially from 1.503 to 0.650 in equation (3), while the coefficient of migration ratio in 1998 is highly significant and 3 times that of neighborhood effect in 2002 in equation (3) and (4). This is consistent with the finding of Munshi (2003): past network is more significant than current network. Intuitively, more information is accumulated in network as time passes. In equation (5), when county dummies are controlled for, current neighborhood effect becomes insignificant and neglectable in magnitude, but the past neighborhood effect is still highly significant and much stronger compared with current neighborhood effect.

We are more interested in the interaction terms of social interactions and neighborhood effect. In equation (2) to (5), all the interaction terms are highly significant. Observing the significance level and the size of the coefficients of the interaction terms, we may get the following conclusions: (1) For information exchange, the more frequently people interact, the stronger the neighborhood effect is in migration decision. (2) For mutual help in the labor market, the more frequently people interact, the weaker the neighborhood effect is in migration decision. The difference in the signs of the two interaction terms can be explained as follows: Information sharing enhances neighborhood effects. However, labor exchange is time consuming, squeezes the time for migration, and mitigates neighborhood effects in migration, although labor exchange reduces social distance.[2] However, the negative sign of the interaction term between labor exchange and neighborhood effect is because many people face time constraint for outward migration when they have helped their neighbors in the busy season. If this is the case, we need to seriously consider how outward migration and mutual help in the labor market are simultaneously determined. Fortunately, with the discrete measurement of mutual help in the busy season, we find that not everybody interacts with his neighbors in the labor market frequently. We will show that modest mutual help in the labor market enhances neighborhood effect in migration decision in the next section.

Nevertheless, the fact that within-community social interaction may play a negative role in labor migration has previously been neglected in the literature except by Narayan (1999) and Alesina and Giuliano (2007). Narayan (1999) separates social capital into within-community "bonding" social capital and between-community "bridging" social capital. He argues that if a community has higher bonding social capital, it will have higher internal welfare, but they will also lose many outside job opportunities. Alesina and Giuliano (2007) find that stronger family ties can decrease the geographical mobility of individuals. However, social interactions in the labor market are perhaps the spontaneous substitutes of an unfledged labor service market in the rural areas. Therefore, we can expect that with the economic development, more and more emerging labor market services will make the interactions of rural residents in the labor exchange more efficient, thus promoting neighborhood effect in outward migration.

Based on equation (4), for a representative person who has a medium level social

[2] In the questionnaire, we also know the number of days that a household is involved in within-village labor exchange. If we substitute the labor exchange frequency measurement by the days of labor exchange, the results won't change. However, the number of observations is reduced to about 10,000, due to missing of information.

interaction and all the other characteristics at the mean level, the marginal neighborhood effect is calculated as 0.1413. Not surprisingly, we can see that a one percentage increase in the village migration ratio increases the individual probability of outward migration by 0.1413%. This is termed a “social multiplier” in the literature (Glaeser *et al.*, 2003).

This result has confirmed the existence of the neighborhood effect in the outward migration decision in rural China. However, the positive relationship between the village migration ratio and individual migration probability also indicates the potential danger of low equilibrium in outward migration: if village peers are less inclined to migrate because of institutional obstacles such as urban–rural segmentation in China, the negative effects will also be amplified by the social multiplier.

All the other coefficients are consistent with the findings in previous studies. We interpret the estimation based on equation (4).

(1) Individual characteristics will significantly affect the migration decision. Women are less inclined to migrate, with a probability that is 4.45% lower than males. Being married will greatly decrease the probability of outward migration by 11.76%. Age has an inverse U-shaped relationship with the migration probability. A laborer has a maximum migration probability at the age of 33, and beyond this age the marginal effect of age is decreasing. All these findings are consistent with existing empirical results. Zhao (2003) finds that all levels of education are insignificant in migration determination, which is in contrast to our result that education levels are all significantly positive with illiteracy as the reference point. However, the influence of education is nonlinear; villagers who receive a junior high school education have the highest probability of migration, 7.13% higher than the illiterate group, while villagers with a primary education have the second-highest probability of migration, 5.25% higher than the illiterate group. If a person has a higher education level, the probability of outward migration is moderately higher than the illiterate group. The probability of

migration for villagers with a technical school education or higher is 4.75% higher and for senior high school education is 4.9% higher. Our findings seemingly imply that higher education for the rural residents may be at the expense of a lower outward migration probability. For the policy makers, there may exist some “optimal” education level for the purpose of rural–urban migration. However, we need to be cautious about this conclusion: one possible explanation for the nonlinear “education return” is that the higher education receivers have permanently stayed in the city areas after gaining their urban *Hukou* (residence registration), so that they are not included as “migrants” in the rural sample. Another explanation is that better-educated workers are more likely to participate in local nonfarm employment (Zhao, 1999a, 1999b; Liang and White, 1997), which is included as nonmigration in our regression.

(2) Family characteristics also significantly affect the migration decision. For an additional laborer in a family, the individual probability of migration increases by 4.01%. Meanwhile, if a family has more arable land, the probability of outward migration declines because of the labor substitution between local farming and migration work. The family structure can also influence the individual migration decision. Families that have one additional child aged between 6 and 12 have a 0.98% lower migration probability. One additional older person does not significantly influence the migration decision because they can be either an effective laborer in the household or a person to be taken care of in rural China. Being aware of the existence of household-specific network besides the village-level network of migration, we also control for other measurement of the household social network that can also promote labor migration. If one’s family has social ties outside the village or has kin as village cadre, the probability of outward migration increases by 1.2% and 1.03%, respectively.

(3) Village characteristics also matter. An increase of village income by 1000 RMB yuan

increases the opportunity costs of migration and decreases the individual migration probability by 0.92%. People living in the mountainous and hilly areas have a higher probability of migration. These two dummies may have captured unobserved poor living conditions regardless of village income. The distance from a village to the county seat and to the nearest transportation terminal does not significantly influence the migration decision.

6 Robustness Check

In the previous section, we use continuous measurement for the frequency of social interactions in information exchange and labor market mutual help. To check the robustness of whether labor market interactions always diminish the neighborhood effects, we utilize the discrete frequency of mutual help in labor market form. Correspondingly, the discrete frequency of information exchange is also used. The hypothesis is that only those people who have frequent mutual help with their neighbors in the labor market have lower neighborhood effects in migration decision, because they face time constraint. However, for those people who only interact a little with their neighbors in the labor market, they may have shorter social distances with their neighbors, and the neighborhood effect in migration decision is amplified through labor market interaction.

<Insert Table 4 here.>

Table 4 reports the results where the interaction terms of social interaction and neighborhood effect are constructed using the discrete measurement of social interaction. Table 4 shows that adding neighborhood effect and its interaction terms with social interaction frequency increased R-squared from 0.1828 in equation (1) to 0.2149 in equation (6). In equations (7) to (10), we find: (1) For information exchange, the more frequently people interact, the stronger the neighborhood effect is in migration decision. When information exchange is few, the neighborhood effect is not significantly increased, if village-level variables are controlled for. (2) For mutual help in the labor market, the more frequently people interact, the weaker the neighborhood effect is in migration decision. Interestingly, when mutual help is at the lowest level, the neighborhood effect may even be increased. The intuition is: when people help mutually, but not frequently, they do not face time constraint to migrate, and social distances among people are closer, so that neighborhood effects are enhanced. Here, since time constraint is not binding for those people who interact with their neighbors in the labor market, but not much, the simultaneity problem is not a major concern.

