Data-Driven Technologies and the Local Information Advantages in Small Business Lending*

Wilbur Chen[1,3], Jung Koo Kang[2,3], and Aditya Mohan[4]

[1] *Hong Kong University of Science and Technology*
[2] *Harvard Business School*
[3] *Digital, Data and Design Institute at Harvard*
[4] *Analysis Group*

July 14, 2023

PRELIMINARY DRAFT, PLEASE DO NOT DISTRIBUTE WITHOUT PERMISSION

*Email: acwilbur@ust.hk, jkang@hbs.edu and aditya.mohan@analysisgroup.com.

We thank helpful suggestions and comments from Shane Greenstein, Jonas Heese, Yimeng Li, Siko Sikochi, Charlie Wang and from seminar participants at Harvard Business School and participants at the CAAA conference. We are grateful for the financial support from our respective institutions. Wilbur Chen and Jung Koo Kang also acknowledges the Digital Value Lab in the Digital, Data, and Design (D^3) Institute at Harvard for research support.

Data-Driven Technologies and the Local Information Advantages in Small Business Lending

Abstract

We investigate whether the adoption of data-driven technologies diminishes information advantages held by local banks in small business lending markets. We start by showing evidence of local banks' information advantages using local newspaper closures as a plausibly exogenous reduction in local information for non-local banks. Consistent with decreases in valuable local information reinforcing information advantages for local banks, we find that banks with higher local concentration of deposits and small business loans, obtain a larger share of the small business loans in the local county following local newspaper closures. Next, we find that these information advantages of local banks diminish after cloud platforms, a key data-driven technology infrastructure, are widely implemented. Employing various proxies to measure data-driven technology adoption, we further document consistent evidence that local banks do not obtain larger share for small business loans in counties where banks exhibit greater AI job demand, engaged in an acquisition of an AI technology firm, adopt web analytics technologies, have higher data storage budget or disclose AI-related terms. Overall, our findings suggest that data-driven technologies have the potential to reduce the private information advantages of local banks, thereby leveling the playing field in lending markets.

Keywords: Data-Driven Technologies, Local Information Advantages, Private Lenders, Local Banks, Small Businesss Loans
JEL Codes: D80, G21, M40, O16, O30

I Introduction

Small business lending is crucial for creating job opportunities, fostering innovation and driving economic growth (Biggs 2002; Beck and Demirguc-Kunt 2006). However, the inherent opacity of small businesses presents a significant challenge, potentially restricting their access to credit. To mitigate this information problem, banks have traditionally relied on relationship lending using repeated interactions with borrowers to accumulate valuable information over time (Petersen and Rajan 1994; Berger and Udell 2002; Berger and Black 2011). Prior studies suggest that local banks, in particular, invest more in the information collection through the lending relationship due to their proximity in the local market and decentralized operating structure resulting in significant information advantages over non-local banks (Stein 2002; Loutskina and Strahan 2011). Yet, the past decade has seen an important shift in how banks and other lenders produce borrower information. Increasingly, these institutions are harnessing a variety of alternative data sources, such as mobile phone usage, internet footprints, online shopping habits, and social media activities. Coupled with the use of machine learning, artificial intelligence, and customer analytics, these data-driven technologies are transforming tasks such as credit scoring, customer profiling and fraud management (Lin et al. 2012; Srivastava and Gopalkrishnan 2015; Jagtiani and Lemieux 2019). Despite these technological advancements, there is limited research on the impact of the data-driven technology adoption on information production in the private lending market.

In this study, we investigate whether data-driven technologies diminish information advantages traditionally held by local banks in the small business lending market. We argue that non-local banks derive more benefit from these technologies in terms of their information production about borrowers, which can mitigate local banks' existing information advantages. Non-local banks serve wider geographical areas and diverse customer bases, granting them access to a larger pool of data for credit analysis. Moreover, non-local banks are larger and possess greater financial resources, which allows them to invest in sophisticated data infrastructure and specialized analytics teams[1]. Thus, in markets where data-driven lending technologies are more prevalent, non-local banks are better positioned to leverage these technologies, which may improve their competitive advantages against local banks. On the other hand, the insights gleaned from data-driven technologies may not endow non-local banks with sufficient information advantages relative to the borrower information obtained through lending relationships by local banks. In addition, the successful implementation of such technologies is a complex process that requires time and resources (Bresnahan and Greenstein 1996; Bresnahan et al. 2012), therefore its impact on information advantages is an open empirical question.

One of the primary empirical challenges in studying this question lies in identifying information advantages of local banks because banks' information acquisition from lending relationships is not directly observable. To address this empirical challenge, we employ local newspaper closures as a plausibly exogenous reduction in valuable local information for non-local banks. Prior studies document that media plays a crucial role in capital markets by uncovering and disseminating information about economic and firm fundamentals (Miller 2006; Tetlock et al. 2008; Bushee et al. 2010). Media also serves as an important information intermediary in mitigating information asymmetry in the private loan market (Bushman et al. 2017). In particular, local newspapers, due to their proximity to local firms and social networks, can effectively identify and engage with local firms' stakeholders such as employees and suppliers and provide useful local information about local economies and firm performance (Engelberg et al. 2011; Peress 2014; Kim et al. 2021; Gao et al. 2020). Therefore, local newspaper

1. Consistent with this view, survey evidence finds that large banks spend more (even as a percentage of expenses) on technology than their smaller peers. Among banks with less than $100 Billion in assets, 70% of respondents noted that budgetary constraints were a key impediment to the successful implementation of a technology strategy. Moreover, larger banks focus technology spending on innovation and revenue enhancement while smaller banks focus on IT enhancement and operational risk reduction (Martinez et al. 2019)

closures may strengthen existing information advantages of local banks because, unlike non-local banks, local banks can substitute the reduction of valuable local information following the local newspaper closures with their own information production through their lending relationships with borrowers.

We begin our analyses by using the local newspaper closures as an information shock to identify local banks' information advantages. We employ a staggered difference in differences design by exploiting variations in the timing of local newspaper closures across counties. For each treatment county that experienced local newspaper closures, we match it with the adjacent control country that never experienced local newspaper closures.[2] To address biases arising from treatment effect heterogeneity (Baker et al. 2022), we stack the regression by the matched county pairs. Moreover, by comparing the changes in bank lending within a treatment country to the corresponding changes in bank lending within adjacent control counties, we control for economic differences across regions. To control for time invariant differences between banks in each county, we include bank-county fixed effects. We further control for time trends for each matched county pair by using year fixed effects.

Next, we measure the extent of a bank's *local* operational presence using the following four proxies—the proportion of deposit in a local county, the local deposit concentration (herfindahl index of the local deposit proportion), the proportion of small business loans in a local county, and the local small business loan concentration (herfindahl index of the local small business loan proportion). To examine the changes in information advantages between local and non-local banks, we focus on the market competition in the information-intensive small business loan market where information advantages play a crucial role in the lending decision (Berger and Udell 2002; Berger and Black 2011; Loutskina and Strahan 2011).

We begin our analysis by showing that local banks obtain greater market share of

2. We use the NBER county adjacency files to identify geographically adjacent county groups.

small business loans in the local country after the local newspaper closures, consistent with reduction in valuable local information strengthening information advantages for local banks. Economically, the standard deviation of the local deposit proportion in a county (0.211) and the bank-level local deposit concentration (0.292), corresponds to a 0.4% and 0.7% increase in the shares of local small business loans issued (or a 4% and 8% increase relative to the sample average of 8.7%). Moreover, we find similar results using the proportion of small business loans in a local county to measure the extent of a bank's local operation. Specifically, the standard deviation of the local small business loan proportion in a county (0.197) and the local small business loan concentration (0.284), corresponds to a 0.43% and 0.9% increase in the shares of local small business loans issued (or a 5% and 10% relative increase). These results are consistent with prior studies that local banks specialize in informationally opaque small-business lending (Berger and Udell 2002; Carter et al. 2004; Carter and McNulty 2005), where they rely more on soft information about borrowers collected through repeated interactions between loan officers and borrowers.[3] Whereas, non-local banks tend to be larger and more centralized organizations, and typically base their lending decisions on hard and verifiable information (Stein 2002).

