

Further, to show more clearly the spatial-temporal variation of the main variables, we plotted Fig. 1. where the left side shows the distribution of *gtfp* from top to bottom for 2011, 2015, and 2019, while the right side corresponds to the spatial pattern of *de*. As can be seen from Fig. 1, both *de* and *gtfp* are gradually increasing over time. *gtfp* increases more significantly in coastal areas and inland central cities. As far as *de* is concerned, *de* values are usually higher in provincial capitals than in other regions.

To test Hypothesis 1, we conducted a benchmark regression, and the results are shown in Table 2. From the left to the right are the results of regressions, with the gradual addition of control variables. The first column shows the results of the regression without adding any control variables, and the coefficient is 0.406 with a t-value of 5.43, which is significant at the 1 percent level. However, this result is not credible because it suffers from the problem of important omitted variables. After we add other variables that may have an impact on green total factor productivity to the model, we find that the regression coefficients of DE remain around 0.38 and are all significant at the 1 percent level. Therefore, we consider the conclusions to be reliable. In other words, the digital economy can contribute to the growth of GTFP. Next, we also need to briefly focus on the coefficients of the control variables, where *nie* and *sav* have a negative effect on GTFP. It may be because the growth in the number of above-scale industries is mainly due to the gradual expansion of some small

and medium-sized enterprises (SMEs), which are still immature and therefore not sufficiently advanced in terms of production technology. As for savings, it may be because too much savings will squeeze the factor inputs for innovation and hinder the progress of GTFP. Overall, however, hypothesis 1 was tested.

Next, to ensure the reliability of the experimental results, we need to test the robustness of the regression equations. Therefore, we conducted panel quantile regression as well as 1% and 99% reduced-tail regression. Thus, we use a panel quantile regression model, exclude the exceptional samples, and shrink the tails by 1% and 99% for robustness testing. As shown in Table 3, the first column shows the results of the panel quantile regression at the 90 percent level, the second column shows the results of the panel quantile regression at the 75 percent level, the third column shows the results of the panel quantile regression at the 60 percent level, the fourth column shows the results of the reduced-tailed regression, and the fifth column shows the results after excluding the four special cities of Shanghai, Beijing, Tianjin, and Chongqing.

We find that all the coefficients of *de* are greater than 0 and remain significant from columns 1 to 5. This indicates that our model construction is robust. Further, we need to consider whether there are important omitted variables that cause the model to be endogenous. In this paper, an instrumental variables approach is used to address this issue. The basic requirement for an instrumental variable is that it is correlated with the

Table 2. Benchmark regression results.

	(1)	(2)	(3)	(4)	(5)
<i>de</i>	0.406***	0.402***	0.399***	0.378***	0.375***
	(5.43)	(5.39)	(5.36)	(5.08)	(5.05)
<i>fe</i>		0.473***	0.548***	0.320**	0.392***
		(3.81)	(3.95)	(2.16)	(2.61)
<i>nie</i>			-0.139	-0.322***	-0.316***
			(-1.21)	(-2.63)	(-2.60)
<i>eco</i>				0.791***	0.842***
				(4.24)	(4.49)
<i>sav</i>					-0.680***
					(-2.90)
<i>_cons</i>	0.785***	-1.320**	-0.73	0.565	5.089***
	(8.39)	(-2.36)	(-0.98)	(0.71)	(2.90)
YEAR	YES	YES	YES	YES	YES
CITY	YES	YES	YES	YES	YES
<i>N</i>	2223	2223	2223	2223	2223
adj. <i>R</i> ²	0.466	0.469	0.47	0.474	0.476

t statistics in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, the same below.

Table 5. Results of the heterogeneity test.

