

Open data and API adoption of U.S. banks*

Xiangyu Lin, S. Sarah Zhang, Markos Zachariadis†

This version: June 2024

Abstract

The drive for open data in financial services is reflected in banks' adoption of external application programming interfaces (APIs), a technology that provides bank customers with greater access to their banking and transaction data and enable them to share their data with other third-party financial institutions and FinTechs, also promoted as “open banking”. We analyze determinants of API adoption by U.S. banks and subsequent changes in bank performance during the period 2007 to 2022. We show that banks that adopt APIs experience an increase in Return on Assets (*ROA*) and *Tobin's Q* and a decrease in loan loss provisions, particularly after President Biden's executive order announcement that encouraged greater data portability and ownership of bank customers. Consistent with existing literature, we find greater effects for smaller banks with lower levels of credit risk and competitive pressure. Overall, our study shows the potential impact of open banking and data portability on bank competition and performance and highlights broader implications for financial inclusion and borrower welfare.

JEL classifications: G21; G28; O32; O33

Keywords: API adoption; open banking; data portability; FinTech; bank performance

* We gratefully acknowledge the support of Seth Benzell, Jonathan Hersh and Mark Boyd for sharing a subsample of their data. We would like to thank Serafeim Tsoukas (discussant), Francesco Vallasca (discussant), Seongjin Park (discussant) Marie Dutordoir, Olga Kolokolova, Maria Marchica, Eirini Konstantinidi, Chi-Yang Tsou, Xiaomeng Shi, Yibing Wang, Yuanjin Zhang, Cheng Xue and Aung Yee for helpful comments, and conference participants at EFMA (2024), Behavioural Finance Working Group (2024), QMUL Economics and Finance Workshop (2024), Lancaster-Manchester-Warwick Joint PhD Workshop on Quantitative Finance and Fintech (2024), CIMSC 2024 ICMA, AMBS Doctoral best paper competition (2024), 2023 FMA European and FMARC for useful comments and suggestions. All errors are ours.

† Alliance Manchester Business School, The University of Manchester, Manchester M15 6PB, UK. Email address: xiangyu.lin@postgrad.manchester.ac.uk (X. Lin), Corresponding author: sarah.zhang@manchester.ac.uk (S. S. Zhang) and markos.zachariadis@manchester.ac.uk (M. Zachariadis).

1. Introduction

The recent decades have seen a transformation of the banking landscape, with respect to the adoption of Information Technology (IT) of traditional banks and the advent of FinTechs which created additional pressure on the retail banking sector. Furthermore, the financial services sector has experienced a drive for open data in recent years, reflected in regulatory open banking initiatives in several countries and jurisdictions that have compelled banks to provide their customers with greater access to their data (such as their transaction and credit history) and more control to share such data with other third-party providers and financial institutions. For example, European Union countries and the United Kingdom are working towards an open financial framework with the implementation of the EU Payment Services Directive 2 (PSD2) and the open banking initiative.¹ While discussions on open banking in the U.S. have been more focused on voluntary market-based initiatives, recent regulatory initiatives aim at enhancing data portability for bank customers. In particular, the Biden administration published an executive order on July 9, 2021, that “encourage[s] the Consumer Financial Protection Bureau (CFPB) to issue rules allowing customers to download their bank data and take it with them.”²

In order to make bank customer data more easily accessible and portable, banks practically adopt external application programming interfaces (APIs) that can be accessed by external third parties. However, adopting APIs can pose a significant threat to the information monopoly of banks over their customers' data under the traditional bank model and result in more competition from other banks and FinTech startups (Marquez, 2002; Zachariadis & Ozcan, 2016). For example, third-party providers can access users' bank account information through HSBC's open banking API. Once users grant permission, the application can securely access account balances, transaction history, and more. Additionally, the HSBC API supports transfers and payment processing with user consent, helping users manage their finances more effectively.³ Furthermore, the adoption can come with considerable investment costs (due to its impact on the entire banking architecture (Dinçkol et al., 2023)) as well as cybersecurity risks

¹ See Babina et al. (2022) for an overview of open banking regulation in different countries.

² See details of President Biden's executive order here: <https://www.whitehouse.gov/briefing-room/statements-releases/2021/07/09/fact-sheet-executive-order-on-promoting-competition-in-the-american-economy/>

³ See the examples of HSBC APIs here: <https://www.hsbc.co.uk/help/open-banking/>; <https://develop.hsbc.com/api-overview/treasury-payment-initiation>

and data privacy challenges (Benzell et al., 2023). However, proponents of open banking argue that the increase of available data and data sharing required by the adoption of an open banking strategy may expand the markets for both lending and payment services. Furthermore, the introduction of APIs by banks can increase their productivity through more efficient data sharing with third parties, facilitate the automation of lending processes, and lead to more accurate risk assessments through access to more data (Babina et al., 2022).

In this paper, we analyze the determinants of API adoption by U.S. banks and the implications of API adoption for information transmission via APIs and bank performance, before and after the announcement of President Biden's executive order. In particular, we aim to answer the following questions: (1) How does banks' competitive pressure and credit risk considerations affect API adoption?, (2) What are the implications of bank API adoption for Tobin's Q, Return on Assets (*ROA*), loan loss provision, and lending costs?, (3) Do effects differ prior to and after the announcement of Biden's executive order in July 2021? (4) Do the effects of APIs on bank performance exhibit heterogeneity across different types of APIs?

We construct a unique panel dataset covering the API profiles of 713 U.S.-listed banks from 2007-2022. We find that 22 banks adopted external APIs from 2007 to 2022, representing 3% of our overall sample of banks. We categorize APIs into two types: Read-only APIs provide access to customer transactions and credit data but don't allow data manipulation, limiting their value creation for banks. Conversely, Read-Write APIs offer more flexibility by enabling proactive data manipulation and real-time sharing, allowing third parties to offer detailed customer insights and potentially improve financial risk assessments.

While larger banks with greater amount of assets and deposits are more likely to adopt APIs, we also find that banks that experience less competitive pressure (measured by product similarity and fluidity, as proposed by Hoberg and Philipps (2014; 2016)⁴) and that have lower levels of loan loss provisions (*LLP*) and non-performing loans have a higher probability of adoption. After adopting their first API, we further find that bank performance improves. In particular, we observe an economically

⁴ Hoberg & Phillips (2016) and Hoberg et al., (2014) proposed measures of product similarity and fluidity that related to the intensity of direct competitors and potential competitors within an industry group based on the similarity of 10-K filings.

and statistically significant increase in ROA and Tobin's Q. In addition, loan loss provisions (*LLP*) significantly decreases, while banks' market value increases. Accounting for the time periods before and after Biden's executive order announcement, we find that many effects stem from the period after the announcement when regulatory pressure for data portability has increased, although early adopting banks experienced an improvement in ROA prior to the announcement.

To address endogeneity concerns, we include the inverse mills ratio suggested by Heckman (1976) to control for self-selection bias. In addition, we calculate the difference between bank performance of banks that adopt APIs and a control group of banks that do not adopt APIs during our sample period using propensity score matching (PSM), finding consistent results for the improvement in performance after API adoption. We further find heterogeneous changes in bank performance after API adoption, as banks with lower levels of non-performing loans, loan loss provision and competition (measured by product similarity and fluidity), and smaller size with respect to total assets and deposits experience greater increases in performance after API adoption.

Our paper contributes to an emerging body of research on open banking and data portability. Existing studies on FinTech competition and digital information show that the increasing availability of digital information are valuable for default prediction, leading to superior lending decisions (Berg et al., 2020). However, open banking and bank data portability might result in some unintended consequences for bank and FinTech competition and borrower welfare, as shown in the complementary theoretical frameworks of He et al. (2023) and Parlour et al. (2022). In particular, Parlour et al. (2022) show that allowing voluntary data porting by consumers can lead to possible unraveling (i.e. customers are compelled to share data as non-disclosure will be perceived negatively by financial service providers) and negative data externality for customers who do not share their data. While FinTechs can benefit from additional access to customer data in an open banking environment, the loss of customer information due to FinTech competition can disrupt information spillover within traditional banks (for example by using payment data to learn about consumers' credit quality). Similarly, He et al. (2023) show that an open banking environment might widen the screening ability gap between lenders and over-empower FinTechs due to greater access to customer data. While we find support for a decrease

in loan loss provision, consistent with the information generation hypothesis (Koetter & Noth, 2013)⁵, our empirical results further support the notion that banks that experience higher levels of competitive pressure and credit risk do not experience an improvement in performance.

Our study is further related to Benzell et al. (2023), who study how API adoption can create value through an “inverted” firm strategy⁶, and Babina et al. (2022), who analyze the effects of open banking regulation on venture capital investment in FinTech firms in an international study. Our study differs from Benzell et al. (2023) as we provide insight into the decision to adopt APIs by banks who might not necessarily follow an inverted firm strategy but rather adopt APIs to enable easier and safer sharing of bank customer data and to gain access to additional customer information and provide wider service offering through activities with third-party providers. Babina et al. (2023) analyze the effects of open banking regulation on venture capital investment in different countries and on households and SMEs in the UK. In contrast, we focus on the implications of API and open banking adoption for traditional banks in the U.S. In contrast to the promotion and introduction of open banking policies studied by Babina et al. (2023), we further distinguish between voluntary adoption by U.S. banks prior to the announcement of President Biden’s executive order, and semi-mandatory adoption (due to with an increase in regulatory pressure) after the announcement, providing insight into the differential effects of API adoption in different regulatory regimes.

Our findings are further related to the impact of technological innovations and FinTechs on the banking sector. Existing studies have examined the effect of technology on bank performance using measures such as the adoption of ATMs, internet and online banking, payment technology, and robo-advice (such as Beccalli, 2007; Berger, 2003; Hernández-Murillo et al., 2010; D’Andrea & Limodio, 2023; D’Acunto et al. 2019), or overall bank IT expenditure and number of personal computers (Pierri & Timmer, 2022). In our paper, we use external API adoption (based on the first adopted API) as an indicator of individual banks’ intention of widening data access and control for their customers. While API adoption can be considered a form of digital investment, API adoption differs from traditional bank

⁵ Koetter & Noth (2013) postulate that the use of information technology can resolve information asymmetries and thereby help to mitigate competitive pressure.

⁶ The inverted firm strategy refers to the notion that firms can create value externally by becoming platforms for external service providers (Benzell et al., 2023).

digitalization efforts as they enable banks to share and access additional data and functionality through external third-party applications and services in accordance with their customers' consent. Therefore, APIs can help facilitate banks' collaboration with third-party providers to offer a broader range of (mostly digital) services and products without needing to develop these services internally, but rather relying on the expertise and reach of their third-party collaborators.

Our paper is further related to the general body of FinTech research and its impact on the lending industry.⁷ Existing FinTech research has extensively researched the impact of the growth of FinTechs on competition in the banking sector and the competitive relationship between traditional bank lending and FinTech lenders. As the role of competition in the decision to adopt an open banking strategy is not well understood, we provide insight into the role of competitive pressures on the decision to adopt an open banking strategy.

Our study has significant practical and policy implications. Our empirical study complements to existing studies on open banking and data portability that have provided qualitative and theoretical evidence regarding the potential impact of open banking and data portability on bank competition and performance, and wider effects on financial inclusion and borrower welfare. We illustrate the changes in credit risk, Tobin's Q and implications for banks' lending costs and revenues after API adoption. In addition, we highlight the effects of regulatory and government interventions on API adoption as we distinguish between voluntary adoption before Biden's executive order announcement, and the shift to a more or less mandatory open banking regime.

The remainder of our paper is structured as follows: In section 2, we provide details of our sample selection and descriptive statistics and present our results in section 3, as well as further robustness checks and cross-sectional analysis in section 4. Section 5 finally concludes.

⁷ Berg et al. (2022) and Vives (2019) provide an overview of the digital disruption for the banking sector and lending industry, respectively. Other studies that focus on the impact of FinTech and P2P lending include Buchak et al. (2018); Fuster et al. (2019); Gopal & Schnabl (2022), and Tang (2019), Fuster et al., 2019; Vives & Ye, 2023.

