Front. Educ China 2009, 4(3): 413–434
DOI 10.1007/s11516-009-0022-1

RESEARCH ARTICLE

YANG Jun, HUANG Xiao, LI Xiaoyu

Educational inequality and income inequality: An empirical study on China

Abstract Based on the endogenous growth theory, this paper uses the Gini coefficient to measure educational inequality and studies the empirical relationship between educational inequality and income inequality through a simultaneous equation model. The results show that: (1) Income inequality leads to educational inequality while the reduction of educational equality does not contribute to the decrease of income inequality, and there is no simple casual effect between them. However education expansion is beneficial to reduce educational inequality and income inequality. (2) Education relates to income inequality through the human capital transmission mechanism, but this mechanism does not automatically translate into a virtuous cycle of "educational equality←→income equality". (3) In the long run, the reduction of educational inequality does not reduce income inequality, but income inequality has a negative instant-impact on educational equality. (4) At present, the level of educational investment and urbanization do not effectively promote educational equality. In addition, the robustness of the model used in this paper has been partly proved.

Keywords educational inequality, income inequality, empirical study

摘要 基于内生增长理论，构建联立方程组模型，采用教育基尼系数衡量教育不平等，研究教育不平等与收入分配的作用机理及方向发现：(1) 收入分配差距导致教育不平等，教育不平等的改进却没能促进收入分配差距的改善，教育不平等与收入分配差距并非简单线性关系，但教育扩展有利于教育和收入不平等的改善；(2) 教育通过人力资本传导机制与收入分配之间发生联系，但其不会自发形成"教育平等←→收入平等"的良性循环；(3) 长期内教育不平等的降低并没有改善收入不平等，但收入不平等在当期就能加剧教育不平等程度；(4) 目前教育投入的水平、城

Translated from *Guanli Shijie* 管理世界 (Management World), 2008, (1): 38–47

YANG Jun (✉), HUANG Xiao, LI Xiaoyu
College of Economics and Business Administration, Chongqing University, Chongqing 400030, China
E-mail: yangjun@cqu.edu.cn

市化进程并未有效地改进教育不平等。另外，模型的稳健性得到部分证实。

关键词 教育不平等，收入不平等，实证分析

1 Introduction

Since China's reform and opening-up, rapid economic development has been accompanied by a larger income inequality, so that income distribution becomes a key hindrance in the construction of a harmonious society. Education is one of the most important key factors affecting the income inequality, with educational inequality playing a decisive role in its formation. At the same time, income inequality in turn also exercises a great influence on inequalities in education.

Most of the previous studies done on the income inequality and educational inequality, such as Schultz (1960), Becker and Chiswick (1966), Mincer (1974), Becker (1975), construct a human capital model with income inequality, arguing that both the AYS (average years of school attained) and education distribution affect income distribution. Supposing everyone has the same initial wealth and educational return, Farre (2000) argues educational inequality (measured by standard variation of AYS) and income inequality (measured by logarithm variance of income) have a linear relationship. Gregorio and Lee (2002) argue income distribution is closely related to people's AYS and its distribution: The more unequal the education, the deeper the income inequality will be. However, given the preset education distribution, the increase of AYS will have an uncertain effect on income distribution. Galor and Zeria (1993) studied the influence of initial wealth distribution on education distribution, arguing that income equality can promote human capital accumulation. According to the incomplete credit market, better income equality will give educational opportunities to people who do not have access to loans through the credit market. Benabou (1996) and Aghion (1998) have done related studies from the perspective of redistribution. They argue that redistribution policies are helpful in preventing severe income inequality. It is obvious that education, as an important channel for the formation of human capital, always plays a great role in explaining income inequality and educational inequality.

Through empirical studies, Londono (1990) and Ram (1990) first propose that education distribution may have an inverted-*U* curve relationship; that is, with the increase of AYS, educational inequality would first reach a peak value then decline. Ram (1990) does research using cross-sectional data from 94 countries, concluding that the peak value of the inverted-*U* curve is seven years of AYS. Thomas, Wang and Fan (2003) also performed a study using cross-sectional data data from 140 countries in 1990. They find that by setting AYS at over 15 years,

there is no inverted-*U* relationship between the Gini coefficient, which measures educational inequality, and AYS. But the standard variation of education does have a significant inverted-*U* relation with AYS, with the peak at six or seven years. At the same time, they also use panel data from 140 countries during 1960–2000 to do a regression analysis, and the empirical result indicates the standard variation of education also has an inverted-*U* relation with AYS. Furthermore, this relationship can be applied for either the estimation of fixed effect or random effect. For multiple countries, Gregorio and Lee (2002) find standard variation of education (1965; 1990) and the AYS, which lagged five years (1960, 1985), have a significant inverted-*U* relationship, with a peak value of 4.2 years. Existing studies mainly focus on testing the *U* curve hypotheses with cross-country data, but few focus specifically on data from within one country.

Beck and Chiswick (1966), Chiswick (1971), Tinbergen (1972) and Winegarden (1979) all use standard variations of AYS to measure educational inequality, and their sample selections include US, Canada, Netherlands and a series of cross-country data. They conclude that there is a positive correlation between educational inequality and income inequality, meaning that the decrease of educational inequality can help reduce income inequality.

Psacharopoulos (1977) measures educational inequality by through varied coefficients of educational enrolment at different levels. The study, which is based on cross-sectional data from 49 countries, finds that educational inequality has a negative relation with the Gini coefficient. In addition, educational inequality can explain up to a twenty-three-percent income distribution.

Ram (1984) also measures educational inequality, but adopts the bottom 80 percent income population and the bottom 40 percent income population as income variation variables to measure educational inequality. Based on data from 28 countries, this paper argues the relationship between education inequality and income inequality has no statistical significance.

Park (1996) uses standard variation of AYS and its difference coefficient to measure educational inequality and uses the Gini coefficient, the bottom 40 percent income population and the bottom 20 percent income to measure income inequality. Through the empirical study on cross-sectional data from 59 countries and the findings of this study indicate the more unequal education is, the deeper the income distribution inequality will be.

Gregorio and Lee (2002) use data from the population with over 15 years of AYS from Barro and Lee (1997), and then compute the educational standard variation of each country with five years interval[1] from 1960 to 1990 as the

[1] There are seven sections in all: 1960, 1965, 1970, 1975, 1980, 1985, and 1990. But the number of samples is not consistent since the data acquisition. For example, in 1965 and 1990, there are only 23 and 71 samples.

educational inequality variable. Moreover, this paper uses the Gini coefficient and five equal divisions of income as income distribution variables. It adopts the seemingly-unrelated-regression method to perform the empirical study, and the results show that educational inequality is an obstacle to the improvement of income inequality.