	(1)	(2)	(3)
<i>de</i>	0.017	0.263*	0.740***
	(0.06)	(1.83)	(4.32)
<i>fe</i>	0.033	0.261*	0.838
	(0.18)	(1.78)	(1.34)
<i>nie</i>	-0.335**	-0.230*	-0.173
	(-2.30)	(-1.67)	(-0.42)
<i>eco</i>	1.475***	0.959***	0.797
	(5.29)	(4.19)	(1.60)
<i>sav</i>	-0.088	0.200	-1.400
	(-0.38)	(0.67)	(-1.52)
<i>_cons</i>	2.669	-0.898	12.267*
	(1.62)	(-0.40)	(1.69)
YEAR	YES	YES	YES
CITY	YES	YES	YES
<i>N</i>	729	945	549
adj. R^2	0.560	0.665	0.493

levels of economic development, the result is only 0.017 and insignificant. This suggests that the digital economy contributes more significantly to green total factor productivity in regions with better economic development, which is a sign of strength. The digital economy can play a more critical role in regions with well-developed industrial systems and high economic development. Conversely, in some lagging regions, its role is limited.

At this point, the validation of Hypothesis 1 is essentially over. We need to further analyze the mechanism issue. That is, what are the mechanisms by which the digital economy affects green total factor productivity? Of course, this hypothesis has already been proposed in the previous section, and we would like to examine the role that green technological innovation may play in this process. Therefore, we ran a regression on equation 2, and the results are shown in Table 6.

Next, observe the regression results in Table 6. In order to present a clearer picture of the regression, we report the results without putting the control variables together. In addition, we also place the first and fifth columns from Table 1 here. Note that later in the hypotheses, we need to ensure the stability of the main variable (DE). If the significance of the main variable and the direction of the coefficients change significantly with the inclusion of the interaction term, then the coefficients of the interaction term are meaningless. Fortunately, the direction of the coefficients on all of the main variables remained consistent, suggesting that the addition of the interaction term did not create a new

Table 6. Regression results for moderating effects.

	(1)	(2)	(3)	(4)
<i>de</i>	0.296***	0.329***	0.375***	0.406***
	(4.13)	(4.56)	(5.05)	(5.43)
<i>de</i> × <i>gtec</i>	0.057***	0.056***		
	(12.56)	(12.50)		
<i>fe</i>	0.157		0.392***	
	(1.08)		(2.61)	
<i>nie</i>	-0.295**		-0.316***	
	(-2.52)		(-2.60)	
<i>eco</i>	0.990***		0.842***	
	(5.48)		(4.49)	
<i>sav</i>	-0.702***		-0.680***	
	(-3.11)		(-2.90)	
<i>_cons</i>	6.016***	0.824***	5.089***	0.785***
	(3.56)	(9.14)	(2.90)	(8.39)
YEAR	YES	YES	YES	YES
CITY	YES	YES	YES	YES
<i>N</i>	2223	2223	2223	2223
adj. R^2	0.515	0.505	0.476	0.466

endogenous problem. Then, we observe the coefficients and significance of the two interaction terms, and they are 0.057 and 0.056, respectively. This means that green technological innovation has a positive moderating effect on the digital economy, regardless of whether control variables are considered or not. In other words, the coefficient of DE grows with *gtec*. In that case, the second hypothesis is also verified. Finally, we want to examine whether green technological innovation acts as a threshold variable in this process. Accordingly, we perform a threshold regression.

We set up 300 Bootstrap tests, performing single-threshold, double-threshold, and triple-threshold tests in that order. The results show that the F-value for the first threshold is 28.42, the F-value for the second threshold is 68.58, and the F-value for the third threshold is 30.60. The P-value is less than 0.1, which suggests that there may be a triple threshold. However, in terms of confidence intervals, the third of the triple thresholds is missing the upper and lower thresholds. This suggests that the third threshold, the one with a threshold value of 8.104, does not exist (this conclusion is, of course, further verified in the LR diagram). Therefore, we believe that there are just two thresholds, which means that this is a two-threshold model. Next, we divided the time period of the study sample into Phase I and Phase II. The first phase covers the period 2011-2015, while the second phase covers the period 2015-2019. Note that 2015 is included in both phases. Of course, again, we

Table 9. Threshold regression results.