In both Table 3 and 4, when village migration ratio in 1998 is controlled for, the significance level and magnitude of the current migration ratio are reduced. Therefore, we want to see whether village migration ratio in 1998 is a better measurement to test heterogeneous neighborhood effect. To do so, we interact village migration ratio in 1998 with social interaction frequency measurements to repeat the estimation. The results are reported in Table 5.

Compared with the results in Table 3, where we measure neighborhood effect using current village migration ratio, most of the results are robust in Table 5. The only difference is that the interaction terms are less significant. In equation (11), where we control interaction terms using discrete social interaction frequency and county dummies, most interaction terms are insignificant, but the signs of the coefficients are the same as those in Table 3. In summary, we have robust evidences of neighborhood effect regardless of using current or historical migration ratio. However, the heterogeneity of neighborhood effect is more significant when using current migration ratio as the measurement of neighborhood effect.

<Insert Table 5 here.>

7 Neighborhood Effect and Public Policy

Our empirical results show that the neighborhood effect exists in rural China's labor migration. Furthermore, the neighborhood effect is nonlinear. Information sharing and modest labor market interaction enhances the neighborhood effect, while too frequent interaction in the labor market reduces the strength of the neighborhood effect. The existence of a nonlinear neighborhood effect has rich implication for policy makers. Theoretically, the neighborhood effect will lead to multiple equilibria in the economic process. When the mean group behavior outcome is at a low level, the economic process may converge to a low-level equilibrium because of interdependences in decision making; however, when the mean group behavior outcome exceeds some threshold, the economic process will converge to a high-level equilibrium with social interaction (Zanella, 2004).

In the context of our paper, China's urbanization would be dampened if there was a low-level equilibrium in rural–urban labor migration. We use the regression parameters in Table 3, Column 4 to simulate the equilibrium condition in the labor migration decision. Figure 1 reflects the relationship between the village migration ratio and the mean individual migration probability. The horizontal axis represents the village migration ratio and the vertical axis represents the individual migration probability. The solid line is the 45 degree line. The dash-dot line, the individual response curve, shows the relationship between individual migration probability and village migration ratio. Here we have only one point of intersection between the individual migration probability curve and the 45 degree line, with a slope less than one that guarantees a stable equilibrium with an average village migration ratio of 8.56%. As the pdf (probability density function) of the probit model is a standard normal distribution and its cumulative distribution function is assumed to be S-shaped, the low-level and high-level equilibriums can be differentiated according to the intersection point between the 45 degree line and the response curve. If the intersection point lies below 50% of the village migration ratio, the equilibrium is a low-level one. In contrast, if it is above 50%, the equilibrium is high level and stable, meaning that any departure within a limited range from the equilibrium will converge to the high equilibrium during dynamic adjustment. From Figure 1, we see that the intersection of the response curve and 45 degree line lies in the lower half of the S curve. That is to say, with the coefficients of the model unchanged, even if an exogenous shock increases the village migration ratio along the response curve, the labor migration ratio still converges to the low-level equilibrium trap under the influence of the neighborhood effect.

<Insert Figure 1 here.>

Promoting rural-to-urban labor migration is not only beneficial to rural residents, but also to China’s economic growth. Thus, our policy design aims to promote labor migration from rural to urban areas. In the following policy scenario analysis, we distinguish policies of three types and simulate their effects.

The first kind of policy is to move the response curve by changing individual characteristics such as education level. This policy can increase the migration probability but has no impact on social interaction among villagers, and thus does not change the slope of the response curve. Among the variables controlled, only the education level can be largely improved through economic policy. In Figure 2, we assume that policies are to improve the education of the villagers so that all villagers that are illiterate or have a primary school education can have the compulsory junior high school education. From the regression, we have already learned that the enhancement of rural residents’ education will increase the probability of outward migration. Figure 2 again shows this result. We find that the individual migration probability curve moves upwards and intersects with the 45 degree line at a higher point where the village migration ratio equals 9.47%. However, it should be noted that the effect of the policy is still limited and the point of intersection resumes the characteristics of a low-level equilibrium.

<Insert Figure 2 here.>

The second policy is to increase the social interaction that contributes to the neighborhood effect and decrease the social interaction that reduces the neighborhood effect. Graphically, this means rotating the curve anticlockwise while holding the intercept of the response curve constant. Figure 3 shows clearly this case. If we create policies to encourage more extensive interactions among villagers about job information sharing (define the state of "exchange information on employment" as "very frequently") and at the same time establish a rural labor service market to decrease the interactions on the labor market (we define the state of "mutual help during busy season" as "none/few"), we may find a significant counter-clockwise twist of the migration curve and a higher point of intersection on the 45 degree curve with a corresponding village migration ratio of 10.51%. In addition, we observe from Figure 3 that the individual migration probability curve becomes S shaped; however, the equilibrium is still at a low level.

<Insert Figure 3 here.>

What if we combine the above two policies? Figure 4 shows that by altering simultaneously the villagers' education and their social interactions strength, the combined policy will increase the migration ratio in equilibrium with a corresponding village migration ratio of 11.89%. However, the labor migration equilibrium is still at a low level even if the two policies are implemented together. In other words, new policies should be found to escape the low-level equilibrium of labor migration.

<Insert Figure 4 here.>

To facilitate the transition from a low-level equilibrium of migration to a high-level one, an important approach is the integration of the urban and rural labor markets through institutional reform, which is also the third kind of policy we could propose to increase labor migration within our analytical framework. Graphically, the policy will further heighten the intercept of the response curve. Although the current migration decision from rural to urban areas is in fact basically a free decision process, the existence of urban–rural segmentation and urban-biased economic policy still exerts extensive discrimination against rural migrants and labor migration is constrained. If we could eliminate this kind of urban-biased economic policy and promote rural–urban social integration, then the expected return of outward migration and thus the probability of outward migration increases. In Figure 5, we conduct a simulation and increase the intercept from –4.6672 to –4.1255, that is, an increase of 0.5417 in absolute value. Combined with the improvement in the rural education level and social interaction, this leads to an equilibrium migration ratio of 50%, which is obviously the threshold point of having a high-level equilibrium of labor migration. If the high-level equilibrium appears in the figure, by relying on the neighborhood effect and social multiplier, a small-scale positive impact to increase the labor mobility can result in the migration ratio converging to an even higher equilibrium. For the transition from a low-level equilibrium to a high-level one, a “big push” in the institutional environment is needed.

<Insert Figure 5 here.>

8 Conclusion

In this paper, we tested the existence and influence of the neighborhood effect on the labor migration decision. Our empirical results suggest the following conclusions. (1) The neighborhood effect exists in migration decision making, after past network and current bridging social network are controlled for. (2) The magnitude of the neighborhood effect is heterogeneous. Families who are more frequently involved in information sharing or modestly have mutual help in the labor market can enhance the neighborhood effect, while more frequent interactions in the labor market will reduce the positive neighborhood effect.

