Internet Appendix
to Does FinTech Innovation Improve Traditional Banks' Operating Efficiency and Risk?
Empirical Evidence from Patent Filings

Libing Fang,* Xindan Li,† Avanidhar Subrahmanyam,‡ and Ke Zhang§

October 10, 2026

A How China's Patent Filings Address Imperfections in U.S. Data

Lerner and Seru (2022) show that patent-level data in the US can cloud inferences because of the following: (a) the truncation of patent data, especially just prior to the end date, because only successful patent applications are observed, and (b) changes in the composition of sector- or region-level innovators, and how this interacts with (a). Under the leadership of Bronwyn Hall and Jim Bessen, a number of updated versions of the NBER data have been compiled, the most recent being an extension through 2006; please refer to Lerner and Seru (2022) for more details on the remaining issues with these data.

Patent filings in China allay the concerns mentioned above. Specifically, data from China National Intellectual Property Administration include successful and unsuccessful applications, and have text field identifiers that uniquely flag each applicant. In addition, the names of the applicants in patent filings are consistent with their official names, typically the names on the business license. Therefore, we are able to reliably match patent applicants with banks.

A minor issue may arise in our patent data from an employee applying for a patent in person. According to the Chinese Intellectual Property Law, applicants must submit a statement of no conflict with their employment if their application is in person. However, the incentives to be dishonest are low; specifically, the benefit that an employee receives from "stealing" such intellectual property is less applicable in our context because: (1) if a patent does have high market value, bank managers would "keep their eyes" on the employee with the patent; (2) if it does not have high market value, the employee would not steal it compared to the benefit

*School of Management and Engineering, Jiangsu Key Laboratory of Digital Finance, Nanjing University, 22 Hankou Road, Nanjing, Jiangsu 210093, China. Email: lbfang@nju.edu.cn.

†School of Management and Engineering, Jiangsu Key Laboratory of Digital Finance, Nanjing University, 22 Hankou Road, Nanjing, Jiangsu 210093, China. Email: xdli@nju.edu.cn.

‡Anderson Graduate School of Management, University of California at Los Angeles, Los Angeles, CA 90095, USA. Email: asubrahm@anderson.ucla.edu.

§Corresponding author, School of Digital Economics and Management, Nanjing University, 1520 Taihu Road, Suzhou, Jiangsu 215163, China. Email: kezhang@nju.edu.cn.

of retaining a steady job in the bank's technology department. Actually, to our knowledge, there are no such reports of intellectual property infringement in the Chinese banking industry.

B Description of the Reference Documents

To implement the natural language processing (NLP) algorithm that automatically identifies FinTech patents, we construct three categories of reference documents: commercial banking business (BB), FinTech (FT), and computer science (CS). Our approach generally follows the spirit of the method proposed by Bellstam, Bhagat, and Cookson (2021). We select reference documents based on the credibility and authority of their source, as well as their ability to represent the relevant domain's core concepts.

B.1 Reference Document on Commercial Banking Business

Following the spirit of Bellstam, Bhagat, and Cookson (2021), we adopt a widely recognized textbook, *Commercial Banking Operations and Management* (Science Press, 2021), as a reference document to capture the semantic features of commercial banking and serve as the benchmark for text similarity analysis.

We selected this textbook with multiple considerations in mind. First, textbooks are typically subject to academic scrutiny, ensuring a high level of authority and accuracy. As an academic resource endorsed by China's Ministry of Education and included in the national "13th Five-Year Plan" for higher education, *Commercial Banking Operations and Management* is used widely in undergraduate and postgraduate programs in finance, economics, and management, and is referenced by industry practitioners.

Second, the textbook is comprehensive in its coverage, with content spanning capital structure and liability management, asset operations, and off-balance-sheet activities. It includes detailed discussions on loan origination and recovery, capital adequacy frameworks, risk control systems, and internal governance procedures. These components form traditional banking's institutional core, and frequently appear in patents that focus on improving operational efficiency, compliance, or service delivery within legacy systems. Its scope enables a more holistic assessment of the similarity between patent texts and traditional banking functions, going beyond isolated keywords or narrow subtopics.

Third, the use of a textbook generally ensures consistency and comparability. In contrast to case-specific or empirical analyses that may vary by context or time period, textbook content remains relatively stable and widely accepted, supporting our classification method's reproducibility and facilitating future research replication.

In summary, using *Commercial Banking Operations and Management* strengthens the semantic basis of our methodology and supports a distinction between traditional banking technologies and FinTech-driven innovations.

B.2 Reference Documents on FinTech

To identify a reliable reference document for FinTech, we follow the same principle used for commercial banking—selecting sources that are both authoritative and comprehensive. However, with the absence of widely accepted textbooks in this rapidly evolving field, we turn to reports from four globally recognized institutions: the PBoC, BCBS, FSB, and World Bank. These institutions offer complementary perspectives—ranging from regulatory innovation (PBoC),[1] prudential supervision (BCBS),[2] and macro-financial stability (FSB),[3] to financial inclusion and development (World Bank).[4]

The selected reports are also thematically complementary. The PBoC's *FinTech Development Plan (2019–2021)* outlines a national roadmap for FinTech regulation and innovation. The BCBS's report *Sound Practices: Implications of FinTech Developments for Banks and Bank Supervisors* focuses on supervisory and risk management concerns. The FSB's *FinTech and Market Structure in Financial Services* assesses FinTech's potential impact on market structure and financial stability. The World Bank's *FinTech and the Future of Finance: Market and Policy Implications* emphasizes how FinTech is reshaping access, development, and regulatory frameworks.

Google's AI overview and search summaries reinforce the authority and relevance of these reports. The PBoC plan is described as "a significant step in China's efforts to regulate and guide the rapid growth of the FinTech sector." The BCBS report is characterized as "guidance on the implications of FinTech developments for banks and bank supervisors." The FSB document is noted for offering "a broad range of FinTech activities and their potential impact on market structure and financial stability," while the World Bank's report is summarized as an analysis of "how financial technology is reshaping the financial sector and outlines the implications for market outcomes and regulatory approaches." Together, these sources provide a comprehensive foundation for understanding FinTech innovation. Next, we provide a more detailed discussion of the documents and their sources.

1) *FinTech Development Plan (2019–2021)*—PBoC

Issued by China's central bank in August 2019, the *FinTech Development Plan (2019–2021)* (hereafter, the "Plan") is the first national, multiyear strategic roadmap for FinTech in China.[5] It reflects China's official strategic framework for promoting FinTech development. This document uses domain-specific terminology (e.g., blockchain, cloud computing, and artificial intelligence) and policy constructs (e.g., inclusive finance or penetrative regulation). It also outlines specific implementation tasks, such as deploying AI for risk control and

[1] As China's central bank, the PBoC plays a leading role in drafting and enforcing financial regulations, with a strong emphasis on advancing regulatory innovation to meet the demands of a rapidly evolving financial landscape (http://www.pbc.gov.cn/en/3688066/3688080/index.html).

[2] The BCBS is the primary global standard setter for the prudential regulation of banks and provides a forum for regular cooperation on banking supervisory matters (https://www.bis.org/bcbs/index.htm.)

[3] The FSB is an international institution that promotes global financial stability and plays a key role in safeguarding macro-financial stability through coordinated regulatory and supervisory policies (https://www.fsb.org/about/).

[4] The World Bank is a global development institution that provides financial and technical assistance in developing countries, with a core mission to promote inclusive and sustainable economic development (https://www.worldbank.org/ext/en/who-we-are?intcid=ecr_hp_about_en_ext).

[5] Please see https://www.caixinglobal.com/2022-01-06/pbocs-new-fintech-development-plan-flags-more-regulation-in-the-pipeline-101826767.html (in Chinese)

expanding cloud infrastructure—concrete features that serve as anchors in identifying FinTech functionalities in patent texts.

2) *Sound Practices: Implications of FinTech Developments for Banks and Bank Supervisors*—BCBS

Published in August 2017 as a "Sound Practices" consultative paper, this report systematically examines the supervisory and risk management implications of FinTech adoption by banks. It incorporates regulatory terms such as operational, cyber, and compliance risks that allow patent classifications at the intersection of banking supervision and digital transformation. Thus, what distinguishes this report is its focus on risk. It treats FinTech not just as a technological evolution but as a driver of operational change with implications for regulatory technology (RegTech), cyber resilience, and compliance frameworks. Its discussion of supervisory technology (SupTech) and institutional adaptation underscores the report's dual relevance to innovation and regulation. The report also organizes key enabling technologies—such as AI, blockchain, cloud computing, and biometric authentication—into functional categories. This structure is effective in building FinTech-related keyword libraries for semantic analysis.

3) *FinTech and Market Structure in Financial Services*—FSB

Released in February 2019, FSB formally defines FinTech in this report as "technology-enabled innovation ... with material effect."[6] It offers a macro-financial perspective, detailing impacts on market dynamics, competition for FinTech innovation, and structural risks. The report highlights FinTech innovation drivers, including API development, cloud services, and regulatory shifts, while drawing on case studies from major economies (e.g., China, Korea, and the UK) to demonstrate functional and geographic diversity. Such variation enhances the utility of the document for cross-sectional patent assessment.

Another notable contribution of this report is its analysis of BigTech's entry into financial services—raising important issues around hybrid business models, regulatory boundaries, and potential market concentration. These themes are crucial for identifying patents that operate at the convergence of finance and general-purpose technologies. The report also considers systemic risks such as market concentration and regulatory arbitrage, providing a foundation for linking patent content with macro-financial concerns.

4) *FinTech and the Future of Finance*—World Bank

Published in 2023 on the World Bank Open Knowledge Repository (OKR), the *FinTech and the Future of Finance* report ranks as the top search result for "FinTech."[7] The document offers a global perspective on financial development and inclusion, incorporating terminology related to access to finance, small- and medium-sized enterprise (SME) support, digital infrastructure, and regulatory trade-offs. This makes it particularly valuable for classifying patents that promote economic development and innovation in financial services.

Unlike other reports, *FinTech and the Future of Finance* emphasizes FinTech's socioeconomic dimensions—particularly its role in extending financial services to underserved populations during periods of systemic stress, such as

[6] https://www.fsb.org/2019/02/fintech-and-market-structure-in-financial-services-market-developments-and-potential-financial-stability-implications.

[7] At the time of writing, this document ranked among the top search results. See https://openknowledge.worldbank.org/entities/publication/b8167e93-e3fb-4d45-a610-a3e774c42f3e.

the COVID-19 pandemic. It discusses innovation domains such as mobile banking, digital identity, and central bank digital currencies (CBDCs), all of which are relevant to identifying patents in service delivery and financial access.

Additionally, the report covers regulatory trade-offs, consumer protection, digital privacy, and market competition, offering interpretive frameworks for patents that align with policy and development goals. This report's global outlook and long-term focus complement the more regulatory and technical orientations alongside the other reference documents discussed in this section.

For practical implementation, we translate all of the documents except PBoC's *FinTech Development Plan (2019–2021)* from English into Chinese, and then concatenate all of the documents with the original Chinese patent texts. We use the resulting hybrid documents as the benchmark for identifying FinTech-related patents.

B.3 Reference Document on Computer Science

We apply the same principle used for commercial banking and FinTech by selecting a source that is both authoritative and comprehensive. To accurately identify patents that embody substantive innovations in computer science (CS)—as distinct from those that simply apply digital tools in financial contexts—we adopt the *Glossary of Computer Science Terminology (Third Edition)* as our benchmark for text similarity analysis.

Given computer science's broadness and complexity, a single report or textbook cannot capture the field's full scope. A glossary, by contrast, offers structured semantic coverage across subfields, making it suitable for our purposes. Thus, we use what is widely regarded as the most authoritative glossary of CS terminology in Chinese,[8] which aligns closely with the linguistic characteristics of our Chinese-language patent dataset.

This glossary was compiled under supervision of the National Committee for the Standardization of Scientific and Technical Terms and led by the Computer Science Terminology Committee of the China Computer Federation (CCF). It contains 8,795 entries, covering key subfields such as theoretical computing, systems architecture, hardware design, software engineering, databases, and artificial intelligence. Each entry includes a standardized Chinese term, its English equivalent, and a concise, technically rigorous definition. The glossary's structure ensures semantic clarity and disambiguation, especially for emerging or overlapping terminologies.

C Identifying FinTech Patents Using an NLP Algorithm

This section details our methodology for identifying bank-specific FinTech patents using an NLP algorithm. In Section C.1, we outline the data collection and text-processing pipeline, followed by a stepwise description of the semantic similarity algorithm. Section C.2 provides excerpts from the texts of patents categorized as FinTech, CS, and BB, to illustrate semantic distinctions embedded in our approach. We also present examples of FinTech and non-FinTech patent texts in full detail, to further clarify the classification process.

[8] https://www.ccf.org.cn/jsjkxjsmcgzjs (in Chinese).

C.1 Data Collection and Identification Procedure

We collect all International Patent Classification (IPC) G&H class patents whose applicant names match banks and IT firms with bank blockholders. Following standard NLP practice, we preprocess the reference documents and banks' patent filings for vectorization. This work includes segmenting and removing stop-words. Segmenting entails partitioning words in each sentence, and is required because Chinese words are not separated by spaces, unlike in English. Stop-words are common nonsense words that are of little value for classification and need to be deleted before applying NLP. Based on these considerations (listed in Qiu, Pei, and Yan 2020), Geng et al. (2021) share a toolkit for Chinese textual analysis on Github.[9] This tool is what we use to segment and remove stop-words.

After preprocessing, we convert the reference documents (FT, CS, and BB) and each bank's patent filing (PF) into numerical vectors. Following prior literature (Hoberg and Phillips 2016, Bellstam, Bhagat, and Cookson 2021, Guzman and Li 2023), we use *doc2vec* (Le and Mikolov 2014) for vectorization. *doc2vec* builds on the word-embedding model *word2vec* (Mikolov et al. 2013), which captures the context of words rather than relying on simple word counts. For instance, in *word2vec*, the word "Kellogg" would have different vectors depending on whether it appears in the context of the management school or the food company, while "Kellogg" and "General Mills," both food companies, would produce similar vectors. This allows *word2vec* to encode the meaning of a word. *doc2vec* extends this approach to entire documents, enabling vectors that capture document-level semantic content.[10]

The *doc2vec* algorithm represents each document—whether a patent filing or a reference document—as a numerical vector in a "semantic space," where the distance between two documents reflects their similarity in meaning, allowing us to detect conceptual similarity even when the wording differs. A key feature of *doc2vec* is that it considers not only individual words, but the *surrounding context* in which they appear. In practice, this allows the model to learn how words interact with their neighbors and collectively form coherent themes for the entire document. By repeatedly predicting words based on both the document's overall context and the words surrounding them during training, *doc2vec* gradually learns the typical patterns and relationships among words throughout the text. Each prediction reinforces the model's understanding of which words tend to appear together and in what contexts, enabling it to capture not only individual word meanings, but also the document's overall themes, topics, and semantic structure. Thus, documents sharing similar themes, tone, and topical emphasis—such as those describing comparable FinTech technologies—are positioned close together in the semantic space.

