Human Development
Research Paper
2009/37

Brain Drain, Brain Gain and Economic Growth in China

Wei Ha, Junjian Yi,
and Junsen Zhang

United Nations Development Programme
Human Development Reports
Research Paper

August 2009

Human Development
Research Paper
2009/37

Brain Drain, Brain Gain, and Economic Growth in China

Wei Ha, Junjian Yi,
and Junsen Zhang

United Nations Development Programme
Human Development Reports
Research Paper 2009/37
August 2009

Brain Drain, Brain Gain, and Economic Growth in China

Wei Ha, Junjian Yi,
and Junsen Zhang

Wei Ha is Policy Specialist at the Human Development Report Office, United Nations Development Programme. E-mail address: wha@unicef.org.

Junjian Yi is Doctoral student at the Economics Department of the Chinese University of Hong Kong. E-mail address: jack.cuhk@gmail.com.

Junsen Zhang is Professor of Economics at the Economics Department of the Chinese University of Hong Kong and President of the Hong Kong Economics Association. E-mail address: jszhang@cuhk.edu.hk.

Comments should be addressed by email to the author(s).

Abstract

This paper examines the effects of both permanent and temporary emigration on human capital formation and economic growth of the source regions. To achieve this end, this paper explores the Chinese provincial panel data from 1980 to 2005. First, the fixed effects model is employed to estimate the effect of emigration on school enrollment rates in the source regions. Relative to this aspect, we find that the magnitude (scale) of permanent emigrants (measured by the permanent emigration ratio) is conducive to the improvement of both middle and high schools enrollments. In contrast, the magnitude of temporary emigrants has a significantly positive effect on middle school enrollment but does not have a significant effect on high school enrollment. More interestingly, different educational attainments of temporary emigrants have different effects on school enrollment. Specifically, the share of temporary emigrants with high school education positively affects middle school enrollment, while the share of temporary emigrants with middle school education negatively affects high school enrollment. Second, the instrumental variable method is applied to estimate the effect of emigration on economic growth within the framework of system Generalized Method of Moments (GMM). The estimation results suggest that both permanent and temporary emigrations have a detrimental effect on the economic growth of the source regions. Our empirical tests provide some new evidence to the "brain drain" debate, which has recently received increasing attention.

Keywords: Brain drain, human capital, emigration, economic growth.
JEL Classification: J22, J24, O12, O15

1 Introduction

Economists have conventionally argued that international migration may deter the development of the source countries (Beine *et al.*, 2001).[1] It is usually concluded that South-to-North migration contributes to the deterioration of the world's income distribution because most of the emigrants are skilled laborers of the source countries. Thus, the emigration of skilled labor is traditionally referred to as "brain drain." However, this argument has been challenged. For example, as was surveyed by Docuier and Rapoport (2004), migration may also have some positive feedback channels, such as remittances, schooling incentives, and return migration after obtaining additional skills, which will certainly contribute to the economic development of the source countries.

One of the possible benefits of migration on source regions pointed out by existing literature is given to the view of schooling incentives, which has been emphasized by a series of recent studies (see e.g., Beine *et al.*, 2008; Di Maria & Stryszowski, 2008). It is called "brain gain," which suggests that the emigration of skilled laborers may provide incentives for those left behind to invest in human capital, and human capital is one of the key determinants in the long-term economic growth according to the endogenous growth theory (e.g. Romer, 1986; Lucas, 1988). When emigration is temporary or the decision on schooling investment is made according to future migration opportunities, this kind of "brain gain" is much more likely to occur in the sending countries (Mayr & Peri, 2008).

Recently, the "brain drain" and "brain gain" issues have attracted increasing attention. In the 1970s, the literature on brain drain and brain gain concluded that skilled workers' emigration adversely affect the welfare of those who remained in the source countries (see e.g., Bhagwati & Hamada, 1974). A series of recent empirical studies, however, provides evidence revealing that emigration may contribute to the long-term development of the sending countries because emigration encourages human capital investment in the sending countries (e.g., Stark *et al.*, 1997; Vidal, 1998; Beine *et al.*, 2001; Beine *et al.*, 2008). For example, Beine *et al.* (2001) found a significantly positive effect of migration on human capital accumulation in their cross-country analysis of 37 developing countries. A recent study (Clemens, 2007) also found that the

[1] See Docuier and Rapoport (2004) for a survey on the brain drain effect.

emigration of a specific group of highly skilled workers—physicians and nurses—causes a greater production of health workers in Africa.

Despite the progress in obtaining evidence for "brain drain" and "brain gain," however, related research has so far been relatively scarce, and the debate remains. To illustrate this scarcity, rare studies have tried to distinguish the effect of permanent emigration on human capital formation from that of temporary emigration. Second, the varying educational composition of emigrants can have varying effects on the human capital formation of the source countries. As Beine *et al.* (2008) pointed out, the recent debate on "brain gain" may be partly due to the absence of reliable cross-country data on international migration by educational levels.[2] Finally, even if the "brain gain" effect exists, it does not automatically imply the positive effect of emigration on economic growth. For instance, if human capital formation is emigration-oriented, the positive effect of human capital investment is not fully earned by the source regions; therefore, it may be logical to say that "brain gain" does not necessarily help the source regions economically.

This paper examines the effects of both permanent and temporary emigration on the human capital formation and economic growth of the source regions. To achieve this end, the paper explores the Chinese provincial panel data, the benefits of the use of which extend to several dimensions. First, using data from one country can help avoid the inconsistency problems with statistic calibers, which prevail in cross-country regressions. Currently, China has 31 regions, and each region uniformly complies with the same statistics caliber.[3] Thus, data from one country are more comparable than cross-country data. In addition, China's regions are large enough to realize the purpose of this study because each region's average population is over 33 million, which is larger than that of most countries in the world.

Second, it is easy to distinguish permanent emigration from temporary migration. Since the 1950s, the Chinese government has been strictly implementing a registration system in which

[2] Only very recently is such kind of migration data available (see Docquier & Marfouk, 2005).

[3] In the Chinese context, we define regions as units at provincial level which includes 23 provinces, 4 municipalities directly under the Central Government, and 5 autonomous regions, respectively. The terms "provincial" and "regional" are used interchangeably in this paper.

each citizen is issued a *hukou* that designates his/her residential place.[4] It is very difficult to change the residential place designated by the *hukou* for the Chinese citizens, especially for those in the rural areas, even today when the economic reform policy has been in effect for 30 years now. Thus, migration with the change of *hukou* is regarded as permanent migration, and migration without the change of *hukou* is regarded as temporary migration. Third, we can compute the educational composition of migration by using the Chinese census data. In doing so, we can compare the different effects of emigrants with different educational levels on the human capital formation and economic growth in the source regions. Fourth, the panel structure of the Chinese data set helps us to remove the cross-province heterogeneities.

Lastly and most importantly, China offers a source of interesting migration experiences for assessment. As is well known, the *hukou* system (household registration system) strictly limited the mobility of the population before it was reformed in 1979. At that time, the Chinese labor market was seriously segmented between rural and urban areas. However, with the loosening of these restrictions, China has been experiencing a large-scale (probably the largest domestic labor migration in human history) migration from rural to urban areas and from inland to coastal areas over the past 30 years. For example, Sicular and Zhao (2002) showed that the magnitude of rural-to-urban migration more than doubled from 8.9 million in 1989 to 23.0 million in 1994. In addition, Cai (2003) estimated that there are about 77.0 million migrants based on the 2000 population census, accounting for more than 11% of the total labor force that year. According to the Intercensal Population Survey 2005, National Bureau of Statistics, the number of gross migrants reached about 150 million, of which 47 million were inter-provincial migrants (NBS, 2006).

As part of the methodology, we estimated both the human capital formation equation and growth equation. For the human capital formation equation, we used the within-group fixed effects model to eliminate cross-province heterogeneities that could bias the OLS estimates. The fixed effects estimation results suggest that permanent emigration has a positive effect on the human capital formation in the source regions. In contrast, we found that the magnitude of temporary emigrants only has a significantly positive effect on middle school enrollment but does not have a significant effect on high school enrollment. More interestingly, the different

[4] Section 2 gives a detailed description of the institutional background of migration in China.

educational attainments of the temporary emigrants have different effects on school enrollments. Specifically, the share of the temporary emigrants with high school education positively affects middle school enrollment, while that of the temporary emigrants with middle school education negatively affects high school enrollment.

