Trends in Social Sciences and Humanities Research
ISSN: 2959-9938
DOI: https://doi.org/10.61784/tsshr3099

HARNESSING THE TRANSFORMATIVE POTENTIAL OF THE DIGITAL ECONOMY FOR HIGH-QUALITY GROWTH: EVIDENCE FROM CHINA

FengJuan Lu, ZhengTao Chen*
College of Economics, Guangxi Minzu University, Nanning 530000, Guangxi, China.
Corresponding Author: ZhengTao Chen, Email: zhengtao.Chen@outlook.com

Abstract: This study investigates the impact of the digital economy on new quality productivity in China from a machine learning perspective. Employing panel data from 30 Chinese provinces spanning 2012-2021, the research utilizes various econometric and machine learning techniques, including fixed effects models, generalized method of moments, random forest, mediation analysis, and threshold regression. The findings reveal a robust positive relationship between digital economy development and new quality productivity, with green innovation playing a crucial mediating role. The random forest model uncovers a nonlinear relationship, where the marginal contribution of the digital economy to productivity exhibits an inverted U-shaped pattern. Furthermore, the threshold regression analysis highlights the moderating effect of innovation, with the productivity-enhancing impact of the digital economy amplified at higher levels of innovation. These results underscore the transformative potential of digital technologies in driving high-quality economic growth, while emphasizing the importance of fostering green innovation and an enabling innovation ecosystem. The study offers valuable insights for policymakers, advocating for a holistic, innovation-centric approach to harnessing the digital economy as a catalyst for sustainable development.
Keywords: Digital economy; New quality productivity; Green innovation; Machine learning; Threshold effect

1 INTRODUCTION

The advent of the digital era has brought about profound changes in the global economic landscape. The accelerated integration of digital technologies, such as big data, cloud computing, artificial intelligence, and blockchain, with the real economy has become a key driving force for high-quality economic development [1]. As a new economic form, the digital economy is not only reshaping traditional industries but also fostering the emergence of new industries and business models, thereby injecting new vitality into economic growth [2]. The digital economy has become an important engine for promoting total factor productivity and cultivating new drivers of economic development [3].

In the context of the rapid development of the digital economy, the connotation of productivity has also undergone profound changes. The traditional productivity concept, which mainly focuses on the efficiency of factor inputs, can no longer fully adapt to the new requirements of high-quality economic development in the digital era [4]. The concept of new quality productivity, which emphasizes innovation-driven, green and low-carbon, and inclusive growth, has become a new benchmark for measuring a country's comprehensive competitiveness [5]. The deep integration of digital technologies with traditional industries can help improve production efficiency, upgrade product quality, and promote the green and intelligent transformation of industries, thereby contributing to the cultivation of new quality productivity [6].

However, the existing research on the relationship between the digital economy and new quality productivity still faces some limitations. On the one hand, most studies focus on the impact of a single dimension of the digital economy, such as e-commerce [7], digital finance [8], and digital inclusion [9], lacking a comprehensive evaluation of the development level of the digital economy from multiple dimensions. This makes it difficult to fully capture the overall impact of the digital economy on new quality productivity. On the other hand, the existing research mostly adopts traditional econometric methods, such as regression analysis and panel data models, which have limitations in dealing with complex nonlinear relationships and high-dimensional data [10]. Machine learning methods, such as random forest and neural networks, have unique advantages in capturing complex relationships and identifying key influencing factors [11], but there is still a lack of relevant research in the field of digital economy and new quality productivity.

To bridge these research gaps, this paper takes 30 provinces in China (excluding Tibet) from 2012-2021 as the research sample, and constructs a comprehensive evaluation index system to measure the development level of the digital economy and new quality productivity using the entropy weight method. The research adopts a variety of machine learning methods, including the double fixed effect model, generalized moment estimation, random forest model, and threshold model, to empirically examine the impact of the digital economy on new quality productivity and its boundary conditions. Specifically, the random forest model is employed to explore the nonlinear relationship between the digital economy and new quality productivity and to identify the key influencing factors. The random forest model integrates multiple decision trees through bagging and random feature selection, which can effectively improve the accuracy and robustness of the model [12]. Moreover, this paper investigates the mediating effect of green innovation and the moderating effect of innovation level on the relationship between the digital economy and new quality productivity, aiming to reveal the internal mechanism and boundary conditions of the digital economy's impact on new quality productivity.

The main contributions of this study are as follows. First, it constructs a multi-dimensional comprehensive evaluation framework for measuring the development level of the digital economy and new quality productivity, providing a new perspective for quantitative research in this field. Second, the application of machine learning methods, especially the random forest model, helps to capture the complex nonlinear relationship between the digital economy and new quality productivity and to identify the key influencing factors, deepening the understanding of the mechanism of the digital economy's impact on new quality productivity. Third, by examining the mediating effect of green innovation and the moderating effect of innovation level, this paper reveals the realization path and boundary conditions of the digital economy's impact on new quality productivity, providing valuable insights for formulating targeted policies to promote high-quality economic development in the digital era.

