

IZA DP No. 8351

**Aging and Migration in a Transition Economy:  
The Case of China**

Örn B. Bodvarsson  
Jack W. Hou  
Kailing Shen

July 2014

# **Aging and Migration in a Transition Economy: The Case of China**

**Örn B. Bodvarsson**

*St. Cloud State University  
and IZA*

**Jack W. Hou**

*California State University, Long Beach  
and Henan University*

**Kailing Shen**

*WISE, Xiamen University  
and IZA*

Discussion Paper No. 8351  
July 2014

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### **Aging and Migration in a Transition Economy: The Case of China<sup>1</sup>**

Post-reform China has been experiencing two major demographic changes, an extraordinary amount of internal migration and an aging population. We present a general migration model which captures the idea that older migrants have shorter durations in the destination but possibly larger general human capital to transfer. Therefore, the incentive to migrate is ambiguously related to age. We test the theoretical implication using an extended modified gravity model, nuanced to fit the case of a transition economy. We find that shifts in China's age distribution have generated significant changes in the country's migration patterns.

JEL Classification: J61, J11

Keywords: internal migration, age distribution, reforms

Corresponding author:

Kailing Shen  
Wang Yunnan Institute for Studies In Economics (WISE)  
Room A307, Economics Building  
Xiamen University  
Xiamen, China 361005  
E-mail: [Kailing.Shen@gmail.com](mailto:Kailing.Shen@gmail.com)

---

<sup>1</sup> We thank Shuming Bao for invaluable assistance with data collection, Ergan Xu, Robert Girtz, Asm Ulah, Li Lu, and Matthew Jensen for assistance with preparation of some of the data, and Chad Sparber and Lei Meng for helpful comments.

## 1. INTRODUCTION

In sheer numbers, the post-reform “Great Chinese Internal Migration” dwarfs all other episodes of human migration. According to the China Census, during 1985-90 approximately 33 million Chinese moved within or across provinces. During 1995-2000 over 121 million persons moved and during 2000-05 nearly 195 million persons moved, approximately one-third across provincial boundaries. These surges primarily include persons moving from rural-to-urban areas and to the Eastern coastal cities. A small literature on this migration episode has emerged and it has identified two broad determinants.<sup>2</sup> (1) Growth in regional income differences due to deepening economic reforms, a boom in China’s export markets, and a surge in foreign and domestic investments; and (2) A dramatic decline in migration costs due to substantial improvements in the country’s transportation infrastructure, deregulation of migration, and rapid growth in migrant communities.

There has been another major demographic change in China: The age distribution is shifting. China is aging because of global prosperity, post-reform structural changes to the economy, improved health care, and other factors, but most prominently, the long term effects of the “one-child” policy. The China Census reveals some important trends. First, mean provincial age has risen by approximately 30%, from just under 21 years in 1982 to nearly 27 years in 2005. Second, while the youth share (share of population aged 15-29) rose very slightly from 1982 to 1990, it has declined continually since then; In 1990 the youth share was just over 30 percent, whereas by 2005 it had fallen to just over 21 percent. One likely major contributor to both trends is the one-child policy, but other contributors could include improvements in health care and accelerating prosperity (which increases the opportunity costs of having children). Third, age dependency ratio

---

<sup>2</sup> The literature can be conveniently divided into studies utilizing micro-data obtained from special household surveys (see, for example, Liang (2001), Liang and White (1996,1997), Zhao (1997,1999a, 1999b, 2002, 2003)) and a few studies utilizing province-level data (see, for example, Fan (2005), Lin, Wang and Zhao (2004), Poncet (2006), Bao, Hou and Shi (2006), and Bao, Bodvarsson, Hou, and Zhao (2008a, 2008b, 2011)). We should also point out that in 2002, an entire issue of the journal *Urban Studies* was devoted to empirical papers on China’s growing migration and urbanization. We particularly wish to highlight the studies of Chen and Coulson (2002) on the determinants of urban migration, Liang, Chen and Gu (2002) on the effects of rural industrialization on internal migration, Li and Zahniser (2002) on the determinants of temporary rural-to-urban migration, and Goodkind and West’s (2002) study on the floating population.

(ADR) has fallen.<sup>3</sup> In 1982, for every 100 persons of working age (15-64 years), there were on average approximately 62 persons too young or too old to participate in the workforce. By 1990, this number had fallen to just under 50, by 2000 it was approximately 42, and by 2005 it was under 40. The decline in ADR is likely due to the long term effects of the one-child policy, which has reduced the fraction of the population comprising children, teenagers, and young adults. While improved health care and long term prosperity have increased the fraction of the elderly population, the rate of decline in the youth share has been greater (in absolute value) than has the rate of increase in the elderly. Over the same period, however, the share of the working age population has risen,<sup>4</sup> as those born before the one-child policy were the main body of the labor force.

Several important questions are inspired by these trends. First, has an aging population, all other things equal, suppressed migration? Second, has a decline in the fraction of the “dependent” population contributed to a change in the scale of migration? Three categories of previous work are helpful in addressing these two questions:

(1) *Western economic theory*. The theoretical relationship between age and migration has actually received relatively little attention in Western migration literature. Becker (1964) argued that the propensity to migrate will tend to decrease with age because the expected net present value of the benefits from relocation will, due to greater duration of stay in the destination, be higher for younger persons. This implies that migration rates for persons from the lower (higher) end of the home region’s age distribution will be higher (lower).<sup>5</sup>

Gallaway (1969) suggested that there are additional reasons for the age-migration relationship and these cause the relationship to be ambiguous. On the one hand, older

---

<sup>3</sup> The ADR measures the relative size of the population that is not working, hence dependent upon the workforce for financial or in-kind support. The ratio is computed for China using the following formula:

$$ADR = \frac{\text{number of persons aged 0 to 14} + \text{number of persons aged 65 or higher}}{\text{number of persons aged 15 - 64}}$$

<sup>4</sup> Note that the average share of provincial population aged 15-64 was 66.79% in 1990, 70.39% in 2000 and 71.91% in 2005.

<sup>5</sup> Schlottmann and Herzog (1984) described this as “age selectivity of migration. They focused on how career and geographic mobility interact in influencing the age-selectivity of migration. Using data on interstate migration for 1965-70, they concluded that failure to account for this interaction will tend to overstate the negative influence of age on the probability of migration. It should be noted that Schlottmann and Herzog focused on this particular empirical issue and did not provide a theoretical model with any novel implications regarding the relationship between age and proclivity to migrate.

workers face higher costs of relocation because: (a) they face a greater expected cost of not being able to fully transfer pension rights accumulated in retirement programs associated with their jobs; (b) the expected costs of liquidating physical investments in the origin, e.g. selling one's home or business, will be higher; and (c) they face greater psychic costs that come from uprooting themselves from long-held jobs and dwellings. All other things equal, this implies that migration rates of older workers will be lower. On the other hand, older workers usually earn more money and have more assets, so migration will tend to be more affordable for them. Consequently, the likelihood of migration may be higher for those types of workers, all other things equal.<sup>6</sup>

Other explanations have been given for the effects of age on the incentive to migrate. David (1974) suggested that seniority rights (which provide protection from the risk of layoffs) will be lost following a move. Schwartz (1976) makes a strong argument for the importance of psychic costs. He characterizes these costs as the “agony of severing relations” with family members and friends. Schwartz argues that as persons get older, they will invest more in relationships with family members and friends and the emotional costs of severing those relationships will be higher. The “agony” of relocation can be assuaged by visits home and older persons will have a greater demand for these visits.

Lundborg (1991) developed Schwartz's point further by suggesting that the demand for return visits will depend on length of time spent at the destination, age at the time of migration, and the stock of prior migrants from the origin residing in the destination. On the one hand, older migrants value return visits more, but on the other, as time passes after the move the migrant invests in new social relations and the demand for return visits will fall. Furthermore, the larger is the migrant network the less homesick the migrant will feel and his/her demand for return visits will be lower. Lundborg thus provides two additional implications: (1) The deterrent effect of distance is higher for older persons;

---

<sup>6</sup> Gallaway estimated the effects of age on internal U.S. migration rates using individual data for 1957-60 from the One Percent Continuous Work History Sample maintained by the Social Security Administration. In one set of regressions, the dependent variable is the proportion of workers in a particular age group who remain in the region, while in another set the dependent variable is the fraction of those leaving. Dummies for different age categories and dummies for labor market earnings in the region for workers within each age category were included. Gallaway generally found that, *ceteris paribus*, older workers migrate less frequently but mobility will be higher for workers who on average earn more. Thus, his results appear to support both the “higher relocation costs” and “greater affordability” hypotheses.

and (2) There will be a U-shaped relationship between the sensitivity of migration to the size of the migrant community and migrant age;

(2) *Empirical Western literature on age and migration.* There have been numerous tests of the effect of age on migration, mostly of the Becker (1964) hypothesis, done in the West since the 1970s. In the interest of space, we do not provide an exhaustive survey here. In testing Becker's hypothesis, the basic strategy has been to test for a negative relationship between the rate of migration and a migrant's age. Generally, results from these studies have been quite mixed.<sup>7</sup> In some studies, it is found that younger migrants have a higher propensity to migrate, whereas other studies find the opposite. Some studies find no statistically significant relationship, while quite a few studies simply omit age as a regressor. It is difficult to tie together the diverse empirical findings with respect to age across these studies because: (i) there is considerable diversity in empirical specifications and the types of data sets used; and (ii) these studies generally lack a unifying theory capable of accounting for the diversity of results.

