Cutting Operational Costs by Integrating Fintech into Traditional Banking Firms*

Linda Allen[†], Yu Shan[‡], Yi Tang[§] and Alev Yildirim[¶]

February 13, 2022

Abstract

Fintech firms mobilize information technology to provide intermediation services using a broker methodology, whereas dealer banks intermediate using leveraged balance sheets. The integration of Fintech into banking may reduce the unit cost of intermediation by shifting the production function from dealer to broker. A "Fintech score" is derived using nonlinear and machine learning algorithms that show on-balance sheet lending for low Fintech score dealer banks versus securitization, brokered deposits, and non-interest income for high score, broker banks. Using Data Envelopment and Stochastic Cost Frontier Analyses, we find that banks with higher Fintech scores are more operationally efficient and resilient in crises.

JEL classification: G20, G21, G23, G24

Keywords: Fintech, operational costs, broker, dealer, market making, match making

*We are grateful to Yakov Amihud, Tetyana Balyuk, Allen Berger, Ayan Bhattacharya, John Paul Broussard, Greg Buchak, Georgia Bush, Curi Claudia, Iftekhar Hasan, Sonali Hazarika, Kose John, Troy Kravitz, Xioayang Li, Roni Michaely, Lin Peng, Anjan Thakor, Anthony Saunders, Stefan Scharnowski, Jun Wang, Yajun Wang, Dexin Zhou, and other participants at the 2022 American Finance Association, 2021 Southern Finance Association, 2021 China International Risk Forum, 2021 World Finance Conference, 2020 Financial Management Association Annual Meeting, 2020 International Risk Management Conference, Fordham University Fintech Conference, Villanova Webinars in Financial Intermediation, and Baruch College Finance Seminars. All remaining errors are our own.

[†]William F. Aldinger Chaired Professor of Banking and Finance, Bert W. Wasserman Department of Economics and Finance, Zicklin School of Business, Baruch College, One Bernard Baruch Way, B10-225, New York, NY 10010, linda.allen@baruch.cuny.edu and acknowledges support of the Wasserman Foundation.

[‡]Corresponding author. Assistant Professor of Finance, Department of Finance, John Molson School of Business, Concordia University, 1450 Guy Street, MB 12.217, Montreal, QC, H3H 0A1, Canada, yu.shan@concordia.ca.

[§]Robert B. McKeon Endowed Chair in Business, Gabelli School of Business, Fordham University, 45 Columbus Avenue, New York, NY, 10023, ytang@fordham.edu.

[¶]Assistant Professor of Finance, Southern Connecticut State University, 501 Crescent Street, New Haven, CT 06515, yildirima1@southernct.edu.

1. Introduction

Inefficiency in the banking system appears to be resistant to decades of technological innovation that has modernized financial delivery mechanisms. Philippon (2015) demonstrates that the unit cost of financial intermediation in the U.S. has remained stable (around 2%) for the past 130 years. Bazot (2018) documents a similar anomaly in Germany, France and the U.K. This astonishing fact suggests that over decades of technological development and innovation, there has been little or no gain in cost efficiency in providing basic capital allocation services to the economy despite the substantial amounts financial firms spend on technology.[1] This arteriosclerosis in the banking system may be the outcome of a regulatory safety net that shields banks from competition and ossifies inefficiency. Therefore, it should not be surprising that the current wave of Fintech disruption of the bank delivery mechanism originated in the shadow banking sector. Competitive pressures from Fintech disruptors may, therefore, force traditional banks to improve efficiency and reduce operating costs. Indeed, Philippon (2019) finds some indication that quality-adjusted costs of financial intermediation have started to decline in recent years. However, despite early predictions of Fintechs' potential to dislodge the structural inefficiencies in financial intermediation, many Fintech upstarts have been acquired and may have been co-opted by traditional banks protecting their monopoly rents. This paper examines whether operational costs in traditional banks have been impacted by Fintech adoption.

To answer this question, we must more precisely state what form of disruption would be sufficiently transformative to shake up the existing financial delivery mechanism, thereby finally reducing the costs of financial intermediation. Historically, non-disruptive albeit valuable innovations to improve the intermediation **delivery** mechanism have been the focus of bank technology investment. However, all of these IT expenditures are consistent with tra-

[1] Feng and Wu (2018) document the substantial and growing role of technology in traditional banking, estimating that technology has contributed an average of 11.3% to the increase in banking value-added over the 2000 to 2016 period. However, the presence of legacy systems that are internally incompatible has contributed to banks' failure to realize efficiency gains from these expenditures.

ditional balance sheet financial intermediation. For example, ATMs and online banking improved the deposit and loan delivery mechanism without changing the structure of bank balance sheets that transform liquid deposits into illiquid loans. That is, these IT expenditures innovate and expedite, but do not disrupt a basic business model that allows financial institutions to earn economic rents via balance sheet intermediation. We hypothesize that bank IT expenditures designed to modernize delivery mechanisms may have failed to substantially reduce the cost of providing financial services because these innovations do not disrupt traditional banks' fundamental form of intermediation and capital allocation.

The basic function of financial intermediaries is to bring together providers and users of capital as counterparties in financial transactions. Traditional financial intermediaries (such as banks) perform this function by employing a "dealer" technology that utilizes the institution's leveraged balance sheet as the fundamental tool of financial intermediation. That is, the financial intermediary obtains cash from providers of capital (e.g., depositors, insurance policyholders, buy-side investors, etc.) by issuing balance sheet liability claims, and then separately transforms those claims into loans and other financial securities held as balance sheet assets. The financial intermediary provides immediacy of liquidity to both counterparties in the transaction by taking a dealer position onto its balance sheet. Despite the ubiquitous presence of technology in banks, this basic business model has been largely unchanged over decades of technological development.

In contrast, Fintech firms act as "brokers." They provide intermediation services by mobilizing information technology to identify both counterparties to each capital transfer, thereby facilitating the exchange (e.g., online marketplace lending, cryptocurrency trading, etc.). For example, Berg et al. (2020), Aggarwal et al. (2020) and Jagtiani and Lemieux (2019) demonstrate how Fintech access to non-traditional credit data and payment mechanisms increases borrowers' cash-flow pledgeability and reduces search costs. Thus, information technology enables Fintech firms to allocate capital without committing substantial balance sheet resources to the intermediated transaction. Therefore, they do not take on the risks and costs of maintaining financial inventories. Instead, a Fintech broker intermediation method uses artificial intelligence to reduce the counterparties' search costs by brokering financial transactions rather than acting as the market maker. We hypothesize that Fintech can disrupt inefficient banking operations only if it enables the bank to perform its capital allocation functions by shifting from reliance on the balance sheet to instead using information technology to match providers and users of financial resources; i.e., a shift from a dealer to a broker model of financial intermediation. In this paper, we examine whether banks that adopt this Fintech broker technology display greater operational cost efficiency as compared to banks relying on traditional balance sheet intermediation.

We contribute to a growing literature about the impact of Fintech adoption on traditional banking institutions (see Allen, Gu, and Jagtiani (2020) for a survey). Many papers focus on a taxonomy of Fintech innovations. For example, Chen et al. (2019) use patent data to identify eight Fintech categories: cybersecurity, mobile transactions, data analytics, blockchain, P2P, robo-advising, IoT and other. Similarly, Crouhy et al. (2021) allocate Fintech innovations to four banking functions: banking and capital markets, asset and wealth management, insurance/reinsurance and fund transfer and payments. Alternatively, He et al. (2021b) differentiates between IT spending to improve transmission of soft information via internal bank communication networks versus hardware and big data IT expenditures designed to improve banks' processing of hard information.[2] However, none of these Fintech classifications identify whether the innovation enables the shift of intermediation from a dealer to a broker production function. This is not a self-evident distinction. Even "decentralized" finance, as exemplified by the innovative smart contracts discussed in Harvey et al. (2021), can be delivered by either a broker or a dealer intermediary.[3] For example, decentralized

[2]Similarly, Pierri and Timmer (2021) show that higher IT investment reduces non-performing loans, thereby improving traditional bank performance.

[3]Decentralized finance does not necessarily imply that Fintech innovations result in disintermediation (or the absence of an intermediary). Even online marketplaces utilize an intermediary (the lending site) to bring together borrowers and lenders (see Thakor (2020)). Instead, Fintechs are intermediaries that act as matchmakers to bring together counterparties in transactions such as crowdfunding, robo-advisory asset management and payment transfer.

autonomous organizations (MakerDAOs) may offer innovative loan contracts that automatically generate transfers between borrowers and lenders via blockchain backing, vaulting and auditing of a designated cryptocurrency. Because of their underwriting risks, these contracts are typically overcollateralized in the 150-200 percent range (Harvey et al. (2021), page 70). However, one could imagine the development of dealer MakerDAOs that control loan underwriting risk (thereby reducing collateral requirements) by bundling these contracts with other innovative relationship banking services (such as cryptocurrency asset management or initial coin offerings). If the MakerDAO holds the assets and liabilities generated from the provision of these innovative contracts on its balance sheet, then it is no different from a traditional bank despite the ingenuity of its products. That is, if the MakerDAO uses a leveraged balance sheet to offer Fintech products, then operating costs will reflect the risks and costs of balance sheet intermediation. In contrast, if advances in information technology allow the intermediary to act as a broker between borrowers and lenders on high tech products (such as the smart loan contract), then our analysis shows that even traditional banks offering low tech products can use the broker production function to reap improvements in their operating efficiency. Thus, this paper focuses on the fundamental intermediation production function, rather than the form of specific financial products and services.[4]

This paper also contributes to a growing theoretical and empirical literature on securities trading that examines the coexistence of both dealer and broker technologies within the same financial intermediary. In particular, we extend to all intermediation activities a literature that currently focuses only on how brokers trading OTC securities (such as corporate bonds) have greater cost efficiency than market makers. In this paper, we generalize the endogenous choice of broker versus dealer intermediation format to include all types of capital allocation

[4] Jimenez (2021) models banks as experts in the medium-of-exchange with frictions related to transportation costs, supply elasticity and asymmetric information in its value. He traces the history of how banks developed this expertise from their initial specialization in the storage of coinage and other money assets, thereby leading banks to develop a comparative advantage based on balance sheet intermediation. Although technological change may undermine this comparative advantage, a shift to a broker technology may preserve the banking enterprise. In the context of his model, the inefficient"two world" equilibrium may evolve into a more efficient "non-banking world" through the conversion of banks into brokered intermediaries using Fintech innovations, not through bank obsolescence.

activities, not merely OTC securities trading. Thus, broker intermediaries may reduce costs across their entire range of operations, consistent with the literature examining the intermediation of OTC securities. For example, Li and Li (2019) find that the broker form of intermediation dominates the dealer approach in OTC securities markets that are transparent and high volume by reducing inventory holding costs. An, Song, and Zhang (2019) find that three quarters of broker-like intermediary chains of transactions reduce search frictions, in contrast to the remaining rent-seeking dealer trades that are associated with intermediary market power resulting from information asymmetry regarding transaction prices. Similarly, the shift from "balance-sheet intensive" dealer activities to the broker form of intermediation is modelled by Saar et al. (2019) and Cimon and Garriott (2019), who identify capital and liquidity regulatory changes (e.g., the Volcker rule and Basel III) as the explanation for the transition of bank dealers in OTC markets from costly market-making, dealer activities to less costly matchmaking, broker intermediation. Empirical studies (e.g., Di Maggio et al. 2017, Macchiavelli and Zhou (2019) and Andersen, Duffie, and Song (2019)) find higher trading costs when OTC transactions are conducted by dealers as compared to brokers. For example, Choi and Huh (2019) find a 35 to 60 percent increase in trading costs when dealers commit capital in order to provide liquidity (the market making function) as compared to when customers provide liquidity (matchmaking). These higher costs of the dealer form of intermediation may arise from balance sheet frictions associated with leverage and systemic vulnerability (Adrian, Boyarchenko, and Shachar (2017)), as well as potential conflicts of interest (An and Zheng (2020)) that are exacerbated by large dealer inventories.

In this paper, we generalize these approaches to include all intermediation functions rather than limiting attention to OTC bond trading. Different product lines may have differing propensities to undergo this shift from market making to matchmaking (e.g., GSE mortgage securitizations are more likely to be brokered by Fintechs than jumbo mortgages held on bank balance sheets). However, advances in artificial intelligence may enable and accelerate this shift across many product lines. For example, Figure 1 shows an accelerating decline in balance sheet loans and an increase in brokered deposits for U.S. banks, corresponding to a shift toward match making across all banking functions during a period of growing Fintech activity. This suggests that the endogenous decision to organize intermediation using either a dealer or a broker approach applies to all capital allocation functions. Thus, we develop an empirical methodology to study traditional banks' transition from a dealer to a broker intermediation function using their adoption of Fintech innovation.

We develop an empirical measure, denoted the "Fintech score" that measures the degree of Fintech (broker-like) activities at traditional (dealer) banking firms. Specifically, we identify commonalities across traditional bank financial statements for banks with and without Fintech joint ventures and/or partnerships to determine whether each bank's financial position is more consistent with a broker or a dealer intermediation approach. Using a two-stage principal components analysis over 115 financial variables obtained from bank Call Report data, we estimate each bank's quarterly Fintech score. We contrast financial positions that are consistent with a broker technology (such as brokered deposits, non-interest income and trading securities) with dealer-like activities (such as on-balance sheet small business lending and core deposit taking). Although we do not pre-specify the relationship in our score derivation, our results demonstrate that higher Fintech score banks have higher levels of broker-like activities, whereas lower Fintech score banks dominate in dealer variables.

We validate the Fintech score in three ways. First, we utilize a supervised machine learning algorithm, the Random Forest model, to derive the Fintech score. The Random Forest model extracts the target-oriented features embedded in a large set of observables through a non-linear learning process without imposing any structural assumptions about the relation between the dependent and independent variables. Our results show that the Fintech scores obtained using the Random Forest model are positively and statistically significantly correlated with the scores obtained using the two-stage nonlinear analysis.

The second approach we use to validate our Fintech score utilizes an unsupervised machine learning algorithm, K-means clustering model, in order to derive two clusters of banks. In this model, the number of clusters is pre-specified. The algorithm uses distances to centroids of each cluster to minimize the inertia, as measured by the within-cluster sum-of-squares criterion. Our results show that the Fintech scores obtained using K-means clustering are also positively and statistically significantly correlated with the scores obtained using the two-stage nonlinear analysis.

In our third approach to validate our Fintech score, we employ textual analysis on banks' SEC filings to conduct a keyword search for each bank's Fintech self-identification. If the inclusion of Fintech buzzwords in bank financial disclosures is not simply marketing or cheap talk, but rather indicative of the bank's incorporation of the Fintech operating methodology, we should observe a positive correlation between our Fintech score and keyword search results. We find a positive and statistically significant correlation between both contemporaneous and lagged keyword searches and the bank-specific Fintech score, thereby further validating the Fintech score derivation.

The dealer approach to financial intermediation incurs the costs and risks associated with leveraged balance sheets that include inventories of assets and liabilities. In contrast, if the revolution in information and search technology enables brokers to offer capital allocation services with minimal inventory costs, then the cost of financial intermediation should decline as banks optimally shift from a dealer to a broker methodology. That is, we hypothesize that the incorporation of Fintech firms' broker production functions should reduce the unit cost of financial intermediation at high Fintech score banks. Alternatively, if traditional banks subvert Fintech firms in order to neutralize their competitive pressure and maintain the inefficient dealer intermediation status quo, then there will not be a reduction in the unit cost of financial intermediation at high Fintech score banks.