For the economic growth equation, we used an instrument variable method to deal with the possible endogeniety of migration within the framework of the system GMM (Generalized Method of Moments) estimator. Conditional on the initial level of the per capital GDP, our system GMM estimates suggest that both permanent and temporary emigration have a negative effect on the growth rate of per capital GDP in the source regions. This result is strong in various specifications and after adding a series of control variables.

de Brauw and Giles (2008) also examined the effect of migration on high school enrollment in rural China by using the timing of ID card distribution to instrument migration. The researchers found a negative relationship between migration and high school enrollment. Our paper differs from theirs in three respects. First and most importantly, we defined emigration rate and enrollment rate by age groups, which is a major improvement in the literature. If the emigration is such defined that it includes young people of school age, emigration implies a mechanically negative effect on school enrollment, which does not have many economic behavioral implications. Second, we examined the educational composition of the emigrants based on the level of school attainments in the source regions. Finally, we examined not only the effect of human capital formation but also the effect of economic growth in the source regions.

The remainder of this paper is organized as follows. Section 2 describes the institutional background of migration in China. Section 3 specifies the empirical strategy. Section 4 introduces the data sets and related variables used in this paper. Section 5 presents the estimated results. Section 6 provides the conclusion.

2 Background

2.1 The *Hukou* System and Migration in China

To achieve rapid industrialization, China adopted the Soviet model of developing the heavy industries as priority since its establishment in 1949. To implement the heavy-industry-oriented strategy, the government not only had to pool many resources (e.g., suppressing the prices of rural products), but it also needed to limit the population migration from rural to urban areas. To intensify the policy, a series of unreasonable regulations was released by the Chinese government to maintain this kind of distorted system artificially. Among these regulations, the *Regulations on Household Registration of the People's Republic of China* issued in 1958 was the most important; its target was to restrict the mobility of the population in the name of the law. [5]

According to the regulation, there were two different types of *hukou*: agricultural and non-agricultural. Every citizen had to be issued a *hukou* status, which designates his/her legal place of residence. Usually, the non-agricultural *hukou* was held by the urban residents, while the peasants in rural areas were the agricultural *hukou* holders. Given the limited resources at that stage, the Chinese welfare system only covered the urban residents, that is, only those holding a non-agricultural *hukou* were allowed to access the various social facilities and social welfare services such as education, medical care, and old-age pensions during the pre-reform period. To protect the employment opportunities and the social welfare of urban residents, an easy solution was to bind these social welfares to one's *hukou* status. Thus, the *hukou* system became the base for achieving this target.

As for the effects of the *hukou* system on China's labor market, related literature documented that the dual segmented labor market was largely caused by the *hukou* system (see e.g., Knight and Song, 1999, 2005; Meng and Zhang, 2001). Under the strict *hukou* system, the rural residents were confined to work locally for life regardless of the low farm productivity induced by the surplus rural labors. Completing a college education was probably the only legitimate way to obtain a non-agricultural *hukou* for the agricultural *hukou* holders at that time. Therefore, although rural laborers had a strong desire for better paying jobs in urban areas,

[5] It should be noted that the *hukou* (household registration) system was indeed originally established in 1951; however, there was no strict limitation on migration at that time.

rural-to-urban migration was unheard of before the economic reform. Aside from the urban-rural segmentation, the *hukou* system also restricted the migration of urban residents, although they had relatively better treatment in social welfare in the *hukou* system. For example, the employees in urban areas indeed had no right to choose their jobs and to move freely between employment *units* or between regions.

Apparently, the urban-rural segmented system was economically inefficient. According to the calculation of Hu *et al.* (2002), the gross loss induced by the labor market segmentation amounted to 20%-60% of the GDP from 1960 to 1978. To improve economic efficiency, the Chinese government reformed the old economic systems in 1978. When the *Household Responsibility System* was introduced in rural areas, agricultural productivity greatly improved, resulting in surplus rural labor. To transfer this surplus labor to increase the productivity and income of rural labor, the government began to loosen the policies of the institutions. In 1983, the government, for the first time, permitted skilled workers and craftsmen who held the agricultural *hukou* to engage in non-farm activities. In addition, in 1985, *Document No.1* of the Central Committee of the Party (CCP) began to allow farmers to search for jobs and establish businesses in nearby towns. However, because of the government's rationed and subsidized food and other necessities (which were only available for non-agricultural holders living in urban areas) in the early stage of the economic reform, the rural labor transfers were limited within the local rural areas, and migration to urban areas or between provinces were not yet common phenomena (Cai, 2000). This situation was referred to as "leaving the land but not the hometown" (*li tu bu li xiang*).

The significant changes occurred in the late 1980s after the government initiated a major reduction in the control of rural-to-urban migration; moreover, the farmers were then permitted to work and to manage businesses in cities as long as they could provide for themselves. The term "*rural migrant wave*" first emerged in 1989, hence reflecting the moderation of the policy relaxation. In effect, the unprecedented magnitude of rural-to-urban migrants was estimated to have reached 8.9 million in that year (Sicular and Zhao, 2002). Furthermore, the end of the food rations in 1992 greatly reduced the migration cost for agricultural *hukou* holders; it also facilitated rural-to-urban migration. The volume of rural-to-urban migrants more than doubled to 23 million in 1994. Based on the 2000 population census, Cai (2003) reported that there were

about 77 million rural migrants in the urban areas in that survey year.

Although after a long-time reform the migration barriers have been eliminated to a certain extent, the *hukou* system still has certain impacts on the welfare availability for the migrants, especially those from rural areas. This situation has led to the classification of migrants into two: permanent and temporary migrants. The first refers to the migrant workers in destination regions who have legally obtained a *hukou* of the destination region. The second refers to the migrants who work in destination regions but have not obtained a *hukou* of the destination regions.[6] Generally, the number of temporary migrants is far bigger than that of the permanent migrants. In the 2000 census, for example, 74.4% of inter-country migrants were temporary migrants (Fan, 2008, p.72).

2.2 Chinese Educational System

The educational system in China is centralized and is mainly classified into four categories: primary school (six years), middle school (three years), high school (three years), and university (four years). There are several exceptions. For example, after finishing the middle school, many students enter technical schools rather than high school. High school is a kind of general education in China. Almost all high school graduates take the university entrance examination. Students who pass the examination enter universities, while others directly enter the labor market. In contrast, technical schools develop students' professional skills, and almost all technical school graduates directly enter the labor market.

In 1986, China issued the Law of Compulsory Education, which prescribes that education is compulsory for both primary and middle schools. However, some areas, especially rural areas, took considerable time to meet this stand in the late 1980s and 1990s. Even in the late 1990s, the law remained unenforced in some remote rural areas because the tuition fee was expensive, hence unaffordable, especially for poor families. Therefore, educational development across the country was unbalanced. In addition, the distribution of higher education institution graduates (university level) was highly unbalanced. In 2008, Fan reported that a majority of university

[6] In China, “temporary migrants” are often regarded as “floating population.” However, there is a slight difference between “temporary migrants” and “floating population”: the former is a flow measure, while the latter is a stock measure of migrants (over a five-year period).

graduates in China were employed in the coastal regions; this phenomenon was referred to "brain drain." In 2005, Zhang et al. found that the number of students who went back to school substantially increased during the period of economic reform. This could be attributed to the rapid rise in income inequalities in China.

3 Empirical Strategy

Following the cross-country analysis of emigration (Beine *et al.*, 2001), two equations need to be estimated. One is human capital formation equation, which examines if emigration induces more investment in human capital and identifies the effect of emigration on school enrollment. The other is economic growth equation that examines the effect of emigration on economic growth. Both regression equations are specified as follows.

$$\ln(\textit{school enrollment rate}_{it}) = \alpha_0 + \alpha_1 \cdot \textit{emigration rate}_{it} + X_{it}^1 \alpha_2 + \eta_i + \tau_t + u_{it} \qquad (1)$$

$$\textit{growth rate}_{it} = \beta_0 + \beta_1 \cdot \textit{emigration rate}_{it} + \beta_2 \cdot \ln(\textit{per capita GDP}_{i,t-1}) + X_{it}^2 \beta_3 + \eta_i + \tau_t + \varepsilon_{it}, \quad (2)$$

where the growth rate refers to the growth rate of (real) per capita GDP; X_{it}^1 and X_{it}^2 are the two vectors of the other control variables; u and ε are the error terms; and α and β are the (vector) coefficients to be estimated. The subscripts *i* and *t* are the index province and year, respectively.