The remainder of this paper proceeds as follows. Section 2 presents the theoretical analysis and research hypotheses. Section 3 describes the research design, including the model specification, variable measurement, and data sources. Section 4 reports the empirical results and discusses the findings. Section 5 concludes the paper and offers policy implications.

2 RESEARCH DESIGN AND DATA SOURCES

2.1 Modeling

To systematically examine the impact of digital economy development on new quality productivity, this study employs multiple econometric models for empirical analysis, including the fixed effects model, Generalized Method of Moments (GMM), mediating effects model, random forest model, and threshold effect model. The rationale for using multiple models is twofold. First, different models have respective strengths in dealing with panel data and endogeneity issues, and combining them helps unveil the relationship between the digital economy and new quality productivity from various angles, aiming to obtain robust conclusions. Second, the relationship between the digital economy and new quality productivity is complex, and traditional linear models may not fully capture its heterogeneity and nonlinearity. Introducing machine learning models like the random forest can help uncover more valuable information.

2.1.1 Fixed effects model

Considering the potential province-level and time-level inherent differences between digital economy development and new quality productivity, traditional pooled OLS regression may be biased. Therefore, we first employ the fixed effects model to control for such unobservable heterogeneity. The model is specified as follows:

$$Np_{it} = \beta_0 + \beta_1 Dig_{it} + \sum \beta_2 control_{it} + \mu_i + \lambda_t + \varepsilon_{it} \tag{1}$$

where Np_{it} represents the new quality productivity level of province i in year t; Dig_{it} represents the digital economy development level of province i in year t; $control_{it}$ includes a series of control variables: the level of openness (open), environmental regulation (ec), industrial structure (ls), innovation level (lninv), and technological marketization level (tm); μ_i is the province fixed effect, controlling for time-invariant heterogeneity across provinces; λ_t is the time fixed effect, controlling for common shocks faced by all provinces; and ε_{it} is the random error term.

The fixed effects model introduces dummy variables μ_i and λ_t to decompose the dependent variable Np_{it} into three parts: individual differences, time differences, and random disturbances. This approach can alleviate endogeneity issues caused by omitted variables to a certain extent. However, it assumes that the omitted variables are uncorrelated with other explanatory variables, which is often hard to satisfy in practice. Moreover, the fixed effects model can hardly identify the dynamic effects of key variables and is powerless for nonlinear relationships.

2.1.2 GMM model

To further mitigate endogeneity and examine the dynamic effects of variables, this study introduces the Generalized Method of Moments (GMM) based on the fixed effects model, adopting both the system GMM and the difference GMM:

$$Np_{it} = \beta_0 + \beta_1 Np_{it-1} + \beta_2 Dig + \sum \beta_3 control_{it} + \varepsilon_{it} \tag{2}$$

In view of the limitations of short panel data and the risk of endogeneity, this paper chooses the generalized moment estimation method, the best method at this time is the generalized moment estimation, which mainly includes the systematic moment estimation and differential moment estimation, so this paper uses the systematic GMM method and differential GMM.

2.1.3 Mediating effects model

To reveal the transmission mechanism through which the digital economy affects new quality productivity, this study further examines the mediating effect of green innovation. The mediation effect testing procedure proposed by Baron and Kenny (1986) requires estimating the following equations:

$$\ln gn_{it} = \beta_0 + \beta_1 Dig_{it} + \sum \beta_2 control_{it} + \mu_i + \lambda_t + \varepsilon_{it} \tag{3}$$

$$Np_{it} = \beta_0 + \beta_1 \ln gn_{it} + \beta_2 Dig_{it} + \sum \beta_3 control_{it} + \mu_i + \lambda_t + \varepsilon_{it} \tag{4}$$

Among them, $\ln gn_{it}$ is the mediating variable, including green patent applications $\ln gt_{it}$, green utility model applications $\ln gu_{it}$ and green invention patent applications $\ln gi_{it}$. It should be noted that the mediating effect model cannot identify the causal mechanism between variables from the perspective of causal inference, and there may be omitted variable bias between the mediating variable and the dependent variable. Therefore, caution is needed when interpreting the mediating effect.

2.1.4 Random forest model

The above models are based on linear assumptions and cannot examine the potential nonlinear relationships between variables. Therefore, this study further employs the random forest model, a bagging ensemble learning method based on decision trees, to explore the nonlinear impact of the digital economy on new quality productivity. By comprehensively training and predicting sample data with multiple decision trees, this model can effectively capture and utilize the data information of each variable, accurately assess the nonlinear relationship between variables, and more closely match the complex associations in the real world. This allows us to quantitatively compare the contribution of the digital economy to new quality productivity relative to other influencing factors. The specific model is set as follows:

$$Np_{it} = \varnothing(X_{it}, controls_{it}, \mu_i, \varepsilon_{it}) \tag{5}$$

Among them, X_{it} is the core explanatory variable digital economic development level, and $\varnothing(.)$ is the nonlinear model constructed under the random forest method.