Chinese scholars have also done some related research on income distribution from the point of view of educational expansion and educational returns, but few studies have been done from the point of view of educational inequality. Lai Desheng (1997) uses cross-country data and regression analysis to verify that educational expansion and income distribution inequality have a significant inverted-*U* curve relationship. Bai Xuemei (2004) estimates the human capital model proposed by Becker and Chiswick (1966), and the empirical result shows that educational expansion and income distribution inequality have an inverted-*U* curve relationship, a conclusion that Du Peng (2005), Yang Jun and Li Xuesong (2007) also affirmed. Chen Yuyu, Wang Zhigang and Wei Zhong (2004) decompose the equation of wage distribution. The result shows that both the increase of educational returns and the enhanced correlation between education and wage lead to the enlargement of wage disparities, but the decrease of educational inequality can help relieve the wage inequality. Li Xuesong and James Heckman (2004) use newly available Chinese micro data to estimate the returns of college education in the late 20th century in China with regard to heterogeneous returns among individuals and differences of selection. Their results demonstrate that heterogeneity among people leads to substantial educational returns. Wan et al. (2004), Chen Zhao, Lu Ming and Jin Yu (2004) propose that educational development inequality among different provinces in China is the key reason in the formation of regional income disparity. At the same time, due to the increase of people pursuing higher education, the sustained development of education will help decrease disparities in income distribution among different provinces. Lu Ming and Chen Zhao (2005) use a distributed lag model to study the relationship between education and income distribution. They argue that the influence of income distribution inequality on education is varying as time goes by, and is not linear. Wang Xiaolu and Fan Gang (2005) find that improvement of AYS in China's urban areas can lead to an increase of income distribution inequality, meaning that people's rights to educational opportunities are also unequal. People with higher incomes have more educational opportunities than those with low incomes, a phenomenon that also leads to more severe income inequality. Researchers including Li Shi and Zhao Renwei (1999) and Chen Zongsheng (2001) have also done studies on income distribution from the perspective of education or human capital.

Although existing studies have done econometric analysis on the relationship between educational inequality and income distribution, most of them use

standard variations of AYS to measure educational inequality. They mainly adopt the following econometric methods: (1) From the perspective of educational returns, some studies try to investigate the relationship between education and income distribution gap by quantile regression or estimating human capital. (2) By using time series data or panel data, other studies have tried to determine whether the inverted-*U* curve between education expansion and income inequality exists, or the causal effect between them through a single-equation econometric model.

This paper uses an education Gini coefficient to measure educational inequality, and is a study on the relationship between educational inequality and income inequality through the simultaneous equations model (SEM). SEM can explore the interaction among a few endogenous variables while the single-equation econometric model can only describe its unidirectional causality. Economic variables always interact with each other through direct or indirect ways. On the basis of relevant economic theories, SEM can help use these endogenous variables to do interactive and dynamic studies that single-equation econometric models cannot. In fact, this paper uses education inequality and income distribution inequality as two endogenous variables in the SEM. At the same time, we also investigate the instant-impacts and cumulative-impacts among endogenous variables, so that the interaction of endogenous variables can be more accurately described.

2 Models, data and econometric methods

2.1 Empirical model

The theoretical research of education inequality influencing income distribution is originated from Schultz (1960), then Becker and Chiswick (1966) build a formal model. With the basis of former study, Gregorio and Lee (2002) construct a theoretical model as follow:

$$Var(\log Y_s) = \bar{r}^2\, Var(S) + \bar{S}^2\, Var(r) + 2\bar{r}\,\bar{S}\, Cov(r,S) + Var(u)\,.$$

$Var(\log Y_s)$ stands for income inequality, S stands for educational expansion (AYS), $Var\,(S)$ stands for educational inequality and r stands for educational returns. On the other hand, based on the presumption that credit market is imperfect and human capital is indivisible, Galor and Zeria (1993) firstly develop initiative research about how income inequality affects educational inequality. Because the credit market is imperfect, the initial income distribution gap leads to different investment levels of human capital. Income inequality can thus influence educational inequality. This also allows educational inequality ($Var\,(S)$)

between individuals to be explained by income inequality ($Var(\log Y_s)$) . Educational inequality and income inequality are two endogenous variables that interact with each other.

This paper quotes the theoretical analysis framework from Gregorio and Lee (2002) and Galor and Zeria (1993), which analyzed educational inequality and income distribution, the study sets educational inequality and income inequality as two endogenous variables into a SEM. The SEM can be constructed as follows[2]:

$$INEQ = f(EDINEQ_{PDL}, EDINEQ, GR, AYS, EDIN, WEST, CENTRAL) \quad (1)$$

$$EDINEQ = f(INEQ_{PDL}, INEQ, AYS, EDIN, URBAN, WEST, CENTRAL) \quad (2)$$

In model (1), $EDINEQ_{PDL}$ stands for the cumulative-impact of education inequality on income distribution, or $EDINEQ_{PDL} = \sum_{i=2}^{m} \alpha_i INEQ_i$. *INEQ* measured by the income Gini coefficient shows income inequality. The education Gini coefficient (*EDINEQ*), which indicates educational inequality, is introduced to test how educational inequality affects income inequality in an endogenous growth environment. Economic growth rate (*GR*) is used to find how economic growth influences income distribution gap after controlling educational inequality and income inequality in the model (1). Through variable AYS we can know whether AYS contributes to reduced income inequality (Knight and Sabot, 1983).[3] In addition, education has the attributes of a public good, and government investment in education always leads to educational expansion, so we introduce the variable *EDIN*. Lastly, because Chinese economic development is regionally unbalanced, we use *WEST* to stand for western areas in China and *CENTRAL* for central areas, to determine if the relationship between educational

[2] Different assumptions of education return lead to inconsistent result. Under the hypothesis of keeping other variables constant, the increasing of education inequality (Var (*S*)) may have the income distribution gap deepened. If the education return (*r*) and education expansion (s) are independent, then the raising of education expansion would create more income inequality; If the education return (*r*) and education expansion (*S*) are negatively correlated, then education inequality may have uncertain relationship with income inequality. In addition, the micro data of education return are not easy to acquire. So we use Gregoro and Lee's (2002) methods as reference, do not consider the effect of education return at this time. But the explanation in later work would consider education return.