The vector of X_{it}^1 includes the educational input (measured as the ratio of educational expenditure to the local GDP), the share of agricultural population to the total population, and income inequality (measured as the ratio of urban income to the rural income). These variables may not only influence enrollment rate but may also be correlated with the emigration rate; hence, it is important that they be controlled. The vector of X_{it}^2 includes physical investment (measured as the ratio of capital formation to the local GDP), human capital stock (both shares of secondary school graduates and high school graduates to the total population), degree of openness (measured as the ratio of international trade to the local GDP), and government size (measured as the ratio of government expenditure to the local GDP). Furthermore, in both

equations, province dummies η_i are used to sweep out cross-province heterogeneities; year dummies τ_t are also used to control for time period effects, which are common across provinces.

The provincial fixed effects η_i may correlate both the dependent variable and the right-hand variables, and they are usually unobservable to researchers. Thus, omitting η_i could bias the OLS estimates. To deal with this problem, we employ the fixed effects model to eliminate the cross-province heterogeneities in Equation (1) above. For the dynamic regression of Equation (2), if the real per capita GDP is first-order serial correlated, the fixed effects estimates are biased because $\text{cov}(growth\ rate_{i,t-1}, \varepsilon_{it}) \neq 0$ in this case. Therefore, the GMM estimator is employed to deal with the possible endogeniety of the lagged dependent variable. First, the idea behind the GMM estimator is to vary Equation (2) to eliminate the fixed effects. Second, an instrumental variable estimation method is applied to the difference equation.[7]

According to Bond *et al.* (2001), there are two GMM estimators. One is the earlier first-differenced GMM estimator developed by Arrellano and Bond (1991), in which the available lags of $\log(growth\ rate_{t-1})$, in this case, $\log(growth\ rate_{t-2})$, and $\log(growth\ rate_{t-3})$ if the lags exist, are used to instrument the first-difference of $\log(growth\ rate_{t-1})$ in the first-differenced equation. However, Bond *et al.* (2001) and Bond (2002) argued that the first-differenced GMM estimator is subject to the weak instrument and finite sample biases. To address these issues, another GMM estimator—system GMM developed by Arelleno and Bover (1995) and Blundell and Bond (1998)—is preferred to be used. The system GMM estimator utilizes all the available lags of the initial real per capita GDP as the instruments of the first difference of the lagged real per capita GDP in the differenced equation. This system also utilizes the lagged first difference of the real per capita GDP to instrument the initial real per capita GDP in the level equation to deal with the finite sample biases as Bond *et al.* (2001) suggested. For this reason, we employ the system GMM model to estimate Equation (2) in this paper.

[7] See Bond *et al.* (2001) for details.

Migration may be endogenous in both the human capital formation equation (Equation [1]) and the economic growth equation (Equation (2)). The panel dataset allows us to employ the fixed effects model to eliminate effectively the bias that may arise from the unobservable time-invariant factors in both equations. However, the possibilities that can induce the endogenous issue of migration rates still exist. For example, as Beine *et al.* (2003) noted, there may be a reverse effect in the human capital equation as educated laborers are more likely to migrate. The reverse effect of human capital investment on emigration is not much of a concern in our empirical design because we define emigration and school enrollment on different age groups. Thus, we use the fixed effects model to estimate the human capital formation equation.

In the economic growth equation, we use an instrumental variable method to deal with the possible endogeneity of emigration. Previous studies have suggested several possible candidates. For example, in a cross-country analysis, Beine *et al.* (2003) used population size, population density, racial tensions, and stock of migrants in the OECD (Organization for Economic Co-operation and Development) countries to measure the migration rates when they estimated the human capital formation equation.

In this paper, we use the wage premium of migration to measure the migration rate in our economic growth equation. Specifically, the wage premium of migration is calculated as the ratio of the local wage rate to the average wage rate in the urban areas of the coastal regions. In China, the coastal regions are host to most of the emigrants. The argument of the exogeneity of this instrument is that, on one hand, the wage differential is certainly one of the major reasons that induce emigration. On the other hand, the wage rate in the urban areas of coastal regions does not seem to affect the economic growth of the local (middle or west) regions directly.

4 Data

This paper uses the province-level panel data in China for our empirical analysis. As stated in the introduction, using the Chinese data has several advantages. First, data from one country is more comparable than cross-country data because there may be different statistic calibers in different countries. China has 31 different regions, with each region complying with the same statistic caliber. Second, distinguishing the effect of permanent migration from that of temporary

migration is easy. Third, we can also compute the educational composition of the emigrants with the census data. Fourth, the panel structure helps us to eliminate the province fixed effects. Finally, the unprecedented magnitude of domestic migration and the substantial variation of migration across provinces and time periods give us precious opportunity to identify the effects of emigration in several dimensions.

Specifically, we have two datasets.[8] One is a panel data (DATA8005) with information on the variables from 1980 to 2005. The permanent migration rate in DATA8005 is measured as the proportion of the emigrants with the change of *hukou* to the local population. As stated previously, migration with the change of *hukou* is regarded as permanent migration in the Chinese context. Furthermore, as the change in the designated residential place by the *hukou* is highly restrictive, and college graduation is a major channel of changing the *hukou*, permanent migration can also be regarded as the migration of the highly educated laborers. This measure is directly collected from the *Comprehensive Statistical Materials of Population of People's Republic of China: 1949-1985* compiled by the National Bureau of Statistics, the Ministry of Public Security, and various issues of *China Statistical Yearbooks*.

The school enrollment pertained to in this paper is the usual enrollment ratio of middle school to high school. This measure is calculated according to the formula *school enrollment rate* = $enrolls_m/graduates_{m-1}$, where $enrolls_m$ is the number of students enrolled in the m^{th} education level, and $graduates_{m-1}$ is the number of students graduated from the $(m-1)^{th}$ education level. For example, the high school enrollment rate of province *i* in year *t* is calculated as the ratio of the total enrollment of high schools over the number of middle school graduates in province *i* and year *t*. The raw data on enrollments and graduates are collected from various issues of the *Educational Statistical Yearbook of China*. It should be noted that the middle schools here include regular middle schools and vocational middle schools, and the high schools include regular high schools, vocational high schools, and secondary technical schools.

Other economic and demographic variables used in the DATA8005 are mainly collected from the *Comprehensive Statistical Data and Materials on 50 Years of New China*, various issues of the *China Statistical Yearbooks*, and the *China Population Statistical Yearbooks*. It should be

[8] For a systematic comparison of the two data sets, see Appendix 1.

noted that there was a significant adjustment in the regional real GDP according to the first China economic census in 2004. We used these adjusted real GDPs (at the constant price of 1978) to compute the growth rates.

Following Islam (1995) and others, we separated the period to five 5-year sub-periods to implement the panel data analysis of human capital formation and economic growth. This means that all the right-hand variables of the equations are five-year averaged except for the variable of the initial output level in the economic growth Equation (1). Therefore, this panel contains information on 29 provinces for the period 1980-2005[9] with five-year intervals.

The use of DATA8005 has two advantages. One is that we can track a long time period of migration, school enrollment, and economic growth. In fact, the period from 1980 to 2005 almost exhausted the whole period after the economic reform in China. The other advantage is that we can clearly define permanent migration by exploiting the change of *hukou*. However, we cannot check the effect of the educational composition of emigrants on human capital formation and economic growth. We also cannot examine the effect of temporary migration given the data set.

The other dataset (DATA9000) used in this paper complements DATA8005 in both respects mentioned above. Migration is directly computed from the 1990 and 2000 Chinese population censuses in DATA9000.[10] Specifically, the temporary migration rate is measured as the proportion of people aged 20-40 who changed their residences in the past five years over the total population of the same age group in the source regions. One advantage of DATA9000 is that it contains detailed information on the educational composition of emigrants.

Specifically, three educational emigration rates (middle school, high school, and college level) are calculated to estimate the compositional effects of emigration on the interested dependent variables. These three variables are defined as the shares of middle school emigrants, high school emigrants, and college graduates out of the total emigrants. The share of migrants with only primary education or lower is used as the base group, which is dropped from both equations.

[9] Tibet is excluded due to incomplete information, while Chongqing is excluded because it was part of Sichuan before 1997.

[10] Since its establishment in 1949, China has had a total of five population censuses. However, only the 1990 and 2000 censuses contain identifying variables of migration.

Using the census data, we can also assign another measure of human capital formation of province *i* in year *t*. In detail, the school enrollment ratio rate here is measured as the proportion of students at a related educational level to the total population at the normal graduation age. For example, given that the normal graduation age in primary school is 12, the enrollment ratio in middle school is the proportion of students in year one in middle school to the total population aged 12 in province *i* and year *t*. It should be noted that the normal graduation age of primary school can either be 12 and 13, or 12 and 13. Therefore, we have three different measures for the middle school enrollment.