The random forest model, as a black-box function with no expression and no tree diagram output, can be used to generate a scatter plot of the bias function f(xs) of the independent variable xs versus xs with the help of the R language to visualize the marginal impact of the digital economy on the new quality productivity. In this process, f(xs), as the bias function of xs, is treated as a marginal effect of other variables $xm^{(i)}$ by controlling their actual values to remain unchanged, so that f(xs) is only related to the independent variable xs and its interaction term with other variables. To obtain the value of the bias function, we can train several classifier models in parallel on the original data, and then average the outputs of all models to obtain:

$$f_{xs}(xs) = \frac{1}{n}\sum_{i=1}^{n} f(xs, xm^{(i)}) \tag{6}$$

where $xm^{(i)}$ represents the value of variables other than xs at the ith sample point. By examining the changing trend of the bias function f(xs) with respect to xs, we can intuitively understand the marginal effect of the digital economy development level on new quality productivity and its changing characteristics.

2.1.5 Threshold effect model

According to the theoretical analysis in the previous section, the impact of digital economy development on new quality productivity may exhibit a nonlinear relationship. To test this research hypothesis, this study adopts the threshold regression model with innovation level as the threshold variable to explore the threshold effect of digital economy development on new quality productivity. The model is specified as follows:

$$Np_{it} = \beta_0 + \beta_1 \ln gn_{it} \times I(Dig_{it} \le \gamma_1) + \beta_2 \ln gn_{it} \times I(Dig_{it} > \gamma_1) + \sum \beta çontrol_{it} + \varepsilon_{it} \tag{7}$$

where Th_{it} is the threshold variable, innovation level, including three indicators: green patent applications (lngt), green utility model applications (lngu), and green invention patent applications (lngi); I(*) is the indicator function, taking the value of 1 when the condition inside the parentheses holds, and 0 otherwise; and γ is the threshold value to be estimated. Equation (7) allows the regression coefficient of digital economy development to exhibit a stepwise change under different innovation level regimes. If the threshold effect is significant, it indicates that the relationship between digital economy development and new quality productivity will undergo structural changes as the innovation level increases. We employ the bootstrap method to test the significance of the threshold effect and determine the confidence interval of the threshold estimate through grid search, aiming to comprehensively characterize the nonlinear impact of digital economy development on new quality productivity.

2.2 Selection of Variables

2.2.1 Explained variables

New quality productivity (Np), from the three aspects of workers, labor objects, and means of production, using entropy weight method to construct the new quality productivity indicator system, specific measurements as shown in Table 1.

Table 1 New Quality Productivity Indicator System

Level 1 indicators	Secondary indicators	Tertiary indicators	Indicator measurement	Indicator properties
labor force	Worker skills	labor force level	Level of education per capita	+
			Human capital structure	+
target audience	Productivity levels	Per capita output	GDP per capita	+
		Wages per capita	Average wages on board	+
		employment level	Share of employees in the three industries	+
		Entrepreneurship level	Entrepreneurial activity	+
	new industry	emerging industry	Share of emerging strategic industries	+
		high and new technology	Number of robots	+
	ecological	Green development	forest cover	+

	environment	pollution prevention and control	Share of environmental expenditures	+
			Pollutant emissions	-
			control of pollution sources	+
means of production	material information	infrastructure	Road mileage	+
			Railroad mileage	+
			Fiber Length	+
			Internet access per capita	+
		energy consumption	energy consumption	-
			Renewable energy consumption	+
	Intangible information	technological innovation	Patents per capita	+
			R&D investment	+
		Level of digitization	Digital Economy Index	+
			Enterprise digitization level	+

2.2.2 Explanatory variables

Digital economy level (Dig), using entropy weight method to establish an evaluation model, quantitatively evaluating the level of digital economy development through the four dimensions of digital carriers, digital industries, industry level and digital environment, the detailed measurements are shown in Table 2.

Table 2 Indicator System for the Level of Development of the Digital Economy

Level 1 indicators	Secondary indicators	Indicator measurement	Indicator properties
digital carrier	Informatization scale	Internet Broadband Access	+
		Internet broadband access	+
		Number of domain names	+
		Number of pages	+
		Long-haul fiber optic cable length	+
		cell phone base station	+
digital industry	electronic information industry	Total assets of the electronic information manufacturing industry	+
		Number of enterprises in the electronic information manufacturing industry	+
	Software and information technology services	Total telecommunication services	+
		Revenue from software products	+
		Number of software developers	+
		Embedded systems revenue	+
		Number of high-tech listed companies	+
industrial level	Digitization of agriculture	Value added of agriculture, forestry, livestock and fisheries	+
		Rural electricity consumption	+
	Industrial digitization	Computers per 100 persons in industrial enterprises	+
		High-tech main business income	+
		Patent situation in high-tech industries	+
	Digitization of the service sector	Share of e-commerce trading companies	+
		E-commerce sales	+
		Digital Inclusive Finance Index	+
digital environment	Intellectual capital environment	Number of general colleges and universities	+
		Expenditure on education	+
	Digital innovation environment	R&D project funding	+
		Number of R&D personnel above scale	+

2.2.3 Mediating variables

Green innovation is measured by taking the logarithm of green patent applications (lngt), green utility model patent applications (lngu), and green invention patent applications (lngi).