The third step in Section 2.1 of mian text is to label which patents are FinTech by comparing how close each patent's meaning is to the FT reference document. More specifically, we calculate the cosine similarity for each bank's patent vector (PF) to each reference document vector, which is defined as $x \cdot y / (|x| \cdot |y|)$, where $x \cdot y$ is the inner product of two vectors x and y, and $|x| \cdot |y|$ is the product of the two vectors' scalar lengths.

[9] See https://github.com/fastnlp/fastHan.

[10] See Li et al. (2021) and Guzman and Li (2023) for recent applications of *doc2vec* in finance and management research.

The pairwise similarity measures for PF, FT, CS, and BB are denoted as $\langle PF, FT\rangle$, $\langle PF, CS\rangle$ and $\langle PF, BB\rangle$, respectively. We label the patent as a FinTech patent if $\langle PF, FT\rangle$ is larger than $\langle PF, CS\rangle$ and $\langle PF, BB\rangle$, and non-FinTech otherwise.

C.2 Examples of FinTech and Non-FinTech Patents

In this section, we first present extracts from texts of patents that are identified as FinTech, CS, and BB, and thus highlight their semantic distinctions. We then provide detailed examples of FinTech and non-FinTech patents, including complete filing texts, to further demonstrate our process.

Central to our patent classification is a semantic-based approach leveraging *doc2vec* and cosine similarity. This method captures contextual meaning by evaluating the alignment between a patent's semantic content and the FSB definition of FinTech. Consequently, patents addressing core FinTech innovations are classified as FinTech-related. This property is particularly relevant for our study, as the language in FinTech patents can vary across time and institutions because of differences in business models, writing styles, or regulatory changes, while still conveying similar technological concepts. For instance, a patent may shift from describing "mobile payment solutions" to "app-based transaction platforms," yet the underlying innovation remains comparable. By focusing on semantic meaning rather than literal word overlap, *doc2vec* provides a consistent basis for identifying FinTech-related patents, reducing the measurement error from superficial linguistic variation. We then use cosine similarity to quantify how closely each patent aligns with our FinTech reference documents.

C.2.1 Patent Excerpts

To illustrate how such alignment is assessed, we present the following excerpted examples of patents that while modestly related to designated keywords, exhibit strong semantic congruence with FinTech innovation.

Example 1. A patent filed by China CITIC Bank (application 201911081115.7) describes "using machine learning algorithms to build a comprehensive credit-scoring model based on identity traits, fulfillment ability, behavioral preferences, credit history, and social attributes." Although the patent does not rely heavily on typical FinTech keywords such as digital banking, it reflects multiple core FinTech characteristics. First, it introduces a novel credit-scoring system that integrates multiple data dimensions and applies machine learning, representing a substantive innovation in financial services. Second, it incorporates key enabling technologies—particularly AI—clearly within a financial context. Third, it explicitly is designed for a financial application, namely credit evaluation. Together, these elements align closely with the FinTech definition and substantiate the classification.

Example 2. A patent filed by the Industrial and Commercial Bank of China (application 201910767405.0) outlines a method for evaluating customer satisfaction through facial expression recognition at the beginning and end of a banking transaction. Specifically, it describes "capturing facial images before and after service and evaluating satisfaction through expression recognition to enable contactless feedback." This patent introduces

AI (computer vision) into the customer experience assessment process, transforming what was traditionally a manual task into an automated, real-time system. Notably, the technical innovation is captured at the level of full-sentence semantics rather than relying on specific keywords. Although common FinTech terminology is eschewed, integrating AI in this financial service context represents a structural enhancement to service delivery and supports its classification as a FinTech patent.

Example 3. A patent filed by Qingdao Rural Commercial Bank (application 201711445726.6) describes "an Internet-based community financial service system for SMEs that provides mobile finance, service linkage, and security authentication, with payment as the core." This patent does contain common FinTech terminology and presents a concrete application scenario in SME finance. By integrating mobile platforms, internet security, and operational analytics, the patent proposes a restructured service delivery model for small businesses. Its use of digital technologies and its focus on financial service innovation strongly support its inclusion in the FinTech category.

The last two examples we present serve to illustrate how our semantic approach identifies patents in CS and BB, distinguishing them from FinTech patents.

Example 4. This patent, filed by the Agricultural Bank of China (application 201610302670.8), relates to CS technologies. It states, "A digital signature scheme based on conic curves over Zn is proposed, with numerical simulation provided." This suggests that the patent focuses on developing a cryptographic algorithm grounded in mathematical theory and information security, rather than on a financial application. It further notes, "An attacker would need to solve two mathematical problems to break the digital signature. . . indicating strong resistance to attacks." This innovation lies primarily in the field of cryptography, without reference to specific financial services or use cases. In this context, the patent does not propose a new financial service model, product, or process. Also, it lacks integration with financial scenarios or terminology, and contains no discussion of account structures, transaction logic, or financial risk considerations. According to the FSB's definition of "technologically enabled financial innovation that could result in new business models, applications, processes, or products with a material effect on financial markets and institutions," it is more appropriate to classify this patent as belonging to the CS domain.

Example 5. This patent, filed by the Bank of China (application 201610788596.5), concerns BB practices. It describes: "monitoring the balance of a designated bank account and determining whether it meets preset conditions for surplus transfer or shortage replenishment." This reflects a rule-based approach to funds management that does not involve algorithmic optimization, AI, or other advanced technologies. According to the FSB's definition, this patent would not be considered FinTech for two reasons. First, while it enhances the efficiency of traditional fund transfer processes, it does not introduce new financial infrastructures, business models, or service paradigms. Second, it does not incorporate, nor explain, the functioning of enabling technologies such as AI, blockchain, cloud computing, big data, or the IoT.

As illustrated by the five examples above, while keyword frequency provides a useful signal, our classification also relies on the broader semantic context of each patent and its alignment with the concept of FinTech

innovation.

C.2.2 Detailed Texts from a FinTech and a Non-FinTech Patent

Finally, we provide two detailed examples of patents with detailed filing text information. For FinTech, Figure C.1 shows that the Agricultural Bank of China has applied for a patent titled "Method and Device for Assessing Credit Risk."

Figure C.1: **Example of a FinTech patent from the Agricultural Bank of China's application.**

Title(CN)	信用评价模型的构建方法及装置
Title(EN)	Method and device for assessing credit risk
IPC	G06Q 40/02(2012.01)
BankName(CN)	中国农业银行股份有限公司
BankName(EN)	AGRICULTURAL BANK OF CHINA LTD
BankType(CN)	国有大型商业银行
BankType(EN)	State-owned bank
ApplicationNo.	201810689255.1
ApplicationDate	2018.06.28
GrantNo.	nan
GrantDate	nan
Abstract(CN)	本申请提供了一种信用评价模型构建方法，该方法可以通过多次筛选及回选的方式，从业务数据的属性特征中选择出对评价信用评分具有影响作用的属性特征，选择出的属性特征用于构建违约概率评价模型，该模型可以计算违约概率，再获得预设的违约概率与信用评分转换模型，该模型可以将违约概率转换为信用评分，因此该两个模型可以作为信用评价模型。另外，本申请还提供了一种信用评价模型构建装置，用以保证所述方法在实际中的应用及实现。

In English, the abstract of this patent states the following: "The present application provides a method and device for assessing credit risk. The method can select features that have a significant impact on credit risk using business data via multiple rounds of screening with replacement, in order to build a model for evaluating and calculating default probability. Then, it uses a predefined conversion model to convert the default probability to a credit score that can be used in business. Thus, the combination of these two models is proposed as the model for assessing credit risk in practice. In addition, this application provides a device that allows for the model's application and implementation in practice."

For an example of a non-FinTech patent, Figure C.2 shows that the Industrial & Commercial Bank of China (ICBC), the largest state-owned bank, has applied for a patent titled "Database Switching Methods, Electronic Devices, Systems, and Storage Media." We translate the patent's abstract as follows: "The present invention discloses a database switching method, electronic device, and storage medium. The invention includes: obtaining status information on execution capability of the master database and at least one standby database; judging whether to execute a switching operation between the master and standby database according to the status information of the master database; executing the switching operation between the master database and the standby database; and and performing data replenishment for the new master database. The present invention can flexibly select a suitable new master database accordingly, and ensure the integrity of the data after the switching operation via data replenishment."

The first patent is identified by our procedure as a FinTech patent (as it automates credit risk screening) and

Figure C.2: **Example of a non-FinTech patent from the Industrial & Commercial Bank of China's application.**

Title(CN)	数据库切换方法、装置、系统、电子设备及存储介质
Title(EN)	Database switching methods, devices, systems, electronic devices and storage media
IPC	G06F 16/27(2019.01)
BankName(CN)	中国工商银行股份有限公司
BankName(EN)	INDUSTRIAL & COMMERCIAL BANK OF CHINA (ICBC)
BankType(CN)	国有大型商业银行
BankType(EN)	State-owned bank
ApplicationNo.	201910732853.7
ApplicationDate	2019.08.09
GrantNo.	nan
GrantDate	nan
Abstract(CN)	本发明公开了一种数据库切换方法、装置、系统、电子设备及存储介质，其中，该方法包括：分别获取主数据库和至少一个备用数据库的状态信息，状态信息包括：数据信息和执行能力信息；根据主数据库的状态信息判断是否执行主备数据库切换操作；响应于判断结果为执行主备数据库切换操作，根据预定切换策略从至少一个备用数据库中选择新的主数据库；根据获取的至少一个备用数据库的状态信息，对新的主数据库进行数据补偿操作；对主数据库和新的主数据库执行切换操作。通过本发明，可以根据实际情况灵活选择合适的新主数据库，并且通过数据补偿操作，可以保证切换操作后数据的完整性。

the second is designated as a non-FinTech patent (because it has no finance application). These two detailed examples thus underscore the ideas behind our classification method, as detailed in Section 2.1 of main text.

D FinTech-related Words in Banks' Annual Reports

In this appendix, we consider whether banks actually utilize their FinTech patents. We investigate this issue by examining FinTech-related words in banks' annual reports. We first collect the annual reports of the listed banks in our sample. The dataset consists of 257 annual reports, with 107 observations in the high FinTech group and 150 observations in the low FinTech group.

We next examine differences in the frequency of FinTech-related words across the high and low FinTech groups. Specifically, after preprocessing the text, such as segmentation and removal of stop words as in the Appendix C.1, we compute the occurrence of FinTech-related words in each report. These words include blockchain, digital banking, mobile banking, online banking, virtual bank, electronic payments, payment gateway, facial recognition, biometric technology, automated processes, financial technology, FinTech, and open banking. Table D.1 indicates that both the ratio and the count of FinTech words are significantly higher for banks in the high FinTech group. This suggests that banks with higher levels of FinTech patent filings indeed utilize more FinTech in their operations.

Table D.1: **Comparing FinTech-related words frequency in annual reports between high and low FinTech innovation group.**

This table reports the frequency of FinTech-related words within annual reports across high and low FinTech innovation groups, as defined in Section 3 of main text. *FT_ratio* is the percentage of FinTech words in the Management Discussions and Analysis (MD&A) section of annual reports. *FT_num* is the count of occurrences of FinTech words in the MD&A section. *MDA_num* is the total number of words in the MD&A section. $\Delta Mean$ tests the difference between high and low FinTech innovation group via t-statistics that are reported in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	FT_ratio (%)	*FT_num*	*MDA_num*
High FinTech group	0.128	11.093	8484.047
Low FinTech group	0.093	7.573	7437.260
$\Delta Mean$	0.035^{***}	3.520^{***}	1046.787^{***}
	(3.885)	(4.449)	(3.318)

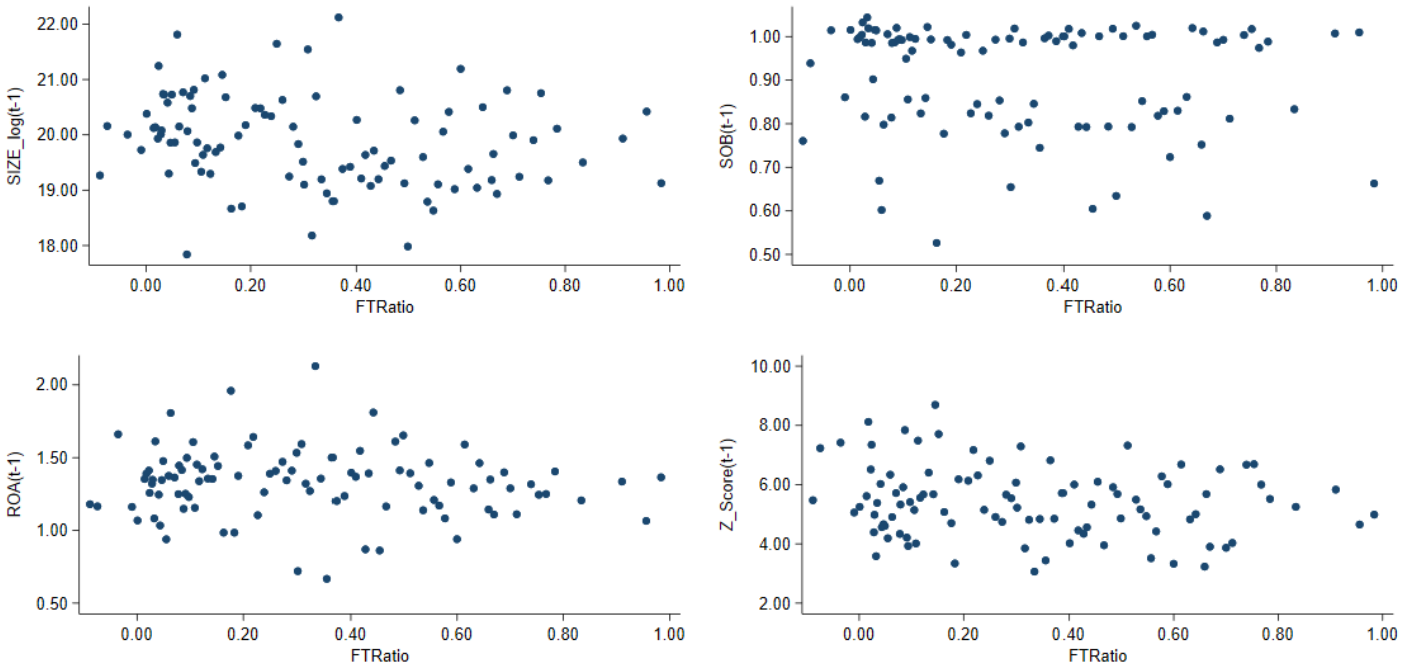
E FinTech Innovation and Control Variables

E.1 FinTech Innovation and Prevailing Bank Characteristics

This subsection examines whether banks' engagement in FinTech innovation is associated with previously-established bank characteristics. To explore this possibility, we analyze four one-year-lagged characteristics: bank size ($SIZE$), state-owned ownership status (SOB), profitability (ROA), and overall risk (Z_score).

