Digital Transformation and Risk Differentiation in the Banking Industry: Evidence from Chinese Commercial Banks

Xinran Cao
Institute of Digital Finance
National School of Development and China Center for Economic Research
Peking University
xrcao2019@nsd.pku.edu.cn

Boyu Han
School of Economics
Peking University
hanboyu1998@163.com

Yiping Huang
Institute of Digital Finance
National School of Development and China Center for Economic Research
Peking University
yhuang@nsd.pku.edu.cn

Xuanli Xie
Institute of Digital Finance
National School of Development and China Center for Economic Research
Peking University
xxl@nsd.pku.edu.cn

Abstract
This paper studies the impact of digital transformation on the ex post risk differentiation of large and small banks, measured by nonperforming loan (NPL) ratios. It uses the Digital Transformation Index of Commercial Banks compiled by the Institute of Digital Finance of Peking University, which contains data on three dimensions—cognition, organization and products—for 97 banks from 2011 to 2018. The three main findings are: (1) the digital transformation of cognition and organization only affects the NPL ratio through the digital transformation of products; (2) the digital transformation of products only increases the NPL ratio of small banks, but not large banks; and (3) the reason for the above results is that, in fulfillment of the mandatory requirement of lending to the micro-, small, and medium-sized enterprises (MSMEs), digital transformation makes it easier for large banks to discount commercial bills held by MSMEs, thereby pushing small banks to extend corporate loans to MSMEs, which have higher risks.

1. Introduction

The banking industry is undergoing digital transformation, and risk is one of the most important indicators for banking businesses. In the process of technological progress, there is sometimes a lag in the devising and implementation of relevant rules and regulations,

Asian Economic Papers 21:3
https://doi.org/10.1162/asep_a_00853

which frequently leads to the accumulation of financial risks. Digital transformation may be no exception. China is one of the world's leading countries in the development of financial technology, and it is also a country where the banking industry dominates the financial system. As lending is the core business of banks, it is of great importance to understand the impact of digital transformation on the loan-side risks in the Chinese banking industry.

The existing literature tentatively concludes that the application of digital finance can strengthen banks' ability to resist and control risks (Jin et al. 2020; Li and Yang 2020). DeYoung et al. (2004) find that due to technological progress, the advantages of small banks in commercial loans, payment services, customer financing, and customer investment have been weakened. In terms of risk control, Li and Yang (2020) find that the level of financial technology development in banks is negatively correlated with the level of risk borne by banks, and there is a U-shaped relationship between bank market power and bank risk. Zhao et al. (2021) find that for banks with a high level of digital transformation, digital transformation is negatively related to their nonperforming loan (NPL) ratio, and for banks with a low level of digital transformation, it is positively related to their NPL ratio.

The literature provides correlation analysis of the relationship between digital transformation and risk in the banking industry. However, it fails to reveal the difference in the changes in loan risks of large and small banks in the process of digital transformation, and it also fails to unveil the essential causes for the changes in banks' loan risks. Starting from the ex post risk of bank loans, we study the heterogeneous impact of the digital transformation of the banking industry on the loan-side risk of large and small banks, as well as the fundamental reasons for the difference. Thus, this study makes three new contributions to the literature.

First, this study proposes that the digital transformation of banks will lead to differentiation between the ex post risks of the loans of large and small banks. The analysis uses the Digital Transformation Index of Commercial Banks (Phase 1), constructed by the Institute of Digital Finance of Peking University (Xie and Wang 2022) with data from banks' annual reports from 2011 to 2018, to show that the NPL ratio of small banks will increase, and the NPL ratio of large banks will remain unchanged.

Second, this study uses instrumental variables for the digital transformation of banks to address the potential endogeneity problem. This approach contrasts with the existing research on digital technology and bank risk, which mainly discusses correlation. The analysis here uses the average closest distance between the local city and the cities in the province along the "eight horizontal and eight vertical" optical fiber trunk network, the number of mobile phone users, and a joint instrumental variable that combines them. Thus, the purpose of using instrumental variables is to identify causality.

Third, this paper also explores the mechanism through which digital transformation affects the ex post risks of the loans of large and small banks differently. The analysis finds that the cause of the differentiation between large and small banks' NPL ratios is that digital transformation has differentiated the loan-side structure of large and small banks.

Policies mandate all banks lending to micro-, small, and medium-sized enterprises (MSMEs) to improve small firms' funding conditions.[1] Both bill discounting and corporate loans are loans to companies, as well as to MSMEs. Banks can use bill discounting and corporate loans to fulfill the policy requirements. However, the risk implications of these two types of loans are quite different. Bills held by MSMEs are endorsed by the credit of large companies or banks.[2] Assessing the risk of bill discounting is straightforward: A bank only needs to check whether the bill is real or not. Meanwhile, bill discounting has a much lower risk and, therefore, implies a lower NPL ratio.[3]

In contrast, corporate loans to MSMEs are dependent purely on MSMEs' credit worthiness. The risk assessment of a corporate loan is much more complicated: A bank needs to perform comprehensive due diligence before making a loan.[4] The NPL ratios of corporate loans to MSMEs are relatively high.[5] In this way, as long as a bank can use bill discounting to fulfill the MSME lending target required by the authorities, it would be reluctant to issue corporate loans endorsed by MSMEs to fulfill the policy requirement.

To fulfill the policy requirements of lending to MSMEs, large banks would first choose bill discounting. Because MSMEs are scattered across the country, however, in the traditional banking environment it is not easy for large banks to conduct bill discounting businesses, given their limited branch networks. Digital transformation has eased this constraint because most of the information can now be available online. This allows large banks to

1 Since the CPC National Congress, the Central Bank has always taken serving small and micro enterprises as the top priority of its work, and financial support for small and micro enterprises has been increasing overtime. For example, the China Banking Regulatory Commission proposed a "three no lower than" policy that efforts should be made to achieve the following three goals: (1) The growth rate of the loans to small and micro enterprises should not be lower than the average growth rate of all items of loans. (2) The number of small and micro enterprise who receive loans should not be lower than that in the same period of the previous year. (3) The loan acquisition rate of small and micro enterprises applying for loans should not be lower than that of the same period last year.

2 We use an example to explain bill discounting. For example, when large company A owes company B some money, it issues a bill to company B by which company B can get cash from large company A on the maturity date of the bill. Company B needs to get this money before the maturity date, so it gives the bill to bank C and receives money from bank C, which means that bank C discount commercial bill (or bill loan) held by company B. Then on the maturity date of the bill, bank C receives cash from large company A by the bill.

