

Article

# Identity, Secondary Vocational Education Options and Return on Investment: Evidence from Children of Rural Chinese Families

Qiang Chen, Zhiming Yu, Chen Liu and Kaihua Zhang \*

School of Business Administration, Zhongnan University of Economics and Law, Wuhan 430073, China;  
15527032018@163.com (Q.C.); 202101080017@stu.zuel.edu.cn (Z.Y.); 18362085816@163.com (C.L.)

\* Corresponding author. E-mail: zkhpr@163.com (K.Z.)

Received: 29 November 2023; Accepted: 19 March 2024; Available online: 1 April 2024

**ABSTRACT:** With the continuous improvement of living standards, the importance of educational choice becomes more and more prominent. Based on the data of China General Social Survey (CGSS), a simultaneous equation model of identity, secondary vocational education choice and investment return is constructed. On the basis of fully considering endogeneity and sample selection bias, this paper analyzes the influence of identity on secondary vocational education choice and investment return by means of instrumental variables and propensity score matching (PSM). It is found that class differentiation is the main factor affecting class identity. The more blurred class differentiation, the higher class identity. Class identity has a significant positive impact on identity. The higher class identity, the easier it is to form identity. Identity has a direct positive impact on personal investment return. The stronger the identity, the higher the investment return. At the same time, identity has a significant positive impact on the choice of secondary vocational education. The stronger the identity, the more inclined to choose secondary vocational education. Compared with individuals with junior high school education, individuals with secondary vocational education have a higher return on education investment. Therefore, identity can not only directly improve an individual's return on investment, but also improve the possibility of an individual's choice of secondary vocational education, thereby improving an individual's return on education investment, and ultimately increasing an individual's return on investment.

**Keywords:** Identity; Class differentiation; Educational choice; Return on investment; Simultaneous equations model; Mediating effect model; Common prosperity



© 2024 The authors. This is an open access article under the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

As the supply-side structural reform in China intensifies, the country's economic and social structures have begun to transform toward those of a skilled society, and the status and role of vocational education have become increasingly prominent. Vocational education is an important foundation for training professional and technical personnel, improving employment skills, and promoting innovation and entrepreneurship. Sustained development and expansion of vocational education promote the training of skilled professionals, improve the whole society's human capital, help build strong technical support, and ultimately facilitate the construction of a great modern socialist country to realize China's rejuvenation.

China's poverty reduction initiative has shown that helping the children of poor rural families invest in vocational education is a key means for poverty alleviation through education and an important experience for China to succeed in poverty reduction. Since China's implementation of its targeted poverty alleviation strategy, more than 8 million children from poor families have received secondary and higher vocational education in China. Among them, the East and West China vocational school enrollment cooperation initiative alone has helped more than 1 million poor students in Western China rise above educational poverty. Since the decisive victory of the poverty alleviation campaign three years ago, China's vocational education participation has expanded with the enrollment of more than 3 million poor students, successfully realizing the goal of "one vocational education, one employment, and one poverty alleviation," effectively blocking the continuation of poverty from one generation to another, and laying a solid foundation for the revitalization of rural talents. Nevertheless, China's vocational education development level, especially secondary

vocational education, remains low. According to the latest data, enrollment in secondary vocational education in China accounts for 42% of the total enrollment of senior high school education, which is far below the national education poverty reduction initiative's target of 60% (Jing et al., 2022) [1]. Compared to other countries, China's per capita years of schooling (9.91 years) are much lower than 13.66 years in the United States, 13.66 years in South Korea, and 12.92 years in OECD countries (Christiaens et al., 2019) [2]. The main reason for this huge difference is that the share of the population invested in secondary vocational education in China is much lower than the average of the United States, South Korea, and OECD countries, both in terms of value and growth. Therefore, there is abundant room for improvement.

With the rapid advancement of industrialization and new-type urbanization, the gap between China's rural and urban classes has become increasingly obvious, and this has seriously weakened rural individual identity, resulting in a lower level of secondary vocational education and lower return on investment. This paper contributes to the empirical research of rural areas in the following ways. First, in terms of theoretical research and based on the classical human capital theory, this paper utilizes the theory of identity and examines why people invest in human capital. On this basis, this paper further analyzes the social and economic utility of identity, thereby enriching the identity theory. Second, this paper investigates the promoting role of identity on rural secondary vocational education choice and investment return. Specifically, this paper not only examines the size of the investment return of identity in rural areas, but also explores the motivation of rural families to invest in secondary vocational education, as well as the impact of identity on the choice of secondary vocational education. In particular, this paper advances the pioneering contributions of Akerlof and Kranton (2000) [3], showing that identity can produce certain economic effects, such as promoting educational choice and return on investment. Finally, this paper empirically tests the theoretical claims of Akerlof and Kranton (2002) [4] and finds that identity has a significantly different impact on the educational investment returns of children from rural families at different income levels, which is a topic rarely considered in the empirical studies of rural areas.

As China endeavors to build a modern nation, identity will play a more important role in influencing individual vocational education choices and investment returns. As such, it is imperative to explore whether identity has an impact on vocational education choice and investment return, how individual identity by income level influences the effect, and how the concept of "great power emphasizes technology" can be realized.

The remainder of this paper is organized as follows. Section 2 presents a review of the relevant literature and theory. Section 3 discusses the data and methods. Section 4 reports the results. Section 5 presents the robustness checks. Finally, Section 6 offers the conclusions and discussion.

## 2. Literature Review and Theory

### 2.1. 2 Identity Theory

Akerlof and Kranton (2000) [3] first defined identity from the perspective of economics. They noted that identity is an individual's perception of and feeling toward the social category to which they belong, and this feeling is closely related to the behavioral norms of that social category. If individuals expect their behavior to conform to the norms of the social category to which they think they belong, they will feel good about themselves and act accordingly. In contrast, if their behavior violates the behavioral norms of their social category, they will feel anxious about their behavior and try their best to avoid similar behavior from happening again (Ye and Zhao, 2018) [5]. Brown (2000) [6] illustrated this identity theory from the perspective of sociological psychology. He found that individuals have internal group preference for their self-categorized social category, and this thus forms their identity. Some attributes of identity are innate, such as gender, race, and ethnicity, while others are acquired, such as social identity, economic status, language, and culture (Tajfel and Turner, 2004) [7]. In addition, identity has external effects. When others in the same social category violate the requirements of the code of conduct, individuals will feel anxious and try to prevent the transgressive behavior of others (Oyserman and Destin, 2010) [8].

When applied to secondary vocational education, identity theory posits when one's choice behavior conforms to the behavioral norms of their social category, the individual will choose secondary vocational education; when one's choice behavior does not conform to the norms of their social category, the individual will reject secondary vocational education. Similarly, when applied to the return on investment, the stronger the identity, the higher the return on investment, and the opposite is true for weaker identity.

