sustainability

MDPI

Article

Research on the Impact of Digital Empowerment on China's Human Capital Accumulation and Human Capital Gap between Urban and Rural Areas

Dapeng Sun [1,†], Bintong Yu [2,†] and Jun Ma [3,*]

[1] College of Marxism, Dalian Maritime University, Dalian 116026, China
[2] School of Agricultural Economics and Rural Development, Renmin University of China, Beijing 100872, China
[3] College of Marxism, Dongbei University of Finance and Economics, Dalian 116025, China
* Correspondence: majun81@dufe.edu.cn
† These authors contributed equally to this work.

Abstract: Human capital is a key factor in the economic growth and sustainable development of a country. However, the impact of digitalization on human capital accumulation and the urban-rural human capital gap in China is seldom a focus in the literature. This paper establishes a three-sector economic growth model surrounding knowledge accumulation, endogenous technology, and labor mobility. It is designed to theoretically explain the influence of digital empowerment on human capital accumulation and its urban-rural gap in an economy. An empirical test is also conducted using CFPS panel data for 25 Chinese provinces from 2016 to 2018. The results show that: Firstly, digital empowerment significantly promotes the accumulation of human capital in rural and urban areas. The number of years of education increases by 1.37% in rural areas and 3.07% in urban areas for each index point of digitalization increase. Secondly, the impact of digitalization on the human capital gap between urban and rural areas demonstrates an inverted U-shaped curve. However, it is still within the Matthew Effect that the human capital gap between urban and rural areas is widening in tandem with digital development. Finally, the Matthew Effect in digitalization on the urban-rural human capital gap is relatively larger in the female group, higher education population, and larger family size.

Keywords: digital empowerment; human capital; urban-rural gap; Matthew Effect; trickle-down effect

check for updates

Citation: Sun, D.; Yu, B.; Ma, J. Research on the Impact of Digital Empowerment on China's Human Capital Accumulation and Human Capital Gap between Urban and Rural Areas. *Sustainability* **2023**, *15*, 5458. https://doi.org/10.3390/su15065458

Academic Editors: Constantina Costopoulou, Sotiris Karetsos, Maria Ntaliani and Konstantinos Demestichas

Received: 8 February 2023
Revised: 10 March 2023
Accepted: 14 March 2023
Published: 20 March 2023

1. Introduction

For any country, human capital is not only an important engine of economic growth and social development, but also a key factor in comprehensive rural revitalization. With the arrival of a new era characterized by information technology and the digital economy, more and more countries and regions have emphasized the role of human capital. Human capital has gradually surpassed the status of material wealth and has become a core factor supporting the sustainable development of any economy [1]. Research results show that since China's reform and opening up, the contribution rate of human capital accumulation to the country's economic growth has reached more than 40% [2], and the comprehensive contribution rate of human capital to farmers' income growth has reached 38.57% [3]. However, China's rural human capital accumulation is still at a low level due to educational concepts, the income gap, development conditions, and other factors, with a huge human capital gap between urban and rural areas. In recent years, the education expenditure of urban and rural students has continued to converge at the stage of receiving compulsory education, but the gap between urban and rural capital education investment and development has continued to expand. In addition, the level of knowledge consumption of urban residents is more than four times that of rural residents, while the proportion of human capital for urban residents in the non-compulsory education stage is more than twice that for rural residents [4,5]. In this context, the human capital gap between urban and rural areas will inevitably further reduce the total productivity factor of an economy, thus widening the urban-rural gap between the rich and the poor, and hindering the sustainable development of the economy.

Since the 21st century, digital development has become an important mover for economic growth and social change, as well as an innovative factor and driving force that enables the development of new industries, formats, and lifestyles. The "14th Five-Year Plan" proposes that it is necessary to create new advantages in the digital economy; accelerate the construction of the digital society; and advance the construction of the digital economy, digital society, and digital government to drive the transformation of production methods, lifestyles, and governance as a whole. According to the 51th *Statistical Reports on Internet Development in China*, as of December 2022, the number of internet users in China reached 1.067 billion, and the internet penetration rate reached 75.6%. China has about one billion internet users, forming the largest and most dynamic digital society in the world. As a result, digital empowerment has brought extensive influence and profound changes to the production and lifestyle of Chinese residents. Telecommuting, intelligent production, online education, and grid governance have become typical aspects of a digital China. Further maturity in the construction of a digital society provided a key starting point in the effective prevention, control, and governance of COVID-19.

Through this background, studies have been conducted on the economic and social effects of digital empowerment in China, which can be summarized in the following four areas: 1. Research on the impact of digital empowerment on economic growth methods, industrial structure upgrading, regional innovation performance, international economic patterns, and other aspects of economic growth from a macro perspective [6–8]; 2. Research and discussion on the impact of digital development on China's urban-rural gap in income and consumption, inclusive financial development, public services, and inclusive economic growth from the welfare perspective of social development [9–13]; 3. Research on the impact of digital transformation on the input-output efficiency of enterprises, the level of division of labor in enterprises, human resource management, corporate governance structure, organizational change, and other aspects of enterprise production from an efficiency perspective [14–16]; 4. Analysis of the impact of digitalization and informatization on rural residents' entrepreneurial performance, production efficiency, land transfer, rural finance, rural governance, and other aspects of rural society from a development perspective [17–20].

At present, academic circles focus on the "digital divide" in education when studying the relationship between digitalization and the development of human capital in urban and rural areas. Research has shown that there is a significant digital divide in basic education between rural and urban areas. Primary and secondary schools in urban and rural areas differ in access to digital information equipment, digital literacy of teachers and students, etc., which has caused and widened the digital divide in basic education [21]. There is also a significant digital divide in family education between urban and rural areas. The family ICT resources obtained by students outside school have an important impact on their educational results, as well as the gap between urban and rural education. ICT resources exert a greater positive impact on the academic literacy of students who have better family backgrounds and enjoy more abundant ICT resources [22]. Online remote learning has been rapidly popularized, particularly due to COVID-19, but it has also deepened the impact of the uneven distribution of digital family education resources on learning results. The access gap, caused by the shortage of equipment, and the use gap, due to the difference in parents' education levels, exert a significant negative impact on the improvement of students' academic performance [23].

Current studies focus on the economic and social effects of digital empowerment from the perspectives of economic growth, the income gap, rural development, and enterprise efficiency; and the relationship between digitalization and human capital development is preliminarily discussed. However, three research limitations remain: First, the impact of digital empowerment on China's human capital accumulation and the human capital gap between urban and rural areas has not been examined from a macro perspective. By starting their research from the use of digital infrastructure, most of the existing studies focused on the impact of digital literacy on students' academic performance. However, they failed to respond to how the continuous improvement of the digitalization level in China affects human capital accumulation and the human capital gap between urban and rural areas, as well as what kind of curve change relationship it presents. Second, there is a lack of theoretical research on the development relationship and general laws between digitalization and human capital accumulation. Most of the existing studies on the impact of digitalization on human capital only analyze the impact of digital resources and their use on academic performance from the perspective of urban and rural education, without any theoretical analysis of the impact of digital empowerment on human capital accumulation and the human capital gap in an economy. Third, studies on the mechanism and situation for the impact of digitalization on the urban-rural human capital gap are still insufficient. The impact of digitalization on human capital accumulation will vary under different cultivation stages and population pressures. Studies on digitalization are related to the orientation and improvement of China's urban-rural development, education, and population policies, but the existing literature is lacking in these areas.

Can digital development promote the accumulation of urban-rural human capital in China? What impact will it have on the human capital gap between urban and rural areas? What are the differences in different education stages and population environments? Studies on the above issues have important theoretical value and practical significance for enhancing the quality of rural human capital, narrowing the human capital gap between urban and rural areas, improving the construction of a digital society, and promoting rural revitalization; but there are few such studies. For this reason, this paper establishes a three-sector economic growth model. It is based on the basic assumptions of new economic growth theories and centers on digital empowerment, human capital investment, and urban-rural labor mobility. It is designed to explain the impact of digital development on human capital accumulation and the human capital gap between urban and rural areas from a theoretical perspective. On this basis, the panel data of 25 Chinese provinces and regions from 2016 to 2018 were used in an empirical test, which combines macro and micro databases such as the *China Digital Index Report* and *China Family Panel Studies* (CFPS).

The contribution of this paper is as follows: First, the promotion effect of digitalization on China's urban-rural human capital accumulation was measured for the first time, providing new research evidence for the "digital dividend" brought by digital empowerment to an economy. Second, based on the basic hypothesis of an endogenous economy, a general sector equilibrium model was established to explain the relationship between digitalization and human capital accumulation. In theory, the influence of digital empowerment on human capital accumulation and the urban-rural gap in an economy was revealed and tested through empirical tests. Third, it is found that the influence of digitalization on the human capital gap between urban and rural areas has an inverted U-shaped curve relationship: in the short term, the human capital gap between urban and rural areas presents a Matthew Effect trend alongside the development of digitalization; and in the long term, digital dividends will be transferred to rural areas through the trickle-down effect, further narrowing the human capital gap between urban and rural areas. Fourth, it is found that the urban-rural gap of human capital accumulation in China is still continuously expanding with the development of digitalization, and this Matthew Effect is greater for larger family sizes and higher support pressure. This trend should be considered when formulating national education, population, and digital empowerment policies.