[3] Knight and Sabot (1983) argue that education has structural effect and wage-compressing effect. Education may increase the proportion of people with high educational background, so structural effect of education background may have income distribution gap deepened at first, then went down later. As another point of view, the wage-compressing effect can reduce income inequality, because the varying of labors' educational background may lead to the variation of supply-demand in labor market. Altogether, the education expansion has uncertain effect on income inequality.

inequality and income inequality varies based upon region.

In model (2), $INEQ_{PDL}$ stands for income inequality's cumulative-impact on educational inequality, or $INEQ_{PDL} = \sum_{i=2}^{m} \alpha_i EDINEQ_i$. Galor and Zeria (1993) were the first to argue that income equality promotes human capital accumulation, so this paper uses *INEQ* as the income inequality variable to test its effect on educational inequality. Because Ram (1990) put forth that educational expansion can reduce educational inequality and that more government investment in education may contribute to education expansion, we incorporate these two variables *AYS* and *EDIN* in this study. Moreover, to test whether urbanization helps the scale effect of education, we introduce the variable *URBAN*. Finally, *WEST* and *CENTRAL* are used again to determine regional differences in educational inequality.

In order to study the cumulative-impact and instant-impact between education inequality and income inequality, this paper uses the distributed lag model, which effectively analyzes the influence of dependent variables on independent variables over time. The model's general expression can be written as follows:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots\cdots + \beta_k X_{t-k} + \mu \tag{3}$$

In equation (3), k stands for lag length; the regression coefficient β_0 is the instant-impact multiplier, which is how a one-unit variation of X influences Y at a given moment. $\beta_1, \beta_2, \cdots \beta_k$ are called delayed-impact multipliers, which measures the lagged impact on Y with sustained variation of X over different periods, So $\sum_{i=1}^{k} \beta_i$ is called the cumulative-impact multiplier. It stands for how a one-unit variation of X affects Y. Model (3) is difficult to successfully estimate because of multicollinearity. As a result, this paper adopts the Almon polynomial to estimate model (3). The Alomon polynomial assumes the distribution of β_i can be approximately expressed as low-order polynomial of i, allowing this method to decrease the number of parameters to be estimated. Its detailed transformation is as follows[4]:

Suppose β_i can be written as,

$$\beta_i = \alpha_0 + \alpha_1 i + \alpha_2 i^2 + \alpha_3 i^3 + \cdots \alpha_m i^m, i = 0,\ 1, \cdots k. \tag{4}$$

Then put (4) into (3), and the finite distribution lag model could be written as:

[4] Quotes from Damodar Gujarati, Basic Econometrics, Mc Graw-Hill, Inc., 1978.

$$
\begin{aligned}
Y_t &= \alpha + \sum_{i=0}^{k} (\alpha_0 + \alpha_1 i + \alpha_2 i^2 + \cdots + \alpha_m i^m) X_{t-i} + \mu_t \\
&= \alpha + \alpha_0 (X_t + X_{t-1} + \cdots X_{t-k}) + \cdots \alpha_m (X_{t-1} + 2^m X_{t-2} + \cdots + k^m X_{t-k}) + \mu_t \\
&= \alpha + \alpha_0 Z_{0t} + \alpha_1 Z_{1t} + \cdots + \alpha_m Z_{mt} + \mu_t . \qquad (5)
\end{aligned}
$$

Using the OLS method to estimate equation (5), we get the estimated value $\alpha_1, \cdots, \alpha_m$. Plug the estimated $\alpha_1, \cdots, \alpha_m$ back into into equation (4) again, so that the regression coefficient of equation (3) can be given.

Of course, before estimating equation (5), it is necessary to identify the lag-length *t* and polynomial order m. Firstly, we use AIC and SC [5] to determine lag-length *t*. Then we add the value of m from 2 to 4 [6], to determine the order of the polynomial (Lu and Chen, 2005). When using educational inequality and its lagged variables to explain income inequality, the lag length *t* is 2 and the polynomial order *m* is 2; when using income inequality and its lagged variables to explain educational inequality, the lag length *t* is 3 and the polynomial order *m* is 2. Because educational inequality and income inequality are included in the SEM model as endogenous variables, the lag length *t*= 3 means the delayed-impact (when *t* >3) can be neglected. Moreover, in order to avoid multicollinearity, it is necessary to subtract some samples from the first three-year period before estimating SEM.

2.2 Data description

The income Gini coefficient (*INEQ*) can be computed as $G = \frac{1}{\mu} \sum_{i=2}^{N} \sum_{j=1}^{i-1} P_i \left| y_i - y_j \right| P_j$ (Thomas, Wang and Fan, 2003), where μ stands for the expected income value, *N* stands for the number of people, and y_i and P_i denote the average income of group *i* and population proportion of group *i*, respectively. To simplify, the formula given below is usually adopted: $G = \sum_{i=1}^{N} W_i Y_i + 2 \sum_{i=1}^{N-1} W_i (1 - V_i) - 1$ (Chen, 2007). This modified formula can be used to calculate the income Gini

---

[5] $AIC = \log(\frac{\sum \hat{\mu}_i^2}{N}) + \frac{2k}{N}$, In order to determine the lag length, through adding the lagged value till the AIC reaches its minimum. $SC = \log(\frac{\sum \hat{\mu}_i^2}{N}) + \frac{k * \log N}{N}$, in order to determine the lag length, through adding the lagged value till the *SC* reaches its minimum. But it strictly limits the number of variables in the right side of equation.

[6] Generally think, when using Almon polynomial, its number of order is always $\leqslant 4$.

coefficient for urban and rural areas. After grouping all samples according to income, Y_i stands for the population proportion of group *i*, $V_i = Y_1 + Y_2 + \cdots Y_i$. Then we use $G = P_1^2 \frac{\mu_1}{\mu} G_1 + P_2^2 \frac{\mu_2}{\mu} G_2 + P_1 P_2 \left| \frac{\mu_2 - \mu_1}{\mu} \right|$ (Sundrum, 1990) to compute the total income Gini coefficient, where $G_1, G_2, P_1, P_2, \mu, \mu_1, \mu_2$ respectively denote the urban income Gini coefficient , the rural income Gini coefficient, the population proportion of urban areas, the population proportion of rural areas, the average income of all people, the average income of urban people, and the average income of rural people[7]. Part of the income Gini coefficients in this paper is quoted from Chen Changbing (2007). Regional data computations were made for the remaining data between 1996 and 2004. The basic data originate from *China Statistical Year Book* and *China Rural Household Survey Year Book* from 1997 to 2005.