These findings have important policy implications. First, apart from the traditional development policies like education and training, rural-urban migration can be enhanced through neighborhood effect if a policy can raise village-level migration ratio. Second, social multiplier can be utilized to promote rural-urban migration. If information sharing is promoted, or if within-village labor exchange can be substituted by more efficient services, the social multiplier can be greater in amplifying the neighborhood effect in rural-urban migration. Third, a society could be trapped in a low equilibrium of rural-urban migration if there is neighborhood effect in migration decision.[3] It is possible that even if the policies like increasing education and enhancing the neighborhood effect can increase the labor migration ratio, however, neither of these policies can shift the low-level equilibrium to a high-level one. Thus, only institutional reform measures that are like a “big push” can change the low-level labor migration equilibrium to a high-level one. In the China case, the elimination of rural–urban labor market segregation policy and promoting social integration between migrants and urban residents are such institutional reform measures.

[3] Please refer to Moffitt (2001) for a more comprehensive discussion of the non-market interactions between individuals that lead to low-level equilibria, or “traps,” and the empirical identification issues and the relevant policies.

References:

Alesina, Alberto, and Giuliano, Paola, 2007. The power of the family. NBER Working Paper 13051.

Araujo, Caridad, de Janvry, Alain, and Sadoulet, Elisabeth, 2004. Peer effects in employment: results from Mexico's poor rural communities. University of California at Berkeley Working Paper.

Ballester, Coralio, Calvó-Armengol, Antoni, and Zenou, Yves, 2006. Who's who in networks. wanted: the key player. Econometrica 74 (5), 1403–1417.

Bao, Shuming, Bodvarsson, Örn B., Hou, Jack W., and Zhao, Yaohui, 2007. Interprovincial migration in China: the effects of investment and migrant networks. IZA Discussion Paper No. 2924.

Bauer, Thomas, Epstein, Gil and Gang, Ira N., 2002. Herd effects or migration networks? The location choice of Mexican immigrants in the U.S. IZA Discussion Paper No. 551.

Cai, Fang, Wang, Meiyan, and Du, Yang, 2003. The political economy of labor migration, (in Chinese), Shanghai Sanlian and Shanghai People's Publishing House.

Calvó-Armengol, Antoni, Patacchini, Eleonora, and Zenou, Yves, 2005. Peer effects and social networks in education and crime. The Research Institute of Industrial Economics, Working Paper No. 645.

Chen, Zhao, and Lu, Ming, 2008. From segmentation to integration: The political economy of urban-rural economic growth and social harmony. (in Chinese), Economic Research Journal (Jingji Yanjiu), No. 1, pp. 21-32.

Brock, William A., and Durlauf, Steven N., 2001. Discrete choice with social interactions. Review of Economic Studies, Vol. 68, No. 2, pp. 235-260.

Durlauf, Steven, 2004. Neighborhood effects. In: Handbook of Regional and Urban Economics Vol. 4, J.V. Henderson and J-F. Thisse (Eds.), Amsterdam: Elsevier Science, pp. 2173-2242.

Evans, William N., Oates, Wallace E., and Schwab, Robert M., 1992. Measuring peer group effects: A study of teenage behavior. Journal of Political Economy, Vol. 100, No. 5, pp. 966-991.

Glaeser, Edward, Sacerdote, Bruce, and Scheinkman, Jose, 2003. The social multiplier. Journal of the European Economic Association, Vol.1, No. 2-3, pp. 345-353.

Gustafsson, Björn, Li, Shi, and Sicular, Terry, (eds.), 2008. Inequality and public policy in China, New York: Cambridge University Press.

Liang, Zai, and White, Michael J., 1997. Market transition, government policies, and interprovincial migration in China: 1983-1988. Economic Development and Cultural Change, 45 (2), 321-339.

Manski, Charles F., 1993. Identification of endogenous social effects: the reflection problem Review of Economic Studies 60 (3), 531-542.

Marmaros, David, and Sacerdote, Bruce, 2006. How do friendships form? Quarterly Journal of Economics 121 (1), 79-119.

Mckenzie, David, and Rapoport, Hillel, 2007. Network effects and the dynamics of migration and inequality: theory and evidence from Mexico. Journal of Development Economics 84, 1-24.

Moffitt, Robert, 2001. Policy interventions, low-level equilibria, and social interactions. in Steven N. Durlauf, H. Peyton Young, (ed.), Social Dynamics, MIT Press.

Munshi, Kaivan, 2003. Networks in the modern economy: Mexican migrants in the U.S. labor market. Quarterly Journal of Economics 118 (2), 549-599.

Narayan, Deepa, 1999. Bonds and bridges: social capital and poverty. World Bank Policy Research Working Paper.

Patacchini, Eleonora, and Zenou, Yves, 2008. Juvenile delinquency and conformism.

Stockholm University Working Paper.
Zanella, Giulio, 2004. Social interactions and economic behavior. University of Siena. Working Paper.
Zhu, Nong, 2002. The impact of income gaps on migration decisions in China. *China Economic Review* 13 (2-3), 213-230.
Zhang, Xiaobo, Li, Guo, 2003. Does *guanxi* matter to nonfarm employment? Journal of Comparative Economics 31,315–331.
Zhao, Yaohui, 1999a. Leaving the countryside: rural-to-urban migration decisions in China. American Economic Association Papers and Proceedings 89 (2), 281-286.
Zhao, Yaohui, 1999b. Labor migration and earnings differences: the case of rural China. Economic Development and Cultural Change 47 (4), 767-782.
Zhao, Yaohui, 2003. The role of migrant networks in labor migration: The case of China. Contemporary Economic Policy 21 (4), 500-511.

Table 1: The Variable Definition

Neighborhood effect	village migration ratio (NE)	village migration ratio (excluding one's own family) in 2002
	village mig ratio 1998	village migration ratio in 1998
Individual characteristics	female	dummy variable, female=1
	age	age
	married	dummy, married=1
	primary school	dummy, if education is primary school, primary school=1
	junior high school	dummy, if education is junior high school, junior high school =1
	senior high school	dummy, if education is senior high school, senior high school =1
	tech school or more	dummy, if education is technical school or college education, tech school or more =1
	communist	dummy, if respondent is communist party member, communist =1
	health good	dummy, if health is very good or good, health good =1
	health bad	dummy, if health is very bad or bad, health bad =1
Family characteristics	household labor force	the number of labor force of a family
	family per capita land	family per capita land
	kids no. under 6	the number of children under age six of a family
	kids no. between 6 and 12	the number of children aging between six and twelve of a family
	elder no. over 65	the number of elders over age 65 of a family
	friends or relatives outside	dummy, if a family has friends and relatives outside village, friends or relatives outside =1
	friends or relatives village cadre	dummy, if a family has friends and relatives as village cadre, friends or relatives village cadre =1
Village characteristics	distance to nearest transportation terminal	the distance from village to a nearest transportation terminal, unit: kilometers
	distance to the country seat	the distance from village to the county seat, unit: kilometers
	village per capita income	village per capita income, unit: hundred Yuan
	mountain area	dummy, if a village locates in the mountain area, mountain area =1
	hill area	dummy, if a village locates in the hill area, hill area =1
Social interaction with other villagers	info exchange	Cardinal order variable, with "exchange information of employment" "very frequently", "often", "just so so", "sometimes" and "none/few" as 5, 4, 3, 2, 1, respectively.
	mutual help	Cardinal order variable, with "mutual-help during busy season" "very frequently", "often", "just so so", "sometimes" and "none/few" as 5, 4, 3, 2, 1, respectively.
	info very frequently	dummy, if "exchange information of employment" is "very frequently", information very frequently =1
	info often	dummy, if "exchange information of employment" is "often", information often =1
	info just so so	dummy, if "exchange information of employment" is "just so so", information just so so =1
	info sometimes	dummy, if "exchange information of employment" is "sometimes", information sometimes =1
	help very frequently	dummy, if "mutual-help during busy season" is "very frequently", help very frequently =1
	help often	dummy, if "mutual-help during busy season" is "often", help often =1
	help just so so	dummy, if "mutual-help during busy season" is "just so so", help just so so =1
	help sometimes	dummy, if "mutual-help during busy season" is "sometimes", help sometimes =1