2.2.4 Control variables

Level of opening to the outside world (open): expressed as the ratio of the total amount of goods imported and exported to the regional GDP; environmental regulation (ec): measured by the proportion of the completed industrial pollution control to the industrial value added; industrial structure (ls): expressed as the ratio of the tertiary industry to the secondary industry; level of technological marketization (tm): expressed as the ratio of the turnover of the technological market to the regional GDP; level of innovation (lninv): the logarithm of the number of invention patent applications received; level of marketization (market): refer to the marketization index constructed by Fan Gang as the replacement variable in this paper. Logarithm of the number of invention patent application acceptance (pcs); Marketization level (market): refer to the marketization index constructed by Fan Gang as the replacement variable in this paper.

2.3 Data Sources

This paper utilizes the dynamic panel data of 30 provinces (except Tibet) in mainland China from 2012-2021 for evaluation and analysis, and the data are mainly obtained from Peking University Digital Inclusive Finance Index, China Statistical Yearbook , China Science and Technology Statistical Yearbook, China Rural Statistical Yearbook, China Industrial Statistical Yearbook, China Energy Statistical Yearbook, and the statistical yearbooks of each province, some of which are missing. data were processed by linear interpolation. Descriptive statistics of specific variables are shown in Table 3:

Table 3 Descriptive Statistics

variable name	notation	sample size	average value	(statistics) standard deviation	minimum value	maximum values
new mass productivity	Np	300	0.137	0.063	0.042	0.477
digital economy	Dig	300	2.135	0.845	0.576	4.547
Egypt's open-door policy towards the outside world	open	300	0.259	0.277	0.008	1.441
environmental regulation	ec	300	0.003	0.004	0	0.031
industrial structure	ls	300	1.283	0.711	0.549	5.297
Level of technology marketability	tm	300	0.017	0.030	0	0.175
Level of marketization	market	300	8.138	1.882	3.359	12.39
Innovation level	lninv	300	2.262	0.146	1.740	2.518

3 EMPIRICAL ANALYSIS

3.1 Benchmark Regressions and Robustness Tests

According to the regression results in Table 4, this paper employs a two-way fixed effects model to systematically examine the relationship between the digital economy and new quality productivity. Column (1) shows that, without including control variables, the level of digital economy development has a significant positive effect on new quality productivity, with a regression coefficient of 0.069, statistically significant at the 1% level. To test the robustness of this conclusion, column (2) incorporates several control variables, such as the degree of outward orientation, environmental regulation, industrial structure optimization, transformation of scientific and technological achievements, and innovation capacity. The results indicate that the positive effect of the digital economy on new quality productivity persists after adding these variables. Column (3) replaces the control variables, and the regression results remain significant. Furthermore, to address potential endogeneity issues, the estimation is re-estimated using the 2SLS method, as shown in column (4). The level of digital economy development remains significant at the 1% statistical level, further confirming that the digital economy can promote new quality productivity. Additionally, the study addresses outliers by removing the extreme values in the 1st and 99th percentiles of the dependent variable and re-estimating the model. The results remain significant at the 1% level, indicating the robustness of the main findings. The study demonstrates that, even after considering other influencing factors, addressing endogeneity, and conducting robustness tests, there is consistently a significant positive correlation between the development of the digital economy and new quality productivity.

Table 4 Benchmark Regression and Robustness Tests

	(1)	(2)	(3)	(4)	(5)
	Np	Np	Np	Np	Np
Dig	0.069^{***}	0.023^{***}	0.021^{***}	0.093^{***}	0.023^{***}
	(0.006)	(0.006)	(0.007)	(0.009)	(0.005)
open		-0.120^{***}	-0.108^{***}	-0.022	-0.120^{***}
		(0.021)	(0.024)	(0.014)	(0.020)
ec		-1.016^{*}	-0.634	-0.105	-1.048^{*}
		(0.522)	(0.583)	(0.709)	(0.548)

ls		-0.003	0.028^{***}	0.005	0.006
		(0.008)	(0.007)	(0.005)	(0.007)
tm		1.951^{***}		0.268^{**}	1.755^{***}
		(0.191)		(0.122)	(0.175)
lninv		0.143^{***}	0.016	-0.159^{***}	0.123^{***}
		(0.039)	(0.046)	(0.045)	(0.037)
market			0.017^{***}		
			(0.003)		
Control	YES	YES	YES	YES	YES
Province	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES
_cons	-0.010	-0.231^{***}	-0.090	0.294^{***}	-0.193^{**}
	(0.013)	(0.082)	(0.094)	(0.089)	(0.080)
N	300	300	300	270	300
adj. R^2	0.237	0.586	0.489	0.682	0.615