3 As is shown in Figure 1.

4 Due diligence includes checking the balance sheet, spot research, and so forth.

5 Loans to MSMEs are often risky because of the lack of collateral.

serve more companies across the country, through bill discounting, given their large pools of cheap funds (since bill discounting businesses, mostly backed by credit worthiness of large corporations, have relatively low risks). Expansion of the large banks' bill discounting business does not necessarily increase their loan risks. In the meantime, the small banks are pushed to extend corporate loans to firms, as the large banks cream-skimmed better quality businesses from them. Their loan risks could increase as a result.

The rest of the paper is structured as follows. Section 2 provides a literature review and points out some directions where this study could make contributions. Section 3 introduces the data and methodology. Section 4 presents the regression results and also conducts some robustness tests. Section 5 offers some concluding remarks.

2. Literature review

2.1 Impact of digital finance on bank risks

As it is a new market force, researchers have focused on the impact of digital finance on commercial bank risks. Guo and Shen (2015) find that the development of Internet finance has an impact on the risk-taking of commercial banks. In the early stage of development, Internet-based finance helps commercial banks to reduce management costs and risk taking. However, Internet finance increases capital costs, which intensifies risk taking. Guo and Shen show that compared with non-systemically important banks, systemically important banks have responded more prudently to the development of Internet finance. Some scholars point out that Internet finance improves depositors' perception of banks' risk taking (Hou et al. 2016). Some have studied the impact of digital finance on banks' risk taking from the perspectives of the liability and asset sides. Wang (2015) shows that digital finance and commercial banks compete directly in the field of business liabilities, dislocate competition in the field of asset business, and compete in the field of intermediate business, which force banks to increase their risk appetite. Specifically, the development of digital finance has promoted the marketization of deposit interest rates in disguised form, changing the structure of bank liabilities, which made commercial banks more reliant on interindustry lending, thereby driving up financing costs (Guo and Shen 2019; Qiu et al. 2018). From the asset side, the rising cost of the liability side prompts banks to choose higher-risk assets to make up for the loss caused by the rising cost of the liability side (Qiu et al. 2018). The previous literature used the data constructed by the Peking University Digital Financial Inclusion Index and Baidu searches of words related to "Internet finance," emphasizing the use of Alipay and the impact of the Internet finance industry in a broad sense on bank risks.

2.2 Impact of the development of digital finance on bank risks

Most scholars believe that banks' use of digital finance has a positive impact on their operations (Berger 2003), which is mainly due to the banks' operational capabilities and risk control. In terms of bank operations, the application of digital finance can greatly reduce the temporal and spatial distance between banks and customers and help banks to maximize customer coverage. At the same time, digital finance can support the diversification of commercial banks' business, channels, and product innovation (Huang and Huang 2018); optimize customer experience (Xing 2016); improve operational efficiency; and increase banks' total factor productivity (Casolaro and Gobbi 2007; Shen and Guo 2015).

DeYoung et al. (2004) describe the past and present of commercial banks through a series of descriptive statistics and envision their future. They find that changes in technology and deregulation have intensified competition in the banking industry, threatening the survival of some banks. With technological progress, small banks have tended to become more involved in commercial loans, payment services, customer financing, and customer investment. The advantages of small banks have weakened, and the market share of large banks has gradually increased with the advance of technology. In 1986, the total assets of the top ten banks in the United States by assets accounted for 28 percent of all bank assets. In 2001, this share had increased to 76 percent. From this, DeYoung et al. (2004) infer that future technological advances and deregulation will differentiate large and small banks. Xie and Gao (2021) use the index compiled by Xie and Wang (2022) to find that the higher the degree of digital transformation of commercial banks, the better is the bank performance, especially when banks carry out transformation in the cognitive and organizational dimensions.

In terms of risk control, with the application of digital finance, commercial banks can better describe customer profiles and serve "long-tail customers." This alleviates the information asymmetry in the traditional banking industry, increasing banks' market power and ability to resist and control risks (Jin et al. 2020; Li and Yang 2020). Li and Yang (2020) crawled all the relevant fintech news in the Advanced Search page of Baidu Information and added up the number of news articles searched by all the key words at each bank level each year to obtain the annual total news articles of the sample banks. Using this information, they measure the level of banks' financial technology development over time. They find that the level of a bank's financial technology development is negatively correlated with the level of risk undertaken by the bank, and there is a relationship between a bank's market power and risk. With the strengthening of bank market power, bank risk presents a trend of "first fall and then rise." Zhao et al. (2021) use data from company annual reports and the index compiled by Xie (2020). They find that the relationship between bank digital transformation and NPL ratio is nonlinear. When the degree of digital transformation is high, the bank digital transformation index and NPL ratio is negatively correlated. When the degree of digital transformation is low, the bank digital transformation and NPL ratio is positively correlated.

Most existing analyses on the relationship between digital transformation and bank risk mainly rely on correlation analysis. They have neither reached a consistent conclusion nor provided a credible analysis of the mechanism of the impact of digital transformation on different types of bank risks. In this paper we use the index compiled by Xie (2020) and data from banks' annual reports from 2011 to 2018. We first propose that the digital transformation of banks will lead to ex post risk differentiation of the loans of large and small banks, which is consistent with the Matthew effect. In addition, we examine the mechanism by which digital transformation differentiates the ex post risks of large and small banks' loans, and put forward specific policy recommendations.

3. Data and variables

3.1 Dependent variables

The analysis uses three dependent variables: (1) bank risk (especially the NPL ratio of bank loans); (2) proportions of bill discounting and corporate loans in all loans; and (3) average yield of bill discounting and corporate loans. The data sources are bank annual reports and the Wind Economic Database, and the data time range is 2011–18. The NPL ratio is generally used as a proxy variable to measure bank asset-side risk, which is bank ex post risk measurement.

3.2 Explanatory variables

We use the Digital Transformation Index of Commercial Banks (Phase 1) (hereinafter referred to as the "Index") compiled by the Institute of Digital Finance of Peking University (Xie and Wang 2022) to measure the level of digital transformation of commercial banks. The time span of the Index is 2011–18. The Index comprises 97 commercial banks, including five large state-owned commercial banks, 12 joint stock commercial banks, and 80 city commercial banks.

The Index has three sub-indexes: the Cognitive Digital Transformation Index, the Organizational Digital Transformation Index, and the Product Digital Transformation Index. In Figure A.1, a brief view of the Index is shown.