As noted, we deliberately define the migration rate and enrollment rate by different age groups. If emigration rate is defined on the age group including the young people of school age, emigration implies a mechanically negative effect on school enrollment, which does not have many economic behavioral implications. Let us take an extreme case for illustration. Assuming that both migration rate and enrollment rate are defined in the same age group (17-18), the dependent variable of the enrollment rate measures the number of students who remain in school. In contrast, the independent variable of emigration measures the number of students who drop out of school and work outside. Therefore, the relationship between the dependent variable and the independent variable would reflect a similar fact that people drop out of school to emigrate or because of emigration itself, rather than the effect of emigration on human capital investment. In this paper, we are interested in whether the emigration of one group *induces* the human capital investment of another (younger) group.

In DATA9000, we also collected other control variables listed in both vectors X^1 and X^2. In contrast to DATA8005, the other control variables are time point valued to be consistent with the definition of migration and enrollment rates, which is based on the 1990 and 2000 censal years.

The statistical descriptions of the variables of DATA8005 and DATA9000 are shown in Tables 1 and 2, respectively. Data show that the average growth rate of real per capita GDP is about 8% with a standard error of 3.4% over the period of 1980-2005. Comparing the two datasets, one can easily find that on average, the temporary emigration rate (about 34%) more than doubled the permanent emigration rate (about 16%), which is consistent with the usual

casual observations. However, school enrollment rates in DATA9000 are much lower than those in DATA8005. For example, whichever of the normal graduation age is adopted, the middle school enrollment in DATA9000 is always lower than that in DATA8005. As for the composition of emigration rates, Table 2 shows that most of the migrants are less educated. This is consistent with the casual observation that many migrants are educated only in middle school or below.

5 Results

In this section, we analyze the effects of emigration on both human capital formation and economic growth of the source regions by exploring the available panel data of China's provinces. Initially, we present the estimation results of Equations (2)-(3) specified in Section 3 using permanent migration data (DATA8005), and then turn to use the temporary migration data (DATA9000) for both robustness checks and further investigation.

5.1 The Effect of Permanent Emigration

In this paper, we employ the within-group fixed effects estimator to regress the human capital equation. To deal with the potential endogeneity of the lagged per capita GDP ($\ln(per\ capita\ GDP_{i,t-1})$), the growth equation estimations are carried out using the system GMM framework developed by Arellano and Bover (1995).

5.1.1 The Effect of Permanent Emigration on School Enrollment

As mentioned earlier, we use two different measures of school enrollment rates, middle school enrollment rate and high school enrollment rate, as the substitutes for human capital formation. Accordingly, we first report the estimates of the effects of emigration on middle school enrollment using the permanent migration data (DATA8005).

Columns (1)-(4) in Table 3 show the fixed effects estimates using the whole sample including all regions in China. Column (1) only includes the involved variables of emigration and time dummies (coefficients are not reported). It also shows that the emigration rate has a positive effect on middle school enrollment and that this effect is statistically significant.

To check for robustness, we add some controls that may influence the enrollment rate and emigration rate step by step. These controls include public education expenditure ratio, urban-rural income inequality, and the proportion of agricultural population. The first two variables are used to control the possible credit constraints for schooling. The last variable is controlled for because there exists a big gap in the educational attainment between rural and urban areas.

Results show that the coefficient of emigration rate is still statistically significant at the 10% level after controlling for the public education expenditure ratio in Column (2), although the estimated magnitude is quite lower than that in Column (1). The next two columns show that the emigration rates still have positive effects on middle school enrollment; however, the coefficients become statistically insignificant after controlling for more variables reflecting income inequality and rural population.

In Columns (5)-(8) of Table 3, the coastal regions are not included in our analysis because these provinces are usually migrant-receiving regions, which, to some extent, are similar to the developed countries based on the cross-country analysis.[11] Despite this fact, we find that emigration rates cause the middle school enrollment to increase, with a coefficient of around 0.01. The coefficients shown in Columns (7) and (8), after controlling for the inequality and share of agricultural population, are marginally significant at the 10% level.

Table 4 presents the estimation of the effects of permanent emigration on high school enrollment in the source regions. Columns (1)-(4) show the full sample estimates. In Columns (5)-(8), we restrict the analysis to the sub-dataset without the coastal regions. Column (1) includes only the emigration rate and period dummies. One may see that the emigration rate has a statistically significant coefficient, with a magnitude of 0.017. This means that on average, one permillage increment in the emigration rate increase the high school enrollment rate by 1.7%. In Table 4, for instance, we add the public education expenditure ratio to the regression model, giving a statistically significant coefficient of emigration with a smaller magnitude of 0.016. The coefficient of public education expenditure ratio also has an expected sign, although it is

[11] The coastal regions are Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, and Hainan. The 2000 census shows that these coastal regions are net immigration regions. In other words, the number of immigrants is larger than the number of emigrants in these regions.

statistically insignificant. Column (3) includes the income inequality, and it shows that the emigration rate still has significantly positive effects on high school enrollment at the 10% level; in addition, the coefficients of the two controls have expected signs. Column (4) controls all the three variables. The result shows that the coefficient of the emigration rate remains positive, but both its magnitude and significance level decline slightly.

Columns (5)-(8) repeat the exercises as Columns (1)-(4) by using a sub-sample dataset, which excludes the 11 coastal regions. Column (5) shows that the emigration rate has a statistically significant positive coefficient with a magnitude of 0.018. When adding the public education expenditure ratio, the emigration rate remains statistically significant. The negative coefficient of the public education expenditure ratio is somewhat unexpected and inconsistent with the first result. However, the effect of public education expenditure is imprecisely measured when its coefficient is statistically insignificant. We further control the other two variables in a step-by-step procedure as shown in Columns (7) and (8); both regressions show that the effect of emigration on high school enrollment is strong that it can control these variables. In addition, the estimated coefficient is positive and statistically significant at the 10% level, although the coefficients have a slight decline. Both income inequality and the proportion of agriculture population have predictable signs and statistically significant coefficients.

Comparing Table 4 with Table 3, we find that the estimated effect of emigration on high school enrollment is larger than its effect on middle school enrollment. To sum up, the estimated results in both Table 3 and Table 4 show that emigration does have a positive effect on school enrollment even if some of the coefficients are not statistically significant. The results presented here suggest a positive effect of permanent emigration on human capital investment in the source regions.

5.1.2 The Effect of Permanent Emigration on Economic Growth

We now analyze the effects of emigration on growth (see Table for the results). Following the literature we reviewed, we adopt the system GMM estimator in this study to carry out the regressions. As mentioned in Section 4, the wage premium of emigration is used to measure the emigration rates in the economic growth equation.

Columns (1)-(4) in Table 5 report the two-step system GMM estimates. Column (1) includes the emigration rate, investment ratio, initial output level, and period dummies (the coefficients are not reported). It also shows that the three variables have statistically significant coefficients at the 1% level. Emigration rate has a negative impact on the economic growth rate, with a magnitude of -0.004. This means that one permillage increase in the emigration rates will decrease the growth rate of that of the province by 0.4%. The negative coefficient of the initial output level implies that there is a convergence in the output per capita between regions.

The results of the diagnostic tests for the GMM model are reported in the second panel of Table 5. Applying the Arelleno-Bond test for autocorrelation in first difference errors, we find that the test for AR(1) rejects the null hypothesis of no first-order autocorrelation. In contrast, the result of the Arelleno-Bond test for AR(2) shows that the null hypothesis of no second-order autocorrelation cannot be rejected. These results support the requisition of implementing the GMM estimation. Furthermore, we conduct a Sargan overidentifying test to examine the validity of the additional instruments.[12] The p-values of the Sargan test suggest that these instruments are statistically valid.

To check the validity and acceptability of the results, we also add some controls to the basic growth regression in Columns (2)-(4). Column (2) includes the budget expense ratio and the degree of openness. Budget expense ratio is used to represent the government size, which is usually argued to have detrimental impacts on growth (see e.g., Barro, 1991). The degree of openness is expected to have beneficial impacts on the economic growth. Column (3) includes the enrollment rates of middle school and high school, both of which are used to represent human capital. According to the endogenous growth theory (see e.g., Romer, 1986; Lucas, 1988), these two variables are expected to have positive effects on the growth rate. Both regressions show that the coefficients of the emigration rates change slightly yet still have high significance levels.