The education Gini coefficient (*EDINEQ*) and AYS between 1996 and 2004 in all provinces of China originate from Yang Jun and Li Xuesong (2007). Formulas are $EL = \left(\frac{1}{\mu} \right) \sum_{i=2}^{n} \sum_{j=1}^{i-1} p_i \mid y_i - y_j \mid p_j$ and $AYS = \sum_{i=1}^{n} p_i y_i$, in which μ stands for AYS, p_i stands for the population proportion of corresponding level to AYS, y_i stands for the different education level through AYS, and *n* stands for the packet number of educational attainment.

Economic growth rate (*GR*) and government investment in education (*EDIN*) orginate from *Chinese Statistical Year Book* from 1997 to 2005. The urbanization rates (*URBAN*) are calculated as "non-agricultural population ÷ total population."

CENTRAL represents the provinces Shanxi, Henan, Anhui, Jiangxi, Hubei, and Hunan, and *WEST* represents Sichuan, Yunnan, Guangxi, Shaanxi, Chongqing, Gansu, Qinghai, Ningxia, Xinjiang, Tibet, Guizhou and Inner Mongolia. The remaining regions are calculated as eastern areas.

2.3 Method of estimation

This paper uses the three stage least squares method (3SLS) to estimate SEM. 3SLS not only estimates all equations of SEM at the same time, but also utilizes the sample information better. When using massive samples, the estimation results of 3SLS are more effective than the two stage least squares method (2SLS) and the limited information maximum likelihood method (LIML). The main idea

[7] In China, the calculation methods for the income Gini coefficient is still in dispute. Because micro survey data are not easily acquired, this paper uses the method we describe in the paragraph to compute the income Gini coefficient.

of 3SLS can be summarized as: 3SLS=2SLS+GLS, meaning that 2SLS is utilized to estimate each equation in the SEM first, then GLS is used to estimate the whole SEM.

In addition, SEM recognition must be considered before estimating SEM. Model recognition is essential in estimating the result of SEM from the estimated inductive coefficient in the model. The importance of model recognition lies in the necessary and sufficient conditions for estimating SEM. In models (1) and (2), we can see *INEQ* and *EDINEQ* are endogenous variables, and other variables are predetermined variables. According to the order conditions and rank conditions of recognition, our SEM can be exactly recognized. Thus results are more effective and consistent.

3 Empirical results and analysis

By using STATA 9.0 software to estimate, the results are listed as follows[8]:

3.1 The instant-impact and cumulative-impact between educational inequality and income inequality

Table 1 shows the estimated result of the distributed lag model. From estimation coefficients, we can see the instant decrease of income inequality causes the reduction of educational inequality immediately. But one year later, the reduction of income inequality instead contributes to the increase of education inequality, and this negative effect lasts for two years. As we know, human capital determines individual income level. The reason that instant income inequality reduction leads to a decrease of educational inequality may be because although the instant income inequality is decreasing, people who want to lessen the income inequality further may increase their investment in education.[9] As a

[8] Because of a few missing values in Hebei and Gansu, their income Gini coefficients are used instead of the average value of eastern and western areas. If estimating SEM without Hebei and Gansu, there is an insignificant difference according to the result with these two provinces. In addition, there are missing values for Chongqing, 1996.

[9] At present, the educational returns of China exhibit the following features: (1) The educational returns of China are increasing rapidly (Huang, 2006). (2) Chinese educational returns from high school or above is significantly higher than that of junior middle school or below (Zhang, 2006), marking an obvious heterogeneity in educational returns. (3) The increase of educational returns leads to an increase of wage gap (Chen, Wang and Wei, 2004). All these features show that achieving higher level education is becoming an important way of getting even higher educational returns. But in 2006, the AYS of China has already reached 8.5 years. From the point of view of society, it is undoubted that through raising educational investment or achieving higher education, people can get high educational returns so that income inequality can be reduced.

result, educational inequality decreases instantly. However, despite a more equal income and a rise in educational investment, inequality in education actually increases. This result is inconsistent with common understanding. There are two main reasons for the result. The first explanation is that rural education return is lower than that of urban areas (Hou, 2004). This phenomenon directly leads to an urban-rural income gap. For unprivileged groups, this phenomenon even has a greater negative effect on access to education. The second reason is that despite the constant increase of AYS, the expansion of educational scale could not alleviate educational inequality stemming from family background [10](Li, 2006). Therefore, because of the macro environment, educational inequality instead increases. In addition, from the estimated coefficient absolute value, educational inequality is significantly influenced by instant income inequality. Lastly, we still cannot be certain of the cumulative-impact of income inequality on educational inequality, because its estimated coefficient is not significant.

Table 1 Regression result

Independent variable	Dependent variable			
	INEQ		*EDINEQ*	
	Coefficient	*T*-statistic	Coefficient	*T*-statistic
INEQ (lag=0)			0.14788	1.79371*
INEQ (lag=1)			−0.15981	2.01328**
INEQ (lag=2)			−0.13112	1.60229*
Sum of lags			−0.14305	−1.34221
EDINEQ (lag=0)	0.01817	0.30758		
EDINEQ (lag=1)	−0.12677	2.64991***		
EDINEQ (lag=2)	−0.14458	3.11216***		
EDINEQ (lag=3)	−0.03525	0.5544		
Sum of lags	−0.28843	2.42361**		

[10] Educational sociology theories argue that the expansion of educational scale cannot reduce educational inequality stemming from family background. Because vulnerable groups do not have complete access to newly added educational opportunities, all students at right age may compete for them. If the distribution mechanism is not changed, the distribution proportion of educational opportunities will also remain unchanged. In fact, for vulnerable groups, educational expansion can increase educational opportunities absolutely, but educational opportunities and educational inequality cannot be improved relatively. This is called the theory of keeping maximum inequality. The theory argues that more educational opportunities has the effect of equality, but such an effect would happen only if some critical points are reached. Only the enrollment rate of privileged groups has been saturated, the influence of family background would become less significant. As a result, under the background of educational expansion, educational inequality will be reduced from primary education, then secondary education and higher education.