First, we plot bin-scatters to visualize the relationship between the FinTech-to-total-patent ratio ($FTRatio$) and each bank characteristic in Figure E.1. The figures display a fairly even dispersion across bins.

Figure E.1: **Banks' FinTech Patenting and Their Characteristics**
This figure shows binned scatters of FinTech-to-total-patent ratio ($FTRatio$) versus four one-year-lagged bank characteristics including $SIZE$, SOB, ROA and Z_score. $SIZE$ is the log of a bank's total assets (in million RMB); SOB is a dummy variable for state-owned banks; ROA is the return on assets ; The overall risk measure $Z_score = (ROA + CAR)/sd(ROA)$, where $CAR = capital/risk\ weighted\ assets$ indicates the extent of bank shareholders' capital covering the risky assets, and $sd(ROA)$ is the standard deviation of ROA during the sample period. Each scatter point represents the mean Size, SOB, ROA, and Z-score against the mean $FTRatio$ in each bin. All plots are residualized with respect to year fixed effects as in Akcigit et al. (2022).



Second, we use the binned means to estimate the slopes by linear regression of $FTRatio$ on each of these characteristics to quantify potential selection bias. The estimated coefficients are -0.060 for SIZE (standard error = 0.036), -0.225 for SOB (standard error = 0.349), 0.020 for ROA (standard error = 0.152), and -0.006 for Z-score (standard error = 0.029).

In sum, these results indicate that FinTech innovation is not related to key prevailing bank characteristics, mitgating selection concerns.

E.2 Definitions of Control Variables

Controls	Interpretation	Data source
IntangibleRatio	The ratio of banks' intangible asset to total asset	Banks' yearly fiscal report table from CSMAR
GradProp	The proportion of employees holding postgraduate degrees	Banks' yearly fiscal report table from CSMAR
SIZE	The logarithm of total assets on each bank's balance sheet in each year	Banks' yearly fiscal report table from CSMAR
BRANCH	The total number of each bank's branches in each year normalized by total assets	Bank branches' basic information from CSMAR
NCITY	Number of unique cities in which the bank has at least one branch	Bank branches' basic information table from CSMAR
HHI	The Hirschman-Herfindahl index in each year, computed by market share using the number of branches of each bank in the city where a bank has the most branches	Bank branches' basic information table from CSMAR
PGDP	The logarithm of yearly per capita GDP of each bank's operating province (that in which the bank has the most branches)	National Bureau of Statistics, China; Bank branches' basic information table from CSMAR
POP	The logarithm of yearly population in each bank's operating province (in which the bank has the most branches)	National Bureau of Statistics, China; Bank branches' basic information table from CSMAR
NFIRM	The logarithm of yearly number of firms that are registered in the province where a bank has the most branches	National Bureau of Statistics, China; Bank branches' basic information table from CSMAR

F Estimating Bank Efficiency Measures via SFA

This section provides details on the estimation of bank efficiency using the SFA framework. The overall objective is to estimate η_i and u_{it} (the profit inefficiency parameters) in Equation (4) of main text and thus calculate *Prof_eff* via Equation (3) of main text. For this purpose, we proceed as follows. First, we define the control variables, i.e., ξ_{it} in Equation (4) of main text, as follows. The first two variables, non-interest cost to total assets (*NIntstCostAssetRatio*) and non-interest cost to total revenues (*NIntstCostRevnRatio*) relate (inversely) to a bank's market power (Huber 2021).[1] The remaining two variables are $PNPL$ and $PNPL^2$, where $PNPL$ is the ratio of nonperforming loans to gross loans in the province where bank i operates (Berger and Mester 1997, He, Ho, and Xu 2020).[2] Such a $PNPL$ in a province is a relatively exogenous control variable for a bank's economic environment. We also include year dummies ($Year_t$) to avoid estimation biases because of technological progress or changes in economic and regulatory environments over the years. The resulting equation is:

$$\begin{aligned}\ln \pi^a_{it} = \ln \tilde{\pi}_{it} + \rho_1 NIntstCostAssetRatio_{it} + \rho_2 NIntstCostRevnRatio_{it} \\ + \rho_3 PNPL_{it} + \rho_4 PNPL^2_{it} + dYear_t - \eta_i - u_{it} + \mu_i + v_{it}.\end{aligned} \tag{F.1}$$

The specification in Equation (F.1) allows the separation of efficiency-related components (η_i and $u_{it} > 0$) from purely random shocks and the effect of other exogenous variables.

When estimating the *Cost_eff* (instead of *Prof_eff*), we adjust Equation (F.1) and calculate $\ln C^a_{it}$ as:

$$\begin{aligned}\ln C^a_{it} = \ln \tilde{C}_{it} + \rho_1 NIntstCostAssetRatio_{it} + \rho_2 NIntstCostRevnRatio_{it} \\ + \rho_3 PNPL_{it} + \rho_4 PNPL^2_{it} + dYear_t + \eta_i + u_{it} + \mu_i + v_{it}.\end{aligned} \tag{F.1$'$}$$

Here, C^a_{it} denotes the actual total cost of bank i observed in year t; $\tilde{C}_{it}$ is the total cost in the best practice to be estimated. Compared with Equation (3′) of main text with Equation (3) of main text, the difference is that *Cost_eff* in the former case captures the idea that the actual cost is statistically larger than in best practice, while for the latter case, the actual profit is lower than in best practice. We replace the corresponding terms in Equation (F.1) by Equation (F.1′) when estimating *Cost_eff*.

The next step involves specifying a function that maps best-practice profits and costs (i.e., $\tilde{\pi}_{it}$ and $\tilde{C}_{it}$) to inputs. Let $\tilde{\mathcal{L}}_{it}$ denote, in turn, the profit or the cost in best-practice. We omit the subscripts (i, t), and following the literature (Kumbhakar, Wang, and Horncastle 2015, Park, Han, and Lee 2023, He, Ho, and Xu 2020), assume

[1] As Huber (2021) states, the non-interest cost captures proportional fixed costs, and this proportion should decrease as the bank gains market power.

[2] The weight of a bank in $PNPL$ is the proportion of total loans issued by the bank in the same province.

a translog function for the best-practice frontier:

$$\begin{aligned}\ln \tilde{\mathcal{L}} = \delta_0 &+ \sum_{j=1}^{3} \delta_j \ln\left(\frac{z_j}{z}\right) + \frac{1}{2}\sum_{j=1}^{3}\sum_{k=1}^{3} \delta_{jk} \ln\left(\frac{z_j}{z}\right)\ln\left(\frac{z_k}{z}\right) \\ &+ \zeta_1 \ln\left(\frac{w_2}{w_1}\right) + \frac{1}{2}\zeta_2 \left[\ln\left(\frac{w_2}{w_1}\right)\right]^2 \\ &+ \sum_{j=1}^{3} \gamma_j \ln\left(\frac{z_j}{z}\right)\ln\left(\frac{w_2}{w_1}\right),\end{aligned} \tag{F.2}$$

where w_1 and w_2 are input factors, z_1 to z_3 are output quantities (that translate to profits), z denotes total assets, and $\tilde{v} \equiv \frac{v}{w_1 z}$.[3] The normalization by w_1 enforces cost homogeneity, that is, a doubling of input costs doubles total costs; for consistency, the same normalization is applied to profits as well. Normalization by z accounts for the bank scale.

As in He, Ho, and Xu (2020), the inputs w_j, $j = \{1, 2\}$ are as follows. First, w_1 is the average funding price, which is the ratio of interest expenses to the sum of deposits and other interest-bearing liabilities. w_2 denotes the average nonfunding price, which is the ratio of noninterest expenses (e.g., labor and other costs) to total fixed assets. Typically, the bulk of the noninterest expense consists of employee and administrative expenses, where the latter mostly involves marketing costs and client maintenance expenses. Both are closely related to tangible assets such as office buildings and the geographical allocation of employees and branches across regions.[4] For output variables z_j, $j = \{1, 2, 3\}$, we consider deposits z_1, loans z_2, and other income-earning assets z_3. We normalize the three outputs (as well as profits) by total assets (z), again to control for bank scale.[5] Annual reports directly provide data on all of the input and output variables.

Combining Equations (F.1) and (F.2), we can estimate η_i and u_{it}, and, in turn, profit efficiency from Equation (3) of main text. We also estimate two alternative efficiency measures—interest income and noninterest income efficiencies—by replacing profit (π^a_{it}) in Equation (F.1), in turn, with these quantities. Similarly, Equations (3′) of main text, (F.1′), and (F.2) yield cost efficiency.

For simplicity, we denote the frontier efficiency ($\ln \tilde{\pi}_{it}$ or $\ln \tilde{C}_{it}$), plus terms for the controls in Equation (F.1), as $f(x_{it}, \beta)$, where β is the coefficient vector and x_{it} is the vector of all independent variables. Then, we can rewrite Equation (F.1) as

$$y_{it} = f(x_{it}, \beta) - \eta_i - u_{it} + \mu_i + v_{it}, \tag{F.3}$$

where y_{it} is, alternatively, the i'th bank's total profit, interest income, and noninterest income in year t. As

[3] The translog production function is an approximation to the constant elasticity of substitution function via Taylor series expansion, and was introduced in Berndt and Christensen (1973). See Chapter 4 of Kumbhakar, Wang, and Horncastle (2015) and Equation (A1) in Appendix 3 of He, Ho, and Xu (2020).

[4] There are outliers in the distribution of w_1 and w_2, hence we winsorize each cross-section of these variables at the 1%-99% levels.

[5] Note from Equation (F.1) that the inefficiency terms are derived from residuals. Since the largest banks' costs and profits are many times larger than those of the smallest banks, the normalization is desirable to make residual variance scale-invariant.

Equation (3′) of main text shows, if y_{it} denotes total cost, we rewrite Equation (F.1′) as

$$y_{it} = f(x_{it}, \beta) + \eta_i + u_{it} + \mu_i + v_{it}. \tag{F.3′}$$

Now, let us turn to the estimation details. The process involves the following steps:

Step 1: Panel Regression with Random Effects

In this step, we estimate

$$y_{it} = f(x_{it}, \beta) + \alpha_i + \epsilon_{it}, \tag{F.4}$$

where y_{it} denotes the output of firm i in year t, and $f(x_{it}, \beta)$ represents the frontier function that relates inputs to outputs and a series of controls. The term α_i captures time-invariant effects, and ϵ_{it} is the error term. Via this step, we get the fitted residual components $\hat{\alpha}_i$ and $\hat{\epsilon}_{it}$.

Step 2: Jondrow et al. (1982)-Based MLE Decomposition to Estimate Inefficiency

The error term ($\hat{\epsilon}_{it}$) and the time-invariant effect ($\hat{\alpha}_i$) are decomposed as:

$$\hat{\epsilon}_{it} = v_{it} - u_{it}, \quad \hat{\alpha}_i = \mu_i - \eta_i \tag{F.5}$$

where v_{it} and μ_i represent statistical noise, while u_{it} and η_i capture inefficiency components. The distributional assumptions are:

$$v_{it} \sim i.i.d.\ N(0, \sigma_v^2), u_{it} \sim i.i.d.\ N^+(0, \sigma_u^2) \tag{F.6}$$

$$\mu_i \sim i.i.d.\ N(0, \sigma_\mu^2), \eta_i \sim i.i.d.\ N^+(0, \sigma_\eta^2). \tag{F.7}$$

Following Jondrow et al. (1982), the procedure first uses maximum likelihood to estimate the parameters σ_u^2, σ_v^2, σ_η^2, and σ_μ^2 of the distributions in (F.6)-(F.7). The estimated components $\hat{\eta}_i$ and $\hat{u}_{it}$ are then the expectations of η_i and u_{it}, respectively, conditional on the observed residual ϵ_{it}. These conditional expectations are computed using the assumed densities in (F.6)-(F.7).[6] Let $\sigma^2 = \sigma_v^2 + \sigma_u^2$ and $\sigma_\alpha^2 = \sigma_\mu^2 + \sigma_\eta^2$. The expressions are then as follows:

$$\hat{u}_{it} = \frac{\sigma_u \sigma_v}{\sigma} \left[\frac{\phi\left(\frac{\sigma_u}{\sigma_v \sigma} \hat{\epsilon}_{it}\right)}{\Phi\left(-\frac{\sigma_u}{\sigma_v \sigma} \hat{\epsilon}_{it}\right)} - \frac{\sigma_u}{\sigma_v \sigma} \hat{\epsilon}_{it} \right]. \tag{F.8}$$

$$\hat{\eta}_i = \frac{\sigma_\eta \sigma_\mu}{\sigma_\alpha} \left[\frac{\phi\left(\frac{\sigma_\eta}{\sigma_\mu \sigma_\alpha} \hat{\alpha}_i\right)}{\Phi\left(-\frac{\sigma_\eta}{\sigma_\mu \sigma_\alpha} \hat{\alpha}_i\right)} - \frac{\sigma_\eta}{\sigma_\mu \sigma_\alpha} \hat{\alpha}_i \right]. \tag{F.9}$$

[6] See Chapter 10 of Kumbhakar, Wang, and Horncastle (2015). The website https://homepage.ntu.edu.tw/~wangh/book-software.html describes the implementation details for the procedure, including the associated code.