2. Data and Sample Selection

We manually collect evidence of banks' external API adoption and analyze their effects on bank performance. Our sample consists of all publicly listed bank firms (i.e., firms with SIC (Standard Industrial Classification) codes between 6020 and 6999) level information from the Compustat and CRSP database for the years covered by our sample period (2007-2022), including banks' financial information. We exclude firm-year observations with missing values for the key variables. Our final data sample consists of 713 banks and 5,692 firm-year observations. API data are manually collected from the bank's official developer websites as well as from a few APIs and open-banking integration platforms, such as Programmable Web, APIdashboard, Openbanking tracker, Platformable, and APItracker. Our key financial data stem from Compustat, CRSP, and BankFocus.

2.1 Sample Selection

We extract financial data for U.S.-listed banks from Compustat annual filings and BankFocus. We exclude companies with year-end share prices below \$1, companies with total assets and book-to-equity annual observations of firms with negative values, and firms with insufficient financial data to calculate the variables of interest in our analysis. After applying these selection criteria, our final sample consists of 5,692 firm-year observations (713 unique banks). To eliminate the potential effect of outliers, we winsorize all continuous variables at the 1% and 99% percentiles to alleviate the potential impact of outliers.

Measures of bank performance

We use *Tobin's Q* and *ROA* to measure bank performance and further analyze changes in banks' market value (*MRKTV*) (similar to Benzell et al. (2023)). Following Hirtle et al. (2020), we analyze different components of bank earnings, in particular net interest margin (*NIM*), the ratio of other noninterest expenses to total assets (*NIE*), the ratio of noninterest income to total assets (*NII*) and the ratio of loan loss provisions to total asset (*LLP*). We further control for several bank-level characteristics as is common in the literature, such as *Bank Size*, *Firm Age*, *Capital Intensity*, *Leverage*, *Tangibility*, *Cash Holding*, *Cost of Capital*, *Deposit Ratio*, *Loan Ratio*, and the number of employees (*Employee*). The specific variable names, definitions, and sources are provided in Table 1.

Measures of banks' API adoption and classification

API implementation is measured by a dummy variable $API_Adopt_{i,t}$ that takes the value 1 if bank i has adopted public APIs in or before year t , and 0 otherwise, during the sample period from 2007-2022.⁸ In addition, we construct the continuous variable $API_Intensity$, which captures the number of APIs owned by bank i in year t . We also incorporate a dummy variable for API_bank , where bank i is classified as an API_bank if it has adopted at least one API during the sample period, and otherwise not. We manually collect the API information (such as API name, categories, description, adoption date, etc.) from the bank's official developer websites as well as from API and open-banking integration platforms, such as Programmable Web, APIdashboard, Openbanking tracker, Platformable and APItracker. Our API adoption dataset contains public APIs that financial institutions use to communicate data and information externally (with third-party providers and partners). We further classify APIs based on the information they process, as suggested by the Financial Conduct Authority's (FCA) definition of third-party providers, i.e. read-only related Account Information Service Providers (AISP) and read-and-write related Payment Initiation Service Providers (PISP).⁹ According to the Financial Conduct Authority (2017), an authorized AISP can access a bank customer's bank account data via bank API if the customer provides explicit consent. This type of "read-only" access means that they can only see the information but cannot access the account. In addition, an authorized PISP can not only view financial account data with customer permission but is also authorized to make payments on behalf of the customer, thus having "read-write" access. We therefore classify APIs depending on whether an API can only request data from a specified resource (*Read-only*) or also send data to a server to create/update a resource (*Read-Write*). While *Read-only* APIs are relevant when access to data is needed, such as bank account, transactional and balance information, *Read-Write* APIs enable a stronger connection and deeper integration between parties, such as for payment initiation processes. In our data, 51.13% of APIs of U.S. public banks are *Read-only* APIs whereas 48.87% of APIs are *Read-Write* APIs, allowing

⁸ Our sample period spans from 2007 to 2022, beginning with the adoption of the first API by a publicly listed U.S. bank and the initial enactment of the European Payment Services Directive (PSD1) in 2007. This directive introduced a new category of cross-border payment services and allowed non-bank institutions to provide financial services securely. Our aim is to capture all the effects associated with open banking and open APIs in U.S. market.

⁹ For further information on AISP and PISP: <https://www.fca.org.uk/consumers/account-information-payment-initiation-services>.

us to separately analyze the influence of the two API types.

Measures of competition and competitive threats

To analyze the influence of competitive factors in the banking industry, we use the product similarity and product fluidity measures for individual banks, from the Hoberg-Phillips Data Library¹⁰. Product similarity measures the availability of alternatives to a company's products and services, while product fluidity measures the frequency of new products and services appearing on the market. Hoberg & Phillips (2016) use textual analysis of product descriptions in companies' 10-K files to identify competitors for each company. They find that firms with higher *product similarity* scores are more likely to mention competitive pressure posed by rivals. In contrast, *product fluidity* measures the ex-ante competitive threats faced by a firm in its product market that captures changes in rival firms' products relative to the firm (Hoberg et al. (2014)). Using both measures, we can capture both competitive pressures by existing and potential rivals based on product descriptions in the 10-K filings.

2.2 Descriptive Statistics

We present the summary statistics of the overall sample (Full Sample) and the subsamples of banks with and without APIs in Table 2. The table reports selected mean and standard deviations (SD) for relevant variables of our analyses for the full sample and two subsamples. The statistics are from 2007 to 2022, covering the major period of growing API adoption.

[Insert Table 2 about here]

The summary statistics in Table 2 give insight into the characteristics of banks with and without APIs. We find that only around 3% of the U.S.-listed banks in our sample have adopted APIs (22 out of 713 banks) by 2022. Banks that have adopted public APIs tend to have larger average total assets (*Bank Size*), have higher levels of *ROA* and non-interest income (*NII*) as well as market value (*MRKTV*) and *Tobin's Q*, compared to banks that have not adopted public APIs. Furthermore, banks with APIs have larger loans and more employees, higher leverage and cash holdings, lower cost of capital, generally lower credit risk (as determined by non-performing loans (*NPL*)), higher operational efficiency

¹⁰ Hoberg and Phillips Data Library can be found here: <https://hobergphillips.tuck.dartmouth.edu>.

(*Efficiency*), relatively lower product similarity (*Similarity*) and product fluidity (*Fluidity*). We formally test the determinants of API adoption using a logistic regression in Section 4.1.

Table A1 in Appendix A further shows the overall correlation between the main independent and dependent variables and bank characteristics. The correlation coefficients show positive correlations of API adoption and API intensity with *Tobin's Q*, market value, and *ROA*, and negative correlations with credit risk (*NPL*, *LLP*, and *Z-Score*) and competition metrics (*Similarity* and *Fluidity*).

2.3 API Adoption by Banks

We analyze the API adoption of banks over time and present our results in Figure 1 and Figure 2.

[Insert Figure 1 and Figure 2 about here]

Figure 1 shows the growth and cumulative APIs and API-banks for the sample period of 2007 to 2022. The bar chart illustrates the cumulative number of API-banks and banking APIs, while the line chart illustrates the number of API-banks and APIs added each year. There is a sizeable increase in API adoption in 2020 and 2021, pointing towards an increasing trend of adopting APIs in the banking industry. Figure 2 displays the distribution of banks that adopted APIs (API banks) by each U.S. state in the sample period.

3. Empirical Results

3.1 Determinants of API Adoption

Studies have shown that bank characteristics such as profitability, bank size, presence in urban markets, membership in bank holding companies, branching intensity, capital-to-asset ratio, non-performing loan ratio, etc., play an essential role in adopting bank websites and online banking services (Furst et al., 2002; DeYoung et al., 2007; Hernandez-Murillo et al. 2010). To analyze the factors that determine API adoption based on bank-specific characteristics, we apply the following logistic regression:

$$API_{bank_i} = \alpha + \beta Bank\ Characteristics_{i,t} + \omega_t + \epsilon_{i,t} \quad (1)$$

where API_bank_i is a time-invariant dummy variable that equals one if bank i has adopted APIs during the sample period and 0 otherwise. Once a bank adopts an API within our sample period (2007-2022), it is defined as an API bank for the entire period. We include the following measures for bank characteristics prior to API adoption: the age of the bank in years (*Firm Age*), the size of the bank (*Bank Size*) measured as the log of bank i 's total assets, the percentage of non-performing loans (*NPL*) and loan loss provision (*LLP*), and the degree of financial leverage (*Leverage*), calculated as the ratio between selected financial assets of the banking sector and its total equity, and measures of liquidity such as cash holdings (*Cash Holding*). We further include measures of operational efficiency (*Efficiency*), *Capital Intensity*, cybersecurity risk exposure (*Cybersecurity*), Tier 1 capital ratio (*Tier 1*), and risk-weighted assets (*RWA*) as well as the competition metrics (*Similarity* and *Fluidity*) (see Table 1 for variable definitions). To examine the factors influencing API adoption, we include bank characteristics based on observations in the years before the API was adopted. For API banks, only the observations prior to API adoption were retained, whereas for non-API banks, the observations before 2017 were included.¹¹ We further control for year-fixed effects ω_t . Table 3 presents the results of the logistic regression in Equation (1) based on a sample of 3,911 firm-year observations.

[Insert Table 3 about here]

We find that bank size (*Bank Size*) and operating efficiency (*Efficiency*) are positive determinants of API adoption, as banks with more resources and funds are more likely to cover the considerable investment costs of adopting APIs. In addition, we observe that younger banks (*Firm Age*) with lower *Capital Intensity* are more likely to adopt API technologies. On the credit side, banks with better credit risk management (as measured by non-performing loan ratio (*NPL*), loan loss provision (*LLP*), and risk-weighted assets (*RWA*)) are more likely to adopt APIs. *Tier 1 capital* has also been found to correlate with increased API adoption. Finally, banks with lower levels of product similarity (*Similarity*) within the banking industry are more likely to adopt and implement APIs.

¹¹ Due to an acceleration in API adoption by publicly listed U.S. banks after 2017, we maintain baseline consistency by only retaining observations up to and including 2017 for banks that did not adopt APIs during the sample period. This approach allows us to capture the impact of bank characteristics on API adoption effectively.

3.2 API Adoption and Bank Performance

We examine the changes in bank performance after API adoption using a two-step regression approach by Heckman (1976) to correct for self-selection. In the first stage regression, we use a Probit model to estimate the selection equation and generate the inverse mills ratio (*IMR*) to account for self-selection which is included in the second stage regression. The Probit model includes control variables and broadband access within the state that the bank is headquartered in (*Broadband*), provided by the Internet Access for High-Speed Services data on broadband adoption¹², as an exogenous variable. As the adoption of APIs is implemented at the bank's headquarters, broadband penetration can have a positive effect on the decision to adopt and develop APIs, with minor consequences to banks' overall performance. In the second stage of regression, shown in equation (3), we include *IMR* to account for self-selection bias. *Controls_{i,t}* are a set of time-varying firm control variables, respectively; and $\epsilon_{i,t}$ is the error term. Specifically, we estimate the following model:

$$Pr(API_Adopt_{i,t}) = \alpha + \beta Broadband_{i,t} + \gamma Controls_{i,t} + \varphi_i + \omega_t + \epsilon_{i,t} \quad (2)$$

$$Perf_{i,t} = \alpha + \beta API_Adopt_{i,t} + \gamma Controls_{i,t} + \delta IMR_{i,t} + \varphi_i + \omega_t + \epsilon_{i,t} \quad (3)$$

where i and t denote bank and year, respectively. The main variable of interest in the baseline model is *API_Adopt_{i,t}*, which is an indicator set to one if bank i has adopted APIs in or before year t , and zero if bank i has not adopted any APIs in year t . The regression model includes year-fixed effects ω_t and firm-fixed effect φ_i and standard errors are clustered at the firm level.