Standard errors in parentheses* $p < 0.1$,** $p < 0.05$,*** $p < 0.01$, same in the following tables

3.2 GMM Test

To further explore the relationship between the digital economy and new quality productivity, this paper employs dynamic panel system generalized method of moments (SYS-GMM) and difference generalized method of moments (DIF-GMM) estimations. These methods effectively address endogeneity issues, yielding more accurate and robust results, as presented in Table 5. The SYS-GMM results indicate that the digital economy is statistically significant at the 1% level for new quality productivity (NQP). The coefficient of the first-order lagged NQP variable is 0.488, reflecting the strong path dependence and continuity of NQP. The DIF-GMM analysis also confirms the significant driving effect of the digital economy on new quality productivity. To assess the validity of the instrumental variables, the study conducts the Arellano-Bond test and Sargan test. The p-values of AR(1) and AR(2) pass the test, indicating serial correlation in the first-order difference perturbation term but not in the second-order difference perturbation term. The Sargan test p-value exceeds 0.1, failing to reject the null hypothesis, thus confirming the validity of the selected instrumental variables. Consequently, the conclusions drawn from the dynamic panel SYS-GMM and DIF-GMM analyses are credible.

Table 5 GMM Regression Results

	System GMM	differential GMM
L.Np	0.488^{**}	0.712^{***}
	(0.213)	(0.172)
Dig	0.116^{***}	0.107^{***}
	(0.033)	(0.030)
open	-0.063	0.272^{**}
	(0.047)	(0.115)
ec	0.273	-0.841
	(0.878)	(0.579)
ls	0.007	-0.029
	(0.014)	(0.029)
tm	0.063	0.342
	(0.222)	(0.489)
lninv	-0.343^{***}	0.072
	(0.096)	(0.076)
Province	YES	YES
Year	YES	YES
_cons	0.612^{***}	
	(0.178)	
N	270	240
AR (1)	0.012	0.043
AR (2)	0.395	0.206
Hansen	0.131	0.549

3.3 Mediation Effects Test

To further illuminate the intrinsic mechanism through which the digital economy impacts new quality productivity, this study introduces green innovation as a mediating variable and empirically examines the relationships among the digital economy, green innovation, and new quality productivity using a mediation effect model. The results, presented in Table 6, reveal a significant transmission mechanism. Employing green patent applications as an indicator of green innovation, the analysis demonstrates that, when controlling for other variables, the level of digital economy development exerts a significant positive impact on the number of green patent applications. Simultaneously, the influence of green patent applications on new quality productivity is significantly positive at the 1% statistical level. Moreover, as shown in Table 7, after incorporating the mediating variable of green innovation, the direct effect coefficient of the digital economy's development level on new quality productivity is 0.035, while the indirect effect coefficient is 0.019. The mediating effect accounts for 42.6%, indicating that green innovation plays a significant and partial mediating role in the process by which the digital economy affects new quality productivity.

To ensure the robustness of these findings, this study also adopts green utility model patent applications and green invention patent applications as proxy variables for green innovation. The estimation results remain consistent with the benchmark regression, confirming that regardless of the type of green patents, the digital economy indirectly enhances new quality productivity by promoting their development. The mediating effect accounts for 45% and 29.8% for green utility model patent applications and green invention patent applications, respectively. These results provide compelling evidence that the digital economy stimulates the vitality of green innovation, which in turn promotes the green transformation of the economy and elevates new quality productivity to a higher level. underscoring the crucial intermediary role of green innovation in the process of the digital economy empowering new quality productivity. This study contributes to the existing literature by elucidating the complex pathways through which the digital economy drives sustainable economic growth and development, highlighting the importance of fostering green innovation as a key mechanism in this process.

Table 6 Mediated Effects Test

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Np	lngt	Np	lngu	Np	lngi	Np
Dig	0.060***	1.337***	0.035***	1.332***	0.033***	1.355***	0.042***
	(0.003)	(0.058)	(0.006)	(0.063)	(0.005)	(0.058)	(0.006)
lngt			0.019***				
			(0.003)				
lngu					0.020***		
					(0.003)		
lngi							0.013***
							(0.003)
Control	YES	YES	YES	YES	YES	YES	YES
Province	YES	YES	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES	YES	YES
_cons	-0.004	5.441***	-0.109***	4.897***	-0.104***	4.525***	-0.064***
	(0.010)	(0.159)	(0.020)	(0.173)	(0.017)	(0.160)	(0.018)
N	300	300	300	300	300	300	300
adj. R^2	0.701	0.815	0.730	0.778	0.740	0.823	0.714