(3) *Empirical work on age and migration in China.* A majority of researchers doing empirical work on the determinants of migration in China have included age as an explanatory variable. The results from these studies have also been mixed.<sup>8</sup> An important result that emerges from a majority of these studies is an *inverse* U-shaped

---

<sup>7</sup> Bowles (1970) studied out-migration rates of black and white workers from the American south and found that for both groups, migration propensities were higher for younger workers, especially blacks. Using 1960 U.S. Census data, Schwarz (1976) found that migration rates were higher for younger persons, especially well-educated ones. Navratil and Doyle (1977) used 1970 U.S. Census data to estimate in-migration rates by race-sex cohorts (an aggregate flow model), as well as the likelihood of an individual migrating (a logit model using microdata), during 1965-70. For the aggregate flow model, they found a positive relationship between age and migration rates for white males, black males and black females, but no relationship for white females. For the logit model the likelihood of migrating was higher for older persons in all sex-race cohorts. Schlottmann and Herzog (1981) found strong evidence of an inverse relationship between the likelihood of migration and age using U.S. Census data for 1965-70, a result echoed by Goss and Paul (1986), who used Panel Study of Income Dynamics (PSID) micro-data for 1974-75. Lundborg (1991) found no evidence of a relationship between migration rates and age for migration between Scandinavian countries. Finally, Clark, Hatton and Williamson (2007) found no relationship between U.S. immigration rates during 1971-98 and the share of source country population aged 15-29.

<sup>8</sup> In a series of studies utilizing small, household surveys in specific areas, Zhao found a negative relationship between age and the propensity to migrate (Zhao (1999a, 1999b, 2003), as well as a positive relationship (Zhao (1997)). A negative effect of age on the propensity to migrate was also found by Zhu (2002), Shi et al (2007) and Wu (2008). We should also mention that Zhao (2002) found a higher tendency for older migrants to return home.

relationship between age and the propensity to migrate.<sup>9</sup> Specifically, the propensity to migrate rises up to 25-30 years of age, then falls thereafter. This type of result has generally not been found in the Western literature. Zhu (2002) has suggested that one reason for observing a negative age coefficient beginning in the late twenties for the Chinese case is that unskilled manual workers and older workers are disadvantaged with respect to their physical strength.

Other researchers provide evidence supporting the views of Lundborg (1991) and Schwartz (1976) that older migrants face higher psychic costs of migration. Zhao (1997) has suggested that the positive relationship between age and migration propensity that occurs through the mid-to-late twenties could be due to Hukou restrictions on migration, which may be especially constraining for very young persons living in rural areas. She points out that these persons may remain in rural areas for a while (or first enlist in the military) to gain favor from local officials in order to be considered for relocation when urban recruitment opportunities arise. Therefore, the apparent quadratic relationship between age and migration suggests that the Chinese case is more complex than the Western one due to certain institutional factors.<sup>10</sup>

In this paper, we seek to understand better the relationship between age and migration. First, we present a theoretical model of the migration decision that is capable of generating the diversity of predictions compatible with what has been found in the empirical literature. We then test numerous implications of the model on two data sets, one comprising cross-section micro data and the other panel aggregate (province-level) data, encompassing three periods of interprovincial migration in China. We wish to emphasize that the theory (section 2) describes the individual migrant's decision, whereas the majority of our tests (found in section 5) are on aggregate data, specifically inter-

---

<sup>9</sup> See Shi and Bao (2007), Liang, Chen and Gu (2002), Zhao (1999a), Liang and White (1996, 1997), Li and Zahniser (2002), Ma and Liaw (1997) and Hare (1999).

<sup>10</sup> We can suggest a number of more casual explanations, applicable to both the Western and Chinese cases, for why migration propensities fall with age. One is that younger persons are on average healthier. Since the act of migration, particularly from a rural to an urban area, involves a relatively substantial investment of resources and there is a greater financial risk from getting sick and requiring hospitalization in the destination, particularly if one is part of the floating population (unauthorized migrants do not have access to free local medical care), older persons may find it much more costly to migrate. Another explanation is that younger persons are less risk averse, more adventuresome, and more entrepreneurial. All these explanations are consistent with and complement the general hypothesis from the basic human capital model that migration rates and age will be inversely related. We have not seen these explanations incorporated into formal theoretical or empirical models of migration, though.



provincial migration rates. We make no claim to directly test the theory. Our empirical emphasis is primarily at the aggregate level, which is not as compatible with the theoretical work as we would like. Our view is that the theory provides important guideposts for understanding the relationship between migrant age and the propensity to migrate, a relationship which should not only be found at the micro level, but should also be detected using data on provincial migration rates and age distributions. It should also be noted that such incompatibility between theory and empirical work is not unusual for studies of migration in China where aggregate data are much more readily available than data obtained from micro surveys. Our most basic finding is that age distribution does affect the scale of migration between provinces and often in unexpected ways.

## **2. AGE AND THE DECISION TO MIGRATE: THEORY**

The theoretical model below incorporates elements of models due to Schwartz (1976), Naskoteen and Zimmer (1990), and Lundborg (1991). While the model describes the behavior of an individual prospective migrant, it has immediate implications for the study of aggregate migration flows. For simplicity, we assume just one potential destination. The decision to migrate is assumed to be influenced by three broad factors: (1) spatial differences in age-earnings profiles; (2) the costs of maintaining investments in social relations at the origin; and (3) direct migration costs, which include the costs of securing local registration in the destination. What makes our model different from most other models is that: (a) we emphasize the influence of spatial differences in earnings profiles, as opposed to levels; and (b) we combine spatial differences in earnings opportunities, the costs of maintaining investments in social relations, and direct migration costs, all in one model. The model below may be viewed as a “Modified Roy (1951) model,” where there are regional variations in age-earnings profiles and these variations drive differences in the age distributions of migrants.<sup>11</sup>

---

<sup>11</sup> The traditional Roy model, further developed by Borjas (1987), explains how spatial differences in age-earnings profiles can influence the skill-selectivity of migration, e.g. migration is positively skill-biased when the destination offers a higher return to skill acquisition. Our approach is different: We argue that when the destination offers a higher return to skill acquisition, this will influence the age distribution of people who move there.

Suppose a risk-neutral worker of age  $a$ , who plans to retire at time  $T$ , is contemplating a move from place  $i$  to place  $j$ . The decision to relocate is facilitated by a calculation of the expected net present value (NPV) of the benefits of relocation:

$$E[NPV(t)] = \int_a^T [E(Y^j) - E(Y^i) - C^{ij}(t)] e^{-rt} dt \quad (1)$$

Where:  $Y^i$  = earnings per period available in the origin province;

$Y^j$  = earnings per period available in the destination province;

$C^{ij}$  = costs of migrating from provinces  $i$  to  $j$ ;

$r$  = discount rate.

Embedded in equation (1) is the idea that the destination offers the migrant an age-earnings profile different from the profile available at home. The profiles differ because: (a) destination earnings with no labor market experience will differ from what would be earned in the origin; (b) there are geographic differences in returns to general and specific human capital; and (c) specific assets are lost following a move. To illustrate how there could be spatial differences in age-earnings profiles, suppose that labor markets in the destination are stronger. Even if the migrant has no labor market experience, at a given education level he/she will likely receive higher compensation in the destination. This implies that the vertical intercept of the destination age-earnings profile will be higher. For the same reason, returns to general and specific human capital are likely to be higher, meaning that the destination age-earnings profile will be steeper. The loss of specific assets could be a significant reason for spatial differences in age-earnings profiles because a migrant could switch occupations following a move. For example, a rural migrant's human capital investments could have been in agricultural trades. If he/she wishes to move to a large city and be employable there, however, it may be necessary to retool and take a job in manufacturing. While he/she may be able to transfer general skills, specific human capital investments will be lost.

Expected earnings each period are assumed to depend upon five components: (i) the wage that would be paid if the worker had zero labor market experience (we call this the "baseline wage,"  $W_B$ ); (ii) the amounts of general and specific human capital accumulated from prior periods; (iii) the price received for supplying a unit of general human capital (call this  $x$ ); (iv) the price received for supplying a unit of specific human

capital (call this  $s$ ); and (v) the probability of securing employment (call this  $1 - \pi$ , where  $\pi$  is the risk of unemployment).

For simplicity, we will assume that, at any working age, the worker acquires one unit of general human capital and one unit of specific human capital each period.<sup>12</sup> General human capital acquired in one location is completely transferrable to another location, while specific human capital cannot be transferred. The prospective migrant is assumed to have graduated from school at age  $a_g$  and has accumulated so far  $(a - a_g)$  units of general human capital. Furthermore, he/she is assumed to have taken the current job (in the origin) at age  $a_k$  ( $a_g < a_k < a$ ) and to have accumulated  $(a - a_k)$  units of specific human capital. While the baseline wage and the returns to general and specific human capital are likely to differ within and across locations, we will assume for simplicity that they are constant over the individual worker's lifetime. It follows then that at age  $a$ , the expected earnings in each location are the following:

$$\begin{aligned}
 E(Y^i) &= \int_a^T [(1 - \pi^i)(W_B^i + x^i + s^i)e^{-rt}] + x^i(a - a_g) + s^i(a - a_k) \\
 E(Y^j) &= \int_a^T [(1 - \pi^j)(W_B^j + x^j + s^j)e^{-rt}] + x^j(a - a_g)
 \end{aligned} \tag{2}$$

The costs of migration are assumed to depend upon the frequency of desired travel back to the origin, the costs of transportation (which vary directly with distance), and the costs of securing Hukou. Following Lundborg (1991, pp. 395-6), we assume that the frequency of desired travel to the origin (call this TRIPS) depends upon the amount of time spent in the destination, the migrant's age at the time of migration, and the relative size of the migrant community in the destination. The function measuring desired return trips may be written as:

$$\text{TRIPS} = g(t - a, a, MS_{ij}) \tag{3}$$

where:  $t - a$  = number of years at the destination

$MS_{ij}$  = the stock of prior migrants from  $i$  residing in  $j$ .

---

<sup>12</sup> This is a departure from the traditional view of on-the-job training, which is that human capital investments taper off with age. As long as the rate of human capital accumulation is the same in either location, the predictions of the model do not depend upon the rate being constant or non-linear. It would be interesting to extend the model to the case where human capital is acquired at a faster rate in the destination, due, for example, to a different industrial structure or more advanced technology.

