The Threshold Effect of High-level Human Capital Investment on China's Urban-Rural Income Gap♣

Yitao Jiang[1], Xiaojun Shi[2*], Shunming Zhang[3] and Jingjing Ji[1]

[1]School of Economics, Xiamen University, 361005
[2]School of Economics and Management, Beijing University of Aeronautics and Astronautics, 100191
[3]School of Finance, Renmin University of China, 100872

This version: January 28, 2011

♣ Financial support from National Natural Science Foundation of China (NSFC Grant Numbers: 70825003, 70972001) and National Social Science Foundation of China (SSFC Grant Number: 07AJL002) is gratefully acknowledged.

* The corresponding author. Phone and Fax: (86 10) 8233 9132; E-mail: sxjstein@gmail.com.

The Threshold Effect of High-level Human Capital Investment on China's Urban-Rural Income Gap

ABSTRACT

Purpose - The purpose of this paper is to shed light on the effect of high-level human capital investment, using tertiary education as the proxy, on the urban-rural income gap in China.

Design/methodology/approach – Using a panel dataset covering 28 provinces of China over the period from 1988 to 2007, this paper employs Hansen's (1999) method and two-step GMM-SYS estimator to estimate the threshold regression model and the dynamic fixed-effect panel model, respectively.

Findings –We find that the urban-rural income gap is related to high-level human capital investment in an inverted U-shaped pattern with respect to economic development level. The estimated threshold turning point is around 20,000 RMB GDP per capita. This estimate is sufficiently robust to model specifications and variants of the dependent variable.

Research limitations/implications – The research is limited to data availability. We use proxies for high-level human capital investment and higher labor migrations, which might affect the precision of our estimations.

Practical implications – For regions such as Shanghai, Tianjin, Beijing, Jiangsu, Zhejiang, Guangdong, Fujian, and Shandong, where economic development surpasses the threshold, high-level capital investments are encouraged to generate further equalizing effects.

Social implications –We forecast that high-level human capital investment could play a role in bridging the urban-rural income gap at the national level by 2014, when China's GDP per capita assumes an annual growth rate of 7.5%.

Originality/value – This, it is believed, is the first research to find an inverted U-shaped pattern for high-level human capital investment and urban-rural income gap nexus in China.

Key words: Human capital investment; Urban-rural income gap; Threshold; Inverted U-shape.

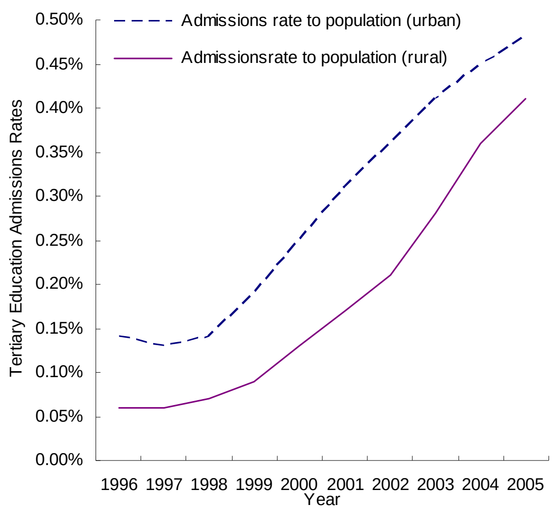
I. Introduction

There is extensive consensus that the urban-rural income gap is the major driver of income inequality in China. To remedy the inequality issue thus necessitates a thorough understanding of factors affecting China's urban-rural income gap. This paper examines the role of high-level human capital investment, particularly, investment in tertiary education, in China's urban-rural income disparity.

Two facts motivate us to explore China's tertiary education investment and urban-rural income gap nexus. First, the last decade witnessed rapid growth in tertiary education investment in both urban and rural China; this raises the question of the economic consequences of this growth in investments, particularly with respect to impacts on the urban-rural income gap. Second, the previous literature regarding human capital and China's rural income overwhelmingly focuses on the role of primary and secondary education, while tertiary education remains less researched. This provides an incomplete picture in the context of rapidly expanding tertiary education in China. We thus need a fuller picture encompassing the impacts of tertiary education.

China's educational transformation, with its emphasis on tertiary education, largely differentiates it from other low wage economies (e.g., India) where primary or secondary education is the central focus (Li, Whalley, Zhang and Zhao, 2010). A major transformation of higher education in China has been underway since 1999. The number of undergraduate and graduate students has been growing at approximately 30% per year since 1999. The number of enrollments for tertiary education has risen quickly, and has approximately ***sextupled*** between 1997 and 2008. A further feature of China's higher education transformation is the considerably improved access to tertiary education for rural households (Li, Whalley, Zhang and Zhao, 2010). Data shows the proportion of urban students in total higher education admissions is decreasing, while the proportion of rural students is increasing (Gou, 2006; Li, Whalley, Zhang and Zhao, 2010). Figure 1 shows both the urban and rural tertiary education admissions rates in China (using respective populations as the denominators) have been growing rapidly, with rural admissions accelerating to make them converge. This notable growth in tertiary education naturally raises the question of its economic consequences. Among these consequences, this paper focuses on its impacts on the urban-rural income disparity.

Figure 1 Tertiary education admissions rates in urban and rural China (1996~2005)



Source: Li, Whalley, Zhang and Zhao (2010).

Previous literature relating rural income in China to education provides a partial picture, with the central focus on primary and secondary education. Basically, this strand of literature lends us insights that education improves lucrative off-farm job opportunities in rural China, which, in turn, could attenuate the enlarging urban-rural income gap to some extent. De Brauw and Rozelle (2008) estimate the return to education in off-farm wage employment in rural China by using hourly wage and controlling for omission variable bias. They find an average return of 6.4% for primary and secondary education. Using a household survey of 2005, Zhang, Zhang, Luo and Li (2009) estimate that this return is 7%, and claim that education in rural China still pays off and even shows an increasing trend in returns over time. These results echo earlier work by Zhao (1997), and Zhang, Huang and Rozelle (2002) who claim that education improves off-farm wage employment opportunities by increasing migration to urban regions, and aiding the formation of a well-functioning rural labor market. Other authors such as Brown and Park (2002), and Zhao and Glewwe (2010) also find an endogenous relationship between rural income and basic school attainment. They confirm that household income has positive impacts on basic education attainment in poor rural villages.

A handful of papers on education and income gap focus on the ***regional*** income gap rather than the urban-rural one. Chi (2008) gives evidence on the indirect role of human capital in China's economic growth through physical capital investment. She finds workers with college education play a more significant role than those with primary or secondary educations. These results imply that *eastern-western* inequality may increase rather than decrease with differentials in high-level human capital endowments. Moreover, Yao and Zhang (2008) quantify that 12 to 47 percent of the *eastern-western* development gap can be attributed to differentials in average education years. A very recent paper by Fleisher, Li and Zhao (2010) gives a careful and illuminating investigation into the human capital and *regional inequality* nexus. They find that while investment in infrastructure generates higher returns in eastern regions than in the interior, investing in human capital generates slightly higher or comparable returns in the interior regions. They claim that human capital investment in less-developed areas is justified on the grounds of efficiency and its contribution to reductions in regional inequality.

However, the literature focusing on the impacts of tertiary education on the ***urban-rural*** income gap in China is largely unavailable, with Guo (2005) and Li, Whalley, Zhang and Zhao (2010) as notable exceptions. Guo (2005) reports that the differences of urban-rural human capital stock and accumulation, together with the birth rate leads to the enlarging of the urban-rural income gap. However, he only presents a descriptive result using one year's data (2003) without quantifying the effects. Li, Whalley, Zhang and Zhao (2010) raise this question in their subsection 3.4 but without further detailed analysis. One may argue that the present paucity of literature on this question is due to its "trivialness". An intuitional answer to this question is that tertiary education would enlarge the urban-rural income gap monotonically, because the vast majority of graduate and post-graduate students from rural areas would migrate to urban areas after their graduation. Our empirical results reject this answer outright, and demonstrate that it is far more complicated than a trivial one.