We next investigate whether the adoption of data-driven lending technologies in a local market influences the existing information advantages held by the local banks. First, we provide descriptive evidence suggesting that the advent of data-driven technologies has reduced the market share gains of more concentrated and local banks following local newspaper closures. Building on prior studies, we identify 2007-2010 as the period when data-driven technologies start to be widely adopted. This shift is largely facilitated by the advent of cloud computing services, which offered a necessary infrastructure for the development and expansion of these technologies (Ewens et

3. Berger and Udell (2002) define soft information as information that "cannot be credibly communicated from one agent to another".

al. 2018).[4] Moreover, prior work that documents local information advantages in lending markets ended their sample in 2007 (Loutskina and Strahan 2011). We examine the market competition between local and non-local banks following newspaper closure in the period before 2007 (pre-cloud period) and after 2010 (post-cloud period).[5] We find that the market share gains of more local banks following local newspaper closure are concentrated in the pre-cloud period.[6] In the post-cloud period, local banks and non-local banks do not exhibit statistically significant differences in small business loan market share after newspaper closures. These results suggest that information advantages of local banks diminishes after the data-driven technology infrastructure are widely implemented.

To further study whether the adoption of data-driven technologies diminishes information advantages of local banks, we perform additional cross-sectional analyses based on proxies for the use of data-driven technology. Given the inherent challenges in directly measuring a bank's use of data-driven technologies, we employ various proxies to capture the extent of the data-driven technology adoption. To measure AI related technology adoption, we assess the extent of a bank's demand for workers with AI skills and engagement in AI technology M&As, following the approach of Acemoglu et al. (2022) and Alekseeva et al. (2021).[7] Next, we measure a bank's deployment of web-based analytics technologies and extent of its data storage budgets to gauge its use of analytics and big data related technologies.[8] We further capture the overall

4. Since 2007, the use of data-driven technologies has scaled dramatically. Specifically, the launch of Amazon Web Services in mid-2006 and Microsoft Azure in 2010 marked significant turning points.

5. We drop the period between 2008-2009 to limit the effects of the global financial crisis.

6. We also investigate trends of changes in small business loan market shares between local and non-local banks prior to local newspaper closures in the pre-cloud period. We do not find statistically significant evidence of pre-trends.

7. We measure the demand for AI worker skills with job postings that contain skills in the AI skill lists of Alekseeva et al. (2021) and the machine-learning related clusters in Goldfarb et al. (2023). We classify a bank as exhibiting high AI demand if its number of AI job postings scaled by the total number of full-time employees are in the upper median of the distribution across banks with AI job postings. For the AI technology M&A, we identify banks that engaged in more than one AI technology M&A in preceding years. We provide detailed description of the methodology in Appendix B

8. We classify a bank as having analytics technologies if it utilizes data-focused analytics technolo-

use of data AI, machine learning, analytics, and big data technologes using AI-related disclosures in the bank's business description of the 10-K or the presentation portion of the conference call following the approach of Chen and Srinivasan (2023).[9] Lastly, we combine these five proxies to create a composite measure of "data-driven" banks, to examine our hypothesis.[10]

Across all measures of data driven technology adoption, we find consistent evidence that local banks do not gain significant market share in small business loans following local newspaper closures in counties where banks utilize these data-driven technologies. On the other hand, in counties where banks do not utilize these technologies, we find consistent evidence that more local banks gain significant market share after newspaper closures. Specifically, we find that the standard deviation of the local deposit proportion in a county and the bank-level local deposit concentration, corresponds to a 1% and 1.3% increase in the shares of local small business loans issued (or a 11% and 14% relative increase). Moreover, we find that the standard deviation of the local small business loan proportion in a county and the local small business loan concentration, corresponds to a 1% and 1.4% increase in the shares of local small business loans issued (or a 11% and 16% relative increase). These results provide further evidence that information advantages traditionally held by local banks diminish when they compete with banks equipped with data-driven technologies.[11]

gies, as detailed in Appendix B, on its website. We consider a bank to have greater data storage expenditure if its data storage budget scaled by the total assets falls within the upper median of its distribution across our sample banks.

9. The AI-related disclosure include terms related to AI, analytics, automation, cloud, big data and machine learning. We provide detailed description of the methodology in Appendix B

10. Specifically, we define a bank as a data-driven bank, if (1) the bank's AI hiring-to-employees is in the upper median of banks that hire workers with AI skills, (2) the bank has engaged in more than one AI technology acquisition in prior years, (3) the bank has analytics technologies on its website, (4) if the bank's data storage budget-to-assets is in the upper median of banks, (5) the bank discloses AI-related terms in its 10-K or conference calls

11. In our online appendix, we conduct additional analyses to provide further insights into the local banks' information advantages. First, we investigate whether local banks increase loan risk to expand market share in small business loans following local newspaper closures. Using non-performing loan (NPL) ratio as a measure for loan risk, we do not find evidence of changes in NPL ratio of local banks relative to non-local banks in the post-newspaper closure period. Second, we further examine whether these banks benefit from an increased market share in small business loans after local newspaper

Our paper makes several contributions. First, we contribute to the growing literature on the role technology in credit markets. Prior studies document that the technology adoption of banks promote entrepreneurship and enhances efficiency in lending process (Ahnert et al. 2021). Moreover, prior work has also shown that the greater usage of tools such as computers, communication equipment and data sharing technologies have allowed banks to reduce reliance on relationship lending and extend physical distances between small firms and lenders (Deyoung et al. 2019; Granja et al. 2022; Sutherland 2018). Our contribution to these studies is to show that technology adoption also plays an important role in generating information on borrowers that could reduce the traditional information advantages of more local banks in informationally opaque small business lending markets.

Second, our paper contributes to the literature on local information advantages. Studies in this literature has shown the existence of local information advantages in the equity market (Coval and Moskowitz 1999) and in credit markets (Carter et al. 2004; Loutskina and Strahan 2011). Our findings contribute to this literature by illustrating that the advent of new technologies could reduce some of the information advantages traditionally enjoyed by local players in credit markets.

Third, our paper contributes to the literature on information asymmetry and competition in lending markets. Existing literature shows that in informationally opaque markets, access to private information enables lenders to gain a competitive advantage through information rents. For instance, Rajan (1992) argue that insider banks have an informational advantage and extract additional rents from borrowers. Others show that when there is an increase in the borrowers' publicly available information, such as an IPO, information asymmetry falls, and outside lenders are better able to com-

closures. Using net interest margins (NIM) as a proxy for profitability, we find some evidence that local banks tend to earn higher NIM compared to non-local banks in the period following newspaper closures. These results indicate that local banks' information advantages not only enable them to expand their market share in small business loan but also potentially enhance profitability without taking additional risks.

pete with insider banks (Schenone 2010). In our study, we focus on the competition between local vs. non-local banks, where local banks arguably exhibit a competitive advantage due to their access to private information on borrowers through relationship lending. Our analysis provides evidence on the impact of reduction in local public information on market competition and lending behavior. And we show that in markets with limited exposure to data-driven technologies, an information shock driven by a local newspaper closure impacts competition in lending markets, as local banks gain market share from peers after these closures.

Lastly, our paper contributes to the literature on the role of media in capital markets. Previous studies have highlighted the important role of media as an information intermediary and as a monitoring agent performing original investigation and analysis. These activities allow media to reduce information asymmetry (Bushee et al. 2010), uncover accounting fraud and other corporate fraud (Miller 2006; Dyck et al. 2008), and shape creditor participation in lending markets (Bushman et al. 2017). We contribute to these studies by showing a relationship between newspapers' activities and information asymmetry in local lending markets. Moreover, we are also among the first to show the declining importance of local media to lenders, as our findings suggest that data driven lending technologies enable non-local banks to substitute information produced by local newspapers.

The rest of the paper proceeds as follows. Section 2 focuses on conceptual underpinnings and institutional details. In section 3, we describe our data and the research design. In section 4 we discuss the results and section 5 concludes.

II Conceptual Underpinnings

II.1 Small Business Lending and Local Banks

In this paper, we focus on the small business lending market.[12] For many communities throughout the US, these loans are an important source of financing for small and medium enterprises (SMEs). According to the small business administration, small businesses are critical to economic growth, as they generate two out of three jobs in the US, and these business play a key role in driving innovation, as well the development of unique products in local economic markets.

While important to local SMEs, the small business lending market, is also one that is characterized by high information problems. Prior work have argued that there tends to be greater asymmetric information problems in these markets (Berger and Udell 1995), as the information asymmetry problems tend to be more acute in the small firms that charactertize this market. Many of these firms are also private companies, and thus provide less disclosure compared to large, public corporations, which often provide rich sets of financial information to potential investors and creditors.

Due to the opacity of the small business lending markets, more local banks have a comparative advantage over their larger, non-local peers. Local banks specialize in soft-information based relationship lending and small-businesses, which often lack hard and verifiable information, are ideal targets for such lenders.[13] Non-local banks, on the other hand, specialize in lending based on hard and verifiable data. Lending decisions at these large, multi-tiered organizations often have to be communicated by loan officers, who interact directly with borrowers, with managers higher up in the organizational hierarchy. Hard, verifiable drivers of lending decisions can be credibly communicated across these hierarchies while soft information cannot (Stein 2002).