The Cognitive Digital Transformation Index refers to the commercial banks' understanding of and emphasis on the "technological change of digital finance," "digital," "digital finance," and other key words.

The Organizational Digital Transformation Index includes the setting of relevant departments of digital finance in banks, the setting of directors and executives with information technology backgrounds, and the banks' investment related to digital finance.

The Product Digital Transformation Index covers four major sectors: e-banking (mobile banking/Wechat banking), Internet wealth management, Internet credit, and e-commerce.

Table 1. Descriptive statistics

Symbol	Variable name	Number	Mean	SD	Min	Max
INDEX	Digital Transformation Index	667	2.8296	2.1411	0	11.7667
PRO	Product Digital Transformation Index	668	2.2769	1.6020	0	5
ORG	Organizational Digital Transformation Index	668	1.6295	1.4068	0	8
COG	Cognitive Digital Transformation Index	667	6.3338	6.5200	0	39.8337
WECHAT	Cognitive digital transformation–Wechat banking {1,0}	670	0.8015	0.3992	0	1
NPL	Nonperforming loan ratio	667	1.3621	0.8571	0.0200	13.2500
SIZE	Total assets, logarithm	631	19.4658	1.6435	16.6249	24.0447
ROA	Return on assets	669	0.9988	0.3844	0.0200	2.7000
CAR	Capital adequacy ratio	668	0.1287	0.0316	0	0.6
DGDP	Deposits/GDP	648	1.9776	4.0478	0.4987	103.4023
LGDP	Loans/GDP	648	1.5482	3.5320	0.1655	90.1567
COM	Competition in the province	669	0.9431	0.0413	0.8210	1
MOBILE	Number of mobile phone users	527	958.9315	711.9078	55.7	4076
DIS	Average shortest distance from the “eight horizontal and eight vertical” optical fiber trunk line network to the cities in the province where the bank’s headquarters is located	672	76188.81	63510.52	933.038	216477
5BIG	Proportion of the number of branches of the five major banks	632	0.3686	0.0964	0.1793	0.8043
LP_COMPANY	Proportion of corporate loans	232	68.6145	8.0334	40.1403	87.1252
LP_BILL	Proportion of discounted bills	231	4.1173	3.6227	0.0284	22.0081
LR_COMPANY	Corporate loans yield	146	5.8071	0.9406	3.99	7.79
LR_BILL	Bill discounting yield	99	5.1161	1.6164	2.06	10.61

The Product Digital Transformation Index is the most direct and core part of measuring the digital finance strategy of commercial banks (Xie and Wang 2022). The Product Digital Transformation Index reflects banks’ activity in product digital transformation and is divided into six levels: none = 0, very low = 1, low = 2, medium = 3, high = 4, and very high = 5.

3.3 Control variables

We control for the individual bank-level factors and city-level factors. The bank-level factors include the return on assets, capital adequacy ratio, and total assets. The city-level controls are the degree of local economic development (province gross domestic product [GDP]), the degree of local financial development (deposits/GDP and loans/GDP), and the degree of local financial competition at the prefecture-level city level (the proportion of the number of branches of the five largest banks). In addition, because the research objective of this paper is to analyze the digital transformation of banks, we use the Digital Financial Inclusion Index (Fintech) of the Peking University Digital Financial Inclusion Index to represent the local level of digital financial development and the level of local competition in digital finance. The analysis also controls time and province-level fixed effects, to control for changes at the policy level and other factors.

Table 1 provides the descriptive statistics of the variables. We drop the sample whose core variables are missing and outliers.

4. Empirical analysis

4.1 Model setting

We focus on the impact of bank digital transformation on bank risk. Referring to Demirgüc-Kunt and Huizinga (2010), we construct the relationship between banks' NPL ratio and banks' Digital Transformation Index.

$$Y_{it} = \beta_0 + \beta_1 tran_{it} + \beta_2 X_{it} + \lambda_t + \delta_i + \varepsilon_{it}, \tag{1}$$

where Y_{it} is banks' NPL ratio; $tran_{it}$ is the bank Digital Transformation Index; and X_{it} is the control variables, including the asset-liability ratio, capital adequacy ratio, total assets (logarithm), GDP per capita (logarithm), deposits/GDP, loans/GDP, proportion of the number of branches of the five largest banks, and the overall Digital Financial Inclusion Index. λ_t is the time fixed effect, and δ_i is the province fixed effect or the bank fixed effect.

The following methods are adopted to solve the endogeneity problem.

First, the next period of the NPL ratio is used as the dependent variable to estimate the impact of the digital transformation of banks on the NPL ratio.

Second, the number of mobile phone users at the prefecture level at the end of the year and the average closest distance to cities in the province along the "eight horizontal and eight vertical" optical fiber trunk line in the province where the bank's headquarters is located are respectively used as instrumental variables for the digital transformation index of banks, and the two are used as a joint instrumental variable for the digital transformation index of banks.

The equations for the mechanism analysis are the following:

$$RATIO_{it} = \beta_0 + \beta_1 TRAN_{it} + \beta_2 X_{it} + \lambda_t + \delta_i + \varepsilon_{it}, \tag{2}$$

$$RETURN_{it} = \beta_0 + \beta_1 TRAN_{it} + \beta_2 X_{it} + \lambda_t + \delta_i + \varepsilon_{it}, \tag{3}$$

where $RATIO_{it}$ represents the proportion of discounted bills, corporate loans, and personal loans to total loans, respectively; $RETURN_{it}$ represents the average yields of discounted bills, corporate loans, and personal loans, respectively; $TRAN_{it}$ is the digital transformation index of the bank; X_{it} is the control variables, including the asset-liability ratio, capital adequacy ratio, total assets (logarithm), GDP per capita (logarithm), deposits/GDP, loans/GDP, proportion of loans from the five major banks, and digital financial inclusion aggregate index. λ_t is the time fixed effect, and δ_i is the province fixed effect or the bank fixed effect.