Table 7 Bootstrap Test for Mediating Effects

variant	effect (scientific phenomenon)	efficiency value	BootSE	95% confidence interval	Percentage of intermediary effects
lngt	direct effect	0.0346	(0.0068036)	(0.0212845,0.0479543)	42.6%
	indirect effect	0.0257	(0.0046217)	(0.0166304,0.0347471)	
lngu	direct effect	0.0332	(0.0060294)	(0.021416,0.045051)	45%
	indirect effect	0.0271	(0.0042727)	(0.0187002,0.0354491)	
lngi	direct effect	0.0423	(0.0075049)	(0.027554,0.0569727)	29.8%
	indirect effect	0.0180	(0.0050853)	(0.0080778,0.0280118)	

3.4 Random Forest Model

To elucidate the nonlinear relationship between the digital economy and new quality productivity, this study employs the random forest model, a machine learning technique, to empirically analyze the underlying mechanisms. Figure 1 illustrates the biased dependence of new quality productivity on the level of digital economy development. The results reveal that as the level of digital economy development increases, its marginal contribution to new quality productivity exhibits a nonlinear characteristic, initially rising and subsequently declining. At low levels of digital economic development, the application and penetration of digital technology remain relatively limited, and the driving effect on new quality productivity is not yet fully manifest. However, as the digital economy enters an intermediate stage of development, the integration of digital technology with the real economy accelerates, fostering a two-way digital transformation of traditional industries. Consequently, the digital economy emerges as a potent engine for enhancing new quality productivity, with the two variables exhibiting a strong positive correlation.

As the digital economy progresses to higher levels of development, the low marginal cost of digital technology becomes ubiquitous across various domains of production and life. The marginal effect of releasing digital dividends diminishes, and the marginal contribution of the digital economy to new quality productivity tends to plateau. Figure 1, which depicts the marginal effect of the digital economy, corroborates these findings. The visualization illustrates that the marginal effect of digital economy development continues to climb from the low level to the medium-high level stage, amplifying its promotional effect on new quality productivity. However, when the digital economy advances to an even higher level, the marginal effect begins to decline, indicating the presence of diminishing marginal returns in the development of the digital economy.

The results presented in Figures 1 provide compelling evidence of a significant nonlinear relationship between the digital economy and new quality productivity. The impact of digital economic development on new quality productivity is characterized by distinct stages, verifying the research hypothesis posited in this paper. These findings contribute to the growing body of literature examining the complex, dynamic interplay between digitalization and economic performance, underscoring the importance of considering nonlinearities and stage-dependent effects in empirical analyses. By employing the random forest model, this study showcases the value of machine learning techniques in uncovering nuanced relationships that may be obscured by traditional linear modeling approaches, thus advancing the frontiers of scholarly understanding in this domain.

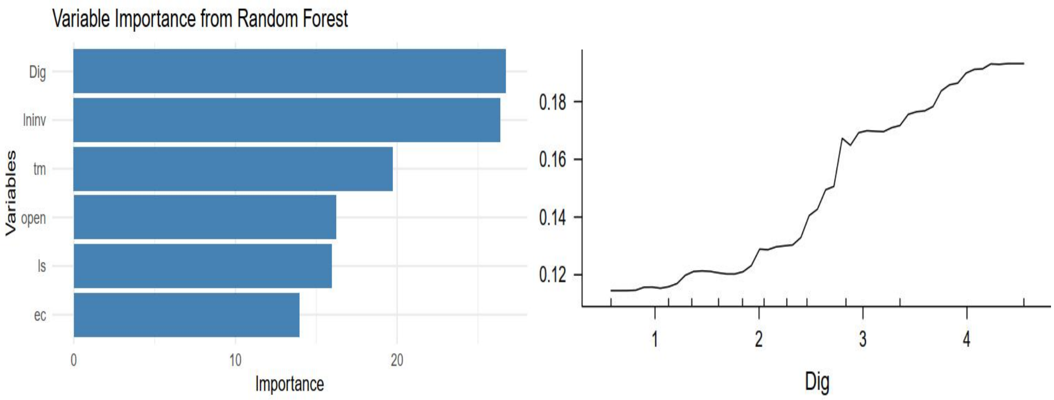

Figure 1 Bias Dependence Diagram and Marginal Effects Diagram

3.5 Threshold Effect Test

To illuminate the heterogeneous effects of the digital economy on new quality productivity across different strata of innovation, this study employs a threshold regression model, introducing innovation level as a threshold variable. The results, presented in Table 8, reveal distinct nonlinear dynamics in the nexus between the digital economy and productivity. When green patent applications serve as the proxy for innovation level, the model identifies a single threshold value, bifurcating the sample into two regimes: high and low innovation. In the low innovation regime (Th≤ q1), the coefficient estimate for the digital economy development level is 0.011, statistically significant at the 5% level. Conversely, in the high innovation regime (Th>q1), the impact of digital economy development on new quality productivity intensifies, with the coefficient rising to 0.019, significant at the 1% level.