We examine the relationship between the Fintech score and the cost of financial intermediation using both the Data Envelopment Analysis (DEA) and the Stochastic Frontier

Analysis (SFA) methodologies.[5] We find that higher Fintech scores are associated with greater operational efficiency as measured by both DEA and SFA. Indeed, our benchmark results show that one standard deviation increases in the Fintech score increase DEA (SFA) efficiency by 9.6 (0.92) percentage points, corresponding to 22.6% (1.25%) of the average DEA (SFA) score.

To address endogeneity concerns, we introduce two instrumental variables (IVs) for the Fintech score. For each bank's headquarter county, we identify: (1) the share of local employees working in the information technology sector, and (2) the share of population older than 65. We find that the Fintech score is higher for banks headquartered in areas with more information technology employment and fewer seniors in the local population. Using these IVs, we find that one standard deviation increase in the FinTech score increases the likelihood of DEA (SFA) efficiency ranked in the top quartile among all the banks by 15.4% (2.6%). Furthermore, using the Global Financial Crisis and the passage of the Dodd-Frank Act as exogenous shocks, we show that banks with higher Fintech scores increase their operational efficiency while adapting to industry shocks and regulatory policy changes. Finally, we examine bank mergers in order to address possible reverse causality between operational efficiency and the Fintech score. We find that banks can improve their Fintech score and reduce operational costs by acquiring higher Fintech score banks.

This paper contributes to a growing theoretical literature about the impact of Fintech entry on traditional banking. Private information production is at the heart of the literature on financial intermediation. An important component is the reusability of information on payments and cash flows to screen and monitor loan applicants. Empirical literature has established that traditional banks use private information obtained from deposit cash flows in

[5] We measure both the Fintech score and operational efficiency at the bank subsidiary level rather than consolidating to the holding company level. The individual bank conforms more closely to the fundamental intermediation unit and permits analysis of both publicly and privately traded banks of all sizes. Further, this approach allows us to examine whether Fintech joint ventures undertaken at the holding company level impact intermediation technology at the individual bank level. We follow the methodology of Barth et al. (2013) in estimating DEA. We estimate the SFA measure with stochastic output distance function using the same variables in our DEA calculation.

lending (e.g., Mester et al. (2007)). Recent empirical papers have also documented a similar process within Fintech firms (e.g., Hau et al. (2019), Ghosh et al. (2021)). Theoretical models incorporate externalities from this information production into models of bank/Fintech competition. For example, Ghosh et al. (2021) and Huang (2021) model Fintechs' comparative advantage in information production vis a vis traditional banks. In contrast, Parlour et al. (2021) assume that the bank generates all payments information, which can be purchased by Fintech competitors. The reality is that both banks and Fintechs each produce different, but valuable proprietary information that can be incorporated into the menu of products and prices offered, thereby impacting competitive equilibria.[6] Vives and Ye (2021) consider IT investment by banks, thereby effectively modelling bank and Fintech integration and incorporate monitoring as well as screening. Further, Boualam and Yoo (2022) consider Fintechs' ability to seize cash flows as a technology-enhanced enforcement (monitoring) mechanism that is superior to traditional banks' dependence upon collateral. The capture of relationship banks' comparative advantages (e.g., lower funding costs) together with access to Fintechs' payments and non-traditional credit data could be the motivation for the integration of Fintech into traditional banks documented in this paper. That is, we find evidence that traditional banks have begun a process of integrating both the banking and Fintech comparative advantages.

The paper proceeds as follows. Section 2 derives and validates the Fintech score. The relationship between the Fintech score and the cost of financial intermediation is estimated in Section 3. Section 4 concludes.

[6]Alternatively, He et al. (2021a) model the welfare effects of consumer ownership of their own data, as proposed by the E.U. and the U.K.

2. Deriving the Fintech Score

2.1. Stage 1: The Principal Component Analysis of the Fintech Score Derivation

We design a two-stage algorithm to identify traditional banks integrated with Fintech applications utilizing their balance sheet positions and income statements from Call Report data as follows: (1) a first stage principal component analysis (PCA) to identify similarities in financial variables and reduce the dimensionality of the problem using a multitude of accounting ratios, and (2) a second stage probit model to derive a Fintech score using principal components that measures commonalities across banks with Fintech bank operations. In the first stage, we designate accounting ratios using a broad range of quarterly Call Report variables measuring bank income, lending activity, deposit-taking (included brokered deposits), charge-offs, past due loans, securitization, real estate lending, and quarterly averages. We identify 11,731 individual, unconsolidated banks over our 2001-Q2 through 2016-Q3 sample period. For each quarter over the sample period, we perform a principal component analysis to extract relevant information from the set of 115 financial variables and accounting ratios listed in Table A1.

We conduct the PCA separately for each of four size groupings since financial variables and accounting ratios (and indeed data availability) differ across banks of different sizes. We follow the Federal Reserve (see Saunders, Schmid, and Walter 2016) and delineate the four size groupings according to total domestic assets as follows: (1) below \$100 million; (2) \$100 million to \$1 billion; (3) \$1 billion to \$10 billion; and (4) over \$10 billion. Table 1 provides descriptive statistics for each size group. Availability of each financial variable differs across bank size and across time, but the maximum number of observations in our sample is 489,173, divided across size classes; 40.19% of the observations comprised the smallest banks (size group 1), 51.3% size group 2, 6.79% size group 3 and only 1.72% of the observations from size group 4. In terms of the number of financial variables used for principal component analysis in each group, size groups 1 and 2 each use 81 variables, size group 3 uses 82 variables, and size group 4 uses 99 variables.

Differences across bank size in terms of reliance on core deposits are consistent with longstanding literature (e.g., Allen, Peristiani, and Saunders 1989) show greater dependence on purchased funds by large banks). Panel B of Table 1 shows that the two smallest bank size classes rely on core deposits to fund around 82% of their assets on average, whereas the two largest bank size classes' mean core deposit ratios are 75.9% and 65.7%, respectively. Moreover, brokered deposits at the largest bank size class are 8.3% of total assets on average, whereas the analogous brokered deposit ratio for the two smallest bank size groups is around 2% on average. Credit card loans and securitization charge-offs are sizable for the largest banks only. In contrast, the largest bank group has the smallest proportion of their assets in real estate loans (1-4 family first lien mortgages), averaging 13.6% as compared to all other size groups' averages around 15% of total assets. This reflects the dominance of securitization at large banks that removes loans from the bank's balance sheet, particularly for real estate loans during our sample period.

To obtain the factor loading on each of the variables in the principal component analysis, we perform estimation for each size subsample individually. Table 2 shows the eigenvalues of the twenty principal components that have eigenvalues over one for each of the size subsamples. As expected, the explanatory power of the first principal component is highest for the largest bank size group, explaining almost 30% of the variance in the raw variables. Moreover, Table 2 shows that the first nine principal components explain two thirds of the variance for the largest size group, whereas the other size groups require 14 principal components. However, since the first 14 principal components account for about two thirds of the variance in the raw variables for all size groups, we utilize this as a heuristic cut-off to determine the number of components in the principal component analysis. Using the coefficients for each size class, we evaluate the fitted value of each of the 14 principal components for each bank using Call Report data for each quarter in our sample period.

2.2. Stage 2: Deriving the Fintech Score Using a Probit Model

The PCA identifies commonalities in operations across banks' production functions without any reference to their Fintech activities. To link bank operational production functions to Fintech activity, we define a dummy variable designated *Fintech*, which equals one if we find any evidence of Fintech activity by each bank at each point in time, zero otherwise.[7] To determine Fintech activity, we searched the websites of known Fintech firms for references to their banking affiliates. As of the quarter that banks invested in or partnered with Fintech companies, we coded the *Fintech* dummy variable as one, zero otherwise. We conducted a manual online search of Fintech companies' websites. We also searched the database PrivCo for the identity of non-public Fintech companies. We examined 1,502 Fintech companies, of which 977 companies are obtained from PrivCo, 287 are obtained from manual online searches, and 238 companies are obtained from both channels. These Fintech companies cover areas of lending, payment and settlement, real estate, regulatory technology, blockchain, data analytics, insurance, personal finance, wealth management, and financial services software. Further, we manually searched the SEC filings, websites and industry reports to identify any traditional banks that integrate with the Fintech companies. Integration could take the form of acquiring Fintech companies (e.g., JPMorgan Chase's acquisition of WePay), setting up venture funds to fund Fintech companies (e.g., Citi Ventures), setting up startup programs to incubate Fintech companies (e.g., JPMorgan Chase's FinLab), partnering with Fintech companies (e.g., WebBank's participation in Lending Club loans and securitizations), or launching its own Fintech subsidiaries (e.g., Citi Fintech). If any of these activities is present, we code the *Fintech* dummy variable as one for each period after the announcement of a connection for each of the banks in our sample. We found that

[7] By designating the *Fintech* dummy variable as one for any bank-Fintech connection of any size, we mitigate any bias stemming from resource constraints on smaller banks' abilities to fund larger Fintech investments.

each of the 1,502 Fintech companies obtained from our search are tied in some way to a total of 245 distinct banks, so that the *Fintech* dummy variable is one for these banks and zero for the remaining banks in the Call Reports database.

In the second stage of our Fintech score derivation, we utilize the *Fintech* dummy variable as the dependent variable in a probit model with independent variables consisting of the fitted values of the first fourteen principal components for each bank in each size group estimated using each bank's quarterly data.[8] The Fintech score for each bank is derived using the coefficients estimated from this probit model, and the quarter-end value of each financial variable is used to compute the value of the bank's principal components. We further standardize the values of the fourteen principal components to have zero mean and unity standard deviation so that we can directly compare the economic significance of the slope coefficients on the principal components from the probit regression. Table 3 presents the results of the probit regression using the top fourteen principal components derived for each size subsample separately. As shown in Table 3, the first principal component has a positive and significant (at the 1% level) coefficient of 0.1322, contributing the most to the Fintech score. The Fintech score is winsorized quarterly at the 1% and 99% levels.

The Fintech score can be interpreted as the probability that the bank will be involved in some way with Fintech operations. Table 4 shows a list of banks with the highest average quarterly Fintech scores over the sample period. These banks have the greatest similarities in terms of their financial operations to banks with known Fintech affiliations. Interestingly, there are banks of all size on the list, suggesting that high Fintech score banks are not limited to the largest banking institutions. Indeed, the sample mean Fintech score is 0.008 for all four size classes (no statistically significant difference across size classes). The Fintech score fluctuates for each bank for each quarter as bank financial positions and Fintech integration (i.e., the *Fintech* dummy variable) change over time. Panel A of Figure 1 offers an overview

[8] We also utilize year-quarter fixed effects and obtain virtually the same Fintech score results as reported here.

covering our sample period of 2001-Q2 through 2016-Q3. As expected, the Fintech score (both cross-sectional minimum and maximum) increases over time as Fintech activities are increasingly integrated into traditional banking. However, Panel B of Figure 1 shows that the cross-sectional standard deviation of the Fintech score has increased over time as banks heterogeneously incorporate Fintech into traditional banking activities.

2.2.1. Fintech Scores Reflect Underlying Shifts in Bank Balance Sheets

If Fintech innovations allow banks to offer intermediation services using a broker technology rather than a dealer technology, we should see the resulting imprint on the bank's financial statements. Traditional dealer banks earn a yield spread on the assets held in their portfolios less the cost of funds on their balance sheets. This is reflected in their net interest income. In contrast, intermediaries relying on a broker technology generate non-interest income from fees and commissions. Thus, we hypothesize that as traditional banks adopt Fintech broker operations, we should observe an increase in the non-interest income component of bank earnings. We do not impose these assertions on the Fintech score analysis, but rather test them.

Our results are consistent with these assertions. Dividing banks of all sizes into quartiles based on their Fintech scores, Panel A of Table 5 shows means for several key financial ratios, whereas Panel B of Table 5 conducts statistical tests of the means differences for the highest (Q4) and lowest (Q1) Fintech score quartiles. Consistent with a shift to a broker technology, Panel A of Table 5 shows that the highest Fintech score subsample has the highest proportion of non-interest income, with the ratio increasing monotonically as the Fintech score quartile increases from 0.11% (Q1 banks) to 0.14% (Q4 banks). Panel B of Table 5 shows that the difference in non-interest income ratios between the highest and lowest Fintech score quartiles is statistically significant at the 1% level.

Bank regulations require the designation of securities and other assets into three categories: (1) held-to-maturity, (2) available for sale, and (3) trading assets.[9] Since dealers utilize their balance sheets as intermediation mechanisms, we would anticipate that they would designate more of their assets as either held-to-maturity or available for sale *ceteris paribus*. In contrast, brokers would be more likely to hold securities temporarily on their balance sheets, and therefore would designate more of their assets as trading securities. Table 5 examines the proportion of these classifications across the Fintech score quartiles. As anticipated, lower Fintech score quartiles have greater proportions of held-to-maturity and available for sale assets as compared to higher Fintech score quartiles. Indeed, Panel B of Table 5 shows that means are significantly (at the 1% level) lower for ratios of both held-to-maturity and available for sale assets for the highest Fintech score quartile as compared to the lowest quartile. In contrast, the highest Fintech score quartile has the highest percentage of trading assets, with the means differences significant at the 1% level.

Further distinctions between broker and dealer operating approaches can be observed by examining securitization and off-balance sheet positions (more broker-like) as compared to on-balance sheet lending (more dealer-like). As hypothesized, Table 5 shows that the highest Fintech score quartiles have significantly (at the 1% level) higher securitization ratios as compared to lower Fintech score quartiles. For example, Table 5 shows that the ownership interest in credit card securitizations as a percent of total credit card loans (variable SCR1) averages 0.05% for the lowest Fintech score quartile, but 0.95% for the highest Fintech score quartile (means difference statistically significant at the 1% level).[10] In contrast, the lowest Fintech score banks are more likely to make balance-sheet small business loans (commercial

[9] "Debt securities that the enterprise has the positive intent and ability to hold to maturity are classified as *held-to-maturity securities*... Debt and equity securities that are bought and held principally for the purpose of selling them in the near term are classified as *trading securities*... Debt and equity securities not classified as either held-to-maturity securities or trading securities are classified as *available-for-sale securities*..." See https://www.fasb.org/summary/stsum115.shtml.

[10] The dominance of credit card loans as a source of broker-like Fintech activity is consistent with Emekter et al. (2015) who find that around 70% of online consumer loans at Fintech lenders represent consolidation of high cost credit card debt. Fintech lenders securitize their originated loans. Further, Panel B of Table 5 shows that allowance for credit losses on off-balance sheet exposures (variables PVR15 and PVR16) is significantly (at the 1% level) higher at highest quartile Fintech score banks as compared to lowest Fintech score banks. This is consistent with Tang (2019) who shows that online lending marketplaces serve marginal, high risk borrowers, formerly excluded from traditional bank lending sources.

and industrial, C&I, loans with original balances less than $100,000) than banks in the highest Fintech score quartile. For example, the ratio of small business loans to total domestic C&I loans declines monotonically from 27.94% to 24.52% to 22.93% to 19.53% as the Fintech score increases (means differences between Q1 and Q4 significant at the 1% level). However, the size of small business loans is largest for the highest Fintech score banks, averaging $25,056 as compared to $22,538 (means differences significant at the 1% level).