Column (4) suggests that we have full control over all the variables. Accordingly, the results

[12] The Sargan test is a test of over-identification restrictions. The joint null hypothesis is that the excluded instruments are correctly excluded from the structural growth equation, and the structural equation is correctly specified. Under the null hypothesis, the test statistic is asymptotically distributed as chi-squared with the degree of freedom equal to the number of over-identify restrictions. For further discussion, see for example, Hayashi (2000, pp.227-228, 407, 417).

are still robust and statistically significant at the 1% level. The coefficients of the other variables have predictable signs, although some are insignificant. Furthermore, the Sargan tests suggest that there is no explicit evidence to reject the validity of IVs (instrumental variables) used in GMM regressions. The Arelleno-Bond tests suggest that the requisition of implementing the GMM estimation suffices.

To examine whether the GMM estimates are biased, Bond *et al.* (2001) proposed a suggestive method. They suggested that the OLS estimate of $\ln(\textit{per capita GDP}_{t-1})$ is usually biased upwards, while the within-group estimate of $\ln(\textit{per capita GDP}_{t-1})$ is biased downwards. Therefore, these estimates provide the upper and lower bounds for the robustness check of the system GMM.[13] Following their procedure, we also conduct these two regressions. The results are presented in Columns (5) and (6). In comparing the coefficient in Column (4) with that in Columns (5) and (6), it is clear that the coefficient of $\ln(\textit{per capita GDP}_{t-1})$ in Column (4) stands between the coefficients in Columns (5) and (6). This indicates that the system GMM estimates are reliable.

5.2 The Effect of Temporary Emigration

As stated in the introduction, although the permanent migration in China may probably be more comparable with the international migration, this measure can only provide the gross migration rate. To obtain the composition of the migration rates, the temporary migration data imputed from the census indeed provide us an alternative.

The census data allow us to separate the gross migration into several groups by educational level, so we could investigate the compositional effect of migration on human capital formation and economic growth. Furthermore, the 1990 census recorded the total migration from 1985 to 1990, and the 2000 Census recorded the total migration from 1995 to 2000. According to the background discussed in Section 2, the large-scale migration between regions indeed happened during these two periods; therefore, the census-based data provide us a better understanding of the migration during these periods. As both censuses recorded only the cross-province migration behavior in a five-year period, the migration rate imputed from the census data is regarded as

[13] For further discussion, see Bond *et al.* (2001).

temporary migration.[14]

However, using the census data also has disadvantages. China has had five censuses since 1953, but only two of them (1990 and 2000) have information regarding migration. Given this, we have the migration rates of only two time points. For this reason, it is impossible to run the GMM regressions for the economic growth equation.

5.2.1 The Effect of Temporary Emigration on School Enrollment

Table 6 presents the estimated effects of emigration rates on middle school enrollment using temporary migration data (DATA9000). In this study, we first run the regression for the full sample, and the results of which are given in Columns (1)-(3). Column (1) shows that the emigration rate has a positive effect on middle school enrollment, although this effect is statistically significant only at the marginal level. The composition of emigration by educational level also has positive coefficients on middle school enrollment. In addition, we find that the effect of emigration on middle school enrollment increases with the share of highly educated emigrants. The coefficients on the share of college-educated temporary emigrants out of the total temporary emigrants are consistently significant at the 1% level. These coefficients mean that compared with the base group (emigrants with either primary education or illiteracy), emigration of workers with high educational attainment will help increase middle school enrollment. Furthermore, the higher the education level, the stronger will such effect be. For example, the temporary emigration rate of workers with college education has the stronger promoting effect on middle school enrollment than that of the emigration rate with high school education.

As discussed in Section 3, the normal graduation age is 12, 13, or both. Therefore, in Column (2), we use another middle school enrollment, which assumes that 13 is the normal graduation age. We still find positive coefficients from all these emigration variables. The coefficients on the variable of the share of temporary emigrants with college education and the variable of total temporary emigrants change little in magnitude. In contrast, for the variables of the shares of temporary emigrants with middle school and high school, notable changes in magnitude have been observed. This indicates that the coefficients on the variables of shares of

[14] To comply with the definition of temporary emigration, we have excluded the emigrants who have had a change of *hukou*.

temporary emigrants with middle school and high school are somewhat sensitive to the selection of normal graduation age, which determines the variation of the dependent variable.

Therefore, we repeat the regression by assuming that both 12 and 13 are the normal graduation ages of primary school pupils. We expect that the estimated coefficients on the variables of shares of temporary emigrants with middle school and high school would stand between those in Columns (1) and (2). The result reported in Column (3) supports this prediction Column (3) also shows that the variable of the share of temporary emigrants with college education still has a statistically significant coefficient, and the magnitude of the coefficient has a small increase. At the same time, the gross emigration rate still exhibits a positive effect on the increase of middle school enrollment, as seen in Table 3. The public education expenditure ratio and income inequality are found to have predictable signs. For the income inequality variable, its coefficients are statistically significant in Columns (2) and (3). The proportion of agricultural population has an unpredictable sign, but the coefficient is statistically insignificant.

In Columns (4)-(6) in Table 6, we repeat the regressions by restricting the sample within the western and central regions. The results show some changes. Each of the three regressions that uses three measures of the middle school enrollment, respectively, reveals that the coefficients of emigration rates with middle school education become negative, but these negative coefficients are statistically insignificant. In addition, the coefficients of emigration rates with high school education and the gross emigration rates are statistically significant. In terms of the magnitude of the coefficients, results suggest that the relative importance of the composition of the emigration rates has changed to some extent. For example, compared with the data in Columns (1)-(3), the emigration of workers with college education is relatively less important in promoting the middle school enrollment in the source regions.

In addition, as seen in Columns (4)-(6), we still find predictable signs for the coefficients of both the public education expenditure ratio and income inequality. The coefficient of the proportion of agricultural population remains unpredictable, although it is statistically insignificant.

Table 7 presents the estimation results of the effects of temporary emigration on high school enrollments by using the fixed effects model. We find that the coefficients of emigration rates of emigrants with middle school education become negative; such coefficients are statistically significant at high levels. When we focus on the sample that excludes the coastal regions (Columns [4]–[6]), the coefficients of emigration rates with middle school education are still statistically significant and negative. This means that the share of temporary emigrants with only middle school education will significantly reduce the high school enrollment.

Based on Tables 6-7, we conclude that not only the total volume of emigrants but also their educational composition affects the human capital investment in the source regions.

5.2.2 The Effect of Temporary Emigration on Economic Growth

We now investigate the relationships between temporary emigration by educational level and economic growth using the temporary migration data. We only have observations for two time points; thus, OLS regressions with robust *t*-statistics are adopted in this research. Table 8 shows the regression outcomes.

For the first three columns in Table 8, we run the regressions using the full sample. In Column (1), the growth rate is regressed on the emigration rates based on the initial output level. The emigration rate of emigrants with middle school education is found to be positively correlated with the growth rate; in contrast, the other two emigration rates are always negatively correlated with the growth rates. This means that compared with the base group (emigrants with only primary school education or those who are illiterate), the emigrations of workers with middle school education will foster the growth of the source regions, while the emigrations of workers with high school education or better will be detrimental to the growth of the sending regions. The gross emigration rate here is found to have a predictable negative effect on growth, which is consistent with the findings in Table 5. The positive and statistically significant coefficient of the lagged per capita GDP suggests that there is a divergence between Chinese regions for the period of 1990-2000.

To check the robustness, we then add some controls to the regressions in Columns (2) and (3). The first is the initial investment rate, which usually needs to be included in the growth

regression model according to the growth theory. The second control is the initial average schooling, which is the proxy of human capital. According to the new growth theory, the initial level of human capital is often conducive to growth. Column (2) shows that there is no significant change in terms of the estimated coefficients after including the initial investment ratio. The negative sign of the coefficient of the initial investment ratio is inconsistent with the theoretical expectation, but it is statistically insignificant. The basic results remain even after controlling for the initial level of human capital, as shown in Column (2).

However, the results in Columns (4)-(6) display a dramatic change after we drop the coastal regions from our analysis data. First, the coefficient of the initial output level now becomes negative as the growth theory is asserted, although this negative coefficient is statistically insignificant. This implies that there could be a "club convergence" in the western and central regions in China.[15] Second, all three compositions of the emigration rates become positively and statistically insignificantly correlated with the growth rates. The only one unchanged is the impact of the gross emigration rate on growth rate, which still exhibits a negative effect on growth. The robustness checks in Columns (5) and (6) reveal that the main results in Column (4) remain.