(*Continued*)

	Equation (1)		Equation (2)	
	Coefficient	*P*-value	Coefficient	*P*-value
EDINEQ	−1.469839	0.029**		
GR	0.0122969	0.00***		
AYS	−0.0683259	0.023**	−0.0459346	0.000***
EDIN	0.9906005	0.336	1.68088	0.000***
WEST	0.04998088	0.00***	−0.0245457	0.021**
CENTRAL	0.017292	0.254	−0.0275631	0.000***
INEQ			0.3646381	0.009***
URBAN			0.0600326	0.000***

Note: a. Because of the elimination of the first three years of samples, the factual number of samples used is 185. b. *** represents significance at the 1 percent level; ** represents significance at the 5 percent level; * represents significance at the 10 percent level.

Secondly, educational inequality has an insignificant instant positive effect on income inequality. That one year later, education would have a negative influence on income inequality means the reduction of educational inequality could lead to the increase of income inequality, and the trend may last for two years. In addition, the estimated results (when lag=1 and lag=2) are significant at the 5 percent level. From another point of cumulative-impact, the correction of educational inequality also does not contribute to the alleviation of income inequality. The main reason for this may lie in the human capital transmission mechanism. This paper will proffer a detailed explanation on the deviation from the mechanism later on.

3.2 The influence of education on income distribution

In model (1), the estimation result shows that educational inequality has a negative relationship with income inequality, which is significant at the 5 percent level. The reason might be that the restriction of model (2) has been considered when estimating model (1). In other words, the interaction between educational inequality and income inequality must be considered together. According to common understanding, the reduction of educational inequality should help decrease income inequality. There are also studies (Beck and Chiswick, 1966; Chiswick, 1971; Tinbergen, 1972; Winegarden, 1979) that use standard deviation of AYS to measure educational inequality, that can confirm this common understanding. However, this paper obtains a different result. Considering the interaction between educational inequality and income inequality, the empirical results for China show that the improvement of educational inequality does not contribute to the decrease of income inequality. A related study done by Muta

(1987) on Japan obtains the same result as this study. The main reason may be that the human capital transmission mechanism does not run as normally as we wish. There is deviation from the mechanism. It is generally thought that the reduction of educational inequality helps promote human capital accumulation and improve wages through raising the marginal productivity of human capital. However, at the same time, the effect of educational structure also plays its role in deepening income inequality. The improvement of educational inequality enlarges the groups with high educational backgrounds and high productivity[11], causing more income inequality. From another point of view, the Chinese labor force market generally exhibits a supply that exceeds demand and low allocation capability, so these economic characteristics can be seen as "catalysts" in deepening income inequality. In fact, this paper argues educational inequality and income inequality may not have a single linear relationship, and may have a nonlinear relationship. Especially in developing countries, educational expansion and educational equality, does not lessen the income inequality, but instead contributes to it.

The AYS has a negative impact on *INEQ* at the 5 percent significant level, meaning that educational expansion helps reduce income inequality. Studies by Tinbergen (1972), Maria and Pshcharopoulos (1986), Tilak (1989), and Zhou Wenxing (2002) also all agree with this result, and most believe educational expansion and income inequality have a linear relationship. But, Ram (1984) and Lai Desheng (1997) suggest that educational expansion and income inequality may have an inverted-*U* curve relationship. This paper's empirical results show that expansion of Chinese education contributes to the decrease of income inequality. No matter what the relationship between education and income inequality, it is clear that China is at a stage at which educational expansion could prompt income equality.

Although the estimated result of *EDIN* is not significant, it seems that the more the government invests in education, the less severe income inequality will be. The reason may be a lower level of educational investment in China coupled with long-term structural defects. Despite the increase of AYS and gradual reduction of educational inequality in recent years, the structural defects of educational investment still have not brought about enough positive impact. Worse still, these structural defects tend to deepen income inequality.

The economic growth rate (*GR*) has a significant positive relationship with

[11] As the *National Statistical Bulletin of Educational Development—2006 AYS*, published by China's Ministry of Education, the total number of people achieving higher education has exceeded 250 million. The gross enrollment rate at institutions of higher education has reached 22 percent. This means Chinese higher education has reached the "popularization stage," which is recognized internationally as a gross enrollment rate of higher education is 15 percent above. The impact of this educational structure has just appeared.

income inequality, meaning that with rapid economic growth, income inequality also increases. Kuznets inverted-*U* theory argues that with economic growth, the income inequality first rises, then falls. However, in China, the validity of this trend of income inequality has yet to be verified (Yang and Zhang, 2003). Since reform and opening-up, economic growth has affected income inequality through the accumulation of human capital and material capital. Human capital reduces income inequality while material capital increases it. But the estimated result of *EDINEQ* shows that a decrease in educational inequality does not help lessen income inequality. The main reason for this may also be a deviation from the human capital transmission mechanism. Therefore, after controlling educational features, economic growth still leads to the increase of income inequality

The estimation coefficient of *WEST* is significant at the one percent level, which means that owing to educational equality, income inequality is more likely in western areas than in eastern areas. Furthermore, in practice, the AYS of western areas just reached 6.99 years in 2004, which is much lower than the central and eastern areas. The western education Gini coefficient and education standard variation is 0.29 and 0.08, both maximum values among the three areas, and the western economic development level, and educational development level both fall behind eastern and central areas. The practical situation above makes it clear that educational inequality in the west is more serious than other regions. Chen Zhao, Lu Ming, Jin Yu (2004) also argue that the difference in educational development is a key factor in the formation of income inequality. The variable *CENTRAL* has a positive but insignificant effect. The reason for this is that central economic development benefits from location advantage, better educational levels and economic policy, compared to some western provinces.

3.3 Income inequality affects educational inequality

In model (2), the estimated coefficient shows income inequality has a positive impact on educational inequality, meaning greater income inequality causes greater inequality in education. This model also proves income inequality is an important determinant of influencing educational inequality, showing that income inequality can determine individual educational levels and educational returns through different human capital investment levels. From this, it can be seen that income inequality is the direct cause of educational inequality. If this circular mechanism cannot be improved, the “The Matthew Effect” begins its vicious cycle.

Government investment in education (*EDIN*) shows significance at the one percent level, seemingly indicating that *EDIN* promotes educational inequality. The reason for this may be explained in three ways.