Table 2. Benchmark regression with the full sample: Impacts of the overall Bank Digital Transformation Index and sub-indexes on banks' ex post risk

	(1)	(2)	(3)
	Nonperforming loan ratio		
	Full sample		
Digital Transformation Index	−0.0229		
	(0.0290)		
Product Digital Transformation Index		0.0446**	0.0630***
		(0.0226)	(0.0232)
Organization Digital Transformation Index			−0.0150
			(0.0238)
Cognitive Digital Transformation Index			−0.0128*
			(0.0076)
Bank control variables	Yes	Yes	Yes
City control variables	Yes	Yes	Yes
Province fixed effect	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes
Robust standard error	Yes	Yes	Yes
Number of observations	633	634	633
R^2	0.3977	0.3991	0.4031

Note: *Standard errors are in parentheses.* ****Statistically significant at the 1 percent level;* ***statistically significant at the 5 percent level;* **statistically significant at the 10 percent level.*

4.2 Benchmark regression without sample division

Table 2 presents the results of the benchmark regression, showing how the Digital Transformation Index and its sub-indexes (Product Digital Transformation Index, Organizational Digital Transformation Index, and Cognitive Digital Transformation Index) affect the NPL ratio, which represents banks' ex post risk. The table shows the results of the full sample regression, which indicates that the overall Digital Transformation Index, the Organizational Digital Transformation Index, and the Cognitive Digital Transformation Index do not affect banks' NPL ratio.

The Product Digital Transformation index has direct and indirect effects on the NPL ratio. In column (2), which includes all the sub-indexes, when the Product Digital Transformation Index increases by 1 unit, the NPL ratio increases by 0.0446 unit. In column (3), which includes only the Product Digital Transformation Index, when it increases by 1 unit, the NPL ratio increases by 0.0630 unit. In other words, on average, the NPL ratio of the bank with the most active product digitization was 0.315 percent higher than that of a bank without product digitization. Table 2 shows that from the perspective of the whole sample, at this stage, only the digital transformation of products has increased the NPL ratio, and the digital transformation of organizations and cognitive digital transformation can only increase banks' NPL ratio through the digital transformation of products.

This may be because only the digital transformation of products will change the banks' main deposit and loan-related behavior, while organizational digital transformation and cognitive digital transformation may not directly affect banks' behavior through channels other than product digital transformation. This also may be because it takes less time for

Table 3. Benchmark regression of large and small banks: Nonperforming loan ratio, bank size, and Product Digital Transformation Index

	(1)	(2)	(3)	(4)
	Nonperforming loan ratio			
	Large banks	Small banks	Large banks	Small banks
Product Digital Transformation Index	0.0444	0.0593**		
	(0.0403)	(0.0298)		
Wechat bank {1,0}			0.0011	0.2593***
			(0.1246)	(0.0887)
Bank control variables	Yes	Yes	Yes	Yes
City control variables	Yes	Yes	Yes	Yes
Province fixed effect	No	No	Yes	Yes
Bank fixed effect	Yes	Yes	No	No
Year fixed effect	Yes	Yes	Yes	Yes
Robust standard error	Yes	Yes	Yes	Yes
Number of observations	80	554	80	556
R^2	0.9102	0.5348	0.8112	0.4189

Note: *Standard errors are in parentheses.* *** *Statistically significant at the 1 percent level;* ** *statistically significant at the 5 percent level.*

product digital transformation to affect banks' risk, while the organization and cognitive digital transformation may take longer. The total duration of the digital transformation of banks is currently only 11 years, so this impact cannot be obtained from the existing data.

4.3 Benchmark regression with large and small banks

Banks of different size may have different ways to conduct business. Therefore, digital transformation may have different effects on large banks and small banks. In this section, we divide banks into subsamples for further analysis.

Table 3 divides the sample into large and small banks. We use two methods to separate large and small banks, with the second method creating the samples for the robustness test of the first method. For the first division method, the top ten banks in asset size are large banks and the others are small banks; for the second method, the top 15 banks in asset size are large banks. The results of the second division method are similar to those of the first division method, so only the regression results of the first division method are discussed in the text. The second division method is regarded as a robustness test, and the results are provided in Table A.2.

In Table 3, the explanatory variable in columns (1) and (2) is the Product Digital Transformation Index.[6] The explanatory variable in columns (3) and (4) is whether a bank has Wechat banking (yes = 1, no = 0). The regression results show that when the product digitization index increases by one unit, the NPL ratio of small banks increases by 0.0593 unit.

6 Product Digital Transformation Index has five sub-indexes: the existence of Wechat banking {1,0}, the existence of mobile banking {1,0}, the existence of Internet loan products {1,0}, the existence of Internet wealth management products {1,0}, and the existence of e-commerce {1,0}. More detailed aspects of the regression can be seen in Table A.1 and Table A.2.

In terms of the direct effect, compared with large banks without Wechat banking, there is no significant change in the NPL ratio of large banks with Wechat banking. Compared with small banks without Wechat banking, the NPL ratio of small banks with Wechat banking increases by 0.2593 unit. We also test the impacts of other four variables: the existence of mobile banking (yes = 1, no = 0), the existence of Internet loan products (yes = 1, no = 0), the existence of Internet wealth management products (yes = 1, no = 0), and the existence of e-commerce (yes = 1, no = 0). We find that the impacts of these four variables on the NPL ratios of large and small banks are neither stable nor significant, as is shown Table A.1.

The heterogeneity analysis in Table 3 shows that the digital transformation of products did not increase the NPL ratio of large banks, but it increased the NPL ratio of small banks. In Table A.2, the regression includes province fixed effects, and the results are still robust. To sum up, from the perspective of the whole sample, the general digital transformation of banks, the first-level cognitive digital transformation, and the digital transformation of organizations do not affect banks' NPL ratios; only the digital transformation of products affects banks' NPL ratios. The digital transformation of products has not increased large banks' NPL ratio (ex post risk), but has increased small banks' NPL ratio (ex post risk).

4.4 Robustness test

So far, the analysis has assumed that the digital transformation of banks is exogenous with respect to the disturbance term after adding fixed effects and other control variables. Although the bank characteristic variables, city characteristic variables, and fixed effects have been controlled as much as possible, it is still difficult for the model to ensure that the above fixed effect regression controls all the variables that may affect the digital transformation of banks and bank risks at the same time. An endogeneity problem caused by omitted variables or reverse causality would affect the consistency of the fixed effect results. Therefore, we conduct further robustness tests on the benchmark regression results for small and large banks.

To solve the endogeneity problem, we adopt the following strategy.

We use instrumental variables to solve the endogeneity problem. We use three instrumental variables. The first is the average closest distance between the city where the bank's headquarters is located and the cities in the province along the "eight horizontal and eight vertical" optical fiber trunk network. The second is the number of mobile phone users at the prefecture level at the end of the year, as the instrumental variables for the banks' digital transformation. The third is the joint instrumental variable of the two variables above.