For the dependent variable *Perf_{i,t}*, we apply *Tobin's Q* and Return on Assets (*ROA*). The results of the baseline regressions are shown in Table 4:

[Insert Table 4 about here]

Table 4 shows the effects of API adoption on bank performance. Column (0) shows the Heckman first-stage regression, and columns (1)-(4) illustrate the second-stage regression results of the API adoption on bank performance, including *IMR* to account for self-selection. Columns (1)- (4) report how the API affects *Tobin's Q* and *ROA*, respectively. We include firm and year-fixed effects in columns

¹²We retrieve the data from FCC's semi-annual report on broadband adoption, this data includes information on the total number of broadband subscribers in a state, the prevalence of different broadband speeds, and adoption rates. Broadband information is gathered from: <https://www.fcc.gov/BroadbandData>.

(2) and (4), while year-fixed effects are not included in columns (1) and (3). We further apply clustered standard error at the firm level.

We observe a consistently significant positive coefficient for API on bank performance, indicating that after banks adopt APIs, ROA and Tobin's Q are on average higher than before adoption. Robustness checks using various fixed effects and clustered standard errors provide consistent results on the positive influence of API implementation on banking performance. More specifically, the results for *Tobin's Q* and *ROA* consistently show statistically and economically significant positive coefficients of *API_Adopt*, indicating positive changes in bank performance after API adoption.

Further, we also examined the impact of API intensity on bank performance, with results presented in Table A3. As we have shown before, the initial adoption of APIs can significantly enhance performance due to technological advancements, improvements in business processes, or the establishment of data portability relationships, increasing the number of APIs did not provide significant additional marginal utility. Banks may have already deployed APIs to meet their business needs, and thus, additional APIs may not yield extra value. While the quantity of APIs increases, their effectiveness may rely more on the quality and integration of the APIs rather than sheer quantity.

3.2.1 Stacked effects over time

While our previous results show that adopting APIs is followed by improvements in bank performance, we further analyze dynamic changes in bank performance in the years before and after API adoption. To address concerns about the reliability of OLS estimators and to consider the heterogeneity of API adoption effects across time and treatment units, we apply robust estimators to estimate staggered effects. We follow the estimation procedure introduced by Callaway and Sant'Anna (2021) that accounts for heterogeneity of treatment effects across time. As APIs are adopted in different years, observed covariates should be taken into account to satisfy the parallel trend assumption (Baker et al., 2022; Callaway and Sant'Anna, 2021), allowing us to deal more flexibly with heterogeneity across time. We thus re-estimate our regression model in Equation (4) to account for these dynamic effects.

$$Perf_{i,t} = \alpha + \sum_{k=-5}^2 \delta_k \times D_{i,t}^k + \gamma Control_{i,t} + \delta IMR_{i,t} + \varphi_i + \omega_t + \epsilon_{i,t} \quad (4)$$

where dummy variables $D_{i,t}^k$ is one up to 5 periods preceding and 2 periods following the API event,¹³ and δ_k is the parameter of interest, with the binary variables $D_{i,t}^k$ capturing a window of 8 years around API adoption, rather than focusing on changes in the same year and after adoption as in Equation (3).

[Insert Figure 3 and Table 5 about here]

To focus on the dynamic changes around API adoption, Figure 3 and Table 5 present the results of the adapted regression of *ROA* and *Tobin's Q* in 5 years prior and 2 years after API adoption [-5, +2], based on Callaway and Sant'Anna (2021). Regardless of the chosen estimator, coefficients for the pre-adoption years of the API remain close to zero, suggesting no significant differences in *Tobin's Q* and *ROA* between the treatment and control groups before APIs were introduced. Consistent with the trend shown in Figures 3a and 3b, column (1) of Table 5 shows a persistent positive trend of *Tobin's Q* after API adoption over time. Results differ from our baseline model, as we take into account the staggered adoption of bank APIs and changes in *ROA* and *Tobin's Q* over time. In contrast, changes in *ROA* are comparably weaker, albeit positive and significant in the year when the API is adopted. Our results imply that increases in *Tobin's Q*, i.e. growth opportunities of the adopting banks, are positive and persistent, while increases in *ROA* are volatile and floating.

3.2.2 Lagged Analysis of Bank Earnings, Costs, and Value

To better understand the impact of APIs on different types of bank earnings, costs, and market value, we analyze bank performance using cost-, income- and value-related components following Hirtle et al. (2020) and Benzell (2023). In particular, non-interest expenses (*NIE*) include staff costs, administrative expenses, branch maintenance, and operating costs, while the cost of loan-loss provision (*LLP*) reflects costs related to the bank's projected loan losses. In terms of bank income, we distinguish between non-

¹³ To explore the dynamic impact of API adoption on bank performance, we set the time interval starting from -5 years to capture pre-adoption performance data. Given that API adoption among publicly listed U.S. banks is concentrated post-2020 and our sample ends in 2022, this limits our ability to observe the long-term effects of API adoption, particularly impacts beyond the third year post-adoption. Considering data availability, selecting a time interval from -5 to +2 years not only maximizes the use of existing data but also allows us to conduct a reasoned comparison and analysis of the effects before and after API adoption.

interest income (*NII*) and net interest margin (*NIM*). Following Benzell et al. (2023), we further analyze banks' market value (*MRKTV*). We regress our performance variables on lagged variables of API adoption.

$$Perf_{i,t} = \alpha + \beta API_Adopt_{i,t-j} + \gamma Controls_{i,t} + \delta IMR_{i,t} + \varphi_i + \omega_t + \epsilon_{i,t} \quad (5)$$

where $API_Adopt_{i,t-j}$ is a binary dummy value that includes up to 2 lags after API adoption (where $j=0, 1, 2$), i.e. measuring the relationship with bank profitability in the year of adoption and in the two years after. We also include year and firm-fixed effects in all models. The results are presented in Table 6.

[Insert Table 6 about here]

Results in Panel A show a significant decline in operating costs (*NIE*) and loan loss provision (*LLP*), while there is an insignificant decline in *NII* and *NIM*. As APIs allow real-time monitoring of borrowers' financial conditions and market dynamics, more accurate and up-to-date credit assessment and reduction in risk of bad debts might lead to a reduction in costs and losses related to non-performing loans. API adoption reduces *NIE* by automating manual processes such as data entry, customer verification, and transaction processing, significantly decreasing reliance on labour and reducing labour costs. Furthermore, APIs facilitate a shift towards digital banking services, reducing dependence on physical branches and further cutting operational costs and staffing requirements. Panel B further shows an increase in banks' market value (in line with previous results of Benzell et al. (2023) who find positive effects of API adoption on firms' market value).

3.3 Impact of President Biden's executive order

To test for variation of effects before and after the announcement of President Biden's executive order to enable data portability of bank customers, we conduct a Difference-in-Differences (DID) analysis for banks that adopted APIs and those that did not, before and after the issuance of President Biden's executive order. We construct a dummy variable *Biden_Post* to capture the impact of President Biden's executive order, issued in July 2021, where the period before Biden's executive order equals 0, and for 2021 and subsequent years the executive order equals 1. We estimate the following regression

model, which extends Equation (3) by incorporating an interaction term between API adoption and *Biden_Post*.

$$Perf_{i,t} = \alpha + \beta API_{i,t} + \partial API_{i,t} \times Biden_Post_t + \gamma Controls_{i,t} + \varphi_i + \omega_t + \epsilon_{i,t} \quad (6)$$

where i and t denote bank and year, respectively. The main variable of interest in the baseline model is $API_Adopt_{i,t}$ and *Biden_Post*. The regression model includes year-fixed effects ω_t and firm-fixed effect φ_i and standard errors are clustered at the firm level. The results of these analyses are presented in Table 7.

[Insert Table 7 about here]

The results indicate that while ROA improved for adopting banks prior to President Biden's executive order, the improvements in Tobin's Q and reduction in LLP were affecting banks mostly after the announcement of the executive order. The positive effects after API adoption are thus at least partially derived from early adopting banks, while the effects on LLP and Tobin's Q seem to mostly stem from the time period after the executive order announcement. We thus confirm existing literature that highlights the beneficial effects of regulatory promotion of open banking (Babina et al, 2023), but also find beneficial effects of early voluntary adopters who see an improvement in ROA after adoption.

4. Robustness checks and cross-sectional variation

4.1 Propensity Score Matching Analysis

To address potential self-selection bias and mitigate systematic differences in characteristics between treatment and control banks, we use propensity score matching (PSM) based on Rosenbaum & Rubin (1983). In particular, we include banks in our treatment group of "API banks" which have at least one API per bank during our sample period, whereas the control group of "non-API banks" has no APIs. We use year-by-year matching, using the bank characteristics from the year before the adoption of an API to match each API bank in the treatment group with a similar non-API bank in the control group. To balance the treatment and control groups, we use the nearest neighbor one-to-five matching with 1% caliper (without replacement) and identify bank-year observations in the treatment and control groups

to match based on *Bank Size*, *Firm age*, *Capital Intensity*, *Cash Holding*, and *Cost of Capital*.

[Insert Table 8 about here]

To ensure that the treatment and control groups are balanced, we compare and report matching statistics based on control variables as in Equation (5), including diagnostic statistics for differences in bank characteristics between the treatment and control groups in Panel A of Table 8. Mean treatment effects are reported in Table A2. To estimate the propensity score of treated firms, we use a logistic regression of the treatment indicator (API_bank_i) on the relevant firm-level variables. We re-estimate the regression model in Equation (5) based on our matched sample, including year-fixed effects and firm-fixed effects.

The results from the OLS regression of PSM analysis are presented in Panel B of Table 8. The results based on the PSM sample are consistent with our baseline finding that API adoption has a positive relationship with bank performance. In particular, we examine the differences in *Tobin's Q* and *ROA* between API banks and non-API banks after API adoption and find that differences in *Tobin's Q* and *ROA* are positive and persistent up to 2 years after API adoption.

4.2 Influence of Different API Categories

We further analyze the differential effects of different types of APIs, in particular Read-only (which facilitate information sharing only) or Read-Write APIs (which allow data to be modified in (near) real-time in addition to allowing access to the data) (see Section 2.1). Read-only APIs and Read-Write APIs can play different roles in the various services that banks offer their customers. Read-only include APIs that contain key terms in their descriptions, such as transaction queries, transaction details, transaction history, account balances, or statements. In contrast, Read/Write APIs include keywords like payment requests, payment transfers, payment types, and authorization of open access. R-APIs generally provide access to broader aggregated transaction and historical information, which can potentially enhance a bank's ability to make more informed lending decisions. On the other hand, RW-APIs facilitate deeper integration between parties, enabling smoother and more efficient payment processes and reducing payment processing costs. Using an OLS regression model based on the baseline model, we analyze the

changes in performance after the adoption of Read-Write APIs and Read-only APIs.

$$Perf_{i,t} = \alpha + \beta API\ category_{i,t} + \gamma Controls_{i,t} + \delta IMR_{i,t} + \varphi_i + \omega_t + \epsilon_{i,t} \quad (7)$$

where $API\ category_{i,t}$ is one of the two dummy variables R_Adopt or R/W_Adopt which is one if bank i has adopted Read-only APIs or Read/Write APIs in or before year t . The results for the adoption variables of the two API types are presented in Panel A of Table 9.

To analyze the influence of the relative proportion of adopted R-APIs and R/W-APIs, we also use the variables R_Ratio and R/W_Ratio which range from 0 to 1 and measure the relative proportion of Read-only APIs and Read-Write APIs relative to the total number of adopted APIs. The regression results for the relative proportion of R- and RW-APIs are shown in Panel B, Table 9.

[Insert Table 9 about here]

The results in Panel A demonstrate that adopting Read-Write APIs can lead to greater changes in *Tobin's Q* and *ROA* than adopting Read-only APIs. In particular, we find that after adopting Read-Write APIs, both *Tobin's Q* and *ROA* improve. As shown in Panel A column (2), this change represents around a 1.6% increase compared to the average *Tobin's Q* of 1.018. The change in *Tobin's Q* and *ROA* is further related to the proportion of R/W APIs relative to the total number, illustrated by the positive coefficient of R/W_Ratio .