Firstly, in 2006, government investment in education accounted for 3.41

percent of GDP. Compared with the average world investment, China's index is not only far lower than that of developed countries, but also lower than some developing countries.

Secondly, educational investment always focuses more on urban and developed areas, and less on rural and undeveloped areas[12].

Thirdly, the investment channel is imperfect and is not efficient (Zhou, 2003). From the experience of developed countries, American education funds always account for 7.1 percent of GDP, and its AYS can reach 13.17; Japanese educational investment accounts for over 5 percent and AYS reaches 12.78. As a developing country, Brazil keeps that ratio above 5.1 percent every year. The common feature these countries have is the promotion of educational expansion and educational equality through long-term high investment in education, but this index in China has just reached 3.41 percent, falling behind the international average of 5 percent. Thus, the government should add investment in education to encourage educational expansion and educational equality.

The AYS has a negative impact on educational inequality, meaning that educational expansion can cause reduced educational inequality. Yang Jun and Li Xuesong (2007) get the same result in their study. Ram (1990) uses the standard variation of AYS to measure educational inequality and shows that educational expansion and educational inequality have an inverted-*U* curve relationship, with a peak value of seven years. However, when using the education Gini coefficient to measure educational inequality, the inverted-*U* relationship has not been comprehensively proven (Yang and Li, 2007). In 2006, the AYS of China had just achieved 8.5 years. According to the empirical result, this paper argues China should be at a stage of educational expansion that is beneficial to educational equality. Therefore, the government should insist on developing education to

[12] In 2003, the investment in education are listed in the table below:

Index	Educational funds divided by different educational levels					Educational funds divided by different regions			Educational funds divided by urban and rural	
	Higher education	Secondary vocational education	Secondary education	Primary education	other	East	Central	West	Urban	Rural
Total funds (RMB billion)	876.87	139.52	1 192.21	1 268.7	189.82	2 417.4	548.27	884.95	2 707.79	1 142.83
Ratio (%)	23.9	3.8	32.51	34.59	5.18	62.78	14.24	22.98	70.32	29.68

Source: *China Education Statistical Year Book*, 2004 and *China Statistical Year Book*, 2004.

From the table above, the Chinese mechanism of education investment shows obvious unbalance regionally and between urban and rural area. Although the investment level in higher education is still less than in secondary and primary education, higher education accounts for a large portion of funding. Moreover, with the development of private education, more and more funds will be put into higher education.

achieve educational equality.

Next, the estimation coefficient of *URBAN* is positive and significant at the one percent level. It shows that the acceleration of urbanization does not contribute to the decrease of educational inequality, supposing other factors are constant. Generally speaking, urbanization should benefit the improvement of education, exert the educational diffusion effect and reduce educational inequality, but the empirical result is not consistent with this pattern. There are two reasons for the inconsistency. Firstly, many small cities are less developed than big cities in each province, especially in western provinces. In small cities, those involved in agriculture always account for a large proportion of the total population. Because agriculture often leads to sparsely populated areas, receiving education, especially high quality education, is much more difficult than in big cities. Secondly, in order to unify the standard of calculating the urbanization rate, this paper uses the computing formula "Agricultural population ÷ Total population." However in most provinces, a rising urbanization rate does not translate into increased city development. It is worth noting that some of the non-agricultural population ratio is still dependent on urban planning so that local people do not fully enjoy the conveniences of urbanization. In addition, the floating population (rural to urban migrant workers) and their offspring do not have access to educational services in urban areas in China ruling out the educational diffusion effect for the Chinese situation, it is necessary to consider the characteristics of urbanization, particularly for "soft environment" issues such as education.

The dummy variables *WEST* and *CENTRAL* are significant, and have the similar estimated coefficient. That is to say that after controlling other related factors, the western and central areas tend to have greater educational inequality than eastern areas. This phenomenon is very similar to the unbalanced economic development of China, where the east leads the central and the western regions. This makes educational equality all the more important for western and central China.

Fig. 1 summarizes the interaction among all variables in the SEM as below. "+" stands for positive correlation and "–" stands for negative correlation.

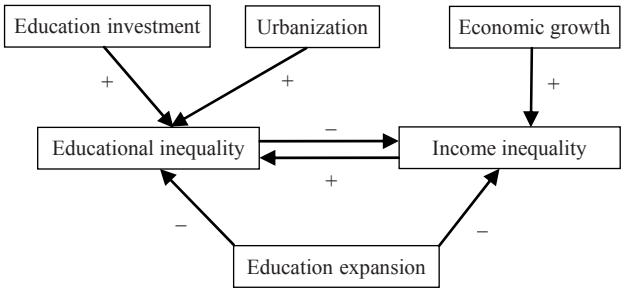

Fig. 1

3.4 Robustness test

As most former studies use SDS to measure educational inequality, this paper will follow suit rather than adopt the education Gini coefficient to test the robustness of the SEM.

Table 2 Robustness test

Independent variable	Dependent variable			
	Result from Table 1		Robustness test	
	INEQ	*EDINEQ*	*INEQ*	*SDS*
EDINEQ	−1.469839**			
	(0.029)			
GR	0.0122969***		0.0126463***	
	(0.00)		(0.003)	
AYS	−0.0683259**	−0.0459346***	−0.0020607	0.0126725
	(0.023)	(0.000)	(0.748)	(0.53)
EDIN	0.9906005	1.68088***	0.0570865	6.917334***
	(0.336)	(0.000)	(0.948)	(0.002)
WEST	0.0499808***	0.0245457**	0.0973188***	−0.039887
	(0.000)	(0.021)	(0.000)	(0.699)
CENTRAL	0.017292	−0.0275631***	0.0151979	−0.1938309**
	(0.254)	(0.000)	(0.488)	(0.009)
INEQ		0.3646381****		2.538858*
		(0.009)		(0.058)
URBAN		0.0600326***		0.3974999**
		(0.000)		(0.011)
SDS			−0.2282895*	
			(0.1001)	

Note: The value in the brackets represents *P*-value, others notes are the same as Table 1.

From the results of robustness test, *INEQ* in model (2) and the coefficient of SDS in model (1) are both significant at the 10 percent level, and they have the same signs as the original results (Shown in Table 1), indicating a strong robustness of interaction between education inequality and income inequality. However, when using SDS, the coefficients of AYS are not significant in either model (1) or (2), meaning the robustness of educational inequality, educational expansion and income inequality has not been proved. The comprehensive model has only been partially proven.