The launch and use of digital banking products rely on fast and stable mobile Internet signals. The National Development and Reform Commission (the former State Planning

Commission) and the Ministry of Industry and Information Technology (the former Ministry of Posts and Telecommunications) built the "eight horizontal and eight vertical" large capacity optical fiber communication trunk network covering cities above the provincial capital and 90 percent of the prefectures in the country from 1986 to 2000. The rapid development of the communications industry in the 1990s laid a solid foundation for the rapid popularization of mobile networks and the development of digital finance. To a certain extent, the distance between the city where a bank's headquarters is located and the optical fiber communication trunk network reflects the level of convenience and cost for the bank, local residents, and companies to obtain fast and stable Internet services.

In general, the closer the city is to the optical fiber trunk line, the higher are the stability and quality of its Internet services, which is more conducive to the digital transformation of local banking products. In addition to affecting the NPL ratio of banks through the digital transformation of banks, the optical fiber network may also affect the behavior of residents and companies through the use of Alipay and other digital inclusive financial services, thereby affecting the NPL ratios of banks. However, our control variables already include the Peking University Digital Financial Inclusion Index, which excludes this possible channel. It is unlikely that the fiber communication trunk network, which was built as early as the 1990s, could affect banks' NPL ratios through other channels.

Similarly, mobile phones, as the most important tool for the use of digital banking products, are closely related to the digital transformation of banking products. In addition, after controlling for the local economic development level, financial development level, and digital financial inclusion index, there is no direct correlation channel between the number of mobile phone users at the prefecture level and bank risk, which make it possible to become an effective instrumental variable. Second, the NPL ratio in the next period is used as the dependent variable, the instrumental variables of the bank's digital transformation remain unchanged, and a robustness test is performed.

Unfortunately, because large banks are spread all over the country, we cannot find instrumental variables for the digital transformation of large banks. The best way is to use explanatory variables to regress the next period of the dependent variables for large banks. The results are shown in Table A.3, and they are consistent with the benchmark regression results.

Table 4 reports the results of the endogeneity analysis, including the regression using lag of the explanatory variable and the instrumental variable regression.

In columns (1) and (2), the dependent variable is the lag of the NPL ratio. Columns (3)–(5) are instrumental variable regressions, and the dependent variable is the NPL ratio. The instrumental variable for column (3) is the average shortest distance from the "eight horizontal and eight vertical" optical fiber trunk line network to the cities in the province

Table 4. Endogeneity: Large banks (dependent variable in the next period) and small banks (dependent variable in the next period and instrumental variables, respectively)

	(1) NPL_lag	(2)	(3) NPL	(4)	(5)
	Big banks	Small banks	Small banks	Small banks	Small banks
Product Digital Transformation Index (Wechat bank)	−0.0865 (0.0657)	0.1543* (0.0912)	2.1655** (0.9425)	1.4569** (0.6767)	1.7646*** (0.6564)
Bank control variables	Yes	Yes	Yes	Yes	Yes
City control variables	Yes	Yes	No	No	No
Province fixed effect	No	No	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes
Robust standard error	Yes	Yes	Yes	Yes	Yes
Number of observations	70	485	556	507	507
R^2	0.9130	0.5709	0.0406	0.2863	0.2075
First-stage F			9.7011	12.2319	8.8161
Prob > F			0.0019***	0.0005***	0.0002***
χ^2					0.4852
Prob > χ^2					0.4861

Note: *Standard errors are in parentheses. ***Statistically significant at the 1 percent level; **statistically significant at the 5 percent level; *statistically significant at the 10 percent level. The control variables include the digital financial inclusion index at the prefecture level.*

where the bank's headquarters is located. The instrumental variable for column (4) is the number of mobile phone users at the prefecture level at the end of the year. The instrumental variable for column (5) is the joint instrumental variable of the two instrumental variables. Due to space limitations, we do not report the regression results for the first stage, but we report the Kleibergen-Paap rk F-statistic and its p-value. The F-statistic of the first stage-regression is high. At the 1 percent level, the null hypothesis of the problem of weak instrumental variables is rejected, indicating that the instrumental variables are not weak. In addition, referring to Dinger and von Hagen (2009), we partially demonstrate the exogeneity of the joint instrumental variable using the overidentification test, reporting the Wald statistic and its p-value. Thus, there is no overidentification problem, which partially proves the exogeneity of the joint instrumental variable. Table 4 shows that the digital transformation of products through Wechat banking indeed increases the NPL ratio of small banks.

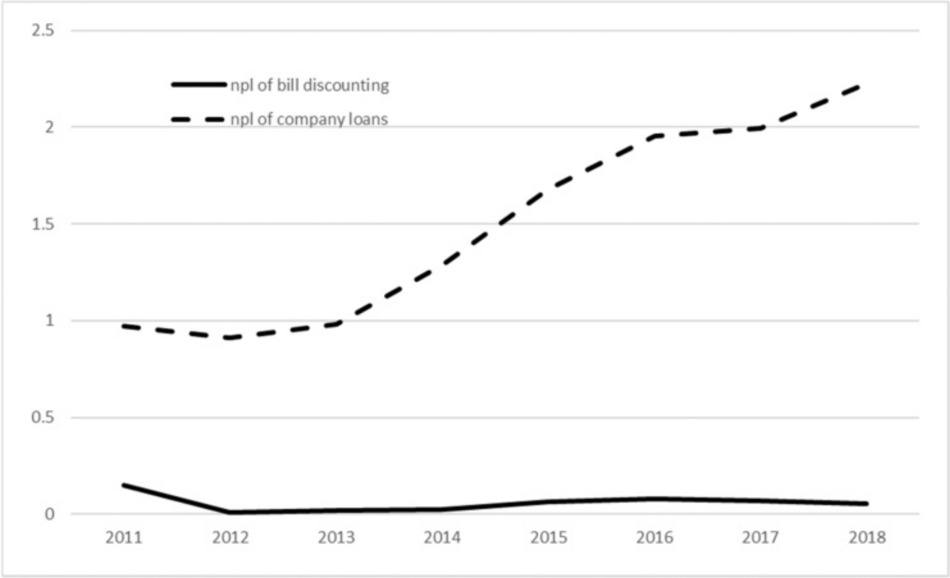
4.5 Mechanism analysis

The previous analysis showed that the digital transformation of banks' products can increase the NPL ratio of small banks, but it does not increase the NPL ratio of large banks. These results are statistically and economically significant. This section proposes and tests a hypothesis for the mechanism for the relationship between digital transformation and the NPL ratio.