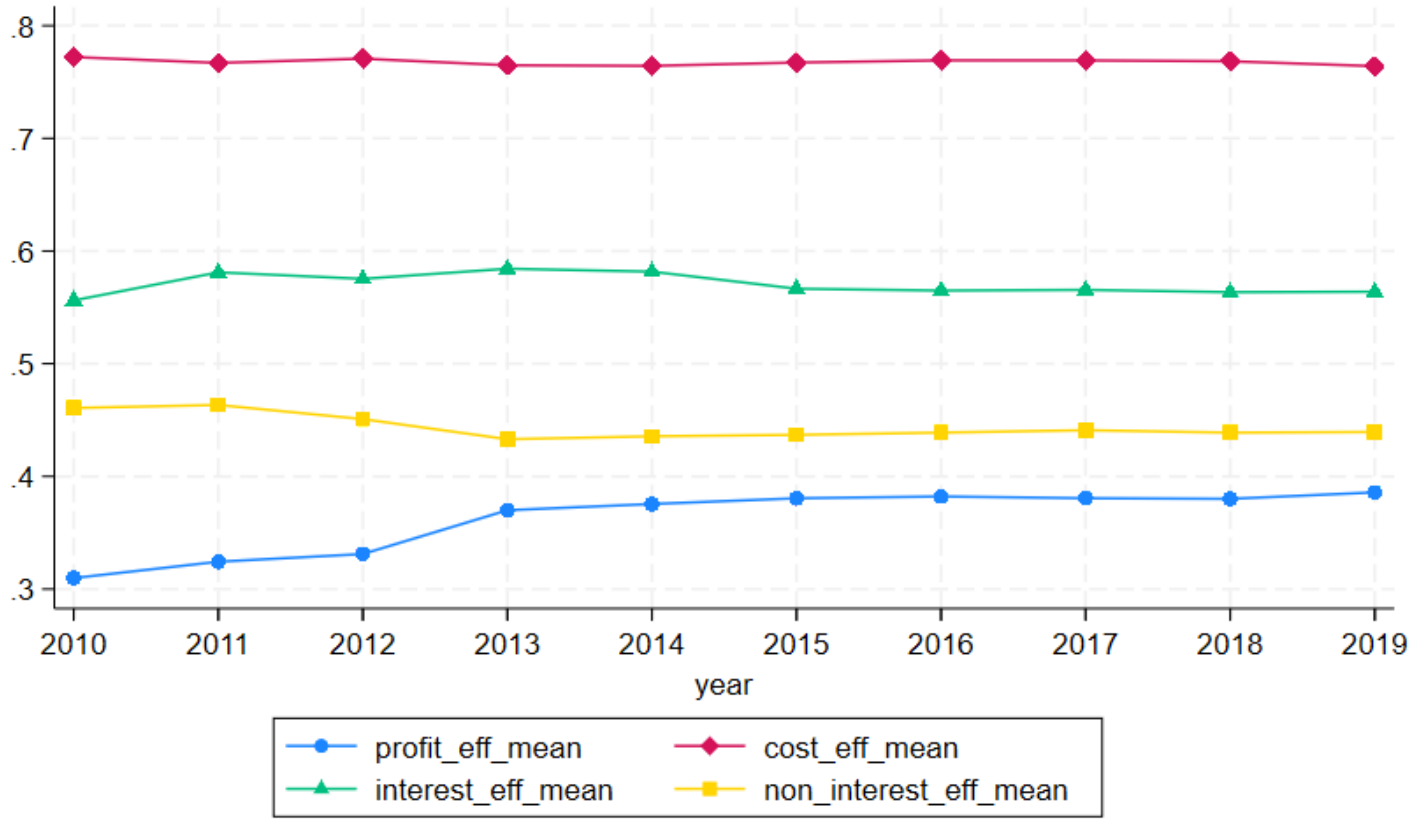
Step 3: Final profit efficiency (*Prof_eff*):

$$\widehat{Prof_eff}_{it} = \exp(-\hat{\eta}_i - \hat{u}_{it}), \tag{F.10}$$

A similar process applies to Equation (F.3′) for estimating *Cost_eff*.

Following the procedure described above, we estimate the average profit, cost, interest income, and non-interest income efficiencies to be 33.7%, 74.9%, 57.0%, and 43.5%, respectively. Figure F.1 plots the yearly mean efficiencies. The results in the figure are consistent with Berger, Hasan, and Zhou (2009) and Jiang, Yao, and Feng (2013), as the relative magnitudes are similar. However, there are some differences in our findings. Specifically, profit efficiency is lower than that reported in their studies. In contrast, average cost efficiency is higher. This suggests that our banks operate in a highly competitive environment, prompting them to focus on reducing operating costs during our sample period. However, the trend of increasing profit efficiency aligns with recent developments in the use of information and communication technologies.

Figure F.1: **Annual means of bank efficiency measures.**
This figure shows trends in banks' four efficiency measures, namely, profit (*Prof_eff*), cost (*Cost_eff*), interest income (*Int_eff*), and noninterest income (*Nint_eff*). The four time series are the yearly cross-sectional averages of the four efficiency measures over the sample period from 2010–2019.



G Pair-wise Correlations between Variables

Table G.1 and G.2 present the Pearson (lower triangle) and Spearman (upper triangle) correlations among the variables used in our analysis. It is constructed for the 1:1 PSM-matched sample, implying that exactly half the sample consists of treated firms, and the other half consists of control firms. The variables are defined as follows: *CIR* is the cost-to-income ratio; *PM* is the profit margin of net income over total revenue; *NIM* is the net interest margin; *dLoan* and *dDeposit* are yearly growth rates of total loans and total deposits; *ROA* is return on assets in percentage. *Prof_eff*, *Cost_eff*, *Int_eff* and *Nint_eff* are the four efficiency measures—profit, interest income, noninterest income, and cost efficiency, estimated by SFA. *Z_score*, *CAR*, *NPL*, and *Liq* are the four risk measures—overall risk, capital asset ratio, nonperforming loan ratio and liquidity ratio. The definitions are as follows: $Z_score = (ROA + CAR)/sd(ROA)$, where *ROA* is the return on assets and $sd(ROA)$ is the standard deviation of *ROA* during the sample period; $CAR = capital/risk\ weighted\ assets$ indicates the extent of bank shareholders' capital covering the risky assets that is reported by each bank according to regulation rules on the weight of risk assessment; *NPL* is the ratio of nonperforming loans to total loans; and *Liq* is the ratio of liquid assets to the sum of deposits and short-term debts. *FinTech* is a dummy variable which equals unity for the treatment group and zero for the control group. The treatment group consists of banks whose number of FinTech patent applications is higher than the 75th percentile of that of all banks in a given year. The control group is the propensity-score-matched sample for the treatment group. *ALL_PAT* is log of the number of a bank's total patent applications in each year. The other variables are the main controls in Equation (1) of main text defined as follows: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; $SIZE$ is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); $PGDP$, POP, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; *NCITY* is the log of the number of cities in which the bank has branches; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city.

Table G.1: **Panel A – Correlations for all variables with the dependent variables.**

This table presents the Pearson (lower triangle) and Spearman (upper triangle) correlations between all variables and the dependent variables after PSM. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	CIR	*PM*	*NIM*	*dLoan*	*dDeposit*	*ROA*	*Prof_eff*	*Cost_eff*	*Int_eff*	*Nint_eff*	*Z_Score*	*CAR*	*NPL*	*Liq*
CIR	–	-0.682***	0.050	0.045	-0.048	-0.382***	-0.264***	-0.214***	-0.084*	-0.135***	-0.266***	-0.198***	0.240***	-0.267***
PM	-0.650***	–	-0.019	-0.081	0.004	0.545***	0.177***	0.132**	0.166***	0.124**	0.431***	0.165***	-0.376***	0.251***
NIM	-0.017	0.037	–	0.267***	0.174***	0.158***	0.123**	-0.070	0.469***	-0.098*	-0.031	-0.013	-0.129***	-0.049
dLoan	0.010	-0.055	0.152***	–	0.472***	0.221***	0.046	-0.074	0.115**	-0.062	-0.011	-0.090*	-0.074	-0.022
dDeposit	-0.091*	0.039	0.109**	0.584***	–	0.345***	0.059	0.033	-0.004	-0.020	0.100*	-0.007	-0.159***	-0.003
ROA	-0.331***	0.545***	0.094*	0.242***	0.357***	–	0.152***	0.068	0.142***	0.173***	0.516***	0.038	-0.298***	0.008
Prof_eff	-0.277***	0.151***	0.106**	0.053	0.074	0.159***	–	-0.173***	0.230***	0.068	-0.084*	-0.090*	-0.016	0.193***
Cost_eff	-0.166***	0.112**	-0.057	-0.113**	-0.020	0.015	-0.165***	–	-0.467***	-0.062	0.118**	0.108**	-0.012	0.079*
Int_eff	-0.023	0.151***	0.494***	0.113**	-0.009	0.144***	0.177***	-0.277***	–	0.067	-0.001	-0.129***	0.022	-0.105**
Nint_eff	-0.165***	0.115**	-0.075*	-0.035	-0.012	0.144***	0.069	-0.072	0.069	–	0.145***	0.089*	0.053	-0.002
Z_Score	-0.250***	0.400***	-0.052	-0.076	0.038	0.429***	-0.073	0.127***	0.003	0.124**	–	0.170***	-0.329***	0.101**
CAR	-0.164***	0.129**	0.046	-0.043	0.035	0.008	-0.056	0.072	-0.140***	0.054	0.151***	–	-0.037	0.064
NPL	0.241***	-0.382***	-0.152***	-0.046	-0.117**	-0.270***	0.022	-0.054	-0.008	0.026	-0.312***	-0.107**	–	-0.215***
Liq	-0.272***	0.131**	0.158***	0.077	0.091*	-0.009	0.122**	0.082*	-0.053	-0.002	0.025	0.109**	-0.134***	–
FinTech	-0.156***	0.133***	0.113**	0.080*	0.146***	0.148***	0.030	0.174***	0.138***	0.158***	0.113**	0.004	-0.110**	0.165***
ALL_PAT	-0.173***	0.076	-0.100**	-0.031	-0.017	0.026	-0.100**	0.085*	0.041	0.214***	0.093*	0.032	-0.014	0.007
GradProp	-0.053	0.019	0.060	0.080*	0.028	0.151***	0.036	-0.036	0.061	0.013	-0.026	0.023	-0.057	-0.072
IntangibleRatio	0.112**	-0.134***	0.002	-0.031	-0.067	-0.113**	0.058	-0.158***	0.138***	-0.046	-0.095*	-0.053	0.054	-0.035
SIZE	-0.142***	0.215***	-0.303***	-0.175***	-0.188***	0.190***	-0.319***	0.110**	-0.015	0.304***	0.354***	0.004	-0.028	-0.192***
BRANCH	-0.047	-0.029	0.036	0.009	0.014	0.048	0.334***	-0.141***	0.124**	-0.124**	-0.162***	-0.223***	0.132***	0.013
PGDP	-0.012	0.020	-0.281***	-0.215***	-0.195***	-0.176***	-0.215***	0.105**	-0.218***	0.168***	0.196***	0.211***	-0.119**	0.106**
POP	-0.048	0.031	-0.189***	-0.161***	-0.156***	0.043	-0.117**	0.046	0.013	0.136***	0.090*	-0.189***	0.075	-0.152***
NFIRM	0.140***	-0.121**	-0.090*	-0.014	0.001	-0.047	0.019	-0.063	-0.017	0.044	0.010	-0.041	-0.072	0.044
NCITY	-0.114**	0.107**	-0.245***	-0.166***	-0.180***	0.081*	-0.350***	0.066	-0.076*	0.310***	0.272***	0.096*	0.001	-0.143***
HHI	-0.194***	0.101*	0.122**	0.073	0.045	0.059	0.147***	-0.017	0.098**	-0.040	-0.087*	-0.014	0.006	-0.062

Table G.2: **Panel B – Correlations for all variables with independent variables.**

This table presents the Pearson (lower triangle) and Spearman (upper triangle) correlations between all variables and the independent variables after PSM. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	FinTech	*ALL_PAT*	*GradProp*	*IntangibleRatio*	*SIZE*	*BRANCH*	*PGDP*	*POP*	*NFIRM*	*NCITY*	*HHI*
CIR	-0.113**	-0.146***	-0.052	0.190***	-0.185***	0.086*	0.008	-0.201***	0.204***	-0.166***	-0.231***
PM	0.145***	0.082*	0.009	-0.192***	0.192***	-0.110**	0.042	0.145***	-0.160***	0.089*	0.128**
NIM	0.052	-0.183***	0.126**	-0.097*	-0.234***	0.168***	-0.421***	-0.189***	-0.063	-0.212***	0.145***
dLoan	0.154***	-0.054	0.137***	-0.105**	-0.108**	0.123**	-0.343***	-0.122**	-0.014	-0.105**	0.085*
dDeposit	0.186***	-0.059	0.097*	-0.127**	-0.123**	0.088*	-0.236***	-0.119**	-0.014	-0.154***	0.043
ROA	0.151***	-0.038	0.075	-0.170***	0.191***	-0.069	-0.157***	0.105**	-0.090*	0.109**	0.062
Prof_eff	0.020	-0.056	-0.044	0.035	-0.281***	0.367***	-0.215***	-0.108**	0.078*	-0.307***	0.045
Cost_eff	0.152***	0.041	0.019	-0.164***	0.067	-0.092*	0.066	0.100*	-0.130***	0.008	0.035
Int_eff	0.076*	-0.012	-0.022	0.083*	-0.049	0.151***	-0.282***	-0.019	-0.005	-0.037	0.102**
Nint_eff	0.148***	0.193***	0.053	-0.028	0.350***	-0.199***	0.177***	0.217***	0.035	0.328***	-0.045
Z_Score	0.110**	0.077*	0.126**	-0.056	0.347***	-0.262***	0.208***	0.196***	-0.024	0.249***	-0.015
CAR	0.008	0.056	0.099**	-0.034	0.114**	-0.246***	0.185***	-0.006	-0.089*	0.140***	0.055
NPL	-0.130***	-0.036	-0.119**	0.101**	-0.033	0.107**	-0.108**	0.038	-0.050	0.049	0.017
Liq	0.094*	-0.058	-0.107**	-0.008	-0.194***	0.078*	0.151***	-0.102**	0.074	-0.226***	-0.046
FinTech	–	0.432***	-0.016	-0.052	0.033	-0.020	0.026	0.012	0.037	0.016	-0.068
ALL_PAT	0.475***	–	0.047	-0.053	0.431***	-0.300***	0.304***	0.254***	0.091*	0.355***	-0.118**
GradProp	0.046	0.067	–	-0.163***	0.193***	-0.451***	0.061	-0.002	0.065	0.251***	0.018
IntangibleRatio	-0.080*	-0.112**	-0.043	–	-0.230***	0.032	0.150***	-0.287***	0.154***	-0.093*	-0.199***
SIZE	0.053	0.468***	0.016	-0.123**	–	-0.555***	0.325***	0.614***	0.002	0.806***	-0.033
BRANCH	-0.039	-0.262***	-0.194***	0.166***	-0.417***	–	-0.544***	-0.132***	-0.016	-0.598***	0.136***
PGDP	0.022	0.297***	0.039	-0.028	0.338***	-0.423***	–	0.186***	0.260***	0.235***	-0.495***
POP	-0.006	0.220***	-0.078*	-0.074	0.537***	0.054	-0.049	–	-0.147***	0.448***	0.165***
NFIRM	0.012	0.081*	0.073	0.052	0.035	-0.034	0.301***	-0.001	–	-0.118**	-0.772***
NCITY	0.045	0.377***	-0.006	-0.073	0.812***	-0.507***	0.238***	0.376***	-0.107**	–	0.079*
HHI	-0.047	-0.101**	-0.041	-0.065	-0.121**	0.201***	-0.555***	0.129***	-0.695***	0.007	–

Table G.1 shows that the cost-to-income ratio (*CIR*) has a strong negative Pearson correlation with profit margin (*PM*) (-0.650) and return on assets (*ROA*) (-0.331), which is intuitive in that banks with lower operational costs relative to income tend to be more profitable. The efficiency of interest income (*Int_eff*) is highly and positively correlated with the net interest margin (*NIM*) for the case of Pearson correlation (0.494), reflecting that efficient generation of interest income supports a bank's lending activity. The nonperforming loan ratio (*NPL*) is negatively correlated with both *PM* (-0.382) and *Z_Score* (-0.312), confirming that higher credit risk erodes profitability and overall stability.