In this paper, we use a panel dataset at the province level to present empirical results on high-level human capital investment and the urban-rural income disparity nexus. Our results highlight that China's urban-rural income gap is related to high-level human capital investment (using tertiary education as the proxy) in a nonlinear way, displaying an inverted U-shape with respect to the economic development level. This implies that there is a threshold effect in the high-level human capital investment and the urban-rural income gap nexus in China. On one side, where the economic development level is above the threshold, high-level human capital investment can narrow the gap; while on the opposite side, below the threshold, high-level human capital investment widens the gap. Furthermore, our results are robust for different proxies for high-level human capital investment and model specifications, controlling for unobservable bias. To our knowledge, we are the first to present this inverted U-shaped pattern. Moreover, we give reliable estimations of the threshold value, which is useful to the policymakers.

The reminder of this paper is organized as follows. Section 2 presents test hypotheses relating the urban-rural income gap to investments in tertiary education. Section 3 explains variables, the sample, and our empirical strategies. Section 4 presents our empirical results with discussions and robustness checks. Section 5 concludes with some policy implications.

II. Test Hypotheses

In this paper, we use tertiary education investment to express high-level human capital investment in China. We argue that there are multiple channels through which high-level human capital investment could attenuate the urban-rural income gap rather than widening it monotonically.

The first channel is through remittances. Permanent rural to urban migrants with tertiary education (also termed as *Hukou* migrants in Fan, 2008) have two counter-effects on rural income. On one hand, such "brain drain" migration may exaggerate the urban-rural human capital disparity; but on the other hand, these *Hukou* migrants send back remittances, which might play a role in attenuating the urban-rural income gap. This remittance channel follows the earlier classic work of such authors as Lucas and Stark (1985).

The second channel is an indirect one — improvement to the level of primary education attainment through the relief of tight credit constraints. This channel is in the spirit of Rosenzweig and Stark (1989). Remittances not only improve the level of rural income directly, but also smooth consumption and relieve the tight credit constraints for receiving households. Many authors (Brown and Park, 2002; Zhao and Glewwe, 2010) present compelling evidence of the central role of credit constraints in determining primary education attainment in rural China. Furthermore, primary education attainment affects off-farm wage job opportunities which are becoming a major source of rural income improvement.

The third channel is a spillover effect of the high-level human capital in terms of enlarged social networks. *Hukou* migrants could use their *Guanxi* to expand the social networks of their original rural relatives and neighbors. These expanding networks have the capacity to present more lucrative off-farm job opportunities, which provide potentially further sources to improve rural income.

We further argue that there is a fourth channel in China due to its increasingly heated competition in urban job markets, large stock of tertiary educated laborers, and attractive opportunities in the rural market of some more developed regions. We term this a “back-flow effect” of high-level human capital.

In describing the “back-flow” effect, we assume that the migration of higher labor, formed by tertiary education investments, depends on two factors: the urban-rural income gap, and the extent of the competition in the urban labor market. As the income gap grows wider, the driving force of higher labor towards cities increases accordingly, whereas when the labor in market is more competitive, counter-driving forces increase.

In less developed areas with wider urban-rural income gaps, the force driving higher labor towards the city dominates. Higher labor thus rarely moves to rural areas in this case. Here, high-level human capital investment cannot narrow the income gap. It might even widen the gap because the prior investment in high-level human capital becomes a pure cost for the rural area, which might therefore become poorer than before.

However, in more developed areas with economic levels above a critical threshold, the income gap is narrower. This results in a weaker driving force towards urban centers. Moreover, because higher labor accumulates in the urban market of such areas during its earlier development stages, the competition in the labor market is more heated, weakening the driving force towards cities even further. Meanwhile, the rural areas in this case, providing more opportunities together with a decent salary, could attract some higher laborers whose migration results in increased productivity in rural areas. With the premise that marginal output determines wage rate, income in rural areas would increase due to consistent higher labor inflows. This suggests that high-level human capital investment could play the role in narrowing the urban-rural income gap in some more developed regions.

The above analysis of the four possible channels through which tertiary education investment may impact the urban-rural income gap provides a far more complicated picture than the simple, intuitional conjecture that this gap would be widened because of brain drain and costs for rural areas caused by tertiary education investment. In a broad sense, the above analysis lends us the insight that tertiary education investment exerts two counter-forces on the urban-rural gap. The force that dominates largely depends on the context of the region in terms of economic development level or human-capital stocks. Combining the four channels and the counter-force exaggerating the disparity would possibly generate a central "threshold effect" hypothesis, describing the impact of the high-level human-capital investment on the urban-rural income gap in China. Specifically, this central hypothesis is expressed as follows:

H1: It is possible to estimate a threshold effect of high-level human capital investment on the urban-rural income gap in China.

H2: In areas where economic development levels are above the critical threshold, high-level human capital investment could possibly narrow the urban-rural income gap. While below the threshold, it widens the gap.

H3: High-level labor migration also follows an inverted U-shaped pattern. Ideally, the two thresholds in H1 and H3 should coincide.

Ⅲ. Variables, Samples and Empirical Models

3.1. Variables

3.1.1. Urban-rural income gap (*gap*)

In China, the disposable income per capita of urban residents and the net income per capita of rural residents reflect their real levels of incomes, respectively (Cai, 1998). These two incomes are well-documented in China's statistical yearbooks. However, two points on the net income of rural households are worth stressing to prevent some confusion regarding the impacts of migrations. The net income of rural households, documented in the statistical yearbooks, covers the remittances from "peasant labor" migrants and the permanent *Hukou* migrants. They are merged into the wage component, which is also well-documented in the yearbooks. In addition, *Hukou* migrants who have changed their rural identities to urban are regarded as urban residents in China, and their incomes are thus not covered by the rural data. This clarification assures these two incomes fit our case and can be used to construct an urban-rural income disparity index. Our central variable, *gap*, is thus defined by the ratio of the real disposable income per capita of urban residents to the net income per capita of rural residents. Since we also use the differentials of the two incomes as our alternative gap index for robustness, the two incomes' series are deflated by consumer price indices in respective year, with 1996 as the base year. The data source for this gap variable is the *China Statistical Yearbooks* between 1989 and 2008.

3.1.2. High-level human capital investment (*hcinv*)

One of the major methodological challenges in empirical research of human capital is determining how to measure human capital *per se*. Literature on human capital reveals multiple indices for human capital. Some earlier work, such as that of Temple (1999) and Kruger and Lindal (2001), employs school enrollment rate or student-teacher ratio. These indices are criticized for their bias in estimating human capital (Chi, 2008). A seemingly more popular measurement uses educational attainment as the proxy of human capital (a long list of literature employs this measurement, such as Chi, 2008; De Brauw and Rozelle, 2008; Yao and Zhang, 2008, among others). Choice among this list of alternatives largely depends on the availability of the data.

In this paper, we use two indicators as the proxy for high-level human capital investment. The first is the ratio of the number of laborers with more than 12 years of schooling to the total population, denoted by $hcinv_1$. This indicator roughly measures the average tertiary educational attainment over the whole population, if we assume constant average schooling years for junior college (i.e., college of 3 years) and over (say, 15.5 years). A similar indicator is also employed in Fleisher, Li and Zhao (2010). However, an accessible data source for this indicator is only available from 1996 through 2007 in China. But our sample covers a much longer period – between 1989 and 2007. This necessitates an alternative proxy for high-level human capital investment. We use the proportion of the number of junior college, college undergraduate, and postgraduate students at school to the population of a province as the alternative proxy for human capital investment, denoted by $hcinv_2$. The data for this proxy indicator from 1989 to 2007 is available in *China Statistical Yearbooks*. Despite extensive criticism on the bias of this enrollment ratio in measuring human capital, our results indicate that it is as reliable a proxy for high-level human capital as the educational attainment indicator.

3.1.3. Higher labor migration (*uralr*)

To test our fourth "back-flow" channel is challenging because labor migration data at any level is largely unavailable in China. What we have at hand is province-level data on higher labor in urban and rural areas from the *China Population Statistical Yearbooks* from 1997 to 2006 and the *China Population and Employment Statistical Yearbooks* from 2007 to 2008.

Subject to the constraints of data availability, we use an approach similar to Zhang and Song's (2003) method to meet this challenge. The premise for our approach is to decompose the numbers of higher labor in a given year in both urban and rural areas into the stock of the previous year, alongside natural change and net migration.