Finally, Table 5, Panel B shows that the highest Fintech score quartile has significantly (at the 1% level) less individual (core) non-transaction deposits (89.85% vs. 91.91% of total non-transaction deposits) and more brokered deposits (7.76% vs. 0.46%) as compared to the lowest Fintech score quartiles. Moreover, highest Fintech score banks hold greater proportions of institutional deposits (commercial banks and other depository institutions) than lower Fintech score banks. Thus, the results of the empirical estimation of Fintech scores are consistent with the conjecture that traditional banks with more Fintech integration (higher Fintech scores) incorporate more broker-type activity as compared to low Fintech score banks relying more on dealer-type technology although the Fintech score derivation did not impose the condition.[11]

2.3. Validating the Fintech Score Using the Random Forest Model

In this section, we employ classification algorithms in Machine Learning (ML) to validate the Fintech score calculated based on the two-stage algorithm. Unlike the two-stage algorithm, an ML algorithm is not susceptible to the empirical difficulties caused by the large number of independent observables. Neither does it impose a structural assumption about the relation between the dependent variable (i.e., the *Fintech* dummy variable) and

[11] As a robustness check, we estimated the Fintech score using a one-step probit model with the *Fintech* dummy variable as dependent and the same 115 financial variables as independent variables. The scores under both the two-step and one-step derivations are positively correlated. Further, this alternative Fintech score also endogenously delineates broker versus dealer activities such that high (low) Fintech score banks have more (less) non-interest income, less (more) held-to-maturity and available for sale assets, more (less) tradeable assets, less (more) on-balance sheet lending and more (less) brokered deposits.

the independent variables. Rather, an ML algorithm can extract the target-oriented features embedded in a large set of observables through a non-linear learning process. However, we should note several drawbacks of Machine Learning. It is hard to identify which characteristics are the most important in determining a bank's Fintech exposure. Furthermore, we lose most of the data in the training stage, which can substantially impact the analysis that compares operational efficiency across banks with different Fintech exposures. Given these considerable disadvantages, we consider Machine Learning an auxiliary method to validate the Fintech score estimated from the two-stage algorithm.

We focus on the Random Forest model, the most used classifier in ML.[12] The Random Forest model can be understood as a collection of many random trees. Specifically, we randomly divide the 10,268 distinct financial institutions in our sample into one training sample and one test sample, with the former accounting for 75% of the banks, and the latter consisting of the remaining 25%. To avoid data snooping, we use the hyperparameters provided by the algorithm package in Python. For each firm-quarter observation in the test sample, we then calculate the probability of its involvement in Fintech operations. We should note that in contrast to a probability in the traditional sense, a probability given by the Random Forest model is an out-of-bag estimate. That is, it is defined as the proportion of the trees that vote for each class (i.e., the *Fintech* dummy is equal to one or zero). If no tree shows that a bank is engaged in Fintech operations, then the probability is zero. Conversely, if every tree points to the bank's involvement in Fintech operations, then the probability is one. As such, a bank's Fintech score increases with the proportion of the trees pointing to its engagement in Fintech operations.

After obtaining the Fintech score based on the Random Forest model, we examine whether it is related to the Fintech score calculated using the two-stage algorithm. If the two Fintech variables contain information pertinent to a bank's engagement in Fintech oper-

[12]The results of Fintech scores estimated from alternative ML algorithms are qualitatively similar, and are available upon request.

ations, we expect them to be positively correlated. Our results show that this is indeed the case. The first row of Internet Appendix Table A2 compares the two-stage Fintech scores with the Fintech scores we obtain from Random Forest ML model. As expected, the difference is positive and statistically and economically significant at about 0.2%, a magnitude that is comparable to the average cross-sectional standard deviation of the Fintech scores calculated based on the two-stage algorithm. Thus, the Fintech score that is independently estimated using a Random Forest approach is positively correlated with the Fintech score obtained using the two-stage algorithm.

Following the analysis in Table 5, we utilize the Random Forest-based Fintech score to examine the differential effects of different characteristics across banks with different Fintech scores. In particular, we perform univariate analysis by dividing quarterly observations in the test sample into one group with the Fintech score given by the Random Forest model equal to zero (i.e., zero possibility of engaging in Fintech operations) and the other group with the Fintech score greater than zero (i.e., positive possibility of engaging in Fintech operations). We then compare the characteristics of these two groups of banks. The results are highly consistent with those of the Fintech score estimated from the two-stage algorithm presented in Table 5. For example, Internet Appendix Table A2 shows that the group of banks with positive Fintech scores using the Random Forest model generally have larger non-interest income, lower proportions of held-to-maturity and available-for-sale assets, higher securitization ratios, and greater proportions of brokered deposits. Therefore, the results based on the Random Forest model provide methodological validation for the two-stage algorithm.

2.4. Validating the Fintech Score Using the Unsupervised Machine Learning Model

In this section, we employ a clustering algorithm of unsupervised Machine Learning (ML) as another validation of our Fintech score calculated based on the two-stage algorithm. In unsupervised learning algorithms, the algorithm is not provided with any pre-assigned labels or scores for the training data. Therefore, this analysis clusters using our 115 financial variables and accounting ratios without including our Fintech score dummy variable. Among alternative clustering algorithms, we utilize K-means clustering as this model is scalable to very large sample sizes. In K-means clustering, each cluster is defined by creating a centroid for each cluster. The centroids capture the closest points closest and add them to the cluster using an iterative clustering algorithm. The K-means algorithm aims to choose centroids that minimise the within-cluster sum-of-squares criterion, or inertia, using the distances between data points. The output of the algorithm is a group of labels.

While unsupervised ML models help identify all kinds of structures or patterns in a collection of uncategorized data without supervision, their biggest drawback is that we cannot get precise information regarding data sorting. This suggests that clusters are not informational. Due to this shortcoming, we restrict the number of clusters to be identified by the algorithm to two. This allows us to compare the Call Report variables of banks clustered in two groups (cluster of zeros and cluster of ones).

After obtaining the Fintech score based on the K-means algorithm, we examine whether it is related to the Fintech score calculated using the two-stage algorithm. Following our comparison of Fintech score based on Random Forest model, if the Fintech scores from K-means algorithm and two-stage process contain information pertinent to a bank's engagement in Fintech operations, we expect them to be positively correlated. Our results show that this is indeed the case. The first row of Internet Appendix Table A3 compares the two-stage Fintech scores with the Fintech scores we obtain from K-means algorithm of unsupervised

ML model. Consistent with results from supervised ML model, the difference is positive and statistically and economically significant at about 0.5%, a magnitude that is also comparable to the average cross-sectional standard deviation of the Fintech scores calculated based on the two-stage algorithm. Thus, the Fintech score estimated using the K-means algorithm provides another independent estimate that is positively correlated with the Fintech score obtained from the two-stage algorithm.

Following the analyses in Table 5 and Internet Appendix Table A2, we utilize the Fintech score from K-means algorithm to examine the different characteristics across banks in different clusters. In particular, we perform univariate analysis using quarterly observations in cluster of zeros and cluster of ones defined by the K-means algorithm. We then compare the characteristics of these two clusters of banks. The results are highly consistent with those of the Fintech score estimated from the two-stage algorithm presented in Table 5 and Random Forest model in Internet Appendix Table A2. For example, Internet Appendix Table A3 shows that the banks in cluster of ones defined by the K-means algorithm generally have larger non-interest income, lower proportions of held-to-maturity and available-for-sale assets, higher proportions of trading assets, higher securitization ratios, and greater proportions of brokered deposits. Therefore, the results based on the K-means algorithm also provide methodological validation for the two-stage algorithm.

2.5. Validating the Fintech Score Using Textual Analysis

In recent years, Fintech has become a buzz word frequently employed by financial institutions. Self-identification with Fintech can be found in public disclosures. We performed textual analysis of traditional bank SEC filings (10-Q, 10-K, etc.) using the Python algorithm to detect each bank's self-identified integration with Fintech operations.[13] Panel A of

[13] In earlier sections, we estimated the Fintech score for both publicly traded and privately held banks. However, since our keyword search relied on Edgar's SEC filings, we were limited to publicly traded banks to perform the validation exercise described in this section of the paper. If financial reports for publicly traded bank holding companies were not available on Edgar, we downloaded them manually using S&P's Capital

Table 6 provides the list of Fintech keywords used in the analysis. Descriptive statistics of the number of keyword (# of Keywords) is provided in Panel B of Table 6. The number of keywords matches ranges between 0 and 255 with a mean of 4.95 and standard deviation of 9.58. The keyword search results are skewed as demonstrated by the sample median of 2.

Table 7 shows the correlation matrix of the Fintech score and number of keywords. According to these results, the Fintech score and the number of keywords from textual analysis are positively correlated at the 1% significance level. Therefore, the keyword search results are consistent with the Fintech scores, offering independent validation of the approach. That is, banks that self-identify with Fintech are shown to have higher Fintech scores.

3. The Fintech Score and Bank Operating Costs

3.1. Calculating Bank Operational Efficiency

In this section, we examine the relationship between the Fintech score and operational efficiency as measured using both the data envelopment analysis (DEA) and the stochastic cost frontier analysis (SFA) in order to investigate whether adoption of the Fintech methodology has reduced costs at traditional banks. Introduced as a single-input/output efficiency measure for the measurement of productive efficiency by Farrell (1957) and generalized into a multiple-input/output case by Charnes, Cooper, and Rhodes (1978) and to a variable returns to scale efficiency measurement model by Banker, Charnes, and Cooper (1984), the DEA is a non-parametric approach that calculates the relative efficiency score of a decision-making unit (banks in our analysis) compared to the Pareto-efficient frontier technology. The DEA's focus on individual observations rather than on the population average is particularly important in our context given the heterogeneity of Fintech adoption across banks (as shown in Figure 1 Panel B). That is, the methodology can accommodate the differences between

IQ.

broker and dealer production (cost) functions. As a robustness check, we also estimate operational efficiency using the SFA, which is a parametric model that assumes a half-normal distribution of the error term (see Aigner et al. (1977) and Meeusen and Van den Broeck (1977)).

Efficiency is measured as the relationship between revenue and costs. For example, if two banks generate the same revenue levels, the one that incurs lower costs has a higher DEA score than the other bank. The DEA measure is continuous between [0,1], with 0 denoting least efficient and 1 denoting frontier efficiency (most efficient) in each quarter. The linear programming method of technical efficiency using input minimization with variable returns to scale is stated by Murillo-Zamorano (2004) as:

$$TE_{VRS} = \min_{\mu} \Psi^0 \tag{1}$$

subject to

$$\sum_{j=1}^{n} \mu_j X_{ij} \leq \Psi X_i^0,\ i = 1, ..., m \tag{2}$$

$$\sum_{j=1}^{n} \mu_j Y_{rj} \geq Y_r^0,\ r = 1, ..., s \tag{3}$$

$$\sum_{j=1}^{n} \mu_j = 1 \tag{4}$$

where X_{ij} are the inputs, Y_{rj} are the outputs, and Ψ is the proportion of consumption of inputs. This method allows for flexibility in the weights (μ_j) assigned to each bank. Following Barth et al. (2013), we employ a financial intermediation model that has four inputs and three outputs. The four inputs (X_i) are: X_1 (total deposits + total money market funds + total other funding); X_2 (personnel expenses as labor input); X_3 (total fixed assets as physical input) and X_4 (loan loss provisions). The three outputs (Y_i) are: Y_1 (total customer loans + total other lending); Y_2 (total other earning assets: other interest generating assets such as bonds and investment securities); and Y_3 (other, non-interest, income). Using the DEA methodology, we calculate the efficiency score of each bank for each quarter for the period between 2001-Q2 and 2016-Q3. Table 6, Panel B shows that the DEA efficiency measure has a mean (median) of 0.424 (0.415) and a standard deviation of 0.255.

We estimate the operational efficiency function for each of our four size classifications separately, thereby allowing production functions to vary with bank size. Further, using the same variables used in the DEA calculation, we estimate the SFA efficiency score for each size class using stochastic output distance with translog production function in each year-quarter. The distance function provides the advantage that it does not require information about prices as a data input.[14] Table 6, Panel B shows that the SFA efficiency measure has a mean (median) of 0.734 (0.766) and a standard deviation of 0.172.

Correlation results presented in Table 7, Panel A show that the Fintech score is positively correlated with both DEA and SFA efficiency scores at the 1% significance level. Sorting the entire bank sample on the DEA efficiency level, Panel B of Table 7 shows that the average Fintech score for banks in the highest DEA quartile is significantly (at the 1% level) higher than for the lowest DEA quartile banks. Similarly, the average Fintech score for banks in the highest SFA quartile is significantly (at the 1% level) higher than for the lowest SFA quartile banks. Finally, Figure 2 shows a positive relationship between the Fintech score and the operational efficiency. Panel A of Figure 2 shows that compared to the top quartile Fintech score sample of banks, the bottom quartile Fintech score sample always has higher DEA efficiency. Similarly, Panel B of Figure 2 shows that top quartile Fintech score sample of banks has higher SFA efficiency than the bottom quartile Fintech score sample, for every year except for the 2008-2009 crisis years. Together, these results are consistent with a higher operational efficiency for higher Fintech score banks.

[14] Our results are robust to estimating the standard profit function in Berger and Mester (2003).

3.2. The Relationship Between the Fintech Score and Operational Efficiency

In this section, we employ multivariate analysis to examine the relationship between the Fintech score and operational efficiency. Using DEA (SFA) efficiency as our dependent variable, we estimate the following:

$$
\begin{aligned}
&DEA\ Efficiency\ (SFA\ Efficiency)_{i,t} = \\
&\alpha + \beta_1 \times FintechScore_{i,t-1} + \beta_2 \times Size_{i,t-1} + \beta_3 ROE_{i,t-1} + \beta_4 \times ROA_{i,t-1} \\
&+ \beta_5 \times BankEquity_{i,t-1} + YearQuarterFE + \varepsilon_{i,t}
\end{aligned} \tag{5}
$$

We control for bank $Size$ to capture the potential size effects on bank efficiency. We measure $Size$ as the logarithm of total bank assets. We also control for Bank equity, $BankEquity$, measured as the ratio of book value of equity to total assets, as well as ROA and ROE profitability measures. ROA is measured as net income divided by total assets. ROE is defined as the ratio of net income to book value of equity.[15]

Table 8 provides the results of estimation of equation (5). Using the OLS results in column (1) of Table 8, the lagged Fintech score coefficient is positive and significant (at the 1% level), implying that a one standard deviation increase in the one quarter lagged Fintech score increases DEA efficiency by 9.6 percentage points (0.004 × 23.982). This effect corresponds to about 22.6% of average DEA score (0.096/0.424). Since the DEA efficiency measure is a continuous variable within the unit interval and OLS estimation does not guarantee that the predicted values of the dependent variable are restricted to the unit interval, we also estimate equation (5) using a fractional response regression. The coefficient on the lagged Fintech score is positive and significant at the 1% level. These results, presented in column (2) of Table 8, show that a one standard deviation increase in the Fintech score increases

[15] Since not all of the banks in our sample are publicly traded, we cannot use market to book and other variables. Instead, we utilize bank fixed effects.

DEA efficiency by about 9.77 percentage points (0.004 × 3.825), corresponding to about 23% of average DEA score (0.0977/0.424). In order to mitigate omitted variable bias, we also estimate fixed effects regressions. Shown in column (3) of Table 8, these results provide evidence (significant at the 1% level) that a one standard deviation increase in the Fintech score increases DEA efficiency by 8.73 percentage points (0.004 × 21.833), corresponding to about 20.6% of average DEA score (0.0873/0.424).