### 2.2. Influencing Factors

A large number of studies have examined the influence of identity, educational choice, and investment return (Luo and Wang, 2020) [9]. Research on individual return on investment began with Mincer's (1974) [10] study on individual

salary, in which he posited that individual salary income was mainly affected by the individual's education level, working years, and the square of the working years. Since there are various factors affecting individual investment returns, subsequent studies on investment returns controlled for the influence of a series of observable variables based on the Mincer equation. For example, the effects of gender, household registration, and urban and rural variables are controlled, and the effects of family income, parental education level, and the family's political, economic, and social status are discussed (Dean and Jolly, 2012) [11].

The impact of identity on educational choice and investment return has also attracted the attention of some scholars (Hatano et al., 2022) [12]. In the 1990s, psychologists showed that psychological factors such as identity have a significant impact on individuals' behavioral decision-making and subjective judgment. In response, Akerlof and Kranton (2000) [3] created the economic utility model of identity from the perspective of economic research. This model suggests that individuals' decisions are affected by the behavioral norms of the social categories that correspond to their unique individual characteristics. Therefore, identity can have a direct impact on an individual's investment return. On this basis, Akerlof and Kranton (2002) [4] further constructed the educational choice model of identity. They found that individuals' self-identification forms the motivation for selection at the psychological level, and then affect individuals' final educational choice. Heckman et al (2006) [13] also found that identity can have a significant impact on individual educational choice. Studies have also shown that identity can have a significant impact on investment returns (Lu and Song, 2006) [14], and Heckman et al. (2013) [15] found that identity can significantly improve individuals' return on investment in education.

In terms of identity, Destin et al. (2019) [16] found that the main factor affecting identity was class identity. Chan (1982) [17] found that China's rural society has realized a high degree of differentiation, resulting in multiple social classes. Actual conflicts and conflicts of interest between different classes have seriously weakened the identity of children in rural families and have greatly affected the harmony and stability of rural society (Browman et al., 2019) [18].

In sum, although existing studies have noted the impact of identity on educational choice, and although some scholars have independently studied the direct impact of identity on investment return, few studies have combined identity, educational choice, and investment return to comprehensively analyze the interaction of these three factors. Therefore, this study constructed a simultaneous equation model of secondary vocational education choice and investment return and explored the impact of identity on education choice and investment return.

### 2.3. Differences with Existing Studies

Due to the lack of broader attention on educational choice behavior, existing studies tend to ignore the impact of identity on educational choice, leading to significant differences in the estimation of investment return. Fleisher et al. (2010) [19] estimated that the average return on investment of rural identity in China is about 26%, after considering factors such as the nature of the employment unit and market level vocational training. Meanwhile, Kuhnen and Melzer (2018) [20] estimated that the average return on investment of rural migrant identity in China was between 15.3% and 26.8%. De Brauw and Rozelle (2008) [21] estimated that the average return on investment of family identity in rural China is about 16.3%. More recently, Guo and Wang (2020) [22] estimated that the return on investment of rural identity in China is about 27%.

In contrast, this paper considers not only the direct impact of identity on investment return, but also the indirect impact of identity on investment return through the choice of secondary vocational education. It also analyzes the influential factors of identity in the process of constructing the identity measurement scale. These are all additions to existing research. Compared with the existing studies, the introduction of identity in the analysis of rural secondary vocational education selection can overcome the endogeneity problem to a certain extent. Through scientific measurement of identity and, on this basis, estimation of the investment return on identity, the estimation bias caused by the omission of important non-observable variables can be effectively alleviated, so as to obtain more accurate, unbiased, and consistent estimation results.

## 3. Data and Methods

### 3.1. Benchmark Model

We use  $y_{ijt}$  to represent the income of children  $i$  of rural families in regions  $j$  in year  $t$  after consumer price index reduction. The purpose is to weaken the influence of inflation and other factors and take the logarithm to avoid the bias caused by excessive variance. The benchmark model is set as follows:

$$\ln y_{ijt} = \beta_0 + \beta_1 \text{Edu}_{ijt} + \beta_2 \text{Iden}_{ijt} + \beta' X_{ijt} + \phi_j + \varphi_t + \varepsilon_{ijt} \quad (1)$$

where  $Edu_{ijt}$  represents the secondary vocational education choice of children from rural families, and  $\beta_1$  represents the return on investment of children from rural families choosing secondary vocational education.  $Iden_{ijt}$  denotes the identity of children from rural families, which is used to control for the impact of identity on the return on investment in the secondary vocational education of children from rural families.  $\beta_2$  represents the direct impact of identity on investment return, which is the income effect of identity.  $X_{ijt}$  denotes the control variables at the individual, household, and labor market levels;  $\phi_j$  and  $\varphi_t$  are regional and time effects, respectively; and  $\varepsilon_{ijt}$  is the random error term.

The control variables in this study are for the individual level, such as work experience; family level, such as the family economic situation (whether in poverty); labor market level, such as the type of the employment of rural families of children (whether engaged in the primary industry), and so on. These are the typical variables affecting individual investment returns. In addition, according to the traditional Mincer equation, the square term of work experience is controlled for in Model 1 to mitigate the impact of omitted variables. Using a benchmark model, this study examines the income effect of identity and whether it has a direct impact on the return on education investment of children from rural families.

### 3.2. The Problem of Sample Selection

In Model (1), whether children from rural families choose secondary vocational education is assumed to be random, implying the assumption that secondary vocational education choice is unrelated to investment return. However, in actuality, the reason children from rural families choose secondary vocational education is usually closely related to the return on investment. In fact, children from rural families first consider their own learning ability and then their preference for investment return before deciding to invest in secondary vocational education (Wang et al., 2014) [23]. As such, whether children from rural families choose secondary vocational education is the result of a comparative selection, which leads to serious sample selection bias if not overcome. Therefore, this study further examines the secondary vocational education choice decision of children from rural families and constructs the intervention model of secondary vocational education choice. The latent variable  $Edu_{ijt}^*$  is used as a dummy variable  $Edu_{ijt}$  in Model (1), expressed as follows:

$$Edu_{ijt}^* = \delta_0 + \delta_1 Iden_{ijt} + \delta' Z_{ijt} + u_{ijt} \quad (2)$$

where  $Edu_{ijt}^*$  represents the net utility of children from rural families choosing secondary vocational education after comparing various educational choices. When  $Edu_{ijt}^* > 0$ , children from rural families choose to invest in secondary vocational education according to their comparative advantages, in which case  $Edu_{ijt} = 1$  and 0 otherwise. The theoretical analysis above shows that identity affects behavioral decision-making by influencing the motivation of secondary vocational education choice of children from rural families; thus, it is an important explanatory variable.  $Z_{ijt}$  denotes a series of other factors affecting the vocational education choice of children from rural families, namely, gender, age, family education background of children from rural families, and the development level of secondary vocational education in the region. In addition to mitigating the sample selection bias, the intervention effect model can effectively mitigate the endogeneity problem caused by unobservable factors  $u_{ijt}$  in Model (1), which affect both the secondary vocational education choice of children from rural families and the investment return of children from rural families.