We denote numbers of higher labor for the urban and rural areas in province i at the end of year t as H_{it}^{u} and H_{it}^{r} respectively. The net higher labor migrations to the urban and rural areas, M_{it}^{u} and M_{it}^{r}, in province i in year t could be derived by the following equation, where $d_{i,t-1}$ is the death rate in province i at year $(t-1)$:

$$H_{it}^{X} = H_{i,t-1}^{X} - H_{i,t}^{X} \cdot d_{i,t} + M_{it}^{X}, X = u, r, \qquad (1)$$

Here, we assume that children born to higher labor cannot become higher labor immediately after birth.

We further denote the annual growth rate of higher labor in urban and rural areas of province i at year t by $g_{it}^{X}, X = u, r,$; we thus have:

$$M_{it}^{X}=\frac{(1+d_{i,t})(1+g_{it}^{X})-1}{1+g_{it}^{X}}H_{i,t}^{X},X=u,r,$$

To make the migrations comparable across provinces, net migrations to the urban and rural areas are normalized by the respective provincial urban and rural populations, $P_{it}^{X},X=u,r,$ in province i in year t: $m_{it}^{X}=M_{it}^{X}/P_{it}^{X},X=u,r.$

One caveat for this migration rate is that the populations of urban and rural areas have not been surveyed every year during our sample period. This leads to underestimation of the populations in the non-survey years. To address this issue, we smooth the higher labor migration ratios for urban and rural areas using a five-year moving average method.

Equipped with higher labor migration ratios, we build an urban-rural higher labor migration gap index, $uralr_{it}$, for province i in year t as follows:

$$uralr_{it}=m_{it}^{u}/m_{it}^{r}=\left[\frac{(1+d_{i,t})(1+g_{it}^{u})-1}{(1+d_{i,t})(1+g_{it}^{r})-1}\cdot\frac{1+g_{it}^{r}}{1+g_{it}^{u}}\right]\cdot\frac{\dfrac{H_{i,t}^{u}}{P_{i,t}^{u}}}{\dfrac{H_{i,t}^{r}}{P_{i,t}^{r}}}.$$

This yields

$$uralr_{it}\propto\frac{h_{i,t}^{u}}{h_{i,t}^{r}}, \tag{2}$$

where $h_{it}^{X},X=u,r,$ denotes the proportion of higher labor in urban and rural areas of province i at year t.

Following Equation (2), we use the urban ratio of higher labor proportion to rural higher labor proportion in a province as our proxy for the higher labor migration gap index.

3.1.4. Other Variables

Economic development (*GDP*). We use provincial real GDP per capita as the proxy for economic development levels in each province.

Inflation rate (*cpi*). Zhou (2009) finds inflation can significantly widen the income gap at the national level. We use consumer price index minus 100 as the proxy for inflation rate.

Financial deepening (*fin*). Greenwood and Jovanovic (1990) claim, with evidence, that financial development can widen the income gap. We express financial deepening in terms of the ratio of outstanding loans to GDP.

Government intervention (*gov*). Lu and Chen (2004) document that governmental fiscal expenditure widens the urban-rural income gap in China. We follow this using the proportion of local fiscal expenditure to GDP as the proxy for government intervention.

Openness (*open*). Wang and Fan (2005) and Zhou (2009) provide empirical evidence supporting the claim that openness, in terms of imports and exports to GDP, widens the income gap. We also use this ratio to express openness.

Infrastructure Investment (*capcon*). Lu and Chen (2004) argue that infrastructure narrows the urban-rural income gap. They argue that farmers get wages from infrastructure construction work, which help narrow the gap. We also introduce this control variable into our empirical models using the ratio of infrastructure expenditure to total expenditure of local government.

3.2. Sample

Our sample covers 28 provinces over the period from 1988 to 2007, which generates a yearly province-level panel dataset. Province-level panel data are extensively employed in the literature to investigate human capital and its effects in China (e.g., Zhang and Song, 2003; Chi, 2008; Fleisher, Li and Zhao, 2010, among others). This is largely because of the difficulty, if not impossibility, of building a panel dataset covering a long period at a more micro level than the province in China. There are other micro-level sources, such as the China Households Income Projects (CHIP), which provides a large number of cross-sections containing rich information on schooling years and incomes. However, these are less informative in terms of variations along the time dimension, because it is impossible to construct panel data using such survey series, as respondents to each survey are not identical. However, for research on growth and the inequality effects of human capital, like ours, cares most about the time-varying behaviors of the effects in a long run. In contrast, we have no such problems using a province-level panel data covering tens of years. The tradeoff between the time dimension and cross-section variations leads us to choose the province-level panel data at the sacrifice of some richness in cross-sectional individual heterogeneity.

In constructing the panel data, we drop two provinces, Sichuan and Chongqing, because there is overlapping in data prior to 1997 when Chongqing became the fourth centrally administered municipality. Tibet is also dropped due to incompleteness of relevant data. The chosen sample period is the longest period with data for high-level human capital investment, urban-rural income gap, and for which the controlling variables are available. Table 1 presents descriptive statistics of the variables.

Table 1 Descriptive Statistics of the Variables (1988-2007)

Variables		Mean	Standard Deviation	Minimum	Maximum	Observations
Urban-rural income gap	gap	2.703	0.687	1.144	4.953	560
High-level human capital investment (enrollment ratio) (%)	$hcinv_1$	0.606	0.626	0.078	3.579	560
High-level human capital investment (tertiary attainment) (%)	$hcinv_2$	5.538	4.348	0.741	30.127	336
Economic development (ten thousand RMB)	GDP	0.844	0.800	0.130	6.354	560
Inflation rate (%)	cpi	6.814	7.970	-3.2	29.4	560
Financial deepening (%)	fin	100.911	29.480	54.669	240.019	519
Government intervention (%)	gov	13.384	5.316	4.917	36.013	560
Openness (%)	$open$	28.337	37.949	2.729	220.293	560
Infrastructure investment (%)	$capcon$	9.671	4.522	3.134	29.509	532

Source: *China Statistical Yearbook* from 1989 to 2008, *China Financial Yearbook* from 1994 to 2008 and *Comprehensive Statistical Data and Materials on 55 Years of New China.*

However, data for higher labor migration is only available from 1996 to 2007. We merge two data sources, *China Population Statistical Yearbook* from 1997 to 2006 and *China Population and Employment Statistical Yearbook* from 2007 to 2008 to calculate the migrations. Moreover, we use an average moving method to smooth the population data, which reduces the sample period for the migration further, to 1998 to 2005. Table 2 presents descriptive statistics of the relevant variables used in migration calculation.

Table 2 Descriptive Statistics of the Variables for Migration Index (1998-2005)

Variables		Mean	Standard Deviation	Minimum	Maximum	Observations
The proportion of higher labor in urban area	*ualr*	10.235	3.45	4.428	26.854	224
The proportion of higher labor in rural area	*ralr*	0.725	0.552	0.122	3.926	224
The ratio of the two proportions of higher labor in urban and rural areas	*uralr*	18.188	9.083	4.744	57.719	224

Source: *China Population Statistical Yearbook* from 1997 to 2006 and *China Population and Employment Statistical Yearbook* from 2007 to 2008.

3.3. Model Specifications

Basically, we need to construct empirical models that estimate the effect of high-level human capital investment on the urban-rural income gap at different levels of economic development. To this end, we estimate two nonlinear models.

Model I is a threshold model aiming at describing the trend of the effect when the level of economic development changes. The threshold is based on economic development level. After setting the proper number of thresholds, this model is able to quantify the effects of high-level human capital investment on the urban-rural income gap in a piecewise manner. By looking at the magnitude of the effects over different intervals defined by the thresholds, we know the changing pattern of the effect when economic development level varies. Equation 3 specifies Model I:

Model I
$$\begin{aligned}&\log gap_{it} = \alpha + \beta \log gap_{i,t-1} + \gamma_1 hcinv_{it} I\{GDP_{it} < \theta_1\} \\ &+\gamma_2 hcinv_{it} I\{\theta_1 \le GDP_{it} < \theta_2\} + \gamma_3 hcinv_{it} I\{GDP_{it} \ge \theta_2\} + \varepsilon_{it},\end{aligned} \tag{3}$$

where *gap* is urban-rural income gap; *hcinv* denotes high-level human capital investment. We use logarithm of *gap* as dependent variable in order to remove heteroskedasticity. *GDP* stands for real GDP per capita (RMB), with 1996 as the base year. $I\{\cdot\}$ is the indication function defined as $I\left\{GDP_{it} < \theta_i\right\} = 1$ if $GDP_{it} < \theta$, otherwise $I\left\{GDP_{it} < \theta_i\right\} = 0$, and $\theta_1 \le \theta_2 \le \theta_3$ are thresholds. Coefficients α, β, γ_1, γ_2, γ_3 and θ are unknown parameters to be estimated.