4.5.1 Proposition of the hypothesis Our hypothesis is as follows.

(1) Digital transformation increases the proportion of large banks' bill loans or the yield of large banks' bill loans, which means that large banks extended the user group of bill discounting.

Figure 1. Fact 1: Discounted bills have a lower NPL ratio compared with corporate loans



(2) Digital transformation increases the proportion of small banks' corporate loans or the yield of small banks' corporate loans, which means that small banks extended the user group of corporate loans.
(3) Digital transformation decreases the proportion of small banks' bill loans and decreases the proportion of large banks' corporate loans.

There are two types of bank loans to companies. One is bill discounting, which is loans to other companies that are endorsed by large company credit or bank credit. The other type is corporate loans, which are loans to companies that are endorsed by their own credit. The average NPL ratio of bill discounting is higher than that of corporate loans (Figure 1). Against the background that banking sector policy mandates that all banks extend their user group of loans, both bill discounting and corporate loans are ways to fulfill the policy requirements. It is easier to check whether companies will default on bill discounting loans because it is only necessary to verify the authenticity of the bills. What's more, the NPL ratio of bill accounting is lower. However, corporate loans are endorsed by the credit of MSMEs and require complex prior risk assessment, such as due diligence. Therefore, as long as a bank can use bill discounting to fulfill the MSME loan target required by the policy, it is reluctant to issue corporate loans endorsed by small business credit to fulfill the policy requirement.

Table 5. Fact 2: In prefecture-level cities with low GDP, the proportion of large bank branches is low

	Number of branches of the five major banks in prefecture-level cities / Number of all bank branches in prefecture-level cities	
Log (GDP) (prefecture level)	0.4436***	2.0881***
	(0.1798)	(0.1984)
Time fixed effect	No	Yes

Note: *Standard errors are in parentheses.* ****Statistically significant at the 1 percent level.*

Table 6. Robustness check of the mechanism

	(1) Bill discounting/ Total loans	(2)	(3) Corporate loans/ Total loans	(4)	(5) Bill discounting yield	(6)	(7) Corporate loans yield	(8)
	Large banks	Small banks	Large banks	Small banks	Large banks	Small banks	Large banks	Small banks
Product Digital Transformation Index	0.8648***	−0.7668***	−1.4903**	0.3734	0.2853**	0.1650	0.0371	0.0807***
	(0.2277)	(0.2291)	(0.6615)	(0.3414)	(0.1407)	(0.2152)	(0.0515)	(0.0380)
Bank control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Robust standard error	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	80	150	80	151	56	43	66	79
R^2	0.5925	0.7340	0.9287	0.8136	0.9314	0.9528	0.9768	0.9699

Note: *Standard errors are in parentheses.* ****Statistically significant at the 1 percent level;* ***statistically significant at the 5 percent level.*

To fulfill the policy requirements for lending to MSMEs, large banks prioritize bill discounting. However, against the background of using traditional financial methods, there is a major obstacle to extend the large bank bill discounting user group: In "poor places," the ratio of large bank branches to small bank branches is relatively small and cannot capture all the local customers, as is shown in Table 5, which has become an obstacle to extend the bill discounting business of large banks. Digital transformation makes it unnecessary for large banks to rely on branch offices for bill discounting business, which is beneficial for large banks extending bill discounting business, thereby pushing small banks to the corporate loan business—the harder and riskier one.

4.5.2 Validation of the hypothesis The following regressions test our hypothesis. The regression results in Table 6 verify hypotheses (1) to (3), using the samples of large and small banks.

According to Table 6, the regression results show that when the Product Digital Transformation Index increases by 1 unit, large banks' proportion of corporate loans decreases by 1.49 percent, large banks' proportion of discounted bills increases by 0.86 percent, small banks' proportion of corporate loans is relatively stable, and small banks' proportion of discounted bills decreases by 0.77 percent. That is, on average, compared with large banks without digital transformation, among large banks with a high degree of digital transformation, the proportion of corporate loans decreases by 7.45 percent, and the proportion of discounted bills increases by 4.3 percent. Transformed small banks, with a high degree of digital transformation, decrease their proportion of discounted bills by 3.85 percent.

When the Product Digital Transformation Index increases by 1 unit, the corporate loan yield of small banks increased by 0.0807 percent, and the discounted bill yields of large banks increases by 0.2853 percent. That is, on average, compared with large banks without product digitization, the discount rate of bills of large banks with a high degree of product digitization increases by 1.427 percent. Compared with small banks without product digitization, small banks with a high degree of product digitization had a 0.404 percent higher loan yield.

Table A.3 provides the results of a robustness check, in which we use the second method to define "large banks."

5. Concluding remarks

We use the Digital Transformation Index of Commercial Banks and data from the annual reports of 97 banks, with a time span of 2011–18, to study the effect of commercial banks' digital transformation on the NPL ratios of large and small banks.

We find that:

(1) At this stage, out of three dimensions of banks' digital transformation—cognition, organization, and products—only transformation of products has significant impacts on banks' current NPL ratios. Changes in organization and cognition are effective only through their influence on changes in products. This may be because only changes in products represent material adjustments in banking businesses, while changes in organization and cognition can only have real impacts if they affect banks' financial products. This may also be because it takes less time for new products to affect banks' risk, while changes in organization and cognition may take a longer time to be effective. The study covers a period of 11 years, which is probably too short for examining the real impact of digital transformation of the banks, let alone that most banks started digital transformation only in recent years.

(2) The digital transformation in terms of the products of large banks did not change their NPL ratios, while that of small banks increased their NPL ratios. This confirms that the digital transformation of banks had an important difference in terms of risk implications between large and small banks.

(3) The broad story of risk differentiation between large and small banks is that all banks are required by the authorities to increase their lending to MSMEs every year—and they all try hard to fulfill this requirement. Against this institutional background, large banks take advantage of digital transformation to expand their bill discounting businesses. One of the consequences of digital transformation is that more information about bills held by MSMEs becomes available online. This allows large banks to extend bill discounting services to more MSMEs than before, using their advantages of large pools of cheap funds. Because the bills are often backed by large corporations' credit worthiness, this expansion of bill discounting businesses by the large banks does not increase loan risks. Since the large banks now serve more MSMEs with relatively higher quality, the small banks have to reach more and lower-quality MSMEs by offering corporate loans, to satisfy the policy requirement. But this inevitably leads to higher loan risks for the small banks.