The estimates in Table 8 do not seem to be robust enough. One possible reason is that the sample we used is very small and limited, hence inducing sensitive results. Another alternative is that the temporary migration may not be a suitable measure in the context of growth analysis. This is because this kind of migration rate contains a large number of the so-called "floating population." Indeed, most of the "floating population" often return to their hometowns even if they were listed as the residents of the destination regions during the census period. Therefore, we argue that this kind of migration has an ambiguous effect on the growth of the sending regions.

6 Discussion and Conclusion

The impact of emigration on the well-being of the source regions is still a hot debate in the

[15] "Club convergence" means that the "per capita" incomes of countries identical in their structural characteristics converge with one another in the long run provided that their initial conditions are similar (Galor, 1996). In other words, countries or regions with similar initial conditions cluster into different groups and converge to different steady states.

literature on this topic. This paper argues that the impact of emigration on the well-being of the source regions depends not only on the emigration type (temporary versus permanent) but also on the educational attainment of the emigrants. By exploring the Chinese provincial panel data from 1980 to 2005, we examined the effects of emigration on both school enrollment and economic growth of the source regions. Unlike the previous studies, this study distinguishes the effects of temporary emigration from those of permanent emigration.

To achieve the goals of this study, we first employed the fixed effects model to estimate the effect of emigration on school enrollment rates in the source regions. We found that permanent emigration improves the enrollment in both middle and high schools. In contrast, we found that the magnitude of temporary emigrants only has a significantly positive effect on middle school enrollment but does not have a significant effect on high school enrollment. More interestingly, different educational attainments of temporary emigrants have different effects on school enrollments. Specifically, the share of temporary emigrants with high school education positively affects middle school enrollment, while the share of temporary emigrants with middle school education negatively affects high school enrollment.

Combining the results from both permanent and temporary migration, we found that the educational level of the migrants is positively associated with school enrollments.[16] As discussed in the subsection of the background on China's educational system, high school level is a kind of general education in which not much technical (professional) skills are trained. The almost exclusive objective of anyone studying in high school is to prepare for further education and then to take the university entrance examination, though the admission rate of going to universities is low. Thus, with more people migrating out through university education, more students will be encouraged to study high school and eventually take university entrance tests. In contrast, with more people migrating out with a low level of education, say, middle school level, more students will prefer to enter the labor market directly for employment rather than enter high schools. This is something like a demonstrative effect.

[16] In China, obtaining a university degree remains the only way to achieve permanent migration or to change one's *hukou* for a long period even after the economic reform, especially for those born in rural areas. Thus, although we have no information on the educational composition of the permanent migrants, the majority of permanent migrants have higher educational attainment.

Second, we applied the instrumental variable method to determine the effect of emigration on the economic growth within the framework of system GMM. The results suggest that both permanent and temporary emigrations have a detrimental effect on the economic growth of the source regions.

In summary, the empirical results of this paper imply that when designing or evaluating migration policies, not only the classification into permanent or temporary migrations should be considered but also the educational composition of the migrants. Furthermore, the fact that emigration affects school enrollment and human capital formation suggests that we should study not only the economic aspect but the human development aspects as well.

References

[1] Arellano Manuel & Olympia Bover (1995), "Another Look at the Instrumental Variable Estimation of Error-Component Models", *Journal of Econometrics* 68:29-51.

[2] Beine Michel, Frederic Docquier & Hillel Rapoport (2001), "Brain Drain and Economic Growth: Theory and Evidence", *Journal of Development Economics*, 64: 275-289.

[3] Beine Michel, Frederic Docquier & Hillel Rapoport (2003), "Brain Drain and LDCs' Growth: Winners and Losers", IZA discussion paper, No.819.

[4] Beine Michel, Frederic Docquier & Hillel Rapoport (2008), "Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers", *Economic Journal*, 118: 631-652.

[5] Bhagwati Jagdish & Koichi Hamada (1974), "The Brain Drain, International Integration of Markets for Professionals and Unemployment", *Journal of Development Economics*, 1(1): 19-42.

[6] Blundell Richard & Stephen Bond (1998), "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models", *Journal of Econometrics* 87:115-143.

[7] Bond Stephen, Anke Hoeffler & Jonathan Temple (2001), "GMM Estimation of Empirical Growth Models", CEPR discussion paper No.3048.

[8] Cai Fang (2003), "Removing the Barriers to Labor Mobility: Labor Market Development

and Its Attendant Reforms", Workshop on National Market Integration, The World Bank Beijing Office.

[9] Cai Fang & Dewen Wang. (1999). Sustainability and Labor Contribution of Economic Growth in China. *Journal of Economic Research (in Chinese)*, 10, 62-68.

[10] Clemens Michael (2007), "Health Effects of African Health Professional Emigration", Center for Global Development, working paper no. 114.

[11] de Brauw Alan & John Giles (2008), "Migration Opportunity and the Educational Attainment of Youth in Rural China", Mimeo.

[12] Docuier Frederic & Hillel Rapoport (2004), "Skilled Migration: The Perspective of Developing Countries", Mimeo.

[13] Di Maria Corrado & Piotr Stryszowski (2008), "Migration, Human Capital Accumulation and Economic Development", *Journal of Development Economics*, in press.

[14] Docquier Frederic & Abdeslam Marfouk (2005), "International Migration by Education Attainment: 1990-2000- Release 1.1", World Bank Policy Research Working Paper.

[15] Fan Cindy (2008), "Migration, Hukou, and the City", in Yusuf Shahid & Saich Tony (eds.), *China Urbanizes: Consequences, Strategies, and Policies*, Washington: The World Bank.

[16] Galor Oded (1996), "Convergence? Inferences from Theoretical Models", *Economic Journal*, 106: 1056-1069.

[17] Hayashi Fumio (2000), *Econometrics*. Princeton: Princeton University Press.

[18] Hu Angang, Yonghong Cheng & Yunxin Yang (2002), *Expand Employment and Challenge Unemployment: the Estimate of China's Employment Policy (1949-2001)*, Beijing: China Labor and Social Security Press.

[19] Islam Nazrul (1995), "Growth Empirics: A Panel Data Approach", *Quarterly Journal of Economics* 110:1127-1170.

[20] Karin Mayr & Giovanni Peri (2008), "Return Migration as a Channel of Brain Gain," NBER Working Papers 14039.

[21] Knight John & Lina Song (1999), *The Rural-Urban Divide: Economic Disparities and Interactions in China,* Oxford, UK: Oxford University Press.

[22] Knight John & Lina Song (2005), *Towards a Labour Market in China*, Oxford, UK: Oxford University Press.

[23] Meng Xin (2000), *Labor Market Reform in China*. New York: Cambridge University Press.

[24] Meng Xin & Junsen Zhang (2001), "The Two-tie Labor Market in Urban China", *Journal of Comparative Economics*, 29(3), 485-504.

[25] National Bureau of Statistics (2006), “Report on the Main Statistics of the 1% Population Survey in 2005”, http://www.stats.gov.cn/tjgb/rkpcgb/qgrkpcgb/t20 06 0 316 _402310923.htm

[26] Stark Oded, Christian Haldenstein & Alexia Prskawetz (1997), "A Brain Gain with a Brain Drain", *Economic Letters*, 55:227-234.

[27] Sicular Terry & Yaohui Zhao (2002), "Employment and Earnings in Rural China", Mimeo.

[28] Vidal Jean Pierre (1998), "The Effect of Emigration on Human Capital Formation", *Journal of Population Economics*, 11(4): 589-600.

[29] Zhang Junsen, Zhao Yaohui, Park Albert & Song Xiaoqing (2005), "Economic Returns to Schooling in Urban China", *Journal of Comparative Economics*, 33 (4), 730-752.

Table 1: Summary Statistics of the Variables

Variable	Obs.	Mean	Std. Dev.	Min	Max
Permanent emigration rate (1/1000)	200	16.483	5.496	4.790	35.186
Real per capita GDP growth rate	199	0.080	0.034	-0.008	0.187
ln (real per capita GDP)	228	6.774	1.049	4.817	9.920
Government expenditure/GDP	144	0.135	0.056	0.046	0.328
Education expenditure/GDP	143	0.032	0.012	0.014	0.074
Rural population/total population	116	0.723	0.139	0.339	0.886
Fixed investment/GDP	145	0.329	0.099	0.154	0.723
Middle school enrollment rate	144	0.840	0.131	0.534	1.012
High school enrollment rate	144	0.452	0.127	0.219	0.903
Income inequality (urban real per capita income/rural real per capital income)	116	2.520	0.617	1.310	4.494
International trade/GDP	143	0.197	0.278	0.009	1.624

Notes: (1) Data source: DATA8005 (see Appendix 1 for details); (2) the permanent emigration rate is defined as the proportion of the emigrants with the change of *hukou* to the local population; (3) the school enrollment ratios are the usual enrollment ratio of middle school and high school (see Section 4 for details); (4) all variables are five-year averaged except for ln (real per capita GDP).