For Model I, we expect to find statistically significant coefficients γ_1, γ_2 and γ_3 with $\gamma_1 > \gamma_2 > \gamma_3$ to support our hypotheses. Such evidence would suggest that high-level human capital investment could gradually bridge the urban-rural income gap as the economy develops.

One drawback of Model I is that it cannot quantify the point at which the effects could be divided into two parts; below this point the gap would widen, while above this point it would narrow. The reason that this point cannot be quantified is that Model I sets the thresholds merely by minimizing the sum of square residual error. This is a purely statistical approach without explicit economic grounds on which to interpret the thresholds in a meaningful way.

Hence, we use Model II to identify that turning point. This is a dynamic panel model with a lag term of the gap and an interaction term of $hcinv$ and GDP, as in Equation 4:

Model II

$$\begin{aligned}\log gap_{it} &= \zeta + \theta \log gap_{i,t-1} + \lambda_1 hcinv_{it} + \lambda_2 GDP_{it} \\ &+ \lambda_3 hcinv_{it} \cdot GDP_{it} + \boldsymbol{\varphi}\mathbf{X} + \mu_i + \eta_t + \varepsilon_{it},\end{aligned} \quad (4)$$

where $\log gdp_{i,t-1}$ is the one-period lagged dependent variable; $\mathbf{X}$ is the vector of controlling variables; μ_i represents time-invariant, unobservable, province-specific effects containing information on heterogeneity; η_t represents time-specific effects which are common to all provinces and change through time; ζ is the constant; and $\theta, \lambda_1, \lambda_2$ and λ_3 are unknown parameters to be estimated. The time varying disturbance term ε_{it} is assumed to be serially uncorrelated with independent variables.

An advantage of this dynamic panel model is its incorporation of the lag term of gap, which largely controls for the lagged effects of human capital investment on the economy. Since controlling for the lagging effects smoothes the gap, we do not include the lagged terms of $hcinv$ and their interaction terms further in the model specification.

The coefficients of central interest in this dynamic panel model are λ_1 and λ_3. We expect the former to be significantly positive and the latter negative, in order to support H1 and H2.

Furthermore, this model enables us to quantify the turning point. From Equation 4, we see the partial effect of high-level human capital investment on urban-rural income gap:

$$\frac{\partial \log gap}{\partial hcinv} = \lambda_1 + \lambda_3 GDP. \quad (5)$$

Hence, if $GDP > -\lambda_1 / \lambda_3$, $\partial \log gap / \partial hcinv < 0$; otherwise, $\partial \log gap / \partial hcinv > 0$, under the conditions that $\lambda_1 > 0$ and $\lambda_3 < 0$. This indicates that the turning point in terms of GDP per capita is located at $(-\lambda_1 / \lambda_3)$. When the economic development level surpasses this threshold, human capital investment bridges the urban-rural income gap (since $\partial \log gap / \partial hcinv < 0$ in this case); otherwise, it widens (because $\partial \log gap / \partial hcinv > 0$ here).

There are a handful of model specification issues associated with Model II that we need to address. The first is the possibility of colinearity and interactions between economic development level and human capital investment. Correlation analysis indicates that the correlation coefficients between GDP and $hcinv$, using enrollment rate and average tertiary education attainment, are both larger than 0.8. This raises the potential colinearity issue. This issue can be addressed to a large extent by the panel model we use for Model II. An advantage of the panel model is its ability to effectively address the problem of multi-colinearity (Baltagi, 2005). Furthermore, we also estimate a variant of Equation 4 without GDP and find that the changes to other estimations are slight, which suggests that the potential colinearity issue is minor. In addition, an interaction term is included in Equation 4 to express the interaction between GDP and $hcinv$.

The second concern is with respect to the potential endogeneity of the central variable, $hcinv$, since urban-rural income disparity apparently affects human capital investment. The Hausman test for the endogeneity of $hcinv$ yields a large value of 17.26, which is significantly (at the level of 5%) greater than the cutoff value $\chi^2_{0.05}(7)$ in our case. Therefore, the endogeneity of $hcinv$ should be addressed.

The third potential problem is unobserved heterogeneity, which should be accounted for in the estimation, otherwise, the estimates may be biased. In our dynamic panel model specification, as in Equation 4, unobserved heterogeneity for each province is documented by the time-invariant term μ_i.

Our panel data structure lends us some advantages in dealing with the second and third concerns. We employ a two-step systematic GMM (GMM-SYS) estimation method for Model II, which is well documented in the literature (Arellado and Bond, 1991; Arellano and Bover, 1995; Blundell and Bond, 1998). This allows us to account for unobserved heterogeneity and to effectively resolve the endogeneity of the independent variable. The next section discusses these elements in more detail.

Moreover, we also construct a fixed-effect panel model for robustness checks to provide more evidence on the robustness of our main results. Such model specification employs fixed-effects of the provinces to control for unobserved heterogeneity, and an instrumental variable method is used to deal with the endogeneity issue of the central regressor, $hcinv$.

3.4. Estimation Methods

For the threshold model (Model I), we employ Hansen's (1999) method and his codes in the estimation. The model is, in essence, a least-square method using fixed-effect transformations to estimate the thresholds and regression slopes. In model specification tests, however, it uses a non-standard asymptotic theory of inference developed by Hansen (1999). The Matlab codes for this method are available on Hansen's personal homepage at www.ssc.wisc.edu/~bhansen/progs/joe_99.html.

For the dynamic panel model (Model II), Hsiao (1985) suggests that the OLS estimation would result in biased coefficients because μ_i is not directly observable and is possibly correlated with other regressors in the model. Furthermore, the correlation of $\log gap_{i,t-1}$ with μ_i would result in inconsistent estimates of coefficients. First differences of the variables could eliminate time-invariant fixed effects, but the OLS estimation here is still inefficient because of the correlation between $\Delta\varepsilon_{it}$ (i.e., $\varepsilon_{it} - \varepsilon_{i,t-1}$) and $\Delta \log gap_{i,t-1}$ (i.e., $\log gdp_{it} - \log gdp_{i,t-1}$) due to the correlation between $\varepsilon_{i,t-1}$ and $\log gdp_{i,t-1}$.

The two-step GMM-SYS method lends us promising solutions to the unobserved bias and endogeneity issue in estimating Model II. This method estimates both levels and first differences equations for Model II. The unobserved heterogeneity problem can thus be effectively resolved by estimating first differences equations, whereby the time-invariant disturbance terms μ_i are eliminated. Also, the extensive instrument variables utilized in this method can solve the endogeneity and even some other model specification problems to a large extent.

The GMM method proposed by Arellano and Bond (1991) not only employs lag term ($\log gdp_{i,t-2}$) and difference ($\Delta \log gdp_{i,t-2}$) as instrument variables, as proposed by Anderson and Hisao (1982), but also uses additional instruments obtained by utilizing the orthogonal conditions that exist between the disturbances and the lagged values of the dependent variable. This yields more efficient estimators than the earlier IV method such as that employed by Anderson and Hisao (1982).

Furthermore, the two-step GMM estimators, which use one-step residuals to construct asymptotically optimal weighting matrices, are more efficient than one-step estimators (Blundell and Bond, 1998). Because it uses orthogonal conditions on the variance-covariance matrix, the two-step GMM method can control for the correlation of errors over time, heteroskedasticity across firms, simultaneity, and measurement errors.

In addition, Arellano and Bover (1995) propose an extended GMM (GMM-SYS) estimation method using instruments in first differences for equations in levels, and instruments in levels for equations in first differences. The GMM-SYS technique estimates the dynamic panel model for both levels and first differences, as level equations are simultaneously estimated using differenced lagged regressors as instruments. This process therefore has the advantage of controlling for individual heterogeneity, while retaining variation partially.