While Read-only APIs enable access to bank customers' transaction and credit information, they do not allow third-party providers to directly manipulate data and entities and are therefore somewhat limited in terms of value creation for the bank. In contrast, Read-Write APIs give banks greater flexibility and interactivity, allowing banks to manipulate data more proactively but also supporting dynamic real-time data sharing, which allows third parties to provide more detailed information about existing customers and possibly enable better assessment of financial credit risks.

4.3 Cross-sectional analysis

We further explore the cross-sectional heterogeneity of the effects depending on bank characteristics. Specifically, we provide further evidence regarding the cross-sectional differences of the effects of API adoption on bank performance measures based on a variety of variables: (i) using the level non-performing loans (*NPL*) and loan loss provisions (*LLP*) to measure banks' credit risk and credit quality,

(ii) banks' product similarity (*Similarity*) and product fluidity (*Fluidity*) to measure competitive pressure for individual banks (Hoberg et al., 2014; Hoberg & Phillips, 2016), (iii) banks' total asset (*Bank Size*) and banks' total deposit (*Deposit*) amount to measure the role of bank size.

In particular, we include a set of dummy variables (*Bank_Character_i*) based on different bank characteristics which equal one if: (i) bank *i*'s *NPL* is higher than the median of sample banks' (*High_NPL=1*) and the bank's *LLP* is higher than the median of sample banks' (*High_LLPL=1*), (ii) bank *i*'s similarity index is higher than the median of sample banks (*High_Similarity=1*) and product fluidity is higher than the median of sample banks (*High_Fluidity=1*). (iii) bank *i*'s total asset is higher than the median of sample banks (*High_Asset=1*) and bank *i*'s total deposit is higher than the median of sample banks (*High_Deposit=1*). We estimate the following regression model, which extends Equation (3) by incorporating an interaction term between API adoption and bank characteristics.

$$Perf_{i,t} = \alpha + \beta API_{i,t} + \partial API_{i,t} \times Bank_Character_i + \gamma Controls_{i,t} + \varphi_i + \omega_t + \epsilon_{i,t} \quad (8)$$

We report on the heterogeneity of results depending on the level of LLP and NPL in Table 10, Panel A.

[Insert Table 10 about here]

The results in Panel A Table 10 show greater increases in performance after API adoption for banks with a high level of non-performing loans (*High_NPL=1*) as well as loan loss provision (*High_LLPL=1*) are lower compared to banks with low NPL (*High_NPL=0*) and LLP (*High_LLPL=0*), particularly for *ROA*. In line with the theoretical predictions of Parlour et al. (2022) and He et al. (2022), our results show that banks with lower credit risk experience greater performance increases.

The results in Panel B Table 10 show greater increases in bank performance for banks with lower levels of product similarity (*High_Similarity*) and lower levels of product fluidity (*High_Fluidity*) but weaker for banks that experience greater competitive pressures, especially for *Tobin's Q*.

In line with previous discussions on banking competition and credit risk considerations, our results illustrate that banks that experience higher levels of competitive pressure and credit risk do not experience significant improvements in performance after API adoption compared to banks with lower levels of competitive pressure and credit risk. This might be due to the disruption of internal information spillover as postulated by Parlour et al. (2022) or alternatively due to the different strategies of banks

with a higher threat of existing competitors (measured by product similarity) as well as potential competitors (measured by product fluidity) that might adopt APIs to retain clients and maintain their market position rather than improve their performance.

Finally, Columns (1) and (2) of Panel C in Table 10 shows the interaction effects of API adoption with the bank size dummy *High_Asset* on *Tobin's Q* and *ROA*, while Columns (3) and (4) show the interaction effects of *High_Deposit* and API adoption on bank performance. The results show that banks with a lower share in bank deposits and total assets are more likely to experience positive changes in *ROA* after API adoption, while the interaction effects are not significant for *Tobin's Q*. Thus, the potential improvement in bank *ROA* after API adoption is captured by smaller banks with fewer total assets and a smaller share in bank deposits, whilst larger banks might have other incentives for adopting APIs other than profitability and performance.

5. Conclusion

This paper provides an empirical analysis of bank performance for banks that adopt an open banking strategy. While there is an open banking mandate from regulators in European countries, our empirical results shed light on the implications of mostly voluntary adoption of open banking strategy and APIs by U.S. banks. We find that bank performance improves after API adoption, supporting recent literature that IT adoption and IT investment by banks can have positive effects on firm performance (Dadoukis et al., 2021). However, we illustrate that effects are heterogeneous, as smaller banks with lower levels of credit risk and competitive pressure experience greater increases in performances compared to banks with higher levels of credit risk and subject to more banking competition.

Our paper contributes to the literatures around the impact of technological innovations on the banking sector (Hernández-Murillo et al., 2010; D'Andrea & Limodio, 2023; Pierri & Timmer, 2022), and provides new evidence of open banking strategies and the adoption of external APIs, which facilitates enhanced connectivity and interaction between banks and a range of third parties. Moreover, our paper is an essential complement to previous theoretical frameworks on open banking, digital portability, and information spillover (He et al., 2022; Parlour et al., 2022). Simultaneously, our findings

complement and expand upon existing research on FinTech and its impact on the lending industry (Berg et al., 2022; Buchak et al., 2018; Fuster et al., 2019; Gopal & Schnabl, 2022; Tang, 2019; Vives, 2019).

Meanwhile, our study has practical and policy implications, as open banking initiatives and policies can have significant consequences for the banking landscape as well as for FinTech competition. Compared with the open banking mandates adopted in Europe and the UK, our findings provide insights into the effects of open banking adoption in the U.S. banking industry. Our results show that open banking adoption has heterogeneous effects on different types of banks, as some banks experience increases in performance whilst other banks might adopt APIs due to competitive or compliance reasons which may or may not result in improvements in performance. Distinguishing between the time period before and after the announcement of President Biden's executive order on July 9, 2021, that demanded data portability for bank customers, we illustrate differential effects for early and later adopting banks. Our results are consistent with existing open banking research which shows that an increasing amount of data sharing and data portability can create certain unintended consequences and adverse effects on some banks. Considering the timeline around the Biden administration's executive order, we noted increased regulatory pressures on data portability post-release, although banks that adopted early saw ROA improvements before the order. Our cross-sectional analysis also implementing excessively stringent regulations may add additional regulatory burdens on banks and increase discrepancies between banks even further. While it is beyond the scope of this paper, future work might explore different levels of open banking regulation in other countries and the implications for other stakeholders.

References

- Babina, T., Buchak, G., & Gornall, W. (2022). Customer Data Access and Fintech Entry: Early Evidence from Open Banking. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4071214>
- Baker, A. C., Larcker, D. F., & Wang, C. C. Y. (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, 144(2), 370–395. <https://doi.org/10.1016/j.jfineco.2022.01.004>
- Beccalli, E. (2007). Does IT investment improve bank performance? Evidence from Europe. *Journal of Banking & Finance*, 31(7), 2205–2230. <https://doi.org/10.1016/j.jbankfin.2006.10.022>
- Benzell, S., Hersh, J. S., Van Alstyne, M. W., & Lagarda, G. (2023). How APIs Create Growth by Inverting the Firm. *Management Science*, 1-22. <https://doi.org/10.1287/mnsc.2023.4968>
- Berg, T., Burg, V., Gombović, A., & Puri, M. (2020). On the Rise of FinTechs: Credit Scoring Using Digital Footprints. *The Review of Financial Studies*, 33(7), 2845–2897. <https://doi.org/10.1093/rfs/hhz099>
- Berg, T., Fuster, A., & Puri, M. (2022). FinTech Lending. *Annual Review of Financial Economics*, 14(1), 187–207. <https://doi.org/10.1146/annurev-financial-101521-112042>
- Berger, A. N. (2003). The Economic Effects of Technological Progress: Evidence from the Banking Industry. *Journal of Money, Credit and Banking*, 35(2), 141–176. <http://dx.doi.org/10.2139/ssrn.332900>
- Buchak, G., Matvos, G., Piskorski, T., & Seru, A. (2018). Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics*, 130(3), 453–483. <https://doi.org/10.1016/j.jfineco.2018.03.011>
- Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. “Difference-in-differences with Multiple Time Periods.” *Journal of Econometrics*, 225 (2): 200–230. <https://doi.org/manchester.idm.oclc.org/10.1016/j.jeconom.2020.12.001>
- Dadoukis, A., Fiaschetti, M., & Fusi, G. (2021). IT adoption and bank performance during the Covid-19 pandemic. *Economics Letters*, 204, 109904. <https://doi.org/10.1016/j.econlet.2021.109904>
- D’Andrea, A., & Limodio, N. (2023). High-Speed Internet, Financial Technology, and Banking. *Management Science*, 70(2), 773-798. <https://doi.org/10.1287/mnsc.2023.4703>
- DeYoung, R., Lang, W. W., & Nolle, D. L. (2007). How the Internet affects output and performance at community banks. *Journal of Banking & Finance*, 31(4), 1033–1060. <https://doi.org/10.1016/j.jbankfin.2006.10.003>
- Dinçkol, D., Ozcan, P., & Zachariadis, M. (2023). Regulatory standards and consequences for industry architecture: The case of UK Open Banking. *Research Policy*, 52(6), 104760. <https://doi.org/10.1016/j.respol.2023.104760>
- Florackis, C., Louca, C., Michaely, R., & Weber, M. (2022). Cybersecurity Risk. *Review of Financial Studies*, 36(1), 351–407. <https://doi.org/10.1093/rfs/hhac024>
- Furst, K., Lang, W. W., & Nolle, D. E. (2002). Internet Banking. *Journal of Financial Services Research*, 22(1/2), 95-117. <https://doi.org/10.1023/a:1016012703620>
- Fuster, A., Plosser, M., Schnabl, P., & Vickery, J. (2019). The Role of Technology in Mortgage Lending. *The Review of Financial Studies*, 32(5), 1854–1899. <https://doi.org/10.1093/rfs/hhz018>
- Gopal, M., & Schnabl, P. (2022). The Rise of Finance Companies and FinTech Lenders in Small Business Lending. *The Review of Financial Studies*, 35(11), 4859–4901. <https://doi.org/10.1093/rfs/hhac034>
- He, Z., Huang, J., & Zhou, J. (2023). Open banking: Credit market competition when borrowers own the data. *Journal of Financial Economics*, 147(2), 449–474. <https://doi.org/10.1016/j.jfineco.2022.12.003>
- Heckman, J. J. (1976). The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models. In *Annals of Economic and Social Measurement*, 5(4), 475–492. NBER.
- Hernández-Murillo, R., Llobet, G., & Fuentes, R. (2010). Strategic online banking adoption. *Journal of Banking & Finance*, 34(7), 1650–1663. <https://doi.org/10.1016/j.jbankfin.2010.03.011>
- Hirtle, B., Kovner, A., & Plosser, M. (2020). The Impact of Supervision on Bank Performance. *Journal of Finance*, 75(5), 2765–2808. <https://doi.org/10.1111/jofi.12964>

- Hoberg, G., & Phillips, G. (2016). Text-Based Network Industries and Endogenous Product Differentiation. *Journal of Political Economy*, 124(5), 1423–1465. <https://doi.org/10.1086/688176>
- Hoberg, G., Phillips, G., & Prabhala, N. (2014). Product Market Threats, Payouts, and Financial Flexibility. *Journal of Finance*, 69(1), 293–324. <https://doi.org/10.1111/jofi.12050>
- Koetter, M., & Noth, F. (2013). IT use, productivity, and market power in banking. *Journal of Financial Stability*, 9(4), 695–704. <https://doi.org/10.1016/j.jfs.2012.06.001>
- Marquez, R. (2002). Competition, Adverse Selection, and Information Dispersion in the Banking Industry. *The Review of Financial Studies*, 15(3), 901–926. <http://dx.doi.org/10.2139/ssrn.264531>
- Parlour, C. A., Rajan, U., & Zhu, H. (2022). When FinTech Competes for Payment Flows. *The Review of Financial Studies*, 35(11), 4985–5024. <https://doi.org/10.1093/rfs/hhac022>
- Pierri, N., & Timmer, Y. (2022). The importance of technology in banking during a crisis. *Journal of Monetary Economics*, 128, 88–104. <https://doi.org/10.1016/j.jmoneco.2022.04.001>
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55. <https://doi.org/10.1093/biomet/70.1.41>
- Tang, H. (2019). Peer-to-Peer Lenders Versus Banks: Substitutes or Complements? *The Review of Financial Studies*, 32(5), 1900–1938. <https://doi.org/10.1093/rfs/hhy137>
- Vives, X. (2019). Digital Disruption in Banking. *Annual Review of Financial Economics*, 11(1), 243–272. <https://doi.org/10.1146/annurev-financial-100719-120854>
- Vives, X., & Ye, Z. (2023). Information Technology and Lender Competition. *Working Paper*. <https://doi.org/10.2139/ssrn.3863988>
- Zachariadis, M., & Ozcan, P. (2016). The API Economy and Digital Transformation in Financial Services: The case of Open Banking. *Working Paper*. <http://dx.doi.org/10.2139/ssrn.2975199>

Figures and Tables

Figure 1. Number of API-banks and APIs by year

This figure shows the growth and cumulative number of API-banks and growth and cumulative number of adopted bank APIs during the sample interval from 2007 to 2022. The bar chart shows the cumulative number of API-bank (APIs), and the line chart shows the number of new API-banks (APIs) for each year.