12. The federal reserve and the CRA defines small business loans as loans of $1 million or less to businesses.

13. Soft information is often defined as information which cannot be credibly communicated from one agent to another and hence cannot be verified.

Thus, loan officers at these large non-local banks tend to place a premium on hard and verifiable information while screening borrowers. And hence, these banking institutions tend to specialize in non-relationship-based lending.

II.2 Local Newspapers and Local News Deserts

The local newspaper act as an information intermediary by disseminating news generated externally to their subscriber base. In other instances, newspapers themselves generate news by monitoring local governments and businesses and by undertaking investigative journalism to uncover fraud and misconduct. Despite their importance to local economies, studies have documented a precipitous decline in the number of local newspapers, due to technological advances and the consolidation of the newspaper industry.[14] This decline has been so acute, that the phrase "newspaper desert" has been coined to describe local communities that are no longer served by a local newspaper. And in a 2018 UNC report, the authors of the report suggest that roughly 200 counties face this issue.

A local newspaper closure can be viewed as a significant information shock, that increases information asymmetry between borrowers and all lenders as all lenders have to work harder to access the information that newspapers would otherwise provide. Moreover, there is also a decline in monitoring of local businesses and consequently lenders face increased adverse selection risk of selecting borrowers indulging in fraud or other forms of misconduct.

We hypothesize that the increase in information asymmetry following a newspaper closure differentially impacts local and non-local banks. Non-local banks do not actively collect local information relative to local banks so non-local banks are affected more by a newspaper closure. Local banks can lean on their existing lending relationships to stay abreast of local economic and business developments thus limiting

14. See for example the UNC report on expanding newspaper deserts: https://www.usnewsdeserts.com/reports/expanding-news-desert/loss-of-local-news/loss-newspapers-readers/

the impact of newspaper closures. In other words, local newspaper closers reinforce information advantages of local banks and exacerbate adverse selection concerns for non-local lenders. These concerns should prompt non-local banks to compete less aggressively in the affected lending markets, reducing competition in these lending markets and increasing local borrowers' lender switching costs. Therefore, we expect local banks to benefit from newspaper closures and perform better than their non-local peers after such closures.

II.3 Data-Driven Technologies in Banking

Next, we focus on how adoption of data driven lending technologies impact bank lending markets. Alternative data sources have upended the market for information by providing some market participants unprecedented ability to track and monitor local economic and business conditions in real time. These data sources have transformed the way in which companies utilize information about current and potential customers enabling them to perform sophisticated customer analytics and helping them better screen and serve their customers. Businesses that were able to quantify their gains from analyzing big data, a type of alternative data, reported an average 8% increase in revenue and a 10% reduction in overall costs, according to a 2015 survey from BARC.[15]

Like other organizations, banks and other private lenders increasingly rely on analytics and alternative data-driven technologies to screen borrowers, identify spending patterns of customers, perform customer segmentation and profiling, sentiment and feedback analysis, and security and fraud management (Srivastava and Gopalkrishnan 2015; Jagtiani and Lemieux 2019). These technologies include customer analytics, machine learning and artificial intelligence-based platforms, among others. The data sources include, among others, data on insurance claims, customer's use of mobile phones and related activities, internet footprints, online shopping habits, social media

15. Source: Business Application Research Center (BARC), https://bi-survey.com/big-data-benefits

activity, social networks and investment choices (Lin et al. 2012; Jagtiani and Lemieux 2019).[16] Driven, in part, by the availability of such data non-traditional lenders such as Lendingclub, Sofi, Kabbage and Prosper Marketplace have either completely or to a large extent reduced their reliance on traditional information sources such as FICO scores.[17] This decline in importance of traditional information sources in lending markets raises the question whether newspapers have also become less relevant as a source of information.

Moreover, the adoption of analytics and other data-driven technologies could also improve the ability of banks to "harden" soft information (Liberti and Petersen 2019), that are traditionally used and processed by local banks. Several studies in the accounting and finance literatures have demonstrated the ability of analytics, machine-learning and other data-driven technologies to convert soft opinions into quantitative formats that provide value to capital markets (Das and Chen 2007; Li 2008; Huang et al. 2017). In the banking setting, prior work has also provided evidence of the role of hardening soft information in reducing the advantages of relationship-based local banking. Studies have shown that banks use credit scores to harden soft information, which reduces the information asymmetries in distant lending (Beck et al. 2017) and reliance on relationship lending (Sutherland 2018).

Banks that adopt data driven lending technologies can also rely on a multitude of data sources to replicate, and in some cases improve upon, the information that a local newspaper would provide. By providing additional data points, these information sources may allow lenders to screen borrowers they previously did not serve because of adverse selection concerns. Thus, lenders who invest in these technologies may continue to compete against local bank with information advantages following a local newspaper closure.

16. Source: Comments by the CFPB director Richard Cordray at the LendIt USA Conference (March 2017).

17. "Will Fintechs kill the FICO score", Penny Crosman, American Banker (June 2016)

Non-local banks are larger and consequently have a higher ability and willingness to adopt these data technologies than their local counterparts. Non-local banks – because of their larger size and wider geographical footprint – have deeper pockets and can invest in these technologies and spread their costs and benefits over the different markets they compete in thus earning a higher return on investment than local banks. Local banks are less likely to afford the large investments that are required to maintain these technologies as data driven lending technologies can be prohibitively expensive to install and maintain. In a survey done by UBS, 70% of respondents from banks with less than $100 billion of assets noted that budgetary constraints represent a key impediment to the successful implementation of a technology strategy (Martinez et al. 2019). On the adoption of such technologies, a similar trend emerged from the survey. 75% of respondents from banks with assets above $100 Billion said they are already implementing cutting edge technologies such as AI strategies, while less than half of respondents from sub $100 Billion asset banks noted that such strategies are being implemented.

The marginal benefits of data driven lending technologies are higher for non-local banks than local banks. Local banks, driven by their focus on relationship lending, focus more on soft information. Soft information is often defined as information that cannot be communicated from one agent to another and also cannot be stored in databases. Borrowers are also more willing to share sensitive private information with relationship lenders (Boot 2000; Bharath et al. 2011). Local lenders might thus find less useful the information advantages obtained from these technologies than their non-local peers.

Non-local banks, on the other hand, rely more on public information and may find it more profitable to invest in data driven lending technologies than local banks. Non-local banks should be more likely to invest in these technologies and hence reduce their reliance on traditional data sources including local newspapers. Thus, in markets where data driven lending technologies play a larger role, non-local banks should be able to compete more aggressively with their local counterparts once a local newspaper shuts down.

Based on these arguments, we hypothesize that in markets where data driven lending technologies play a larger role, non-local banks compete more aggressively with their local peers following a newspaper closure than in those markets where such technologies play a smaller role. This increased competition in lending markets reduces information advantages that otherwise accrue to local banks following a local newspaper closure. Specifically, for our study, We expect that local banks experience a lower increase in the market share of information-intensive loans made in markets where non-local banks utilize data driven lending technologies to substitute the information provided by local newspapers.

III Data

We collect data on small business and and farm loans issued by deposit-taking institutions at the county-level from the FFIEC's CRA dataset.[18] We obtain deposit data from the summary of deposit database provided by the FDIC which contains yearly branch-level records of the amount of deposits taken by depository institutions.[19] We measure bank characteristics using the call reports. We identify local newspaper closures using data from Heese et al. (2022).[20] In Figure 6, we describe the distribution of local newspaper county closure over our sample period.

18. This database collates information from disclosures of depository institutions' small business loan activities, from 1997-2021. This disclosure is mandatory for depository institutions with above roughly 1 B. USD in assets.

19. We aggregate the deposits measure at the bank, county and year-level.

20. Heese et al. (2022) obtain data on closures of U.S. daily newspapers from three data sources. First, data on daily newspapers comes from the United States Newspaper Panel constructed by Gentzkow et al. (2011). Second, data on newspaper closures is collected using UNC's Center for Innovation and Sustainability in Local Media's Database of Newspapers. Lastly, data on newspaper closures is collected by scraping content from the U.S. Newspaper Directory of Chronicling.