With subscripts as derivatives, it is assumed that  $g_{t-a} < 0$  because, as suggested by Schwartz (1976) and Lundborg (1991), as the migrant becomes more familiar and comfortable with the destination, investments in new social relationships occur in the destination and return trips to the origin have increasingly less value. In contrast,  $g_a > 0$  because the older one is at the time of migration the more that will have been invested in social relationships at the origin and the greater will be the desire to make return visits there.<sup>13</sup> We also adopt Lundborg's assumption that  $g_{MS} < 0$  because the presence of a relatively larger community of migrants from one's place of origin will reduce the psychic costs of migration and thus the frequency of desired return visits.<sup>14</sup>

The costs of return trips over the migrant's lifetime will equal the number of trips made times the direct cost of a return trip, which discounted to the present are

$$\int_a^T [g(t-a, a, MS_{ij}) C_{ij}(D_{ij})] e^{-rt} dt, \quad (4)$$

where  $C_{ij}$  = the round-trip direct cost of travelling between origin and destination;

$D_{ij}$  = distance between origin and destination ( $\frac{\partial C}{\partial D} > 0$ ).

Equation (4) quantifies the psychic costs of migration. There will also be various direct costs, e.g. the costs of moving oneself and one's possessions, switching dwellings, and securing local registration. Direct costs of the initial move will vary directly with distance. Older migrants will generally face higher costs of securing Hukou because they will have a lower likelihood of being included in those groups that are considered strong candidates for local registration. Those groups include students admitted to university, military conscripts, and marriage partners.<sup>15</sup> The direct costs of the initial move (M) are:

---

<sup>13</sup> Compared to those who migrate when they are relatively young, older migrants are more likely to have developed longer and deeper friendships in the origin, cultivated relationships with extended family members more extensively, and to have left children behind at the origin. To use Schwartz's (1976) terminology, the "agony" of severing those sorts of ties will be more intense for older migrants and they will feel a greater yearning to return home more frequently.

<sup>14</sup> Lundborg points out that for unmarried migrants, a larger migrant network in the destination could make it more likely that a marriage partner with ties to the origin can be found at the destination. This will tend to reduce the number of desired trips back home.

<sup>15</sup> Persons admitted to universities in another province are granted local registration, as are military conscripts. One convenient way of obtaining Hukou is to marry someone already registered in the destination. University students, military conscripts, and candidates for marriage are usually younger than the mean. Consequently, because registration is a requirement for having access to social medical insurance in the destination, the costs of being unregistered will be higher for older persons because they are at greater risk of needing medical care.

$$M = m(D_{ij}) + H_{ij}(a) \quad (5)$$

where:  $H_{ij}$  = the costs of securing local registration in place  $j$  ( $\frac{\partial H}{\partial a} > 0$ ).

Combining equations (2) through (5), the expected net present value of migration is

$$\begin{aligned} E[NPV(t)] = & \int_a^T [(1 - \pi^j)(W_B^j + x^j + s^j) - (1 - \pi^i)(W_B^i + x^i + s^i)]e^{-rt} dt - \\ & \int_a^T [g(t - a, a, MS_{ij})C_{ij}(D_{ij})]e^{-rt} dt - [m(D_{ij}) - H_{ij}(a)] + \\ & (1 - \pi^j)(x^j - x^i)(a - a_g) - (1 - \pi^i)(s^i)(a - a_k) \quad (6) \end{aligned}$$

Equation (6) illustrates that the net benefits to migration are influenced by five factors, each corresponding to the respective term on the right-hand side (RHS) of equation (6): (i) spatial differences in expected earnings, driven by destination/source differences in baseline wages, returns to general human capital, returns to specific human capital, and unemployment risk (these are all captured in the first term on the RHS of equation (6)); (ii) expected psychic costs of migration, which are influenced by the needed frequency of return trips to the origin (second term); (iii) the direct costs of moving and securing local registration (third term); (iv) the expected extra value the migrant can enjoy from transferring accumulated general human capital to the destination (fourth term); and (v) the expected loss in specific human capital taken due to relocation (fifth term). Since there is just one destination, migration is a yes/no decision, thus migration occurs if equation (6) is positive.

How exactly does age influence the migration decision? While our model is not set up to generate comparative static results regarding the relationships between the probability of migration and its determinants, those relationships are implied by comparative statics on expected net present value. For example, if variable  $X$  and expected net present value are positively related, then there should be a positive relationship between  $X$  and the probability of migration. Since the five terms in equation (6) vary in sign, the sign of the marginal effect of age on expected net present value will depend upon the sizes (in absolute value) of these terms. The marginal effect of age on expected net present value may be obtained by integrating equation (6) and differentiating integrated expression with respect to age:

$$\begin{aligned}
\frac{\partial E[NPV(t)]}{\partial \alpha} = & - \left[ \frac{(1-\pi^j)(W_B^j + x^j + s^j) - (1-\pi^i)(W_B^i + x^i + s^i)}{r} \right] \left[ e^{-r\alpha} \left( \frac{1}{\alpha^2} + \frac{r}{\alpha} \right) \right] + \\
& \left[ \frac{g(t-\alpha, \alpha, MS_{ij})(C_{ij}(D_{ij}))}{r} \right] \left[ \left( \frac{1}{\alpha^2} + \frac{r}{\alpha} \right) e^{-r\alpha} \right] - \left[ \frac{\partial g}{\partial \alpha} C_{ij}(D_{ij}) \right] \left( \frac{1}{\alpha} e^{-r\alpha} - \frac{1}{T} e^{-rt} \right) + \\
& (1 - \pi^j)(x^j - x^i) - (1 - \pi^i)(s^i) - \frac{\partial H_{ij}}{\partial \alpha} \tag{7}
\end{aligned}$$

The first term (negative) on the RHS of expression (7) embodies Becker's (1964) explanation that older migrants will have a lower incentive to migrate because, owing to shorter duration in the destination, the present value of earnings gains will be smaller. The second term (positive) captures the idea that because of shorter duration in the destination, an older migrant will make fewer total trips home over his/her lifetime, hence return migration costs will be smaller;

The third term (negative) measures the increase in return migration costs facing older migrants because, all other things equal, they will have stronger ties to family and friends in the origin;

The fourth term (positive) reflects the idea that older migrants will bring more general human capital to the destination. If the returns to general human capital are greater there, then older migrants will benefit more from transferring their general human capital. The gains from transfer will be larger the larger is the premium to general human capital in the destination and the lower is unemployment risk there.<sup>16</sup>

The fifth term (negative) represents the loss to the migrant from abandoning origin investments in specific human capital. The loss will be greater the lower is unemployment risk in the destination and the higher is the return to specific human capital in the origin;

Finally, the sixth term (negative) measures the extra cost to an older migrant of securing local registration.<sup>17</sup>

The theme in the literature up to this point has been that older migrants will always have a lower incentive to migrate. Equations (6) and (7) reveal that older migrants could

<sup>16</sup> Note that this effect has gone unmentioned in the literature.

<sup>17</sup> There is a seventh and positive effect of age on migration, which is not accounted for in the model, and which was suggested by Gallaway (1969): Older migrants, by virtue of having more assets, may find migration relatively more affordable.

sometimes have a stronger incentive. There are terms of different signs in equation (7) but this may not necessarily mean that NPV is not monotonous in age. Whether or not there is monotonicity in the relationship will depend upon how large, in absolute value, the combined effect of the second and fourth terms on the RHS of equation (7) are relative to the combined effect of the other four (negative) terms. While the impact of age on the incentive to migrate due to duration in the destination (captured in the first term on the RHS of equation (7)) may dominate all other terms in the United States, due to the relative ease of relocation between states and cities, for example, that may not be the case in China. In China, travel between two remotely located provinces may be very high, causing the savings in direct return travel costs to be very high for older migrants (the second term on the RHS of equation (7) would be high). Furthermore, differences in the returns to general human capital between those two provinces could be very high, causing the fourth term on the RHS of (7) to be high. The result is that equation (7) would be positive for that case. Ultimately, the sign of the marginal effect of age is an empirical issue.

Our model provides an explanation for why migrants moving long distances, and who switch occupations following a move, tend to be younger. The long distance implies high costs of return migration, hence only those with relatively small investments in ties to family and friends back home will move. It is generally more costly, for example, for rural migrants to obtain urban Hukou and those costs will often rise rapidly with age. These conditions will also tend to discourage older persons from moving;

### **3. SOME PRELIMINARY MICRO EVIDENCE**

The ideal data to examine the relationship between age and the decision to migrate will be longitudinal micro panel data, where we track individuals' migration patterns over time. But such data are often very difficult to collect. Alternatively, though potentially subject to cohort effects, cross sectional micro data might be more realistic. We are able to provide some preliminary evidence on our theoretical model using cross sectional micro data, in this case from the 2005 1% Chinese census. We calculated the percentage of migrants at each age for persons with rural Hukou status. As Figure 1 shows, the migration rate increases sharply from about 6% to 25% till around age 19/20, then it

declines gradually till age 30, after which the decline is relatively more pronounced. In general, the relationship between age and migration rate is highly non-linear.

[INSERT FIGURE 1 ABOUT HERE]

Figure 1 also gives the age distribution of all rural Hukou individuals as of 2005. Two features are quite visible: 1) the sharp dip around age 45 due to the great Chinese Famine (1959-1961); and 2), the gradual but consistently shrinking younger cohorts, in particular in the 20-30 age range.<sup>18</sup> If we consider the two curves of Figure 1 together, then it appears that there is a long run declining trend of rural out-migration due to a declining fertility rate, or, in other words, aging in rural areas.