2. Theoretical Explanation

This paper established a regional economic growth model comprising the rural sector, urban sector, and government sector, that is based on the basic assumptions of new economic growth theories such as knowledge endogenesis, knowledge accumulation, and technology endogenesis. This was combined with the labor transfer theory and the development practice scenario in China to theoretically explain the effect and mechanism of digital empowerment on human capital accumulation and the urban-rural human capital gap in an economy.

2.1. Rural Sector

Suppose that in an economy under the continuous iteration of production-consumption cycles, there is a rural sector composed of several rural households whose utility function is $U(c_t^r)$, with the consumption level of the first period c_t^r. In a Chinese rural management situation, with the continuous development of the market economy and the cyclical nature of agricultural production, rural residents demonstrate both agricultural production and non-agricultural employment, where agricultural operation is exempt from taxation, but non-agricultural labor income is subject to certain income tax. Therefore, it is assumed that rural residents spend n_t and m_t in agricultural production and non-agricultural employment every year; the unit non-agricultural labor wage is w_t, subject to the government tax rate of π; the total working hours throughout the year are standardized to be 1, i.e., $m_t + n_t \leq 1$. In addition, under agricultural production support and protection, rural residents can receive a certain amount of government transfer payment TR_t in each period. Further, based on the basic assumption of endogenous economic growth theory, agricultural production is assumed to be a function of human capital and production technology, namely $f(\cdot) = f(A_t^r, N_t)$, where $f(\cdot)$ is the agricultural production function, A_t^r is the level of agricultural technology, and N_t is the level of human capital (In order to simplify the analysis and focus on the output contribution of human capital, elements such as land, agricultural capital, and other capital are not included in the agricultural production function herein. Of course, taking the above factors into account in the production function does not affect the derivation and conclusion of the model, as does the urban sector). Based on the knowledge endogenesis hypothesis, education is the main method of human capital accumulation. Suppose $N_t = e_t^r n_t$, with e_t^r as the education level of rural residents in period t, which mainly comes from education investment E_t^r in each period [24–26]. Based on this, according to the endogenous economic growth theory, it is generally assumed that the education level e_t^r in period t is a function of the education investment in the previous period. According to the cumulative characteristics of knowledge, this hypothesis was further developed in this paper, which is to say, the education level of rural residents in period t, e_t^r, is a function of education investment in all previous periods, i.e., $e_t^r = f(\sum_{t=1}^{n} E_t^r)$, which is more in line with actual human capital development practices.

Based on the above, it is assumed that the economy has improved its degree of digitalization through the construction of internet infrastructure, data management platforms, service platforms, and other investment means, thus further enabling industrial development and factor allocation. In view of the endogenous characteristics of digitalization, productivity can be further improved by changing production methods and developing management methods. Therefore, it is assumed that the level of agricultural production technology is a function of the degree of digital development, namely $A_t^r = A_t^r(D_t)$, where D_t is the digitalization level of the economy in period t. Suppose that the utility function of rural residents is $CRRA$ utility function $U(c_t^r) = (c_t^{r\,1-\theta} - 1)/(1-\theta)$, where θ is the risk aversion coefficient; the production function is the simplified CD production function $f(\cdot) = A_t^r(D_t) e_t^r n_t$; the return function of education investment is the concave function $e_t^r = \ln(\sum E_t^r)$ that satisfies the law of diminishing marginal returns; and the human capital level of the non-agricultural labor force is $M_t = m_t e_t^r$. At the same time, it is suitable to be characterized by index because digital transformation drives production and technological progress and has the economy of increasing marginal returns [27]. Thus, the agricultural technology function can be set as $A_t^r(D_t) = a^{D_t^2}$, and $a > 1$ is the technical facility constant. With the market price of agricultural products as P_t^r, and the discount rate of rural residents' intertemporal consumption as β, then the maximum utility of the rural sector can be expressed as:

$$\max\sum_{t=1}^{n}\beta^{t}U(c_t^r) \tag{1}$$

$$s.t.c_t^r + E_t^r \leq P_t^r f(\cdot) + (1-\pi)w_t m_t e_t^r + TR_t \tag{2}$$

2.2. Urban Sector

In contrast to the rural sector, it is assumed that there is an urban sector composed of several urban residents who produce and consume in each period. Suppose that its production function is $F(\cdot) = F(A_t^u, H_t, M_t) = A_t^u(D_t)[e_t^u h_t + e_t^r m_t]$ and its utility function is $U(c_t^u) = (c_t^{u\,1-\theta} - 1)/(1-\theta)$, where c_t^u is the consumption level of urban residents in period t; A_t^u is the level of industrial production technology; and e_t^u and h_t represent the education level and working time of urban residents, respectively. The product of the two, i.e., H_t, is the total human capital of urban residents. The urban sector absorbs both the urban labor force and labor force from the rural sector for production, and pays wages w to the rural labor force in each period, which is subject to certain income taxes. The market price of industrial products is assumed to be P_t^u, and the income tax rate is π. In order to simplify the analysis, the wage level of the rural labor force is assumed to be equal to the value of marginal products from its labor under the assumption of complete market competition [9]. Then, the function of returns to urban residents' education investment is assumed to be $e_t^u = \ln(\sum E_t^u)$, where E_t^u is the education investment of each period, meeting the characteristics of knowledge endogenesis and knowledge accumulation.

Based on the technology endogenesis characteristics of digitalization, it is also assumed that industrial production technology A_t^u is a function of the digital development level D_t of the economy and has the production characteristics of increasing marginal returns. Without loss of generality and relying on better technological infrastructure than the rural sector, urban enterprises can more rapidly apply digital technology to industrial production and enterprise management, thus fully enjoying technology-empowered dividends brought by the transformation of digital production and management in the early stage of digital development [28]. However, when the digital empowerment level of the urban sector develops to a certain stage, the rural sector is able to implement a relatively mature digital application when compared with the urban sector, and can release higher marginal productivity to improve production efficiency [9,27]. Therefore, the technical coefficient of industrial production can be assumed as $A_t^u(D_t) = (2a)^{D_t}$. Thus, the total working hours of urban residents in each period are standardized to be 1, i.e., $h_t \leq 1$, and the discount rate of urban residents' intertemporal consumption is γ.

Then the maximum utility of the urban sector can be expressed as:

$$\max\sum_{t=1}^{n}\gamma^{t}U(c_t^u) \tag{3}$$

$$s.t.c_t^u + E_t^u \leq (1-\pi)[P_t^u F(\cdot) - w_t m_t e_t^r] \tag{4}$$

2.3. Government Sector

It is assumed that the economy has a government sector to carry out macroeconomic regulation and control, mainly providing good social public services for the production, life, and social redistribution of residents. Without loss of generality, digital development is based on the construction of information infrastructure, such as the internet and blockchain, and the establishment of corresponding digital management and service platforms. Thus, it is assumed that the digital development level D_t of the regional economy mainly depends on the government investment in digital construction DI_t, i.e., $D_t = D_t(DI_t), \partial D_t/\partial DI_t > 0$. The government is also assumed to have other public service expenditures of G_t and transfer payments of TR_t to the rural sector, and its income mainly comes from the taxes of urban and rural sectors. Based on the previous assumptions, the total tax revenue of the government in period t is $\pi P_t^u F(\cdot)$. Therefore, the equation for government revenues and expenditures is:

$$DI_t + G_t + TR_t \leq \pi P_t^u F(\cdot) \tag{5}$$

When the economy achieves equilibrium, it can be combined with the constraints of the urban and rural sectors, so that the urban-rural dual economy has the following optimal budget constraints:

$$c_t^r + E_t^r + DI_t + G_t \leq P_t^r f(\cdot) + \pi P_t^u F(\cdot) + (1-\pi) w_t e_t^r m_t \tag{6}$$

$$c_t^u + E_t^u + DI_t + G_t + TR_t \leq P_t^u F(\cdot) - (1-\pi) w_t e_t^r m_t \tag{7}$$