Adopting an alternative measure of innovation, namely green utility model patent applications, uncovers two thresholds, partitioning the innovation spectrum into three intervals: low, medium, and high. As the level of innovation ascends across these intervals, the coefficient estimate for digital economy development exhibits a monotonic increase from 0.009 to 0.023, accompanied by a concomitant enhancement in statistical significance from the 10% to the 1% level. Consonant results emerge when innovation level is proxied by green invention patent applications, with the model again identifying two distinct thresholds. The coefficient on digital economy development climbs from 0.014 in the low innovation regime to 0.028 in the high innovation regime, corroborating the proposition that the productivity-enhancing effects of the digital economy are accentuated as the level of innovation escalates.

The empirical evidence garnered from the threshold regression analysis lends credence to the hypothesis of a significant threshold effect of innovation on the relationship between the digital economy and new quality productivity. As economies ascend the innovation ladder, the marginal contribution of digital economy development to productivity growth amplifies, and the magnitude of the positive association between the two variables strengthens commensurately. These findings underscore the catalytic role of innovation in unleashing the potential of the digital economy to galvanize productivity growth. The results suggest that the innovation-driven development strategy will be instrumental in stimulating the digital economy's capacity to enhance new quality productivity. Therefore, expediting the construction of an innovation-oriented economy emerges as a policy imperative of paramount significance for realizing China's high-quality development objectives in the digital era.

Table 8 Threshold Effect Model

	(1)		(2)	(3)
variant	lngt	variant	lngu	lngi
$(Th \leq q)_1$	0.011**	$(Th \leq q)_1$	0.009*	0.014***
	(0.004)		(0.005)	(0.004)
$(Th > q_1)$	0.019***	$(q_1 < Th \leq q_2)$	0.017***	0.020***
	(0.004)		(0.005)	(0.004)
		$(Th > q_2)$	0.023***	0.028***
			(0.004)	(0.004)
_cons	-0.052**		0.000	-0.042**
	(0.021)		(0.019)	(0.018)
Control	YES		YES	YES
Province	YES		YES	YES
Year	YES		YES	YES
N	300		300	300
adj. R^2	0.759		0.738	0.770

4 CONCLUSIONS AND RECOMMENDATIONS

4.1 Main Findings

Using panel data from 30 Chinese provinces from 2012 to 2021, this study employs econometric and machine learning techniques to elucidate the impact of digital economy development on new quality productivity. The results reveal a robust positive correlation between digital economy development and new quality productivity, underscoring the digital economy's pivotal role in propelling high-quality economic growth.

Mediation analysis highlights the indispensable role of green innovation in the digital economy's impact on new quality productivity. The integration of digital technologies with green innovation stimulates eco-friendly industries and green transformation of traditional sectors, paving the way for a synthesis of economic progress and ecological civilization.

The random forest model uncovers a nonlinear relationship between the digital economy and new quality productivity, with the marginal contribution of digital technologies exhibiting an inverted U-shaped pattern. This finding calls for a targeted approach to digital economy development that harnesses digital technologies' multiplier effects while expanding their application domains.

Threshold regression analysis reveals the contingent nature of the digital economy's impact on new quality productivity, dependent on the prevailing level of innovation. As an economy ascends the innovation ladder, digital technologies' productivity-enhancing effects are amplified, underscoring the imperative of fostering a vibrant innovation ecosystem alongside the digital economy.

This study contributes to the literature on the digital economy and productivity by providing a comprehensive investigation of the complex mechanisms and boundary conditions shaping this critical relationship. The findings highlight the transformative potential of digital technologies in igniting new quality productivity growth and the indispensable role of green innovation and the broader innovation milieu. The research emphasizes the need for a holistic, innovation-centric approach to harnessing the digital economy as an engine of high-quality development.

4.2 Recommendations

First,catalyzing the Digital Economy as an Engine of High-Quality Growth To harness the digital economy's potential as a catalyst for high-quality development, policymakers must prioritize a multifaceted approach integrating innovation, sustainability, and institutional reform. Accelerating digital economy development requires reconfiguring the innovation landscape by deepening strategic layout, optimizing knowledge creation and diffusion mechanisms, and amplifying investments in research and translational activities. Fostering a confluence of innovation actors, economic sectors, financial resources, and enabling policy frameworks is crucial for nurturing a vibrant innovation ecology. A commitment to frontier research in core digital technologies is vital for seizing opportunities unleashed by the scientific and technological revolution, propelling industrial upgrading and economic transformation. Underpinning these endeavors is the cultivation of specialized human capital in digital intelligence, providing the intellectual bedrock for building a robust digital economy.

Second, strengthening the Wellsprings of Innovation-Driven Productivity To fortify innovation-driven productivity growth, a holistic strategy interweaving innovation systems, research investment, and talent development is indispensable. Policymakers must implement an innovation-driven development paradigm characterized by continuous refinement of institutional mechanisms for knowledge generation, diffusion, and application. Augmenting investment in basic and applied research is crucial for expanding scientific understanding and catalyzing the translation of discoveries into tangible benefits. Facilitating seamless integration of innovation networks, industrial ecosystems, financial resources, and policy support fosters an enabling environment conducive to innovation-driven entrepreneurship. A resolute focus on surmounting technological bottlenecks and achieving strategic breakthroughs is paramount for capturing opportunities from the unfolding scientific and technological revolution. Underpinning these efforts is the systematic cultivation of a digitally-savvy, innovation-oriented talent pool with requisite skills and competencies to thrive in a knowledge-intensive economy.