Table 2: Summary Statistics of the Variables

Variable	Obs	Mean	Std. D	Min	Max
Temporary emigration rate (1/1000)	61	34.424	32.517	7.829	166.647
Share of temporary emigrants with middle school education	61	39.353	11.574	15.796	64.656
Share of temporary emigrants with high school education	61	18.227	5.967	6.750	34.639
Share of temporary emigrants with college education	61	10.994	10.265	1.466	48.303
Middle school enrollment rate (1): age group 12	61	0.637	0.289	0.000	0.984
Middle school enrollment rate (2): age group 13	61	0.723	0.234	0.000	0.984
Middle school enrollment rate (3): age groups 12 and 13	61	0.699	0.252	0.000	0.984
High school enrollment rate (1): age group 15	61	0.205	0.206	0.000	0.890
High school enrollment rate (2): age group 16	61	0.275	0.196	0.000	0.845
High school enrollment rate (3): age groups 15 and 16	61	0.248	0.200	0.000	0.849
Income inequality	58	2.411	0.608	1.140	4.280
Education expenditure/GDP	57	0.030	0.010	0.006	0.057
Rural population/total population	61	0.730	0.146	0.330	0.932
ln (real per capita GDP)	58	7.265	0.731	6.013	9.439
Fixed investment/GDP	58	0.323	0.104	0.162	0.606
Average schooling years	61	4.897	1.227	1.017	7.818

Notes: (1) Data source: DATA9000 (see Appendix 1 for details); (2) the temporary migration rate is directly computed from the 1990 and 2000 census of population. The temporary migration rate is defined as the share of emigrants with the age of 20-40 of the total population of the same age cohort migrated out the source regions during the past five years; (4) the school enrollment ratios are measured as the proportion of students of related grade of education to total population at the normal graduation age.

Table 3: Fixed Effect Estimates of the Effects of Permanent Emigration on Middle School Enrollment

	Dependent Variable: ln (Middle school enrollment rate)							
	All regions				Coastal regions excluded			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Permanent emigration rate	0.010	0.008	0.005	0.005	0.015	0.013	0.008	0.008
	(2.03)*	(1.76)*	(1.02)	(0.94)	(2.24)**	(1.83)*	(1.12)	(1.02)
ln (education expenditure/GDP)	—	-0.194	-0.144	-0.135	—	0.026	0.050	0.085
		(2.75)**	(1.28)	(1.07)		(0.22)	(0.32)	(0.52)
ln (income inequality)	—	—	-0.222	-0.210	—	—	-0.199	-0.159
			(1.64)	(1.63)			(1.20)	(1.24)
ln (rural population/total population)	—	—	—	-0.102	—	—	—	-0.757
				(0.21)				(0.68)
Constant	-0.508	-1.115	-0.704	-0.712	-0.315	-0.185	-0.182	-0.281
	(5.10)***	(4.51)***	(1.53)	(1.58)	(3.45)**	(0.39)	(0.29)	(0.41)
# of Observations	144	142	115	115	90	88	71	71
# of Regions	29	29	29	29	18	18	18	18
R-squared	0.70	0.72	0.72	0.72	0.77	0.76	0.77	0.77

Notes: (1) Data source: DATA8005 (see Appendix 1 for details); (2) Table 1 presents the summary statistics; (3) robust t statistics in parentheses; (4) * significant at 10%; ** significant at 5%; *** significant at 1%; (5) all regressions include year dummies.

Table 4: Fixed Effect Estimates of the Effects of Permanent Emigration on High School Enrollment

	Dependent Variable: ln (High school enrollment rate)							
	All regions				Coastal regions excluded			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Permanent emigration rate	0.017	0.016	0.012	0.010	0.018	0.014	0.012	0.010
	(2.33)**	(2.10)**	(1.75)*	(1.56)	(2.55)**	(2.10)*	(1.86)*	(1.82)*
ln (education expenditure/GDP)	—	0.002	0.140	0.222	—	-0.274	-0.097	0.012
		(0.02)	(0.86)	(1.28)		(1.54)	(0.39)	(0.05)
ln (income inequality)	—	—	-0.239	-0.129	—	—	-0.361	-0.236
			(1.87)*	(1.03)			(2.78)**	(1.88)*
ln (rural population/total population)	—	—	—	-0.960	—	—	—	-2.380
				(1.67)				(2.14)**
Constant	-1.251	-1.210	-0.473	-0.553	-1.099	-2.037	-1.190	-1.502
	(8.46)***	(3.08)***	(0.77)	(0.95)	(11.06)***	(3.15)***	(1.36)	(1.72)
# of Observations	144	142	115	115	90	88	71	71
# of Regions	29	29	29	29	18	18	18	18
R-squared	0.55	0.54	0.61	0.64	0.55	0.57	0.63	0.68

Notes: (1) Data source: DATA8005 (see Appendix 1 for details); (2) Table 1 presents the summary statistics; (3) robust t statistics in parentheses; (4) * significant at 10%; ** significant at 5%; *** significant at 1%; (5) all regressions include year dummies.

Table 5: GMM Estimates of the Effects of Permanent Emigration on Economic Growth

	Dependent Variable: Real per capita GDP growth rate					
	SYS-GMM (1)	SYS-GMM (2)	SYS-GMM (3)	SYS-GMM (4)	POLS (5)	FE (6)
ln (real per capita GDP_{t-1})	-0.019	-0.038	-0.016	-0.029	-0.025	-0.083
	(5.74)***	(8.93)***	(4.87)***	(5.04)***	(3.74)***	(7.55)***
Permanent emigration rate	-0.004	-0.005	-0.003	-0.003	0.000	0.001
	(11.22)***	(8.00)***	(9.70)***	(13.57)***	(0.74)	(0.76)
ln (fixed investment/GDP)	0.038	0.051	0.017	0.024	0.013	0.021
	(6.17)***	(10.52)***	(4.41)***	(4.19)***	(1.51)	(1.92)*
ln (middle school enrollment rate)	—	—	0.014	0.030	0.016	0.019
			(2.39)**	(3.20)***	(0.84)	(0.87)
ln (high school enrollment rate)	—	—	0.028	0.037	0.016	0.034
			(4.34)***	(5.03)***	(1.45)	(2.54)**
ln (government expenditure/GDP)	—	-0.005	—	-0.008	-0.024	-0.001
		(1.41)		(2.28)**	(3.67)***	(0.12)
ln (international trade/GDP)	—	0.012	—	0.007	0.010	-0.001
		(5.41)***		(3.82)***	(2.87)***	(0.09)
Constant	0.344	0.522	0.312	0.421	0.301	0.781
	(10.08)***	(11.89)***	(9.15)***	(7.50)***	(4.56)***	(9.31)***
AR(1) test	0.002	0.002	0.004	0.005		
AR(2) test	0.553	0.380	0.723	0.735		
Sargan test	0.278	0.409	0.476	0.501		
# of Observations	144	142	144	142	142	142
R-squared					0.54	0.68

Notes: (1) Data source: DATA8005 (see Appendix 1 for details); (2) Table 1 presents the summary statistics; (3) absolute value of *t* or *z* statistics in parentheses; (4) * significant at 10%; ** significant at 5%; *** significant at 1%; (5) all regressions include year dummies; (6) AR(1) test and AR(2) test denote Arelleno-Bond test for zero autocorrelation in first difference errors. Sargan test denotes the test of overidentifying restrictions for additional instruments in level equations. P-values for AR(1), AR(2), and Sargan tests are reported. (7) There are 29 regions in the regressions.