For the reasons explained above, our examination of the threshold turning point is based on Equation 4 using the two-step GMM-SYS estimates.

Ⅳ. Empirical Results

4.1. Correlation and Graphic Analysis

Table 3 reports the correlation matrix. We learn from Table 3 that the correlation coefficients between gap and two types of proxies for high-level human capital investment are positive, at around 0.06, but neither is significant (both have large p-values greater than 0.15). A possible explanation for this result is that a simple linear correlation analysis cannot account for the complicated high-level human capital investment and urban-rural income gap nexus; and a nonlinear pattern for this nexus may exist.

Table 3 Correlation Matrix

	gap	$hcinv_1$	$hcinv_2$	*GDP*	*cpi*	*fin*	*capcon*	*open*	*gov*
gap	1								
$hcinv_1$	0.0558 P-value: 0.3238	1							
$hcinv_2$	0.0756 P-value: 0.1856	0.8616***	1						
GDP	-0.3009***	0.8033***	0.7942***	1					
cpi	0.0606	0.0546	0.0028	0.0054	1				
fin	-0.0144	0.3958***	0.6120***	0.3475***	-0.1223**	1			
capcon	0.1707***	0.1856***	0.3342***	0.3191***	-0.1761***	0.3556***	1		
open	-0.3200***	0.5966***	0.6766***	0.8130***	0.0056	0.4204***	0.3127***	1	
gov	0.6934***	0.0753	0.1374**	-0.1026	-0.001	0.3364***	0.5564***	-0.1964***	1

Note: *** and ** denote the statistical significance at the levels of 0.01 and 0.05, respectively.

Figure 2 illustrates this nonlinear pattern graphically. Panel A of Figure 2 plots the scatter graph of the urban-rural gap and high-level human capital investment (using $hcinv_2$) for the poorer provinces in which GDP per capita is below the mean value. The fractional polynomial fitted curve for the poorer provinces indicates the gap-widening effects of high-level human capital investment. Whereas Panel B of Figure 2, plotting for the richer provinces in which GDP per capita is above the mean, shows the possibility that high-level human capital investment can reduce the gap.

Moreover, correlation coefficients between the economic development level and high-level human capital investment are very large and also significant. This justifies our inclusion of the interaction term of these two variables in our model.

Among all pairs of variables in Table 3, our two proxies for high-level human capital investment have the largest correlation coefficient. Figure 3 plots the scatter graph for theses two variables. We can see that they are highly positively and linearly correlated. This confirms the usability of the enrollment rate proxy ($hcinv_1$) for high-level human capital investment, despite many critiques.

Figure 2 Urban-rural Gap vs. High-level Human Capital Investment

Panel A: For richer provinces

Panel B: For poorer provinces

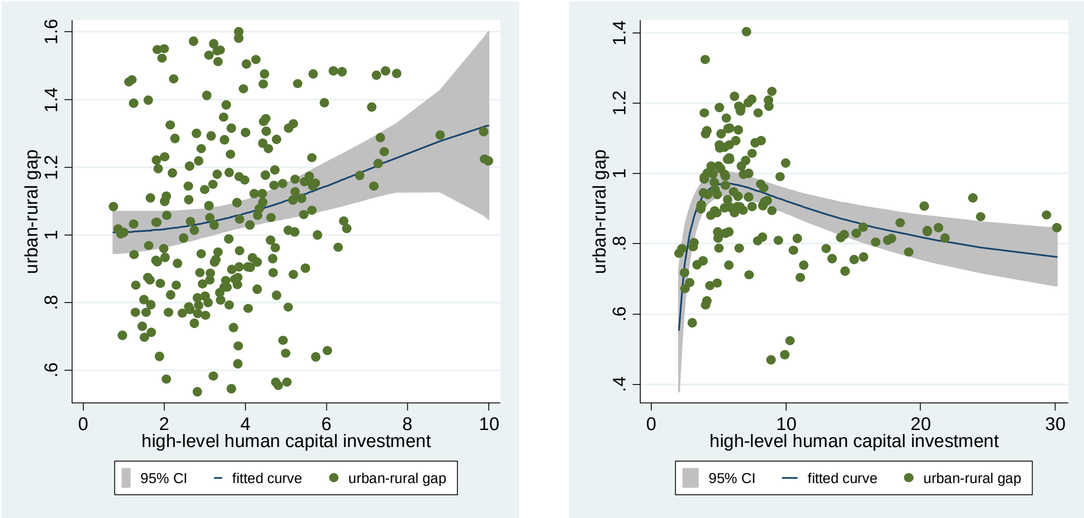
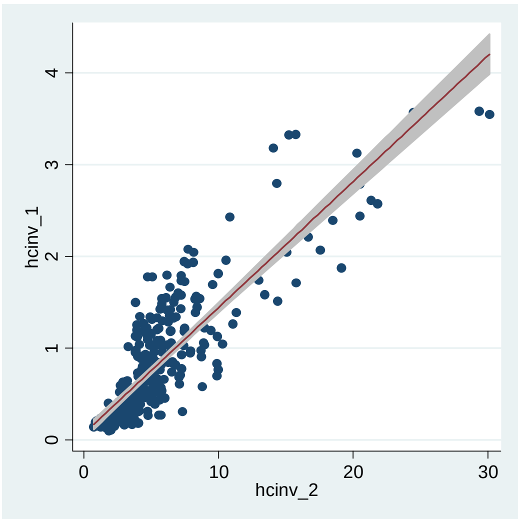


Figure 3 Positive and Linear Correlation Between Two Types of High-level Human Capital Investments Proxies



Note: The shadowed area around the fitted line presents the confidence interval of 5% level.

4.2.Results of the Threshold Model

Table 4 presents the estimation results of the threshold model. It identifies two thresholds, 3266 and 7572, in terms of real GDP per capita. Three intervals are thus generated: $[\min, 3266), [3266, 7572), [7572, \max]$. Over these three intervals, estimated coefficients of high-level human capital investment are $\gamma_1 = 0.263 > \gamma_2 = 0.116 > \gamma_3 = 0.059 > 0$, respectively. The positive coefficients indicate that high-level human capital investment widens the urban-rural income gap. However, these coefficients decrease sharply as the threshold increases. Moreover, the coefficient rapidly converges to zero. This gives evidence of the shrinking trend of the widening impacts of high-level human capital investment on the urban-rural income gap with improving economic development levels. This further implies that the widening effect might cease at a certain point of economic development and possibly reverse to narrow the gap after that. H1 and H2 are partially supported by this result.

Table 4 Estimation Results for Model I

Parameters	α	β	γ_1	γ_2	γ_3
Estimation value	0.211^{***} (0.024)	0.748^{***} (0.029)	0.263^{***} (0.069)	0.116^{***} (0.020)	0.059^{***} (0.008)
H_0:No threshold，H_1:Single threshold			F_1=6.188**(2.714, 3.792, 6.502)		
H_0:Single threshold，H_1:Double threshold			F_2=8.902***(2.641, 3.728, 7.081)		
Bootstrap count	10000	Threshold θ_1	3266	Threshold θ_2	7572
Descriptive statistics of the variables in each interval					
Intervals	Statistics	$\log gap$	*hcinv*	*GDP*	Observations
Interval 1	Mean	1.039	0.164	2484	104
	Standard Deviation	0.186	0.067	469	
Interval 2	Mean	1.002	0.327	5113	212
	Standard Deviation	0.261	0.210	1188	
Interval 3	Mean	0.917	1.136	15273	216
	Standard Deviation	0.246	0.703	7572	

Note: *** , ** and * denote the statistical significance at the levels of 0.01, 0.05 and 0.1, respectively. Standard error is bracketed below the estimated parameter.

4.3.The Threshold Effect of High-level Human Capital Investment on the Urban-Rural Income Gap

Table 5 presents the estimation results for Model II specified in Equation 4. Panel A uses the enrollment rate proxy for high-level human capital investment

($hcinv_1$), and Panel B employs the tertiary attainment proxy ($hcinv_2$). Both Panels A and B also estimate a variant of Model II using urban-rural differential in incomes as the dependent variable to check robustness.