FinTech's correlations with performance measures in Table G.1 are consistently in line with our expectations. For example, this variable is negatively correlated with *CIR* (-0.156) and *NPL* (-0.110), and positively correlated with *Cost_eff* (0.174), *Nint_eff* (0.158), and *Liq* (0.165). These patterns imply that FinTech innovation improves operational efficiency and reduces risk. Notably, the overall patent count (*ALL_PAT*) shows a weaker and less consistent correlation with these measures. For example, this count is positively correlated with *Nint_eff* (0.214) but negatively correlated with *Prof_eff* (-0.100), reinforcing that FinTech patents provide unique value distinct from general innovation.

Table G.2 presents the relationships between all variables and the independent variables. The correlation between *FinTech* and *ALL_PAT* in Table G.2 is about 0.475, which is positive but far from perfect. It confirms the necessity of isolating FinTech-specific patents rather than relying on total patent counts. Importantly, *FinTech* has low correlations with most control variables, with the strongest being a modest -0.047 with *HHI*, indicating that FinTech adoption is not strongly driven by observable bank or regional characteristics. This low correlation reduces concerns about confounding factors. Additionally, *ALL_PAT* is positively correlated with *SIZE* (0.468) and *NCITY* (0.377), showing that larger, more geographically diverse banks tend to file more patents overall, but this pattern does not hold for FinTech specifically.

Regarding the control variables, Table G.2 reveals that bank size (*SIZE*) has an extremely high Pearson correlation with the number of cities with branches (*NCITY*) (0.812), indicating that larger banks naturally expand their geographic footprint across more cities (Berger, Hasan, and Zhou 2009, Cai, Xu, and Zeng 2016). *BRANCH* (branches relative to assets) is strongly positively correlated with *Prof_eff* (0.367) but negatively correlated with *SIZE* (-0.417), suggesting that smaller banks may achieve greater profit efficiency through more focused branch allocation. The Herfindahl-Hirschman Index (*HHI*) is strongly negatively correlated with *NFIRM* (-0.695) and *PGDP* (-0.555), meaning more competitive markets (lower *HHI*) are associated with more firms and higher per capita income in a bank's operating region. We address potential collinearity concerns between controls in Appendix H.3.

H Additional Robustness Checks

H.1 Alternative Measure of FinTech Innovation

In this appendix, we conduct additional robustness checks using alternative measures of banks' FinTech innovation to assess the robustness of our main findings.

First, we assign indicators to each quartile of FinTech innovation, rather than considering the banks in the top 75th percentile of all banks as high FinTech firms. We denote quartiles *FinTech_Q*4, *FinTech_Q*3, *FinTech_Q*2, *FinTech_Q*1, to represent the 75th–100th, 50th–75th, 25th–50th, and bottom 25th percentiles of FinTech patent numbers in each year, respectively.

Table H.1: **Average financial outcomes, efficiency, and risk within FinTech quartile groups.**
This table reports the average values of each dependent variable (financial outcome, efficiency and risk measures) for different FinTech quartile groups.

Dep.Variables	*FinTech_Q*4	*FinTech_Q*3	*FinTech_Q*2	*FinTech_Q*1
CIR	0.680	0.728	0.705	0.731
PM	0.188	0.173	0.179	0.168
NIM	0.087	0.078	0.064	0.127
dLoan	0.236	0.160	0.414	0.182
dDeposit	0.218	0.138	0.255	0.149
ROA	1.326	1.129	1.245	1.192
Prof_eff	0.361	0.358	0.330	0.356
Cost_eff	0.777	0.756	0.755	0.760
Int_eff	0.592	0.590	0.512	0.553
Nint_eff	0.483	0.442	0.456	0.451
Z_Score	5.527	4.962	4.734	4.930
CAR	0.257	0.258	0.272	0.250
NPL	1.465	1.674	1.720	1.546
Liq	0.279	0.233	0.194	0.230

Table H.1 reports the average values of each dependent variable (financial outcomes, efficiency and risk measures) for different FinTech quartile groups. From Table H.1, we can see that the group with the highest number of FinTech patents (*FinTech_Q*4) demonstrates best performance across financial outcome measures with exception of net interest margin (*NIM*). It is intuitive that *NIM* in *FinTech_Q*4 is lower than that of *FinTech_Q*1 but outperforms *FinTech_Q*2 and *FinTech_Q*3. FinTech-enabled banks use more refined risk-aligned pricing (Jiang, Jørring, and Xu 2023, He et al. 2023). They charge lower spreads to creditworthy borrowers and screen out high-risk applicants who demand higher interest rates. This may slightly compress their net interest margins compared to banks using traditional lending models with bundled risk premiums, especially the banks with extremely low FinTech innovation. We note that *FinTech_Q*4 exhibits higher average efficiency measures compared to the other groups. Additionally, its risk measures are relatively better.

Furthermore, we conduct a panel regression as Equation (1) of main text that includes indicators of each quartile group except *FinTech_Q*1 (used as a baseline). Table H.2 presents the results. We can see that the *FinTech_Q*4 group exhibits the highest outcomes, efficiency and risk measures. The results are consistent with our main findings in Tables 3, 4 and 5 of main text.

Finally, Table H.3 changes the FinTech innovation measure to be $\ln(1 + \#FinTech)$, where $\#FinTech$ is the number of FinTech patent applications filed by a bank in a given year. We re-estimate the baseline regressions of Equation (1) of main text using this measure in full sample. The results remain broadly consistent with our main findings. FinTech innovation continues to be positively associated with bank performance and efficiency and negatively associated with bank risk, with the exception of the Z-score and *CAR*. These results provide further support for the robustness of our main findings.

Table H.2: **The impact of FinTech in quartile groups on banks' financial outcome, efficiency and risk.**

This table reports the impact of FinTech in quartile groups on banks' outcomes, efficiency and risk measures as in Equation (1) of main text (without PSM). *FinTech_Q*4, *FinTech_Q*3 and *FinTech_Q*2 represent the 75th-100th, 50th-75th and 25th-50th percentiles of FinTech patent numbers in each year, respectively. The main control variables are defined as follows: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; $SIZE$ is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); $PGDP$, POP, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; *NCITY* is the log of the number of cities in which the bank has branches; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	CIR	*PM*	*NIM*	*dLoan*	*dDeposit*	*ROA*	*Prof_eff*	*Cost_eff*	*Int_eff*	*Nint_eff*	*Z_score*	*CAR*	*NPL*	*Liq*
*FinTech_Q*4	-0.043***	0.020***	0.018***	0.048***	0.064***	0.120***	0.017*	0.016***	0.031***	0.031***	0.318***	0.016	-0.127***	0.033***
	(0.012)	(0.005)	(0.004)	(0.012)	(0.016)	(0.030)	(0.009)	(0.004)	(0.008)	(0.007)	(0.093)	(0.010)	(0.046)	(0.009)
*FinTech_Q*3	-0.002	0.006	-0.001	0.011	0.020	0.014	-0.003	0.003	-0.008	-0.003	0.080	0.013	-0.027	-0.005
	(0.006)	(0.004)	(0.002)	(0.011)	(0.016)	(0.026)	(0.007)	(0.004)	(0.008)	(0.005)	(0.079)	(0.009)	(0.046)	(0.005)
*FinTech_Q*2	-0.005	0.005	0.002	0.010	0.023	-0.008	0.000	0.002	-0.001	-0.003	0.063	0.010	0.122**	0.007
	(0.008)	(0.005)	(0.003)	(0.015)	(0.018)	(0.034)	(0.007)	(0.004)	(0.009)	(0.006)	(0.104)	(0.011)	(0.054)	(0.005)
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1061	1086	1393	1195	1195	1387	1378	1393	1393	1393	1383	1393	1393	1280
$Adj.R^2$	0.212	0.279	0.133	0.064	0.143	0.272	0.069	0.039	0.057	0.077	0.245	0.047	0.055	0.050

Table H.3: **The impact of FinTech on banks' financial outcome, efficiency and risk.**

This table reports the results of Equation (1) of main text using $\ln(1 + \#FinTech)$ as an alternative measure of banks' FinTech innovation, where *#FinTech* denotes the number of FinTech patent applications filed by a bank in a given year. The main control variables are defined as follows: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; *NCITY* is the log of the number of cities in which the bank has branches; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	CIR	*PM*	*NIM*	*dLoan*	*dDeposit*	*ROA*	*Prof_eff*	*Cost_eff*	*Int_eff*	*Nint_eff*	*Z_score*	*CAR*	*NPL*	*Liq*
$\ln(1 + \#FinTech)$	-0.019***	0.008***	0.004***	0.013***	0.021***	0.037***	0.011**	0.005***	0.006**	0.012***	0.059	0.005	-0.046***	0.010***
	(0.005)	(0.002)	(0.001)	(0.005)	(0.007)	(0.013)	(0.005)	(0.002)	(0.003)	(0.003)	(0.037)	(0.004)	(0.017)	(0.003)
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1061	1086	1393	1195	1195	1387	1388	1393	1393	1393	1383	1393	1393	1393
$Adj.R^2$	0.243	0.307	0.128	0.056	0.135	0.256	0.033	0.030	0.093	0.081	0.263	0.046	0.044	0.039

H.2 PSM Regressions with Gradually Added Controls

We redo the regression analysis presented as in Tables 8 of main text while maintaining the inclusion of FinTech (the key variable of our interest), and sequentially adding each control variable in Equation (1) of main text. The results are displayed sequentially in Tables H.4–H.17, corresponding to the dependent variables of financial outcomes, efficiency and risk measures. The results of all these tables confirm our main findings.

Table H.4: **The impact of FinTech on banks' cost-to-income ratio (*CIR*).**
This table reports the impact of FinTech on banks' cost-to-income ratio by estimating Equation (1) of main text (after PSM) via gradually adding controls: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FinTech	-0.035***	-0.034***	-0.031***	-0.031***	-0.031***	-0.032***	-0.031***	-0.031***	-0.033***	-0.033***
	(0.012)	(0.012)	(0.012)	(0.011)	(0.012)	(0.011)	(0.012)	(0.012)	(0.012)	(0.012)
IntangibleRatio		-0.041**	-0.039**	-0.039**	-0.038**	-0.045**	-0.044**	-0.043**	-0.044**	-0.040**
		(0.018)	(0.019)	(0.019)	(0.019)	(0.018)	(0.019)	(0.019)	(0.021)	(0.020)
GradProp			-0.882***	-0.882***	-0.888***	-0.896***	-0.917***	-0.918***	-0.946***	-0.928***
			(0.189)	(0.189)	(0.189)	(0.189)	(0.193)	(0.193)	(0.198)	(0.198)
SIZE				-0.000	-0.004	-0.010	-0.008	-0.007	-0.014	-0.018
				(0.035)	(0.036)	(0.036)	(0.037)	(0.037)	(0.036)	(0.037)
BRANCH					-0.009	-0.007	-0.011	-0.011	-0.007	-0.009
					(0.011)	(0.011)	(0.012)	(0.012)	(0.013)	(0.013)
PGDP						3.107	3.423	3.048	2.518	1.451
						(2.523)	(2.611)	(2.740)	(2.804)	(3.113)
POP							-0.014**	-0.014**	-0.015**	-0.014**
							(0.006)	(0.006)	(0.006)	(0.006)
NFIRM								0.139	0.097	0.100
								(0.418)	(0.413)	(0.404)
NCITY									0.208	0.202
									(0.136)	(0.138)
HHI										-0.064
										(0.054)
Constant	0.670***	0.674***	0.753***	0.754	0.832	-6.511	-7.225	-6.698	-5.451	-2.782
	(0.017)	(0.017)	(0.023)	(0.673)	(0.690)	(5.979)	(6.205)	(6.226)	(6.359)	(7.263)
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	611	611	601	601	601	601	570	570	570	570
$Adj.R^2$	0.122	0.124	0.169	0.169	0.169	0.171	0.178	0.178	0.185	0.188

Table H.5: **The impact of FinTech on banks' profit margin (*PM*).**

This table reports the impact of FinTech on banks' profit margin by estimating Equation (1) of main text (after PSM) via gradually adding controls: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FinTech	0.013***	0.013***	0.012***	0.012***	0.011***	0.012***	0.012***	0.012***	0.014***	0.014***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
IntangibleRatio		0.012	0.013	0.012	0.016	0.025**	0.026**	0.025**	0.025**	0.022**
		(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	(0.011)	(0.010)
GradProp			0.321***	0.322***	0.304***	0.314***	0.318***	0.320***	0.347***	0.330***
			(0.101)	(0.101)	(0.102)	(0.099)	(0.102)	(0.100)	(0.103)	(0.100)
SIZE				0.005	-0.004	0.004	0.002	0.002	0.008	0.012
				(0.018)	(0.020)	(0.020)	(0.020)	(0.020)	(0.019)	(0.020)
BRANCH					-0.023	-0.025	-0.021	-0.021	-0.025*	-0.023*
					(0.015)	(0.017)	(0.019)	(0.019)	(0.015)	(0.014)
PGDP						-3.925**	-3.511*	-2.823	-2.298	-1.277
						(1.842)	(1.809)	(1.779)	(1.853)	(1.962)
POP							0.008***	0.008***	0.008***	0.008***
							(0.003)	(0.003)	(0.003)	(0.003)
NFIRM								-0.254	-0.211	-0.215
								(0.293)	(0.284)	(0.275)
NCITY									-0.205***	-0.200***
									(0.054)	(0.054)
HHI										0.061*
										(0.032)
Constant	0.209***	0.208***	0.179***	0.075	0.280	9.556**	8.541*	7.571*	6.332	3.778
	(0.014)	(0.014)	(0.018)	(0.352)	(0.395)	(4.388)	(4.323)	(4.160)	(4.356)	(4.615)
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	617	617	607	607	607	607	575	575	575	575
$Adj.R^2$	0.180	0.181	0.211	0.212	0.222	0.238	0.237	0.241	0.275	0.285