In Table 1, we conducted several regressions to examine the effect of age on migration for the rural Hukou population using the following linear probability model:

$$\begin{aligned} Migrate_{ij} = & \alpha_0 + \alpha_1 Age_{i1} + \alpha_2 Age_{i2} + \alpha_3 Age_{i3} \\ & + \sum_k \beta_k Edu_{ik} + \Omega_1 Demo_i + \Omega_2 Sibling_i + v_j + \mu_{ij} \end{aligned} \quad (8)$$

where  $Migrate_{ij}$  is dummy equaling one if individual  $i$  from Hukou region  $j$  is currently living outside her/his Hukou region. Our key explanatory variable (age) enters the regression using a piecewise-linear construction. Ages 19 and 30 are used as the kink points here. This way, we highlight the nonlinear relationship between age and migration in a transparent and parsimonious fashion. Let individual  $i$ 's age be  $Age_i$ , then the first three variables on the RHS of equation (8) are defined as,

$$Age_{i1} = \begin{cases} Age_i & \text{if } Age_i \leq 19; \\ 19 & \text{otherwise.} \end{cases} \quad (9)$$

$$Age_{i2} = \begin{cases} 0 & \text{if } Age_i \leq 19; \\ Age_i - 19 & \text{if } 19 < Age_i \leq 30; \\ 30 - 19 = 11 & \text{if } Age_i > 30; \end{cases} \quad (10)$$

$$Age_{i3} = \begin{cases} 0 & \text{if } Age_i \leq 30 \\ Age_i - 30 & \text{if } Age_i > 30 \end{cases} \quad (11)$$

We constructed 4 categories of educational attainment, with the default group being those that attended up through junior/ middle school. The three categories in the regression as  $Edu_{ik}$  are: 1) never attended school; 2) attended up through elementary school; and 3) attended up through high school. There is only a very small proportion of

<sup>18</sup> Note, since we exclude all students, the age distribution for the 15-20 years age range is not representative of the pool of potential migrants.



individuals with rural *Hukou* who attended up through college or above, so those observations were eliminated. We also included dummies for female, minority nationality, and never married, in some specifications. These are referred to as  $Demo_i$  in equation (8). In order to investigate the impact of fertility on migration, we also included number of siblings in the regression equation. The set of sibling-related variables is referred to as  $Sibling_i$  in equation (8). This set contains 5 variables: 1) a dummy for no brothers; 2) dummy for no sisters; 3) number of brothers; 4) number of sisters; and 5) a dummy if there was missing information on number of siblings. Finally, given the high likelihood of correlation of migration with regional factors, we included the regressor  $v_j$  to capture hukou registration areas' fixed effects.

[INSERT TABLE 1 ABOUT HERE]

Table 1 shows regressions results. In specification (the results which are shown in column 1), we start with a basic specification that includes only three age variables and Hukou region fixed effects. Based on that, we add the education dummies  $Edu_{ik}$  to specification (2). Then, we further add the three most commonly considered demographic variables of  $Demo_i$  in specification (3). Finally, in specification (4), we add the five  $Sibling_i$  variables. As the table shows, the impact of age on migration is highly nonlinear, even after controlling for various demographic personal characteristics and regional fixed effects. The likelihood of migration increases with age dramatically for teenagers, flattens in the twenties, and then decreases after age 30. Table 1 also shows that the presence of brothers and sisters is positively related to migration likelihood.

Overall, our examination using census micro data implies that aging can have ambiguous effects on migration. When the peak of the age distribution is still in the lower 20s, there can be an increasing trend for migration if the overall population is only starting to experience aging, lower fertility rates, and longer life expectancy (probably as individuals have larger general human capital endowments to transfer). But as the population gets older and older, there will be less and less migration (probably as other factors become more relevant).

#### 4. EVIDENCE FROM PROVINCE-LEVEL DATA

We now test our theoretical model using province-level data. At the aggregate level, the theoretical model implies a double-log empirical specification where the dependent variable is the log of the migration rate, defined as the number of persons moving from place  $i$  to place  $j$  as a percentage of all persons moving out of  $i$ .<sup>19</sup> In contrast to most previous studies of China, ours is a panel study spanning multiple periods of migration. Bao, Bodvarsson, Hou and Zhao (2011) is the only other study that examines China's migration using panel data.<sup>20</sup> Here, we extend their original specification by adding measures of the origin province's age distribution as follows,

$$\ln\left(\frac{Migrants_{ijt}}{\sum_k Migrants_{ikt}}\right) = \alpha_0 + \alpha_1 \ln(Age_{it_0}) + \alpha_2 \ln(Age_{it_1}) + \alpha_3 \ln(Hukou\ Percentage_{jt}) + \alpha_4 \ln(Distance_{ij}) + \alpha_5 \ln(1 - u_{it}) + \alpha_6 \ln(1 - u_{jt}) + \alpha_7 \ln\left(\frac{Per\ capita\ Income_{jt}}{Per\ capita\ Income_{it}}\right) + \alpha_8 1(neighbour_{ij}) + \beta Z_{ij} + v_j + \mu_t + \varepsilon_{ijt} \quad (12)$$

Where  $\ln\left(\frac{Migrants_{ijt}}{\sum_k Migrants_{ikt}}\right)$  is the log of the proportion of migrants from origin province  $i$  to destination province  $j$  to all migrants from province  $i$  in year  $t$ . Our key variable of interest is  $\ln(Age_{it_1})$  which measure the age distribution of the origin province at the end of the period. We also control for the same age measure for province  $i$  at the start of the period  $\ln(Age_{it_0})$ . Controlling for the initial age distribution, we are able to directly examine the impact of aging on migration. The age distribution measure here can take three forms: (i) the mean age of the province's population; (ii) provincial youth share and (iii) provincial age dependency ratio (ADR).

In general, our choice of control variables follows Lin, Wang and Zhao (2004), Bao, Hou and Shi (2006), Poncet (2006), and Bao, Bodvarsson, Hou, and Zhao (2011). Following Bao, Bodvarsson, Hou and Zhao (2011), we control for the probability of

<sup>19</sup> This specification, widely used in the literature, is due originally to Greenwood (1969).

<sup>20</sup> These authors in turn base their specification on earlier work by Greenwood (1969), Lin, Wang and Zhao (2004) and Poncet (2006).

securing Hukou in the destination using  $\ln(Hukou\ Percentage_{jt})$ , the lagged relative frequency of registered households. It is assumed that prospective migrants know the historical relative frequencies of registered households in destination provinces and have adaptive expectations about barriers to entry. When the likelihood of securing Hukou rises, perceived benefits to migration will rise and that will lead to a higher migration rate. Following Poncet (2006), we also control for migration flows between adjacent provinces by including a dummy ( $neighbour_{ij}$ ), which equals one if migration flows are between adjacent provinces.

The other control variables are<sup>21</sup> the following:

- (1)  $\ln(Distance_{ij})$ : log of railroad distance between the two provinces in kilometers;
- (2)  $\ln(1 - u_{it})$ : log of one minus unemployment rate in the source province in period t;
- (3)  $\ln(1 - u_{jt})$ : log of one minus unemployment rate in the destination province in period t;
- (4)  $\ln\left(\frac{\text{per capita income}_{jt}}{\text{per capita income}_{it}}\right)$ : log of the ratio of destination province income to origin province income in period t;
- (5) log ratio of real FDI per capita in the destination province to real FDI per capita in the origin province;
- (6) log ratio of real domestic fixed asset investment per capita in the destination province to real domestic fixed asset investment in the original province;
- (7) log percentage of population enrolled in the origin province's universities;
- (8) log percentage of population enrolled in the destination province's universities;
- (9) log ratio of the share of manufacturing employment in the destination province to the share of manufacturing employment in the origin province;
- (10) log ratio of the urban share of the destination province's population to the urban share of the origin province's population;
- (11) log ratio of the destination province's minority population share to the origin province's minority population share
- (12) log ratio of mean annual temperature in the capital city of the destination province to the mean annual temperature in the capital city of the origin province.

Finally, we include  $v_j$  and  $\mu_t$ , the destination province fixed and period fixed effects, respectively, in all of our regressions.

Although we have a panel data set, there could still be omitted variables, which would lead to biased estimation of the coefficients on our age measures. For each of the three age measures, we estimate three different specifications: OLS without destination

---

<sup>21</sup> Due to limited space, we did not report the estimated coefficients for some of the variables in the tables.

province fixed effects, OLS with destination province fixed effects, and 2SLS with destination province fixed effects.<sup>22</sup> The two instruments used in 2SLS are: (1) the number of doctors per 10,000 persons in the source province; and (2) the share of minority (non-Han) population in the source province. We conjecture that doctor density is an indicator of the level of public health spending in the province. When doctor density is higher, residents will ultimately be healthier, infant mortality will be lower, and death rates, particularly among the elderly, will be lower. This could result in a shift in the age distribution. If the non-Han population has different fertility rates due to either cultural differences or/and exemption from family planning policies, then the share of minority population could also affect the age distribution of source province.<sup>23</sup>

## 5. DATA

Our data set includes the 1985-90, 1995-2000, and 2000-05 periods.<sup>24</sup> All data are drawn from two major sources. For the 1985-90 and 1995-2000 periods, we expand the data set used by Lin, Wang and Zhao (2004) in their study of interprovincial migration.<sup>25</sup> Data for 2000-05 are taken from University of Michigan's *China Data Online* website (<http://www.chinadataonline.org/>). For all regressions, only observations with positive migration rates are kept. Like Lin, Wang and Zhao (2004) and others, we exclude Tibet

---

<sup>22</sup> Given our definition of the dependent variable, the sum of the dependent variable for each origin province will be a constant (one). Therefore, it is mechanically not meaningful to include origin province fixed effects here. Also, given that our key variable here is origin province-specific, it is also not feasible to include origin province fixed effects.

<sup>23</sup> Note that we assume both instruments will not affect the share of migration from source to destination provinces through channels other than their impacts on the age structure of the source province, which seems reasonable given that we have also included log ratio of the destination province's minority population share to origin province's minority population share in our regression.