Overall, it is not a bad outcome for the whole economy, as large banks increase their bill discounting business, and therefore push small banks to extend corporate loans to firms who have higher risks. Overall, MSMEs receive more funding, as a result. The logic of the large and the small banks' lending to small businesses is different: Large banks are more distant from the MSMEs, in general, and therefore, do not have enough soft information to support uncollateralized loans to MSMEs (Williamson 1967). On the other hand, small banks tend to have much closer relationships with the owners of the small banks, or even the owners themselves. As a result the owner of the bank may allow the lender to use more soft information (information on local businesses and business owners, etc.) in lending decisions. In this way, small banks rely more on soft information for making loans (Cole 2004). This gives small banks a natural comparative advantage in lending to small businesses. In contrast, large banks have comparative advantages in the bill discounting business. This is because part of the business in bill discounting is the loan endorsed by the credit of large companies. Compared with small banks, large banks often have stronger relationships with large companies. Therefore, development of the bill discounting business by large banks will help more MSMEs to obtain short-term financing through bill discounting. Thus, we believe that the rise of NPL itself may not totally be a bad thing, since more entities can have access to financial resources after bank digital transformation. At the same time, the differentiation of NPL ratio may be a good phenomenon, because it means that both large banks and small banks do the business where they have comparative advantages.

The NPL ratio is an important indicator for monitoring bank risks, but its role is limited. For example, the NPL ratio cannot monitor off-balance-sheet risks. The NPL ratio of bill discounting must be lower. However, this does not mean that the off-balance-sheet risks and financial system risks brought by bill discounting are lower than those of corporate loans. Discounted bills are short-term loans, generally no more than 180 days, and essentially money market funds, which may lead to problems such as the term mismatch.[7] At the State Council executive meeting on 20 February 2019, Li Keqiang said: The rise of bill financing and short-term loans may not only cause “arbitrage” and “fund idling,” but also may bring new potential risks.

Based on these findings, we also offer some policy recommendations.

First, it is important to look objectively at the optimization of the real economic services behind the risk differentiation of large and small banks, and not to carry out blind control on the differentiation of NPL ratio and the rise of NPL ratio of small banks.

Second, we recommend targeted prevention of possible systemic financial risks in the process of digital transformation of the banking industry. As mentioned herein, bill discounting can cause risks off the balance sheet. This requires regulators to pay special attention to where the funds of MSMEs obtained from large banks go. The government can use digital technology to integrate the information of logistics, capital flow, information flow, and other information on MSMEs and connect them to the digital bill discounting business of large companies. In this way, the digital bill discounting business of large banks is not just an online discount of traditional bills, but also allows for better supervision of the flow of funds from large banks to MSMEs, which can improve the quality of large banks’ services for financing private and MSMEs through discounted bills.

Third, we recommend scientific guidance of the banking industry to use digital technology to serve the real economy. Both discounted bills and corporate loans are essentially ways to expand loans to MSMEs. Regulators should think about how to guide the future pattern of the banking industry, in terms of the method for achieving the goal of enabling digitalization to assist in financing MSMEs. It is not necessary for regulators to impose the same requirements on large and small banks. Instead, the requirements should comply with the comparative advantages of large and small banks.[8] For example, regulators should not require large and small banks to lend to MSMEs in the same way. What’s more, regulators should not only look at the NPL ratio as a general indicator. Instead, they can set “red lines” for the NPL ratios of bill discounting and corporate loans separately. Regulators can tolerate higher NPL ratios of small banks, thus encouraging them to give more corporate loans to MSMEs, which uses soft information.

7 In many cases, companies use short-term company loans to finance for their long-term investment and research and development expenditure.

8 As we have mentioned above, small banks have a natural comparative advantage in lending to MSMEs because they can rely more on soft information for making loans (Cole 2004). In contrast, the large banks have comparative advantages in the bill discounting business.

References

Berger, Allen N. 2003. The Economic Effects of Technological Progress: Evidence from the Banking Industry. *Journal of Money, Credit and Banking 35*(2):141–176. 10.1353/mcb.2003.0009.

Casolaro, Luca, and Giorgio Gobbi. 2007. Information Technology and Productivity Changes in the Banking Industry. *Economic Notes 36*(1):43–76. 10.1111/j.1468-0300.2007.00178.x.

Cole, Matthew A. 2004. Trade, the Pollution Haven Hypothesis and the Environmental Kuznets Curve: Examining the Linkages. *Ecological Economic 48*(1):71–81. 10.1016/j.ecolecon.2003.09.007.

Demirgüç-Kunt, Asli, and Harry Huizinga. 2010. Bank Activity and Funding Strategies: The Impact on Risk and Returns. *Journal of Financial Economic 98*(3):626–650.

DeYoung, Robert, William C. Hunter, and Gregory F. Udell. 2004. The Past, Present, and Probable Future for Community Banks. *Journal of Financial Services Research 25*(2):85–133. 10.1023/B:FINA.0000020656.65653.79.

Dinger, Valeria, and Jurgen von Hagen. 2009. Does Interbank Borrowing Reduce Bank Risk? *Journal of Money, Credit and Banking 41*(2–3):491–506. 10.1111/j.1538-4616.2009.00217.x.

Guo, Pin, and Yue Shen. 2015. Does Internet Finance Increase Commercial Banks′ Risk-taking? Empirical Evidence from Chinese Banks [in Chinese]. *Nankai Economic Studies 4*:80–97.

Guo, Pin, and Yue Shen. 2019. Internet Finance, Deposit Competition, and Bank Risk-Taking [in Chinese]. *Journal of Financial Research 8*:58–76.

Hou, Xiaohui, Zhixian Gao, and Qing Wang. 2016. Internet Finance Development and Banking Market Discipline: Evidence from China. *Journal of Financial Stability 22*:88–100. 10.1016/j.jfs.2016.01.001.

Huang, Yiping, and Zhuo Huang. 2018. Digital Finance Development in China: Present and Future [in Chinese]. *Economics (Quarterly) 17*(4):1489–1502.

Jin, Hongfei, Hongji Lin, and Yinlu Liu. 2020. FinTech, Banking Risks and Market Crowding-Out Effect [in Chinese]. *Journal of Finance and Economics 5*:52–65.

Li, Xuefeng, and Panpan Yang. 2020. Fintech, Market Power and Bank Risk [in Chinese]. *Modern Economic Science 1*:45–57.