2.4. Solution of Regional Economic Equilibrium and the Effect of Digitalization

When the economy achieves equilibrium, the following Lagrange equations were established by combining the objective function of urban and rural sectors with the optimal budget constraints:

$$\ell^r = \sum_{t=1}^{n} \beta^t U(c_t^r) + \sum_{t=1}^{n} \lambda_t^r [P_t^r f(\cdot) + \pi P_t^u F(\cdot) + (1-\pi) w_t e_t^r m_t - c_t^r - E_t^r - DI_t - G_t] \tag{8}$$

$$\ell^u = \sum_{t=1}^{n} \gamma^t U(c_t^u) + \sum_{t=1}^{n} \lambda_t^u [P_t^u F(\cdot) - (1-\pi) w_t e_t^r m_t - c_t^u - E_t^u - DI_t - G_t - TR_t] \tag{9}$$

where, ℓ_t^u and ℓ_t^r represent the Lagrange equation for maximizing the utility of urban and rural sectors, respectively, under the continuous circulation of the economic cycle and balanced operation. λ_t^u and λ_t^r represent the Lagrange multipliers of the urban and rural sectors, respectively. Based on this, the respective human capital accumulation equations for the rural and urban sectors under optimal conditions can be obtained by seeking the partial derivatives of the urban and rural sectors against their decision variables: consumption and education investment:

$$e_t^r = \ln[\frac{P_t^r A_t^r(D_t) n_t + (1-\pi) w_t m_t + \pi P_t^u A_t^u(D_t) m_t}{1 - \beta(c_t^r / c_{t+1}^r)^{\theta}}] \tag{10}$$

$$e_t^u = \ln[\frac{P_t^u A_t^u(D_t) h_t}{1 - \gamma(c_t^u / c_{t+1}^u)^{\theta}}] \tag{11}$$

The urban-rural human capital gap equation $e^{gap} = e_t^u - e_t^r$ can be further expressed as follows:

$$e_t^{gap} = \ln\{\frac{[1 - \beta(c_t^r / c_{t+1}^r)^{\theta}] P_t^u A_t^u(D_t) h_t}{[1 - \gamma(c_t^u / c_{t+1}^u)^{\theta}][P_t^r A_t^r(D_t) n_t + (1-\pi) w_t m_t + \pi P_t^u A_t^u(D_t) m_t]}\} \tag{12}$$

Then, through the partial derivation against D_t, the digital development degree of the economy, and using the urban-rural human capital accumulation equation and the urban-rural human capital gap equation, we can get:

$$\frac{\partial e_t^r}{\partial D_t} = \frac{P_t^r n_t \partial A_t^r(D_t)/\partial D_t + (1-\pi) m_t \partial A_t^u(D_t)/\partial D_t + \pi P_t^u m_t \partial A_t^u(D_t)/\partial D_t}{P_t^r n_t A_t^r(D_t) + m_t A_t^u(D_t)[(1-\pi) + \pi P_t^u]} > 0 \tag{13}$$

$$\frac{\partial e_t^u}{\partial D_t} = \frac{\partial A_t^u(D_t)/\partial D_t}{A_t^u} > 0 \tag{14}$$

$$\frac{\partial e_t^{gap}}{\partial D_t} = \frac{A_t^r(D_t) P_t^r n_t (\ln 2a - D \ln a^2)}{A_t^r(D_t) P_t^r n_t + A_t^u(D_t) m_t [(1-\pi) + \pi P_t^u]} \tag{15}$$

It can be concluded that the education levels of rural and urban residents are positively correlated with the level of digital development. In other words, digital empowerment can significantly promote the human capital accumulation of urban and rural residents. It can be further proved that:

$$\partial e_t^{gap}/\partial D_t > 0, s.t.D < \ln 2a/\ln a^2 \tag{16}$$

$$\partial e_t^{gap}/\partial D_t < 0, s.t.D > \ln 2a/\ln a^2 \tag{17}$$

$$T = \ln 2a/\ln a^2 > 0, s.t.a > 1 \tag{18}$$

It can be seen that the relationship between the urban-rural human capital gap and the level of digital development in the economy shows an inverted U-shaped curve: When the level of digital development is lower than the threshold T, the urban-rural human capital gap is positively correlated with the level of digital development. At this time, the impact of digitalization on the urban-rural human capital gap is at the "digital divide" stage; i.e., there is a Matthew Effect under which the urban-rural human capital gap will expand. However, with the continuous digital development of the economy, when the digitalization level is higher than the threshold T, the urban-rural human capital gap is negatively correlated with the level of digital development. At this time, the impact of digitalization on the urban-rural human capital gap is entering the digital dividend stage; i.e., there is a trickle-down effect under which the urban-rural human capital gap will narrow, as shown in Figure 1.

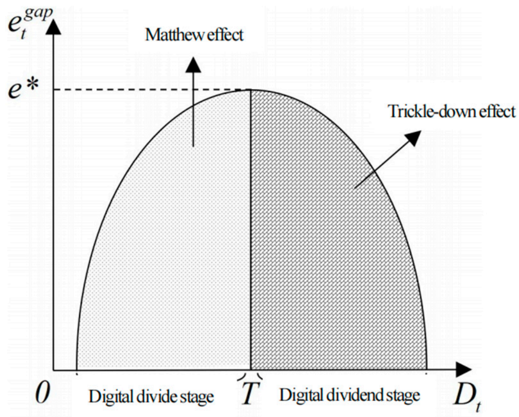

Figure 1. Digitalization and the urban-rural human capital gap (e* represents the peak of the urban-rural human capital gap at the inflection point from the expansion to convergence of the impact of digitization on the urban-rural human capital gap).

According to the derivation process of the above theoretical model and the following relevant research findings, digital empowerment can significantly promote the human capital accumulation of urban and rural residents. It is mainly subject to internal motivation as follows: First, digital development can effectively promote production progress and economic growth, and improve the income of urban and rural residents–through the development of the digital economy, digital industry, and technology empowerment–so as to further enhance the education investment of urban and rural residents in the economy and promote education level growth in the urban and rural labor force [7,29]. Second, digital development can bring greater information capital for urban and rural residents to promote the information symmetry and scale sharing of formal education and skill training resources, enrich and optimize human capital cultivation resources, accelerate the formation of a learning society, and advance the accumulation of human capital in urban and rural societies [30]. Third, educational means and concepts should be further improved through digital technology. Digital means such as network platforms and online teaching are used to accelerate knowledge dissemination and reduce the marginal cost of education. Based on digital technology, education efficiency and the human capital cultivation effect are improved. In this process, the transformation of education concepts is promoted, enabling urban and rural families to focus more on human capital investment, thus further boosting the human capital growth of urban and rural residents [31,32].

Further, the impact of digital development on the urban-rural human capital gap presents an inverted U-shaped curve, which is possibly due to the internal motivation of a significant "digital divide" between urban and rural residents in the ability to access and use digital resources. In addition, the "first-level digital divide" caused by the differences in digital infrastructure construction and the "second-level digital divide" caused by the differences in digital application capabilities, such as information awareness and information skills, will further lead to the "third-level digital divide," which is caused by the widening urban-rural gap between the rich and the poor, the education gap, and other development gaps [21,33,34]. In this context, urban residents rely on "first-hand advantages" to enjoy the welfare effect of digital development, such as improving the education effect and education level. As a result, the impact of digital development on the urban-rural human capital gap will enter the stage of an ever-expanding "digital divide," which may lead to the Matthew Effect of those who are already better becoming even better, as shown on the left side of Figure 1. Then, as the digitalization level of an economy is further developed, especially through the popularization of mobile internet technology, the ability of rural residents to acquire, manage, and process digital information will continue to improve. At the same time, rural areas can learn from the experience of urban areas in the development and application of early digital technology to better apply digital technology to increase farmer income, industrial development, and educational progress [35,36]. Thus, the role of digitalization in promoting the flow of urban and rural factors and resource sharing will become increasingly mature, which will release relatively higher marginal growth potential [9,11]. Based on this, digital development will create a trickle-down effect of digital dividends that will flow from urban residents to rural residents. Rural residents will have more development potential in exploiting digital dividends. This "later-advantage" will significantly inhibit the expansion of the urban-rural human capital gap, so that the balanced development of urban-rural education will enter the "digital dividend" stage, as shown on the right side of Figure 1.