<sup>24</sup> We skipped the 1990-95 period for a number of important reasons. First, there are some incompatibilities between the data for that period and for the other periods. Most notably, data on the migration rate for 1990-95 are less accurate, plus there are missing series for other variables we need to include in our regressions. Second, one benefit of studying 1985-2005 in China is that, in estimating a panel equation, one can observe structural changes in migration due to the intensification of reforms. Imposing a five-year break in the data makes it easier to ascertain whether or not structural change occurred. The structure of China's post-reform labor market has been changing at an exponential pace, with at least as much (if not more) structural change occurring between 1995-2000 and 2000-05 as between 1985-90 and 1995-2000. Therefore, imposing the five-year break when comparing the 1980s and 1990s, as opposed to no break between the 1990s and 2000s, should not bias the coefficient estimates.

<sup>25</sup> Note that we replaced Lin, Wang and Zhao's (2004) calculations of the dependent variable with our own calculations. The reason is that there are some inaccuracies in the series used by Lin, Wang and Zhao, which they acknowledged in communications with us.

from our sample because of data gaps and treat Chongqing as part of Sichuan. Our sample thus includes 29 provinces, each a prospective destination and point of origin.

A major drawback of regressing 1985-90 migration rates is that historical data on the relative size of the migrant network in the destination are not available. The reason is that the 1990 semi-decennial census was the first to include questions about change in residence, so if one wanted to study how migration during 1985-90 was influenced by migrant stock, information on pre-1985 migration would have had to be available. However, information about migrant networks is available when estimating migration for 2000 and later. One could, for example, regress migration rates during 1995-2000 on the stock of migrants that accumulated in the destination since 1985-90. Consequently, we produced two sets of estimates: (i) estimates for the full panel with no control for past migration; and (ii) estimates for a smaller panel comprising the latter two periods only, which do include a control for past migration. There are 2,383 usable observations in the full panel, of which 765, 810 and 808 are from period 1985-1990, 1995-2000 and 2000-05, respectively. The smaller panel of the latter two periods has 1,535 observations and includes 784 observations for 1995-2000 and 751 observations for 2000-05. The reason that fewer observations for each period are used in the smaller panel is, that we further dropped observations for which past migration is zero.

[INSERT TABLE 2 ABOUT HERE]

Tables 2 shows summary statistics for all variables used in our regressions for each of the three migration periods. Starting from the top of each table, we describe the variable, the data source from which the variable is taken, and trends apparent in the data:

(i) *Gross interprovincial migration rate*. For the 1985-90, 1995-2000 and 2000-05 periods, respectively, migration rates are calculated from samples comprising 1% of the 1990 population census, 0.95% of the 2000 census,<sup>26</sup> and 1% of the 2005 census. In the

---

<sup>26</sup> As pointed out by Lin, Wang and Zhao (2004), there is a small difference between the 1990 and 2000 censuses with respect to how migration is defined. If a person is observed to change residence *and* to change their household registration (a situation officially called “Hukou migration”), then this movement is officially classified as “migration” in both censuses. If, however, the person is observed to change residence without changing registration (“non-Hukou migration”), then the movement is classified as “migration” only if the migrant has been away from the place of registration for a minimum period of time. In the 2000 census, this period is 6 months, but in the 1990 census it is one year. To account for this change in classification between the two periods, the migration numbers in both periods were standardized by

1990 (2000, 2005) census, respondents were asked to report on migration activities during 1985-90 (1995-2000, 2000-05, respectively). Consequently, migration rates during each decade were calculated for only the second half of each decade. The mean volume of migration from one province to another surged from over 355,000 persons during 1985-90 to over 1,075,000 during 1995-2000 and to over 2,200,000 during 2000-05.<sup>27</sup> Note that mean provincial population rose by 9.44% between 1990 and 2000 and by 5.86% between 2000 and 2005;

(ii) *The three age distribution measures.* Data for the three age distribution measures (mean age, youth share, and age dependency ratio (ADR)) were calculated from census information on provincial age distributions. We used 1982 (1995, 2000) data as the start-of-period measures for the 1985-90 (1995-2000, 2000-05) period of migration. We used 1990 (2000, 2005) data as the end-of-period measures for the 1985-90 (1995-2000, 2000-05) periods of migration. Note that for mean provincial age, census information for 1982 provides the percentage of population for each year of age. Consequently, we computed a weighted mean age, where the weights are the shares of the population of a particular age. For subsequent years, census information provides the percentage of population only for 5-year intervals. In those cases, we computed weighted mean age by using the midpoint of age in each interval;

(iii) *Size of the community of migrants from the origin who reside in the destination.* An ideal measure of the size of the destination's migrant community is the current stock of migrants from the origin as a percentage of current population in the destination. Unfortunately, unlike many Western data sets, this type of migrant stock measure is unavailable for China. Therefore, we used relative migrant flows during the half-decade ending five years prior to the migration period. For the 1995-2000 (2000-05) period, size of migrant network was calculated by the ratio of total flows from origin to destination during 1985-90 (1990-95) to the destination's population in 2000 (2005). There are several reasons for this approach. First, it is presumed that the stock of previous migrants is proportional to the size of the previous flow of migrants. Second, by lagging past flows

---

discounting the 2000 numbers by a small amount, approximately 5%. For further details, see Lin, Wang and Zhao (2004, page 593).

<sup>27</sup> There are likely to be discrepancies in the calculations of these numbers between decades, for the reasons discussed in the preceding footnote.

by 5 years, we hopefully reduce the risk of serial correlation. As Table 2 shows, the estimated provincial migrant stock averaged approximately 1.35 million persons for the 1995-2000 migration period and approximately 1.22 million persons for the 2000-05 period. The reduction in migrant stock could be due to return migration;

(iv) *The historical relative frequency of persons with local Hukou.* This is the ratio of the registered population to total (registered + unregistered) population at year's end. For the 1985-90 (1995-2000, 2000-05) period, we use the mean annual proportion of persons with Hukou during 1980-84 (1990-94, 1995-99). By using the lagged proportion of persons with Hukou, we avoid potential problems with simultaneous equations bias;

(v) *Mean real per capita income.* Due to lack of available data for consecutive years during the 1980s and 1990s, income data only for 1989 (1999) were used to measure average annual income for the 1985-90 (1995-2000) periods. For 2000-05, though, we use annual mean incomes. All income data are adjusted for cost of living differences using provincial CPI measures;

(vi) *The share of the province's population that is Minority (non-Han).* Because data on Han population shares for 1990 are not available, we used 2000 data to proxy minority shares for the first two migration periods. For the most recent migration period, we used information on Han population shares from the 2005 census;

(vii) *Real annual FDI and FAI per capita.* For each period, we used mean annual real FDI (FAI) per capita during 1980-84 when regressing 1985-90 migration flows, 1990-94 mean annual real FDI (FAI) per capita when regressing 1995-2000 migration flows, and 1995-99 mean real FDI (FAI) per capita when regressing 2000-05 migration flows. We lagged investment spending because it typically takes time for migration to respond to changes in spending on investment projects. Furthermore, since there is very likely to be two-way causality between investment and migration, by regressing migration rates on lagged investments we hoped to avoid possible simultaneous equations bias. We adjusted the investment series for cost of living differences between the two decades, as well as across provinces within each decade, using national government measures of provincial CPI and calculating both series at 1985 price levels. For most of the provinces, FDI numbers were available for each year, but for some there were missing years. For several provinces, no investment data were available for 1980-84, so

we used the earliest year available as a proxy for that period. Therefore, our coefficient estimates for the early period may be influenced by measurement error in some parts of the investment series. Note that the FDI series is in USA dollars, whereas the fixed asset investment series is in Chinese Yuan;

(viii) *The share of manufacturing employment.* Manufacturing is classified in China as a “Secondary” industry and construction is one of its components. There is considerable variation across the country with respect to the dominance of manufacturing in the provincial labor market;

(ix) *Mean level of educational attainment.* Educational attainment was measured as the percentage of the population aged 22-60 enrolled in universities in 1990 (for the 1985-90 period), in 2000 (for the 1995-2000 period), and in 2005 (for the latest period). For all three periods, a large majority of a typical province’s adult population was not enrolled in universities, due to substantial barriers to post-secondary education in China. However, as reforms deepened and barriers to access fell, the percentage of the population enrolled at universities rose at an increasing rate, from over 3% in 1990 to nearly 9% in 2005. Note also that the variance of enrollment rose at an increasing rate, a likely explanation for rising income inequality in China.

Data on the remaining variables are from Lin, Wang and Zhao (2004). Please refer to their paper for details on data sources and measurement of these variables.

## 6. COEFFICIENT ESTIMATES

Table 3 includes all the estimated effects of log mean age on the interprovincial migration rate in 6 specifications, 3 for the full panel and 3 for the last two periods only. Column (1) of Table 3 serves as a benchmark. It corresponds to OLS without destination province fixed effects included. Then, we add destination province fixed effects in column (2). Finally, we consider the possibility of endogeneity in our key measure at the end of the period.<sup>28</sup> Columns (4)-(6) replicate the first three specifications with two changes: (1) the sample is limited to the last two periods only; and (2) previous migration flows are added as a control variable. Given the high likelihood of path dependence in

---

<sup>28</sup> In other words, we take the start-of-period age measure as a given and use the two instruments to extract the exogenous change of the age measure over this period. Both IV estimates are statistically significant at the 1% level in the first stage.



migration flows, we consider the 2SLS estimation based on the later 2 periods as more robust. The other specifications serve as robustness checks. Tables 4 and 5 replicate Table 3's design with the key variables changed to youth share and provincial ADR respectively.

[INSERT TABLE 3 ABOUT HERE]

From Table 3, we find strong evidence that *ceteris paribus*: (i) interprovincial migration flows are stronger from provinces that have on average *older* populations at the start of the period; and (ii) these flows are also stronger from provinces that have on average younger populations at the end of the period. In other words, aging in the origin province is correlated with less migration, while provinces endowed with relatively mature populations are sending out more migrants in general. Considering that average age increases from around 21 to 27 over our study periods, as well as our micro evidence, these two aspects of the age-migration relationship seem quite intuitive. Recall that our micro data show that it is persons in the 20-30 year old age range that comprise the bulk of migrants.