The last three columns of Table 8 present results of regressions using the SFA efficiency score as the dependent variable. OLS results in column (4) of Table 8 show that a one standard deviation increase in the one quarter lagged Fintech score increases the SFA efficiency by about 0.92 percentage points (e.g., 0.004 × 2.299), significant at the 1% level. This effect corresponds to about 1.25% of average SFA score (0.0092/0.734). Results using fractional response regressions in column (5) show that the same increase in Fintech score increases SFA efficiency score by about 1.13 percentage points (e.g., 0.004 × 2.832), significant at the 1% level. This effect corresponds to about 1.54% of average SFA score (0.0113/0.734). Finally, in the fixed effects regression in column (6), one standard deviation increase in the one quarter lagged Fintech score increases the SFA efficiency by about 0.63 percentage points (e.g., 0.004 × 1.584), significant at the 1% level. It corresponds to about 0.86% of average SFA score (0.063/0.734).

Among the control variables, bank size has a positive and statistically significant (at the 1% level) relationship with bank efficiency, indicating that larger banks operate more efficiently *ceteris paribus*. The coefficients of ROA and ROE are also positive and statistically significant at the 1% level, showing that more profitable banks also operate more efficiently. Finally, banks with higher equity capital have significantly higher DEA and SFA efficiency levels.

3.3. Instrumental Variable Analysis

In this section, we address endogeneity by identifying two instrumental variables (IVs) to explain the Fintech score. For each bank's headquarters county, we identify: (1) the share of local employees working in the information technology sector, and (2) the share of population older than 65. We obtain data on employment in the information technology sector from the Quarterly Census of Employment and Wages of the U.S. Bureau of Labor Statistics. The data on population age across U.S. counties are obtained from the National Center for Health Statistics. We hypothesize that a higher share of information technology sector workers in the county where the bank is headquartered increases the bank's propensity to adopt Fintech innovations. In contrast, a higher percentage of senior customers in the bank's headquarters county may increase customer resistance to novel technologies, thereby limiting the bank's propensity to adopt Fintech innovation. In addition to instrument relevance, our IVs meet the exclusion restriction since they should not directly affect operational efficiency except via Fintech adoption.

Table 9 reports the results of the instrumental variable method. Column 1 shows the first stage regression. As expected, the first stage results show that the bank's Fintech score significantly (at the 10% level) increases with information technology employment and significantly (at the 1% level) decreases with the share of senior citizens in the bank's headquarters county. The weak identification test rejects the null hypothesis that the first stage equation is weakly identified. The Hansen test also does not reject the null hypothesis that the instruments are exogenous. Column 2 of Table 9 presents the results for the second stage of the analysis using the fitted Fintech score and a linear probability model. The dependent variable is a dummy variable that equals one if a bank's DEA efficiency is in the top quartile among all banks in each quarter, and zero otherwise. The results confirm our findings in the baseline regression of Equation 5. The coefficient on lagged Fintech score is positive and significant at the 1% level, suggesting that the higher the bank's Fintech score, the greater the bank's operational efficiency. Economically, a one standard deviation increase in the FinTech score increases the likelihood of the bank's DEA efficiency ranking in the top quartile across all banks by 15.4% (0.162 × 0.952).[16]

3.4. The Impact of the Global Financial Crisis

The Global Financial Crisis impacted every aspect of banking. Therefore, we consider the impact of the crisis on the relationship between the Fintech score and DEA efficiency.[17] We construct a quasi difference-in-differences setting to analyze whether the operational efficiency of banks with high Fintech scores is different from the operational efficiency of those with low Fintech scores before and after the crisis. We define our control and treatment groups using the first and fourth quartiles of the Fintech score variable, respectively, for each quarter during our sample period. As a robustness test, we use below and above median Fintech scores as our control and treatment groups.

In Table 10, we define the $Crisis$ dummy variable as 1 for the post-crisis period (2009Q3-2011Q1) and as 0 for the pre-crisis period (2006Q1-2007Q3). We use NBER's recession dates to define crisis quarters. The coefficient on the $Crisis$ dummy shows that banks experienced decreases in their operational efficiency after the crisis compared to the pre-crisis period (i.e., the coefficient estimate on the $Crisis$ dummy variable is negative and significant at the 1% level in both columns of Table 10). However, as the positive and statistically significant (at the 1% level) coefficient on the High fintech score × $Crisis$ dummy interaction variable in column (1) of Table 10 shows, after the crisis banks with high Fintech scores experienced higher increases (or lower decreases) in their operational efficiency as compared to banks with low Fintech scores. Specifically, banks with high Fintech scores have 20.1 percentage points higher DEA efficiency after the crisis than banks with low Fintech scores. The results

[16] Analogous results using SFA are presented in Internet Appendix Table A4.

[17] As a robustness test, we estimate these regressions using the SFA efficiency measure as the dependent variable. Our results are qualitatively and quantitatively very similar and provided in Internet Appendix Table A5.

with below and above median control and treatment groups shown in column (2) of Table 10 are qualitatively similar, but lower in magnitude (with 9.2 percentage points higher DEA efficiency for above median Fintech score banks).

3.5. The Impact of the Passage of the Dodd-Frank Act

Another major event in the banking industry, in the aftermath of the Global Financial Crisis, is the passage of the Dodd-Frank Act during the third quarter of 2010 (announced during Q2 of 2010). We examine the impact of this regulatory policy change on the relationship between DEA efficiency and the Fintech score.[18] Similar to the analysis presented in the previous section, we define our control and treatment groups using the 1st and 4th quartiles of the Fintech score variable, respectively, for each bank in each quarter during our sample period. As a robustness check, we also use below and above median Fintech score as our control and treatment groups. We investigate the long-run effects of the Dodd-Frank Act using 24 quarters before and after the passage of law, defining the $Dodd-Frank$ dummy as 0 for the period 2004Q3-2010Q2 and 1 for the period 2010Q3-2016Q2.

The results of this analysis, presented in Table 11, show that higher Fintech score banks were better able to adapt to the new post-Dodd Frank regulatory environment and experienced higher increases (or lower decreases) in their operational efficiency as compared to lower Fintech score banks. The positive and statistically significant (at the 1% level) coefficient on the difference-in-differences term presented in column (1) of Table 11 shows that top quartile Fintech score banks have 12.5 percentage points higher DEA efficiency in the post-Dodd-Frank Act period as compared to low Fintech score banks during the same time period. Similarly, column (2) of Table 11 shows that above median Fintech score banks have a statistically significant (at the 1% level) 9.3 percentage points higher DEA efficiency in the post-Dodd-Frank Act period as compared to below median Fintech score banks.

[18]We present similar results using the SFA efficiency measure as the dependent variable in Internet Appendix Table A6.

3.6. The Impact of Fintech Score Increasing Mergers on DEA Efficiency

We have focused on the integration of Fintech into traditional banking by examining joint ventures, acquisitions and mergers between banks and standalone Fintech companies. However, Fintech integration can also be accomplished if low Fintech score banks acquire high Fintech score banks. In this section, we analyze how Fintech integration through bank acquisitions affects banks' efficiency by analyzing banks with low Fintech scores that acquire banks that had higher Fintech scores before the merger. We compare these observations to a matched sample of non-merging banks.

From the Federal Reserve Bank of Chicago, we obtained bank subsidiary-level mergers that were completed between 2001-Q2 through 2016-Q3. We include only those bank mergers where the target bank ceases to exist after the merger (i.e., merger code 1: charter discontinued). Using our sample of merged survivor banks, we perform both panel IV regression and quasi difference-in-differences analysis. We define the variable, $Merger$ dummy, as zero for two quarters before the merger completion date and one for the quarter of and the quarter after the merger completion date. $Treated$ dummy variable equals one for bank mergers in which the acquirer (survivor) banks' pre-merger Fintech score is lower than the target banks' pre-merger Fintech score.[19] This definition allows us to observe the integration of Fintech through acquisition of higher Fintech score banks by acquirer banks that have lower Fintech scores in the pre-merger quarters. However, in addition to the integration of Fintech via acquisition, the acquirer can improve its post-merger efficiency if the target is run more efficiently than the acquirer. In order to rule out this confounding argument, we restrict the acquirer-target pairs in the treated sample to only those merger events where the target banks' pre-merger DEA score is lower than the pre-merger DEA score of the acquirer. As a result of imposing these restrictions, we define 189 acquirer-target pairs in the treated

[19] Our robustness tests using the SFA efficiency measure are available in Internet Appendix Table A7.

sample.

Table 12 validates the assertion that the acquiring banks in our merger sample increase their Fintech score by merging with higher Fintech score target banks. The table shows univariate tests of Fintech score changes from the pre-merger period (two quarters before the merger completion date) to the post-merger period (the quarter of and following the merger completion date). According to these tests, low Fintech score banks' acquisitions of higher Fintech score targets increase the survivor banks' average Fintech score in the post- as compared to the pre-merger period (with the mean difference significant at the 1% level). On the other hand, bank acquisitions of lower or same level Fintech score targets do not result in a statistically significant change in the survivor bank's Fintech score.

To define the control group banks, we match the treated sample banks to non-merging banks in the pre-merger quarter. Following Iacus, King, and Porro (2012), we employ a coarsened exact matching method. As a result of one-to-one matching, we have 118 acquirer-target bank pairs in the treated group and 118 non-merger banks in the control group. The top rows of Table 13 show the tests of the balance of covariates between treated and control samples in the pre-merger quarter. In the bottom rows of Table 13, we present parallel trend tests. Both sets of results provide evidence that our analysis is not driven by unbalanced samples, and that treated and control samples had parallel trends in the pre-merger quarter.

To address possible endogeneity from the acquirer's choice of merger target, we employ a two-stage analysis. In the first stage, we limit the merger's impact on the Fintech score to the case where acquirers have lower Fintech scores than targets. We utilize the Merger $\times$ Treated interaction variable as an instrument for the post-merger change in Fintech score in the first stage. The first column of Table 14 shows the results of the first stage of the panel IV regression. The coefficient on the Merger $\times$ Treated instrumental variable is positive (significant at the 1% level), suggesting that treated banks experience a statistically significant increase in their Fintech score as compared to matched non-merger banks over the same time period. The second stage analysis utilizes the fitted value of the first stage (i.e., the predicted post-merger Fintech score) in order to analyze the impact on efficiency for the treated mergers. Column (2) of Table 14 shows the positive and significant (at the 1% level) impact of the predicted Fintech score on the DEA efficiency. The economic significance suggests that a one standard deviation increase in the predicted Fintech score (0.104) increases the DEA efficiency by 2.84 percentage points during the post-merger period as compared to matched control sample. (i.e., 0.104×0.274). This represents a substantial economic impact since the sample mean DEA is 0.424.

Columns (3) of Table 14 presents the results of quasi difference-in-differences estimation of DEA around the merger event for acquirers with lower Fintech scores than their targets. The dependent variable is the DEA score from pre- to post-merger. The positive and significant (at the 5% level) coefficient on Merger $\times$ Treated in column (3) shows that the acquisition of a bank with a higher Fintech score increases the DEA efficiency of the acquirer (surviving) bank. The coefficient estimate in Column (3), with bank level time-varying control variables, implies a substantial economic impact of a 2.7 percentage point increase in DEA for banks acquiring a higher Fintech score target bank. The results of this analysis are consistent with the finding that DEA efficiency improves when low Fintech score banks substantially increase their Fintech score by acquiring higher Fintech score target banks.[20]

4. Conclusion

We empirically develop a novel conceptual definition of Fintech as financial intermediation using information technology to support a broker (match maker), as compared to a dealer (market maker) methodology. *Ceteris paribus*, the cost of production for brokers is less than the cost for dealers who rely on leveraged balance sheets and bear inventory risk. To

[20] Our results are robust when we replace the DEA efficiency measure with the SFA efficiency measure, as shown in Internet Appendix Table A9.

the extent that many traditional banks have substantially integrated Fintech companies and applications into their operations, we investigate whether these Fintech-adopting banks' unit costs have declined.

We develop an empirical Fintech score to measure the degree of Fintech integration into traditional bank operations. The Fintech score identifies commonalities in financial statements across banks with and without Fintech joint ventures or acquisitions. The results of analysis of high and low Fintech score banks indicate that high Fintech score banks are more likely to engage in broker-like forms of intermediation, whereas low Fintech score banks use more dealer-like approaches. That is, high Fintech scores banks have higher levels of securitization, trading assets and non-interest income consistent with a broker technology. In contrast, low Fintech score banks have higher levels of on-balance sheet small business lending and core deposit taking.

We utilize the Fintech score to examine the relationship between the broker versus dealer technology and the cost of intermediation services. We find that higher Fintech score banks have greater operational efficiency, as measured using the non-parametric Data Envelopment Analysis (DEA) and the parametric Stochastic Frontier Analysis (SFA) methodologies. We address endogeneity using instrumental variables that establish causality in Fintech adoption by examining the percentage of IT employment and senior citizens in bank headquarters areas. Our IV analysis confirms the OLS findings of a positive relationship between the bank's Fintech score and operational efficiency. Analysis of the Global Financial Crisis and the Dodd-Frank regulatory policy change suggests that traditional banks with higher Fintech integration experience increases in their operational efficiency when adapting to industry shocks and regulatory policy changes. Finally, bank acquisitions of higher Fintech score bank targets increase the post-merger banks' Fintech scores and improve operational efficiency.

References

Adrian, Tobias, Nina Boyarchenko, and Or Shachar, 2017, Dealer balance sheets and bond liquidity provision, *Journal of Monetary Economics* 89, 92–109.

Aggarwal, S., S. Alok, P. Ghosh, and S. Gupta, 2020, Financial inclusion and alternate credit scoring for the millenials: Role of big data and machine learning in fintech.

Aigner, D.J., C.A.K. Lovell, and P. Schmidt, 1977, Formulation and estimation of stochastic frontier production functions, *Journal of Econometrics* 6, 21–37.

Allen, Franklin, Xian Gu, and Julapa Jagtiani, 2020, A survey of fintech research and policy discussion, Federal Reserve Bank of Philadelphia Working Paper No. 20-21, June.

Allen, Linda, Stavros Peristiani, and Anthony Saunders, 1989, Bank size, collateral, and net purchase behavior in the federal funds market: empirical evidence, *Journal of Business* 501–515.

An, Yu, Yang Song, and Xingtan Zhang, 2019, Reducible intermediation chains, working paper.

An, Yu, and Zeyu Zheng, 2020, Conflicted immediacy provision, working paper.

Andersen, Leif, Darrell Duffie, and Yang Song, 2019, Funding value adjustments, *The Journal of Finance* 74, 145–192.

Banker, Rajiv D, Abraham Charnes, and William Wager Cooper, 1984, Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science* 30, 1078–1092.

Barth, James R, Chen Lin, Yue Ma, Jesús Seade, and Frank M Song, 2013, Do bank regulation, supervision and monitoring enhance or impede bank efficiency?, *Journal of Banking & Finance* 37, 2879–2892.

Bazot, Guillaume, 2018, Financial consumption and the cost of finance: Measuring financial efficiency in europe (1950–2007), *Journal of the European Economic Association* 16, 123–160.

Berg, T., V. Burg, A. Gombovic, and M. Puri, 2020, On the rise of fintechs: Credit scoring using digital footprints, *The Review of Financial Studies* 33, 2845–2897.

Berger, Allen N, and Loretta J Mester, 2003, Explaining the dramatic changes in performance of us banks: technological change, deregulation, and dynamic changes in competition, *Journal of Financial Intermediation* 12, 57–95.

Boualam, Y., and P. Yoo, 2022, Fintech disruption, banks and credit (dis-)intermediation: When do foes become friends?