3. Research Design

3.1. Model Parameters

In this paper, a measurement model was constructed to empirically analyze the impact of digital empowerment on China's human capital accumulation and the urban-rural human capital gap. First, the following regression equation was established:

$$Edu_{it} = \beta + \alpha_1 Dig_{it-1} + \alpha_2 UR_{it} + \alpha_3 UR_{it} \cdot Dig_{it-1} + \alpha_4 UR_{it} \cdot Dig_{it-1}^2 + \sum \gamma_i x_{it} + \sigma_i + \varepsilon_t + \mu_{it} \quad (19)$$

where, Edu_{it} is the human capital level of the respondent, measured by the number of years of education. Dig_{it-1} is the level of digitalization in the last period of the region where the respondent is located. In order to avoid the problem of simultaneous causality, the main explanatory variables were delayed for one period using the research methods of Zhang Xun et al. [12]. UR_{it} is the location of the respondent (urban = 1; rural = 0). α_1, α_2, α_3, and α_4 represent the influence coefficients: digitalization, urban/rural classification, interaction between digitalization and urban/rural classification, and interaction between digitalization^2 and urban-rural classification. x_{it} is a control variable that may affect education level; γ_i is the influence coefficient of the control variable; σ_i represents the individual effect that does not change with time; ε_t represents the time effect that does not change with individual time; β and u_{it} represent intercept term and random error, respectively; i refers to the i^{th} respondent; and t refers to period t. Then, Equation (19) can be further transformed into:

$$Edu_{it} = \beta + \alpha_1 Dig_{it-1} + (\alpha_4 Dig_{it-1}^2 + \alpha_3 Dig_{it-1} + \alpha_2) UR_{it} + \sum \gamma_i x_{it} + \sigma_i + \varepsilon_t + \mu_{it} \quad (20)$$

It can be seen that if $\alpha_1 > 0$ is significant, digitalization can effectively promote the human capital accumulation of urban and rural residents. If $\alpha_2 > 0$ is significant, there is a significant human capital gap between urban and rural areas, i.e., the human capital level of urban residents is significantly higher than that of rural residents. If $\alpha_4 < 0$ is significant, the impact of digitalization on the urban-rural human capital gap presents an inverted U-shaped curve. On this basis, if $\alpha_3 > 0$ is significant, the impact of digitalization on the urban-rural human capital gap is expanding; if $\alpha_3 < 0$ is significant, the impact of digitalization on the urban-rural human capital gap is narrowing. To control as much as possible the impact of individual characteristic factors that do not change with time and time trend factors that do not change with the individual on the human capital level of urban and rural residents, a two-way fixed effect model was used in this paper for benchmark regression estimation.

Two endogenous estimation errors may still exist in the above model. The first is reverse causality. As the quality of human capital is an important factor affecting technological innovation, the years of education of respondents in a certain region may affect the level of digital development in that region. Therefore, there may be endogenous errors caused by reverse causality. The main explanatory variables were delayed for one period in this paper. However, since the CFPS database is generally surveyed in the middle of the year, the information collected is related to the respondents in the past 12 months. The measurement of the degree of digitalization was calculated based on the year-end data of the region in the previous period. Thus, there may still be an overlap period between the main explanatory variables and the observed explained variables. This observation detail was considered in this paper, and the problem of simultaneous causality was not completely excluded. The second error is the problem of missing variables. The explained variable is the level of human capital, while the factors that affect the development of the human capital of Chinese residents are numerous and difficult to fully observe. Although the influences of the individual effect and the time effect are controlled in the model, historical factors and environmental factors may have an impact on the years of education. Therefore, endogenous problems caused by missing variables will inevitably occur. In this paper, the following two methods were used for the robustness test:

The first method is to substitute the explained variable. In this paper, the net difference between the average years of education of urban samples and the average years of education of rural samples was used as the proxy variable of the urban-rural human capital gap to establish the following measurement model:

$$e_{it}^{gap} = \beta + \lambda_1 Dig_{it-1} + \lambda_2 Dig_{it-1}^2 + \sum \gamma_i x_{it} + \mu_{it} \tag{21}$$

where, e_{it}^{gap} is the urban-rural human capital gap. λ_1 and λ_2, respectively, represent the influence coefficients of digitalization and its quadratic term on the urban-rural human capital gap. If $\lambda_2 < 0$ is significant, the impact of digitalization on the urban-rural human capital gap presents an inverted U-shaped curve. On this basis, if $\lambda_1 > 0$ is significant, the urban-rural human capital gap is expanding with the development of digitalization; or conversely, it is narrowing.

The second is an instrumental variable method. In order to eliminate the endogenous estimation bias caused by missing variables and the simultaneous causality from the measurement method level, an appropriate instrumental variable was selected to "substitute" digitalization, the explanatory variable, for regression in this paper; based on the fact that it is highly related to the explanatory variables, not directly related to the dependent variables, and not significantly related to other explanatory variables. On this basis, using the research method of Yu Bintong et al. [37], the two-stage least squares (2SLS) and the limited information maximum likelihood (LIML) methods were used in the regression of instrumental variables to ensure the robustness and credibility of the regression results.

3.2. Variable Design

1. Explained variables. Based on the existing literature, the years of education were used to measure the human capital level of urban and rural residents in this paper [26,38]. Specifically, the years of education of all urban and rural samples were used as an explained variable in this paper for benchmark regression analysis. The highest degree completed by the respondents at the time of the survey was converted into the corresponding years of education to measure their human capital level through continuous variables. Then, the urban-rural human capital gap was used as an explained variable for robustness analysis. The urban-rural human capital gap is the actual value of the average years of education of urban samples minus the average years of education of rural samples, so as to obtain a continuous variable for the urban-rural human capital gap. The larger the net difference, the greater the urban-rural human capital gap, or the inverse of this is true.
2. Main explanatory variables. In this paper, the level of digital development in China's provinces and regions was taken as the main explanatory variable. Using the theory and index system of Fan Hejun et al. [39] to measure China's regional digitalization degree, the digitalization level of an economy was considered from four aspects: production, distribution, exchange, and consumption. These comprise an organic system in social production and reproduction according to Marxist political economics. Therefore, the digital empowerment level of 25 Chinese provinces (autonomous regions and municipalities directly under the Central Government) was depicted and measured using 23 secondary indexes in total, from the four dimensions of production digitalization (production), circulation digitalization (exchange), consumption digitalization (consumption), and government digitalization (distribution) (For the specific index system, see the China Digital Index Report, Fan Hejun et al., Beijing; Economy & Management Publishing House, 2020). The measurement system fully considers all links of economic operation, social exchange, and resident life and can comprehensively and objectively measure the degree of digitalization in various regions of China. As the proxy variable of digital empowerment, it has good pertinence and credibility.
3. Instrumental variables. Based on the research design of Nunn and Qian [40], Zhao Tao et al. [41], and Yuan Chun et al. [28], the interaction between the number of fixed-line telephones in China's provinces and regions in 1984, and the number of internet users in China in a period behind in the sample observation period, was taken as the instrumental variable for regional digital development level in this paper. On the one hand, the development and application of digital technology began with the public switched telephone network (PSTN). The early level of telecommunication infrastructure construction in a region will affect the subsequent digital construction due to its technical basis and technical inertia. The number of fixed-line telephones reflects the early level of telecommunication infrastructure construction in each region and can better meet the correlation principle of instrumental variables. On the other hand, 1984 is far from the sample observation period, and the number of fixed-line telephones belongs to the category of communication. As fixed-line telephones mainly provide communication services for the public, this variable cannot have a direct impact on the years of education of urban and rural residents in the sample observation period from 2016 to 2018, and it can well meet the exogenous principle of instrumental variables. The CFPS data used in this paper include two periods of panel data. Therefore, based on the existing research and the number of fixed-line telephones in each region in 1984, the interaction for the numbers of internet users in 2015 and 2017 was introduced as an instrumental variable of the panel data. The instrumental variable formed using this method is called Shift-Share. Its exogenous nature is mainly provided by Share (the number of fixed-line telephones), while Shift (the number of Chinese internet users) provides a unified time change trend that is not at the same observation level as the Share portion, nor heterogeneous with the

observation values of each sample. New endogenous factors will not be introduced, so as to form an effective instrumental variable [40–42].

4. Control variables. The existing literature was first reviewed. Then, to better observe the impact of digitalization on the accumulation of human capital and the gap between urban and rural areas—while also trying to control factors that affect the explained variables—individual characteristics of respondents, their parents' education levels, and family environment factors were introduced into the regression equation as control variables. These included the respondents' sex, age, marital status, mother's education level, father's education level, family size, family economic status, etc. [43–45] Based on this, the individual effect and time effect were further controlled using a fixed effect model to minimize the estimation bias caused by missing variables.

3.3. Data Source and Descriptive Statistics

The data used in this study are from the 2016–2018 China Family Panel Studies (CFPS), China Digital Index Report, China Statistical Yearbook (1985), and Statistical Reports on Internet Development in China. The reasons for selecting the above four data sources: firstly, the main explanatory variables used in this study are from the China Digital Index Report; secondly, the explained variables and control variables used in this study mainly come from the CFPS database; thirdly, in this study, instrumental variables are constructed on the basis of two data sources—the China Statistical Yearbook (1985) and Statistical Reports on Internet Development in China. Based on the above data sources, the following data matching and processing work was conducted: According to the research method used by Zhang Xun et al. [12], the degree of digital development in China's provinces and regions in 2015 and 2017 was matched with the relevant data, such as the years of education of residents, in the CFPS for the years 2016 and 2018 to accurately estimate the impact of digital empowerment on human capital growth. Then, CFPS samples who were born from 1988–2002 were selected as the research objects for empirical analysis. This was based on the idea that as residents are generally continuously educated, the number of years of formal education is less likely to increase after the age of 30. At the same time, samples born after 2002 were not included because those samples received compulsory education in the earliest observation period of 2016. The increase in the number of years of education during this period mainly depends on the national compulsory education policy. The impact of this policy was eliminated in this paper. Based on this, samples born from 1988–2002 were selected. This period covers the complete human capital growth cycle from primary school to graduate school and can accurately define the audience range and estimate the impact of digital empowerment on the accumulation of human capital in urban and rural areas. In addition, the family questionnaire in the CFPS database was matched with the individual questionnaire to obtain the corresponding control variable system. The null values, extreme values, and non-Chinese samples were removed. After strict data matching and cleaning, 17,558 samples were obtained from 25 mainland Chinese provinces (autonomous regions and municipalities directly under the Central Government), including 9704 rural samples and 7854 urban samples.