According to column (6) of Table 3, a one percentage point increase in the end-of-period average age of residents in origin province  $i$  is associated, all other things equal, with a 22% drop in the emigration rate from province  $i$  to province  $j$ . Assume for the sake of illustration that the average fraction of  $i$ 's population migrating to province  $j$  is 4% per year. Then, the coefficient estimate implies that that, all other things equal, the emigration rate to province  $j$  will drop to approximately 3.12%, probably due to the aging of the potential migration pool. On the other hand, a one percentage point increase in the start-of-period average age in province  $i$  will lead to an emigration rate of 4.52%, probably due to the province having a more mature labor endowment.<sup>29</sup> Such a two-sided pattern highlights the nonlinear relationship between age and migration, as discussed both in the theoretical section and the section on preliminary micro evidence. In the long run, we would then expect the migration rate to be decreasing as the population ages.

The effect of age on migration is quite consistent across the various specifications in Table 3, with the magnitude considerably larger for 2SLS specifications and for the later

---

<sup>29</sup> While not incorporated into our theoretical model, there is also Gallaway's (1969) point that older migrants may find migration more affordable, owing to greater assets.

2 periods. Table 3 also shows strong path dependence with consistently and significantly positive coefficients for previous migration flows. The estimated coefficients on other variables are in general quite intuitive, although we are reluctant to interpret them as causal, given that the focus of our empirical strategy is on age structure. It is worth noting that the coefficients on the log odds of Hukou status are negative, contrary to our prior assumptions. One possibility could be that the migration decision is less motivated by acquiring Hukou than by other factors. Provinces with larger shares of persons with Hukou status could be less friendly to temporary migrants in terms of labor market opportunities, as they might have greater incentive to protect the employment of local population.

[INSERT TABLE 4 ABOUT HERE]

Just as Table 2 shows, the mean age is increasing while the youth share and ADR are decreasing during our study period. Table 4 examines how youth share and migration rate are linked. Again, we include log of youth share at the start of the period together with that at the end of the period. According to column (6), the coefficient implies that a one percentage point decrease in youth share will lead to an 8.4% drop in the migration rate. Assuming the migration rate is 4%, this implies a reduction to 3.67%, *ceteris paribus*. Table 5 examines how ADR affects migration using exactly the same specifications as Table 3 and Table 4. The results are highly consistent across all these three tables.

[INSERT TABLE 5 ABOUT HERE]

Our aggregate results are also consistent with the micro data results. All suggest that aging of the population will lead to less migration in the long run. Although, our theoretical section does not predict any monotonic relationship between age and migration, our micro evidence does suggest a highly non-linear relationship between these two variables. Our results suggest that the particular age structure of China is such that aging will lead to decreasing migration, because the population of persons aged 20-30 will eventually experience a downward trend.

## 7. CONCLUDING REMARKS

We view this study as making two important contributions, one to the literature on migration in China and the other to the Western literature on migration theory. First, we link together two major long term demographic shifts in China – the post-reform surge in internal migration and the shifting of China’s age distribution. We find strong evidence that the Great Chinese Internal Migration has been influenced by shifts in the national age distribution. We believe our results for the effects of age distribution are important because they clearly demonstrate that shifts in the age distribution in China are capable of generating often sizeable changes in migration rates.

China’s aging appears to have very important implications for labor mobility. This is compounded by the fact that China’s labor force is expected to shrink starting around 2015-16, as those that entered the labor force at the introduction of the one-child policy are slated to retire *en masse* and will be replaced by a much smaller group of new entrants. The industrial base that has concentrated on the eastern coastal provinces will be under extra pressure, as the migrant workers they have become reliant upon will start to dwindle. This is why, for example, the World’s largest electronic components manufacturer, Foxconn, has left its home base of Longhua, Shenshen (across from Hong Kong), where it employs up to 430,000 workers, to build plants in Zhenzhou, Henen (China’s most populous province), employing 120,000 workers and in Chengdu, Szechuan (China’s second most populous province). Clearly, if our long run projection of the decline in migrant flows as the population ages is correct, a geographic reshuffling of China’s industrial facilities will be in her future, which will have major public policy implications, e.g. implications for taxation, infrastructure, and urban planning. Though beyond the scope of this study, this implication cannot be ignored and research into this aspect is most certainly warranted.

The second contribution of the paper, which is important for both the case of China and the Western migration literature at large, is that we analyzed the theoretical relationship between migrant age and the net benefits of migration. With the exception of Lundborg (1991), who did derive some testable implications for age in his migration model, we believe our study goes the farthest in examining how age affects the expected benefits and costs of relocation. We contend that the traditional Becker (1964) view that younger migrants always have a higher likelihood of moving is just one set of important,

conflicting explanations. We argue that it is important to also consider the effects of age on psychic costs, out-of-pocket costs, the loss of specific assets, and spatial differences in the returns to different types of human capital. Furthermore, for the China case one must also consider the effects of age on the endowment of general human capital. Our model demonstrated that when one blends all these explanations together, it is not generally true that older migrants have a lower propensity to migrate. This may explain the diversity of results for the age variable across both Western and Chinese studies.

We recommend that the next steps in this research are: (1) to apply the theoretical model to Western cases of internal migration; (2) test the model on micro data from carefully designed household surveys in different regions and in different policy environments; and (3) use the results obtained from estimating our migration equations to carefully construct a model and test of age distribution with the ultimate goal of testing a general equilibrium theory of migration and age distribution. Furthermore, as new waves of census data become available in China, studies of structural change in migration, particularly as it relates to aging, will become more feasible.

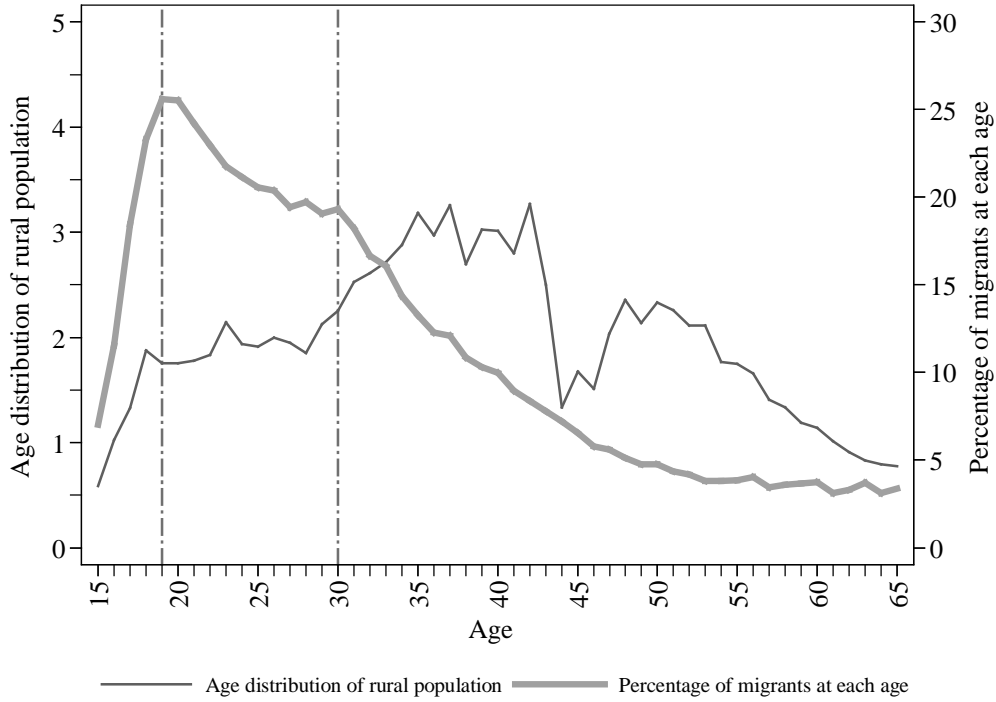
## REFERENCES

- Bao, S., Bodvarsson, Ö., Hou, J., & Y. Zhao (2008a). The Deregulation of People Flows in China: Did the Structure of Migration Change? *The Indian Journal of Labour Economics*, 51, 587-602.
- Bao, S., Bodvarsson, Ö., Hou, J., & Y. Zhao (2008b). Migration in China During 1985-2000: The Effects of Past Migration, Investments and Deregulation. *The Chinese Economy*, 42, 7-28.
- Bao, S., Bodvarsson, Ö., Hou, J., & Y. Zhao (2011). The Regulation of Migration in a Transition Economy: China's Hukou System. *Contemporary Economic Policy*, 29, 564-79.
- Bao, S., Hou, J., & Shi, A. (2006). Migration and Regional Development in China. In S. Bao, S. Lin and C. Zhao (Eds.), *Chinese Economy after WTO Accession*. Aldershot, UK: Ashgate.
- Becker, G. (1964). *Human Capital: A Theoretical and Empirical Analysis*. New York: Columbia University Press.
- Borjas G. (1987). Self-Selection and the Earnings of Immigrants. *American Economic Review*, 77, 531-53.
- Bowles, S. (1970). Migration as Investment: Empirical Tests of the Human Investment Approach to Geographical Mobility. *Review of Economics and Statistics*, 52, 356-62.
- Chen, A., & Coulson, N.E. (2002). Determinants of Urban Migration: Evidence from Chinese Cities. *Urban Studies*, 39, 2189-97.
- Clark, X., Hatton, T., & Williamson, J. (2007). Explaining U.S. Immigration, 1971-1998. *Review of Economics and Statistics*, 89, 359-73.
- David, P.A. (1974). Fortune, Risk and the Micro-Economics of Migration. In P.A. David and M.W. Reder (Eds.), *Nations and Households in Economic Growth*. New York, NY: Academic Press.
- Fan, C.C. (2005). Modeling Interprovincial Migration in China, 1985-2000. *Eurasian Geography and Economics* 46, 165-84.
- Galloway, L. (1969). Age and Labor Mobility Patterns. *Southern Economic Journal*, 36, 171-80.
- Goodkind, D., & West, L.A. (2002). China's Floating Population: Definitions, Data and Recent Findings. *Urban Studies*, 39, 2237-50.
- Goss, Ed., & C. Paul (1986). Age and Work Experience in the Decision to Migrate. *Journal of Human Resources*, 21, 397-405.
- Greenwood, M.J. (1969). An Analysis of the Determinants of Geographic Mobility in the United States. *Review of Economics and Statistics*, 51, 189-94.