Table 2: Statistical Description of Variables

	Full sample		Migrants		Non-migrants	
	16401		2675		13726	
Variable	Mean	s. d.	Mean	s. d.	Mean	s. d.
Individual Characteristics:						
female	0.4459	0.4971	0.3727	0.4836	0.4602	0.4984
age	34.6344	12.4495	27.1166	8.34829	36.09952	12.5880
married	0.6993	0.4586	0.4501	0.4976	0.7479	0.4343
primary school	0.2649	0.4413	0.1806	0.3847	0.2813	0.4496
junior high school	0.5033	0.5000	0.6191	0.4857	0.4807	0.4996
senior high school	0.1321	0.3386	0.1140	0.3179	0.1356	0.3424
tech school or more	0.0659	0.2481	0.0789	0.2696	0.0634	0.2437
communist	0.0710	0.2568	0.0303	0.1714	0.0789	0.2696
health good	0.8688	0.3376	0.9458	0.2265	0.8539	0.3533
health bad	0.0328	0.1781	0.0120	0.1087	0.0368	0.1884
Family Characteristics:						
household labor force	2.7678	1.2718	3.3544	1.2583	2.6534	1.2427
family per capita land	2.0937	2.3302	1.6258	1.7300	2.1849	2.4196
kids no. under 6	0.1818	0.4285	0.2011	0.4534	0.1780	0.4234
kids no. between 6 and 12	0.3354	0.6014	0.2819	0.5630	0.3458	0.6081
elder people no. over 65	0.1806	0.4535	0.1966	0.4794	0.1775	0.4482
friends or relatives outside	0.5726	0.4947	0.5922	0.4915	0.5688	0.4953
friends or members village cadre	0.2240	0.4169	0.2456	0.4305	0.2198	0.4141
Village Characteristics:						
distance to the country seat	25.2382	21.6849	27.1437	20.3367	24.8668	21.9194
distance to nearest transportation terminal	5.4653	8.3177	5.3916	7.9651	5.4797	8.3849
village per capita income	2.3886	1.3957	2.1802	1.1521	2.4292	1.4349
mountain area	0.2187	0.4134	0.2426	0.4287	0.2140	0.4102
hill area	0.3436	0.4749	0.4426	0.4968	0.3243	0.4681
Neighborhood effect:						
village migration ratio (NE)	0.1703	0.1474	0.2297	0.1533	0.1588	0.1434
village mig ratio 1998	0.0882	0.0786	0.1204	0.0814	0.0819	0.0764
Social Interaction Strength:						
info exchange	2.4816	1.2039	2.6860	1.2080	2.4418	1.1991
mutual help	2.9152	1.2424	2.8277	1.2639	2.9322	1.2375
info very frequently	0.0465	0.2106	0.0587	0.2351	0.0442	0.2054
info often	0.1722	0.3776	0.2213	0.4152	0.1627	0.3691
info just so so	0.2855	0.4517	0.2916	0.4546	0.2844	0.4511
info sometimes	0.2077	0.4057	0.2041	0.4031	0.2084	0.4062
help very frequently	0.1138	0.3176	0.0983	0.2978	0.1169	0.3213
help often	0.2198	0.4141	0.2329	0.4228	0.2173	0.4124
help just so so	0.3057	0.4607	0.2662	0.4420	0.3134	0.4639
help sometimes	0.1890	0.3915	0.2034	0.4026	0.1862	0.3893

Table 3: Probit Regression Result (Continuous Social Interactions)
Dependent Variable: Migrant or not (Migrant=1, Non-migrant=0)

Migrant (yes or no)	(1)	(2)	(3)	(4)	Marginal Effect Based on (4)	(5)
NE		1.503*** (0.163)	0.650*** (0.194)	0.656*** (0.195)	11.21%***	0.002 (0.228)
village mig ratio 1998			2.024*** (0.250)	1.903*** (0.252)	32.56%***	1.111*** (0.317)
info exchange* NE		0.262*** (0.046)	0.260*** (0.046)	0.241*** (0.046)	4.13%***	0.187*** (0.053)
mutual help* NE		-0.176*** (0.044)	-0.164*** (0.044)	-0.184*** (0.044)	-3.15%***	-0.138*** (0.050)
female	-0.286*** (0.027)	-0.275*** (0.028)	-0.274*** (0.028)	-0.264*** (0.028)	-4.45%***	-0.258*** (0.030)
married	-0.574*** (0.048)	-0.592*** (0.049)	-0.603*** (0.049)	-0.591*** (0.049)	-11.76%***	-0.649*** (0.054)
age	0.198*** (0.011)	0.201*** (0.011)	0.200*** (0.011)	0.199*** (0.011)	3.40%***	0.209*** (0.012)
age squared	-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)		-0.003*** (0.000)
communist	-0.170*** (0.065)	-0.160** (0.066)	-0.163** (0.067)	-0.163** (0.067)	-2.54%***	-0.062 (0.072)
health good	0.168*** (0.055)	0.186*** (0.057)	0.209*** (0.057)	0.229*** (0.058)	3.51%***	0.230*** (0.063)
health bad	-0.053 (0.112)	-0.001 (0.114)	0.020 (0.115)	0.021 (0.115)	0.36%	0.001 (0.123)
primary school	0.287** (0.115)	0.291** (0.118)	0.303*** (0.118)	0.282** (0.118)	5.25%**	0.293** (0.126)
junior high school	0.358*** (0.114)	0.383*** (0.117)	0.397*** (0.117)	0.415*** (0.118)	7.13%***	0.416*** (0.127)
senior high school	0.183 (0.119)	0.220* (0.122)	0.227* (0.122)	0.254** (0.123)	4.90%**	0.223* (0.132)
tech school or more	0.118 (0.124)	0.194 (0.126)	0.196 (0.127)	0.243* (0.127)	4.75%*	0.180 (0.138)
household labor force	0.240*** (0.011)	0.242*** (0.011)	0.242*** (0.011)	0.234*** (0.011)	4.01%***	0.248*** (0.013)
family per capita land	-0.084*** (0.007)	-0.070*** (0.007)	-0.066*** (0.007)	-0.073*** (0.008)	-1.25%***	-0.035*** (0.010)
kids no. under 6	-0.040 (0.032)	-0.033 (0.033)	-0.030 (0.033)	-0.047 (0.034)	-0.81%	0.011 (0.037)
kids no. between 6 and 12	-0.056** (0.024)	-0.047* (0.024)	-0.045* (0.025)	-0.057** (0.025)	-0.98%**	-0.033 (0.028)
elder people no. over 65	0.051* (0.027)	0.050* (0.028)	0.049* (0.028)	0.040 (0.028)	0.69%	0.051* (0.031)
friends or relatives outside	0.030 (0.029)	0.047 (0.030)	0.051* (0.030)	0.071** (0.030)	1.20%**	0.095*** (0.034)
friends or members village cadre	0.065* (0.034)	0.062* (0.034)	0.070** (0.035)	0.059* (0.035)	1.03%*	0.047 (0.038)
distance to the country seat				0.001 (0.002)	0.01%	-0.002 (0.002)
distance to nearest transportation terminal				0.000 (0.001)	0.01%	0.002** (0.001)