### 3.3. The Problem of Sample Endogeneity

Endogeneity problems, such as causation and reciprocity, may also exist. In the intervention model, different from Akerlof and Kranton's (2002) [4] hypothesis regarding the exogeneity of identity, this study holds that a typical reverse causality exists between identity and educational choice, and identity affects the educational choice of children from rural families. Moreover, it is conjectured that in the process of education selection, children from rural families reidentify and reevaluate their identity according to the relevant information of education selection. In the baseline model, identity affects the individual's investment return through the income effect. Similarly, the investment return also affects identity in that it is important information for individuals to evaluate their self-identity. Therefore, identity is not completely exogenous, and effective instrumental variables for identity are necessary. As an important index for constructing identity, class identity has a strong correlation with identity. The China General Social Survey (CGSS) specifically asked respondents about their class identity at age 14. This was just before the respondents made their educational choices and considered the return on their investment. Therefore, in theory, the class identity of respondents at age 14 is not affected by their educational choice and investment return, thus satisfying the independence of the instrumental variables. Here, differences in the development level of vocational education at the regional level are controlled for to mitigate the bias caused by omitted variables. The influence of institutional higher education expansion

policy is also further controlled for, to remove as much as possible the direct relationship between class identity at age 14 and children’s educational choice and the investment return of rural families, and to strengthen the exogeneity of the instrumental variables. Therefore, the instrumental variable estimation equation of identity is expressed as follows:

$$Iden_{ijt} = \theta_0 + \theta_1 Stratum_{ijt} + \theta' \lambda_{ijt} + \phi_{ijt} \tag{3}$$

where  $Stratum_{ijt}$  denotes the class identity of an individual at age 14, and  $\lambda_{ijt}$  denotes other key factors affecting identity.

In summary, examining the influence of identity on vocational education choice and return on investment is a systematic problem, which must be described through multiple equations. According to Zhang et al.’s (2013) [24] work, the simultaneous equations model is set as follows:

$$\begin{cases} \ln y_{ijt} = \beta_0 + \beta_1 Edu_{ijt} + \beta_2 Iden_{ijt} + \beta' X_{ijt} + \phi_j + \varphi_t + \varepsilon_{ijt} \\ Edu_{ijt}^* = \delta_0 + \delta_1 Iden_{ijt} + \delta' Z_{ijt} + u_{ijt} \\ Iden_{ijt} = \theta_0 + \theta_1 Stratum_{ijt} + \theta' \lambda_{ijt} + \phi_{ijt} \end{cases} \tag{4}$$

### 3.4. Data

In addition to the above-mentioned data at the regional and institutional levels, all other variables in this study were derived from the CGSS. Given that the survey takes a five-year cycle and is not a follow-up survey, this study can only select samples from 2012 to 2017 to form a mixed set of cross-sectional data. Because this study focuses on the secondary vocational education of children in rural families, it only includes the data of junior high school graduates and secondary vocational education groups. Regarding return on investment, only the data of individuals receiving wages in the labor market can reveal the corresponding return on investment in education. Therefore, samples without labor wage compensation are further excluded, and the final effective observation included in the regression analysis relates to 4966 individuals. Table 1 shows the descriptive statistics of all variables.

**Table 1.** Description of the log of labor income, vocational education options, identity, and control variables.

Variable	Mean	Identity Level	
		High Identity	Low Identity
Log of labor income	9.538	9.671	9.416
Vocational education options	0.110	0.124	0.096
Identity	0.001	N.A.	N.A.
Gender	0.639	0.621	0.656
Age	41.18	40.22	42.06
Work experience	26.85	25.84	27.77
Work experience square	830.4	779.8	876.7
Absolute poverty	0.074	0.060	0.086
Family education	2.753	2.860	2.656
Type of employment	0.419	0.378	0.457
Secondary vocational development level	0.598	0.609	0.588
Expansion of higher education	0.211	0.231	0.193
Class identity at age 14	2.958	3.432	2.524

Source: Calculated by the authors based on CGSS.

Secondary vocational education choice and individual annual labor income are explained variables in the education choice model and the return on investment model respectively. In this paper, secondary vocational education selection variables are constructed according to whether the respondents choose to accept secondary vocational education, where the value of receiving secondary vocational education is 1, otherwise 0. The results in Table 1 show that on average, only about 1 out of every 10 people choose to receive secondary vocational education, indicating that there is still much room for improvement in secondary vocational education. Compared with individuals with a low sense of identity, individuals with a high sense of identity are more likely to choose secondary vocational education. In order to eliminate the effect of inflation, this paper takes 2010 as the base period to deflate the annual income of individual labor according to the CPI of that year. At the same time, in order to avoid the estimation bias caused by excessive variance, this paper takes the logarithm of individual annual labor income after adjustment. The results in Table 1 show that individuals with a higher sense of identity have a relatively higher annual labor income than those with a lower sense of identity. The above results show that individuals with a higher sense of identity are more likely to choose secondary vocational education, and the future return on education investment is also higher, which preliminarily proves the conjecture of this study.

To avoid estimation bias due to missing important explanatory variables, several control variables were selected at the individual, household, and industry levels. At the individual level, traditional research has shown that an individual's return on educational investment is related to work experience as well as the square of work experience (Browman et al., 2019) [19]. In this paper, the duration from the first job to the latest job is used as a proxy variable for personal work experience, and on this basis, the square term of work experience is constructed. Oosterbeek and Webbink (2007) [25] found gender and age differences in vocational education choices. In this paper, the gender variable is constructed according to the gender of the survey object, and the value of male is 1, and the value of female is 0. At the same time, this paper constructs the age variable according to the difference between the date of birth and the time of the survey. The type of employment also plays an important role in the estimates. This paper constructs the employment type variable according to whether the respondents work in the public sector or the private sector. The value of working in the public sector is 1, otherwise the value is 0. Ferguson et al. (2007) [26] showed that individuals born in poor families have a significantly higher willingness to work than to go to school. According to the setting of China's national poverty line, this paper defines a family whose per capita household income is below the poverty line as a poor family, and the value of a poor family is 1, and the value of a poor family is 0. This paper also selects the highest degree of father or mother as the proxy variable of family education level. In addition, this paper also selects the ratio of secondary vocational education and general high school education enrollment in an individual's province as a proxy variable for the development level of secondary vocational education to control the influence of the development level of secondary vocational education at the provincial level. Since the expansion of higher education has a great impact on China's education system, this paper constructs the variable of higher education expansion according to whether the education choice of the survey object has been affected by the expansion of higher education. If the individual faces the choice of education after 1998, the value is 1, and if the opposite value is 0.

### 3.5. Calculation of Identity

Based on the research presented in the second part, this paper maintains that identity includes self-identity and social identity. Self-identity primarily manifests as an individual's willingness to identify with a certain social category, and this includes psychological structures such as behavior, perception, and emotion (Li et al., 2022) [27]. Social identity mainly reflects whether an individual has integrated into a certain social category, including integration and identification with specific cultural organizations, relatives, friends, and other social groups (Weiner and Tatum, 2021) [28]. Therefore, when measuring the size of identity, it is necessary to consider not only an individual's perception of their own economic and social status and emotional structure, but also whether they have identified with and integrated into their surrounding social groups.