- Hare, D. (1999). 'Push' versus 'Pull' Factors in Migration Outflows and Returns: Determinants of Migration Status and Spell Duration Among China's Rural Population. *Journal of Development Studies*, 35, 45-72.
- Li, H., & Zahniser, S. (2002). The Determinants of Temporary Rural-to-Urban Migration in China. *Urban Studies*, 39, 2219-35.
- Liang, Z. (2001). The Age of Migration in China. *Population and Development Review*, 27, 499-524.
- Liang, Z., Chen, Y.P., & Gu, Y. (2002). Rural Industrialisation and Internal Migration in China. *Urban Studies*, 39, 2175-87.
- Liang, Z., & White, M.J. (1996). Internal Migration in China, 1950-88. *Demography*, 33, 375-84.
- Liang, Z., & White, M.J. (1997). Market Transition, Government Policies, and Interprovincial Migration in China: 1983-1988. *Economic Development and Cultural Change*, 45, 321-39.
- Lin, J., Wang, G., & Zhao, Y. (2004). Regional Inequality and Labor Transfers in China. *Economic Development and Cultural Change*, 52, 587-603.
- Lundborg, P. (1991). An Interpretation of the Effects of Age on Migration: Nordic Migrants' Choice of Settlement in Sweden. *Southern Economic Journal*, 58, 392-405.
- Ma, Z., & Liaw, K. (1997). Explaining Hierarchical and Interprovincial Migrations of Chinese Young Adults by Personal Factors and Place Attributes: A Nested Logit Analysis. *Mathematical Population Studies*, 6, 217-39.
- Naskoteen, R.A., & Zimmer, M. (1980). Migration and Income: The Question of Self-Selection. *Southern Economic Journal*, 46, 840-51.
- Navratil, F., & J. Doyle (1977). The Socioeconomic Determinants of Migration and the Level of Aggregation. *Southern Economic Journal*, 43, 1547-59.
- Poncet, S. (2006). Provincial Migration Dynamics in China: Borders, Costs and Economic Motivations. *Regional Science and Urban Economics*, 36, 385-98.
- Roy, A. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, 3, 135-46.
- Schwarz, A. (1976). Migration, Age, and Education. *Journal of Political Economy*, 84, 701-19.
- Schlottmann, A., & H. Herzog (1981). Employment Status and the Decision to Migrate. *Review of Economics and Statistics*, 63, 590-98.
- Schlottmann, A., & H. Herzog (1984). Career and Geographic Mobility Interactions: Implications for the Age Selectivity of Migration. *Journal of Human Resources*, 19, 72-86.
- Shi, A., & Bao, S. (2007). Migration, Education and Rural Development: Evidence from China 2000 Population Census Data. *Journal of Chinese Economic and Business Studies*, 5, 163-77.

- Shi, X., Heering, N., & Futian, Q. (2007). Choices Between Different Off-farm Employment Sub-categories: An Empirical Analysis for Jiangxi Province, China. *China Economic Review*, 18, 438-55.
- Wu, Z. (2008). Self-selection and Earnings of Migrants: Evidence from Rural China. *Discussion Papers in Economics and Business from Osaka University*, No. 08-24.
- Zhao, Y. (1997). Labor Migration and Returns to Rural Education in China. *American Journal of Agricultural Economics*, 79, 1278-87.
- Zhao, Y. (1999a). Labor Migration and Earnings Differences: The Case of Rural China. *Economic Development and Cultural Change*, 47, 767-82.
- Zhao, Y. (1999b). Leaving the Countryside: Rural-to-Urban Migration Decisions in China. *American Economic Review*, 89, 281-86.
- Zhao, Y. (2002). Causes and Consequences of Return Migration: Recent Evidence from China. *Journal of Comparative Economics*, 30, 376-94.
- Zhao, Y. (2003). The Role of Migrant Networks in Labor Migration: The Case of China. *Contemporary Economic Policy*, 21, 500-11.
- Zhu (2002). The Impact of Income Gaps on Migration Decisions in China. *China Economic Review*, 13, 213-30.

**Figure1.**  
**Age Distribution and Migration Likelihood for Rural Population**



Note: The data set used is the micro 2005 1% census data of China. Individuals with rural Hukou, aged 15-65, healthy, not currently in school, and with education levels below or equal to high school, are included here. Individuals with a high school education but younger than 18 years of age are dropped. Migrants are those currently not living in the Hukou registration area, the type 2 and 3 individuals in the census.



**Table 1. Effect of Age on Migration for Rural Hukou, Linear Probability Model**

Dependent Variable: Whether migrate

	(1)	(2)	(3)	(4)
Age <sub>1</sub> (when 15≤age≤19)	0.0450*** (0.0019)	0.0405*** (0.0018)	0.0415*** (0.0019)	0.0372*** (0.0017)
Age <sub>2</sub> (when 19<age≤30)	0.0049*** (0.0003)	0.0048*** (0.0003)	0.0048*** (0.0003)	-0.0002 (0.0005)
Age <sub>3</sub> (when 30<age≤65)	-0.0102*** (0.0006)	-0.0094*** (0.0005)	-0.0076*** (0.0005)	-0.0049*** (0.0005)
<i>Education (the default group is junior middle school)</i>				
never attended	--	-0.0160*** (0.0033)	-0.0189*** (0.0034)	-0.0198*** (0.0034)
elementary school	--	-0.0311*** (0.0023)	-0.0320*** (0.0022)	-0.0312*** (0.0022)
high school	--	-0.0287*** (0.0023)	-0.0299*** (0.0022)	-0.0305*** (0.0022)
<i>Other demographics</i>				
Female	--	--	0.0098*** (0.0015)	0.0088*** (0.0015)
Minority	--	--	-0.0415*** (0.0053)	-0.0427*** (0.0052)
Single	--	--	-0.0415*** (0.0053)	-0.0423*** (0.0052)
<i>Number of Siblings</i>				
no brothers	--	--	--	-0.0218*** (0.0034)
no sisters	--	--	--	-0.0039 (0.0031)
# of brothers	--	--	--	0.0101*** (0.0025)
# of sisters	--	--	--	0.0046*** (0.0018)
Adjusted R2	0.0778	0.0846	0.0858	0.0871

Note: The data set used is the micro 2005 1% census data of China. The dependent variable is the dummy for migration defined as currently not living in the Hukou registration area, in other words, the type 2 and 3 individuals in the census. Individuals with rural Hukou, aged 15-65, healthy, not currently in school, and with education levels below or equal to high school are included here. Individuals with high school educations but younger than 18 are dropped. The number of observations is 1,149,090 for all four specifications. The impacts of age on migration are captured here using  $Age_1$ ,  $Age_2$  and  $Age_3$  for the age ranges [15,19], (19,30] and (30,65], respectively. The detailed specification is discussed in the text. All specifications also consider Hukou registration areas' fixed effects. Specification (4) also includes a dummy for missing information on number of siblings. Standard errors are clustered by 345 Hukou registration areas. \*, \*\*, \*\*\* refers to significant at 10%, 5% and 1% levels, respectively.

**Table 2. Summary Statistics**

	1985-1990		1995-2000		2000-2005	
# of observations	765		810		808	
Variable	Mean	s.d.	Mean	s.d.	Mean	s.d.
Migration rate (%) <sup>30</sup>	3.77	6.98	3.51	7.16	3.59	7.32
Start year mean provincial age	20.84	1.81	23.60	1.69	25.04	1.81
End year mean provincial age	22.42	1.75	25.04	1.81	26.98	1.98
Start year youth share (%) <sup>31</sup>	29.55	2.26	26.52	2.97	25.83	2.29
End year youth share (%)	30.83	2.44	25.83	2.29	21.76	2.81
Start year ADR (%) <sup>32</sup>	62.66	11.48	49.92	7.07	42.55	6.85
End year ADR (%)	50.07	7.32	42.55	6.85	39.66	7.88
Migration flows 10 years earlier ('0000) <sup>33</sup>			135.14	343.93	121.81	353.18
Mean annual % of households with <i>Hukou</i> status during 1980-84	98.40	1.41	90.39	5.40	88.61	6.68
Railway distance between capital cities (kilometers)	1,902	1,002	1,963	1,027	1,959	1,025
Real annual per capita income (Yuan) <sup>34,35</sup>	510.95	183.11	1,963	1,027	5091.80	2617.52
Unemployment rate (%)	1.18	0.70	5.90	3.54	3.13	1.50
Minority population share (%)	11.91	15.89	14.09	5.08	13.08	16.63
Doctors per 10,000 persons	5.40	2.33	12.57	16.18	6.54	3.26
Real Mean Annual FDI Per Capita during previous 5 years (Yuan) <sup>35</sup>	2,757	6,985	55,247	93,960	145,620	231,548
Real Mean Annual Fixed Asset Investment (FAI) Per Capita during previous 5 years (Yuan) <sup>36</sup>	43.74	22.41	306.92	235.60	849.57	635.37
% of adult population enrolled in universities	3.05	3.01	1,063.06	440.21	8.77	4.89
Manufacturing share of employment (%)	24.44	12.09	4.40	2.41	22.64	9.83
Urban share of population (%)	31.03	16.17	22.61	9.83	40.04	18.43
Mean yearly temperature (Celsius)	14.04	4.99	40.05	18.41	14.13	5.26

<sup>30</sup> Here, a migrant is defined as someone who has moved from another town or township and has lived there for between one and five years.

<sup>31</sup> This is defined as the percentage of population aged 15-29.