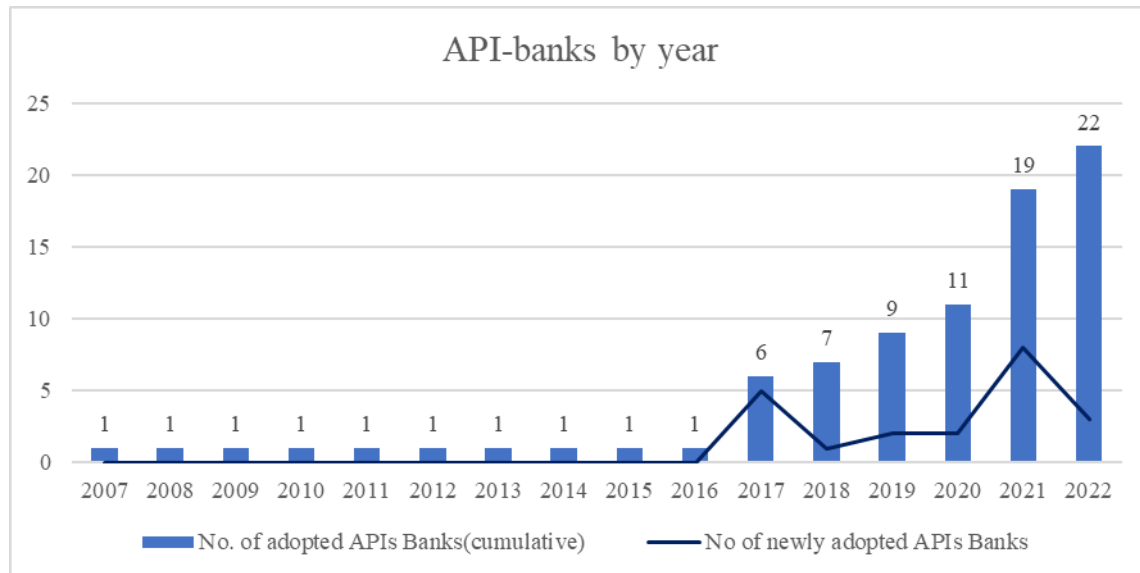


Figure 1a. Number of API-banks over time

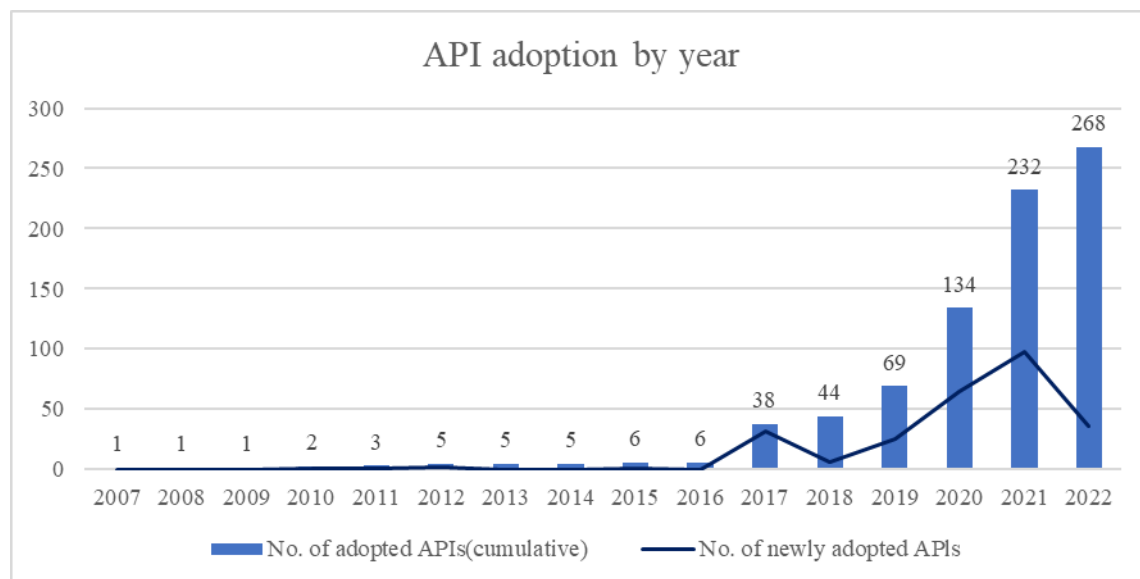


Figure 1b. Number of adopted APIs over time

Figure 2. Distribution of API banks in the U.S.

This map displays the distribution of banks that adopted APIs (API banks) by each state. The banks in 14 states (California, New York, North Carolina, Pennsylvania, Tennessee, Kansas, Michigan, Minnesota, Missouri, Nevada, Ohio, Rhode Island, Texas) adopted at least one API from 2007 to 2022.

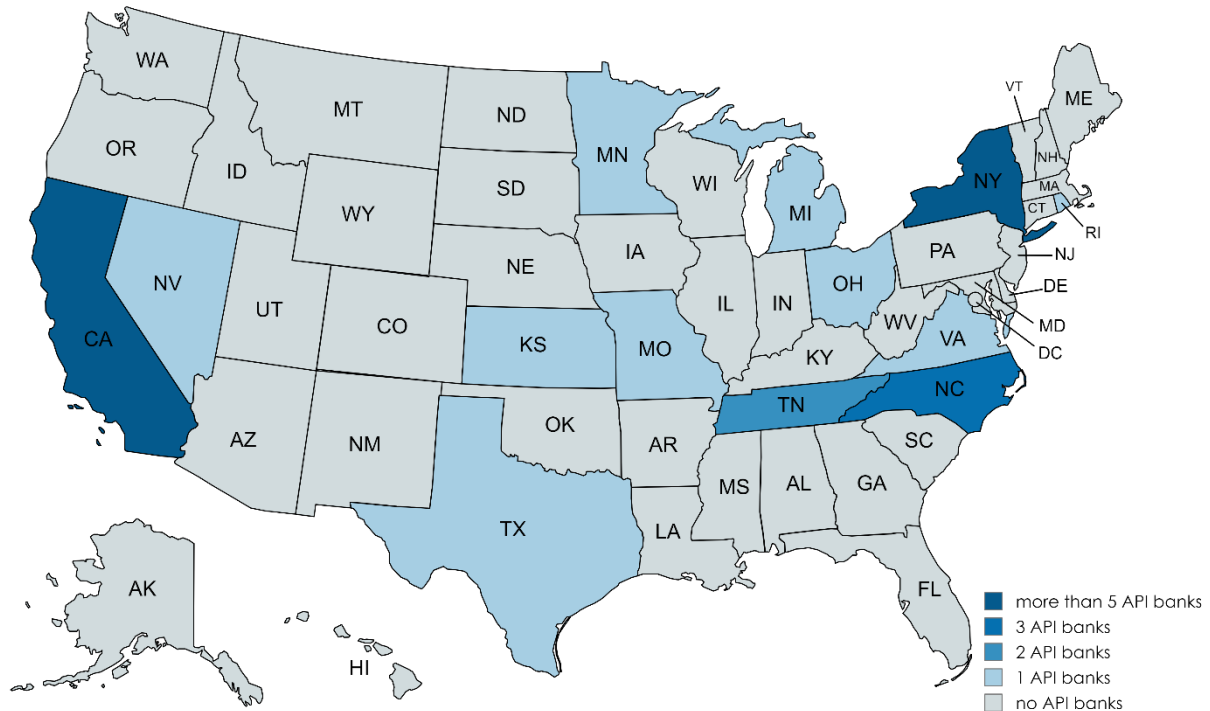


Figure 3. Changes in Tobin's Q and ROA around API adoption

This figure presents the changes in *Tobin's Q* and *ROA* around API adoption during our sample period from 2007-2022. We report the lag and lead for a binary variable that captures the change in bank performance before and after API adoption, and the graph shows the coefficient from -5 to +2 years. The outcome variable of the upper figure is the bank's overall growth metric *Tobin's Q*, while the lower figure is *ROA*. The control variables and fixed effects are the same as in column (2) of Table 4 and are not reported here for brevity. The bar chart represents a 95% confidence interval. The Y-axis represents the average changes in *Tobin's Q* and *ROA* of API banks after adopting APIs relative to non-API banks.