Charnes, Abraham, William W Cooper, and Edwardo Rhodes, 1978, Measuring the efficiency of decision making units, *European Journal of Operational Research* 2, 429–444.

Chen, Mark, Qinxi Wu, and Baozhong Yang, 2019, How valuable is fintech innovation?, *The Review of Financial Studies* 32, 2062–2106.

Choi, Jaewon, and Yesol Huh, 2019, Customer liquidity provision: Implications for corporate bond transaction costs, working paper.

Cimon, David, and Corey Garriott, 2019, Banking regulation and market making, *Journal of Banking & Finance* 109, 105653.

Crouhy, Michel, Dan Galai, and Zvi Wiener, 2021, The impact of fintechs on financial intermediation: A functional approach, *The Journal of FinTech* 1, 1–52.

Di Maggio, Marco, Amir Kermani, and Zhaogang Song, 2017, The value of trading relations in turbulent times, *Journal of Financial Economics* 124, 266–284.

Emekter, Riza, Yanbin Tu, Benjamas Jirasakuldech, and Min Lu, 2015, Evaluating credit risk and loan performance in online peer-to-peer (p2p) lending, *Applied Economics* 47, 54–70.

Farrell, Michael James, 1957, The measurement of productive efficiency, *Journal of the Royal Statistical Society: Series A (General)* 120, 253–281.

Feng, Zifeng, and Zhonghua Wu, 2018, Bank technology: Productivity and employment, working paper.

Ghosh, P., B. Vallee, and Y. Zeng, 2021, Fintech lending and cashless payments.

Harvey, Campbell, Ashwin Ramachandran, and Joey Santoro, 2021, *DeFi and the Future of Finance* (John Wiley and Sons).

Hau, Y., H.and Huang, H. Shan, and Z. Sheng, 2019, How fintech enters china's credit market, *The American Economic Review Papers and Proceedings* 109, 60–64.

He, Z., J. Huang, and Z. Zhou, 2021a, Credit market competition when borrowers own the data.

He, Z., S. Jiang, D. Xu, and X. Yin, 2021b, Investing in lending technology: It spending in banking, University of Chicago Working Paper.

Huang, J., 2021, Fintech expansion.

Iacus, Stefano M, Gary King, and Giuseppe Porro, 2012, Causal inference without balance checking: Coarsened exact matching, *Political Analysis* 1–24.

Jagtiani, J., and C. Lemieux, 2019, The roles of alternative data and machine learning in fintech lending, *Financial Management* 48, 1009–1029.

Jimenez, Mathias, 2021, Are banks technologically obsolete? a new monetarist approach, Stanford University working paper.

Li, Jiacui, and Wenhao Li, 2019, Brokers or dealers? trading intermediation across markets and over time, working paper.

Macchiavelli, Marco, and Xing Alex Zhou, 2019, Funding liquidity and market liquidity: the broker-dealer perspective, working paper.

Meeusen, W., and J. Van den Broeck, 1977, Efficiency estimation from cobb-douglas production functions with composed error, *International Economic Review* 18, 435–444.

Mester, L., L. Nakamura, and M. Renault, 2007, Transactions accounts and loan monitoring, *The Review of Financial Studies* 20, 529–556.

Murillo-Zamorano, Luis R, 2004, Economic efficiency and frontier techniques, *Journal of Economic surveys* 18, 33–77.

Parlour, C., U. Rajan, and H. Zhu, 2021, When fintech competes for payment flows.

Philippon, Thomas, 2015, Has the us finance industry become less efficient? on the theory and measurement of financial intermediation, *American Economic Review* 105, 1408–38.

Philippon, Thomas, 2019, On fintech and financial inclusion, NBER Paper Number 26330.

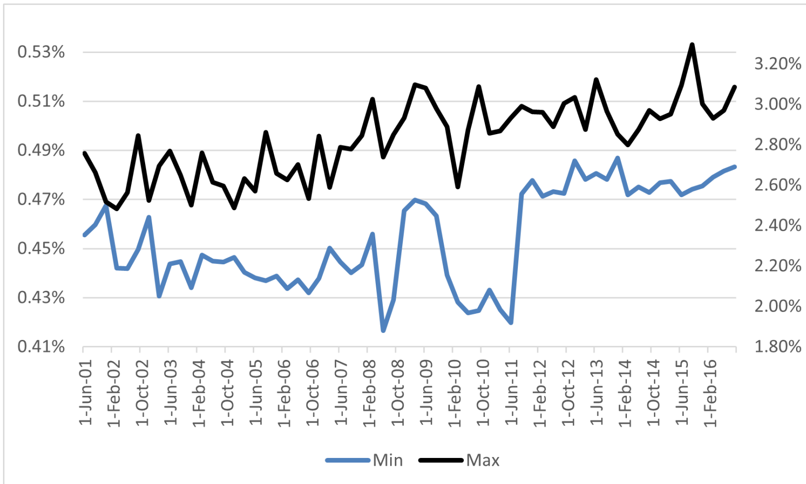
Pierri, N., and Y. Timmer, 2021, The importance of technology in banking during a crisis, ESRB Working Paper Series No. 117, March.

Saar, Gideon, Jian Sun, Ron Yang, and Haoxiang Zhu, 2019, From market making to matchmaking: Does bank regulation harm market liquidity?, working paper.

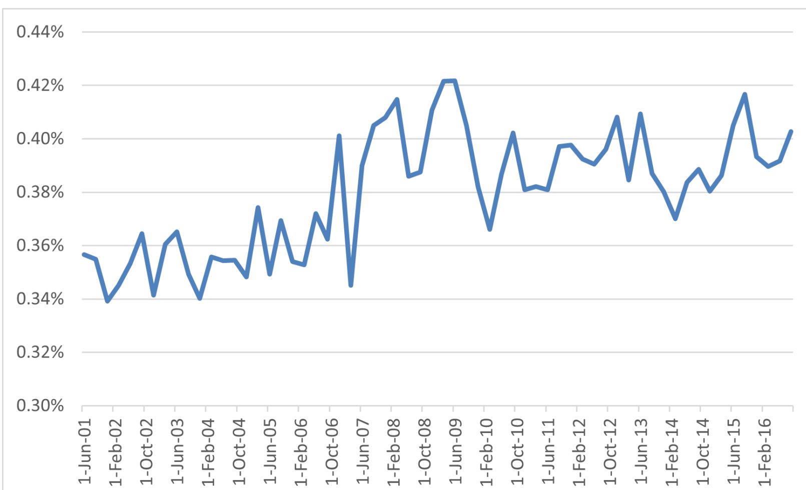
Saunders, Anthony, Markus Schmid, and Ingo Walter, 2016, Non-interest income and bank performance: Does ring-fencing reduce bank risk, working paper.

Thakor, Anjan, 2020, Fintech and banking: What do we know?, *Journal of Financial Intermediation* 41, 1–13.

Vives, X., and Z. Ye, 2021, Information technology and bank competition.



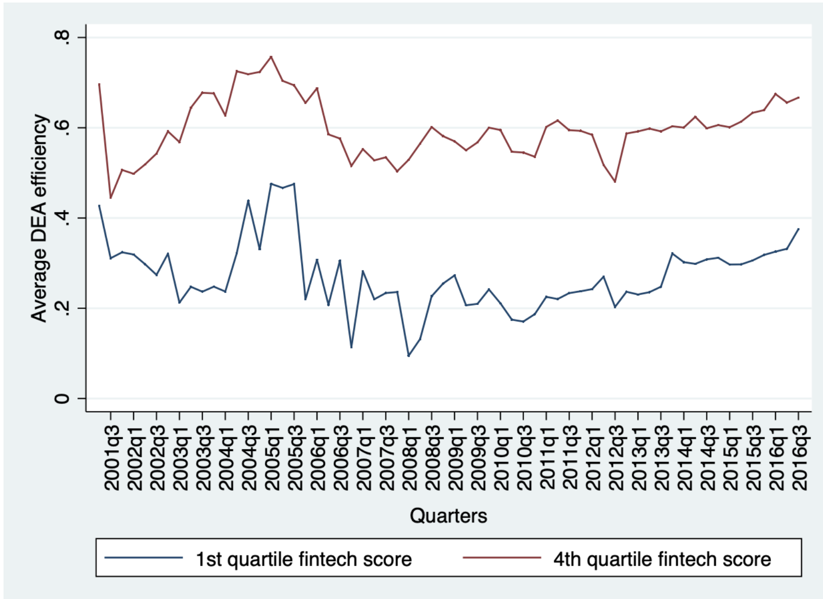
Panel A Cross-sectional Minimum and Maximum of the Fintech Score



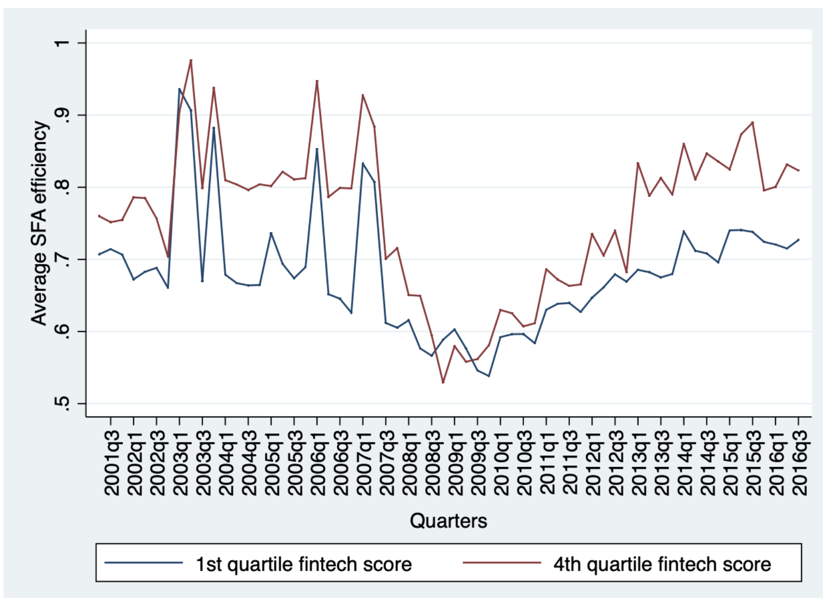
Panel B Cross-sectional Standard Deviation of the Fintech Score

Figure 1: Cross-sectional Statistics of the Fintech Score

This figure presents the cross-sectional statistics of the Fintech score. Panel A presents the cross-sectional minimum and maximum value of the Fintech score. Panel B presents the cross-sectional standard deviation of the Fintech score.



Panel A Average DEA Efficiency for Top and Bottom Quartile Fintech Score



Panel B Average SFA Efficiency for Top and Bottom Quartile Fintech Score

Figure 2: Fintech Score and and Bank Efficiency

This figure shows the relationship between the Fintech score and operational efficiency. In Panel A (B), we compare the average DEA (SFA) efficiency measure for banks with Fintech scores in the bottom quartile to banks with Fintech scores in the top quartile.

Table 1: Descriptive Statistics

Panel A

	All Firms		Size 1 Total Domestic Assets ∈ (0 Mil, \$100 Mil]		Size 2 Total Domestic Assets ∈ (\$100 Mil, \$1 Bil]		Size 3 Total Domestic Assets ∈ (\$1 Bil, \$10 Bil]		Size 4 Total Domestic Assets ∈ (\$10 Bil, + ∞)	
Variable	N	Mean (\$k)	N	Mean (\$k)	N	Mean (\$k)	N	Mean (\$k)	N	Mean (\$k)
Total assets (RCON2170)	489,173	1,546,177.41	196,611	52,524.48	250,941	294,762.38	33,227	2,815,503.82	8,394	68,918,589.55
C&I loans to U.S. Addresses (RCON1763)	41,589	1,723,510.34	.	.	.	.	33,207	326,444.48	8,382	7,258,271.23
Quarterly average credit cards (RCONB561)	468,485	56,910.20	187,642	57.38	244,958	945.85	29,941	45,667.79	5,944	4,214,631.96
Loans to indv.: Credit cards loans (RCONB538)	473,316	56,640.03	192,452	54.29	244,979	935.43	29,941	46,377.79	5,944	4,236,276.44
Brokered deposits (RCON2365)	473,316	68,956.99	192,452	833.38	244,979	9,832.55	29,941	142,925.57	5,944	4,338,822.59
Domestic deposits (RCON2200)	473,316	927,058.17	192,452	44,321.84	244,979	238,922.54	29,941	1,973,253.65	5,944	52,599,166.92
Total int. & fee income on loans (RIAD4010)	473,315	29,919.59	192,452	1,362.12	244,978	7,489.01	29,941	64,011.58	5,944	1,707,273.32
Charge-offs on credit card loans (RIADB514)	473,314	2,248.86	192,451	1.62	244,978	48.61	29,941	1,651.82	5,944	168,697.89
Recoveries on credit card loans (RIADB515)	473,315	332.37	192,452	0.24	244,978	7.76	29,941	265.95	5,944	24,799.57
Total noninterest income (RIAD4079)	473,315	18,265.95	192,452	471.97	244,978	2,106.69	29,941	26,515.54	5,944	1,218,829.10
Interest income: credit cards (RIADB485)	468,479	4,347.59	187,637	4.48	244,957	97.83	29,941	4,123.42	5,944	317,713.70
Trading revenue (RIADA220)	473,315	1,550.49	192,452	2.36	244,978	1.81	29,941	163.50	5,944	122,489.78
Total interest income (RIAD4107)	473,315	39,049.79	192,452	1,713.29	244,978	9,037.23	29,941	78,881.14	5,944	2,284,223.55
Indv.: non-txn. accounts incl. MMDA (RCONB550)	473,316	740,160.88	192,452	28,003.62	244,979	170,199.02	29,941	1,591,729.24	5,944	42,999,240.25
Txn. accounts of indv., corps., ptr. (RCONB549)	473,316	4,409.65	192,452	65.32	244,979	663.69	29,941	6,430.09	5,944	289,278.87
Total non-txn. accounts incl. MMDAs (RCON2385)	489,104	857,227.09	196,569	30,457.78	250,915	183,661.44	33,226	1,703,413.27	8,394	37,003,231.24
Total txn. accounts incl. demand dep. (RCON2215)	473,316	142,739.69	192,452	13,303.30	244,979	53,030.87	29,941	257,613.33	5,944	7,452,234.57
Total C&I loans domestic (RCON1766)	489,094	173,479.76	196,552	4,751.13	250,931	29,709.08	33,226	358,012.20	8,385	7,699,919.19
Net loans and leases (RCONB529)	463,828	210,197.66	192,344	31,546.35	243,443	187,987.16	28,041	1,628,460.16	.	.
Indv.: non-txn. accounts incl. MMDA divided by total non-txn. dep. (RCONB550/RCON2385)	470,235	0.67	189,773	0.63	244,674	0.68	29,855	0.79	5,933	0.82
Qtr. avg. total loans in dom. offices (RCON3360)	473,316	783,070.52	192,452	31,574.36	244,979	190,308.25	29,941	1,717,906.78	5,944	44,836,118.17
Closed end 1-4 family mtg. first lien (RCON5367)	473,316	184,091.34	192,452	8,138.59	244,979	45,715.82	29,941	364,604.77	5,944	10,674,805.10
Charge offs on credit card sold (RIADB749)	464,553	1,263.00	187,532	0.56	241,572	42.53	29,587	887.16	5,862	93,842.67
Recoveries on credit card sold (RIADB756)	464,553	133.42	187,532	0.10	241,572	8.07	29,587	80.27	5,862	9,832.72

Panel B

Variable	N	Mean (% of TA)	N	Mean (% of TA)	N	Mean (% of TA)	N	Mean (% of TA)	N	Mean (% of TA)
Core Deposits	473,314	0.817	192,450	0.820	244,979	0.826	29,941	0.759	5,944	0.657
Brokered Deposits	473,314	0.025	192,450	0.014	244,979	0.029	29,941	0.048	5,944	0.083
Interest Income on Loans	473,313	0.025	192,450	0.025	244,978	0.026	29,941	0.024	5,944	0.023
Noninterest Income	473,313	0.040	192,450	0.087	244,978	0.007	29,941	0.010	5,944	0.014
Credit Card Loans	473,314	0.003	192,450	0.001	244,979	0.003	29,941	0.013	5,944	0.065
Credit Card Charge off Securitizations	464,551	0.000	187,530	0.000	241,572	0.000	29,587	0.000	5,862	0.002
Real Estate Loans	473,314	0.152	192,450	0.147	244,979	0.159	29,941	0.134	5,944	0.136
C&I Loans	488,117	0.095	195,575	0.086	250,931	0.098	33,226	0.123	8,385	0.122

Table 2: Principal Component Analysis

This table reports the results for the principal component analysis for the four asset size subsamples.