Many studies have measured identity and the measurement indicators differ based on the research purposes. Jiang and Jiang (2022) [29] measured personal identity from four perspectives: occupation, hometown, community, and organization. From the perspective of urban and rural differences, Li et al. (2022) [27] divided identity into urban identity and rural identity and found a significant difference between the two. Poll and Critchley (2022) [30] used social identity theory to build a multidimensional framework for measuring female identity, including 12 dimensions and 36 measurement indicators. A few studies have directly measured identity using the question of "whether one belongs to a certain class or social category" (Tong, 2021) [31]. Thus, it can be seen that different studies have different definitions of identity, and it follows that the selected measurement indicators are also very different. Generally speaking, measurement indicators can be roughly divided into two categories: the willingness to pay attention to self-identity and the results of social identity.

This paper argues that self-identity and social identity are two different concepts, and there is a progressive relationship between them, so identity is a multi-level and sub-dimensional concept. Self-identity is a prerequisite for social identity. Without self-identity, there is no social identity, but self-identity does not necessarily determine social identity. Therefore, it is unscientific to use a single index to measure identity. At the very least, we should measure identity from the two dimensions of self-identity and social identity.

A large number of studies use factor analysis to measure identity from multiple dimensions (Duan et al., 2021) [32]. Christiaens et al. (2022) [2] explored the application of factor analysis in the measurement of teenagers' educational identity and found that about 40% had a high degree of identity identification. Using factor analysis and the identity framework, Wang et al. (2022) [33] calculated that the proportion of people with high identity identification was 55.14%. Therefore, this paper uses the factor analysis method to measure identity.

This paper selects the following questionnaire items as measurement indicators in Table 2: “Which class do you think you are in?” “What do you consider your socioeconomic status to be?” “Do you feel happy with your life?” “Do you often read in your spare time?” “Do you often participate in cultural activities?” “Do you often get together with your relatives?” “Do you often get together with your friends?” These items are the proxy variables of class identity, status identity, emotional identity, cultural organization identity, family identity, and friend identity, respectively. These items were chosen for the following reasons: The higher the class, the more one is admired by others, resulting in higher identity (Darity, 2022) [34]. Similarly, when one’s status is higher, they tend to receive more attention from others (Chelwa et al., 2022) [35]. When positive emotion is more frequent, the easier it is for one to influence others. When one participates more frequently in cultural activities, they experience more gatherings of relatives and friends, resulting in richer social capital (Kanzola et al., 2021) [36]. All of these situations are known to result in higher identity. Therefore, the index selection in this paper is credible.

Using factor analysis, this paper calculated that 44.78% of individuals had high identity, which was within the reasonable range of related research (Xiong et al., 2021) [37]. The results of factor loading showed that the identity index could explain 63.3% and 93.9% of the total variation of these seven indicators, showing that identity could well represent the common characteristics of these indicators. Therefore, the measurement results in this paper are valid. In sum, the identity index measurement system constructed in this paper has both a theoretical basis and has passed the corresponding reliability and validity tests.

**Table 2.** The index system for measuring identity.

Dimension	Indicator	Rotational	Interpretation	Explanation
Self-identification	Class-identification	0.5983	The higher the value, the stronger the class identity	What class do you think you are? Lowest=1 → Highest=10
	Status-identification	0.5899	The higher the value, the higher the social status	What do you think of your social status? Lowest=1 → Highest=3
	Affective-identification	0.2901	The higher the value, the more the happiness	Do you think you are happy? Unluckiest=1 → Happiest=5
		0.0156	The higher the value, the more erudite the individual	Do you often read in your spare time? Never=1 → Everyday=5
Social-identification	Cultural-identification	0.0578	The higher the value, the stronger the cultural sense	Do you often take part in cultural activities? Never=1 → Everyday=5
	Relatives-identification	0.0677	The higher the value, the more the number of relatives of the individual	Do you often get together with your relatives? Never=1 → Everyday=5
	Friends-identification	0.0874	The higher the value, the more the number of friends of the individual	Do you often get together with your friends? Never=1 → Everyday=5

## 4. Results

Next, this study focuses on the impact of identity on secondary vocational education choices and investment returns of children from rural families. We start with the intervention effect model, then consider the endogeneity problem, after which we examine the benchmark model, and finally conduct the mechanism analysis and robustness test.

### 4.1. Identity and Educational Choices

Column (1) of Table 3 reports the estimation results of the intervention effect model, where the dependent variable is the secondary vocational education choice of children from rural families. The table shows that the estimated coefficient of identity is significantly positive, indicating that identity can promote children from rural families to choose secondary vocational education on the whole, which is consistent with most studies in the literature. Observing the control variables shows that at the individual level, age is positively correlated with educational choice, whereas the coefficient of gender is not significant, indicating that with the enhancement of the concept of gender equality between men and women in rural areas, the gender difference in educational investment between men and women gradually weakens. At the family level, parents with a higher education level are more inclined to invest in secondary vocational education for their children. At the regional level, the coefficient of the development level of secondary vocational education is not significant. Because of data limitations, this study can only collect information on the regional education development level from 2012 to 2017, but most of the samples made the education choice decision before this timeframe. At the institutional level, higher education enrollment expansion is positively correlated with educational choice, which further indicates that educational choice decision is closely related to the time node of making the choice. The coefficient of the class identity variable at age 14 is not significant, indicating that the exclusive constraint variables, such as the

development level of secondary vocational education and the expansion of higher education, can strengthen the exogeneity of the instrumental variables to a certain extent.

**Table 3.** Estimation results: Probit and IV Probit regression for educational choices during 2012–2017.

VARIABLES	Secondary Vocational Education Options (Secondary Vocational = 1, Middle School = 0)				
	(1) Probit	(2) IV Probit	(3) First Step	(4) IV Probit	(5) First Step
Identity	0.111 ***	0.296 ***		0.327 **	
Age	0.007	0.009 **	−0.003 **	0.009 **	−0.004 **
Gender	0.067	0.064	−0.016	0.066	−0.028
Family education	0.123 ***	0.121 ***	0.022 ***	0.119 ***	0.026 ***
Expansion of higher education	0.875 ***	0.888 ***	−0.017	0.886 ***	−0.020
Secondary vocational development level	0.577	−0.124	0.158	−0.118	0.061
Class identity at age 14	0.026		0.133 ***		
Integrated class identity at age 14					0.363 ***
Provincial fixed effect	Control	Control	Control	Control	Control
Constant	−2.272 ***	−1.696 ***	−0.359 **	−1.702 ***	−0.368 **
Sample size	4966	4966	4966	4966	4966

Significance relationships are indicated by the *p*-values: \**p* < 0.10, \*\**p* < 0.05, \*\*\**p* < 0.01.