III.1 Measuring Data-Driven Technologies

In addition, we employ five different datasets to create proxies for data-driven technology adoption. First, we measure the corporate demand for AI workers using the job posting data following the literature on the AI adoption (Acemoglu et al. 2022; Alekseeva et al. 2021; Babina et al. 2022). We use the skills list in Alekseeva et al. (2021) and the skills clusters that are complementary to machine learning in Goldfarb et al. (2023) to measure corporate demand for workers with AI skills at the parent bank-level from 2010-2021. We then code banks that are in the upper median of AI demand-to-employees for banks with AI skills demand, as banks with high AI demand.[21][22] We then code treated counties that are served by such banks as counties with high AI demand banks. Figure 1 presents the distribution of AI worker demand-to-employees for the banks in our sample, and the figure shows a clear upward trend in the demand of workers with AI-related skills.

Second, we proxy for the data-driven technology adoption by the presence of a prior AI technology M&A using a dataset of technology M&As provided by S&P 451 (also used in, Jin et al. 2023). To construct a sample of AI technology M&A, we reviewed the business descriptions of the target firms from 2001-2021, and coded target firms as AI-related if they mentioned AI-related terms (list of terms presented in Appendix B) in the business description. To proxy for the extent of AI M&As, we code banks that have acquired in more than 1 AI technology target firm in the preceding years, as banks with significant AI M&A activity.[23] We then code treated counties that are served by

21. We merge the Burning Glass dataset to our bank dataset, by fuzzy matching with bank names in the Call reports, and the employer name in Burning Glass. In our fuzzy matching algorithm, we keep the top 3 matches by levenshtein distance and manually validate each of these potential matches. In cases where the employer names match to multiple banks, we keep the observation that matches with the same county or zipcode location. For multiple matches, after this step, we discard the matches.

22. For the list of skills and skill clusters, see Appendix B

23. We merge the technology M&A dataset to our bank dataset, by fuzzy matching with bank names in the call reports, and the acquirer name in technology M&A dataset. In our fuzzy matching algorithm, we keep the top 3 matches by levenshtein distance and manually validate each of these potential matches.

such banks as counties with AI M&A banks. Figure 2 presents the distribution of AI technology M&A, and like Figure 1, we see an upward trend in the investment in AI technologies through M&As.

Third, we measure data-driven technology adoption with the extent of website analytics technologies in banks (also used in, Charoenwong et al. 2022). We use Builtwith to measure website analytics technologies from with some data collection capabilities from 2002-2021 (see Appendix A for more details on the specific analytics technology categories), and we study the presence of these technologies as a proxy for the extent of web analytics technologies in the bank.[24] We then code treated counties that are served by such banks as counties with analytics banks. Figure 3 plots the cumulative counts of web analytics technologies in companies, and we also see an increasing trend of the use of these technologies in banks starting from 2010.

Fourth, we measure data-driven technology adoption by the extent of expenditure related to data storage. Conceptually, this variable captures the investment in big data, and we measure data storage with the data storage budgets reported in the Aberdeen database from 2007-2019. We code banks that are in the upper median of data storage budget-to-assets for all banks as banks with high data storage budgets.[25] We then code treated counties that are served by such banks as counties with high data storage banks. Figure 4 presents the distribution of data-storage budget-to-assets across time, and we again see a clear upward trend in expenditures on data storage.

Lastly, we measure the adoption of the broader field of analytics, big data and AI technologies, by identifying banks that have made AI-related disclosures in the 10-K and conference calls. To do so, we build on the methodology in Chen and

24. We extract the analytics technology data from Builtwith by searching the bank's corresponding website information (as reported in the call reports) from Builtwith.

25. We merge the Aberdeen data to our bank dataset, by fuzzy matching with bank names in the Call reports, and the employer name in Burning Glass. In our fuzzy matching algorithm, we keep the top 3 matches by levenshtein distance and manually validate each of these potential matches. We also further require that the website reported in the call report (where available) matches the website reported in Aberdeen. If the website is not available, we further require that the zipcode reported in Aberdeen matches the zipcode reported in the Call reports. All other matches are discarded.

Srinivasan (2023) to measure the extent of AI-related activity in banks through their disclosure on 10-K filings and earnings conference calls from 1997-2021 (we focus only on terms relating to analytics, automation, big data, cloud, machine learning and AI categories).[26][27][28] Specifically, we count the number of AI-related terms in the business description section of the 10-K report and the presentation section of the conference call, and code banks with at least 1 AI-related terms as AI-related disclosure banks. We then code treated counties that are served by such banks as AI-related disclosure bank counties. Figure 5 presents the distribution of digital terms and digital banks over time. Consistent with the trends reported in Chen and Srinivasan (2023), we find that there is an increasing disclosure of digital terms, particularly in the time-period between 2015-2021.

III.2 Summary Statistics

The sample selection of our study is as follows. We begin with all banks that make call reports to the Federal Reserve, CRA disclosures, and FDIC summary of deposits. We then keep reporting entities that are commercial banks (200), savings bank (300), savings and loans association (310), corporate banks (320), credit union (330), industrial banks (340) and corporate credit unions (370) We then restrict the sample to those that file with the FDIC and the FFIEC. We then construct a treated and control sample for counties surrounding a newspaper closure, and we drop bank-county observations that are more than 3 years from the newspaper closure event. In total we obtain a bank-county-year sample of nearly 36,000 observations, across 1997-2021. To reduce the effects of outliers, we also winsorize the continuous variables in our sample

26. We merge the 10-K and conference call data to CRSP by the wrds-sec link table and company tickers respectively. We then merge the text-based dataset with the main dataset using the crsp-frb link table provided by the New York Federal Reserve (https://www.newyorkfed.org/research/banking_research/datasets.html)

27. For the conference calls, our data runs from 2000-2021 and we code AI-related terms from conference call as zero if the conference call transcript is not available.

28. For more details on the termsthat we use to identify AI-related technologies, see Appendix B

at the 1% and 99% level.

In Table 1 we examine the key summary statistics of the sample. In Panel A, we examine the statistics of the full sample, and show that the average bank in our sample has assets of $23 Billion USD. The distribution of assets is highly skewed with the median bank's assets ($2 Billion) well below the mean assets; consequently, in our regressions we use log of assets as a control variable. The average net interest margin is 4%. On average a bank accounts for 7.1% of a local market's small business loans while a local market accounts for 4% of total loans of an average bank, or 163 million USD on average. In Panels B-C of the table, we present the sample statistics of the banks in treatment and control counties (defined with the local newspaper closure variable in Appendix A) respectively.

III.3 Research Design

Our primary research design is a staggered difference-in-difference design. The main source of exogenous variation is the closure of local newspapers, which we use as an exogenous shock to the availability of local information. Thus, to implement this analysis, we first classify counties that have been impacted by a newspaper closure as treated counties.[29][30] With these treated counties, we then construct a matched sample (with replacement) of adjacent counties as a control group. To increase the precision of our tests, we keep treated and control observations that are no more than 3 years from the newspaper closure.[31]

As several economic factors might be associated with newspaper closures, we utilize an array of fixed effects to control for these confounding factors. We control for time

29. Data on the adjacent counties is taken from the NBER website: https://www.nber.org/research/data/county-adjacency.

30. We restrict analysis to treated countries where there are at least 3 banks in the control group, and we restrict the counties in the control group to those that were not treated throughout the sample. For control-counties that are matched to different treatment-counties, we match the banks in the control-counties to each treatment county with a different identifier (that is, we weight the control observation by the number of treated county matches).

31. For counties with multiple closures, we limit the analysis to only the first closure.

invariant differences between banks in a county by using bank-county fixed effects. To control for time-trends and biases arising from treatment effect heterogeneity (Baker et al. 2022), we stack the regression by matched pairs (in our specification with matched-county-pair-year FE) and control for the interaction of matched-pair FE and year FE. Moreover, to further address concerns that the staggered DiD estimator exhibits bias in the presence of treatment effect heterogeneity (Baker et al. 2022), in our regressions we only use those counties that never get treated as controls.

The key estimate of interest in this study is the relative importance of information provided by local newspapers for non-local banks compared to local banks. Thus, we interact the traditional staggered difference-in-differences estimator with the bank's extent of local operations, which os proxied by several variables that measure the concentration of the bank's local operations. Specifically, we examine four different proxies, namely, the local deposit proportion in a county (relative to the bank's total deposits), the local deposit concentration, the local SBL lending proportion in a county (relative to the bank's total SBLs) and the local SBL lending concentration (See Appendix A for more details on the variable construction).

Equation 1 below presents the regression model that is used in the main analysis:

$$SBL\ M/S_{i,c,t} = \alpha_{i,c} + \alpha_{s,t} + \beta_1 Close_{c,t} + \beta_2 Conc_{i,c,t} + \beta_3 Close_{c,t} \times Conc_{i,c,t} + \sum_k \gamma_k X_{i,t} + \epsilon_{i,c,t} \tag{1}$$

where $Var_{i,c,t}$ is the market share of local SBL loans, indexed by bank i, county c and year t. $Close_{c,t}$ and $Conc_{i,c,t}$ are the indicators for newspaper closure and one of the four variables that measures the concentration of the bank's operations in a county.