The five controlling variables are introduced into the model in a sequential manner for the purpose of robustness check. The introduction order of these variables proceeds in the following order: *inflation rate, financial deepening*, *government intervention*, *openness*, and *infrastructure investment*. Furthermore, a variant of Model II dropping *GDP* is also estimated to check colinearity.

Columns 3 to 7 in Panels A and B of Table 5 present estimation results using the urban-rural income ratio as the dependent variable. All estimates in these columns are statistically significant, at the 5% level, and their signs remain unchanged across different columns. This shows that our results are quite robust to various model specifications. Column 8 in both Panels present the estimation results when *GDP* is dropped in order to check the impacts of a potential colinearity issue. Comparing numbers and significance in Column 8 with those in Column 7 in both Panels shows only a slight difference. This lends evidence to the effectiveness of the dynamic panel model in controlling colinearity bias (Baltagi, 2005).

Results in Table 5 give compelling evidence about the threshold effect of high-level human capital investment on the urban-rural income gap. In all columns, the estimated coefficients of high-level human capital investment are significantly positive, at the 1% level, and those of the intersection terms are negative and statistically significantly, at the 1% level in most cases, except in Column 3 of Panel A. Our results thus confirm the threshold effect and its robustness to model specifications.

Table 5 Estimation Results of Model II

Panel A: Using enrollment ratio as the proxy for high-level human capital investment (1988~2007)

Dependent Variable	Gap in ratio							Gap in difference
1	2	3	4	5	6	7	8	9
One-period Lag	L_1	0.592*** (0.240)	0.477*** (0.025)	0.501*** (0.021)	0.367*** (0.027)	0.375*** (0.027)	0.448*** (0.036)	0.803*** (0.027)
High-level human capital investment	$hcinv_1$	0.082*** (0.012)	0.120*** (0.015)	0.072*** (0.009)	0.134*** (0.026)	0.137*** (0.025)	0.171*** (0.026)	0.176*** (0.024)
Economic development	GDP	0.060** (0.025)	0.100*** (0.027)	0.084*** (0.018)	0.081** (0.036)	0.085** (0.035)		0.173*** (0.042)
Intersection term	$hcinv_1 * GDP$	–0.021** (0.008)	–0.040*** (0.010)	–0.030*** (0.006)	–0.060*** (0.014)	–0.059*** (0.014)	–0.040*** (0.013)	–0.074*** (0.018)
Inflation rate	cpi	0.002*** (0.0003)	0.003*** (0.0004)	0.003*** (0.0004)	0.004*** (0.0004)	0.004*** (0.0004)	0.003*** (0.0003)	0.003*** (0.0007)
Financial deepening	fin		0.001*** (0.0001)	0.001*** (0.0001)	0.001*** (0.0002)	0.001*** (0.0002)	0.001*** (0.0002)	0.004*** (0.0003)
Government intervention	gov			0.010*** (0.001)	0.012*** (0.001)	0.010*** (0.001)	0.007*** (0.001)	0.005** (0.002)
Openness	$open$				0.003*** (0.001)	0.002*** (0.001)	0.002* (0.001)	–0.0004 (0.001)
Infrastructure investment	$capcon$					0.003*** (0.001)	0.004*** (0.001)	0.004* (0.002)
Constant	$cons$	0.300*** (0.020)	0.273*** (0.022)	0.178*** (0.019)	0.144*** (0.037)	0.118*** (0.043)	0.148*** (0.040)	0.966*** (0.174)
Observations		532	500	500	500	472	472	472
m_2		0.165	0.190	0.222	0.723	0.570	0.350	0.150
Difference–Hansen		0.963	0.951	0.776	0.606	0.260	0.668	0.285
Threshold (RMB)		39683	29912	24209	22545	23101		23736

Note: *** , ** and * denote the statistical significance at the levels of 0.01, 0.05 and 0.1, respectively. Standard error is bracketed below the estimated parameter.

Panel B: Using tertiary attainment as the proxy for high-level human capital investment (1996~2007)

Dependent Variable	Gap in ratio							Gap in difference
1	2	3	4	5	6	7	8	9
One-period Lag	L_1	0.480*** (0.076)	0.500*** (0.080)	0.571*** (0.069)	0.430*** (0.128)	0.064 (0.222)	0.236 (0.210)	0.782*** (0.132)
High-level human capital investment	$hcinv_2$	0.040*** (0.004)	0.039*** (0.004)	0.021*** (0.003)	0.039*** (0.006)	0.046*** (0.007)	0.066*** (0.009)	0.032*** (0.007)
Economic development	GDP	0.158*** (0.028)	0.157*** (0.028)	0.108*** (0.017)	0.083** (0.041)	0.109*** (0.041)		0.240** (0.108)
Intersection term	$hcinv_2 * GDP$	–0.012*** (0.002)	–0.012*** (0.002)	–0.008*** (0.001)	–0.014*** (0.003)	–0.018*** (0.003)	–0.014*** (0.004)	–0.013*** (0.004)
Inflation rate	cpi	–0.007*** (0.001)	–0.007*** (0.001)	–0.005*** (0.001)	–0.007** (0.003)	–0.005* (0.003)	–0.004 (0.003)	–0.010*** (0.003)
Financial deepening	fin		0.021 (0.021)	0.025* (0.014)	0.005 (0.033)	0.020 (0.039)	0.00002 (0.0004)	0.073 (0.050)
Government intervention	gov			0.007*** (0.001)	0.010*** (0.002)	0.017*** (0.005)	0.014*** (0.005)	0.009 (0.009)
Openness	$open$				0.394** (0.186)	0.622*** (0.229)	0.003 (0.002)	0.039 (0.081)
Infrastructure investment	$capcon$					–0.004 (0.003)	–0.003 (0.002)	0.002 (0.005)
Constant	$cons$	0.241*** (0.041)	0.204*** (0.056)	0.151*** (0.044)	0.145* (0.083)	0.355*** (0.134)	0.257* (0.140)	1.348 (0.893)
Observations		308	308	308	308	280	280	280
m_2		0.398	0.421	0.281	0.628	0.624	0.287	0.256
Difference–Hansen		0.686	0.681	0.587	0.804	0.551	0.806	0.127
Threshold (RMB)		33171	32487	27042	28127	26139		24056

Note: *** , ** and * denote the statistical significance at the levels of 0.01, 0.05 and 0.1, respectively. Standard error is bracketed below the estimated parameter.

We next estimate the value of the threshold turning point. Using Equation 5, we calculate the threshold values for each model specification and report them in the last row of Table 5. Since we introduce the control variables sequentially, the estimated values of λ_1 and λ_3, which determine the threshold values, vary considerably from Column 3 to Column 6 in both Panels. As the full inclusion of the control variables improves the model specification relative to the parsimonious ones, we thus use the threshold estimation in Column 7 as the reliable variable.

In Panel A of Table 5, Column 7 gives the partial effect of high-level human capital investment on the urban-rural income gap by $0.137 - 0.059 * gdp$. This yields the threshold value of the turning point in terms of the economic development level of 23,101 RMB GDP per capita. Similarly, Column 7 in Panel B yields a threshold estimation of 26,139 RMB GDP per capita. These two estimation results are very close, with the latter using attainment proxy being slightly larger, because it is estimated over the period from 1996 to 2007, while the former is estimated over the period from 1988 to 2007. This closeness is consistent with the high and linear correlation between the enrollment and attainment proxies, and also confirms the robustness of our threshold estimation.

The estimated threshold value, either by enrollment or attainment proxy, is higher than that of the national level over the sample period. The national real GDP per capita is 14,815 RMB in 2007, significantly less than the threshold values. This explains why high-level human capital investment has been widening the urban-rural income gap in China at the national level during our sample period of 1988 to 2007.

To check the robustness of this estimated threshold value presented in Column 7, we re-estimate the model in Column 7, for both Panels A and B, using instead the urban-rural differential in incomes as the dependent variable. The results are presented in Column 9 in both panels of Table 5. For Panel A, we find the estimation of λ_1 in Column 9 increases relative to that in Column 7, while λ_3 decreases. Whereas for Panel B, the estimations of both λ_1 and λ_3 decrease. However, changes in the threshold values calculated from Column 9 and Column 7 are slight, within the range of 10%. This lends further evidence to the robustness of our estimation of the threshold value.