#### 4.2. Endogeneity Treatment

Class identity at age 14 is used as an instrumental variable of identity to alleviate the endogeneity of identity and educational choice. Columns (2) and (3) of Table 3 report the regression results of the second and first stages, respectively. The former shows that the endogeneity problem does exist. Therefore, instrumental variables are necessary. From the estimation results, compared with the benchmark results of the intervention effect model, the coefficient of identity remains significant and the value is larger, indicating that ignoring the endogeneity problem will lead to serious measurement errors. The latter shows that the coefficient of class identification at age 14 is positive and significant, as expected; therefore, the instrumental variables are valid. The coefficient of the development level of secondary vocational education at the regional level is significantly positive because, by controlling the recruitment ratio of secondary vocational education, a social norm encouraging individuals to invest in secondary vocational education can be established, and a balanced education development environment is more conducive to the formation of individual identity. Studies have shown that stratification can also have a significant impact on individual educational behavior (Liu, 2015; Zhang et al., 2013) [24,38], that is, the higher the degree of class differentiation, the more obvious the class difference and the lower the overall social identity. Therefore, this study conjectures that class differentiation weakens individual identity and thus reduces individual educational investment intention. To verify this conjecture, this study reclassifies the class identity of individuals at age 14 by categorizing the classes from 1 to 3 as low class, from 4 to 7 as middle class, and from 8 to 10 as high class. The results are presented in columns (4) and (5) of Table 3. The results show that the integrated class identity has a stronger impact on individuals' identity, which helps in promoting rural family children to choose secondary vocational education.

#### 4.3. Identity and Return on Investment

According to the results in Table 3, the family education background influences both identity and education choice. The results of the auxiliary regression show that the influence of family education background on the labor remuneration of children from rural families is not significant (Because of space limitation, the results of the auxiliary regression are not reported in the main text, and interested readers can contact the author for this information). To mitigate the influence of omitted variables and enhance the exogeneity of instrumental variables, this study uses family education background and class identification at age 14 as instrumental variables of identity in the regression equation of the benchmark model. Owing to space limitation, this study only presents the regression results of the final stage of each model.

Columns (1) and (2) of Table 4 report the results using only the instrumental variables method to control for the endogeneity problem. The results show that the instrumental variables are valid. The F-statistic value of the weak instrumental variables in the first-stage test is much higher than 19.93, and the exogeneity test based on the Hansen J statistic does not reject the null hypothesis. The former shows that identity has a significant impact on individual income level. Each standard unit increase in identity increases the individual's return to labor investment by 25.7% while controlling for other factors. The coefficient of the work experience variable at the individual level is significantly positive, whereas the coefficient of the square of work experience variable is significantly negative, indicating that the return on individual labor investment first increases and then decreases with the improvement of working years when other variables are controlled for. The coefficient of poverty status at the household level is significantly negative,



indicating that on average, the better the economic condition of the family, the more financial support it can obtain, and the more conducive it is to find a higher-paying job. The coefficient of the variable of industry type of employment at the market level is significantly negative, indicating that under the same conditions, the labor remuneration of the group engaged in non-farm industry is higher. The latter shows that the income effect of identity remains significant after adding the education variable. Moreover, the individual’s secondary vocational education’s return on investment is significantly positive. When other influential factors remain unchanged, the individual’s return on investment of secondary vocational education increases by 22.4%, on average.

Because of the problem of self-selection of samples, ignoring the bias of sample selection will lead to a serious bias in the estimation results. Column (3) of Table 4 reports the results of using both the intervention effect model and the instrumental variable method to mitigate sample selection bias and endogeneity. The results show that the average return on investment of secondary vocational education for children from rural families is approximately 22.1%, which is close to the estimated result of Velde 2009, The test results show that the instrumental variables remain valid, and sample selection bias does exist (Even though the inverse Mills ratio is not considered significant at the 5% level here, a p-value of 0.050 is just around the critical value and can be considered significant). Compared with Model (2), after controlling for sample selection bias and reverse causation, the secondary vocational education investment returns of the children from rural families significantly decrease by 0.3%, while identity significantly lowers the labor return on investment by 3.5%. Thus, ignoring sample selection bias and the reverse causation problem will lead to overestimation. Further examining the effect of class differentiation on investment returns, Column (4) of Table 4 reports the results using integrated class identification at age 14 as an instrumental variable. The results show that compared with Model (3), the return on education investment increases by 0.2 percentage points to 22.3%, which is significant at the 1% level, indicating that class stratification can indeed reduce the return on secondary vocational education investment of children from rural families.

**Table 4.** Estimation results: IV and Intervention Effect Model.

VARIABLES	(1) IV without Education Choice	(2) IV with Education Choice	(3) Intervention Effect Model	(4) IV and Intervention Effect Model
Identity	0.257 ***	0.249 ***	0.215 ***	0.182 ***
Work experience	0.026 ***	0.031 ***	0.039 ***	0.038 ***
Work experience square	−0.001 ***	−0.001 ***	−0.001 ***	−0.001 ***
Absolute poverty	−1.084 ***	−1.084 ***	−1.089 ***	−1.095 ***
Type of employment	−0.584 ***	−0.568 ***	−0.570 ***	−0.575 ***
Education choice		0.224 ***	0.221 ***	0.223 ***
Inverse Mills ratio			−0.683 *	−0.626 *
Constant	9.724 ***	9.623 ***	10.016 ***	9.993 ***
Provincial fixed effect	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Sample size	4966	4966	4966	4966
R-square	0.3419	0.3467	0.3499	0.3514
F-statistic	192.579	191.856	180.928	140.822
Hansen/Sargan	0.9159	0.8808	0.2171	0.2996

Significance relationships are indicated by the p-values: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

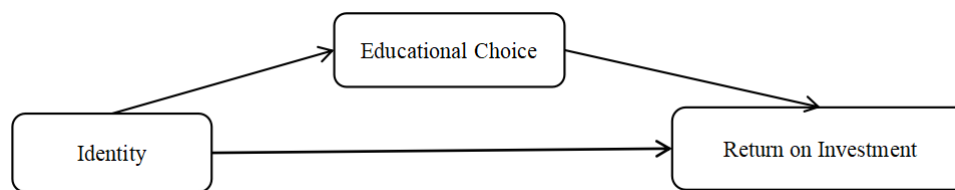
#### 4.4. Mechanism Analysis

The results discussed above indicate that identity has a significant impact on the choice of secondary vocational education and the return on investment of children from rural families. Regarding education choice, identity influences the vocational education choice of children from rural families by influencing the individual’s psychological level with respect to vocational education investment motivation. Using the individual’s class identity at age 14 as an instrumental variable, under the condition that other influential factors remain unchanged, the probability of choosing secondary vocational education of children from rural families increases by 6%, on average, for each standard unit of identity increase, and is significant at the 1% level.