The preliminary statistical characteristics of the data sample are shown in Figure 2. In 2016, the average number of years of education for rural samples was 9.23 years, and that of urban samples was 10.69 years, showing an urban-rural human capital gap of 1.46 years. In 2018, the average number of years of education for rural samples was 9.92 years, and that of urban samples was 11.56 years, showing an urban-rural human capital gap of 1.64 years. Compared with 2016, the years of education for urban and rural residents increased significantly in 2018, but the growth rate of urban residents was 0.87 years, which was higher than that of rural residents, 0.69 years. In contrast, the average digital development index of the 25 Chinese provinces and regions was 56.74 in 2015, which rose to 66.10 in 2017, indicating that China's digital development level has been growing rapidly in recent years. This growth rate is highly consistent with the growth scale of China's internet users, the construction of internet infrastructure, and other internet development rates published in the *Statistical Reports on Internet Development in China*. The above statistical analysis shows that the human capital gap between urban and rural areas in China has a weak expansion trend during the sample observation period, and the degree of digital development has a synchronous change trend with the years of education of rural residents, the years of education of urban residents, and the urban-rural human capital gap. With the continuous development of digitalization in China, the growth rate of the years of education of urban residents is significantly higher than that of rural residents, reflecting certain first-hand advantages of the digital dividend.

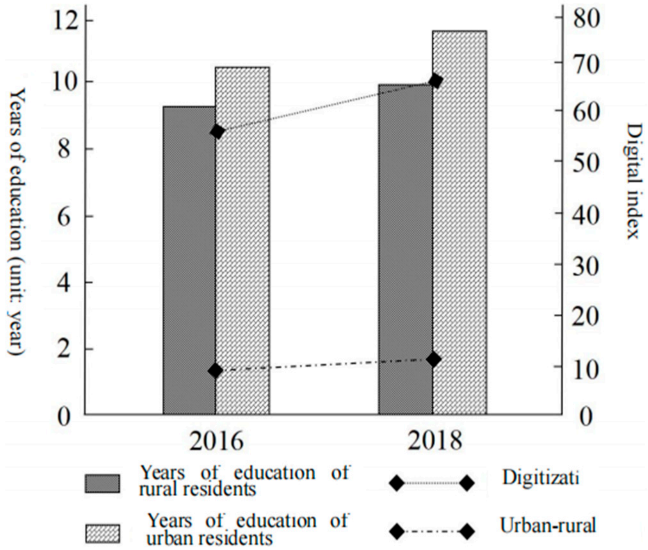

Figure 2. Digital development and the growth of years of education in urban and rural areas.

In this regard, the effect and mechanism of digital empowerment on human capital accumulation and the human capital gap between urban and rural areas were accurately measured through the rigorous measurement test in this paper. Specifically, the variable assignment and statistical characteristics of data samples are shown in Table 1.

Table 1. Variable Assignment and Statistical Characteristics.

Variables	Definition	Mean	SD
	Explained variables		
Human capital accumulation	Measured by the number of years; the highest degree obtained by the respondent is converted into the corresponding number of years: Illiterate/Semi-illiterate = 1; Primary school = 6; Junior high school = 9; High school/Technical secondary school = 12; College = 15; Bachelor = 16; Master = 19; Doctorate = 22	10.27	3.51
Urban-rural human capital gap	Net difference between the average years of education of urban residents and the average years of education of rural residents among all samples	1.44	3.50
	Main explanatory variables		
Digitalization	Digital development index, a continuous variable of 1–100	61.42	12.94
Urban/rural classification	Actual residence of the respondent: Rural = 0; Urban = 1	0.45	0.50
	Control variables		
Sex	Sex of the respondent: Female = 0; Male = 1	0.49	0.50
Age	Actual age of the respondent	23.03	4.44

Table 1. *Cont.*

Variables	Definition	Mean	SD
Marital status	Whether the respondent is married: No = 0; Yes = 1	0.36	0.48
Mother's education level	The highest degree obtained by the respondent's mother: Illiterate = 1; Primary school = 6; Junior high school = 9; High school/Technical secondary school = 12; College = 15; Bachelor = 16; Master = 19; Doctorate = 22	4.92	4.13
Father's education level	The highest degree obtained by the respondent's father: Illiterate = 1; Primary school = 6; Junior high school = 9; High school/Technical secondary school = 12; College = 15; Bachelor = 16; Master = 19; Doctorate = 22	6.26	4.23
Family income	ln value of actual income of the respondent's family in the past 12 months	10.90	1.05
Family size	Actual number of household population of the respondent's family	4.52	2.08
	Instrumental variable		
Number of fixed-line telephones in 1984	Interaction between the number of fixed-line telephones in 1984 and the number of Chinese internet users in 2015/2017 in the province where the respondent is located, as an instrumental variable of the digitalization level in 2015/2017	84.10	37.26

4. Empirical Test

4.1. Benchmark Regression Analysis

In this paper, a two-way fixed effect model was used to estimate the impact of urban-rural differences on human capital accumulation (Model 1) and the impact of digitalization on the urban-rural human capital gap (Model 2), as shown in Table 2. In order to ensure the validity of the estimation results, the Hausman test was carried out for Model 1 and Model 2. The Prob > chi2 test values of the two models were both contrary to the original hypothesis at the 1% statistical level, indicating that the fixed effect model was suitable for estimation. In addition, the statistical values of Prob > F in the two models were both significant at the 1% statistical level, indicating that they passed the overall significance test and had statistical significance.

According to the specific estimation results: First, the impact of urban-rural classification on the human capital level among all samples was estimated separately in Model 1. It is found that the urban-rural classification is positively correlated with the years of education at the 1% statistical level, indicating that there is a significant human capital gap between urban and rural areas, and the human capital level of urban residents is significantly higher than that of rural residents. Second, digitalization, the interaction between digitalization and urban/rural classification, and the interaction between digitalization^2 and urban-rural classification, were further included in this paper for estimation in Model 2. The results show that: 1. The impact of digitalization on the years of education of urban and rural residents is positively correlated at the 1% statistical level, indicating that digitalization can significantly promote human capital accumulation of urban and rural residents, which is consistent with the theoretical analysis results. 2. The interaction between digitalization^2 and urban-rural classification is negatively correlated at the 1% statistical level, indicating that the impact of digitalization on the urban-rural human capital gap presents an inverted U-shaped curve. In other words, along with digital development, the human capital gap between urban and rural areas will rise first and then fall, which further proves the theoretical point in this paper. 3. The interaction between digitalization and urban-rural classification is positively correlated at the 5% statistical level, which indicates that digitalization is exerting a greater impact on the urban-rural human capital gap, demonstrating the Matthew Effect.

Table 2. Benchmark regression analysis.

	Model 1 (Dependent Variable: Human Capital Level) (Regression of All Urban and Rural Samples)	Model 2 (Dependent Variable: Human Capital Level) (Regression of All Urban and Rural Samples)
Urban/rural classification	0.1391 (0.0640) **	0.1642 (0.0711) **
Digitalization	–	0.0157 (0.0039) ***
Digitalization * urban/rural	–	0.0059 (0.0030) **
Digitalization^2 * urban/rural	–	−0.0004 (0.0001) ***
Sex	0.4428 (0.2634) *	0.4347 (0.2631) *
Age	0.0171 (0.0599)	0.0151 (0.0598)
Marital status	−0.2233 (0.0472) ***	−0.2215 (0.0472) ***
Mother's education level	0.0258 (0.0124) **	0.0252 (0.0124) **
Father's education level	0.0022 (0.0099)	0.0032 (0.0099)
Family size	−0.0091 (0.0104)	−0.0026 (0.0106)
Family income	0.0680 (0.0146) ***	0.0662 (0.0146) ***
Time effect	0.7924 (0.1209) ***	0.6629 (0.1255) ***
Intercept term	8.4425 (1.3364) ***	7.5665 (1.3506) ***
Observed value	17,558	17,558
R^2	0.225	0.227
F value	282.80	214.63

Note: ***, **, and * represent the significance at the statistical levels of 1%, 5%, and 10%, respectively, and indicate the same in the tables below.