Table H.6: **The impact of FinTech on banks' net interest margin (*NIM*).**

This table reports the impact of FinTech on banks' net interest margin by estimating Equation (1) of main text (after PSM) via gradually adding controls: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FinTech	0.010***	0.010***	0.010***	0.011***	0.011***	0.011***	0.010***	0.010***	0.010***	0.010***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
IntangibleRatio		0.006	0.006	0.008	0.011	0.010	0.011	0.010	0.010	0.013
		(0.011)	(0.010)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)
GradProp			-0.005	-0.002	-0.002	-0.002	-0.001	-0.002	-0.002	-0.001
			(0.008)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
SIZE				-0.029	-0.034	-0.034	-0.034	-0.034	-0.034	-0.036
				(0.023)	(0.022)	(0.023)	(0.023)	(0.023)	(0.023)	(0.022)
BRANCH					-0.022**	-0.022*	-0.023*	-0.023*	-0.022*	-0.023**
					(0.011)	(0.011)	(0.012)	(0.012)	(0.012)	(0.012)
PGDP						0.084	0.286	0.558	0.523	-0.297
						(1.800)	(1.825)	(1.744)	(1.716)	(1.740)
POP							-0.005	-0.005	-0.005	-0.005
							(0.004)	(0.004)	(0.004)	(0.004)
NFIRM								-0.101	-0.103	-0.097
								(0.152)	(0.151)	(0.160)
NCITY									0.010	0.005
									(0.050)	(0.050)
HHI										-0.048
										(0.033)
Constant	0.063***	0.063***	0.063***	0.615	0.722*	0.521	0.074	-0.320	-0.244	1.767
	(0.004)	(0.004)	(0.004)	(0.446)	(0.431)	(4.221)	(4.292)	(4.144)	(4.088)	(4.160)
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	692	692	674	674	674	674	642	642	642	642
$Adj.R^2$	0.134	0.134	0.130	0.162	0.176	0.176	0.174	0.175	0.175	0.183

Table H.7: **The impact of FinTech on banks' yearly change of total loans (*dLoan*).**

This table reports the impact of FinTech on banks' yearly change of total loans by estimating Equation (1) of main text (after PSM) via gradually adding controls: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FinTech	0.042***	0.043***	0.050***	0.050***	0.051***	0.051***	0.049***	0.048***	0.047***	0.047***
	(0.014)	(0.015)	(0.013)	(0.013)	(0.014)	(0.014)	(0.012)	(0.012)	(0.012)	(0.011)
IntangibleRatio		-0.151	-0.147	-0.149	-0.150	-0.124	-0.111	-0.108	-0.107	-0.117
		(0.111)	(0.105)	(0.103)	(0.104)	(0.104)	(0.087)	(0.086)	(0.083)	(0.082)
GradProp			-0.206***	-0.201***	-0.198***	-0.194***	-0.189***	-0.187***	-0.189***	-0.190***
			(0.013)	(0.026)	(0.024)	(0.026)	(0.025)	(0.026)	(0.026)	(0.026)
SIZE				0.022	0.031	0.046	0.056	0.055	0.049	0.054
				(0.088)	(0.088)	(0.092)	(0.094)	(0.094)	(0.094)	(0.095)
BRANCH					0.025	0.020	0.019	0.019	0.023	0.028
					(0.054)	(0.056)	(0.049)	(0.049)	(0.050)	(0.049)
PGDP						-8.692	-7.917	-8.857	-9.382	-6.992
						(6.319)	(6.431)	(6.644)	(6.646)	(6.257)
POP							-0.047	-0.047	-0.047	-0.047
							(0.072)	(0.072)	(0.072)	(0.072)
NFIRM								0.350	0.334	0.343
								(0.522)	(0.528)	(0.523)
NCITY									0.216	0.229
									(0.138)	(0.139)
HHI										0.172
										(0.120)
Constant	0.250***	0.262***	0.281***	-0.145	-0.327	20.269	18.499	19.858	20.994	15.052
	(0.029)	(0.030)	(0.030)	(1.686)	(1.696)	(14.828)	(15.052)	(15.187)	(15.263)	(14.216)
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	627	627	612	612	612	612	582	582	582	582
$Adj.R^2$	0.059	0.065	0.106	0.106	0.107	0.115	0.126	0.126	0.130	0.135

Table H.8: **The impact of FinTech on banks' yearly change of total deposits (*dDeposit*).**

This table reports the impact of FinTech on banks' yearly change of total deposits by estimating Equation (1) of main text (after PSM) via gradually adding controls: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FinTech	0.061***	0.062***	0.063***	0.063***	0.061***	0.061***	0.056***	0.055***	0.057***	0.058***
	(0.017)	(0.017)	(0.017)	(0.018)	(0.019)	(0.019)	(0.017)	(0.017)	(0.017)	(0.017)
IntangibleRatio		-0.124	-0.114	-0.117	-0.112	-0.092	-0.069	-0.066	-0.067	-0.083
		(0.161)	(0.149)	(0.148)	(0.154)	(0.152)	(0.131)	(0.131)	(0.135)	(0.132)
GradProp			-0.277***	-0.265***	-0.271***	-0.267***	-0.260***	-0.258***	-0.256***	-0.257***
			(0.020)	(0.032)	(0.030)	(0.031)	(0.031)	(0.032)	(0.034)	(0.033)
SIZE				0.052	0.031	0.043	0.058	0.057	0.063	0.071
				(0.081)	(0.077)	(0.078)	(0.078)	(0.078)	(0.081)	(0.081)
BRANCH					-0.062	-0.066	-0.065	-0.065	-0.069	-0.061
					(0.053)	(0.053)	(0.043)	(0.044)	(0.046)	(0.047)
PGDP						-6.765	-5.329	-6.196	-5.642	-1.472
						(6.528)	(6.970)	(7.455)	(7.723)	(8.358)
POP							-0.085	-0.086	-0.086	-0.086
							(0.054)	(0.054)	(0.054)	(0.054)
NFIRM								0.323	0.339	0.353
								(0.769)	(0.766)	(0.763)
NCITY									-0.229	-0.206
									(0.192)	(0.188)
HHI										0.300**
										(0.144)
Constant	0.266***	0.276***	0.299***	-0.703	-0.246	15.783	12.539	13.792	12.592	2.223
	(0.035)	(0.038)	(0.037)	(1.558)	(1.488)	(15.654)	(16.685)	(17.160)	(17.759)	(19.303)
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	627	627	612	612	612	612	582	582	582	582
$Adj.R^2$	0.145	0.147	0.179	0.181	0.184	0.187	0.208	0.208	0.210	0.219

Table H.9: **The impact of FinTech on banks' return on asset (*ROA*).**

This table reports the impact of FinTech on banks' return on asset by estimating Equation (1) of main text (after PSM) via gradually adding controls: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FinTech	0.139***	0.145***	0.140***	0.119***	0.118***	0.121***	0.116***	0.116***	0.123***	0.123***
	(0.038)	(0.038)	(0.038)	(0.036)	(0.036)	(0.035)	(0.032)	(0.033)	(0.033)	(0.033)
IntangibleRatio		-0.309*	-0.310*	-0.344**	-0.341**	-0.237	-0.205*	-0.207*	-0.204	-0.217*
		(0.157)	(0.157)	(0.148)	(0.154)	(0.147)	(0.116)	(0.118)	(0.127)	(0.126)
GradProp			0.531***	0.480***	0.480***	0.479***	0.479***	0.476***	0.472***	0.468***
			(0.072)	(0.092)	(0.091)	(0.094)	(0.094)	(0.097)	(0.101)	(0.100)
SIZE				0.460**	0.454**	0.491**	0.494**	0.495**	0.501**	0.508**
				(0.219)	(0.227)	(0.215)	(0.208)	(0.210)	(0.201)	(0.201)
BRANCH					-0.026	-0.056	-0.071	-0.071	-0.094	-0.090
					(0.175)	(0.185)	(0.165)	(0.165)	(0.147)	(0.145)
PGDP						-40.980*	-32.134	-30.940	-27.634	-23.614
						(20.863)	(21.824)	(22.638)	(23.275)	(23.931)
POP							-0.198**	-0.197**	-0.196**	-0.197**
							(0.076)	(0.076)	(0.078)	(0.078)
NFIRM								-0.444	-0.333	-0.360
								(1.415)	(1.407)	(1.398)
NCITY									-0.920	-0.897
									(0.664)	(0.659)
HHI										0.233
										(0.325)
Constant	1.187***	1.216***	1.175***	-7.559*	-7.431*	90.420*	70.286	68.560	61.377	51.515
	(0.063)	(0.067)	(0.067)	(4.149)	(4.326)	(48.872)	(51.373)	(52.521)	(54.221)	(56.186)
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	691	691	674	674	674	674	642	642	642	642
$Adj.R^2$	0.270	0.275	0.292	0.342	0.342	0.358	0.374	0.374	0.381	0.382

Table H.10: **The impact of FinTech on banks' profit efficiency (*Prof_eff*).**

This table reports the impact of FinTech on banks' profit efficiency by estimating Equation (1) of main text (after PSM) via gradually adding controls: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FinTech	0.024**	0.024**	0.023**	0.024**	0.024**	0.025**	0.026**	0.027***	0.026**	0.025**
	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
IntangibleRatio		0.039	0.037	0.040	0.041	0.051*	0.048	0.046	0.046	0.043
		(0.026)	(0.026)	(0.026)	(0.025)	(0.030)	(0.030)	(0.029)	(0.030)	(0.030)
GradProp			0.172	0.176	0.176	0.176	0.171	0.169	0.169	0.168
			(0.122)	(0.124)	(0.124)	(0.125)	(0.121)	(0.121)	(0.120)	(0.120)
SIZE				-0.035	-0.037	-0.033	-0.029	-0.028	-0.029	-0.027
				(0.030)	(0.031)	(0.029)	(0.027)	(0.028)	(0.029)	(0.029)
BRANCH					-0.008	-0.011	-0.007	-0.007	-0.003	-0.002
					(0.020)	(0.020)	(0.020)	(0.020)	(0.023)	(0.023)
PGDP						-4.024	-3.553	-2.575	-3.115	-2.239
						(4.510)	(4.518)	(4.928)	(4.703)	(4.700)
POP							0.005	0.006	0.006	0.006
							(0.009)	(0.009)	(0.009)	(0.009)
NFIRM								-0.364	-0.382	-0.388
								(0.655)	(0.660)	(0.659)
NCITY									0.150	0.155
									(0.171)	(0.171)
HHI										0.051
										(0.055)
Constant	0.344***	0.340***	0.325***	0.992*	1.030*	10.638	9.392	7.979	9.151	7.003
	(0.020)	(0.021)	(0.023)	(0.563)	(0.575)	(10.887)	(10.911)	(11.342)	(10.795)	(10.783)
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	692	692	674	674	674	674	642	642	642	642
$Adj.R^2$	0.027	0.029	0.060	0.066	0.066	0.069	0.072	0.074	0.077	0.078

Table H.11: **The impact of FinTech on banks' cost efficiency (*Cost_eff*).**

This table reports the impact of FinTech on banks' cost efficiency by estimating Equation (1) of main text (after PSM) via gradually adding controls: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FinTech	0.021***	0.022***	0.022***	0.021***	0.020***	0.021***	0.023***	0.023***	0.023***	0.023***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
IntangibleRatio		-0.016	-0.015	-0.016	-0.011	-0.010	-0.006	-0.008	-0.009	-0.011
		(0.023)	(0.023)	(0.022)	(0.022)	(0.022)	(0.018)	(0.019)	(0.019)	(0.019)
GradProp			-0.024***	-0.026***	-0.027***	-0.027***	-0.026***	-0.030***	-0.029***	-0.030***
			(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
SIZE				0.016	0.008	0.009	0.010	0.012	0.011	0.013
				(0.016)	(0.014)	(0.015)	(0.016)	(0.015)	(0.015)	(0.015)
BRANCH					-0.037***	-0.037***	-0.034***	-0.034***	-0.033***	-0.033***
					(0.012)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)
PGDP						-0.685	0.683	2.314	2.177	2.802
						(2.040)	(1.963)	(1.986)	(1.951)	(2.116)
POP							-0.011*	-0.011*	-0.011*	-0.011*
							(0.007)	(0.006)	(0.006)	(0.006)
NFIRM								-0.607***	-0.611***	-0.616***
								(0.174)	(0.173)	(0.174)
NCITY									0.038	0.042
									(0.057)	(0.058)
HHI										0.036
										(0.038)
Constant	0.742***	0.743***	0.744***	0.438	0.618**	2.254	-0.993	-3.351	-3.054	-4.586
	(0.013)	(0.013)	(0.013)	(0.309)	(0.269)	(4.834)	(4.633)	(4.636)	(4.557)	(4.955)
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	692	692	674	674	674	674	642	642	642	642
$Adj.R^2$	0.083	0.084	0.091	0.096	0.116	0.116	0.143	0.160	0.160	0.163

Table H.12: **The impact of FinTech on banks' interest income efficiency (*Int_eff*).**