Table 6: Fixed Effects Estimates of Temporary Emigration on Middle School Enrollment

	Dependent variable: Middle school enrollment rates					
	Coastal regions included			Coastal regions excluded		
	MID 1	MID 2	MID 3	MID 1	MID 2	MID 3
	(1)	(2)	(3)	(4)	(5)	(6)
Share of temporary emigrants with	0.377	0.588	0.541	-0.519	-0.171	-0.344
middle school education	(0.84)	(1.31)	(1.11)	(1.14)	(0.38)	(0.73)
Share of temporary emigrants with	0.937	0.241	0.498	2.594	1.871	2.239
high school education	(1.14)	(0.35)	(0.65)	(2.60)**	(2.08)*	(2.23)**
Share of temporary emigrants with	1.211	1.269	1.317	0.575	0.853	0.785
College education	(2.99)***	(3.14)***	(3.25)***	(0.83)	(1.52)	(1.31)
Temporary emigration rate	1.083	1.099	1.172	2.269	2.102	2.299
	(1.61)	(1.74)*	(1.70)	(3.39)**	(3.13)***	(3.18)***
ln (education expenditure/GDP)	0.189	0.175	0.193	0.428	0.516	0.525
	(1.29)	(0.97)	(1.12)	(3.41)**	(4.05)***	(4.05)***
ln (income inequality)	-0.242	-0.314	-0.311	-0.026	-0.106	-0.088
	(1.44)	(1.93)*	(1.76)*	(0.16)	(0.73)	(0.54)
ln (rural population/total population)	-0.009	0.249	0.197	0.246	0.557	0.526
	(0.02)	(0.57)	(0.42)	(0.40)	(1.00)	(0.88)
Constant	0.786	1.096	1.051	1.512	2.092	2.047
	(1.75)*	(1.91)*	(1.88)*	(4.31)**	(5.53)***	(5.43)***
p-values of the joint F test	0.020	0.000	0.001	0.034	0.004	0.012
# of Observations	57	57	57	35	35	35
# of Regions	29	29	29	18	18	18
R-squared	0.95	0.89	0.9	0.97	0.96	0.96

Notes: (1) Data source: DATA9000 (see Appendix 1 for details); (2) Table 2 presents the summary statistics; (3) absolute values of robust t statistics in parentheses; (4) * significant at 10%; ** significant at 5%; *** significant at 1%; (5) F-statistics present the p-values of the joint tests for variables relating to migration (the first four variables); (6) all regressions include year dummies; (7) middle school enrollment rates are calculated as the share of middle school students to the total population of the same age group (age of 12, 13, or 12-13). MID1: age 12 is assumed as the normal graduation age of primary school, MID2: age 13 is assumed as the normal graduation age of primary school; MID3: 12 or 13 is assumed as the normal graduation age of primary school pupils.

Table 7: Fixed Effects Estimates of the Effect of Temporary Emigration on High School Enrollment

	Dependent variable: High school enrollment rates					
	Coastal regions included			Coastal regions excluded		
	HIG 1	HIG 2	HIG 3	HIG 1	HIG 2	HIG 3
	(1)	(2)	(3)	(4)	(5)	(6)
Share of temporary emigrants	-0.935	-0.530	-0.678	-0.775	-0.367	-0.536
middle school education	(2.15)**	(2.36)**	(2.35)**	(4.33)*	(2.51)**	(3.66)**
Share of temporary emigrants	0.638	0.190	0.299	-0.160	-0.497	-0.390
high school education	(0.74)	(0.39)	(0.50)	(0.49)	(1.20)	(1.11)
Share of temporary emigrants	-1.315	-0.850	-1.002	-0.824	-0.635	-0.726
College education	(2.31)**	(3.45)**	(3.08)**	(1.02)	(1.10)	(1.12)
Temporary emigration rate	0.700	0.419	0.542	1.492	0.603	0.952
	(0.91)	(1.09)	(1.09)	(1.24)	(1.60)	(0.86)
ln (education	0.131	-0.050	-0.017	-0.179	-0.129	-0.151
	(0.52)	(0.73)	(0.16)	(2.93)*	(2.21)**	(2.69)**
ln (income inequality)	-0.013	-0.035	-0.038	-0.096	-0.039	-0.053
	(0.11)	(0.60)	(0.52)	(2.06)*	(0.58)	(0.90)
ln (rural population/total	0.649	-0.238	0.019	0.196	-0.190	-0.041
	(1.54)	(1.35)	(0.08)	(1.57)	(1.99)*	(0.40)
Constant	1.066	0.159	0.370	-0.097	-0.093	-0.103
	(1.36)	(0.91)	(1.20)	(0.44)	(0.56)	(0.56)
	0.172	0.024	0.060	0.000	0.000	0.000
p-values of the joint F test	57	57	57	35	35	35
# of Observations	29	29	29	18	18	18
# of Regions	0.86	0.96	0.94	0.98	0.99	0.98
R-squared	0.95	0.89	0.9	0.97	0.96	0.96

Notes: (1) Data source: DATA9000 (see Appendix 1 for details); (2) Table 2 presents the summary statistics; (3) absolute values of robust t statistics in parentheses; (4) * significant at 10%; ** significant at 5%; *** significant at 1%; (5) F-statistics present the p-values of the joint tests for variables relating to migration (the first four variables); (6) all regressions include year dummies; (7) high school enrollment rates are calculated as the share of high school students to the total population of the same age group (age of 15, 16, or 15-16). HIG1: age 15 is assumed as the normal graduation age for middle school students, HIG2: age 16 is assumed as the normal graduation age for middle school students; HIG3: 15 or 16 is assumed as the normal graduation age of middle school students.

Table 8: OLS Estimates of the Effect of Temporary Emigration on Economic Growth

	Dependent Variable: ln (real per capita GDP_t) - ln (real per capita					
	Coastal regions included			Coastal regions excluded		
	(1)	(2)	(3)	(4)	(5)	(6)
Share of temporary with middle school	0.122	0.116	0.113	0.060	0.054	0.039
	(2.49)**	(2.27)**	(1.94)*	(1.39)	(1.22)	(0.63)
Share of temporary with high school education	-0.042	-0.027	-0.029	0.002	0.016	0.012
	(0.70)	(0.40)	(0.41)	(0.03)	(0.27)	(0.20)
Share of temporary with college education	-0.070	-0.073	-0.076	0.058	0.052	0.045
	(1.09)	(1.12)	(1.05)	(1.03)	(0.90)	(0.71)
Temporary emigration rate	-1.393	-1.234	-1.238	-1.070	-0.849	-0.854
	(2.41)**	(1.87)*	(1.83)*	(2.07)*	(1.43)	(1.38)
ln (real per capita GDP_{t-1})	0.026	0.029	0.028	-0.011	-0.008	-0.014
	(2.10)**	(2.13)**	(1.76)*	(0.70)	(0.51)	(0.60)
ln (fixed investment/GDP)	—	-0.009	-0.009	—	-0.013	-0.007
		(0.53)	(0.45)		(0.79)	(0.31)
ln (average schooling	—	—	0.004	—	—	0.012
			(0.11)			(0.35)
Constant	-0.084	-0.072	-0.072	0.154	0.172	0.180
	(1.10)	(0.88)	(0.86)	(1.55)	(1.67)	(1.64)
p-values of the joint F test	0.019	0.043	0.058	0.103	0.310	0.484
# of Observations	29	29	29	18	18	18
R-squared	0.40	0.41	0.41	0.47	0.50	0.51

Notes: (1) Data source: DATA9000 (see Appendix 1 for details); (2) Table 2 presents the summary statistics; (3) absolute values of robust t statistics in parentheses; (4) * significant at 10%; ** significant at 5%; *** significant at 1%; (5) F-statistics present the p-values of the joint tests for variables relating to migration (the first four variables); (6) all regressions include year dummies.

Appendix 1

Comparing the two data sets: DATA8005 vs. DATA9000

DATA8005

Period: 1980-2005

Intervals: 5 years

Source: All variables are collected from *Comprehensive Statistical Materials of Population of Peoples' Republic of China: 1949-1985* and from various issues of the *China Statistical Yearbooks, Educational Statistical Yearbook of China,* and *China Population Statistical Yearbooks.*

Migration type: Permanent migration

Description: This data set contains only the information on migrants whose *hukou* has been changed. As discussed in the Background section, migrants who have had a change of *hukou* are regarded as permanent migrants in the Chinese context.

DATA9000

Period: 1990-2000

Intervals: 10 years

Source: Variables relating to migration and school enrollments are directly generated from the 1990 and 2000 censuses of population (1% sample). Other socioeconomic variables are taken from the statistical book *Comprehensive Statistical Materials of Population of Peoples' Republic of China: 1949-1985* and from the various issues of *China Statistical Yearbooks*.

Migration type: Temporary migration

Description: Temporary migrants are those who have changed their residences during the last five years. The main advantage of DATA9000 is that it allows the classification of the migrants into different education groups. We can explore the different effects of the different educational attainments of temporary emigrants on the human capital formation and economic growth in the source regions.