4.2. Robustness Test

1. Substitution of the explained variable. In this paper, the years of education of all samples used in the benchmark regression analysis were substituted by the difference in years of education. Based on the rural samples, the impact of digitalization on the urban-rural human capital gap was estimated. The results are shown in Table 3. To ensure the validity of the estimated results, the Hausman test, overall significance test, and time trend test were carried out. The Hausman test results show that the Prob > chi2 test values of the two models were both contrary to the original hypothesis at the 1% statistical level. In the time trend test, the explained variable has a time effect that does change with the individual. Then, a fixed effect model was used to determine the estimate. The statistical values of Prob > F in the two models were both significant at the 1% statistical level, indicating that they passed the overall significance test and the estimated results had statistical significance.

The estimated results show that: First, in Model 3, the impact of digitalization on the urban-rural human capital gap is positively correlated at the 10% statistical level, indicating that digital development can promote the urban-rural human capital gap, which means that it is currently in the Matthew Effect stage and the urban-rural human capital gap is expanding. Second, in Model 4, the quadratic term of digitalization was further included. It is found that the impact of digitalization on the urban-rural human capital gap is still positively correlated at the 5% significance level, and the impact of digitalization^2 on the urban-rural human capital gap is also negatively correlated at the 5% significance level, which indicates that the impact of digitalization on the urban-rural human capital gap shows an inverted U-shaped curve. In other words, the urban-rural human capital gap is expanding with digital development. The above estimated results are consistent with the results of benchmark regression analysis.

Table 3. Estimation of substituting explained variables.

	Model 3 (Dependent Variable: Urban-Rural Human Capital Gap) (Regression of Rural Samples)	Model 4 (Dependent Variable: Urban-Rural Human Capital Gap) (Regression of Rural Samples)
Digitalization	0.0126 (0.0072) *	0.0147 (0.0073) **
Digitalization^2	–	−0.0002 (0.0001) **
Sex	−0.1289 (0.3565)	−0.1211 (0.3564)
Age	−0.0331 (0.0355)	−0.0438 (0.0359)
Marital status	0.2198 (0.0645) ***	0.2177 (0.0645) ***
Mother's education level	−0.0514 (0.0179) ***	−0.0525 (0.0179) ***
Father's education level	0.0136 (0.0148)	0.0129 (0.0148)
Family size	−0.0281 (0.0150) *	−0.0286 (0.0150) *
Family income	−0.0858 (0.0207) ***	−0.0872 (0.0207) ***
Intercept term	2.6139 (0.5127) ***	2.7958 (0.5195) ***
Observed value	9426	9426
R^2	0.012	0.013
F value	7.09	6.81

***, ** and * represent the significance at the statistical levels of 1%, 5%, and 10%, respectively.

2. Endogenous processing. In this paper, the instrumental variable method was used for endogenous processing. First, the instrumental variable test was conducted for the benchmark regression analysis, and the 2SLS and LIML models were used for estimation. The results are shown in Table 4. The impact of digitalization on the years of education of urban and rural residents is still significant at the 10% statistical level. The interaction between digitalization and urban/rural classification, and the interaction between digitalization^2 and urban-rural classification are both positively and negatively correlated with the years of education at the 1% statistical level. The above results are consistent with the benchmark regression results, indicating that while digitalization significantly promotes the accumulation of human capital in urban and rural areas, its impact on the urban-rural human capital gap also presents an inverted U-shaped curve. The urban-rural human capital gap is expanding with digital development. The estimated regression coefficient and significance level results in the 2SLS model and LIML model are highly consistent, indicating that there is no weak instrumental variable problem and the instrumental variable test results are robust. In addition, the weak identification test for instrumental variables shows that the Kleibergen-Paap rk LM item is significant at the 1% statistical level, which is contrary to the original hypothesis: insufficient identification of instrumental variables. The Cragg-Donald Wald F item is greater than the critical value at the 10% level in the identification test of Stock-Yogo weak instrumental variables, which is contrary to the original hypothesis of weak instrumental variables. The above test results show that the selection and estimation of instrumental variables are reliable.

Further, the instrumental variable test was carried out for the regression model in which the explained variable was substituted, and the 2SLS model and LIML model were used for estimation at the same time (Table 5). In Model 7 and Model 8, the instrumental variable test was carried out for the impact of digitalization on the urban-rural human capital gap, and the impact of digitalization^2 was controlled. The results show that the impact of digitalization on the urban-rural human capital gap is still positively correlated at the 1% significance level, while digitalization^2 is negatively correlated with the urban-rural human capital gap at the 1% level, which is consistent with the results of the benchmark regression analysis and its instrumental variable test, indicating that the measurement results have good robustness. In addition, in order to test whether the inverted U-shaped relationship of the impact of digitalization on the urban-rural human capital gap is robust, the quadratic term of the instrumental variable was taken as the instrumental variable of digitalization^2 in Model 9 and Model 10, and the instrumental variable test was carried out on the impact of digitalization^2. The estimated results are highly consistent with Model 7 and Model 8. Digitalization^2 is still negatively correlated with the urban-rural human capital gap at the 1% level, which indicates that the inverted U-shaped relationship of the impact of digitalization on the urban-rural human capital gap is relatively robust. In addition, the estimated results of the 2SLS model and the LIML model in the above estimation are consistent, and the weak instrumental variable test results are contrary to the original hypothesis of "insufficient identification of instrumental variables," indicating that the estimated results of instrumental variables are reliable.

Table 4. Test of instrumental variables for benchmark regression analysis.

	Model 5 (2SLS Estimation) (Regression of All Urban and Rural Samples)	**Model 6 (LIML Estimation) (Regression of All Urban and Rural Samples)**
Urban/rural classification	1.0085 (0.0649) ***	1.0085 (0.0649) ***
Digitalization	0.0056 (0.0032) *	0.0056 (0.0032) *
Digitalization * urban/rural	0.0249 (0.0047) ***	0.0249 (0.0047) ***
Digitalization^2 * urban/rural	−0.0026 (0.0003) ***	−0.0026 (0.0003) ***
Control variable	Controlled	Controlled
Time effect	Controlled	Controlled
Intercept term	−2.1425 (0.2922) ***	−2.1425 (0.2922) ***
Observed value	17,558	17,558
R^2	0.2645	0.2645
Wald test value	6293.15	6293.15

Note: In order to conserve space, the estimated results of control variables have been omitted, and the same is done in the tables below. *** and * represent the significance at the statistical levels of 1% and 10%, respectively.

Table 5. Test of instrumental variables substituting explained variables.

	Model 7 (2SLS Estimation) (Regression of Rural Samples)	**Model 8 (LIML Estimation) (Regression of Rural Samples)**	**Model 9 (2SLS Estimation) (Regression of Rural Samples)**	**Model 10 (LIML Estimation) (Regression of Rural Samples)**
Digitalization	0.0825 (0.0073) ***	0.0825 (0.0073) ***	0.0322 (0.0093) ***	0.0322 (0.0093) ***
Digitalization^2	−0.0036 (0.0003) ***	−0.0036 (0.0003) ***	−0.0015 (0.0006) ***	−0.0015 (0.0006) ***
Control variable	Controlled	Controlled	Controlled	Controlled
Time effect	Controlled	Controlled	Controlled	Controlled
Intercept term	7.6525 (0.4935) ***	7.6525 (0.4935) ***	9.8281 (0.5626) ***	9.8281 (0.5626) ***
Observed value	9426	9426	9426	9426
R^2	0.1636	0.1636	0.1678	0.1678
Wald test value	1932.48	1932.48	1788.72	1788.72

*** represent the significance at the statistical levels of 1%, respectively.

In summary, the robustness test based on substituting explained variables and instrumental variables shows that: First, digitalization can significantly promote the human capital accumulation of urban and rural residents. Second, the impact of digitalization on the urban-rural human capital gap presents an inverted U-shaped curve. Third, the impact of digitalization on the urban-rural human capital gap is currently expanding. For the above three measurement conclusions, the estimation results are robust.

3. Estimation by sample. The above test results show that the impact of digitalization on the urban-rural human capital gap currently presents a Matthew Effect trend, i.e., the urban-rural human capital gap is expanding with the development of digitalization. Based on this, the robustness of this conclusion was tested by grouping the estimation for urban and rural samples. The estimated results are shown in Table 6. First, in Model 11, the impact of digitalization on the years of education of rural residents is significant at the level of 10%, with an influence coefficient of about 1.37%, which indicates that the years of education of rural residents will increase by 1.37% for each index point of digitalization. In contrast, in Model 12, the impact of digitalization on the years of education of urban residents is significant at the level of 1%, with an influence coefficient of 3.07%, which indicates that the years of education of urban residents will increase by 3.07% for each index point of digitalization. The promotion effect of digital empowerment on human capital accumulation in urban areas is more than twice that of rural areas, and the significance level is relatively higher. It can be seen that the digital dividend enjoyed by urban residents in the current process of digital development is far higher than that enjoyed by rural residents, further widening the urban-rural human capital gap. Therefore, the impact of digital empowerment on the urban-rural human capital gap shows the Matthew Effect at this stage, and this estimation result is robust.