Qiu, Han, Yiping Huang, and Yang Ji. 2018. How Does FinTech Development Affect Traditional Banking in China? The Perspective of Online Wealth Management Products [in Chinese]. *Journal of Financial Research 11*:17–29.

Shen, Yue, and Pin Guo. 2015. Interest Finance, Technology Spillover and Commercial Banks TFP [in Chinese]. *Journal of Financial Research 3*:160–175.

Wang, Jing. 2015. Analysis on the Form of Internet Finance and Its Impact on Commercial Banking from the Perspective of Financial Functions [in Chinese]. *Finance and Economics 3*:56–65.

Williamson, Oliver E. 1967. Hierarchical Control and Optimum Firm Size. *Journal of Political Economy 75*(2):123–138. 10.1086/259258.

Xie, Xuanli, and Jue Gao. 2021. The Impact of Digital Transformation of Commercial Banks on Bank Performance. Working Paper, Institute of Digital Finance, Peking University.

Xie, Xuanli, and Shihui Wang. 2022. Digital Transformation of Commerical Banks: Measurement, Progress and Impact [in Chinese]. *Economics (Quarterly)* (forthcoming).

Xing, Yan. 2016. Inclusive Finance: A Basic Theoretical Framework. *Studies of International Finance* *9*:21–37.

Zhao, Jiaqi, Hongyi Jiang, Shiyun Hu, and Yan Shen. 2021. The Matthew Effect under Digital Inclusive Finance—Based on the Perspective of Digital Transformation. Working Paper, Institute of Digital Finance, Peking University.

Appendix A. Additional material

Figure A.1 A brief view of the Digital Transformation Index

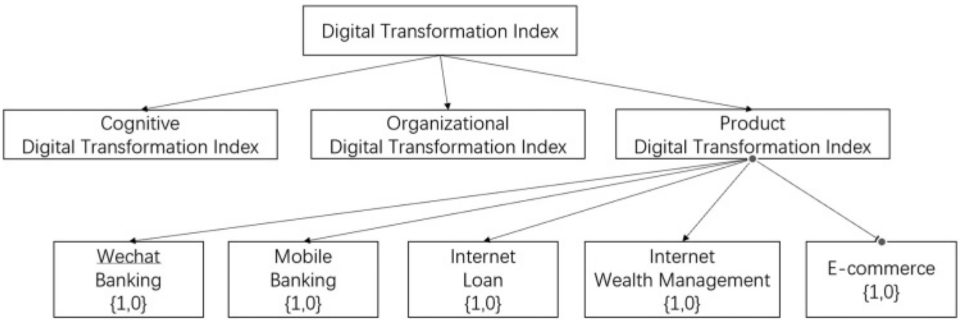


Table A.1. Robustness checks: Full versions of subcategories of the Product Digital Transformation Index

	(1) NPL	(2)	(3)	(4)
	Large banks	Small banks	Large banks	Small banks
Product Digital Transformation Index	0.0444	0.0593**		
	(0.0403)	(0.0298)		
Wechat {1,0}			−0.0387	0.2569***
			(0.1106)	(0.0911)
Mobile bank {1,0}			—	0.0606
			—	(0.0868)
Internet loan {1,0}			0.0974	−0.0135
			(0.0594)	(0.0622)
Internet finance {1,0}			−0.0422	0.1066*
			(0.0451)	(0.0622)
E-commerce {1,0}			0.1343**	−0.0629
			(0.0568)	(0.0719)
Bank control variables	Yes	Yes	Yes	Yes
City control variables	Yes	Yes	Yes	Yes
Province fixed effect	No	No	Yes	Yes
Bank fixed effect	Yes	Yes	No	No
Year fixed effect	Yes	Yes	Yes	Yes
Robust standard error	Yes	Yes	Yes	Yes
Number of observations	80	554	80	556
R^2	0.9102	0.5348	0.8298	0.4211

Note: *Standard errors are in parentheses.* *** *Statistically significant at the 1 percent level;* ** *statistically significant at the 5 percent level;* * *statistically significant at the 10 percent level.*

Table A.2. Robustness tests for large and small banks in the benchmark regression: Another way to divide large and small banks

	(1) Nonperforming loan ratio	(2)	(3)	(4)
	Large banks (Method II)	Small banks (Method II)	Large banks (Method II)	Small banks (Method II)
Product Digital Transformation Index	0.0016	0.0634*		
	(0.0255)	(0.0332)		
Wechat bank {1,0}			−0.1274*	0.3015***
			(0.0616)	(0.0987)
Mobile bank {1,0}			0.1752	0.0487
			(0.1472)	(0.0886)
Internet loans {1,0}			0.1042**	−0.0292
			(0.0464)	(0.0724)
Internet finance {1,0}			−0.0908	0.0779
			(0.0562)	(0.0667)
E-commerce {1,0}			0.1276**	−0.0796
			(0.0505)	(0.0776)
Bank control variables	Yes	Yes	Yes	Yes
City control variables	Yes	Yes	Yes	Yes
Province fixed effect	No	No	Yes	Yes
Bank fixed effect	Yes	Yes	No	No
Year fixed effect	Yes	Yes	Yes	Yes
Robust standard error	Yes	Yes	Yes	Yes
Number of observations	119	515	120	516
R^2	0.8762	0.5338	0.8081	0.4240

Note: *Standard errors are in parentheses. ***Statistically significant at the 1 percent level; **statistically significant at the 5 percent level; *statistically significant at the 10 percent level.*

Table A.3. Robustness checks for mechanism analysis: Method II to divide large and small banks

	(1) Bill discounting/ Total loans	(2)	(3) Corporate loans/ Total loans	(4)	(5) Bill discounting yield	(6)	(7) Corporate loans yield	(8)
	Large banks	Small banks	Large banks	Small banks	Large banks	Small banks	Large banks	Small banks
Product Digital Transformation Index	0.5484***	−0.9094***	−1.4011***	0.5800	0.1688	−0.1522	0.0368	0.1008***
	(0.1491)	(0.2917)	(0.4546)	(0.4038)	(0.1492)	(0.3197)	(0.0345)	(0.0335)
Bank control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Robust standard error	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	111	119	111	120	66	33	87	58
R^2	0.6418	0.7290	0.9469	0.7448	0.9179	0.9683	0.9785	0.9739

Note: *Standard errors are in parentheses. ***Statistically significant at the 1 percent level.*