This table reports the impact of FinTech on banks' interest income efficiency by estimating Equation (1) of main text (after PSM) via gradually adding controls: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FinTech	0.028***	0.027***	0.026***	0.026***	0.027***	0.027***	0.025***	0.025***	0.025***	0.025***
	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
IntangibleRatio		0.043	0.043	0.045	0.044	0.048	0.047	0.046	0.046	0.048
		(0.052)	(0.051)	(0.050)	(0.050)	(0.051)	(0.049)	(0.049)	(0.048)	(0.048)
GradProp			0.106**	0.108**	0.108**	0.108**	0.108***	0.107**	0.107**	0.108**
			(0.041)	(0.042)	(0.042)	(0.042)	(0.041)	(0.041)	(0.041)	(0.042)
SIZE				-0.021	-0.020	-0.019	-0.020	-0.020	-0.019	-0.021
				(0.018)	(0.019)	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)
BRANCH					0.002	0.001	-0.004	-0.004	-0.004	-0.005
					(0.023)	(0.024)	(0.023)	(0.023)	(0.024)	(0.025)
PGDP						-1.542	-2.223	-1.890	-1.800	-2.401
						(4.639)	(4.776)	(4.676)	(4.598)	(4.791)
POP							-0.001	-0.001	-0.001	-0.001
							(0.012)	(0.012)	(0.012)	(0.012)
NFIRM								-0.124	-0.121	-0.117
								(0.536)	(0.537)	(0.544)
NCITY									-0.025	-0.028
									(0.138)	(0.137)
HHI										-0.035
										(0.057)
Constant	0.570***	0.566***	0.559***	0.955***	0.945***	4.626	6.295	5.813	5.618	7.093
	(0.019)	(0.021)	(0.021)	(0.336)	(0.360)	(11.016)	(11.338)	(11.066)	(10.898)	(11.294)
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	692	692	674	674	674	674	642	642	642	642
$Adj.R^2$	0.057	0.060	0.080	0.082	0.083	0.083	0.083	0.083	0.084	0.084

Table H.13: **The impact of FinTech on banks' non-interest income efficiency (*Nint_eff*).**

This table reports the impact of FinTech on banks' non-interest income efficiency by estimating Equation (1) of main text (after PSM) via gradually adding controls: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FinTech	0.038***	0.038***	0.039***	0.040***	0.041***	0.041***	0.041***	0.041***	0.042***	0.042***
	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
IntangibleRatio		0.008	0.007	0.010	0.007	0.016	0.018	0.017	0.018	0.012
		(0.022)	(0.022)	(0.020)	(0.020)	(0.021)	(0.022)	(0.023)	(0.022)	(0.022)
GradProp			0.016	0.020	0.020	0.020	0.020	0.019	0.019	0.017
			(0.016)	(0.018)	(0.019)	(0.019)	(0.018)	(0.018)	(0.018)	(0.018)
SIZE				-0.037	-0.032	-0.029	-0.030	-0.030	-0.029	-0.026
				(0.026)	(0.025)	(0.024)	(0.025)	(0.025)	(0.024)	(0.023)
BRANCH					0.019	0.017	0.018	0.018	0.015	0.017
					(0.012)	(0.012)	(0.013)	(0.013)	(0.012)	(0.012)
PGDP						-3.475	-3.436	-3.105	-2.790	-0.898
						(2.699)	(2.773)	(2.959)	(2.939)	(2.845)
POP							0.009	0.010	0.010	0.009
							(0.011)	(0.011)	(0.011)	(0.011)
NFIRM								-0.123	-0.113	-0.125
								(0.359)	(0.362)	(0.360)
NCITY									-0.087	-0.077
									(0.088)	(0.089)
HHI										0.110***
										(0.036)
Constant	0.431***	0.430***	0.431***	1.124**	1.030**	9.328	9.209	8.730	8.047	3.406
	(0.020)	(0.020)	(0.020)	(0.493)	(0.485)	(6.472)	(6.642)	(6.809)	(6.780)	(6.577)
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	692	692	674	674	674	674	642	642	642	642
$Adj.R^2$	0.073	0.073	0.075	0.086	0.089	0.093	0.095	0.095	0.098	0.107

Table H.14: **The impact of FinTech on banks' risk measure of Z-score (*Z_score*).**

This table reports the impact of FinTech on banks' risk measure of Z-score by estimating Equation (1) of main text (after PSM) via gradually adding controls: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FinTech	0.214***	0.225***	0.234***	0.211**	0.193**	0.196**	0.208**	0.209**	0.224***	0.224***
	(0.081)	(0.081)	(0.084)	(0.084)	(0.082)	(0.082)	(0.081)	(0.083)	(0.082)	(0.082)
IntangibleRatio		-0.633	-0.614	-0.652*	-0.570	-0.458	-0.376	-0.381	-0.375	-0.402
		(0.396)	(0.380)	(0.362)	(0.363)	(0.353)	(0.289)	(0.294)	(0.315)	(0.313)
GradProp			0.404*	0.347	0.340	0.340	0.319	0.313	0.305	0.297
			(0.232)	(0.230)	(0.221)	(0.223)	(0.214)	(0.218)	(0.224)	(0.221)
SIZE				0.517**	0.373*	0.412*	0.444**	0.447**	0.458*	0.474**
				(0.207)	(0.196)	(0.211)	(0.211)	(0.215)	(0.232)	(0.238)
BRANCH					-0.652**	-0.684***	-0.726***	-0.726***	-0.772***	-0.763***
					(0.251)	(0.262)	(0.248)	(0.248)	(0.231)	(0.234)
PGDP						-44.107	-29.543	-27.075	-20.499	-12.038
						(39.247)	(39.749)	(41.956)	(41.657)	(41.775)
POP							-0.357***	-0.356***	-0.354***	-0.355***
							(0.084)	(0.083)	(0.085)	(0.085)
NFIRM								-0.918	-0.697	-0.754
								(5.304)	(5.338)	(5.332)
NCITY									-1.830	-1.783
									(1.332)	(1.339)
HHI										0.491
										(0.767)
Constant	5.322***	5.382***	5.349***	-4.465	-1.261	104.061	70.555	66.987	52.699	31.944
	(0.187)	(0.197)	(0.199)	(3.900)	(3.715)	(92.594)	(93.826)	(95.850)	(95.202)	(95.086)
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	690	690	673	673	673	673	641	641	641	641
$Adj.R^2$	0.401	0.404	0.408	0.417	0.429	0.431	0.435	0.435	0.439	0.439

Table H.15: **The impact of FinTech on banks' risk measure of capital adequacy ratio (*CAR*).**

This table reports the impact of FinTech on banks' risk measure of capital adequacy ratio by estimating Equation (1) of main text (after PSM) via gradually adding controls: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FinTech	0.006	0.006	0.009	0.006	0.005	0.005	0.008	0.008	0.008	0.008
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
IntangibleRatio		0.014	0.014	0.008	0.013	0.029	0.030	0.029	0.028	0.026
		(0.024)	(0.024)	(0.025)	(0.026)	(0.028)	(0.030)	(0.030)	(0.030)	(0.030)
GradProp			-0.005	-0.014	-0.014	-0.014	-0.017	-0.019*	-0.019*	-0.020*
			(0.010)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
SIZE				0.079***	0.070***	0.075***	0.078***	0.079***	0.079***	0.081***
				(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
BRANCH					-0.040	-0.044	-0.041	-0.041	-0.041	-0.040
					(0.042)	(0.040)	(0.039)	(0.039)	(0.040)	(0.041)
PGDP						-6.183	-6.372	-5.303	-5.344	-4.508
						(4.904)	(4.955)	(5.243)	(5.215)	(5.125)
POP							0.016**	0.016**	0.016**	0.016**
							(0.008)	(0.008)	(0.008)	(0.008)
NFIRM								-0.398	-0.399	-0.404
								(0.523)	(0.529)	(0.527)
NCITY									0.011	0.016
									(0.189)	(0.189)
HHI										0.049
										(0.109)
Constant	0.301***	0.300***	0.299***	-1.192**	-0.996**	13.767	14.074	12.529	12.618	10.566
	(0.019)	(0.020)	(0.020)	(0.488)	(0.491)	(11.580)	(11.711)	(12.023)	(11.990)	(11.702)
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	692	692	674	674	674	674	642	642	642	642
$Adj.R^2$	0.055	0.055	0.062	0.082	0.086	0.091	0.094	0.095	0.095	0.096

Table H.16: **The impact of FinTech on banks' non-performing loan ratio (*NPL*).**

This table reports the impact of FinTech on banks' risk measure of non-performing loan ratio by estimating Equation (1) of main text (after PSM) via gradually adding controls: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FinTech	-0.167**	-0.155**	-0.171**	-0.166**	-0.160**	-0.162**	-0.174**	-0.175***	-0.181***	-0.180***
	(0.069)	(0.063)	(0.066)	(0.065)	(0.064)	(0.065)	(0.067)	(0.067)	(0.067)	(0.067)
IntangibleRatio		-0.730**	-0.733***	-0.725***	-0.754***	-0.836***	-0.849***	-0.846***	-0.848***	-0.812***
		(0.283)	(0.279)	(0.276)	(0.273)	(0.260)	(0.247)	(0.247)	(0.256)	(0.248)
GradProp			0.402***	0.413***	0.415***	0.416***	0.411***	0.416***	0.419***	0.430***
			(0.095)	(0.096)	(0.098)	(0.098)	(0.096)	(0.093)	(0.092)	(0.096)
SIZE				-0.101	-0.051	-0.080	-0.043	-0.045	-0.050	-0.072
				(0.092)	(0.091)	(0.093)	(0.086)	(0.087)	(0.087)	(0.093)
BRANCH					0.227**	0.251**	0.186	0.186	0.205*	0.193*
					(0.102)	(0.111)	(0.113)	(0.113)	(0.112)	(0.109)
PGDP						32.571*	27.641	25.638	23.013	11.653
						(19.463)	(18.312)	(17.890)	(18.214)	(18.515)
POP							-0.153***	-0.154***	-0.155***	-0.153***
							(0.048)	(0.047)	(0.048)	(0.048)
NFIRM								0.745	0.657	0.733
								(2.800)	(2.784)	(2.723)
NCITY									0.730	0.668
									(0.716)	(0.724)
HHI										-0.659
										(0.458)
Constant	1.577***	1.646***	1.620***	3.541**	2.427	-75.345	-63.269	-60.374	-54.670	-26.802
	(0.128)	(0.137)	(0.139)	(1.714)	(1.711)	(46.683)	(44.164)	(42.642)	(43.251)	(43.703)
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	692	692	674	674	674	674	642	642	642	642
$Adj.R^2$	0.068	0.083	0.090	0.092	0.097	0.102	0.112	0.112	0.114	0.119

Table H.17: **The impact of FinTech on banks' risk measure of liquidity ratio (*Liq*).**

This table reports the impact of FinTech on banks' risk measure of liquidity ratio by estimating Equation (1) of main text (after PSM) via gradually adding controls: *IntangibleRatio* is the ratio of banks' intangible asset to total asset; *GradProp* is the proportion of employees holding postgraduate degrees; *SIZE* is the log of a bank's total assets (in million RMB); *BRANCH* is the ratio of the total number of branches to the total assets (in million RMB); *PGDP*, *POP*, and *NFIRM* are respectively the log of per capita income, the log of the population, and the number of firms in the provinces where banks principally operate; and *HHI* is the Herfindahl-Hirschman Index of bank branches in each bank's principal city. We include dummy variables to control fixed effects of individual banks and years, as the row of TWFE indicates. Standard errors clustered at the bank level are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
FinTech	0.028***	0.028***	0.026***	0.024***	0.023***	0.023***	0.027***	0.027***	0.026***	0.026***
	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.007)
IntangibleRatio		0.009	0.009	0.007	0.013	0.019	0.018	0.017	0.016	0.019
		(0.015)	(0.014)	(0.015)	(0.015)	(0.020)	(0.021)	(0.021)	(0.022)	(0.022)
GradProp			-0.043*	-0.047**	-0.047**	-0.047**	-0.049**	-0.051***	-0.050***	-0.049**
			(0.023)	(0.020)	(0.020)	(0.020)	(0.019)	(0.019)	(0.019)	(0.019)
SIZE				0.032	0.021	0.023	0.023	0.024	0.023	0.021
				(0.031)	(0.029)	(0.030)	(0.031)	(0.031)	(0.029)	(0.028)
BRANCH					-0.047***	-0.049***	-0.045**	-0.045**	-0.041**	-0.042**
					(0.018)	(0.018)	(0.018)	(0.018)	(0.016)	(0.017)
PGDP						-2.370	-2.195	-1.358	-1.969	-2.888
						(3.283)	(3.306)	(3.392)	(3.391)	(3.658)
POP							0.008	0.008	0.008	0.008
							(0.010)	(0.010)	(0.010)	(0.010)
NFIRM								-0.311	-0.332	-0.325
								(0.244)	(0.249)	(0.251)
NCITY									0.170*	0.165
									(0.100)	(0.101)
HHI										-0.053
										(0.049)
Constant	0.140***	0.139***	0.142***	-0.456	-0.224	5.436	4.964	3.755	5.082	7.337
	(0.008)	(0.008)	(0.008)	(0.600)	(0.550)	(7.711)	(7.773)	(7.868)	(7.891)	(8.578)
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	692	692	674	674	674	674	642	642	642	642
$Adj.R^2$	0.052	0.052	0.051	0.057	0.069	0.071	0.078	0.080	0.086	0.088

H.3 Other Robustness Checks

We now perform further robustness analyses. First, Lerner and Seru (2022) emphasize that patent grants observed in the U.S. often undergo a review of more than two years. Thus, considering only the number of granted patents may lead to data truncation. However, the total number of patent applications could overstate the prevalence of innovation activities, as also suggested by Lerner and Seru (2022). To address these concerns, we use the number of granted FinTech patents ($FinTech2$) in place of total patents. Panel A of Table H.18 reports the results, which are consistent with those in Tables 3, 4 and 5 of main text. The findings confirm that our results on the effects of FinTech innovation on banks' outcomes, efficiency and risk measures are robust to concerns regarding potential truncation or inflation of the proxy variable.

The variables of $SIZE$, $BRANCH$, $PGDP$, POP, $NFIRM$, $NCITY$ and HHI are closely related (see Appendix G). To address the potential multicollinearity issues, we employ PCA to extract the principal components of these variables at the bank level. In untabulated results, we find that the first two components (denoted by $Comp1$ and $Comp2$) explain 94% of the total variance among the seven variables. We replace the aforementioned seven variables with $Comp1$ and $Comp2$ in Equation (1) of main text. Panel B of Table H.18 reports the findings, which are consistent with those in Tables 3, 4 and 5 of main text.

According to our method, a bank's efficiency level is the difference between actual bank-specific costs (or profits) and best-practice costs. This metric could be cross-sectionally influenced by different exposures of banks to aggregate trends in costs and profits. To address this issue, we use scaled bank rankings in place of the actual financial outcomes, efficiency and risk measures. Specifically, for each year, banks are ranked so that those with the lowest values receive the lowest scaled rank, and those with the highest values receive the highest rank. We then scale each ranking by the number of banks in each year. This measure has the benefit of being more comparable across time. We report the results in Panel C of Table H.18. The results are similar to those in Tables 3, 4 and 5 of main text.