Threshold variable	<i>Double(2011-2019)</i>	<i>Double(2011-2015)</i>	<i>Double(2015-2019)</i>
$de (gtec \leq \gamma_1)$	0.202*** (2.82)	0.180 (0.48)	0.256*** (2.91)
$de (\gamma_1 < gtec \leq \gamma_2)$	0.420*** (8.54)	0.411*** (3.33)	0.437*** (2.96)
$de (\gamma_2 < gtec)$	0.994*** (10.09)	0.669*** (5.41)	1.020*** (6.35)
<i>fe</i>	0.670 (0.44)	0.209* (1.90)	0.240 (1.40)
<i>nie</i>	-0.892*** (-7.08)	-0.287*** (-2.62)	-0.048 (-0.24)
<i>eco</i>	0.868*** (8.69)	1.201*** (13.58)	0.336 (1.14)
<i>sav</i>	0.173** (2.09)	0.126*** (2.94)	0.336*** (5.14)
<i>_cons</i>	-7.877*** (-6.67)	-7.532*** (-6.88)	-13.877*** (-9.29)
YEAR	YES	YES	YES
CITY	YES	YES	YES
<i>N</i>	2223	1235	1235
<i>F</i>	255.94	197.83	82.20

coefficient of *de* becomes 0.994. It can be said that the difference between the three regression coefficients is still relatively large. Next, we look at the regression results for the first period. We find that the driving effect of the DE on GTFP is insignificant when *gtec* is smaller than the first threshold. The regression coefficient is 0.411 when *gtec* is between the two thresholds, and this result is 0.669 when *gtec* is larger than the second threshold. The overall regression coefficient is lower than that for the full sample period, which is certainly in line with the true picture. This period is in the early stages of the development of the digital economy the concept of green total factor productivity is not yet widely recognized and pollution control in China is still in its infancy. Finally, we observe the regression results between 2015-2019. At less than the first threshold, the regression coefficient is 0.256, after crossing the first threshold, this result comes to 0.437, and after crossing the second threshold, this result becomes 1.020, with all three regression coefficients significant. This result is similar to the full sample, which indicates that the results of the threshold regression are plausible. Note, however, that their thresholds change after distinguishing between the before and after phases (of course, you can easily spot this in Fig. 2). Here, we are looking to examine the existence of the thresholds at different time periods rather than the changes in the threshold values. Because we are using panel data, its experimental results are only responsible for China from 2011-2019. That is, if the data are available, we assume that the interval of the study extends backward from 2019 to 2022 or even 2023, where the results become different.

In fact, the reasons for the difference in thresholds are quite understandable. In different time periods, either DE, GTFP, or *gtec*, their means as well as their statistical distributions are different, and in general, the second stage is slightly higher than the first. From the regression coefficients themselves, the coefficients are relatively higher in the second stage, which suggests that the growth rate of GTFP is higher than that of DE in the second stage. Of course, this is literally true. What we are trying to convey is that the thresholds are not likely to be uniform because of the stochastic nature of the movement of the mechanism variable, the independent variable, and the dependent variable. But one thing can be proved: in any case, we are sure that *gtec* can act as a threshold variable in this interval and that the positive effect of the digital economy on GTFP is stronger when the level of *gtec* exceeds a specific value. At this point, hypothesis 3 is verified.

Discussion

In this study, we constructed the DEA-SBM-ML model to measure GTFP. Here, we integrated the non-expected outputs into the model and, unlike the static model, included the ML index. The indicator system here is also relatively more suitable for industrial use. Of course, as we mentioned earlier, the increase in green total factor productivity is not necessarily generated by advances in clean technology. Therefore, we need to fully consider the mechanism of action while examining the impact of the digital economy on green total factor productivity. In this way, we introduced green