In addition, we implement the fixed effect structure that we outlined in the previous paragraph, as well as a vector of controls. Specifically, we control for the bank-county fixed effect ($\alpha_{i,c}$) and the interaction between matched pair fixed effect (indexed by s) and year fixed effects ($\alpha_{s,t}$). To control for size, we include the logarithm of total book assets as a control. To control for the loan profile of the bank, we control for loan loss provision relative to total loans and loan growth. Finally, to control for the extent to which banks engage in small business lending, we control for the proportion of total small business loans at the bank-level relative to the total of all loans. To address concerns of serial correlation in the error terms, in all our regression analyses, we cluster standard errors at the bank, county and year level.

IV Results

IV.1 Local Newspaper Closure and Local Information Advantage

We begin our analysis by studying whether local newspapers closures increase the information advantage of local banks relative to non-local banks. Over the 1997-2021 time period, we find some evidence that local banks have leveraged from the exit of a local newspaper, by increasing their market share in small business loans. These results are presented in Table 2, where we examine the changes in the market share of local small business loans, for local and non-local banks in the years after the newspaper closure. In Panel A, of this table we measure the extent of local operations with the local deposit proportion in a county, and the bank-level local deposit concentration (HHI index). Across both types of concentration measures, we find consistent evidence that more local banks tend to gain more market share in the local small business loan market, when a newspaper closes in the county. In the first column, our estimates suggest that a one standard deviation difference in the local deposit proportion (0.211) corresponds to a 0.4% increase in market share (or a 4% relative increase relative to the sample average of 8.8%). In the following column, we also measure deposit concentration at the bank-level, with the local deposit concentration. In these regression analysis, presented in the following column, our estimates suggest that a one standard deviation difference in the HHI index of deposits

(0.296) corresponds to a 0.7% change in market share (or a 8% relative increase).

In Panel B, we conduct the same analysis but with concentration of local operations measures based on small business loans (SBLs). Across the local SBL lending proportion and concentration (HHI-index) measures, we find consistent and similar results. In the first column of Panel B, we find that a one standard deviation difference in the local SBL lending proportion (0.197) corresponds to a 0.4% increase in market share (or a 5% relative increase). In the next column, we study the extent of local operations of a bank by the local SBL lending concentration. For this regression analysis, our estimates show that a one standard deviation difference in the HHI index of small business loans (0.289) increases 0.9% (or a 10% relative increase). Hence, consistent with our predictions, our analysis suggest that more local banks tend to gain more market share in the period after newspaper closures.

IV.1.1 Temporal Variation in Local Information Advantage: Before and After the Advent of Cloud Computing Services

To first provide some descriptive evidence of the differential impact of newspaper closure on the market competition between local and non-local banks, we analyze the main regressions in sub-samples across time.

We choose two time-periods for analysis, namely, the pre-2007 period and the post-2010 period (we drop 2008 and 2009 from the sample due to the GFC). And based on the literature, we expect the information advantage for local banks to be smaller in the period after 2010, as the period between 2007-2010 coincides with major developments in cloud computing services (Ewens et al. 2018), which served as an important platform and infrastructure for the development of data-driven technologies.[32] More-

32. Specifically, the advent of Amazon Web Services in 2006 and other cloud platform services marks an important turning point in the use of data-driven technologies follows the arguments articulated in Ewens et al. (2018). The development of AWS helped many entrepreneurs to experiment and create new digital and data-driven technologies, as it enabled these developers to rent "hardware" space on the cloud. Moreover, Microsoft Azure, another key cloud services provider, came online in 2010. Thus in the period between 2007-2010, there were substantive improvements in the availability

over, the choice of studying the pre-2007 sample follows prior work on the information advantages in local banks which have ended their sample in 2007 (Loutskina and Strahan 2011). Thus, studying the pre-2007 landscape of information advantages for local banking institutions allows us to validate prior work, and analyzing the post-2007 landscape of local information advantages, allows us to tackle an open empirical question.

In Table 3 we investigate the aforementioned time-periods using the same regression framework as Table 2. In Panel A, we examine the two sub-samples with the extent of local operations proxied by the local deposit proportion in a county, and the bank-level local deposit concentration (HHI index). Our analysis shows that the post-newspaper closure gains in SBL market share for more local banks, tends to be significant only in the period before the advent of cloud-technologies in the 2007-2010 period. In particular, our estimates show that a one standard deviation difference in the local deposit proportion (0.211) corresponds to a 0.7% increase in market share (or a 8% increase relative to the sample average of 8.8%), in the sub-sample time period from 1997-2007. Furthermore, in the same sub-sample time-period, a standard deviation difference in the local deposit concentration (0.296) corresponds to a 1.2% change in market share (or a 14% relative increase). On the other hand, in the time-period from 2010-2021, we find no significant difference in market share between local and non-local banks after newspaper closures.

To further assess the validity of the difference-in-difference estimator in our regression analysis in Table 2, we examine regressions with indicators for pre-newspaper closures and post-newspaper closures to check for pre-trends in the period before the newspaper closures. Specifically, we study the following regression model in Equation

of platforms that facilitated data-driven technologies.

2:

$$SBL\ M/S_{i,c,t} = \alpha_{i,c} + \alpha_{s,t} + \beta_1 Close_{c,t,t+2} + \beta_2 Close_{c,t-2,t+5} + \beta_4 Conc_{i,c,t} \qquad (2)$$
$$+\beta_5 Close_{c,t,t+2} \times Conc_{i,c,t} + \beta_6 Close_{c,t-2,t+5} \times Conc_{i,c,t} + \sum_k \gamma_k X_{i,t} + \epsilon_{i,c,t}$$

where $Close_{c,t,t+2}$ and $Close_{c,t-2,t+5}$ are the indicators for years t to $t+2$ relative to the newspaper closure, and $t-2$ to $t-5$ relative to the newspaper closure.

In Panel B of Table 3, we study the pre and post-trends of small business loan market share for local and non-local banks in counties that experience newspaper in the pre-cloud and post-cloud time-periods in Panel A. In the first two columns of this panel, we find that in the period after the advent of cloud computing services, there is no pre-trend in small business loan market share differences across local and non-local banks, but there are also no significant changes in market share across both types of banks in the post newspaper closure period. On the other hand, we find that there is a positive change in market share for banks with greater local deposit proportion for the pre-cloud sub-sample. Moreover, we also show that the these banks exhibit no significant pre-trend in differences of small business loan market share, as the coefficient on $Closure\ Event_{i,t-2,t-5} \times Conc_{i,t}$ is statistically insignificant. In addition, the estimates on the $Closure\ Event_{i,t,t+2} \times Conc_{i,t}$ variable suggests that a one standard deviation difference in the local deposit proportion corresponds to a 0.6% increase in market share (or a 7% relative increase) after the newspaper closure in the pre-cloud period.

Next, we perform the same set of analysis with the concentration of local operations defined at the bank-level, using the local deposit concentration (HHI-index) variable. Our analysis finds similar results. In the post-cloud period, we find statistically insignificant differences in local and non-local banks in the pre and post-newspaper closure periods. On the other hand, in the pre-cloud period, we find that local banks gain market share in the post-newspaper closure period, and these banks exhibit statistically insignificant differences in market share in the pre-newspaper closure period. Specifically, our estimate shows that a one standard deviation difference in the local deposit concentration corresponds to a 1.2% increase in market share (or a 14% relative increase) after the newspaper closure in the pre-cloud period.

In Panels C and D of Table 3, we examine the two sub-samples with the extent of local operations proxied by the local SBL lending proportion in a county, and the bank-level local SBL lending concentration (HHI-index). In Panel C, we find similar results as Panel A. We find that a one standard deviation difference in the local SBL lending proportion (0.197) corresponds to a 0.7% increase in market share (or a 8% relative increase), in the sub-sample time period from 1997-2007. Moreover, in the same sub-sample time-period, a standard deviation difference in the local SBL lending concentration (0.289) increases 1.2% (or a 14% relative increase). On the other hand, in the time-period from 2010-2021, we find no significant difference in market share between local and non-local banks after newspaper closures.

In Panel D of Table 3, we study the same set of analysis but with indicator for the pre- and post-newspaper closure period. In the first and second column, we study the changes in small business loan market share of local and non-local banks defined using the local SBL lending proportion at the bank-county-level in the post- and pre-cloud period, respectively. Our analysis in the first column shows no statistically significant difference across local and non-local banks in the pre- and post-newspaper closure periods. In the second column, we find that in the pre-cloud period, local banks gain market share in the post-newspaper closure period. Specifically, our estimate shows that a one standard deviation difference in the local deposit concentration corresponds to a 0.7% increase in market share (or a 8% relative increase) after the newspaper closure in the pre-cloud period.