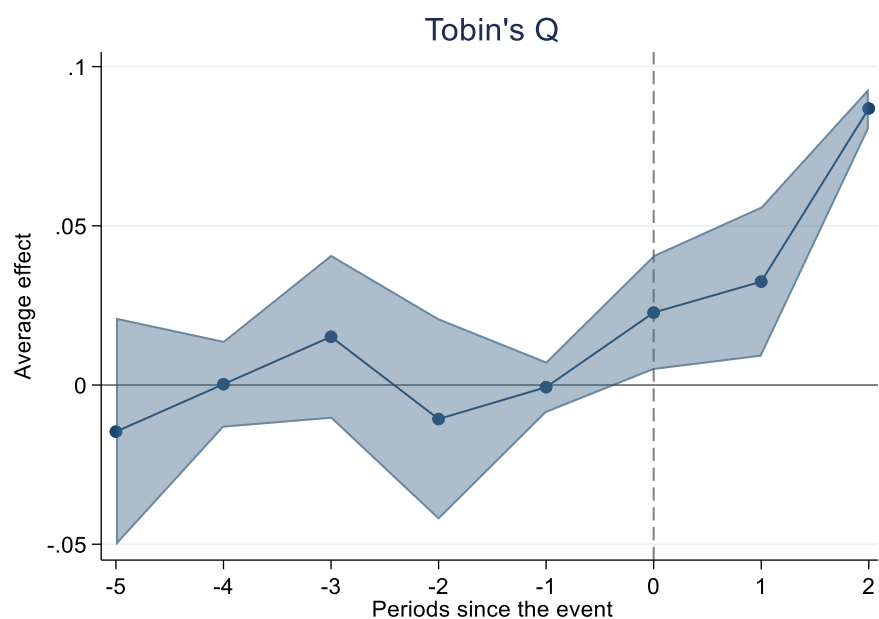


Figure 3a. Change in Tobin's Q around API adoption

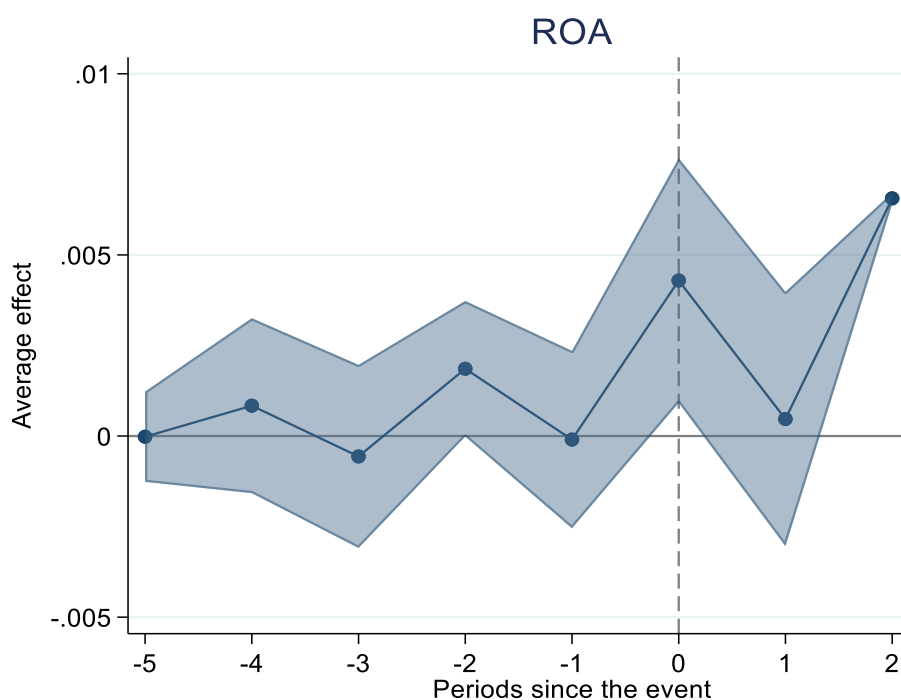


Figure 3b. Change in ROA around API adoption

Table 1. Variable description

Variable name	Definition	Source
<i>Variables for API Adoption</i>		
<i>API_Adopt_{i,t}</i>	The dummy variable equals 1 if the bank <i>i</i> has operating APIs in period <i>t</i>	Bank websites and API platforms
<i>API_Intensity_{i,t}</i>	Number of APIs owned by bank <i>i</i> in year <i>t</i>	
<i>API_bank_i</i>	Dummy variable indicating if bank <i>i</i> has operating APIs during the sample period (2007-2022)	
<i>R_Ratio</i>	The number of adopted Read-only APIs to the total number of APIs in the bank <i>i</i> in <i>t</i> years	
<i>R/W_Ratio</i>	The number of Read-Write APIs to the total number of APIs in the bank <i>i</i> in <i>t</i> years	
<i>Measures of Bank Performance</i>		
<i>Tobin's Q</i>	The ratio of the market value of bank shares to the replacement cost of the assets represented by the shares	Compustat, BankFocus
<i>MRKTV</i>	Natural logarithm of market value of equity	
<i>NIM</i>	Net interest margin (<i>NIM</i>)	
<i>NIE</i>	The ratio of other noninterest expenses to total assets (<i>NIE</i>)	
<i>NII</i>	The ratio of noninterest income to total asset (<i>NII</i>)	
<i>LLP</i>	The ratio of loan loss provision to total loans	
<i>Bank and Industry Characteristics</i>		
<i>Firm Age</i>	Year minus the IPO year or the first year Compustat reports data for the firm	Compustat, BankFocus
<i>Bank Size</i>	Logarithm of total assets	
<i>Capital Intensity</i>	The ratio of capital expenditures to total assets	
<i>Leverage</i>	Long-term debt plus debt in current liabilities, divided by total assets	
<i>Cash Holding</i>	Cash and short-term investments are divided by total assets	
<i>Cost of Capital</i>	Total interest expense divided by the total amount of debt	
<i>Tier 1</i>	Tier 1 capital ratio of a bank's equity capital with its total risk-weighted assets (RWAs)	
<i>Deposit Ratio</i>	The ratio of total deposit to total assets	
<i>Loan Ratio</i>	The ratio of total loans to total assets	
<i>Employee</i>	We use the natural logarithm of 1 plus the number of employees in the regression, the number of employees in thousands.	
<i>Efficiency</i>	Operational efficiency is the ratio of the bank's inputs to outputs, (input measured by deposits, non-interest expenses (net of loan losses), fixed assets, and loan loss provisions) and output by total loans and leases, other earning assets	Hoberg-Phillips Database
<i>RWA</i>	The logarithm of the bank's Risk-weighted asset	
<i>NPL</i>	The ratio of non-performing loans to total loans	
<i>Similarity</i>	The pairwise product similarities for each firm with competitor firms in 10-k filings. See Hoberg and Phillips (2016)	Cybersecurity score FCC's semi-annual report
<i>Fluidity</i>	The overlap between the product descriptions of firms and corresponding rivals. See Hoberg et al. (2014)	
<i>Cybersecurity</i>	Based on textual analysis and a comparison of cybersecurity risk disclosures of banks. See Florackis (2022)	
<i>Broadband</i>	State-level broadband subscription rates for Internet access within the range [0,1], with values near 0 indicating low broadband subscription and near 1 indicating high residential broadband subscription	

Table 2. Summary statistics

This table presents summary statistics for the main variables in the regressions for 5,692 firm-year observations for 713 banks from 2007 to 2022. Variable descriptions are provided in Table 1. Column (1) shows the descriptive statistics number of observations (N), mean, and standard deviation (SD) for the overall sample, columns (2) and (3) show the descriptive statistics for the 691 banks without APIs and the 22 banks with APIs, respectively.

	(1) Full-Sample (N=713)			(2) Banks with APIs (N=22)		(3) Banks without APIs (N=691)	
No. of banks	N	Mean	SD	Mean	SD	Mean	SD
<i>VARIABLES</i>							
<i>Panel A: Variables for API adoption</i>							
<i>API_Adopt</i>	5692	0.008	0.091	0.222	0.416	0.000	0.000
<i>API_Intensity</i>	5692	0.060	0.969	1.590	4.762	0.000	0.000
<i>API_bank</i>	5692	0.037	0.189	1.000	0.000	0.000	0.000
<i>R_Ratio</i>	5692	0.034	0.532	0.896	2.618	0.000	0.014
<i>R/W_Ratio</i>	5692	0.028	0.473	0.745	2.347	0.000	0.000
<i>Panel B: Measures of bank performance</i>							
<i>Tobin's Q</i>	5657	1.018	0.049	1.037	0.043	1.017	0.049
<i>MRKTV</i>	5657	2.195	9.147	31.18	33.43	1.072	3.163
<i>ROA</i>	5692	0.005	0.012	0.008	0.005	0.005	0.012
<i>NIM</i>	5636	3.508	0.687	3.249	0.828	3.518	0.679
<i>NIE</i>	5641	0.032	0.013	0.030	0.010	0.032	0.014
<i>NII</i>	5641	0.010	0.009	0.017	0.009	0.010	0.009
<i>LLP</i>	5211	0.587	1.020	0.638	0.849	0.585	1.026
<i>Panel C: Bank and industry characteristics</i>							
<i>Firm Age</i>	5692	2.704	0.696	3.448	0.687	2.675	0.68
<i>Bank Size</i>	5692	7.947	1.612	11.43	2.235	7.812	1.422
<i>Capital Intensity</i>	5692	0.107	0.038	0.103	0.019	0.107	0.038
<i>Leverage</i>	5692	1.207	1.234	1.488	1.139	1.196	1.236
<i>Cash Holding</i>	5692	0.056	0.053	0.096	0.083	0.054	0.051
<i>Cost of Capital</i>	5692	3.269	40.80	0.477	1.359	3.377	41.58
<i>Tier 1</i>	5692	12.480	4.119	11.640	2.120	12.520	4.174
<i>RWA</i>	5692	4.753	6.273	8.404	8.306	4.612	6.138
<i>Employee</i>	5692	6.275	1.485	9.359	2.339	6.156	1.303
<i>Efficiency</i>	5692	0.122	0.010	0.024	0.016	0.012	0.009
<i>Broadband</i>	5692	0.722	0.125	0.759	0.100	0.721	0.126
<i>Deposit Ratio</i>	5692	0.765	0.096	0.711	0.106	0.767	0.095
<i>Loan Ratio</i>	4164	0.043	0.267	0.060	0.048	0.043	0.271
<i>NPL</i>	5328	1.539	1.060	1.468	0.781	1.542	1.070
<i>Similarity</i>	5692	51.21	16.68	41.77	15.38	51.58	16.62
<i>Fluidity</i>	5692	10.63	3.521	9.404	3.453	10.68	3.516
<i>Cybersecurity</i>	5215	0.094	0.032	0.112	0.021	0.093	0.032

Table 3. Logistic regression of API adoption on bank characteristics

This table shows the results of a logistic regression of banks' API adoption (*API_bank*) on bank characteristics:

$$API_bank_i = \alpha + \beta Bank\ Characteristics_{i,t} + \omega_t + \epsilon_{i,t} \quad (1)$$

The table reports regression coefficients and t-statistics (in parentheses). The regression model is shown in Equation (1), with the dependent variable *API_bank_i* (a dummy variable indicating if bank *i* adopts APIs during the sample period) regressed on variables of bank characteristics before the adoption of the first API. The independent variable definitions are provided in Table 1. All models include year-fixed effects. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

<i>VARIABLES</i>	<i>API bank</i>	
	(1)	(2)
<i>NPL</i>	-0.573*** (-5.10)	
<i>Similarity</i>	-0.023*** (-2.93)	
<i>LLP</i>		-0.509*** (-4.98)
<i>Fluidity</i>		-0.034 (-0.93)
<i>Firm Age</i>	-0.665*** (-2.94)	-0.608*** (-2.81)
<i>Bank Size</i>	1.232*** (10.12)	1.271*** (10.52)
<i>Cash Holding</i>	-1.052 (-0.63)	-0.718 (-0.45)
<i>Efficiency</i>	24.800*** (2.72)	20.284** (2.08)
<i>Capital Intensity</i>	-18.775*** (-3.30)	-17.212*** (-3.21)
<i>Leverage</i>	-0.002 (-0.02)	0.075 (1.08)
<i>Cybersecurity</i>	-3.078 (-0.50)	-1.368 (-0.24)
<i>Tier 1</i>	-0.101** (-2.08)	-0.074* (-1.65)
<i>RWA</i>	-0.085*** (-3.13)	-0.094*** (-3.40)
<i>Constant</i>	-8.195*** (-8.87)	-10.669*** (-12.41)
Year FE	Yes	Yes
Observations	3,911	3,911

Table 4. Regression of bank performance on API adoption

The table reports baseline regression coefficients and t-statistics (in parentheses) using the Heckman two-stage analysis.

$$Pr(API_{i,t}) = \alpha + \beta Broadband_{i,t} + \gamma Controls_{i,t} + \varphi_i + \omega_t + \epsilon_{i,t} \quad (2)$$

$$Perf_{i,t} = \alpha + \beta API_Adopt_{i,t} + \gamma Controls_{i,t} + \delta IMR_{i,t} + \varphi_i + \omega_t + \epsilon_{i,t} \quad (3)$$

This table shows the effect of API adoption on the bank's performance. Column (0) shows the Heckman first-stage Probit regression of API adoption by the bank i (API_i) on control variables and variable $Broadband_i$ that measures the strength of Internet access in the state where the bank i is headquartered. IMR denotes the inverse Mills ratio generated from the first step regression. In the second-stage OLS regression, we regress *Tobin's Q* and *ROA* on variable $API_Adopt_{i,t}$ (a dummy variable representing whether bank i adopted APIs in or before year t) and control variables, as defined in Table 1. Firm FE and Year FE indicate whether the model includes firm- or year-fixed effects. Standard errors clustered at the firm level. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