Table 6. Impact of digitalization on the urban-rural human capital gap.

	Model 11 (Dependent Variable: Years of Education of Rural Residents) (Regression of Rural Samples)	**Model 12 (Dependent Variable: Years of Education of Urban Residents) (Regression of Urban Samples)**
Digitalization	0.0137 (0.0078) *	0.0307 (0.0069) ***
Control variable	Controlled	Controlled
Individual effect	Controlled	Controlled
Time effect	Controlled	Controlled
Intercept term	5.5027 (1.7949) ***	10.4111 (2.2056) ***
Observed value	9704	7854
R^2	0.217	0.234
F value	139.29	122.76

*** and * represent the significance at the statistical levels of 1% and 10%, respectively.

4.3. Heterogeneity Analysis

Based on the above analysis, and taking into consideration the national education policy and population policy, a heterogeneity analysis was further conducted on the impact of digitalization on the urban-rural human capital gap in different situations and groups (Table 7). The estimated results show that: First, the Matthew Effect of digitalization on the urban-rural human capital gap is more significant among women, which indicates that women are still relatively vulnerable in rural family education. In rural households with relatively scarce digital resources and educational resources, women are more likely to become the givers under the "lifeboat effect". Second, based on the school-age characteristics after compulsory education, urban and rural residents aged 14–18 and above were divided into secondary education and higher education stages, and two respective regression analyses were carried out. The results show that the expansion effect of digitalization on the urban-rural human capital gap is more significant in the stage of higher education, indicating that digitalization has expanded the human capital gap between urban and rural residents receiving higher education, namely, the accumulation of high-level human capital.

Therefore, it may be necessary to enable rural residents to fully enjoy the digital dividend from high schools to universities, i.e., the stage of higher education, so as to improve the development level of rural human capital and narrow the human capital gap between urban and rural areas.

Table 7. Heterogeneity analysis.

Dependent Variables: Years of Education among All Samples (Regression of All Urban and Rural Samples)		
	Digitalization * urban/rural	$Digitalization^2$ * urban/rural
	Sex	
Male	0.0025 (0.0044)	−0.0001 (0.0002)
Female	0.0089 (0.0041) **	−0.0007 (0.0002) ***
	Educational phase	
Secondary education	−0.0039 (0.0094)	−0.0008 (0.0005) *
Higher education	0.0095 (0.0031) ***	−0.0003 (0.0001) *
	Family members	
Three or less	0.0177 (0.0065) ***	−0.0007 (0.0003) **
Four	0.0198 (0.0103) *	−0.0002 (0.0005)
Five	0.0191 (0.0107) *	−0.0004 (0.0005)

Note: In order to conserve space, the estimated results of control variables have been omitted. ***, ** and * represent the significance at the statistical levels of 1%, 5%, and 10%, respectively.

Finally, considering that low-income families may experience the "lifeboat effect" in terms of educational resources and development resources after the enactment of China's three-child policy, the impact of digitalization on the urban-rural human capital gap under different family sizes was further estimated. This was done to look at the further expansion in opportunity inequality for the human capital development of the new generation in urban and rural areas. The results show that the expansion effect of digitalization on the urban-rural human capital gap is relatively higher in a family of four or family of five than in a family of three, which indicates that the Matthew Effect of digitalization on the urban-rural human capital gap will be further intensified with larger family sizes in urban and rural areas. Therefore, while China promulgates the population stimulation policy, it is necessary to consider the balanced distribution of education resources and development resources, to properly adjust education resources and digital dividends to rural groups, and to avoid the widening human capital gap between urban and rural areas caused by the population policy. A country should improve the quantity of human capital, enhance the quality of new human capital, and pay attention to the balanced development of human capital in urban and rural areas, so as to effectively improve total factor productivity and promote sustainable economic and social development.

5. Conclusions and Enlightenment

Based on the above research, four main conclusions are summarized in this paper. First, digital empowerment can significantly promote the human capital accumulation of urban and rural residents. Theoretical studies and empirical tests show that digitalization has actively promoted the higher education level of urban and rural residents, thus promoting the accumulation of human capital in urban and rural areas through the improvement of the income mechanism, information mechanism, and education methods and concepts. Second, the impact of digital empowerment on the human capital gap between urban and rural areas shows an inverted U-shaped curve. In the early stage of digital development, urban residents enjoy the first-hand advantage of digital dividends under the existence of the first and second-level digital divide, which further leads to the continuous expansion of the human capital gap between urban and rural areas as digital developments progress. This is a clear demonstration of the Matthew Effect. With further digital development, digital dividends will be transferred from urban areas to rural areas, further releasing the digital empowerment potential of rural residents. In this case, the human capital gap between urban and rural areas will be further narrowed through the trickle-down effect. Third, the impact of digitalization on the human capital gap between urban and rural areas is still in the Matthew Effect stage. Based on the empirical analysis of panel data from 25 Chinese provinces and regions, the promotion effect of digital empowerment on the education growth of urban residents is more than twice that of rural residents; i.e., the human capital gap between urban and rural areas is still expanding with the development of digitalization. Finally, the expansion effect of digitalization on the human capital gap between urban and rural areas reflects certain demographic characteristics and characteristics of the education stage, especially in the female group and higher education stage, and presents the features of further expansion with the growth of family size. In other words, compared with a family of three, the Matthew Effect is relatively greater in a family of four or a family of five in urban and rural areas.

Based on the above research conclusions, the following four policy implications were proposed in this paper. First, digital construction should be strengthened in underdeveloped regions, and the regional human capital development level should be improved. Research shows that digitalization has significantly promoted an increase in the number of years of education and human capital accumulation in urban and rural societies. Therefore, in terms of the construction of the internet and the digital society, a balanced and optimized path for improving the development level of regional human capital is to strengthen the cultivation capacity and accumulation potential of human capital in underdeveloped regions through mechanisms such as income, information, and resources. In the end, it is necessary to strengthen the establishment of information infrastructure and digital technology platforms in underdeveloped regions, improve the internet penetration rate, accelerate the construction of a digital society, and enhance the spillover effect of digital empowerment on human capital accumulation. The second point is to narrow the digital gap between urban and rural areas and accelerate the arrival of the turning point of the inverted U-shaped curve. The digital infrastructure construction and use gaps between urban and rural areas are the main reasons for the widening human capital gap between urban and rural areas in terms of digital development. Therefore, to release the digital dividend potential of rural society as soon as possible, the urban-rural integration strategy should be adopted to construct digital infrastructure in a timely manner. At the same time, it is necessary to strengthen the rural application of digital technology in production and life, improve digital teaching methods, and enhance the digital literacy of educators, thus accelerating the release of the technical potential of digital elements for human capital cultivation. This will also mark the turning point in the narrowing of the human capital gap between urban and rural areas. Third, with the liberalization of population policies, attention should be given to the further expansion of the human capital gap between urban and rural areas. Digital technology should be used to bridge the gap in education resources and development resources between urban and rural areas, as well as to promote the coordinated development of urban and rural human capital. The equalization of urban and rural education capacity should be improved through corresponding digital teaching and high-quality online resource sharing. A focus should be placed on appropriate policy preference for digital resources and educational resources in rural areas. Fourth, the digital use ability and digital dividend level of rural residents should be improved in the high school, higher enrollment, and higher education stages. Research shows that the expansion effect of digitalization on the human capital gap between urban and rural areas is mainly reflected in its promotion effect on urban residents to receive higher education, which is far higher than that of rural residents. For this reason, it is necessary to improve the digital use ability and digital support level in rural areas at the corresponding school-age stage, to enhance the sharing status of digital education resources and development resources, and to promote the coordinated development of high-level human capital in urban and rural areas.