Regarding investment returns, identity has a direct impact on the investment return of children from rural families by influencing the endogenous development motivation of these individuals. After controlling for the influence of endogeneity, the investment return of children from rural families increases by 25.7%, on average, for each standard unit of identity, that is, the total income effect of identity is 0.257. Because education choice directly determines the return on education investment, identity has a significant promoting effect on the secondary vocational education choice of children from rural families. Therefore, identity can not only directly affect an individual’s return on investment, but also

indirectly affect the return on investment by influencing the educational choice of individuals (Figure 1). The results of the Sobel test show that some mediating effects exist and are significant. According to the transmission effect and chain rule, after solving for endogeneity problems, such as selection bias and reverse causality, the return on investment in secondary vocational education of children from rural families increases by 1.32%, on average, for each standard unit of identity increase. The mediating effect of secondary vocational education accounts for 5.14%.

A further analysis shows that compared with social identity, self-identity, especially class identity, has a greater impact on individual identity. In terms of education choice, when controlling for other influential factors, every one standard unit increase in class identification at age 14 increases the sense of identity by 0.134 units, resulting in a 0.8% increase in the probability of investment in secondary vocational education. In terms of investment return, after controlling for the influence of other variables, when an individual is 14 years old, every one standard unit increase in class identification increases identity by 0.1 unit, resulting in a 2.57% increase in educational investment return, among which, the return on education investment generated by influencing secondary vocational education choice increases by 0.132%.



**Figure 1.** An analytical framework for influencing mechanisms.

## 5. Robustness Checks

In order to enhance the reliability of the research results, this paper uses different estimation methods to test the robustness of the results. One such method is the propensity score matching (PSM) method, which estimates the impact of identity on educational choice and investment return. Although the PSM method cannot show the direct influence of explanatory variables on the explained variables, it can accurately match samples with similar characteristics, thus improving the accuracy of estimation results. This paper also uses the sample quantile regression method to estimate the impact of identity on education choice and investment return. The sample quantile regression method can accurately describe the influence of identity and education choice on investment return under different income levels, and to some extent eliminate the estimation bias caused by the difference in income levels.

### 5.1. Propensity Score Matching Estimate

In addition to using the Heckman two-step method, researchers usually use the propensity score matching (PSM) method (Titus, 2007) [39] to mitigate the estimation bias of educational investment returns caused by model-setting errors. The advantages and disadvantages of the two estimation methods have been briefly analyzed in theory. Referring to the work of Heckman and Li (2004) [40], this study focuses on the influence of identity on education choice and investment return under the PSM method. To facilitate the comparison, this study constructs a propensity score model based on the intervention effect model to estimate the propensity score of children from rural families.

To satisfy the common support condition in this study, the predictive value of identity is first obtained by the first-stage result of the IV Probit estimation, then brought into its second stage, and finally, the predictive value of individual higher education choice is obtained. Here, the four most common matching methods are used for matching estimation (see Table 5). The results show that the average treatment effect (ATT) of participants is similar among different matching methods. Although the final estimation results are not sensitive to the selection of matching methods, compared with K-nearest neighbor matching and caliper matching, local linear regression matching uses more sample information when constructing counterfactual results, and the variance of the matching results is the smallest (Heckman et al., 1998) [41]. Compared with the results of kernel matching, local linear regression matching also uses all sample information, but the value of the ATT of participants is similar to that of the K-nearest neighbor matching and caliper matching; therefore, its conclusion is more general and representative. Hence, after strict screening and comparison, this study finally selects the local linear regression matching method for matching.

To obtain more robust matching estimation results, this study adopts the bootstrap method to perform 500 local linear regression estimates with the regression model. The results show that all the observed values using the local linear regression method are within the common value range, which satisfies the common support hypothesis. After the bootstrap method is adopted, the sample standard error is further reduced, and the matching estimation accuracy is significantly

improved. In comparing the standardized deviation of the sample data before and after matching, it is found that the standardized deviation of each variable after matching is not more than 11%, which passes the balance test. This result further proves the assumption of conditional independence among covariates after controlling for the endogeneity of the education choice equation. The bootstrap results show that for children from rural families who have chosen secondary vocational education, the average rate of return on secondary vocational education investment is 35.27%, which is higher than the result estimated through the Heckman two-step method, indicating that the endogeneity problem must be overcome in the process of estimating the return on education investment. Further comparative analysis shows that the order  $ATU > ATE > ATT$  exists no matter which matching method is used. This result shows that the return of investment in secondary vocational education is higher for junior high school graduates. These results highlight the urgency for the central decision-making authorities to issue relevant policies to encourage, support, and guide junior middle school graduates to continue to invest in vocational education, so as to improve the income level of children with junior middle school education and coming from rural families, narrow the gap between the rich and the poor, and achieve common prosperity.

**Table 5.** Robustness checks: Treatment effects under different matching methods.

VARIABLES	Treatment Group	Control Group	ATT	Standard Error	t-Value
K-nearest neighbor matching (1:4)					
before	10.037	9.477	0.560	0.046	12.15 ***
after	10.037	9.683	0.354	0.051	6.97 ***
Caliper matching (calipers scope = 0.04)					
before	10.037	9.477	0.560	0.046	12.15 ***
after	10.035	9.684	0.351	0.046	7.55 ***
Kernel matching					
before	10.037	9.477	0.560	0.046	12.15 ***
after	10.037	9.658	0.379	0.045	8.35 ***
Local linear regression matching (bandwidth = 0.4)					
before	10.037	9.477	0.560	0.046	12.15 ***
after	10.037	9.684	0.353	0.064	5.46 ***

Significance relationships are indicated by the  $p$ -values: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## 5.2. Sample Quantile Regression

Arias et al. (2002) [42] proposed that a sample quantile regression can be used to mitigate the estimation bias of educational investment return caused by individual labor remuneration. To facilitate comparison, this study adopts the benchmark model and uses the same instrumental variables to mitigate the impact of endogeneity. On this basis, the investment return difference between secondary vocational education and identity of children from rural families at each percentile is investigated.

The results show obvious differences in the return on investment of secondary vocational education among children from rural families with different income levels. As the income level improves, education investment returns show an accelerated decline. Providing instructions to low-income families, especially children of poor families in rural areas, regarding investment in secondary vocational education is an important way to enhance the level of income. This is also an effective way to narrow the gap between the rich and poor to realize common prosperity. For the children of rural families who have a high income level, vocational education does little to improve their income. As the identity of the children from rural families strengthens, the future income level receives a boost, which indicates that identity plays an important role in the process of raising the income of rural families. Especially in this new era of comprehensive poverty alleviation, the impact of identity on the return on investment of children from rural families cannot be ignored.

Because the traditional quantile regression estimation ignores the direct impact of identity on educational choice and neglects the indirect effect of identity on educational investment return, its estimation results of educational investment return are usually downward biased, which is also verified in the results of Supplementary Figure S1. In conclusion, Heckman's two-step method has certain comparative advantages in estimating the investment returns of the secondary vocational education of children from rural families, and its estimation results are robust and reliable.