Moreover, we find statistically stronger results in the specification with concentration of local operations defined using the bank-level local SBL lending concentration. As before, we analyze the post and pre-cloud sub-samples separately. In the post-cloud sub-sample, we find no statistically significant difference across local and non-local banks in the pre and post-newspaper closure period. On the other hand in the pre-cloud sub-sample, we find a significantly positive difference in loan market share between local and non-local banks in the post-newspaper closure period. Additionally, we also find no statistically significant differences between local and non-local banks in the pre-newspaper closure period. To provide some economic magnitudes of this result, our estimates in this regression analysis show that one standard deviation difference in the local SBL lending concentration corresponds to a 1.2% increase in market share (or a 14% relative increase) after the newspaper closure in the pre-cloud sub-sample.

Taken together, our analysis in Table 3 suggests that there are limited pre-trends in our analysis, as the differences in small business loan market shares between local and non-local banks before newspaper closures are statistically insignificant. Moreover, our analysis also provides some evidence that the main effects of newspaper closure on the competition between local and non-local banks occur in the pre-cloud period, which is before the widespread use of many data-driven technologies.

IV.2 Data-Driven Technologies and Local Information Advantage

An empirical challenge arises from interpreting the previous set of analyses, is that while 2007-2010 might mark the point in time after which data driven lending technologies start getting widely adopted, it also coincides with the GFC and the beginning of an era of significant regulatory changes in the banking industry enacted in response to the GFC. These regulatory changes might confound our results as they can reasonably be expected to have an effect on the competitive dynamics within the banking industry. While many of the provisions of the Dodd Frank act, passed in response to the GFC, were designed in such a way so as to limit their impact on small banks, survey and empirical evidence suggests that small banks did feel the burden of the extra regulations imposed as a part of this act (Peirce et al. 2014; Bordo and Duca 2018). Moreover, prior work has also shown that distant small business lending and risk-taking shifts along with boom and bust cycles, such as the 2008 GFC (Granja et al. 2022).

To dig deeper into the role that data-driven technologies play in reducing the local information advantage that local banks have traditionally enjoyed after newspaper closures, we study several cross-sectional analysis, that reduces the reliance on a time-trend comparison. Specifically, we measure several proxies of data-driven technologies at the bank-level, and examine whether the presence of such banks in the treated county, limits the extent to which local banks gain market share after a newspaper closure.

As the concept of data-driven technologies encompasses many different technologies across analytics, big data and artificial intelligence, we measure data-driven technologies with five proxies. To measure AI technologies, we use the demand for workers with AI skills and the presence of AI-related technology M&As. To measure analytics technologies, we examine the presence of analytics technologies with some data collection capabilities, in the website of banks. To measure big data, we examine the data storage budget in banks. Lastly, to measure the aggregate use of these technologies, we study the disclosure of AI, analytics and big data terms in the business description of 10-K and presentation portion of the conference call of banks.

IV.2.1 AI Technologies

We examine the first proxy, the demand for workers with AI skills, in Table 4. In Panel A, we begin our analysis with the concentration of local operations defined using deposit-based measures. Across the two measure of concentration, namely, the local deposit proportion and concentration, we find consistent evidence that more local banks tend to gain greater market share after newspaper closures, in counties that are not served by banks with high AI demand. Specifically, in this sub-sample, our estimates show that a standard deviation difference in the local deposit proportion and concentration, are associated with a 47% and 26% increase in market share respectively. Moreover, these estimates are also statistically more positive than the estimates from the sub-samples of treated counties that are served by banks with high AI demand.

Next, in Panel B, we examine the same analysis with the concentration of bank operations at the local-level defined using small business loan-based measures. Like our analysis in the previous panel, we study two measures of concentration, namely, the local SBL lending proportion and concentration. With these concentration measures, we also find evidence that more local banks tend to gain greater market share after newspaper closure, in counties that are not served by banks with high AI demand. Specifically, in these sub-samples, our estimates show that a standard deviation difference in the local SBL lending proportion and concentration, are associated with a 77% and 15% increase in market share respectively. In addition, these estimates are also statistically more positive than the estimates from the sub-samples of treated counties that are served by banks with high AI demand.

To further examine the role of AI technologies in influencing the local information advantage of local banks following newspaper closures, we examine the second proxy, AI-related technology M&A, in Table 5. In Panel A, we examine the same analysis in Table 4 Panel A, but with the presence of a past AI-related technology M&A as the cross-sectional variable. We study the concentration of local operations with the local deposit proportion in the first two columns. And we find some consistent but not statistically significant results that suggest that more local banks tend to gain more market share in treated counties that are not served by high AI demand banks in the post-newspaper closure period. In the following two columns, which measures the concentration of local operations with the local deposit concentration, we find stronger statistical evidence that suggest that local banks gain market share after newspaper closures in counties that are not served by high AI demand banks. Specifically, our estimates suggest that a standard deviation difference in the local deposit concentration, is associated with a 0.6% increase in market share (or a 6% relative increase). Moreover, this estimate is also statistically more positive than the estimates from the sub-samples of treated counties that are served by banks with AI technology M&As.

We turn to the same set of analysis with the concentration variables defined using small business loans in Panel B. Like our previous set of analyses, we study two measures of concentration of local operations, namely, the local SBL lending proportion and concentration. With these concentration measures, we find further evidence that more local banks tend to gain greater market share after newspaper closure, in treated counties that are not served by banks with AI technology M&As. Specifically, in these sub-samples, our estimates show that a standard deviation difference in the local SBL lending proportion and concentration, are associated with a 0.4% and 0.9% increase in market share respectively (or a 4% and 10% relative increase, respectively). In addition, these estimates are also statistically more positive than the estimates from the sub-samples of treated counties that are served by banks with AI technology M&As.

Thus overall, our analysis across Tables 4 and 5 suggest that the use of AI technologies in banking has decreased the market share gains that local banks have traditionally enjoyed following newspaper closures.

IV.2.2 Analytics Technologies

In Table 6, we examine the role of analytics with the website analytics technologies measured in Builtwith. To ensure greater precision of our tests, we focus on specific categories of technologies that are likely to enable the bank to glean insights on customers, from user behavior on their website (see Appendix A for more details on the categories of technologies). The structure of our analysis follows the previous analysis in Tables 4 and 5, but we use the presence of banks with web analytics technologies in the treated county as the cross-sectional variable in Table 6. Panel A reports the estimates from the regression with concentration of local operations variables defined using deposit data. Across the two measures of concentration of local operations, we find consistent evidence that more local banks tend to gain greater market share after newspaper closure, in treated counties that are not served by banks with web analytics technologies. Specifically, in these sub-samples, our estimates show that a standard deviation difference in the local deposit proportion and concentration, are associated with a 0.9% and 1.4% increase in market share respectively (or a 10 and 16% relative increase respectively). Moreover, these estimates are also statistically more positive than the estimates from the sub-samples of treated counties that are served by banks with website analytics technologies.

Next, we turn to the analysis with the concentration of local operations variables defined using small business loans in Panel B. With the small business loan-based concentration measures, we continue to find further evidence that more local banks tend to gain greater market share after newspaper closure, in treated counties that are not served by banks with website analytics technologies. Specifically, in these sub-samples, our estimates show that a standard deviation difference in the local SBL lending proportion and concentration, are associated with a 1% and 1.3% increase in market share respectively (or a 11 and 15% increase respectively). In addition, we also find that these estimates are also statistically larger than the estimates from the sub-samples of treated counties that are served by banks with high AI demand. Hence, our overall analysis in Table 6 suggest that the presence of banks with analytics technologies also diminishes the market share gains that local banks have traditionally enjoyed following newspaper closures.

IV.2.3 Big Data Technologies

We examine the fourth proxy, the extent of data storage budgets in Table 7. As discussed in Section 3, we measure data storage budget with Aberdeen, and for comparability across banks, we scale the data budgets by the bank's total assets. Panel A, reports the cross-sectional analysis across treated counties that are served by banks with high and low data storage budget, for the specification with the concentration of local operations variables defined using bank deposits. In the first two columns, we study the local deposit proportion and we find that banks that are more concentrated and local tend to exhibit higher market share after newspaper closures, in treated counties that are not served by banks with high data storage budgets. Specifically, our estimates show that in these counties, a standard deviation difference in the local deposit proportion, corresponds to a 87% increase in small business loan market share. Moreover, this increase is also higher than the estimate in the sub-sample of treated counties with high data storage budgets.