Size 1				Size 2				Size 3				Size 4			
Total Domestic Assets $\in$ (0 Mil, \$100 Mil]				Total Domestic Assets $\in$ (\$100 Mil, \$1 Bil]				Total Domestic Assets $\in$ (\$1 Bil, \$10 Bil]				Total Domestic Assets $\in$ (\$10 Bil, $+\infty$)			
Principal Component	Eigenvalue	Proportion	Cumulative	Principal Component	Eigenvalue	Proportion	Cumulative	Principal Component	Eigenvalue	Proportion	Cumulative	Principal Component	Eigenvalue	Proportion	Cumulative
1	12.1472	0.1687	0.1687	1	13.6648	0.1775	0.1775	1	14.6770	0.1835	0.1835	1	26.6700	0.2807	0.2807
2	6.7480	0.0937	0.2624	2	7.2809	0.0946	0.272	2	8.2390	0.103	0.2865	2	8.4944	0.0894	0.3702
3	4.1149	0.0572	0.3196	3	4.4696	0.058	0.3301	3	4.3492	0.0544	0.3408	3	5.3261	0.0561	0.4262
4	3.1085	0.0432	0.3628	4	3.2093	0.0417	0.3717	4	3.8816	0.0485	0.3893	4	4.1055	0.0432	0.4694
5	2.6236	0.0364	0.3992	5	3.0341	0.0394	0.4112	5	3.4133	0.0427	0.432	5	3.9121	0.0412	0.5106
6	2.5419	0.0353	0.4345	6	2.3987	0.0312	0.4423	6	2.5820	0.0323	0.4643	6	3.7012	0.039	0.5496
7	2.4092	0.0335	0.468	7	2.1413	0.0278	0.4701	7	2.3120	0.0289	0.4932	7	3.6487	0.0384	0.588
8	2.2997	0.0319	0.4999	8	2.1355	0.0277	0.4978	8	2.1368	0.0267	0.5199	8	3.3212	0.035	0.6229
9	2.1394	0.0297	0.5296	9	2.0451	0.0266	0.5244	9	2.0675	0.0258	0.5457	9	2.9145	0.0307	0.6536
10	2.0342	0.0283	0.5579	10	1.9532	0.0254	0.5498	10	1.9579	0.0245	0.5702	10	2.4286	0.0256	0.6792
11	1.8064	0.0251	0.583	11	1.9293	0.0251	0.5748	11	1.8895	0.0236	0.5938	11	2.2889	0.0241	0.7033
12	1.6792	0.0233	0.6063	12	1.8796	0.0244	0.5992	12	1.8025	0.0225	0.6164	12	2.0474	0.0216	0.7248
13	1.5608	0.0217	0.628	13	1.8598	0.0242	0.6234	13	1.7817	0.0223	0.6386	13	1.8816	0.0198	0.7446
14	1.5215	0.0211	0.6491	14	1.7399	0.0226	0.646	14	1.5957	0.0199	0.6586	14	1.7139	0.018	0.7627
15	1.4201	0.0197	0.6688	15	1.6071	0.0209	0.6669	15	1.5765	0.0197	0.6783	15	1.5996	0.0168	0.7795
16	1.3743	0.0191	0.6879	16	1.4232	0.0185	0.6853	16	1.4063	0.0176	0.6959	16	1.5190	0.016	0.7955
17	1.3163	0.0183	0.7062	17	1.4115	0.0183	0.7037	17	1.3291	0.0166	0.7125	17	1.3022	0.0137	0.8092
18	1.1793	0.0164	0.7226	18	1.3758	0.0179	0.7215	18	1.2545	0.0157	0.7282	18	1.2069	0.0127	0.8219
19	1.1213	0.0156	0.7381	19	1.2993	0.0169	0.7384	19	1.2292	0.0154	0.7435	19	1.1307	0.0119	0.8338
20	1.0223	0.0142	0.7523	20	1.1553	0.015	0.7534	20	1.1722	0.0147	0.7582	20	1.0505	0.0111	0.8449

Table 3: Probit Estimation of the Fintech Score

This table reports the results of the probit regression of the *Fintech* dummy on the main principal components estimated from the principal component analysis. The *Fintech* dummy equals one from the date at which the bank has any Fintech connection, and zero otherwise.

Variable	Estimate	Standard Error	P-value
Intercept	-2.4173	0.0062	<0.0001
Prin1	0.1322	0.0044	<0.0001
Prin2	0.0130	0.0030	<0.0001
Prin3	0.0481	0.0038	<0.0001
Prin4	-0.0078	0.0038	0.0390
Prin5	-0.0010	0.0040	0.8001
Prin6	0.0117	0.0038	0.0021
Prin7	0.0119	0.0038	0.0015
Prin8	0.0115	0.0039	0.0029
Prin9	0.0121	0.0039	0.0017
Prin10	0.0197	0.0037	<0.0001
Prin11	0.0217	0.0036	<0.0001
Prin12	0.0148	0.0039	0.0001
Prin13	0.0104	0.0040	0.0094
Prin14	0.0060	0.0042	0.1513
N	449,405		
Pseudo R^2	0.0320		

Table 4: Top Fintech Scoring Banks

This table lists the banks with the highest FinTech score estimated from the probit regression.

Bank Name	Fintech Score	Bank Name	Fintech Score
BANKERS BK NORTHEAST	2.99%	PRUDENTIAL B&TC	2.45%
FARM BUREAU BK FSB	2.98%	MIDFIRST BK	2.44%
PACIFIC COAST BKR BK	2.87%	MBNA AMERICA BK DE	2.43%
BANK OF AMER NA	2.83%	MBNA AMERICA DE NA	2.43%
UNITED BKR BK	2.81%	BANK HOLLAND	2.42%
JPMORGAN CHASE BK NA	2.77%	ADVANTA BK CORP	2.41%
CHASE MANHATTAN BK	2.77%	INDEPENDENT ST BK OF OH	2.37%
JPMORGAN CHASE BK	2.77%	GREAT LAKES BKR BK	2.37%
WELLS FARGO HSBC TRADE B	2.76%	LCA BK CORP	2.37%
PACIFIC CENTURY BK NA	2.76%	ALOSTAR BK OF CMRC	2.34%
DROVERS & MECHANICS BK	2.74%	AMERICAN INV BK NA	2.33%
FIRST NORTH AMER NB	2.71%	WEBBANK CORP	2.33%
SEARS NB	2.70%	WEBBANK	2.33%
WORLD FNCL NETWORK NB	2.70%	DIRECT MRCH CR CARD BK N	2.33%
COMENITY BK	2.70%	TOYOTA FNCL SVG BK	2.32%
WORLD FNCL NETWORK	2.70%	MORRIS PLAN CO TERRE HAU	2.31%
AMERIPRISE NAT TR BK	2.68%	GOLETA NB	2.31%
AMERIPRISE BK FSB	2.68%	COMMUNITY WEST BK NA	2.31%
MEDALLION BK	2.67%	COMMUNITY W BK NA	2.31%
PRIORITY BK	2.67%	BMW BK OF NORTH AMER	2.31%
FIRST UNION NB	2.65%	BMW BK OF N AMER	2.31%
NEXTBANK NA	2.65%	COMMUNITY BK OF NORTHERN	2.30%
BANK & TR OF PR	2.64%	MUSKEGON CMRC BK	2.29%
CELTIC BK CORP	2.64%	PLANTATION FED BK	2.28%
WORLD FNCL CAP BK	2.64%	ADMIRALS BK	2.28%
COMENITY CAP BK	2.64%	ARKANSAS BKR BK	2.26%
UNIVERSAL FNCL CORP	2.64%	LIBERTY BK FSB	2.26%
UNIVERSAL FC	2.64%	TRANSPORTATION ALLIANCE	2.25%
FLEET BK RI NA	2.63%	DBA TAB BANK	2.25%
FEDERAL SVGS BK	2.62%	TRANSPORTATION ALLI BK	2.25%
VOLKSWAGEN BK USA	2.62%	HARBOR BK NA	2.25%
WORLDS FOREMOST BK NA	2.59%	TARGET NB	2.24%
WORLD'S FOREMOST BK NA	2.59%	RETAILERS NB	2.24%
WORLDS FOREMOST BK	2.59%	FIRST BUS BK	2.24%
MERRICK BC	2.57%	AMERICAN INV FNCL	2.23%
BANKERS' BK	2.54%	WRIGHT EXPRESS FS CORP	2.22%
FIRST FS&L BK	2.52%	WEX BK	2.22%
INFIBANK	2.51%	WRIGHT EXPRESS FNCL SVC	2.22%
INFIBANK NA	2.51%	USAA SVG BK	2.21%
INDEPENDENT BKR BK OF FL	2.51%	FLEET NA BK	2.19%
FIRST CONSUMERS NB	2.50%	MORGAN GUARANTY TC OF NY	2.19%
CITIBANK NA	2.50%	WELLS FARGO BK NA	2.19%
ASSOCIATES CAP BK	2.49%	MONOGRAM CREDIT CARD BK	2.18%
ENERBANK USA	2.48%	MONOGRAM CR CARD BK	2.18%
HURLEY ST BK	2.47%	MIDWEST INDEPENDENT BK	2.17%
CITIBANK USA NA	2.47%	MIDWEST INDEP BK	2.17%
BANKERS BK OF THE WEST	2.47%	1ST FNCL BK USA	2.17%
PITNEY BOWES BK	2.47%	LIVE OAK BKG CO	2.16%
SALLIE MAE BK	2.46%	KENT CMRC BK	2.16%
MARLIN BUS BK	2.46%	FLORIDA CMNTY BK	2.16%

Table 5: Comparison of Call Report Variables Between Fintech Score Quartiles

Panel A: Variable Means by FinTech Score Quartiles

Fintech Quartile	Q1		Q2		Q3		Q4	
Variable	N	Mean	N	Mean	N	Mean	N	Mean
Non-interest Income:								
NII2	90,777	0.1139	90,755	0.1198	90,775	0.1207	90,742	0.1396
Trading Assets:								
SEC1	90,799	0.0534	90,763	0.0337	90,785	0.0285	90,747	0.0195
SEC2	90,799	0.2162	90,763	0.1915	90,785	0.1810	90,747	0.1487
SEC3	90,799	0.0539	90,763	0.0340	90,785	0.0287	90,747	0.0197
SEC4	90,799	0.2178	90,763	0.1929	90,785	0.1823	90,747	0.1496
BL31	90,799	0.0086	90,763	0.0093	90,785	0.0147	90,747	0.0532
Securitization and Off-Balance Sheet Activities:								
PVR13	90,799	0.0086	90,763	0.0093	90,785	0.0147	90,747	0.0532
SCR1	15,021	0.0005	20,822	0.0004	21,072	0.0016	24,233	0.0095
SCR2	6,807	0.0000	13,544	0.0000	28,514	0.0000	42,559	0.0008
PVR9	34,418	0.3839	35,183	0.4344	35,173	0.4419	34,992	0.4333
PVR11	29,724	9.8604	29,684	18.0513	29,508	34.5868	28,718	52.0146
PVR15	33,597	0.0031	34,965	0.0041	34,804	0.0055	34,787	0.0165
PVR16	90,799	0.0000	90,763	0.0000	90,785	0.0001	90,747	0.0001
On-Balance Sheet Lending:								
PVR7	41,508	0.2794	46,255	0.2452	46,334	0.2293	46,852	0.1953
PVR8	37,169	22.5384	44,198	23.3974	44,225	24.1831	45,168	25.0556
Brokered Deposits:								
PVR21	90,799	0.0046	90,763	0.0184	90,785	0.0353	90,747	0.0776
BSF2	90,799	0.0016	90,763	0.0071	90,785	0.0147	90,747	0.0367
BSF3	36,058	0.0023	36,046	0.0095	36,053	0.0158	36,038	0.0272
PVR22	30,451	0.0049	30,441	0.0115	30,447	0.0185	30,434	0.0337
PVR24	90,799	0.0050	90,763	0.0131	90,785	0.0230	90,747	0.0418
PVR20	90,799	0.0006	90,763	0.0007	90,785	0.0008	90,747	0.0052
PVR23	90,799	0.9191	90,763	0.9184	90,785	0.9107	90,747	0.8985

Panel B: Univariate Tests by FinTech Score Quartiles

Variable Description	Variable Name	Q4−Q1	t-stat
Non-interest Income			
Non-interest income/Total income	NII2	0.026	13.55***
Trading Assets			
Total Held to Maturity Securities (Amortized Cost, Cons.)/Total Domestic Assets	SEC1	-0.034	-80.00***
Total Available for Sale Securities (Amortized Cost, Cons.)/Total Domestic Assets	SEC2	-0.068	-98.85***
Held to Maturity Securities (Fair Value, Cons.)/Total Domestic Assets	SEC3	-0.034	-80.10***
Total Available for Sale Securities (Amortized Cost, Cons.)/Total Domestic Assets	SEC4	-0.068	-99.25***
Total Trading Assets/Total Domestic Assets	BL31	0.001	15.30***
Securitization and Off-Balance Sheet Activities			
Other Assets Serviced for Others (Excl. Closed End 1−4 Family Mortgage)/Total Assets	PVR13	0.044	12.45***
Ownership Interest in Securitized Credit Card Loans/Total Credit Card Loans	SCR1	0.009	11.10***
Ownership Interests in Securitized C&I loans/Total Domestic C&I Loans	SCR2	0.001	2.55**
Unused Off-balance Sheet C&I Loan Commitments/Total C&I Loans Domestic	PVR9	0.050	5.55***
All Other Unused OBS Commitments/Total Other Domestic Consumer Loans	PVR11	42.154	3.85***
Allowance for Credit Losses on OBS Exposures/Other Unused OBS Loan Commit	PVR15	0.013	3.10***
Allowance for Credit Losses on OBS Exposures/Total Domestic Assets	PVR16	0.000	49.35***
On-Balance Sheet Lending			
C&I loans <$100,000 Original Amt/Total Domestic C&I Loans	PVR7	-0.084	-62.25***
Average Loan Size for C&I loans <$100,000 Original Amount	PVR8	2.517	34.80***
Brokered Deposits			
Brokered Deposits/Total Domestic Deposits	PVR21	0.073	62.45***
Brokered Deposits<$100,000/Total Domestic Deposits	BSF2	0.035	98.55***
Brokered Deposits 100k−250k & Some IRAs/Total Domestic Deposits	BSF3	0.025	68.10***
Non-brokered Deposits Obtained via Deposit Listing Services/Total Domestic Deposits	PVR22	0.029	56.10***
Bank&depository Inst: Non-transaction Accts/Total Non-transaction Deposits	PVR24	0.037	95.75***
Bank&depository Inst: Transaction Accts/Total Domestic Deposits	PVR20	0.005	25.40***
Individuals Non-transaction Accounts/Total Non-transaction Deposits	PVR23	-0.021	-41.35***