4 Conclusions

This paper uses the education Gini coefficient to measure educational inequality.

It then uses the SEM to study the interaction between educational inequality and income inequality, and the distributed lag model to find their instant-impact and cumulative-impact. From provincial panel data from 1996–2004 in China, the following conclusions were reached.

Firstly, the improvement of educational inequality does not lessen the income inequality. From the perspective of estimated results by the distributed lag model, the reduction of educational inequality also cannot promote the reduction of income inequality in the long term, perhaps caused by deviation from the human capital transmission mechanism. According to common understanding, the reduction of educational inequality should benefit the accumulation of human capital and enhance wages through the improvement of marginal productivity of human capital, all steps that should lead to the reduction of income inequality. However, because of the Chinese labor force market, the effects of Chinese educational structure, and a supply that exceeds demand, reduction of educational inequality instead leads to more severe income inequality. The solution instead should be to work towards higher quality economic development to correct the human capital transmission mechanism, so that the reduction of educational inequality will then transfer to a lessened the income inequality.

Secondly, the decreasing of income inequality may reduce education inequality significantly. According to the Hysteresis effect, inequality in education decreases instantly as the income inequality is reduced and education inequality is an instant-impact of income inequality. However, in our study in the following two years, the decrease of income inequality actually leads to more unequal education. This phenomenon may be a result of lacking educational investment and the current situation of educational returns. Income inequality affects individual human capital investment level directly, then through the income effect of educational return, the educational inequality and income inequality are influenced again. Here it is obvious that income inequality is the direct cause of educational inequality. If this circular mechanism is not improved, we will see the vicious cycle of the "The Matthew Effect" take place. Therefore, the government should pay more attention to the improvement of educational inequality.

Thirdly, the human capital transmission mechanism on the basis of educational equality does not have an automatic positive effect as it should in theory. Generally speaking, more educational investment in education should contribute to educational expansion, which should help decrease the educational inequality. At the same time, because of educational expansion, individual human capital accumulation and educational returns can be improved, and the income inequality can be reduced, further lessening education inequality. Based on the Chinese situation and our empirical results however, this virtuous cycle does not form automatically. The most important reason is that China is in a period of

economic transformation period, if the country is to construct a virtuous cycle involving educational equality, economic development and reduced income inequality, policies and measures must be adopted. On the one hand, full use must be made of the mechanism that educational expansion corrects educational inequality and income inequality, by increasing educational investment. On the other hand, because of disadvantages such as the a flawed labor market mechanism, urban-rural differences in educational return and imbalances educational investment, measures must be taken to promote the mutual promotion of educational equality and income equality. For example, more educational investment must be paid in compulsory education to improve access to education among unprivileged groups while reforming the current distribution system.

Fourthly, the empirical results show that owing to education inequality, the western China is more likely to have an income gap than the eastern China. After controlling other related factors, the western and central areas tend to have more unequal education than the eastern area. This phenomenon follows the pattern of unbalanced economic development among provinces in China. Therefore, the government should pay more attention to the west and central areas when considering educational investment and economic policy. Educational resources and distribution are both are particularly scarce in the west. Moreover, many professionals from the west have been migrating east, causing a further strain in resources for economic development in the region. In addition, this study also finds that urbanization also does not reduce educational inequality, perhaps caused by the levels of urban development and the educational investment system. For this reason, the government needs to pay more attention to soft environments such as education in cities.

Finally, according to the research progress, the relationship between educational inequality and income inequality still requires further study. Further points of study include whether the relationship between variables for China will develop as developed countries did, or if will have its special pattern, or whether the development of economic growth will correspond with the stage of the country's economic development. Unfortunately, doing further research will also require data sets that span longer periods. China is, after all, in a stage of economic transition, so there will be even more uncertainties during continued research. In addition, access to micro survey data is still difficult, and the quality of some data used in this paper needs to be improved, perhaps another direction of research for the future.

Acknowledgements The authors are grateful to the National Natural Science Funds (No. 70703038), National Educational Science Funds (No. EFA080248) and China Postdoctor Science Funds (No. 20070420478) for their financial assistances to this study.

Glossary

compulsory education 义务教育
cumulative-impact multiplier 累积影响乘数
delayed-impact multiplier 延期影响乘数
distributed lag model 分布滞后模型
education Gini coefficient 教育基尼系数
endogenous growth theory 内生增长理论
endogenous variable 内生变量
harmonious society 和谐社会
heterogeneity 异质性
human capital 人力资本
income Gini coefficient 收入基尼系数
instant-impact multiplier 即期影响乘数
Kuznets Inverted-*U* theory 库兹涅茨倒 U 理论
limited information maximum likelihood method 有限信息极大似然法
marginal productivity 边际生产力
multicollinearity 多重共线性
order condition 阶条件
panel data 面板数据
predetermined variable 前定变量
quantile regression 分位回归
random effect 随机效应
rank condition 秩条件
robustness 稳健性
seemingly unrelated regression 似不相关回归
simultaneous equation model 联立方程模型
the Matthew Effect 马太效应
three-stage least square method 三阶段最小二乘法
two-stage least square method 两阶段最小二乘法
wage-compressing effect 工资压缩效应

References

Aghion P, Howitt P (1998). Endogenous Growth Theory. Cambridge: MIT Press
Bai X M 白雪梅 (2004). Education and income inequality: An empirical study on China 教育与收入不平等: 中国的经验研究. Management World 管理世界, (6):54–55
Barro Robert J, Lee J W (1997). International measures of schooling years and schooling quality. American Economic Review, 86(2): 218–223
Becker G S (1975). Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education (2nd). New York: National Bureau of Economic Research
Becker G S, Chiswick B R (1966). Education and the distribution of earning. American Economic Review, (56): 358–369
Benabou R (1996). Equity and efficiency in human capital investment: The local connection. Review of Economic Studies, 62(2): 237–264
Chen C B 陈昌兵 (2007a). Calculation of various Gini coefficients from different regions in China and analysis using the nonparametric model 各地区居民收入基尼系数计算及其非参数计量经济模型分析. Quantitative Economy and Technical Economy Studies 数量经济技术经济研究, (1): 136–137
Chen C B 陈昌兵 (2007b). Mechanism between growth and income distribution—an empirical study based in VECM 经济增长与收入分配间的相互作用机制及其实证分析. Modern Economic Science 当代经济科学, (1): 62
Chen Y Y 陈玉宇, Wang Z G 王志刚, Wei Z 魏众 (2004). The varying of Chinese urban

income inequality since 1990— regional factors, human capital 中国城镇居民20世纪90年代收入不平等及其变化——地区因素、人力资本在其中的作用. Economic Science 经济科学, (6): 22–24