village per capita income				-0.054*	-0.92%*	-0.041
				(0.031)		(0.058)
village per capita income squared				0.001		-0.005
				(0.004)		(0.006)
mountain area				0.163***	2.97%***	-0.027
				(0.039)		(0.081)
hill area				0.189***	3.35%***	0.017
				(0.032)		(0.058)
constant	-4.180***	-4.632***	-4.706***	-4.667***		-10.784***
	(0.215)	(0.221)	(0.222)	(0.228)		(0.572)
County Dummy						**Y**
Number of obs	16401	16401	16401	16401		15730
Pseudo R2	0.1828	0.2136	0.2181	0.2232		0.3030

Note: NE=neighborhood effect. *, **, ***: Coefficient different from zero at 10, 5, 1 percent significance levels respectively. Standard Errors are in parentheses.

Table 4: Robustness Check (Discrete Social Interactions)

Dependent Variable: Migrant or not (Migrant=1, Non-migrant=0)

Migrant (yes or no)	(6)	(7)	(8)	(9)
NE	1.422***	0.584***	0.584***	0.138
	(0.148)	(0.181)	(0.182)	(0.213)
village mig ratio 1998		2.032***	1.907***	1.089***
		(0.251)	(0.253)	(0.318)
info very frequently*NE	0.976***	0.984***	0.930***	0.646**
	(0.244)	(0.245)	(0.246)	(0.283)
info often*NE	0.920***	0.931***	0.832***	0.512***
	(0.170)	(0.170)	(0.171)	(0.190)
info just so so*NE	0.424***	0.414***	0.368**	0.144
	(0.156)	(0.156)	(0.157)	(0.174)
info sometimes*NE	0.271*	0.339**	0.259	-0.154
	(0.165)	(0.165)	(0.166)	(0.181)
help very frequently*NE	-0.705***	-0.598***	-0.696***	-0.496**
	(0.218)	(0.219)	(0.220)	(0.246)
help often *NE	-0.226	-0.240	-0.305*	-0.224
	(0.166)	(0.166)	(0.167)	(0.187)
help just so so*NE	-0.300*	-0.330**	-0.373**	-0.306*
	(0.157)	(0.157)	(0.158)	(0.175)
help sometimes*NE	0.371**	0.349**	0.291*	0.155
	(0.169)	(0.170)	(0.170)	(0.183)
female	-0.275***	-0.273***	-0.264***	-0.258***
	(0.028)	(0.028)	(0.028)	(0.031)
married	-0.593***	-0.604***	-0.592***	-0.650***
	(0.049)	(0.049)	(0.049)	(0.054)
age	0.201***	0.200***	0.199***	0.209***
	(0.011)	(0.011)	(0.011)	(0.012)
age squared	-0.003***	-0.003***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
communist	-0.159**	-0.162**	-0.161**	-0.060
	(0.067)	(0.067)	(0.067)	(0.072)
health good	0.188***	0.210***	0.230***	0.231***
	(0.057)	(0.057)	(0.058)	(0.063)
health bad	-0.006	0.014	0.015	-0.003
	(0.115)	(0.115)	(0.115)	(0.123)
primary school	0.298**	0.310***	0.288**	0.303**
	(0.118)	(0.118)	(0.119)	(0.127)
junior high school	0.389***	0.403***	0.420***	0.426***
	(0.117)	(0.118)	(0.118)	(0.127)
senior high school	0.223*	0.231*	0.256**	0.231*
	(0.122)	(0.123)	(0.123)	(0.133)
tech school or more	0.202	0.204	0.248*	0.189
	(0.127)	(0.127)	(0.128)	(0.138)
household labor force	0.243***	0.242***	0.235***	0.249***
	(0.011)	(0.011)	(0.011)	(0.013)
family per capita land	-0.070***	-0.066***	-0.073***	-0.034***
	(0.007)	(0.007)	(0.008)	(0.010)
kids no. under 6	-0.037	-0.034	-0.051	0.009
	(0.033)	(0.033)	(0.034)	(0.037)
kids no. between 6 and 12	-0.045*	-0.044*	-0.055**	-0.030
	(0.024)	(0.025)	(0.025)	(0.028)
elder people no. over 65				

friends or relatives outside	0.050* (0.028)	0.049* (0.028)	0.040 (0.028)	0.049 (0.031)
friends or members village cadre	0.045 (0.030)	0.050* (0.030)	0.069** (0.030)	0.095*** (0.034)
have_relative/friend_as_cadre	0.066* (0.035)	0.074** (0.035)	0.062* (0.035)	0.048 (0.038)
distance to the country seat			0.001 (0.002)	-0.002 (0.002)
distance to nearest transportation terminal			0.000 (0.001)	0.002** (0.001)
village per capita income			-0.050 (0.032)	-0.046 (0.058)
village per capita income squared			0.000 (0.004)	-0.005 (0.006)
mountain area			0.162*** (0.039)	-0.022 (0.081)
hill area			0.186*** (0.033)	0.019 (0.058)
constant	-4.640*** (0.222)	-4.718*** (0.223)	-4.684*** (0.229)	-10.784*** (0.546)
County Dummy				**Y**
Number of obs	16401	16401	16401	15730
Pseudo R2	0.2149	0.2193	0.2242	0.3037

Note: NE=neighborhood effect. *, **, ***: Coefficient different from zero at 10, 5, 1 percent significance levels respectively. Standard Errors are in parentheses.