Table 6: List of Fintech Keywords

Panel A: List of Fintech Keywords

fintech	automate	application programming interface	cyber security
technology	automation	api	social invest
mobile banking	automating	startup	social commerce
mobile wallet	robo-advising	regtech	crowdfunding
mobile payments	robo-advisor	regulation technology	p2p
mobile point-of-sale	robo-adviser	regulation technologies	peer-to-peer
mobile	machine learning	cloud	micro-insurance
online	artificial Intelligence	distributed ledger technology	insurtech
ecommerce	deep learning	distributed ledger technologies	digital cash
e-commerce	internet	dlt	digital wallet
electronic trading	initial coin offering	data monetization	digital credit
suptech	Crypto	telematic	digital lending
open banking	crypto currency	crypto-asset	Iot device
virtualisation	cryptocurrency	cryptosecurity	blockchain
innovation	crypto asset	crypto security	big data
aggregator comparison engine			

Panel B: Descriptive Statistics

Variable	Obs.	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
DEA efficiency score	361,238	0.424	0.255	0.007	0.203	0.415	0.633	1
SFA efficiency score	351,062	0.734	0.172	0.000	0.651	0.766	0.847	1
Fintech score	363,094	0.008	0.004	0.004	0.006	0.007	0.009	0.033
# of Keywords	16,576	4.953	9.585	0	0	2	6	255
Size	363,094	11.98	1.354	7.238	11.098	11.829	12.67	21.474
ROA	363,094	0.005	0.015	-0.461	0.002	0.005	0.008	2.309
ROE	363,094	0.046	0.072	-0.317	0.019	0.044	0.080	0.228
Bank equity	362,975	0.112	0.055	0.000	0.086	0.101	0.123	0.987

Table 7: Bank Efficiency and Fintech Score: Univariate Analysis

Panel A: Correlation Matrix

	SFA efficiency	DEA efficiency	Fintech score
DEA efficiency	0.2384		
	(0.0000)		
Fintech score	0.1278	0.4305	
	(0.0000)	(0.0000)	
# of Keywords	0.1288	0.1864	0.0289
	(0.0000)	(0.0005)	(0.0002)

Panel B: Univariate Tests of Fintech Score

By DEA Quartiles						
Variable	N (Q4)	N (Q1)	Mean (Q4)	Mean (Q1)	Q4−Q1	t_stat
Fintech score	90,282	90,336	0.010	0.006	0.004	243.03***

By SFA Quartiles						
Variable	N (Q4)	N (Q1)	Mean (Q4)	Mean (Q1)	Q4−Q1	t_stat
Fintech score	87,329	87,786	0.009	0.007	0.002	109.27***

Table 8: The Fintech Score and Bank Efficiency

The dependent variable is DEA efficiency Score in columns 1, 2 and 3 and SFA efficiency Score in columns 4, 5 and 6. The estimations in column 1 and 4 are done with OLS using year-quarter fixed effects. In column 2 and 5 we use fractional response regression with year-quarter fixed effects. In columns 3 and 6 we use fixed effects regressions. Size is the logarithm of total assets. Bank Equity is defined as the ratio of book value of equity to total assets. ROA is defined as net income divided by total assets. ROE is defined as the ratio of net income to book value of equity. The period of analysis is 2001−Q2 and 2016−Q3. The standard errors are clustered at the bank level. T-stats are in parenthesis and *, **, *** denote 10%, 5% and 1% significance levels, respectively.

Dependent Variable	DEA Efficiency Score			SFA Efficiency Score		
	(1)	(2)	(3)	(4)	(5)	(6)
Fintech Score$_{i,t-1}$	23.982***	24.444***	21.833***	2.299***	2.832***	1.584***
	(44.18)	(41.27)	(53.01)	(13.79)	(16.40)	(9.11)
Size$_{i,t-1}$	0.038***	0.039***	-0.006	0.029***	0.030***	0.049***
	(27.11)	(27.75)	(1.38)	(44.18)	(48.62)	(22.60)
ROA$_{i,t-1}$	0.110	0.159	0.073	0.087	1.096**	0.217*
	(0.80)	(0.97)	(0.94)	(0.96)	(2.12)	(1.74)
ROE$_{i,t-1}$	0.226***	0.229***	0.226***	1.088***	0.909***	0.761***
	(9.89)	(9.27)	(15.08)	(73.24)	(16.94)	(46.11)
Bank Equity$_{i,t-1}$	0.298***	0.318***	0.075***	0.201***	0.178***	0.129***
	(10.92)	(11.10)	(2.87)	(8.36)	(6.92)	(4.87)
Constant	-0.279***		0.257***	0.297***		0.101***
	(16.67)		(4.98)	(34.83)		(3.97)
Bank FE	NO	NO	YES	NO	NO	YES
Year-quarter FE	YES	YES	YES	YES	YES	YES
Observations	350,166	350,166	350,166	341,231	341,231	341,231
Adjusted R^2	0.30	.	0.22	0.44	.	0.40

Table 9: Fintech Score and Bank Efficiency: Instrumental Variable Method

This table reports the results of the instrumental variable analysis. Column 1 reports the first stage regression results, and column 2 reports the second stage regression results. DEA_High is a dummy variable that equals one if a bank's DEA efficiency is in the top quartile among all banks in a quarter, and zero otherwise. IT_Worker is the share of workers in the information technology sectors in the bank's headquarter county. Senior is the share of population older than 65 years in the bank's headquarter county. Size is the logarithm of total assets. Bank Equity is defined as the ratio of book value of equity to total assets. ROA is defined as net income divided by total assets. ROE is defined as the ratio of net income to book value of equity. The standard errors are clustered at the bank level. *T*-stats are in parenthesis and *, **, *** denote 10%, 5% and 1% significance levels, respectively.

Dependent Variable:	Finscore	DEA_High
	(1)	(2)
IT_Worker$_{i,t-1}$	0.596*	
	(1.869)	
Senior$_{i,t-1}$	-0.902***	
	(-5.496)	
Fintech Score$_{i,t-1}$		0.952***
		(5.851)
Size$_{i,t-1}$	0.083***	-0.043***
	(11.616)	(-3.275)
ROA$_{i,t-1}$	0.391*	-0.079
	(1.760)	(-0.437)
ROE$_{i,t-1}$	-0.178***	0.321***
	(-5.907)	(8.064)
HHI$_{i,t-1}$	-0.016	0.032
	(-0.357)	(0.629)
Equity$_{i,t-1}$	-0.387***	0.402***
	(-8.166)	(4.845)
Bank FE	YES	YES
Year-Quarter FE	YES	YES
Observations	347,482	347,482
Adjusted R2	0.530	0.117

Table 10: Differences-in-Differences Analysis Around the Global Financial Crisis

The Crisis dummy equals zero for period 2006Q1−2007Q3 and equals one for period 2009Q3−2011Q1. High Fintech score dummy equals zero for the first quartile of Fintech score in each quarter and equals one for the fourth quartile of Fintech score in each quarter in the first column. In the second column, high Fintech score dummy equals zero for the below median Fintech score in each quarter and equals one for above median Fintech score in each quarter. The standard errors are clustered at the bank level. *T*-stats are in parenthesis and *, **, *** denote 10%, 5% and 1% significance levels, respectively.

	DEA efficiency score	
	1st and 4th Quartile Fintech Score	Below and Above Median Fintech Score
High Fintech score × Crisis	0.201***	0.092***
	(16.86)	(16.18)
Crisis	-0.189***	-0.135***
	(17.95)	(20.91)
Fintech Score$_{i,t-1}$	7.483***	12.306***
	(5.52)	(9.05)
Size$_{i,t-1}$	0.130***	0.108***
	(3.13)	(4.39)
ROA$_{i,t-1}$	0.890	0.748
	(1.46)	(1.55)
ROE$_{i,t-1}$	0.150**	0.095*
	(2.15)	(1.87)
Bank Equity$_{i,t-1}$	0.599***	0.749***
	(2.76)	(6.14)
Constant	-1.222**	-1.011***
	(2.51)	(3.49)
Bank FE	YES	YES
Year-quarter FE	YES	YES
Observations	10,008	34,883
Adjusted R^2	0.29	0.23

Table 11: Differences-in-Differences Analysis Around the Dodd-Frank Act

The Dodd-Frank dummy equals zero for period 2004Q3−2010Q2 and equals one for period 2010Q3−2016Q2. High Fintech score dummy equals zero for the first quartile of Fintech score in each quarter and equals one for the fourth quartile of Fintech score in each quarter in the first column. In the second column, high Fintech score dummy equals zero for the below median Fintech score in each quarter and equals one for above median Fintech score in each quarter. The standard errors are clustered at the bank level. T-stats are in parenthesis and *, **, *** denote 10%, 5% and 1% significance levels, respectively.

	DEA efficiency score	
	1st and 4th Quartile Fintech Score	Below and Above Median Fintech Score
High Fintech score × Dodd_Frank	0.125***	0.093***
	(7.60)	(15.37)
Dodd_Frank	-0.219***	-0.228***
	(14.68)	(21.86)
Fintech Score$_{i,t-1}$	8.264***	14.292***
	(4.25)	(14.83)
Size$_{i,t-1}$	0.131***	0.068***
	(4.15)	(3.88)
ROA$_{i,t-1}$	-0.667	-1.268*
	(0.99)	(1.78)
ROE$_{i,t-1}$	0.142**	0.280***
	(2.03)	(4.19)
Bank Equity$_{i,t-1}$	0.217	0.455**
	(0.73)	(2.47)
Constant	-1.113***	-0.388*
	(3.08)	(1.90)
Bank FE	YES	YES
Year-quarter FE	YES	YES
Observations	12,396	58,309
Adjusted R^2	0.37	0.33

Table 12: Fintech Score Comparisons of Acquirer and Target Banks

This table presents univariate tests within the treated sample, consisting of acquirer−target bank pairs in the pre-merger quarter such that the Fintech score of the acquirer is lower than the Fintech score of the target bank. *, **, *** denote 10%, 5% and 1% significance levels, respectively.

Fintech Scores of	Quarter of merger		Quarter before merger		During-Before Merger	
Acquirer vs. Target Banks	N	Mean (Fintech score)	N	Mean (Fintech score)	Diff.	t-stat
FT (Acquirer) <FT (Target)	189	0.0077	189	0.0070	0.0007	2.75***
FT (Acquirer) $\geq$ FT (Target)	481	0.0102	481	0.0099	-0.0002	-0.87
Fintech Scores of	Quarter after merger		Quarter before merger		After-Before Merger	
Acquirer vs. Target Banks	N	Mean (Fintech score)	N	Mean (Fintech score)	Difference	t-stat
FT (Acquirer) <FT (Target)	189	0.0078	189	0.0070	0.0007	2.80***
FT (Acquirer) $\geq$ FT (Target)	481	0.0102	481	0.0099	-0.0002	-0.77

Table 13: Balance Tests Between Treated and Matched Control Groups in the Pre-merger Quarter

The Treated group consists of mergers in which a lower Fintech score bank acquires a higher Fintech score target. We only include merger events in which the target bank's pre-merger DEA score is lower than the pre-merger DEA score of the acquirer. We perform the analysis by comparing the pre-merger period (two quarters before the merger completion date) to the post-merger period (for the quarter of and the quarter after the merger completion date). To construct the Control group, we match non-merging banks to the treated sample banks in the pre-merger quarter using the coarsened exact matching method. The period of analysis is 2001−Q2 and 2016−Q3.

	Treated	Control	Difference	St. Err.	*t*-stat	p-value
Balance of Covariates						
DEA Efficiency	0.525	0.528	-0.004	0.028	-0.15	0.896
Size	13.426	13.168	0.258	0.212	1.20	0.225
ROA	0.005	0.005	0.001	0.001	0.60	0.547
ROE	0.050	0.045	0.005	0.007	0.60	0.537
Bank Equity	0.116	0.109	0.007	0.006	1.15	0.254
Parallel Trends						
Δ DEA$_{i,t-1}$	-0.003	0.000	-0.002	0.011	-0.20	0.861
Δ Fintech Score$_{i,t-1}$	0.000	0.000	0.000	0.000	-0.50	0.607

Table 14: Analysis of Bank M&As and DEA Efficiency

Merger takes the value zero for two quarters before the merger completion date and one for the quarter of and the quarter after the merger completion date. *Treated* takes the value one for bank mergers where the acquirer bank's pre-merger Fintech score (when *Merger* dummy is zero) is lower than the target bank's pre-merger Fintech score. We only include merger events in which the target bank's pre-merger DEA score is lower than the pre-merger DEA score of the acquirer. To construct the control sample, we match the treated sample banks in the pre-merger quarter with banks that did not acquire other banks in the same quarter using the coarsened exact matching method. The period of analysis is 2001−Q2 and 2016−Q3. The standard errors are clustered at the bank level. *T*-stats are in parenthesis and *, **, *** denote 10%, 5% and 1% significance levels, respectively.

Dependent variable	First stage Fintech score	Second stage DEA score	Diff-in-diff DEA score
Merger × Treated	0.100***		0.027**
	(4.33)		(2.11)
Fintech score (from first stage)		0.274***	
		(2.11)	
Merger	0.007	-0.013	-0.012
	(0.33)	(1.20)	(1.08)
Size	-0.008	-0.050	-0.052
	(0.13)	(1.21)	(1.27)
ROA	3.577	-1.022	-0.044
	(0.91)	(0.32)	(0.01)
ROE	-0.692	0.342	0.153
	(1.26)	(0.89)	(0.42)
Bank equity	-0.261	-0.611	-0.683
	(0.39)	(0.89)	(1.00)
Constant	0.764	1.158*	1.367**
	(0.81)	(1.80)	(2.14)
Bank FE	YES	YES	YES
Year-Quarter FE	YES	YES	YES
Observations	944	944	944
Adjusted R^2	0.16	0.32	0.32

Appendices

Table A1: List of Financial Variables Used in PCA

This table reports the variables used in the principal component analysis. Item numbers are from the Call Reports.