Chen Z 陈钊, Lu M 陆铭, Jin Y 金煜 (2004). Regional difference of human capitals and education development in China: An estimation of the panel data 中国人力资本和教育发展的区域差异：基于面板数据的估算. World Economy 世界经济, (6): 28–31

Chen Z S 陈宗胜 (2001). Population characteristics as education level affecting the income of Chinese urban people 文化程度等人口特征对城镇居民收入及收入差别的影响. Nankai Economic Studies 南开经济研究, (4): 38–41

Chiswick B R (1971). Earning inequality and economic development. Quarterly Journal of Economics, (85): 21–39

Du P 杜鹏 (2005). Chinese educational development and its influence on income distribution gap: An empirical study 我国教育发展对收入差距影响的实证分析. Nankai Economic Studies 南开经济研究, (4): 50–52

Farre A E (2000). Education and income distribution. EC426 Public Economics, London School of Economics

Galor O, Zeira J (1993). Income distribution and macroeconomics. Review of Economic Studies, (60): 35–52

Gregorio J D, Lee J W (2002). Ducation and income distribution: New evidence from cross-country data. Review of Income and Wealth, (48): 395–416

Hou F Y 侯风云 (2004). The study on rates of investment return to human capital in rural areas 中国农村人力资本收益率研究. Economic Research Journal 经济研究, (12): 78–82

Huang G H 黄国华 (2006). Unobserved effects and return to education 非观测效应和教育收益率. Quantitative Economy and Technical Economy Studies 数量经济与技术经济研究, (4): 72–74

Knight J B, Sabot R H (1983). Educantion expansion and the Kuznets effect. American Economics Review, (73): 1132–1136

Lai D S 赖得胜 (1997). Education expansion and income inequality 教育扩展与收入不平等. Economic Research Journal 经济研究, (10): 48–50

Li S 李实, Zhao R W 赵人伟 (1999). A study on income distribution of Chinese residents 中国居民收入分配再研究. Economic Research Journal 经济研究, (4): 14–15

Li X S 李雪松, James Heckman (2004). Heterogeneity, selection bias and the return to education: A empirical analysis based on Chinese micro-data 选择偏差、比较优势与教育的异质性回报：基于中国微观数据的实证研究. Economic Research Journal 经济研究, (4): 92–98

Li Y 李煜 (2006). Institutinal change and educational inequality: Mechanism in educational stratification in urban China: 1966–2003 制度变迁与教育不平等的产生机制：中国城市子女教育的获得 1966–2003. Chinese Social Science 中国社会科学, (4): 106–109

Londono J L (1990). Kuznetsian tales with attention to human capital. Paper presented at the Third Inter-American Seminar in Economics, Rio de Janeiro, Brazil

Lu M 陆铭, Chen Z 陈钊 (2005). Equality for the sake of growth: The nexus of inequality, investment, education and growth in China 因患寡，而患不均：中国的收入差距、投资、教育和增长的相互影响. Economic Research Journal 经济研究, (12): 9–12

Mincer (1974). Schooling, Experience and Earnings. Cambridge, National Bureau of Economic Research

Muta H (1987). Equlization Potential of Education in Income Distribution in Japan. [D]Ria De Janerio: Sixth World Congress of comparative Education (mimeo)

Park K H (1996). Education expansion and educational inequality on income distribution. Economics of Education Review, (15): 51–58

Psacharopoulos G (1977). Unequal access to education and income distribution: An international comparison. De Economist, (25): 383–392

Psacharopoulos G, Arriagada Ana-Maria (1986). The educational attainment of the labor force: An international comparison. The World Bank, Report No. EDT38

Ram R (1984). Population increase, economic growth, education inequality, and income distribution: Some recent evidence. Journal of Development Economics, (14): 419–428

Ram R (1990). Educational expansion and schooling inequality: International evidence and some implications. The Review of Economics and Statistics, 72(2): 266–274

Schultz T W (1960). Capital formation by education. Journal of Political Economy, 68(12): 571–583

Sundrum R M (1990). Income distribution in less development countries. London and New York: Routledge

Thomas V, Wang Y, Fan X B (2003). Measuring education inequality: Gini coefficients of education for 140 countries, 1960–2000. Journal of Education Planning and Administration, 17(1): 5–33

Tilak J B G (1989). Rates of return to education and income distribution. De Economist, (137): 454–465

Tinbergen J (1972). The impact of education on income distribution. Review of Income & Wealth, 16(2): 221–234

Wan G H, Lu M, Chen Z (2004). Golabalization and regional income inequality in China. UNU/WIDER working paper

Wang X L 王小鲁, Fan G 樊纲 (2005). The trend and influencing factors of income gap in China 中国收入差距的走势和影响因素分析. Economic Research Journal 经济研究, (10): 32–33.

Winegarden C R (1979). Schooling and income distribution: Evidence from international data. Economica, (46): 83–87

Yang J 杨俊, Li X S 李雪松 (2007). Economic inequality, human capital and economic growth: An empirical study on China 教育不平等、人力资本积累与经济增长：基于中国的实证研究. Quantitative Economy and Technical Economy Studies 数量经济与技术经济研究, (3): 40–43

Yang J 杨俊, Zhang Z Y 张宗益 (2003). A study on the inverted-*U* relationship between economic growth and income distribution in China 中国经济发展中的收入分配及库兹涅茨倒U假设再探讨. Quantitative Economy and Technical Economy Studies 数量经济与技术经济研究, (2): 32–34

Zhang C W 张车伟 (2006). Human capital return and income disparity: “Mathew Effect” and its implication 人力资源回报率变化与收入差距：“马太效应”及其政策含义. Economic Research Journal 经济研究, (12): 63–67

Zhou W X 周文兴 (2002). Economic growth and income distribution: An empirical study on Chinese urban residents 中国城镇居民收入分配与经济增长关系实证分析. Economic Science 经济科学, (1): 44–46

Zhou Y N 周运浓 (2003). The disadvantage of the current compulsory education investment system and its reform thoughts 现行义务教育投入体制的弊端及改革思考. Education and Economics 教育与经济, (4): 38–39