VARIABLES	API_i	<i>Tobin's Q</i>		<i>ROA</i>	
	(0)	(1)	(2)	(3)	(4)
<i>Broadband</i>	0.545 (0.63)				
API_Adopt_i		0.018* (1.93)	0.015* (1.65)	0.004*** (2.84)	0.002* (1.66)
<i>IMR</i>		-0.010*** (-3.59)	-0.010*** (-3.93)	-0.001 (-1.23)	-0.002* (-1.90)
<i>Firm Age</i>	-0.177 (-1.50)	0.027*** (6.78)	0.007 (1.32)	0.012*** (6.99)	0.010*** (6.27)
<i>Bank Size</i>	0.735** (2.52)	-0.016*** (-4.13)	-0.028*** (-5.60)	-0.003*** (-2.78)	-0.005*** (-3.35)
<i>Cash Holding</i>	1.749 (1.29)	-0.143*** (-7.37)	-0.020 (-1.01)	-0.026*** (-3.99)	-0.028*** (-3.95)
<i>Cost of Capital</i>	-0.008 (-1.12)	0.000** (2.16)	0.000*** (3.58)	0.000 (1.02)	0.000* (1.88)
<i>Employee</i>	-0.284 (-0.93)	0.014*** (2.95)	0.019*** (3.96)	-0.000 (-0.03)	0.001 (0.45)
<i>Leverage</i>	-0.124 (-1.47)	-0.001 (-1.03)	0.001 (0.98)	0.140*** (2.64)	0.140*** (2.64)
<i>Capital Intensity</i>	-10.023*** (-2.64)			0.307*** (6.37)	0.289*** (6.26)
<i>Deposit Ratio</i>				-0.015*** (-3.41)	-0.020*** (-3.87)
<i>Loan Ratio</i>				0.004 (1.33)	0.004 (1.33)
<i>Constant</i>	-4.923*** (-5.58)	1.021*** (43.90)	1.131*** (35.13)	-0.144*** (-2.73)	-0.120** (-2.31)
Firm FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	No	Yes	No	Yes
Observations	5,613	5,613	5,613	4,098	4,098
R-squared	0.3635	0.584	0.718	0.551	0.618

Table 5. Dynamic regression of bank performance on lagged API adoption

This table reports the changes in *Tobin's Q* and *ROA* around API adoption during our sample period from 2007-2022 using the estimation procedure introduced by Callaway and Sant'Anna (2021) (variables are defined as in Table 4):

$$Perf_{i,t} = \alpha + \sum_{k=-5}^2 \delta_k \times D_{i,t}^k + \gamma Control_{i,t} + \delta IMR_{i,t} + \varphi_i + \omega_t + \epsilon_{i,t} \quad (4)$$

We report the lag and lead for a binary variable that captures the change in bank performance before and after API adoption, and the table shows the coefficient from -5 to +2 years. Specifically, the coefficient for Pre^5 reflects the average effect of API adoption five years before implementation, $Post$ represents the average effect in the year of API adoption, and $Post^2$ indicates the average effect two years following API adoption. For the dependent variables, we apply *Tobin's Q* in column (1) and *ROA* in column (2). We report t-statistics in parentheses. We include control variables as defined in Table 1. All models include year-fixed and firm-fixed effects, and the standard errors are clustered by firm level. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

<i>VARIABLES</i>	<i>Tobin's Q</i>	<i>ROA</i>
	(1)	(2)
Pre_avg	-0.002 (-0.47)	0.000 (1.17)
$Post_avg$	0.047*** (8.15)	0.003*** (4.40)
Pre^5	-0.015 (-0.80)	-0.000 (-0.03)
Pre^4	0.000 (0.04)	0.001 (0.68)
Pre^3	0.015 (1.15)	-0.001 (-0.44)
Pre^2	-0.011 (-0.66)	0.002* (1.95)
Pre^1	-0.001 (-0.16)	-0.000 (-0.08)
$Post$	0.023** (2.48)	0.004** (2.49)
$Post^1$	0.032*** (2.70)	0.000 (0.26)
$Post^2$	0.087*** (26.63)	0.007*** (80.97)
<i>Control Variables</i>	Yes	Yes
Firm FE	Yes	Yes
Year FE	Yes	Yes
Observations	5,534	4,305

Table 6. Regression of bank earnings, costs, and market value on lagged API adoption

This table shows the changes in bank earnings and costs up to three years after API adoption based on the OLS regression below (variables are defined as in Table 4):

$$Perf_{i,t} = \alpha + \beta API_Adopt_{i,t-j} + \gamma Controls_{i,t} + \delta IMR_{i,t} + \varphi_i + \omega_t + \epsilon_{i,t} \quad (5)$$

We apply contemporaneous and lagged variables of API adoption: API_Adopt_t , API_Adopt_{t-1} , and API_Adopt_{t-2} . The dependent variables are net interest margin (NIM), other noninterest expenses to total asset (NIE), noninterest income to total asset (NII), loan loss provisions to total asset (LLP), and market value ($MRKTV$). We report regression coefficients and t-statistics in parentheses. The control variables are defined in Table 1. All models include year and firm fixed effects, and standard errors are clustered by firm level. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Panel A. Regression of bank earnings and costs on lagged API adoption												
VARIABLES	<i>NIM</i>		<i>NIE</i>			<i>NII</i>			<i>LLP</i>			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<i>API_Adopt_t</i>	-0.040 (-0.18)			-0.002 (-1.09)			-0.001 (-0.79)			-0.363** (-2.36)		
<i>API_Adopt_{t-1}</i>		-0.214 (-1.01)			-0.003** (-2.15)			-0.002** (-2.02)			-0.317 (-1.64)	
<i>API_Adopt_{t-2}</i>			-0.275 (-1.49)			-0.004*** (-3.39)			-0.004*** (-3.06)			-0.342* (-1.76)
<i>Controls</i>	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm cluster	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,071	3,799	3,528	4,073	3,799	3,528	4,073	3,799	3,528	4,038	3,767	3,494
R-squared	0.789	0.805	0.818	0.766	0.778	0.778	0.727	0.73	0.76	0.615	0.633	0.616

(continued next page)

Table 6. Regression of bank earnings, costs, and market value on lagged API adoption (cont'd)

Panel B. Regression of bank value on lagged API adoption			
<i>VARIABLES</i>	<i>MRKTV</i>		
	(1)	(2)	(3)
<i>API_Adopt_t</i>	3.441* (1.83)		
<i>API_Adopt_{t-1}</i>		1.962 (1.11)	
<i>API_Adopt_{t-2}</i>			-0.361 (-0.54)
<i>Controls</i>	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Firm cluster	Yes	Yes	Yes
Observations	5,613	5,100	4,602
R-squared	0.960	0.960	0.962

Table 7. The role of President Biden's executive order announcement on API adopting banks

This table reports Differences in difference regression results of bank performance on API adoption before and after President Biden's executive order. We estimate the following equation (6), which extends Equation (3) by incorporating an interaction term between API adoption and Biden_Post.

$$Perf_{i,t} = \alpha + \beta API_{i,t} + \partial API_{i,t} \times Biden_Post_t + \gamma Controls_{i,t} + \varphi_i + \omega_t + \epsilon_{i,t} \quad (6)$$

The table below presents the results of the extended regression model from equation (6), including the interaction effect of Biden Policy and API_Adoption. We include *API_Adopt* (a dummy variable indicating whether bank *i* adopted API by year *t*). Additionally, we construct a dummy variable, *Biden_Post*, to capture the impact of President Biden's executive order issued in July 2021, where the period before the executive order is equalled as 0, and 2021 and subsequent years are equalled as 1. The dependent variables are consistent with those in Tables 4 and 6, comprising *Tobin's Q*, *ROA*, and the four components of ROA: Net Interest Margin (*NIM*), Non-Interest Expense (*NIE*), Net Interest Income (*NII*), and Loan Loss Provisions (*LLP*). The regression results based on matched samples are shown here and OLS regression coefficients and t-statistics are in parentheses. Variables are defined in Table 1. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	<i>Tobin's Q</i>	<i>ROA</i>	<i>NIM</i>	<i>NIE</i>	<i>NII</i>	<i>LLP</i>
<i>VARIABLES</i>	(1)	(2)	(3)	(4)	(5)	(6)
<i>API_Adopt</i>	0.011 (1.13)	0.002* (1.88)	-0.016 (-0.07)	-0.002 (-0.96)	-0.001 (-1.10)	-0.226 (-1.35)
<i>API_Adopt * Biden_Post</i>	0.012* (1.76)	-0.001 (-0.39)	-0.065 (-0.34)	-0.001 (-0.26)	0.001 (0.80)	-0.366*** (-2.79)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm cluster	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,613	4,098	4,071	4,073	4,073	4,038
R-squared	0.718	0.618	0.789	0.766	0.727	0.615

Table 8. Propensity scores matching results

This table reports OLS regression results of bank performance on API adoption based on the matched sample using propensity scores matching (PSM). The model is shown in Equation (3), adopted for the PSM sample. The covariate balance checks are presented in Panel A. We balance the treated and control banks using firm-level controls, namely, *Bank Size*, *Firm age*, *Capital Intensity*, *Cash Holding*, and *Cost of Capital*. The regression results based on matched samples are shown here and OLS regression coefficients and t-statistics are in parentheses. We regress *Tobin's Q* in columns (1) - (3) and *ROA* in columns (4) - (6) on three-year lagged terms of API adoption dummy API_Adopt_t , API_Adopt_{t-1} and API_Adopt_{t-2} (which represent whether bank i adopted APIs in or before year t , $t-1$, and $t-2$). Variables are defined in Table 1. ***, ** and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Results of covariate balance checks						
	Treatment group		Control group			
	Banks with APIs		Banks without APIs		Difference	t-stat
<i>VARIABLES</i>	Mean		Mean			
<i>Bank Size</i>	10.068		10.251		-0.183	-0.26
<i>Firm Age</i>	2.5656		2.472		0.094	0.25
<i>Capital Intensity</i>	0.103		0.1002		0.003	0.12
<i>Cash Holding</i>	0.1022		0.0817		0.021	0.68
<i>Cost of Capital</i>	1.4667		0.7966		0.67	0.58
Panel B. Regression results using PSM procedure						
	<i>Tobin's Q</i>				<i>ROA</i>	
<i>VARIABLES</i>	(1)	(2)	(3)	(4)	(5)	(6)
<i>API_Adopt_t</i>	0.015** (2.02)			0.003** (2.21)		
<i>API_Adopt_{t-1}</i>		0.016* (1.85)			0.003 (1.61)	
<i>API_Adopt_{t-2}</i>			0.013 (1.00)			0.005* (1.85)
<i>Control Variables</i>	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm cluster	Yes	Yes	Yes	Yes	Yes	Yes
Observations	786	728	675	586	546	513
R-squared	0.774	0.780	0.798	0.812	0.815	0.779

Table 9. API category and bank performance

The table reports OLS regression coefficients of bank performance on adoption and relative adoption variables for API types R and R/W (see Section 3) as follows:

$$Perf_{i,t} = \alpha + \beta API\ category_{i,t} + \gamma Controls_{i,t} + \delta IMR_{i,t} + \varphi_i + \omega_t + \epsilon_{i,t} \quad (7)$$

The regression model is shown in Equation (6). For the dependent variables, we apply *Tobin's Q* in columns (1) and (2) and *ROA* in columns (3) and (4). Panel A and Panel B illustrate the influence of the adoption of each specific API type on bank performance and the ratio of different API types relative to the total number of APIs, respectively. Panel A shows the results for the regression of bank performance on dummy variables *R_Adopt* and *R/W_Adopt* to represent whether bank *i* has adopted Read-only or Read/Write APIs in or before year *t*. Panel B shows the regression results with continuous variables *R/W_Ratio* and *R_Ratio* which measure the number of R/W- and R-APIs relative to the total number of APIs of bank *i* in year *t*, ranging from 0 to 1. All models control for a year and firm fixed effects. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