<sup>32</sup> ADR refers to age dependency ratio, the ratio of (1) number of individuals age less or equal to 14 or more or equal to 65 to (2) number of individuals age 15 to 64.

<sup>33</sup> Computed for the periods of 1985-90 and 1995-2000, respectively.

<sup>34</sup> Computed using average annual income and average annual CPI for 1989, 1999 and 2000-05, respectively.

<sup>35</sup> Computed for the periods of 1985-90 and 1995-2000, respectively.

<sup>36</sup> Computed using average annual FDI, FAI and CPI for the periods of 1980-84, 1990-94, 1995-99, respectively.

**Table 3. Effects of Log Mean Age on Interprovincial Migration Rate**

	Full Sample 1985-90, 1995-2000 and 2000-05			Later 2 Periods 1995-2000 and 2000-05		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	2SLS	OLS	OLS	2SLS
log(mean age at the start of period)	3.814*** (0.820)	2.203*** (0.384)	6.968** (2.881)	1.891*** (0.592)	0.944* (0.530)	13.353*** (2.580)
log(mean age at the end of period)	-2.921*** (0.826)	-4.355*** (0.548)	-10.015*** (3.698)	-5.278*** (0.763)	-7.080*** (0.653)	-21.951*** (2.985)
Log(# of migration flows from origin to destination province 10 years earlier)	---	---	---	0.496*** (0.035)	0.432*** (0.035)	0.449*** (0.035)
log(Odds of Hukou)	-7.993*** (1.554)	-4.739*** (1.300)	-4.686*** (1.364)	-7.476*** (1.296)	-6.989** (2.669)	-6.564** (2.643)
log(Railroad distance in kilometers)	-0.885*** (0.127)	-1.143*** (0.072)	-1.189*** (0.071)	-0.487*** (0.064)	-0.644*** (0.063)	-0.760*** (0.063)
log(Odds of job in dest.)	10.097** (3.690)	2.911 (3.977)	3.383 (3.840)	-0.761 (4.456)	7.284 (8.152)	8.064 (7.933)
log(Odds of job in origin)	-4.004* (2.128)	-0.410 (1.133)	-0.599 (1.144)	-5.180*** (1.757)	1.645 (1.316)	0.175 (1.212)
log(Dest./orig. income)	-0.385 (0.246)	-0.963*** (0.212)	-0.963*** (0.210)	-0.682*** (0.219)	-0.582** (0.243)	-0.576** (0.226)
Dummy = 1 if flow to adjacent province	0.790*** (0.171)	0.559*** (0.157)	0.514*** (0.165)	0.124 (0.097)	0.095 (0.097)	-0.046 (0.104)
Destination province fixed effect	No	Yes	Yes	No	Yes	Yes
Observations	2,383	2,383	2,383	1,535	1,535	1,535
R-square	0.508	0.654	0.645	0.711	0.776	0.709

Note: The dependent variable is the log of gross interprovincial migration rate. In all specifications, period fixed effects are accounted for, and log ratio of real FDI per capita in the destination province to real FDI per capita in the origin province, log ratio of real domestic fixed asset investment per capita in the destination province to real domestic fixed asset investment in the origin province, log percentages of population enrolled in the origin province's universities, log percentages of population enrolled in the destination province's university, log ratio of the share of manufacturing employment in the destination province to the share of manufacturing employment in the origin province, log ratio of the urban share of the destination province's population to the urban share of the origin province's population, log ratio of the destination province's minority population share to the origin province's minority population share, log ratio of mean annual temperature in the capital city of the destination province to the mean annual temperature in the capital city of the origin province, are also included as control variables. The 2SLS specifications used in (3) and (6) used the number of doctors per 10,000 persons in the origin province and the share of minority (non-Han) population in the origin province as instruments. Standard errors are in parentheses, clustered by destination provinces. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

**Table 4. Effects of Log Youth Share on Interprovincial Migration Rate**

	Full Sample 1985-90, 1995-2000 and 2000-05			Later 2 Periods 1995-2000 and 2000-05		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	2SLS	OLS	OLS	2SLS
log(youth share at the start of period))	-0.582* (0.334)	0.196 (0.180)	-0.356 (0.418)	-1.289*** (0.220)	-0.357** (0.172)	-4.259*** (0.672)
log(youth share at the end of period	1.162*** (0.296)	1.303*** (0.259)	3.073** (1.458)	3.567*** (0.320)	3.056*** (0.276)	8.409*** (0.961)
Log(# of migration flows from origin to destination province 10 years earlier)	---	---	---	0.512*** (0.033)	0.449*** (0.037)	0.521*** (0.033)
log(Odds of Hukou)	-8.142*** (1.604)	-4.656*** (1.186)	-4.617*** (1.164)	-6.469*** (1.202)	-6.969*** (2.445)	-6.741*** (2.393)
log(Railroad distance in kilometers)	-0.930*** (0.122)	-1.110*** (0.074)	-1.135*** (0.072)	-0.469*** (0.056)	-0.576*** (0.056)	-0.612*** (0.049)
log(Odds of job in dest.)	10.178** (3.731)	2.831 (3.939)	3.313 (3.748)	1.654 (3.801)	7.163 (7.606)	6.713 (7.362)
log(Odds of job in origin)	-1.673 (2.595)	-0.094 (1.379)	2.487 (1.954)	-6.054*** (1.706)	-2.452* (1.328)	-4.330*** (1.319)
log(Dest./orig. income)	-0.342 (0.255)	-0.789*** (0.190)	-0.610*** (0.216)	-0.457** (0.197)	-0.252 (0.211)	-0.066 (0.208)
Dummy = 1 if flow to adjacent province	0.746*** (0.168)	0.593*** (0.158)	0.567*** (0.164)	0.130 (0.086)	0.144 (0.097)	-0.030 (0.076)
Destination province fixed effect	No	Yes	Yes	No	Yes	Yes
Observations	2,383	2,383	2,383	1,535	1,535	1,535
R-square	0.508	0.654	0.642	0.729	0.771	0.707

Note: The dependent variable is the log of gross interprovincial migration rate. In all specifications, period fixed effects are considered, and log ratio of real FDI per capita in the destination province to real FDI per capita in the origin province, log ratio of real domestic fixed asset investment per capita in the destination province to real domestic fixed asset investment in the original province, log percentages of population enrolled in the origin province's universities, log percentages of population enrolled in the destination province's university, log ratio of the share of manufacturing employment in the destination province to the share of manufacturing employment in the origin province, log ratio of the urban share of the destination province's population to the urban share of the origin province's population, log ratio of the destination province's minority population share to the origin province's minority population share, log ratio of mean annual temperature in the capital city of the destination province to the mean annual temperature in the capital city of the origin province, are also included as control variables. The 2SLS specifications used in (3) and (6) used the number of doctors per 10,000 persons in the origin province and the share of minority (non-Han) population in the origin province as instruments. Standard errors are in parentheses, clustered by destination provinces. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

**Table 5. Effects of Log Age Dependency Ratio (ADR) on the Interprovincial Migration Rate**

	Full Sample 1985-90, 1995-2000 and 2000-05			Later 2 Periods 1995-2000 and 2000-05		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	2SLS	OLS	OLS	2SLS
log(ADR at the start of period)	0.147 (0.179)	0.595*** (0.156)	-2.354*** (0.642)	1.845*** (0.244)	1.484*** (0.242)	-3.639*** (0.870)
log(ADR at the end of period)	-1.719*** (0.280)	-0.588*** (0.173)	4.674*** (1.308)	-2.017*** (0.235)	-0.265 (0.189)	7.842*** (1.273)
Log(# of migration flows from origin to destination province 10 years earlier)	---	---	---	0.466*** (0.030)	0.405*** (0.034)	0.379*** (0.031)
log(Odds of Hukou)	-7.928*** (1.437)	-4.658*** (1.181)	-4.968*** (1.421)	-7.171*** (1.167)	-6.959** (2.560)	-6.528** (2.545)
log(Railroad distance in kilometers)	-0.848*** (0.121)	-1.073*** (0.076)	-1.168*** (0.069)	-0.406*** (0.063)	-0.538*** (0.071)	-0.606*** (0.065)
log(Odds of job in dest.)	11.758*** (3.194)	2.435 (3.995)	2.071 (4.137)	2.450 (3.582)	7.311 (7.913)	7.989 (8.203)
log(Odds of job in origin)	-0.798 (2.295)	0.873 (1.180)	-6.734*** (2.331)	-1.569 (1.690)	3.717*** (1.297)	6.357*** (1.356)
log(Dest./orig. income)	-0.296 (0.240)	-0.896*** (0.204)	-1.272*** (0.233)	-0.650*** (0.214)	-0.598** (0.247)	-0.537** (0.230)
Dummy = 1 if flow to adjacent province	0.817*** (0.162)	0.625*** (0.161)	0.553*** (0.168)	0.253** (0.093)	0.257** (0.109)	0.282** (0.135)
Destination province fixed effect	No	Yes	Yes	No	Yes	Yes
Observations	2,383	2,383	2,383	1,535	1,535	1,535
R-square	0.516	0.647	0.581	0.700	0.746	0.653

Note: The dependent variable is the log of gross interprovincial migration rate. In all specifications, period fixed effects are considered, and log ratio of real FDI per capita in the destination province to real FDI per capita in the origin province, log ratio of real domestic fixed asset investment per capita in the destination province to real domestic fixed asset investment in the original province, log percentages of population enrolled in the origin province's universities, log percentages of population enrolled in the destination province's university, log ratio of the share of manufacturing employment in the destination province to the share of manufacturing employment in the origin province, log ratio of the urban share of the destination province's population to the urban share of the origin province's population, log ratio of the destination province's minority population share to the origin province's minority population share, log ratio of mean annual temperature in the capital city of the destination province to the mean annual temperature in the capital city of the origin province, are also included as control variables. The 2SLS specifications used in (3) and (6) used the number of doctors per 10,000 persons in the origin province and the share of minority (non-Han) population in the origin province as instruments. Standard errors are in parentheses, clustered by destination provinces. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.