Many of the sample provinces have already crossed the threshold. These include Shanghai, Tianjin, Beijing, Jiangsu, Zhejiang, Guangdong, Fujian, and Shandong[1]. In these regions, increasing high-level human capital investment narrows their urban-rural income gaps.

We can also postulate that China would cross the above threshold by 2014, if China's real GDP per capita increases at an annual rate of 7.5%. By then, the high-level human capital investment would narrow the urban-rural income gap at the national level.

In addition, our results regarding the control variables are largely consistent with the literature, with the exception of one. We find that inflation rate, financial deepening, government intervention, and economic openness all play a role in widening the urban-rural income gap, as predicted in the literature. However, our results suggest that infrastructure investment widens the income gap as well, which departs from the results of Lu and Chen (2004).

4.4. Further Robustness Checks

In order to confirm that our results are robust to alternative model specifications dealing with the endogeneity issue of the central regressor, $hcinv$, we further estimate a fixed-effect panel model using an instrumental variable (IV) approach. This model specification uses fixed-effects to account for the unobserved heterogeneity of each province explicitly.

We choose the state government fiscal appropriation for education ($sgae$) as our instrumental variable. This can be justified as follows: the state fiscal appropriation for education is allocated to each province within a state planning framework, which is very weakly connected to the provincial urban-rural income disparity; but it is an important driver for education enrollment or attainments. We thus find $sgae$ is, perhaps weakly, exogenous to the dependent variable, gap, while it is a determinant of the central regressor, $hcinv$. Therefore, $sgae$ can be used as an instrumental variable.

[1] Shanghai, Beijing, Tianjin, and Zhejiang succeeded in crossing above the critical value of economic development in 1997, 2002, 2003, and 2005 respectively. Jiangsu and Guangdong succeeded in 2006. Fujian and Shandong did so in 2007.

Table 6 Further Robustness Check Using Fixed-effect Model Specification

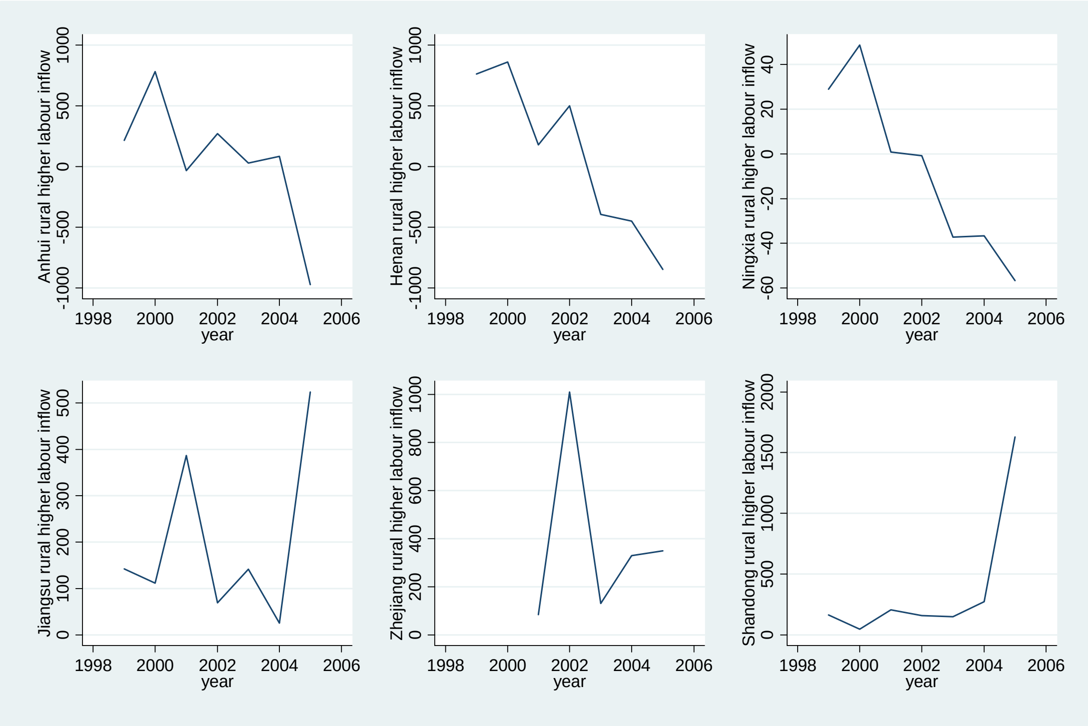
1	2	3	4
Dependent variable		Gap in ratio	Gap in ratio
One-period Lag	L_1	0.544*** (0.064)	0.565*** (0.080)
High-level human capital investment	*Enrollment proxy for column (3); Attainment proxy for column (4)*	0.165*** (0.035)	0.064*** (0.019)
Intersection term	$hcinv*GDP$	-0.040*** (0.008)	-0.011*** (0.003)
Inflation rate	cpi	-0.005*** (0.002)	-0.009*** (0.003)
Financial deepening	fin	0.018 (0.020)	-0.053* (0.029)
Government intervention	gov	0.005* (0.002)	-0.001 (0.004)
Openness	$open$	0.146*** (0.045)	0.178*** (0.062)
Infrastructure investment	$capcon$	0.003* (0.001)	0.002 (0.002)
Constant	$cons$	0.228*** (0.047)	0.191*** (0.056)
Threshold (RMB)		41250	59787
R^2 (overall)		0.7119	0.2698
F test that all u_i=0		2.56 [Pvalue: 0.0001]	1.55 [Pvalue: 0.0446]
Wald chi2(8)		157812.94 [Pvalue: 0.000]	91781.19 [Pvalue: 0.000]

Table 6 reports the estimation results of the fixed-effect model specification using IV estimator. Column 3 presents the estimation results using enrollment rate as the proxy for high-level human capital; and Column 4 presents the results when tertiary attainment is the proxy. Comparing the coefficients of the central regressors of interest, $hcinv$ and $hcinv \times GDP$, in Columns 3 and 4 of Table 6 with Column 7 in Panel A, and Column 6 in Panel B of Table 5, respectively, we find that they are within the same scale and at same significance level. This lends evidence that our main result is robust and the unobserved bias is effectively addressed in the GMM-SYS estimation.

3. Evidence on the Migration Hypothesis

We further test the migration hypothesis (H3) explaining the threshold effect of high-level human capital investment on the urban-rural income gap in China. Figure 4 plots higher labor migration in the rural regions for some representative provinces. It shows that the provinces with higher GDPs per capita than the estimated threshold value, such as Jiangsu, Zhejiang, and Shandong, all have higher labor inflows to the rural regions with an increasing trend. But for the poorer provinces falling below the estimated threshold value, such as Anhui, Henan, and Ningxia, the results show increasing outflows of the higher labor from the rural regions. This graphic comparison gives some primary evidence on our migration hypothesis.

Figure 4 Higher Labor Migration in Rural Regions in Typical Provinces



To test the migration hypothesis formally, we construct a model in the same spirit of Equation 4, with the higher labor migration gap index as the dependent variable. We expect evidence of similar threshold effects on the higher labor migration gap to support hypothesis 3. This empirical model is:

$$uralr_{it} = \zeta + \theta uralr_{i,t-1} + \lambda_1 hcinv_{it} + \lambda_2 hcinv_{it} \cdot GDP_{it} + \varepsilon_{it} \quad \textbf{(6)}$$

where $uralr$ denotes the higher labor migration gap index defined in Section 3.1.

We expect the estimation of $\lambda_1 > 0$ and $\lambda_2 < 0$ to support H3. If so, such results are coincident with our estimation results of Equation 4. This would indicate that the pattern of higher labor migration (produced by high-level human capital investment) duplicates that of high-level human capital investment on the urban-rural income gap. This would suggest that labor migration is among the drivers of the threshold effect of human-capital investment on the income gap. The estimation results of Equation 6 are presented in Table 7:

Table 7 Estimation Results of the Migration Model

Dependent variable		the ratio of the proportions of higher labor in urban and rural areas
Lagged one period	L_1	0.805*** (0.013)
High-level human capital investment	$hcinv_1$	0.657*** (0.164)
Interaction term	$hcinv_1 \times GDP$	-0.304*** (0.016)
Constant	$cons$	1.929*** (0.428)
Observations		196
m_2		0.139
Difference-Hansen		0.759

Note: *** , ** and * denote the statistical significance at the levels of 0.01, 0.05 and 0.1, respectively. Standard error is bracketed below the estimated parameter.