In Panel B, we perform the same analysis but with the concentration of local operations variables defined using small business loans. Overall we find similar results — banks with higher proportion and concentration of SBLs tend to gain greater market share after newspaper closures, in treated counties that are not served by banks with high data storage budgets. Specifically, we find that a standard deviation difference in the local SBL lending proportion and concentration, are associated with a 60% and 20% increase in market share respectively. These estimates are also larger and more positive compared to the estimate reported in the sub-sample of treated counties that are served by banks with high data storage budgets. Thus, taken together, our cross-sectional analysis of the data storage-based sub-samples suggest that the use of big data technologies also reduces the market share gains that local banks have traditionally enjoyed after newspaper closures.

IV.2.4 AI, Analytics and Big Data Technologies

Our final proxy, the extent of AI-related disclosure, is presented in Table 8. As discussed in Section 3, we measure the AI-related disclosure using the analytics, big data, cloud, AI and automation terms in Chen and Srinivasan (2023), and we code banks with at 1 term from these categories as AI-related disclosure banks. Panel A, reports the cross-sectional analysis across treated counties that are served by banks with/without AI-related disclosures, for the specification with the concentration of local operations variables defined using bank deposits. Across the two measures of concentration of local operations, namely, the local deposit proportion and concentration, we find consistent evidence that more local banks tend to gain greater market share after newspaper closure, in treated counties that are not served by banks with AI-related disclosure. Specifically, in these sub-samples, our estimates show that a standard deviation difference in the local deposit proportion and concentration, are associated with a 0.7% and 1.2% increase in market share respectively (or a 8 and 13% relative increase). Moreover, these estimates are also statistically larger than the estimates from the sub-samples of treated counties that are served by banks with high AI demand.

We turn to the same analysis but with concentration of local operations variables defined using small business loans in Panel B. With the small business loan-based concentration measures, we continue to find further evidence that more local banks tend to gain greater market share after newspaper closure, in treated counties that are not served by banks with website analytics technologies. Specifically, in these sub-samples, our estimates show that a standard deviation difference in the local SBL lending proportion and concentration, are associated with a 0.68% and 1.4% increase in market share respectively (or a 8 and 15% increase respectively). In addition, we also find that these estimates are also statistically larger than the estimates from the sub-samples of treated counties that are served by banks with high AI demand. Much like our previous set of analyses, our results in Table 8 suggest that the presence of banks with that discuss analytics, big data and AI technologies in the county also diminishes the market share gains that local banks have traditionally enjoyed following newspaper closures.

To examine the combined effects of the proxies that we have study in the preceding five tables, we combine the proxies into one measure and analyze the cross-section with this measure in Table 9. Specifically, we create a measure that identifies banks as data-driven banks if the bank either (1) discloses AI-related terms in the 10-K, (2) exhibits above median demand for workers with AI skills, (3) exhibits above median data storage budget, (4) has engaged in more than 1 AI technology M&A in the preceding years and (5) has analytics technologies on its website. We then use this measure to split the sample into treated counties that have been served by data-driven banks, and those that have not been served by such banks.

Table 9 presents the results of cross-sectional analysis with this definition of data-driven banks. In Panel A, we study the concentration of local operations variable defined using deposits, and we find consistent evidence that more concentrated and local banks tend to gain market share in the post-newspaper closure period, in treated counties that are not served by data-driven banks. Notably, our estimates show that, in these sub-samples, a standard deviation difference in the local deposit proportion and concentration, are associated with a 1% and 1.3% increase in market share respectively (or a 11 and 14% increase respectively). Moreover, these estimates are also statistically more positive than the estimates from the sub-samples of treated counties that are served by banks with data-driven technologies.

Next, we study the cross-sectional analysis with the concentration of local operations variable defined using small business loans in Panel B. Consistent with the previous panel, we continue to find evidence that more concentrated and local banks tend to gain market share in the post-newspaper closure period in treated counties that are not served by data-driven banks. In particular, our estimates show that, in these sub-samples, a standard deviation difference in the local SBL lending proportion and concentration, are associated with a 1% and 1.4% increase in market share respectively (or a 11 and 16% increase respectively). In addition, these estimates are also statistically more positive than the estimates from the sub-samples of treated counties that are served by banks with data-driven technologies.

Overall, our analysis across Tables 4 to 9 paint a consistent picture — that, the advent and use of analytics, big data and AI technologies have diminished the market share gains that more local and concentrated banks have traditionally enjoyed after newspaper closures.

IV.3 Robustness Analysis

Lastly, we conduct and analyze several robustness analysis in the internet appendix. First, we address the concern of whether the market share gains that we document for more local banks in treated counties with less data-driven technology exposure, is due to the issuance of riskier loans. We address this point by studying the changes in non-performing loans-to-total loans for these banks in the post-newspaper closure period. Our findings show that more local banks (defined using deposit and SBL-based measures) exhibit statistically insignificant differences in NPL-to-loans in the period after the newspaper closure. Thus, we find some suggestive evidence that changes in the riskiness of the loan portfolio is unlikely to explain our main findings.

Second, we examine whether the increase in market share of small business loans for more local banks in treated counties with less data-driven technology, benefits these banks. We study this question by examining the changes in net interest margins (defined at the bank-level), for these banks in the post-newspaper closure period. Overall, we find consistent results that suggest that banks with a higher local operations tend to earn higher net interest margins in treated counties that are not served by data-driven banks. Thus, we find some suggestive evidence that more local banks in treated counties with less data-driven technologies, tend to benefit from local newspaper closures.

V Conclusion

Advancements in data collection and storage technologies have enabled certain market participants to track information about businesses and economies on a real time basis. From studying large macroeconomic data like the billion prices project to minute details such as tracking the traffic flow in a retailer's store parking lot, alternative data and data driven lending technologies allow market participants to supplant traditional information intermediaries like newspapers. Our paper is one of the first to show how availability of these data sources and lending technologies diminishes local information advantages of local banks. Our paper also speaks to the increasing role data driven technologies and alternative data are playing in the banking market. Traditional measures of borrower quality are being supplemented by non-traditional data points which allow lender unprecedented ability to screen borrowers. How these changes continue to impact usefulness of traditional information sources and the lending markets remains a question of widespread interest.

References

Acemoglu, D., D. Autor, J. Hazell, and P. Restrepo. 2022. "Artificial Intelligence and Jobs: Evidence from Online Vacancies." *Journal of Labor Economics* 40, no. S1 (April): S293–S340.

Ahnert, T., S. Doerr, N. Pierri, and Y. Timmer. 2021. "Does IT help startups? Information Technology in Banking and Entrepreneurship." *Working Paper.*

Alekseeva, L., J. Azar, M. Giné, S. Samila, and B. Taska. 2021. "The demand for AI skills in the labor market." *Labour Economics* 71 (August): 102002.

Babina, T., A. Fedyk, A. X. He, and J. Hodson. 2022. "Artificial Intelligence, Firm Growth, and Product Innovation." *SSRN Electronic Journal* (May).

Baker, A. C., D. F. Larcker, and C. C. Wang. 2022. "How much should we trust staggered difference-in-differences estimates?" *Journal of Financial Economics* 144 ([2]): 370–395.

Beck, T., and A. Demirguc-Kunt. 2006. "Small and medium-size enterprises: Access to finance as a growth constraint." *Journal of Banking & Finance* 30 ([11]): 2931–2943.

Beck, T., V. Ioannidou, and L. Schäfer. 2017. "Foreigners vs. Natives: Bank Lending Technologies and Loan Pricing." *https://doi.org/10.1287/mnsc.2016.2706* 64, no. 8 (April): 3792–3820.

Berger, A. N., and L. K. Black. 2011. "Bank size, lending technologies, and small business finance." *Journal of Banking & Finance* 35 ([3]): 724–735.

Berger, A. N., and G. F. Udell. 1995. "Relationship Lending and Lines of Credit in Small Firm Finance." *The Journal of Business* 68 ([3]): 351–381.

———. 2002. "Small Business Credit Availability and Relationship Lending: The Importance of Bank Organisational Structure." *The Economic Journal* 112, no. 477 (February): F32–F53.

Bharath, S. T., S. Dahiya, A. Saunders, and A. Srinivasan. 2011. "Lending Relationships and Loan Contract Terms." *The Review of Financial Studies* 24, no. 4 (April): 1141–1203.

Biggs, T. 2002. "Is small beautiful and worthy of subsidy? Literature review." *International Finance Corporation (IFC).*

Boot, A. W. 2000. "Relationship Banking: What Do We Know?" *Journal of Financial Intermediation* 9, no. 1 (January): 7–25.