The development of nonbank financial institutions in China, including Ant Financial Services Group, WeBank of Tencent Company, and Du Xiaoman Financial (or Baidu Financial Services Group), has been a major phenomenon in recent years. These institutions provide financial services (e.g., loans and savings products), and compete with traditional banks. In response, traditional banks are motivated to increase their financial inclusion efforts to compete (e.g., Chauvet and Jacolin 2017, Banna et al. 2022), and the observed impact of FinTech on bank performance measures might be due in part to this competition. To address this concern, we include the Digital Financial Inclusion Index (published by Peking University)[7] as a control variable. Panel D of Table H.18 shows that our main findings are unchanged.

[7]The index is available at https://tinyurl.com/yvcekbya.

Table H.18: **Additional robustness tests.**

This table reports results from further robustness checks for Table 8 of main text. In Panel A, we substitute the number of FinTech patent applications (*FinTech*) with the number of grants (*FinTech*2) to address the concern of whether the FinTech proxy is truncated or inflated. In Panel B, we address the possible issue of multicollinearity among *SIZE*, *BRANCH*, *PGDP*, *POP*, *NFIRM*, *NCITY* and *HHI*. We employ principal component analysis (PCA) and extract the first two components (*Comp*1 and *Comp*2) across these variables; these components explain 94% of the total variance. We check the robustness by replacing $(SIZE, BRANCH, PGDP, POP, NFIRM, NCITY, HHI)$ with $(Comp1, Comp2)$ in Equation (1) of main text. In Panel C, we use the scaled rankings of banks' outcomes, efficiency and risk measures as the dependent variables. Each indicator is ranked annually across banks and then scaled by the total number of banks in that year, resulting in values ranging from 0 to 1. In Panel D, we address the concern of omitted variable bias and add the Digital Financial Inclusion Index to the controls in Equation (1) of main text. This index (*Inclusion*) measures financial digitization in the province of the bank's headquarters and covers 2011-–2018. The estimated coefficients of other controls are similar to those in Tables 3, 4 and 5 of main text, so we omit them here. Standard errors are in parentheses. Here ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	CIR	*PM*	*NIM*	*dLoan*	*dDeposit*	*ROA*	*Prof_eff*	*Cost_eff*	*Int_eff*	*Nint_eff*	*Z_score*	*CAR*	*NPL*	*Liq*
Panel A: Patent granted as the measurement of FinTech														
*FinTech*2	-0.019***	0.007***	0.002*	0.018***	0.030***	0.027*	0.009**	0.005*	0.007*	0.021***	0.028	0.003	-0.096**	0.011***
	(0.006)	(0.002)	(0.001)	(0.006)	(0.009)	(0.016)	(0.004)	(0.003)	(0.004)	(0.006)	(0.045)	(0.006)	(0.039)	(0.004)
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	569	574	641	582	582	641	641	641	641	641	640	641	641	641
$Adj.R^2$	0.189	0.286	0.168	0.124	0.219	0.368	0.076	0.123	0.071	0.109	0.432	0.097	0.121	0.084
Panel B: PCA														
FinTech	-0.033***	0.013***	0.010***	0.047***	0.057***	0.121***	0.025**	0.023***	0.020**	0.040***	0.199**	0.007	-0.190***	0.023***
	(0.012)	(0.004)	(0.003)	(0.011)	(0.017)	(0.033)	(0.010)	(0.006)	(0.008)	(0.011)	(0.082)	(0.011)	(0.067)	(0.007)
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	570	575	642	582	582	642	642	642	642	642	641	608	642	642
$Adj.R^2$	0.184	0.269	0.165	0.134	0.212	0.318	0.074	0.163	0.084	0.107	0.426	0.084	0.117	0.088
Panel C: Alternative measures of bank efficiency and risk														
FinTech	-0.042*	0.065***	0.036*	0.107***	0.102***	0.067***	0.071***	0.095***	0.031**	0.108***	0.015*	0.020	-0.071***	0.061**
	(0.023)	(0.021)	(0.020)	(0.018)	(0.028)	(0.018)	(0.021)	(0.026)	(0.010)	(0.031)	(0.008)	(0.023)	(0.022)	(0.024)
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	570	575	642	582	582	642	631	642	631	642	641	608	642	642
$Adj.R^2$	0.259	0.250	0.254	0.159	0.261	0.375	0.084	0.105	0.183	0.108	0.374	0.113	0.112	0.116

To be continued...

Table H.18: **Additional robustness tests.**-Continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	CIR	*PM*	*NIM*	*dLoan*	*dDeposit*	*ROA*	*Prof_eff*	*Cost_eff*	*Int_eff*	*Nint_eff*	*Z_score*	*CAR*	*NPL*	*Liq*
Panel D: Additional control variable of financial inclusion														
FinTech	-0.033***	0.014***	0.010***	0.045***	0.057***	0.121***	0.024**	0.023***	0.020**	0.040***	0.199**	0.007	-0.189***	0.023***
	(0.012)	(0.004)	(0.003)	(0.011)	(0.017)	(0.032)	(0.010)	(0.006)	(0.008)	(0.011)	(0.081)	(0.012)	(0.067)	(0.007)
Inclusion	-0.033	-0.071***	-0.081***	-0.243**	-0.062	-0.771***	0.008	0.051	-0.113**	-0.022	-0.416	0.066	0.307	-0.006
	(0.047)	(0.022)	(0.028)	(0.100)	(0.107)	(0.232)	(0.050)	(0.037)	(0.053)	(0.033)	(0.498)	(0.090)	(0.511)	(0.045)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TWFE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	570	575	642	582	582	642	642	642	642	642	641	608	642	642
$Adj.R^2$	0.188	0.302	0.214	0.150	0.219	0.400	0.078	0.169	0.094	0.108	0.440	0.102	0.121	0.088

I The Use of Youden's Index in FinTech Patent Classification

As a robustness check on classifying patents as FinTech-related, we employ Youden's index, a diagnostic measure that jointly considers sensitivity and specificity. This metric helps identify the optimal cut-off point that maximizes classification performance and mitigates biases from false positive rates (FPR) or false negative rates (FNR). The relevant concepts are defined as follows:

- TP: the number of true positives, referring to patents manually labeled as FinTech that are correctly classified as FinTech by the machine learning-based method,
- FN: the number of false negatives, referring to patents manually labeled as FinTech that are incorrectly classified as non-FinTech,
- TN: the number of true negatives, referring to patents manually labeled as non-FinTech that are correctly classified as non-FinTech,
- FP: the number of false positives, referring to patents manually labeled as non-FinTech that are incorrectly classified as FinTech.

Based on these definitions, the key metrics used to assess classification performance are specified as follows:

$$\text{Sensitivity (True Positive Rate, TPR)} = \frac{TP}{TP + FN}, \tag{I.1}$$

$$\text{Specificity (True Negative Rate, TNR)} = \frac{TN}{TN + FP}, \tag{I.2}$$

$$\text{False Positive Rate (FPR)} = \frac{FP}{TN + FP} = 1 - \text{Specificity}, \tag{I.3}$$

In practice, simultaneously achieving high sensitivity and high specificity is challenging, as both measures depend on the choice of threshold in classification. Lowering the threshold increases sensitivity by classifying more patents as FinTech, thereby reducing false negatives, but this typically comes at the cost of reduced specificity and an elevated false positive rate. Conversely, raising the threshold improves specificity by correctly excluding more non-FinTech patents, yet this gain is offset by a decline in sensitivity, as some true FinTech patents are overlooked. To address this trade-off, Youden's index provides a criterion for selecting the threshold that optimally balances sensitivity and specificity. Specifically, the index, denoted as J, is formally defined as:

$$J = \text{Sensitivity} + \text{Specificity} - 1 = TPR - FPR, \tag{I.4}$$

where TPR represents the true positive rate and FPR represents the false positive rate.

By construction, J ranges from -1 to 1. A value of $J = 1$ indicates perfect classification, $J = 0$ reflects no

discriminative power (equivalent to random guessing), and negative values imply performance worse than random classification, typically caused by systematic misclassification.

The threshold that maximizes J is considered optimal, achieving the best balance between sensitivity and specificity. To determine the optimal classification threshold for Youden's index, we follow the procedure outlined in Chen, Wu, and Yang (2019), and randomly select 500 patents from the raw sample of 26,168 patents with G & H classes (IPC). For the labeling task, we recruited three postgraduate students with academic backgrounds in finance and computer science. To ensure a shared understanding of what constitutes a FinTech patent, all coders carefully reviewed the definition and related discussions in Chen, Wu, and Yang (2019). For example, a patent is a FinTech one if it ultimately consists of digital computing technologies that are applied (or that will likely be applied in the future) to financial services.

Consistent with the application of *doc2vec* in the reference documents (FT, BB, and CS) (Appendix C.1), the coders also briefly read these documents to better understand the implications and differences between FT, BB, and CS. Then, the three coders labeled whether each patent is FinTech independently. We measure the consistency of their labeling results using Fleiss's Kappa coefficient, which quantifies the degree of agreement among three coders (Fleiss 1975, Cao, Li, and Zhang 2025). This coefficient is 0.8556, indicating that the three pairs are all at the level of almost perfect agreement (the scale generally used is Poor: < 0; mild: 0-0.2; Fair: 0.21-0.4; Moderate: 0.41-0.6; Substantial: 0.61-0.8; Almost perfect: 0.81-1.0). Among the 500 patents, 463 were labeled consistently by all three coders. We use the majority voting rule to determine whether the 37 remaining ones are FinTech. Ultimately, we acquired 500 labeled patents which include 103 FinTech ones and 397 others.

We apply the identification method on these patents, and calculate Youden's index across a range of values for cosine similarity between a patent and its FT reference. We then identify the threshold that maximizes Youden's index, which is 0.0468. This procedure yields Precision: 0.6202, Recall: 0.7767, F1 Score: 0.6897, and AUC: 0.8295. Compared to the results from *text-embedding-ada-002* (Precision: 0.6241, Recall: 0.8058, F1-Score: 0.7034, AUC: 0.8584), we see that the performance of *doc2vec* is similar to *text-embedding-ada-002*, although the F1-score and AUC indicate marginally better performance of the latter method. We apply Youden's index based on *doc2vec* in the three-rule method described in Section 4.4 to classify the full set of patents, but demonstrate robustness of the main conclusions to *text-embedding-ada-002* in Table 13 of main text.

References

Akcigit, U., J. Grigsby, T. Nicholas, and S. Stantcheva (2022). Taxation and innovation in the twentieth century, *The Quarterly Journal of Economics* **137**(1): 329–385.

Banna, H., M. R. Alam, R. Ahmad, and N. M. Sari (2022). Does financial inclusion drive the Islamic banking efficiency? A post-financial crisis analysis, *Singapore Economic Review* **67**(01): 135–160.

Bellstam, G., S. Bhagat, and J. A. Cookson (2021). A text-based analysis of corporate innovation, *Management Science* **67**(7): 4004–4031.

Berger, A. N. and L. J. Mester (1997). Inside the black box: What explains differences in the efficiencies of financial institutions?, *Journal of Banking and Finance* **21**(7): 895–947.

Berger, A. N., I. Hasan, and M. Zhou (2009). Bank ownership and efficiency in China: What will happen in the world's largest nation?, *Journal of Banking and Finance* **33**(1): 113–130.

Berndt, E. and L. Christensen (1973). The translog function and the substitution of equipment, structures, and labor in us manufacturing 1929-68, *Journal of Econometrics* **1**(1): 81–113.

Cai, W., F. Xu, and C. Zeng (2016). Geographical diversification and bank performance: Evidence from China, *Economics Letters* **147**: 96–98.

Cao, J., X. Li, and L. Zhang (2025). Is relevancy everything? a deep-learning approach to understand the effect of image-text congruence, *Management Science* **71**(12): 10579–10602.

Chauvet, L. and L. Jacolin (2017). Financial inclusion, bank concentration, and firm performance, *World Development* **97**: 1–13.

Chen, M. A., Q. Wu, and B. Yang (2019). How valuable is FinTech innovation?, *Review of Financial Studies* **32**(5): 2062–2106.

Fleiss, J. (1975). Measuring agreement between two judges on the presence or absence of a trait, *Biometrics* pp. 651–659.

Geng, Z., H. Yan, X. Qiu, and X. Huang (2021). fasthan: A BERT-based multi-task toolkit for Chinese NLP, *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: System Demonstrations*, pp. 99–106.

Guzman, J. and A. Li (2023). Measuring founding strategy, *Management Science* **69**(1): 101–118.

He, D., C.-Y. Ho, and L. Xu (2020). Risk and return of online channel adoption in the banking industry, *Pacific-Basin Finance Journal* **60**: 101268.

He, Z., S. Jiang, D. Xu, and X. Yin (2023). Investing in lending technology: IT spending in banking, *Working Paper 30403*, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w30403*

Hoberg, G. and G. Phillips (2016). Text-based network industries and endogenous product differentiation, *Journal of Political Economy* **124**(5): 1423–1465.

Huber, K. (2021). Are bigger banks better? Firm-level evidence from Germany, *Journal of Political Economy* **129**(7): 2023–2066.

Jiang, C., S. Yao, and G. Feng (2013). Bank ownership, privatization, and performance: Evidence from a transition country, *Journal of Banking and Finance* **37**(9): 3364–3372.

Jiang, S., A. Jørring, and D. Xu (2023). Bank technology adoption and loan production in the U.S. mortgage market, *Available at SSRN: https://ssrn.com/abstract=4455105* .

Jondrow, J., C. Knox Lovell, I. S. Materov, and P. Schmidt (1982). On the estimation of technical inefficiency in the stochastic frontier production function model, *Journal of Econometrics* **19**(2): 233–238.

Kumbhakar, S. C., H.-J. Wang, and A. P. Horncastle (2015). *A Practitioner's Guide to Stochastic Frontier Analysis Using Stata*, Cambridge University Press, New York.

Le, Q. V. and T. Mikolov (2014). Distributed representations of sentences and documents, *Proceedings of the 31st International Conference on Machine Learning*, Beijing.

Lerner, J. and A. Seru (2022). The use and misuse of patent data: Issues for finance and beyond, *Review of Financial Studies* **35**(6): 2667–2704.

Li, K., F. Mai, R. Shen, and X. Yan (2021). Measuring corporate culture using machine learning, *Review of Financial Studies*

34(7): 3265–3315.

Mikolov, T., K. Chen, G. Corrado, and J. Dean (2013). Efficient estimation of word representations in vector space, *arXiv preprint* **arXiv:1301.3781**.

Park, J., K. Han, and B. Lee (2023). Green cloud? An empirical analysis of cloud computing and energy efficiency, *Management Science* **69**(3): 1639–1664.

Qiu, X., H. Pei, and H. Yan (2020). Multi-criteria Chinese word segmentation with transformer.
URL: *https://arxiv.org/abs/1906.12035v1*