Results in Table 7 present an inverted U-shaped pattern of higher labor migration with respect to economic development levels. Using the same method in Equation 5, we find the threshold value of the inverted U-shaped turning point in migration is 21,614 RMB GDP per capita. This threshold value for migration is very close to that of human-capital investment on the urban-rural gap (23,101 RMB GDP per capita).

The coincidence of these two thresholds predicts that the two processes, higher labor migration to rural areas and the shrinking of the urban-rural income gap, would arise simultaneously when GDP per capita exceeds the threshold, around 20,000 RMB. The former is the driving process, while the latter is the resultant one, as high-level human capital investment is the source of producing higher labor. This gives strong evidence supporting H3.

Ⅴ. Conclusions

This paper finds a threshold effect of high-level human capital investment (using tertiary education investment as the proxy) on the urban-rural income gap in China. This finding is of substance in the context of rapidly increasing tertiary education investments in both urban and rural China.

Estimation results of two non-linear models provide supporting evidence of this threshold effect. We estimate that the threshold value is located at around 20,000 RMB GDP per capita. Many provinces or regions, such as Shanghai, Tianjin, Beijing, Jiangsu, Zhejiang, Guangdong, Fujian, and Shandong, have already crossed this threshold. In these regions, high-level human capital investment serves to narrow the urban-rural income gap. Furthermore, we forecast that China could cross the threshold by 2014, assuming an annual real GDP per capita growth rate of 7.5%. After crossing the threshold, China's high-level human capital investment would narrow the urban-rural income gap at the national level.

References

Anderson, T.W. and Hsiao, C. (1982), "Formulation and Estimation of Dynamic Models Using Panel Data", *Journal of Econometrics*, No. 18, pp. 47–82.

Arellano, M. and Bond, S. (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and An Application to Employment Equations", *Review of Economic Studies*, No. 58, pp. 277–297.

Baltagi, B.H. (2005), "Econometric Analysis of Panel Data(Third Edition)", Wiley.

Blundell, R. and Bond, S. (1998), "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models", *Journal of Econometrics*, No. 87, pp. 115–143.

Brown, P.H., and Park, A. (2002), "Education and poverty in rural China", *Economics of Education Review*, No.21, pp.523-541.

Cai, F. (2005), "Why did not labor migration bridge urban-rural income gap?" *Theoretical front*, No. 20, pp. 18-20. (in Chinese)

Cheng, K. and Li, J. (2007), "Analysis of Mechanism and Dynamic Econometric Relationships between Urban Bias, Urbanization and Urban-rural Inequality", *The Journal of Quantitative and Technical Economics*, No. 7, pp. (in Chinese)

Chi, W. (2008), "The Role of Human Capital in China's Economic Development: Review and New Evidence", *China Economic Review*, No. 19, pp. 421-436.

De Brauw, A., and Rozelle, S. (2008), "Reconciling the Returns to Education in Off-Farm Wage Employment in Rural China", *Review of Development Economics*, No. 12, pp.57-71.

Fan, L. (2008), "Measuring Interprovincial Flows of Human Capital in China: 1995–2000", *Population Research and Policy Review*, No. 28, pp. 367-387.

Fleisher, B., Li, H. and Zhao, M. Q. (2009), "Human Capital, Economic Growth, and Regional Inequality in China", Forthcoming in *Journal of Development Economics*.

Gou, R. (2006), "Explore on China's High Education Equality from Admission to Urban and Rural Students", *Research on Education Development*, No.5A, pp.29-31. (In Chinese)

Greenwood, J. and Jovanovic, B. (1990), "Financial Development, Growth, and the Distribution of Income", *Journal of Political Economy*, No. 98, pp. 1076–1107.

Guo, J.X. (2005), "Human Capital, the Birth Rate and the Narrowing of the Urban-Rural Income Gap", *Social Sciences In China*, No. 3, pp. 27-37.

Hansen, B.E. (1996), "Inference When A Nuisance Parameter Is Not Identified Under the Null Hypothesis", *Econometrica*, No. 64, pp. 413–430.

Hansen, B.E. (1999), "Threshold Effects in Non-Dynamic Panels: Estimation, Testing, and Inference", *Journal of Econometrics*, No. 93, pp. 345–368.

Kruger, A.B., and Lindal, M. (2001), "Education for Growth: Why and for Whom?", *Journal of Economic Literature*, No. 39, pp. 1101-1136.

Li, F. (2008), "Heterogeneous Labor Migration and Regional Income Disparity: Theoretical Analysis and Empirical Study", PhD Dissertation of Zhejiang University. (in Chinese)

Li, Y., Whalley, J., Zhang, S. and Zhao, X. (2010), The Higher Educational Transformation of China and Its Global Implications, Forthcoming in *The World Economy*.

Lin, J.Y., Wang, G. and Zhao, Y. (2004), "Regional Inequality and Labor Transfers in China", *Economic Development and Cultural Change*, No. 52, pp. 587–603.

Lu, M. and Chen, Z. (2004), "Urbanization, Urban-Biased Economic Policies and Urban-Rural Inequality", *Economic Research Journal*, No. 6, pp. 50-58. (in Chinese)

Lucas, R.E., and Stark, O. (1985), "Motivations to Remit: Evidence from Botswana", *The Journal of Political Economy*, No.93, pp.901-18.

Rosenzweig. M.R., and Stark. O. (1989), "Consumption Smoothing, Migration and Marriage: Evidence from Rural India", *The Journal of Political Economy*, No.97, pp.905-918.

Scully, G.W. (1969), “Interstate Wage Differentials: A Cross Section Analysis”, *American Economic Review*, No. 59, pp. 757–773.

Temple. J. (1999), “The New Growth Evidence”, *Journal of Economic Literature*, No.37, pp.112-156.

Wang, X.L. and Fan, G. (2005), “Income Inequality in China and Its Influential Factors”, *Economic Research Journal*, No. 10, pp. 24-36. (in Chinese)

Whalley, J. and Zhang, S. (2007), “A Numerical Simulation Analysis of (Hukou) Labour Mobility Restrictions in China”, *Journal of Development Economics*, No. 83, pp. 392–410.

Windmeijer, F. (2005), “A Finite Sample Correction for the Variance of Linear Efficient Two-step GMM Estimators”, *Journal of Econometrics*, No. 126, pp. 25–51.

Yang, D.T. (1999), “Urban-Biased Policies and Rising Income Inequality in China”, *American Economic Review*, No. 89, pp. 306–310.

Yao, X. and Zhang, H. (2008), “Education, Human Capital and Regional Economic Differentials”, *Economic Research Journal*, No. 5, pp. 47-57. (in Chinese)

Zhang, H., Zhang, L., Luo, R., and Li, Q. (2009), “Does Education Still Pay Off in Rural China: Revisit the Impact of Education on Off-farm Employment and Wages”, *China & World Economy*, No. 16, pp. 50-65.

Zhang, K.H. and Song, S. (2003), “Rural-urban Migration and Urbanization in China: Evidence from Time-Series and Cross-Section Analyses”, *China Economic Review*, No. 14, pp. 386–400.

Zhang. L.X., Huang. J.K., and Rozelleb. S. (2002), “Employment, Emerging Labor Markets, and The Role of Education in Rural China”, *China Economic Review*, No.13, pp.313-328.

Zhao, Y. (1997), “Labor Migration and Returns to Rural Education in China”, *American Journal of Agricultural Economics*, No.79, pp.1278 – 1287.

Zhao. M., and Glewwe. P.W. (2010), “What Determines Basic School Attainment in Developing Countries? Evidence from Rural China ” *Economics of Education Review*, No.29, pp.451-460.

Zhou, Y.B. (2009), “Urbanization, Urban-Rural Income Gap and Overall Income Inequality in China:An Empirical Test of the Inverse-U Hypothesis”, *China Economic Quarterly*, No. 4, pp. 1239-1256.

Zhu, N. (2002), “The Impacts of Income Gaps on Migration Decisions in China”, *China Economic Review*, No. 13, pp. 213–230.