Asset Sales Ratio	RIADB514	RCON1460	RCFDB800
RCFDB795/RCON1763	RIADB515	**Securitization**	RCFDB801
RCFDB802/RCON1763	RIADB493	RCFDB735	RCFDB802
Balance Sheet	RIAD4012	RCFDB742	**Securitization Ratio**
RCFD2170	RIAD5416	RCFDB736	RCFDB502/RCON1763
RCON2170	RIAD4079	RCFDB743	**Other Variables**
RCFD1773	RIAD3196	RCFDB737	RCONA591
RCON5369	RIAD4230	RCFDB744	RCONB550
RCONB528	RIADB485	RCFDB738	RCONB552
RCON1763	RIADB486	RCFDB745	RCONB549
RCON3123	RIADA220	RIADB749	RCONB551
RCFD3545	RIAD4069	RIADB756	RCON2385
RCON1545	RIAD4107	RIADB750	RCON2200
RCON3387	RIAD4185	RIADB757	RCONB557
RCONB561	**Pass-due**	RIADB751	RCON1766
RCONB538	RCON5459	RIADB758	RCONB529
RCONB539	RCON5460	RIADB752	RCONA591/RCON2170
Brokered Source of Funds	RCON5461	RIADB759	RCONB557/RCON2170
RCON2365/RCON2200	RCFDB707	RCFDB501	RCONB550/RCON2200
RCON2343/RCON2200	RCFDB708	RCFDB502	RCONB549/RCONB550
Charge-offs	RCFDB709	RCFDB765	RCONB552/RCON2200
RIAD4605	RCFDB710	RCFDB766	RCONB551/RCON2200
RIADC079	RCFDB711	RCFDB768	RCONB550/RCON2385
RIAD5523	RCONB575	RCFDB769	RCONB552/RCON2385
RIAD3123	RCONB576	RIADB771	RCONB551/RCON2385
RIADB522	RCONB577	RIADB772	RCON2365/RCON2385
Deposits	**Quarterly Average**	RIADB774	RCON3123/(RCON5369+RCONB528)
RCON2365	RCON3360	RIADB775	RCON5369/RCONB528
RCON2343	RCONB562	RCFDB792	
RCON2200	RCON3485	RCFDB793	
Income	**Real Estate**	RCFDB794	
RIAD4010	RCON5367	RCFDB795	
RIAD4073	RCON5368	RCFDB799	

Table A2: Comparison of Call Report Variables Between Banks with Zero and Positive Fintech Scores Estimated from the Random Forest Model

Variable Description	**Variable Name**	**Positive − Zero**	**t-stat**
Fintech score from two-stage model		0.002	31.24***
Non-interest Income			
Non-interest income / Total income	NII2	0.045	10.15***
Trading Assets			
Total held to maturity securities (amortized cost, cons.) / Total domestic assets	SEC1	-0.009	-4.95***
Total available for sale securities (amortized cost, cons.) / Total domestic assets	SEC2	-0.032	-11.25***
Held to maturity securities (fair value, consolidated) / Total domestic assets	SEC3	-0.009	-4.98***
Total available for sale securities (amortized cost, cons.) / Total domestic assets	SEC4	-0.032	-11.25***
Total trading assets / Total domestic assets	BL31	0.004	31.75***
Securitization and Off-Balance Sheet Activities			
Other assets serviced for others (excl. closed end 1-4 family mortgage) / Total domestic assets	PVR13	0.164	16.05***
Ownership interest in securitized credit card loans / Total credit card loans	SCR1	0.021	16.30***
Ownership interests in securitized C&I loans / Total domestic C&I loans	SCR2	0.002	3.28***
Unused off-balance sheet C&I loan commitments / Total C&I loans domestic	PVR9	0.075	2.55***
All other unused OBS commitments / Total other domestic consumer loans	PVR11	23.628	0.450
Allowance for credit losses on OBS exposures / Other unused OBS loan commit	PVR15	152.308	14.95***
Allowance for credit losses on OBS exposures / Total domestic assets	PVR16	0.000	17.75***
On-Balance Sheet Lending			
C&I loans <$100,000 original amount / Total domestic C&I loans	PVR7	-0.068	-12.45***
Average loan size for C&I loans <$100,000 original amount	PVR8	1.562	5.37***
Brokered Deposits			
Brokered deposits / Total domestic deposits	PVR21	0.098	57.40***
Brokered deposits <$100,000 / Total domestic deposits	BSF2	0.043	38.60***
Brokered deposits $100k−$250k & some IRAs / Total domestic deposits	BSF3	0.021	16.79***
Non-brokered deposits obtained via deposit listing services / Total domestic deposits	PVR22	0.018	9.32***
Bank & depository institutions: non-transaction accounts / Total non-transaction deposits	PVR24	0.018	13.55***
Bank & depository institutions: transaction accounts / Total domestic deposits	PVR20	0.004	7.55***
Individuals non-transaction accounts / Total non-transaction deposits	PVR23	0.006	3.05***

Table A3: Comparison of Call Report Variables Using Fintech Scores Estimated from Unsupervised Machine Learning Model

Variable Description	**Variable Name**	**C1 − C0**	**t-stat**
Fintech score from two-stage model		0.005	68.70***
Non-interest Income			
Non-interest income / Total income	NII2	0.049	5.80***
Trading Assets			
Total held to maturity securities (amortized cost, cons.) / Total domestic assets	SEC1	-0.023	-12.08***
Total available for sale securities (amortized cost, cons.) / Total domestic assets	SEC2	-0.070	-21.88***
Held to maturity securities (fair value, consolidated) / Total domestic assets	SEC3	-0.023	-12.15***
Total available for sale securities (amortized cost, cons.) / Total domestic assets	SEC4	-0.071	-22.10***
Total trading assets / Total domestic assets	BL31	0.001	8.37***
Securitization and Off-Balance Sheet Activities			
Other assets serviced for others (excl. closed end 1-4 family mortgage) / Total domestic assets	PVR13	0.075	5.72***
Ownership interest in securitized credit card loans / Total credit card loans	SCR1	0.065	26.02***
Ownership interests in securitized C&I loans / Total domestic C&I loans	SCR2	-0.000	-0.44
Unused off-balance sheet C&I loan commitments / Total C&I loans domestic	PVR9	-0.019	-0.40
All other unused OBS commitments / Total other domestic consumer loans	PVR11	-0.463	-0.01
Allowance for credit losses on OBS exposures / Other unused OBS loan commit	PVR15	0.007	0.58
Allowance for credit losses on OBS exposures / Total domestic assets	PVR16	0.000	7.33***
On-Balance Sheet Lending			
C&I loans <$100,000 original amount / Total domestic C&I loans	PVR7	-0.107	-18.52***
Average loan size for C&I loans <$100,000 original amount	PVR8	6.195	19.09***
Brokered Deposits			
Brokered deposits / Total domestic deposits	PVR21	0.088	21.49***
Brokered deposits <$100,000 / Total domestic deposits	BSF2	0.019	13.88***
Brokered deposits $100k−$250k & some IRAs / Total domestic deposits	BSF3	0.046	34.93***
Non-brokered deposits obtained via deposit listing services / Total domestic deposits	PVR22	0.234	127.30***
Bank & depository institutions: non-transaction accounts / Total non-transaction deposits	PVR24	0.533	395.50***
Bank & depository institutions: transaction accounts / Total domestic deposits	PVR20	0.109	186.60***
Individuals non-transaction accounts / Total non-transaction deposits	PVR23	-0.604	-312.16***

Table A4: Fintech Score and SFA Bank Efficiency: Instrumental Variable Method

This table reports the results of the instrumental variable method using SFA efficiency as the outcome variable. Column 1 reports the first stage regression results, and column 2 reports the second stage regression results. SFA_High is a dummy variable that equals one if a bank's DEA efficiency is in the top quartile among all banks in a quarter, and zero otherwise. IT_Worker is the share of workers in the information technology sectors in the bank's headquarter county. Senior is the share of population older than 65 years in the bank's headquarter county. Size is the logarithm of total assets. Bank Equity is defined as the ratio of book value of equity to total assets. ROA is defined as net income divided by total assets. ROE is defined as the ratio of net income to book value of equity. The standard errors are clustered at the bank level. T-stats are in parenthesis and *, **, *** denote 10%, 5% and 1% significance levels, respectively.

Dependent Variable:	Finscore	SFA_High
	(1)	(2)
$\text{Information}_{i,t-1}$	0.592*	
	(1.833)	
$\text{Senior}_{i,t-1}$	-0.840***	
	(-5.126)	
$\text{Finscore}_{i,t-1}$		1.006***
		(5.075)
$\text{Size}_{i,t-1}$	0.077***	0.037***
	(10.676)	(2.665)
$\text{ROA}_{i,t-1}$	0.490*	0.076
	(1.745)	(0.274)
$\text{ROE}_{i,t-1}$	-0.103***	0.811***
	(-3.079)	(21.203)
$\text{HHI}_{i,t-1}$	-0.018	0.022
	(-0.384)	(0.392)
$\text{Equity}_{i,t-1}$	-0.351***	1.263***
	(-3.785)	(9.746)
Bank FE	YES	YES
Year-Quarter FE	YES	YES
Observations	338,610	338,610
Adjusted R^2	0.530	0.041

Table A5: Difference-in-Differences Analysis Around the Global Financial Crisis: SFA Efficiency

The *Crisis* dummy equals zero for period 2005Q2−2007Q2 and equals one for period 2009Q4−2011Q4. High Fintech score dummy equals zero for the first quartile of Fintech score in each quarter and equals one for the fourth quartile of Fintech score in each quarter in the first column. In the second column, high Fintech score dummy equals zero for the below median Fintech score in each quarter and equals one for above median Fintech score in each quarter. The standard errors are clustered at the bank level. T-stats are in parenthesis and *, **, *** denote 10%, 5% and 1% significance levels, respectively.

	SFA Efficiency Score	
	1st and 4th Quartile Fintech Score	Below and Above Median Fintech Score
High Fintech score × Crisis	0.004***	0.001***
	(4.48)	(3.81)
Crisis	0.048***	0.046***
	(54.13)	(104.99)
Fintech Score$_{i,t-1}$	-0.346***	-0.206***
	(3.18)	(3.27)
Size$_{i,t-1}$	0.007***	0.005***
	(3.75)	(5.88)
ROA$_{i,t-1}$	0.017	0.014
	(0.46)	(0.37)
ROE$_{i,t-1}$	0.005	0.012***
	(0.88)	(2.70)
Bank Equity$_{i,t-1}$	0.025**	0.035***
	(2.26)	(4.14)
Bank FE	YES	YES
Year-quarter FE	YES	YES
Observations	9,950	37,374
Adjusted R^2	0.85	0.85

Table A6: Differences-in-Differences Analysis Around the Dodd-Frank Act: SFA Efficiency

The $Dodd-Frank$ dummy equals zero for period 2004Q3−2010Q1 and equals one for period 2010Q2−2016Q1. High Fintech score dummy equals zero for the first quartile of Fintech score in each quarter and equals one for the fourth quartile of Fintech score in each quarter in the first column. In the second column, high Fintech score dummy equals zero for the below median Fintech score in each quarter and equals one for above median Fintech score in each quarter. The standard errors are clustered at the bank level. T-stats are in parenthesis and *, **, *** denote 10%, 5% and 1% significance levels, respectively.

	SFA Efficiency Score	
	1st and 4th Quartile Fintech Score	Below and Above Median Fintech Score
High Fintech score × Dodd-Frank	0.005***	0.003***
	(4.31)	(6.42)
Dodd-Frank	-0.083***	-0.081***
	(43.23)	(64.17)
Fintech Score$_{i,t-1}$	-0.051	-0.008
	(0.39)	(0.20)
Size$_{i,t-1}$	0.005***	0.003***
	(3.60)	(5.19)
ROA$_{i,t-1}$	0.149***	0.087***
	(4.01)	(2.68)
ROE$_{i,t-1}$	0.008	0.010**
	(1.17)	(2.56)
Bank Equity$_{i,t-1}$	0.026**	0.026***
	(1.99)	(3.84)
Bank FE	YES	YES
Year-quarter FE	YES	YES
Observations	12,448	57,923
Adjusted R^2	0.85	0.85

Table A7: Fintech Score Comparisons of Acquirer and Target Banks: SFA Efficiency

This table presents univariate tests within the treated sample, which includes acquirer−target bank pairs in pre-merger quarter in which the Fintech score (FT) of the acquirer is lower than the fintech score of the target bank. We also restrict the acquirer-target pairs in the treated sample to only those merger events where the target banks' pre-merger SFA score is lower than the pre-merger SFA score of the acquirer. *, **, *** denote 10%, 5% and 1% significance levels, respectively.

Fintech Scores of	Quarter of merger		Quarter before merger		During-Before Merger	
Acquirer vs. Target Banks	N	Mean (Fintech score)	N	Mean (Fintech score)	Diff.	t-stat
FT (Acquirer) <FT (Target)	248	0.0080	248	0.0073	0.0007	3.08***
FT (Acquirer) $\geq$ FT (Target)	509	0.0101	509	0.0100	0.0001	0.25
Fintech Scores of	Quarter after merger		Quarter before merger		After-Before Merger	
Acquirer vs. Target Banks	N	Mean (Fintech score)	N	Mean (Fintech score)	Difference	t-stat
FT (Acquirer) <FT (Target)	248	0.0081	248	0.0073	0.0008	3.36***
FT (Acquirer) $\geq$ FT (Target)	509	0.0101	509	0.0100	0.0001	0.26

Table A8: Balance Tests Between Treated and Matched Control Groups in the Pre-merger Quarter: SFA EFficiency

The Treated group consists of mergers in which a lower Fintech score bank acquires a higher Fintech score target. We only include merger events in which the target bank's pre-merger SFA score is lower than the pre-merger SFA score of the acquirer. We perform the analysis by comparing the pre-merger period (two quarters before the merger completion date) to the post-merger period (for the quarter of and the quarter after the merger completion date). To construct the Control group, we match non-merging banks to the treated sample banks in the pre-merger quarter using the coarsened exact matching method. The period of analysis is 2001−Q2 and 2016−Q3.

	Treated	Control	Difference	St. Err.	*t*-stat	p-value
Balance of Covariates						
SFA Efficiency	0.829	0.826	0.003	0.011	0.25	0.801
Size	13.264	13.258	0.007	0.154	0.05	0.968
ROA	0.007	0.007	0.001	0.001	0.60	0.540
ROE	0.063	0.063	0.000	0.005	0.01	0.991
Bank Equity	0.113	0.113	0.001	0.004	0.1	0.925
Parallel Trends						
Δ SFA$_{i,t-1}$	0.015	0.017	0.002	0.009	0.15	0.880
Δ Fintech Score$_{i,t-1}$	0.000	0.000	0.000	0.000	1.35	0.180

Table A9: Analysis of Bank M&As and SFA Efficiency

Merger takes the value zero for two quarters before the merger completion date and one for the quarter of and the quarter after the merger completion date. *Treated* takes the value one for bank mergers where the acquirer bank's pre-merger Fintech score (when *Merger* dummy is zero) is lower than the target bank's pre-merger Fintech score. We only include merger events in which the target bank's pre-merger SFA score is lower than the pre-merger SFA score of the acquirer. To construct the control sample, we match the treated sample banks in the pre-merger quarter with banks that did not acquire other banks in the same quarter using the coarsened exact matching method. The period of analysis is 2001−Q2 and 2016−Q3. The standard errors are clustered at the bank level. *T*-stats are in parenthesis and *, **, *** denote 10%, 5% and 1% significance levels, respectively.

Dependent variable	First stage Fintech score	Second stage SFA score	Diff-in-diff SFA score
Merger × Treated	0.058***		0.016**
	(3.04)		(2.10)
Fintech score (from first stage)		0.283**	
		(2.10)	
Merger	0.011	-0.023***	-0.020***
	(0.61)	(3.00)	(2.99)
Size	0.069	0.008	0.027
	(1.49)	(0.38)	(1.52)
ROA	-2.136	1.087	0.483
	(0.68)	(1.59)	(0.75)
ROE	-0.280	0.047	-0.032
	(0.77)	(0.42)	(0.31)
Bank equity	-1.254***	0.529	0.174
	(3.45)	(1.38)	(0.52)
Constant	0.045	0.376	0.389
	(0.07)	(1.39)	(1.43)
Bank FE	YES	YES	YES
Year-Quarter FE	YES	YES	YES
Observations	1608	1608	1608
Adjusted R^2	0.11	0.30	0.30