Panel A. Results of API category Adoption				
VARIABLES	<i>Tobin's Q</i>		<i>ROA</i>	
	(1)	(2)	(3)	(4)
<i>R_Adopt_t</i>	0.021 (1.43)		0.000 (0.21)	
<i>R/W_Adopt_t</i>		0.016* (1.68)		0.002* (1.84)
<i>Control Variables</i>	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm cluster	Yes	Yes	Yes	Yes
Observations	5,613	5,613	4,098	4,098
R-squared	0.718	0.718	0.618	0.618
Panel B. Results of API category ratio				
VARIABLES	<i>Tobin's Q</i>		<i>ROA</i>	
	(1)	(2)	(3)	(4)
<i>R_Ratio</i>	0.021 (1.01)		-0.000 (-0.08)	
<i>R/W_Ratio</i>		0.018* (1.92)		0.004*** (3.45)
<i>Control Variables</i>	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm cluster	Yes	Yes	Yes	Yes
Observations	5,613	5,613	4,098	4,098
R-squared	0.718	0.718	0.618	0.618

Table 10. Cross-sectional differences after API adoption

$$Perf_{i,t} = \alpha + \beta API_{i,t} + \partial API_{i,t} \times Bank_Character_i + \gamma Controls_{i,t} + \varphi_i + \omega_t + \epsilon_{i,t} \quad (8)$$

This table shows results of the extended regression model for bank performance after API adoption in Equation (3), including interaction effects with the level of non-performing loans (NPL) and loan loss provision (LLP). We include interaction effects of *API_Adopt* (a dummy variable representing whether bank *i* adopted APIs in and before year *t*) and dummies that split the sample of API banks based on cross-sectional differences, in particular (1) non-performing loans (*High_NPL* equals one for banks with *NPL* above the average of all sample banks) and loan loss provisions *High_LL*P (which equals one for banks with *LLP* above the average of all sample banks), (2) product market fluidity (*High_Fluidity* equals one for banks with *Fluidity* above the average of all sample banks) and product market similarity (*High_Similarity* equals one for banks with *Similarity* above the average of all sample banks), and (3) total assets (*High_Asset* equals one for banks with *Total Assets* above the average of all sample banks) and deposits (*High_Deposit* equals one for banks with *Deposits* above the average of all sample banks). The dependent variables are *Tobin's Q* in columns (1) and (2) and *ROA* in columns (3) and (4). The control variables are defined in Table 1. We report t-statistics in parentheses. All models include year- and firm-fixed effects. The standard errors are clustered by firm level. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

<i>Panel A: The role of Non-Performing Loans and Loan Loss Provisions</i>				
<i>VARIABLES</i>	<i>Tobin's Q</i>		<i>ROA</i>	
	(1)	(2)	(3)	(4)
<i>API_Adopt</i>	0.024** (2.08)	0.015 (1.43)	0.005*** (4.94)	0.005*** (4.19)
<i>API_Adopt * High_NPL</i>	-0.015 (-0.85)		-0.005*** (-3.07)	
<i>API_Adopt * High_LL</i> P		0.001 (0.06)		-0.004** (-2.17)
<i>Control Variables</i>	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	5,613	5,613	4,098	4,098
R-squared	0.718	0.718	0.618	0.618

<i>Panel B: The role of Product Similarity and Fluidity</i>				
<i>VARIABLES</i>	<i>Tobin's Q</i>		<i>ROA</i>	
	(1)	(2)	(3)	(4)
<i>API_Adopt</i>	0.028*** (3.36)	0.027*** (2.96)	0.002 (1.61)	0.002*** (2.62)
<i>API_Adopt * High_Fluidity</i>	-0.035** (-2.44)		0.000 (0.15)	
<i>API_Adopt * High_Similarity</i>		-0.027* (-1.83)		0.000 (0.24)
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	5,613	5,613	4,098	4,098
R-squared	0.719	0.718	0.618	0.618

Table 10. Cross-sectional differences after API adoption (cont'd)

<i>Panel C: The role of bank size and deposits</i>				
<i>VARIABLES</i>	<i>Tobin's Q</i>		<i>ROA</i>	
	(1)	(2)	(3)	(4)
<i>API_Adopt</i>	0.016*** (2.87)	0.026** (2.25)	0.006*** (3.93)	0.006*** (4.97)
<i>API_Adopt * High_Asset</i>	-0.001 (-0.07)		-0.004** (-1.97)	
<i>API_Adopt * High_Deposit</i>		-0.012 (-0.78)		-0.005*** (-2.63)
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	5,613	5,613	4,098	4,098
R-squared	0.718	0.718	0.618	0.618

“The effects of voluntary open banking adoption”

In this online appendix, we provide additional discussions and robustness tests that supplement the empirical analysis in the main body of the manuscript.

Figure A1. Deviations Before and After PSM

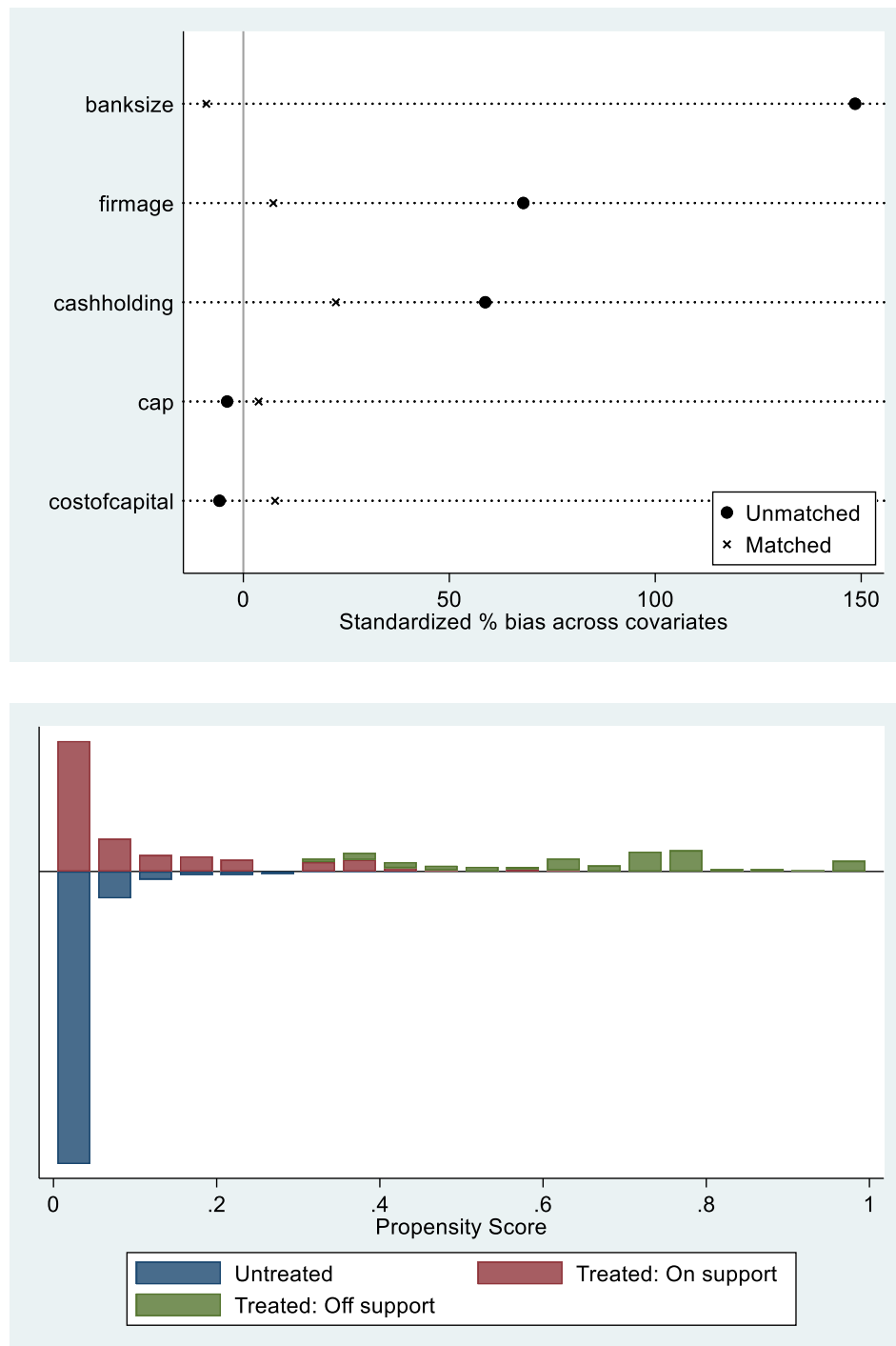


Table A1. Correlation matrices of the key variables

We calculated the correlation coefficient for each pair of key variables and demonstrated its significance. We have included independent and dependent variables and interaction variables in our analysis. (1)-(2) respectively represent the bank's ownership of the API; (3)-(9) represent the bank performance. Among the key variables of our cross-sectional analysis are (10)-(12). All variables are defined in Table 1. * indicates statistical significance at the 1% level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) <i>API_bank</i>	1											
(2) <i>API_Adopt</i>	0.4837*	1										
(3) <i>Tobin's Q</i>	0.0715*	0.0446*	1									
(4) <i>MRKTV</i>	0.5059*	0.4147*	0.1424*	1								
(5) <i>ROA</i>	0.0014	0.0032	0.4311*	0.0915*	1							
(6) <i>NIM</i>	-0.1088*	-0.0942*	0.1008*	-0.1754*	0.2317*	1						
(7) <i>NIE</i>	-0.0466*	-0.0410*	-0.1605*	-0.0375*	-0.3541*	0.2383*	1					
(8) <i>NII</i>	0.0375*	0.0103	0.0580*	0.0807*	0.1537*	0.0903*	0.7779*	1				
(9) <i>LLP</i>	-0.0009	-0.0184	-0.2185*	0.0290*	-0.4385*	0.0462*	0.4102*	0.0561*	1			
(10) <i>NPL</i>	-0.0056	-0.0141	-0.1439*	0.016	-0.2687*	0.0127	0.1835*	-0.0161	0.3936*	1		
(11) <i>Similarity</i>	-0.0783*	-0.0554*	-0.2310*	-0.1775*	-0.3174*	-0.0399*	0.0066	-0.1685*	0.0786*	0.1833*	1	
(12) <i>Fluidity</i>	-0.0474*	-0.0371*	-0.2643*	-0.1291*	-0.2356*	-0.1083*	0.0704*	-0.0553*	0.1781*	0.1488*	0.5191*	1

Table A2. Propensity Scores Matching

The Propensity score matching indicates that all univariate differences in bank characteristics are statistically insignificant, suggesting that any differences in bank performance between the treatment and control groups should be driven by the presence or absence of an API rather than observable bank characteristics. The means of the five measures of bank performance for the treated banks: *Tobin's Q* and *ROA* are significantly higher than those of the control banks, which are consistent with our main results. In addition, we also show by visualization in Appendix Figure 1 that there are no significant differences between the means of the variables before and after matching (minor deviations from the mean, small t-values are not significant), i.e., the matched data pass the balance test.

	Treatment group (Banks with APIs)	Control group (Banks without APIs)		
<i>VARIABLES</i>	Mean	Mean	Difference	t-stat
<i>Tobin's Q</i>	1.0455	1.0265	0.0189***	4.30
<i>ROA</i>	1.5572	1.1846	0.3725***	3.17

Table A3. Bank performance and API number

When using the number of APIs as an independent variable, we found that the continued adoption of APIs has no significant impact on Tobin's Q and ROA. We experimented with various methods, including different combinations of fixed effects and clustering, but the results remained consistent. This suggests that the initial adoption of APIs has a certain effect on banks because it establishes a partnership or data-sharing model. Subsequent adoptions of additional APIs do not seem to produce a significant impact. In Table Appendix 3, we present results that are methodologically consistent with the baseline model. This includes controlling for firm fixed effects, year fixed effects, and clustering the standard errors at the firm level.

<i>VARIABLES</i>	(1)	(2)
	<i>Tobin's Q</i>	<i>ROA</i>
<i>API_Intensity</i>	0.001 (0.76)	-0.000 (-0.61)
<i>Constant</i>	1.129*** (34.69)	-0.206*** (-5.93)
<i>Controls</i>	Y	Y
Firm FE	Y	Y
Year FE	Y	Y
Firm cluster	Y	Y
Observations	5,595	4,073
R-squared	0.712	0.629