Author Contributions: Conceptualization, D.S.; B.Y. and J.M.; methodology, D.S.; B.Y. and J.M.; formal analysis, D.S.; B.Y. and J.M.; data curation, B.Y.; writing—original draft preparation, B.Y.; writing—review and editing, D.S. and J.M.; supervision, D.S.; project administration, J.M.; funding acquisition, D.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The National Social Science Fund of China grant number 21BSH055.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are available from the authors upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, J.; Gao, K.; Lu, W. Human Capital Accumulation, Innovation and Economic Growth. *Dongyue Trib.* **2020**, *41*, 39–48.
2. Huang, M.; Mu, G.; Wang, X. Re-measurement of China's Talent Contribution Rate in 1978–2017. *Stat. Decis.* **2020**, *20*, 39–43.
3. Cheng, M.; Gai, Q.; Jin, Y.; Shi, Q. Human Capital Accumulation and Farmers' Income Growth. *Econ. Res. J.* **2016**, *1*, 168–181+192.
4. Guo, L.; Guo, X. Impact of the Duality of Human Capital Investment on the Income Gap between Urban and Rural Areas. *J. Tech. Econ. Manag.* **2017**, *1*, 96–101.
5. Huang, Z.; Liu, Z. Capital Accumulation, Urban-Rural Income Gap and Rural Residents' Investment in Education. *Chin. J. Popul. Sci.* **2019**, *6*, 71–83+127–128.
6. Su, J.; Su, K.; Wang, S. Does the digital economy promote industrial structural upgrading?—A test of mediating effects based on heterogeneous technological innovation. *Sustainability* **2021**, *18*, 10105. [CrossRef]
7. Xu, X.; Zhang, M. Research on the Scale Measurement of China's Digital Economy—From the Perspective of International Comparison. *China Ind. Econ.* **2020**, *5*, 23–41.
8. Song, P.; Bai, X.; Li, L. Digital empowerment, factor substitution and industrial structure transformation. *J. Shanxi Univ. Financ. Econ.* **2023**, *1*, 69–84.
9. Cheng, M.; Zhang, J. Internet Popularization and Income Gap between Urban and Rural Areas: Theory and Demonstration. *Chin. Rural Econ.* **2019**, *2*, 19–41.
10. Cheng, M.; Zhang, J. Internet Development and the Consumption Gap between Urban and Rural Residents in the New Era. *J. Quant. Tech. Econ.* **2019**, *7*, 22–41.
11. Chen, W.; Wang, Q.; Zhou, H. Digital Rural Construction and Farmers'Income Growth: Theoretical Mechanism and Micro Experience Based on Data from China. *Sustainability* **2022**, *18*, 11679. [CrossRef]
12. Zhang, X.; Wan, G.; Wu, H. Bridging the Digital Divide: Development of Digital Finance with Chinese Characteristics. *Soc. Sci. China* **2021**, *8*, 35–51+204–205.
13. Xu, Y.; Yan, Q. Obstacles of technology embedding in the construction of digital government and its resolution path. *Jianghai Acad. J.* **2022**, *6*, 151–159.
14. Liu, S.; Yan, J.; Zhang, S.; Lin, H. Can Digital Transformation of Enterprise Management Improve Input and Output Efficiency. *Manag. World* **2021**, *5*, 170–190+13.
15. Qu, Y. The organizational basis of digital innovation and China's heterogeneity. *Manag. World* **2022**, *10*, 158–174.
16. Xie, X.; Zuo, Y.; Huang, Q. Human Resource Management in the Digital Era: From the Perspective of Interaction between Human and Technology. *Manag. World* **2021**, *1*, 200–216+13.
17. Su, L.; Kong, R. Does Internet Use Promote Farmers' Entrepreneurship Gain? An Empirical Analysis Based on Endogenous Transformation Regression Model. *Chin. Rural Econ.* **2020**, *2*, 62–80.
18. Zhang, J.; Zhang, X. Research on the Impact and Mechanism of Internet Use on Farmland Transfer-out Decision—Micro-evidence from CFPS. *Chin. Rural Econ.* **2020**, *3*, 57–77.
19. Zhu, Q.; Bai, J.; Peng, C.; Zhu, C. Does Informatization Improve Agricultural Productivity? *Chin. Rural Econ.* **2019**, *4*, 22–40.
20. Liu, X.; Zhao, K. Research on the role of digital economy in promoting high-quality development of rural revitalization. *Agric. Econ.* **2023**, *1*, 42–44.
21. Zhang, H.; Wu, J.; Liu, Y.; Song, N. The 'Digital Divide' of Basic Education in Urban and Rural Areas: Representation, Causes and Elimination Policies—An Empirical Research Based on Online Teaching. *Educ. Econ.* **2021**, *4*, 20–28.
22. Zheng, L.; Zheng, Y. Digital Family Gap and Discipline Literacy Gap between Urban and Rural Students: A Study Based on PISA 2018 Data from Four Provinces and Cities in China. *China Educ. Technol.* **2021**, *7*, 43–51.
23. Liu, J. Education Equity under the Digital Divide—An Analysis Based on PISA 2018 Data from Four Provinces and Cities in China. *J. Natl. Acad. Educ. Adm.* **2020**, *9*, 35–43.

24. Lucas, R.E. On The Mechanics of Economic Development. *J. Monet. Econ.* **1988**, *1*, 3–42. [CrossRef]
25. Romer, P.M. Increasing Returns and Long-run Growth. *J. Political Econ.* **1986**, *5*, 1002–1037. [CrossRef]
26. Lucas, R.E. Human Capital and Growth. *Am. Econ. Rev.* **2015**, *5*, 85–88. [CrossRef]
27. Zhao, G. Data Elements: New Dynamics of Global Economic and Social Development. Posts & Telecom Press: Beijing, China, 2021.
28. Yuan, C.; Xiao, T.; Geng, C.; Sheng, Y. Digital Transformation and Division of Work in Enterprises: Specialization or Vertical Integration. *China Ind. Econ.* **2021**, *9*, 137–155.
29. Han, W.; Chen, H. New Production Factors of Digitization and Income Distribution. *Financ. Econ.* **2021**, *3*, 56–68.
30. Yang, Y. Educational Changes and Policy Responses Brought by 'Internet + Education. *Educ. Res.* **2016**, *6*, 64–68.
31. Jiang, F.; Wu, F. Internet Dividend and the Leap-forward Development of Human Resources in China. *Educ. Res.* **2018**, *12*, 118–126.
32. Chen, W.; Cheng, Z.; Cai, Q. Education Equity from the Perspective of Family Education Expectations—The Mitigation Effect of Digital Inclusive Finance on Non-self-induced Family Factors. *Educ. Res.* **2021**, *10*, 122–137.
33. Lv, P. Digital Village and Information Empowerment. Social Sciences in Chinese. *High. Educ. Inst.* **2020**, *2*, 69–79+158–159.
34. Wei, K.K.; Teo, H.H.; Chan, H.C.; Tan, B. Conceptualizing and Testing a Social Cognitive Model of the Digital Divide. *Inf. Syst. Res.* **2011**, *1*, 170–187. [CrossRef]
35. Scheerder, A.; Deursen, A.V.; Dijk, J.V. Determinants of Internet Skills, Uses and Outcomes. A Systematic Review of The Second-and Third-level Digital Divide. *Telemat. Inform.* **2017**, *8*, 1607–1624. [CrossRef]
36. Xue, E.; Li, J.; Li, X. Sustainable development of education in rural areas for rural revitalization in China: A comprehensive policy circle analysis. *Sustainability* **2021**, *23*, 13101. [CrossRef]
37. Yu, B.; Wang, Z.; Zhu, J.; Hou, Y. Assistance Structure, Leadership and Industrial Poverty Alleviation Performance. *China Ind. Econ.* **2021**, *6*, 23–41.
38. Li, S.; Yin, H. Talent Allocation and Economic Growth between Government and Enterprises: An Empirical Study Based on the Data of China's Prefecture-level Cities. *Econ. Res. J.* **2017**, *4*, 78–91.
39. Fan, H.; Wu, T.; Wang, S.; Du, W. *China Digital Index Report*; Economy & Management Publishing House: Beijing, China, 2020.
40. Nunn, N.; Qian, N. U.S. Food Aid and Civil Conflict. *Am. Econ. Rev.* **2014**, *6*, 1630–1666. [CrossRef]
41. Zhao, T.; Zhang, Z.; Liang, S. Digital Economy, Entrepreneurial Activity and High-quality Development: Evidence from Chinese Cities. *Manag. World* **2020**, *110*, 65–76.
42. Goldsmith-Pinkham, P.; Sorkin, I.; Swift, H. Baztik Instruments: What, When, Why and How. *Am. Econ. Rev.* **2020**, *8*, 2586–2624. [CrossRef]
43. Hou, H.; Wu, Y.; Song, Y. Low-age Boarding and Human Capital Accumulation of Rural Primary School Students: Evidence from 'Cancellation and Concentration. *Chin. Rural Econ.* **2018**, *7*, 113–129.
44. Zhu, J.; Xu, L.; Wang, H. A Study of Urban-Rural Differences in the Intergenerational Transmission of Education—Verification Based on China's Comprehensive Social Survey Data. *Educ. Econ.* **2018**, *6*, 45–55.
45. Wu, Y. Unequal Educational Opportunities for Urban and Rural Residents in China and Their Evolution (1978–2008). *Soc. Sci. China* **2013**, *3*, 4–21+203.