Table 5: Robustness Check (using village migration ratio in 1998 as NE)
Dependent Variable: Migrant or not (Migrant=1, Non-migrant=0)

Migrant (yes or no)	(10)	(11)	(12)	(13)
NE	0.749***	0.109	0.777***	0.107
	(0.136)	(0.163)	(0.135)	(0.162)
village mig ratio 1998	1.694***	1.211***	1.729***	1.013**
	(0.349)	(0.415)	(0.377)	(0.449)
info very frequently	1.636***	1.012*		
* village mig ratio 1998	(0.468)	(0.532)		
info often	1.389***	0.712**		
* village mig ratio 1998	(0.321)	(0.356)		
info just so so	0.795***	0.327		
* village mig ratio 1998	(0.295)	(0.327)		
info sometimes	0.678**	-0.173		
* village mig ratio 1998	(0.323)	(0.351)		
help very frequently	-0.990**	-0.625		
* village mig ratio 1998	(0.442)	(0.488)		
help often	-0.800**	-0.531		
* village mig ratio 1998	(0.323)	(0.366)		
help just so so	-0.937***	-0.749**		
* village mig ratio 1998	(0.303)	(0.338)		
help sometimes	0.336	0.203		
* village mig ratio 1998	(0.329)	(0.354)		
info. exchange			0.415***	0.278***
* village mig ratio 1998			(0.088)	(0.099)
mutual help			-0.332***	-0.231**
* village mig ratio 1998			(0.087)	(0.099)
female	-0.264***	-0.259***	-0.264***	-0.259***
	(0.028)	(0.030)	(0.028)	(0.030)
married	-0.595***	-0.652***	-0.594***	-0.651***
	(0.049)	(0.054)	(0.049)	(0.054)
age	0.199***	0.209***	0.199***	0.209***
	(0.011)	(0.012)	(0.011)	(0.012)
age squared	-0.003***	-0.003***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)
communist	-0.158**	-0.059	-0.160**	-0.059
	(0.067)	(0.072)	(0.067)	(0.072)
health good	0.228***	0.229***	0.227***	0.230***
	(0.058)	(0.063)	(0.057)	(0.063)
health bad	0.009	-0.006	0.015	-0.001
	(0.115)	(0.123)	(0.115)	(0.123)
primary school	0.284**	0.298**	0.280**	0.291**
	(0.118)	(0.127)	(0.118)	(0.126)
junior high school	0.417***	0.421***	0.414***	0.414***
	(0.118)	(0.127)	(0.117)	(0.127)
senior high school	0.253**	0.225*	0.251**	0.220*
	(0.123)	(0.133)	(0.122)	(0.132)
tech school or more	0.244*	0.183	0.241*	0.177
	(0.127)	(0.138)	(0.127)	(0.137)
household labor force	0.234***	0.249***	0.234***	0.248***
	(0.011)	(0.013)	(0.011)	(0.013)
family per capita land	-0.073***	-0.034***	-0.073***	-0.034***
	(0.008)	(0.010)	(0.008)	(0.010)
kids no. under 6	-0.051	0.009	-0.047	0.011
	(0.034)	(0.037)	(0.033)	(0.037)

kids no. between 6 and 12	-0.054**	-0.030	-0.056**	-0.032
	(0.025)	(0.028)	(0.025)	(0.028)
elder people no. over 65	0.041	0.048	0.040	0.051*
	(0.028)	(0.031)	(0.028)	(0.031)
	0.070**	0.095***	0.070**	0.096***
friends or relatives outside	(0.030)	(0.034)	(0.030)	(0.034)
friends or members village	0.063*	0.050	0.061*	0.049
cadre	(0.035)	(0.038)	(0.035)	(0.038)
	0.001	-0.002	0.001	-0.002
distance to the country seat	(0.002)	(0.002)	(0.002)	(0.002)
distance to nearest	0.000	0.002**	0.000	0.002**
transportation terminal	(0.001)	(0.001)	(0.001)	(0.001)
village per capita income	-0.051	-0.049	-0.053*	-0.041
	(0.032)	(0.058)	(0.031)	(0.058)
village per capita income	0.001	-0.005	0.001	-0.005
squared	(0.004)	(0.006)	(0.004)	(0.006)
mountain area	0.164***	-0.013	0.167***	-0.022
	(0.039)	(0.081)	(0.039)	(0.081)
hill area	0.188***	0.021	0.190***	0.016
	(0.033)	(0.058)	(0.032)	(0.058)
constant	-4.686***	-10.786***	-4.669***	-10.787***
	(0.228)	(0.572)	(0.228)	(0.546)
County Dummy		**Y**		**Y**
Number of obs	16401	15730	16401	15730
Pseudo R2	0.2237	0.3032	0.2228	0.3027

Note: NE=neighborhood effect. *, **, ***: Coefficient different from zero at 10, 5, 1 percent significance levels respectively. Standard Errors are in parentheses.

Figure 1: Simulation of Labor Migration Equilibrium

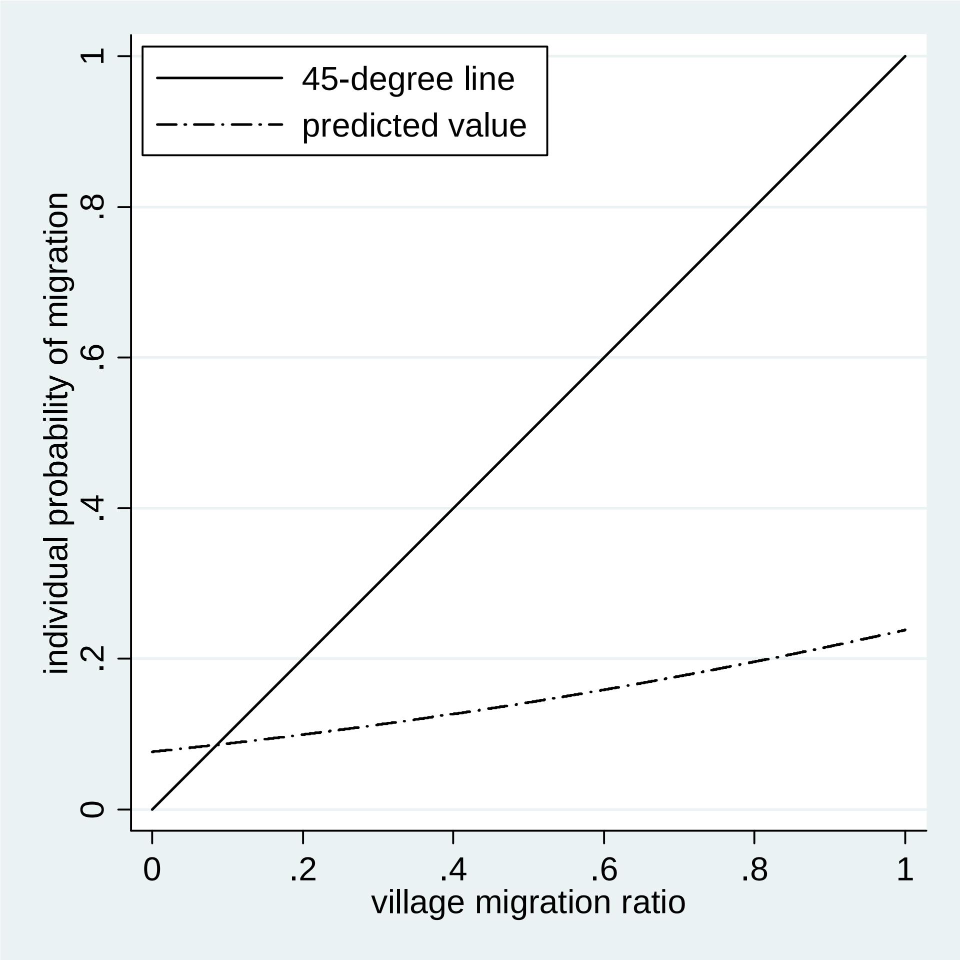

Figure 1 shows the relationship between village migration ratio and mean individual out migration probability (simulation parameters are from table 3). When the two values equal (cut the 45 degree line), it is the equilibrium migration ratio. As shown in the graph, the equilibrium migration ratio is 8.56%.