## 6. Discussion and Conclusions

As China strives to become a great modern socialist country, investment in secondary vocational education plays a decisive role not only in personal income but also in national development. Vigorously developing secondary vocational education is critical for training artisans in major countries, to narrow the income gap and realize common prosperity. By encouraging investment in secondary vocational education, children from rural families can not only continue to

improve the level of vocational technology but also promote income growth, assist rural children to rise above absolute poverty, prevent intergenerational continuation of poverty, help improve the level of human capital, promote social transformation, and upgrade social and economic structures, thereby successfully overcoming the middle-income trap.

From the perspective of identity, it is helpful to understand the mechanism of farmers' secondary vocational education choice and investment return. Based on identity theory, this study constructs a systematic analysis framework of identity, secondary vocational education choice, and investment return, and creates a brand new identity index system, which provides a novel perspective for the related literature on traditional human capital. Using the latest CGSS data (2012–2017), facilitated by the instrumental variable method and the intervention effect model of endogenous problems and sample selection bias, this study's findings confirm that identity has a significant promoting effect on secondary vocational education for children from rural families and not only directly and significantly impacts the return on investment but also indirectly affects the investment returns that these children may obtain in the labor market, and finally affects the speed of rural human capital accumulation.

Through analyses of the influence of identity, secondary vocational education choice, and investment return, this study finds the following:

- (1) Identity can significantly enhance the motivation of farmers to invest in secondary vocational education, which confirms the promoting effect of identity on farmers' secondary vocational education choices;
- (2) Identity can not only have a direct and significant promoting effect on the return on investment by influencing the endogenous development motivation of farmers but also has an indirect and significant impact on the return on education investment by influencing the educational choices of farmers;
- (3) The results of the instrumental variable model analysis show that as an important part of identity, class identity can significantly improve the choice of secondary vocational education and the return on investment of farmers. By contrast, class stratification restrains farmers' willingness to choose secondary vocational education and reduces the return on education investment; and
- (4) In addition to weakening class differentiation and strengthening class identity, expanding the enrollment quota of secondary vocational education and increasing the enrollment proportion of secondary vocational education can also enhance the identity of farmers. This is because the expansion of secondary vocational education enhances the sense of identity of children from rural families who invest in secondary vocational education by establishing norms of behavior that encourage investment in secondary vocational education, and indirectly enhances the sense of identity of those who benefit from the expansion policy of secondary vocational education through the fair effect of the policy, thereby finally achieving the policy goal of promoting rural family children to invest in secondary vocational education.

In view of the aforementioned conclusions, the following policy recommendations are provided:

- (1) Enhancing the identity of farmers can help them invest in secondary vocational education, obtain higher education returns, improve the overall human capital level of the society, narrow the income gap, and achieve common prosperity. Therefore, this goal deserves attention from society, especially from policymakers;
- (2) It is necessary to pay special attention to the direct and indirect impact of identity on investment returns, which is related to the income effect brought by identity—an important path to consolidate and expand the achievement of poverty alleviation and overcome the middle-income trap;
- (3) Focusing on the influence of class identification and class differentiation in identity is important. With the increasing differentiation and solidification of social classes in China, the income gap and wealth gap of different classes are gradually increasing, which must be addressed. It is suggested to formulate and introduce relevant policies to strengthen the protection of the rights and interests of low-income groups, narrow class differences, break class solidification, accelerate class integration, and achieve common prosperity; and
- (4) Other important factors affecting identity, higher education choice, and investment return should not be ignored, especially the influence of the expansion policy of secondary vocational education. Compared with developed countries, the enrollment ratio of secondary vocational education in China is still at a low level. It is suggested that we continue to expand enrollment targets for secondary vocational education, optimize the distribution of enrollment targets for secondary vocational education, and give more weight to low-income groups and poor areas, to achieve a balanced economic and social development.

In view of these conclusions, this paper recommends the following suggestions:

In theory, we should strengthen the research on the measurement of identity and form a scientific and unified index system as soon as possible. We recommend including more measures and expanding the measurement dimensions, for example by adding indicators related to social exclusion and broadening the dimensions of social inclusion. We also should strengthen the research on the economic and social effects caused by identity, such as the impact of identity on consumption

and welfare effect. Finally, we should further research human capital investment motivation, for example by conducting a comprehensive study of identity and non-cognitive ability on the impact of human capital investment motivation.

In practice, individuals should strengthen their class identity, so as to enhance their self-identity and social identity. Families should create an encouraging environment for identification; in particular, parents should actively improve their cultural literacy and educational levels. Policy makers should strengthen the publicity of social identity, provide more learning and employment opportunities, and break the rigid class categorizations. Excessive stratification should be avoided.

Although we have revealed the influence of identity on secondary vocational education choice and return on investment through theoretical and empirical analysis, this study still has some limitations. First, this study primarily focused on the impact of individual identity, but it failed to consider the efforts made by the government in regard to identity and secondary vocational education, such as social identity construction, ethnic culture publicity, and vocational education and training, all of which can directly affect identity, education choice, and return on investment. Future research could explore the effects and mechanisms of the Chinese government's social identity and vocational education policies, especially in how they pertain to identity, educational choice, and return on investment. Second, due to data limitations, we failed to discover the relationship and mechanism between identity, vocational education choice, and return on investment in other countries. If more comprehensive data are available, future studies could focus on the impact of national identity and secondary vocational education choices on return on investment.

### Supplementary Materials

The supporting information can be found at: <https://www.sciepublish.com/article/pii/158>.

### Acknowledgments

We would like to thank Editage ([www.editage.cn](http://www.editage.cn)) for English language editing.

### Author Contributions

Conceptualization, K.Z., Q.C.; data curation, C.L.; formal analysis, Q.C.; funding acquisition, K.Z.; methodology, Q.C. and Z.Y.; project administration, Q.C.; software, Q.C. and C.L.; validation, Q.C. and Z.Y.; writing—original draft preparation, Q.C., Z.Y., and C.L.; writing—review and editing, K.Z. and Q.C.; All authors have read and agreed to the published version of the manuscript.

### Ethics Statement

Not applicable.

### Informed Consent Statement

Not applicable.

### Funding

This research was funded by the National Social Science Foundation of China (Grant number 19ZDA151) and the National Statistical Research Project (Grant number 2020LY065).

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Data Availability Statement

Publicly available datasets were analyzed in this study. This data can be found here: <http://cnsda.ruc.edu.cn/>.

### References

1. Jing T, Chung E, Gregory ML. Vocational education in China: Its history, roles, challenges and the way forward. *J. Cogn. Sci. Hum. Dev.* **2022**, *8*, 112–121.
2. Christiaens AH, Nelemans SA, De Moor EL, Erentaitė R, Vosylis R, Branje S. Psychometric qualities of the educational identity processes scale (EIPS). *Front. Psychol.* **2022**, *13*, 861220.
3. Akerlof GA, Kranton RE. Economics and identity. *Q. J. Econ.* **2000**, *115*, 715–753.
4. Akerlof GA, Kranton RE. Identity and schooling: Some lessons for the economics of education. *J. Econ. Lit.* **2002**, *40*, 1167–1201.