Thrid,harmonizing Digital Transformation with Environmental Sustainability To achieve sustainable and inclusive digital transformation, policymakers must promote the harmonious integration of digital technologies with ecological conservation and low-carbon development. The ethos of environmental stewardship should permeate every facet of the digital economy, from technological solution design to data-driven market governance. Harnessing digital technologies' capabilities in optimizing resource allocation, enhancing energy efficiency, and minimizing environmental externalities is crucial for facilitating green upgrading of traditional industries. Policymakers should encourage the development of eco-friendly digital solutions, such as smart grids, precision agriculture, and intelligent transportation systems, which leverage data analytics and artificial intelligence to drive sustainable outcomes. Simultaneously, establishing dedicated green technology innovation ecosystems and investing in clean technologies and low-carbon infrastructure is vital for accelerating the deployment of cutting-edge solutions. By fostering a symbiotic relationship between the digital economy and ecological preservation, policymakers can chart a path towards a resilient, equitable, and sustainable future.

Fourth,deepening Institutional Reform and Global Integration To fully unlock the digital economy's potential, policymakers must commit to a comprehensive agenda of institutional reform and international cooperation. Accelerating the construction of a high-standard market system with fair competition, transparent regulation, and effective intellectual property protection is essential for fostering an enabling environment for digital innovation and entrepreneurship. Dismantling market entry barriers, combating monopolistic practices, and promoting a level playing field should be at the forefront of reform efforts. Concurrently, establishing an agile, adaptive governance framework for the digital economy is vital for ensuring the trustworthiness, security, and ethical integrity of digital technologies and data-driven business models. Policymakers should proactively engage in global dialogue and norm-setting processes to shape international governance of the digital domain, promoting principles of openness, inclusivity, and mutual benefit. Deepening cross-border collaboration in scientific research, technology transfer, and digital infrastructure development is crucial for harnessing the network effects and economies of scale inherent in the digital economy. By embracing a proactive stance towards institutional reform and international cooperation, policymakers can position their economies to thrive in an increasingly interconnected and knowledge-driven world.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Brennen, JS, Kreiss, D. Digitalization. The International Encyclopedia of Communication Theory and Philosophy, 2016, 1-11. DOI: https://doi.org/10.1002/9781118766804.wbiect111.

[2] Sturgeon, TJ. Upgrading strategies for the digital economy. Global Strategy Journal, 2021, 11(1): 34-57. DOI: https://doi.org/10.1002/gsj.1364.

[3] Goldfarb, A, Tucker, C. Digital economics. Journal of Economic Literature, 2019, 57(1): 3-43. DOI: https://doi.org/10.1257/jel.20171452.

[4] Janowski, T. Digital government evolution: From transformation to contextualization. Government Information Quarterly, 2015, 32(3): 221-236. DOI: https://doi.org/10.1016/j.giq.2015.07.001.

[5] Vial, G. Understanding digital transformation: A review and a research agenda. The Journal of Strategic Information Systems, 2019, 28(2): 118-144. DOI: https://doi.org/10.1016/j.jsis.2019.01.003.

[6] Raj, A, Dwivedi, G, Sharma, A, et al. Barriers to the adoption of industry 4.0 technologies in the manufacturing sector: An inter-country comparative perspective. International Journal of Production Economics, 2020, 224, 107546. DOI: https://doi.org/10.1016/j.ijpe.2019.107546.

[7] Fan, J, Tang, L, Zhu, W, et al. The Alibaba effect: Spatial consumption inequality and the welfare gains from e-commerce. Journal of International Economics, 2018, 114, 203-220. DOI: https://doi.org/10.1016/j.jinteco.2018.07.002.

[8] Zhu, X, Xie, L, Li, J. Digital finance, green technological innovation and eco-efficiency: Evidence from China's regional economies. Journal of Environmental Management, 2021, 296, 113199. DOI: https://doi.org/10.1016/j.jenvman.2021.113199.

[9] Kou, G, Olgu Akdeniz, Ö, Dinçer, H, et al. Fintech investments in European banks: A hybrid IT2 fuzzy multidimensional decision-making approach. Financial Innovation, 2021, 7(1): 1-28. DOI: https://doi.org/10.1186/s40854-021-00256-y.

[10] Coccia, M. A theory of the general causes of long waves: War, general purpose technologies, and economic change. Technological Forecasting and Social Change, 2018, 128, 287-295. DOI: https://doi.org/10.1016/j.techfore.2017.11.013.

[11] Athey, S. Beyond prediction: Using big data for policy problems. Science, 2017, 355(6324): 483-485. DOI;https://doi.org/10.1126/science.aal4321.

[12] Breiman, L. Random forests. Machine Learning, 2001, 45(1): 5-32. DOI: https://doi.org/10.1023/A:1010933404324.