Figure 2: Policy Effect: Increasing Educational Level

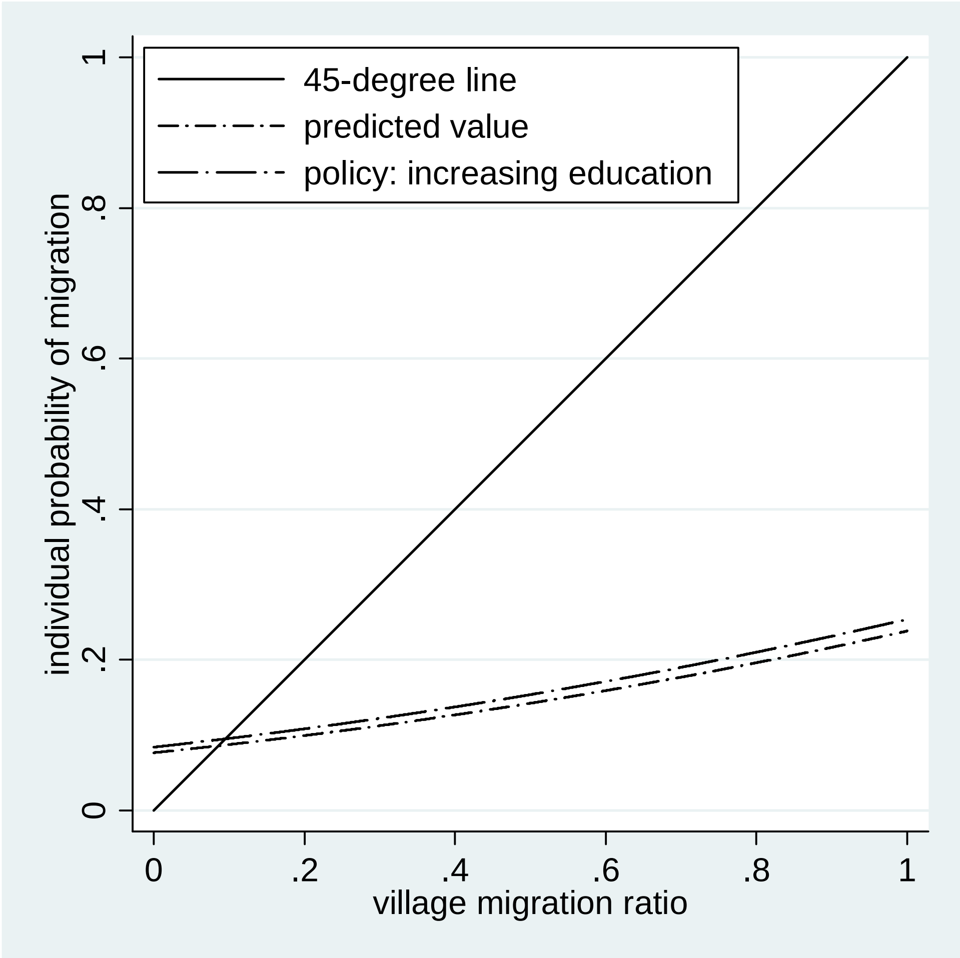

Figure 2 shows the policy effect of increasing education investment on out migration decision. We assume every sample individual receives at least nine year compulsory education (junior high school level). The equilibrium migration ratio increases to 9.47%.

Figure 3: Policy Effect: Increasing Pro-Peer Effect Social Interaction

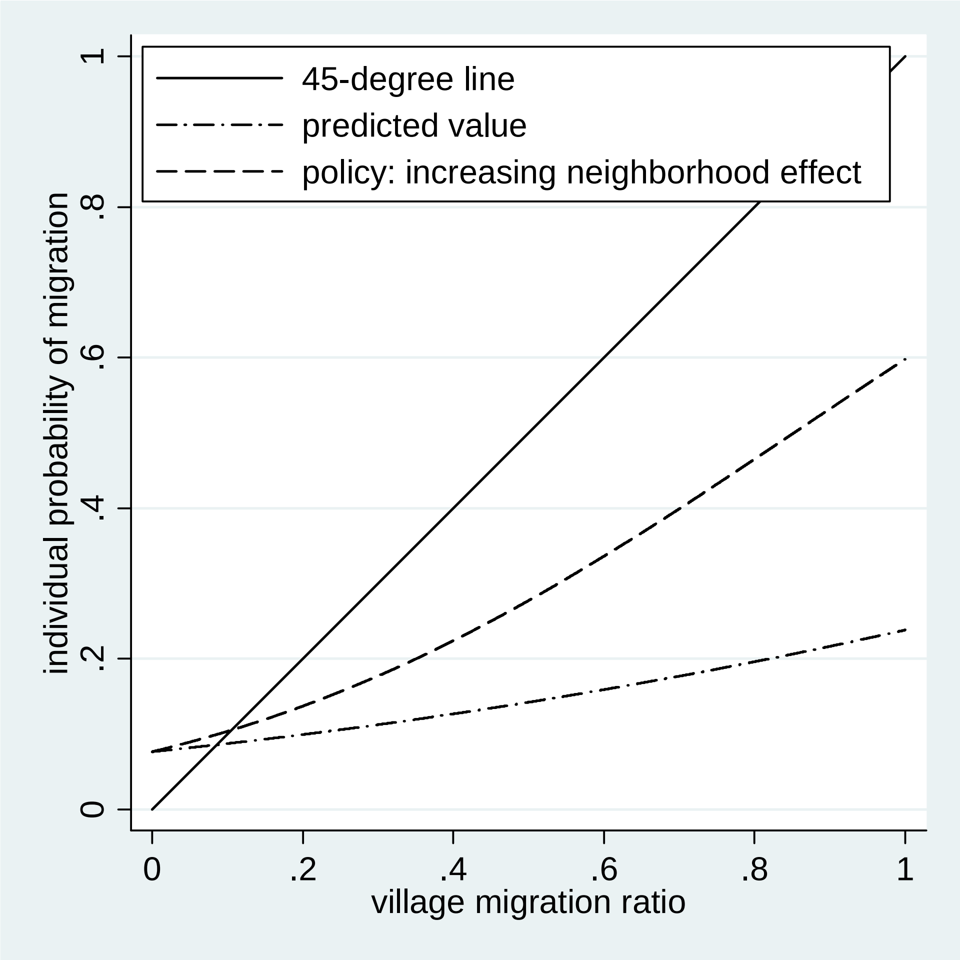

Figure 3 demonstrates the policy effect of increasing pro-peer effect social interaction on migration decision. In here, we control the information sharing interaction at “very frequently” while set the labor market interaction at “none/few”. The intuitive policy measures are establishing formal job information broadcasting institution and labor service enterprises in rural areas. For such policies, the equilibrium migration ratio reaches 10.51%.