5. Ye B, Zhao Y. Women hold up half the sky? Gender identity and the wife's labor market performance in China. *China Econ. Rev.* **2018**, *47*, 116–141.
6. Brown R. Social identity theory: past achievements, current problems and future challenges. *Eur. J. Soc. Psychol.* **2000**, *30*, 745–778.
7. Tajfel H, Turner JC. The social identity theory of intergroup behavior. In *Political Psychology: Key Readings*; Jost, J.T., Sidanius, J., Eds; Psychology Press: London, UK, 2004; pp. 276–293.
8. Oyserman D, Destin M. Identity-based motivation: Implications for intervention. *Couns. Psychol.* **2010**, *38*, 1001–1043.
9. Luo J, Wang X. Hukou identity and trust—Evidence from a framed field experiment in China. *China Econ. Rev.* **2020**, *59*, 101383.
10. Mincer J. *Schooling, Experience, and Earnings (Human Behavior & Social Institutions No. 2)*; National Bureau of Economic Research: New York, NY, USA, 1974.
11. Dean KL, Jolly JP. Student identity, disengagement, and learning. *Acad. Manag. Learn. Educ.* **2012**, *11*, 228–243.
12. Hatano K, Hihara S, Sugimura K, Crocetti E. Direction of associations between personality traits and educational identity processes: Between-and within-person associations. *J. Adolesc.* **2022**, *94*, 763–775.
13. Heckman JJ, Stixrud J, Urzua S. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *J. Labor Econ.* **2006**, *24*, 411–482.
14. Lu Z, Song S. Rural–urban migration and wage determination: The case of Tianjin, China. *China Econ. Rev.* **2006**, *17*, 337–345.
15. Heckman J, Pinto R, Savelyev P. Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *Am. Econ. Rev.* **2013**, *103*, 2052–2086.
16. Destin M, Hanselman P, Buontempo J, Tipton E, Yeager DS. Do student mindsets differ by socioeconomic status and explain disparities in academic achievement in the United States? *AERA Open.* **2019**, *5*, 10.
17. Chan A. Images of China's social structure: The changing perspectives of Canton students. *World Polit.* **1982**, *34*, 295–323.
18. Browman AS, Destin M, Kearney MS, Levine PB. How economic inequality shapes mobility expectations and behaviour in disadvantaged youth. *Nat. Hum. Behav.* **2019**, *3*, 214–220.
19. Fleisher B, Li H, Zhao MQ. Human capital, economic growth, and regional inequality in China. *J. Dev. Econ.* **2010**, *92*, 215–231.
20. Kuhnen CM, Melzer BT. Noncognitive abilities and financial delinquency: The role of self-efficacy in avoiding financial distress. *J. Finance.* **2018**, *73*, 2837–2869.
21. De Brauw A, Rozelle S. Migration and household investment in rural China. *China Econ. Rev.* **2006**, *19*, 320–335.
22. Guo D, Wang A. Is vocational education a good alternative to low-performing students in China. *Int. J. Educ. Dev.* **2020**, *75*, 102187.
23. Wang YS, Lin HH, Li CR, Lin SJ. What drives students' knowledge-withholding intention in management education? An empirical study in Taiwan. *Acad. Manag. Learn. Educ.* **2014**, *13*, 547–568.
24. Zhang JH, Yu ML, Wu FW, Chen W. Human capital of family and social mobility in rural areas—Evidence from China. *J. Integr. Agric.* **2013**, *12*, 1905–1915.
25. Oosterbeek H, Webbink D. Wage effects of an extra year of basic vocational education. *Econ. Educ. Rev.* **2007**, *26*, 408–419.
26. Ferguson HB, Bovaird S, Mueller MP. The impact of poverty on educational outcomes for children. *Paediatr. Child Health.* **2007**, *12*, 701–706.
27. Li Y, Chen J, Xie S, Huang X. Acculturation, urban identity, and psychological well-being of rural–urban migrants in China. *Appl. Psychol. Health Well-Being.* **2022**, *14*, 1129–1150.
28. Weiner S, Tatum DS. Rethinking Identity in Political Science. *Political Stud. Rev.* **2021**, *19*, 464–481.
29. Jiang C, Jiang S. Mechanisms from person–environment fit to professional identity of social workers in China: The roles of person-organization value congruence and collective psychological ownership. *J. Soc. Serv. Res.* **2022**, *48*, 535–544.
30. Poll A, Critchley C. Development and validation of the feminist social identity scale. *Curr. Psychol.* **2022**, 1–16. doi:10.1007/s12144-021-02590-x.
31. Tong C. The “New Urban Poor” identity: Evidence from China. *Am. J. Econ. Sociol.* **2021**, *80*, 977–990.
32. Duan W, Li Y, Kong Y. Construction of the index of social work professional identity. *Res. Soc. Work Pract.* **2021**, *31*, 653–661.
33. Wang P, Liu H, Wang X, Yuan X, Sun Y, Xiao L, et al. A preliminary validation of the Chinese version of the information technology identity scale. *Psychol. Res. Behav. Manag.* **2022**, *15*, 751.
34. Darity WA Jr. Position and possessions: Stratification economics and intergroup inequality. *J. Econ. Lit.* **2022**, *60*, 400–426.
35. Chelwa G, Hamilton D, Stewart J. Stratification economics: Core constructs and policy implications. *J. Econ. Lit.* **2022**, *60*, 377–399.
36. Kanzola AM, Papaioannou K, Petrakis PE. Social identity, rationality, creativity. *Int. J. Entrep. Behav. Res.* **2021**, *28*, 136–150.
37. Xiong M, Xiao L, Ye Y. Relative deprivation and prosocial tendencies in Chinese migrant children: Testing an integrated model of perceived social support and group identity. *Front. Psychol.* **2021**, *12*, 658007.
38. Liu Y. Geographical stratification and the role of the state in access to higher education in contemporary China. *Int. J. Educ. Dev.* **2015**, *44*, 108–117.
39. Titus MA. Detecting selection bias, using propensity score matching, and estimating treatment effects: An application to the private returns to a master's degree. *Res. High. Educ.* **2007**, *48*, 487–521.
40. Heckman JJ, Li X. Selection bias, comparative advantage and heterogeneous returns to education: Evidence from China in 2000. *Pac. Econ. Rev.* **2004**, *9*, 155–171.
41. Heckman JJ, Ichimura H, Todd P. Matching as an econometric evaluation estimator. *Rev. Econ. Stud.* **1998**, *65*, 261–294.
42. Arias O, Hallock KF, Sosa-Escudero W. Individual heterogeneity in the returns to schooling: Instrumental variables quantile regression using twins data. In *Economic Applications of Quantile Regression*; Physica: Heidelberg, Germany, 2002; pp. 7–40.