Figure 4: Policy Effects: Increasing Educational Level and Pro-Peer Effect Social Interaction

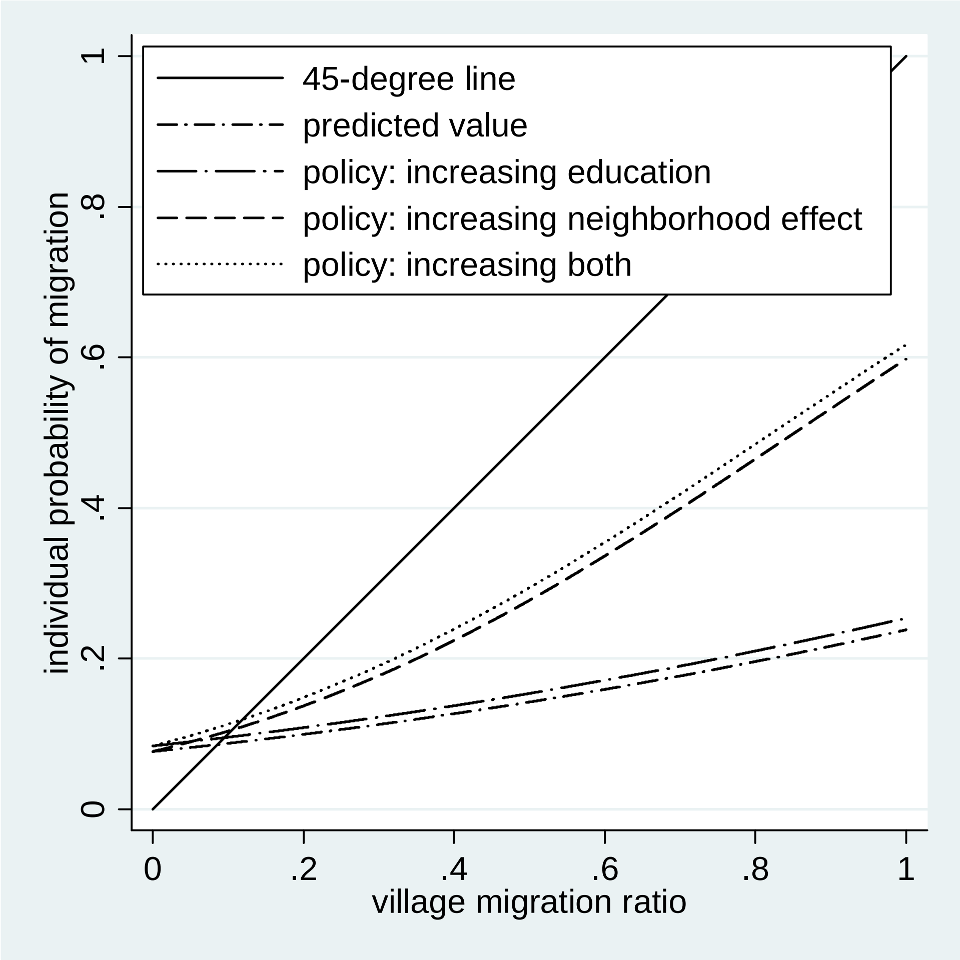

Figure 4 combines Figure 2, 3 and additionally shows the overall policy effect of increasing both education level and pro-peer effect social interaction. The combining policy will lift up equilibrium migration ratio to 11.89%.

Figure 5: Policy Effect: Institutional "Big Push" in Rural-Urban Integration

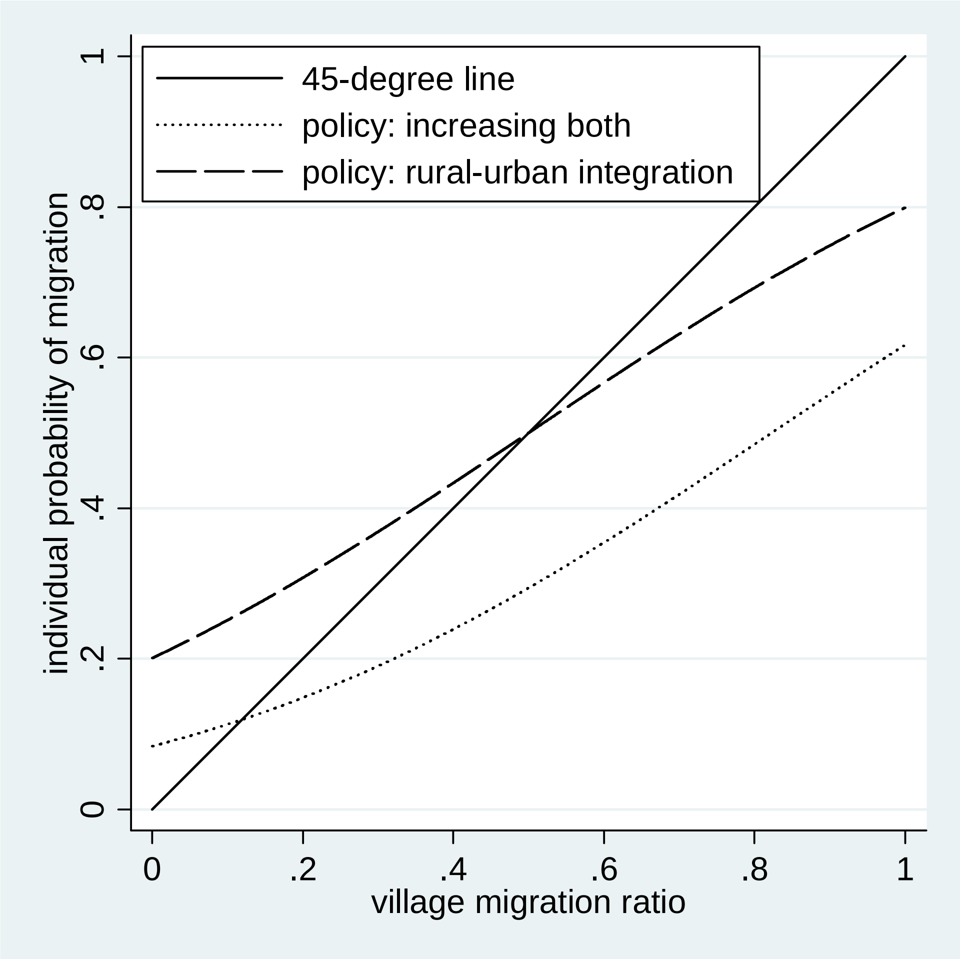

Figure 5 shows the effect of rural-urban labor market integration on out migration decision (long dash line). Though in our framework we do not have explicit parameters to measure the extent of labor market discrimination against rural migrants, we increase the intercept term, which is exogenous and homogenous to every sample individual and thus can represent the "institutional change", to demonstrate the effect of market integration. We increase intercept from -4.6672 to -4.1255 and the equilibrium migration ratio reaches 50%.