Bordo, M. D., and J. V. Duca. 2018. "The Impact of the Dodd-Frank Act on Small Business." *NBER Working Paper* (Cambridge, MA) 24501.

Bresnahan, T. F., and S. Greenstein. 1996. "Technical Progress and Co-invention in Computing and in the Uses of Computers." *Brookings Papers on Economic Activity: Microeconomics,* 1–83.

Bresnahan, T. F., R. M. Henderson, and S. Greenstein. 2012. "Schumpeterian Competition and Diseconomies of Scope: Illustrations from the Histories of Microsoft and IBM." In *The Rate and Direction of Inventive Activity Revisted,* edited by J. Lerner and S. Stern. National Bureau of Economic Research, Inc.

Bushee, B. J., J. E. Core, W. Guay, and S. J. Hamm. 2010. "The Role of the Business Press as an Information Intermediary." *Journal of Accounting Research* 48, no. 1 (March): 1–19.

Bushman, R. M., C. D. Williams, and R. Wittenberg-Moerman. 2017. "The Informational Role of the Media in Private Lending." *Journal of Accounting Research* 55, no. 1 (March): 115–152.

Carter, D. A., and J. E. McNulty. 2005. "Deregulation, technological change, and the business-lending performance of large and small banks." *Journal of Banking & Finance* 29, no. 5 (May): 1113–1130.

Carter, D. A., J. E. McNulty, and J. A. Verbrugge. 2004. "Do small banks have an advantage in lending? An examination of risk-adjusted yields on business loans at large and small banks." *Journal of Financial Services Research* 25 ([2-3]): 233–252.

Charoenwong, B., Z. T. Kowaleski, A. Kwan, and A. Sutherland. 2022. "RegTech." *SSRN Electronic Journal* (September).

Chen, W., and S. Srinivasan. 2023. "Going digital: implications for firm value and performance." *Review of Accounting Studies* (March): 1–47.

Coval, J. D., and T. J. Moskowitz. 1999. "Home Bias at Home: Local Equity Preference in Domestic Portfolios." *The Journal of Finance* 54, no. 6 (December): 2045–2073.

Das, S. R., and M. Y. Chen. 2007. "Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web." *https://doi.org/10.1287/mnsc.1070.0704* 53, no. 9 (September): 1375–1388.

Deyoung, R., D. Glennon, and P. Nigro. 2019. "Small business lending and social capital: Are rural relationships different?" *The Journal of Entrepreneurial Finance* 22 ([1]): 99–136.

Dyck, A., N. Volchkova, and L. Zingales. 2008. "The Corporate Governance Role of the Media: Evidence from Russia." *The Journal of Finance* 63, no. 3 (June): 1093–1135.

Engelberg, J. E., C. A. Parsons, G. Altı, N. Barberis, S. Dellavigna, M. Fishman, M. Gentzkow, et al. 2011. "The Causal Impact of Media in Financial Markets." *The Journal of Finance* 66, no. 1 (February): 67–97.

Ewens, M., R. Nanda, and M. Rhodes-Kropf. 2018. "Cost of experimentation and the evolution of venture capital." *Journal of Financial Economics* 128, no. 3 (June): 422–442.

Gao, P., C. Lee, and D. Murphy. 2020. "Financing dies in darkness? The impact of newspaper closures on public finance." *Journal of Financial Economics* 135, no. 2 (February): 445–467.

Gentzkow, M., J. M. Shapiro, and M. Sinkinson. 2011. "The Effect of Newspaper Entry and Exit on Electoral Politics." *American Economic Review* 101, no. 7 (December): 2980–3018.

Goldfarb, A., B. Taska, and F. Teodoridis. 2023. "Could machine learning be a general purpose technology? A comparison of emerging technologies using data from online job postings." *Research Policy* 52, no. 1 (January): 104653.

Granja, J., C. Leuz, and R. G. Rajan. 2022. "Going the Extra Mile: Distant Lending and Credit Cycles." *The Journal of Finance* 77, no. 2 (April): 1259–1324.

Heese, J., G. Pérez-Cavazos, and C. D. Peter. 2022. "When the local newspaper leaves town: The effects of local newspaper closures on corporate misconduct." *Journal of Financial Economics* 145, no. 2 (August): 445–463.

Huang, A. H., R. Lehavy, A. Y. Zang, and R. Zheng. 2017. "Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach." *Management Science* 64, no. 6 (June): 2833–2855.

Jagtiani, J., and C. Lemieux. 2019. "The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform." *Financial Management* 48, no. 4 (December): 1009–1029.

Jin, G. Z., M. Leccese, and L. Wagman. 2023. "M&A and Technological Expansion." *NBER Working Paper* (Cambridge, MA) 31126 (April).

Kim, M., D. Stice, H. Stice, and R. M. White. 2021. "Stop the presses! Or wait, we might need them: Firm responses to local newspaper closures and layoffs." *Journal of Corporate Finance* 69 (August): 102035.

Li, F. 2008. "Annual report readability, current earnings, and earnings persistence." *Journal of Accounting and Economics* 45 ([2-3]): 221–247.

Liberti, J. M., and M. A. Petersen. 2019. "Information: Hard and Soft." *The Review of Corporate Finance Studies* 8, no. 1 (March): 1–41.

Lin, M., N. R. Prabhala, and S. Viswanathan. 2012. "Judging Borrowers by the Company They Keep: Friendship Networks and Information Asymmetry in Online Peer-to-Peer Lending." *Management Science* 59, no. 1 (September): 17–35.

Loutskina, E., and P. E. Strahan. 2011. "Informed and Uninformed Investment in Housing: The Downside of Diversification." *The Review of Financial Studies* 24, no. 5 (May): 1447–1480.

Martinez, S., J. S. Lowe, J. Roy, F. Gonzalez, and A. Chapa. 2019. "UBS Report on US Bank Technology Spending." *UBS Global Research.*

Miller, G. S. 2006. "The Press as a Watchdog for Accounting Fraud." *Journal of Accounting Research* 44, no. 5 (December): 1001–1033.

Peirce, H., I. Robinson, and T. Stratmann. 2014. "How are small banks faring under Dodd-Frank?" *Pratt's Bank Law and Regulatory Report* 4 ([1]).

Peress, J. 2014. "The Media and the Diffusion of Information in Financial Markets: Evidence from Newspaper Strikes." *The Journal of Finance* 69, no. 5 (October): 2007–2043.

Petersen, M. A., and R. G. Rajan. 1994. "The benefits of lending relationships: Evidence from small business data." *The Journal of Finance* 49 ([1]): 3–37.

Rajan, R. G. 1992. "Insiders and Outsiders: The Choice between Informed and Arm's-Length Debt." *The Journal of Finance* 47, no. 4 (September): 1367–1400.

Schenone, C. 2010. "Lending Relationships and Information Rents: Do Banks Exploit Their Information Advantages?" *The Review of Financial Studies* 23, no. 3 (March): 1149–1199.

Srivastava, U., and S. Gopalkrishnan. 2015. "Impact of Big Data Analytics on Banking Sector: Learning for Indian Banks." *Procedia Computer Science* 50 (January): 643–652.

Stein, J. C. 2002. "Information Production and Capital Allocation: Decentralized versus Hierarchical Firms." *The Journal of Finance* 57, no. 5 (October): 1891–1921.

Sutherland, A. 2018. "Does credit reporting lead to a decline in relationship lending? Evidence from information sharing technology." *Journal of Accounting and Economics* 66, no. 1 (August): 123–141.

Tetlock, P. C., M. Saar-Tsechansky, and S. Macskassy. 2008. "More Than Words: Quantifying Language to Measure Firms' Fundamentals." *The Journal of Finance* 63, no. 3 (June): 1437–1467.

Appendix A: Variable Definitions

Variable Name	Variable Description
Data Tech - AI Disclosure	An indicator variable equal to one for a matched county pair in our sample if the treated county, defined as a county that experienced local newspaper closures, includes at least one bank that discloses analytics, big data and AI terms in the business description of the annual report or presentation portion of the earnings conference call, zero otherwise. We define the analytics, big data and AI terms using the list of words under the analytics, AI, big data, cloud and machine learning technologies categories in Chen and Srinivasan (2023) (See Appendix B for the detailed description).
Data Tech - AI Job Posting	An indicator variable equal to one for a matched county pair in our sample if the treated county, defined as a county that experienced local newspaper closures, includes at least one bank whose number of AI job postings scaled by the total number of full-time employees are in the upper median of its distribution across banks with AI job postings over 2010-2021, zero otherwise. AI job postings are defined following the BurningGlass skill lists in Alekseeva et al. (2021) as well as the complementary machine-learning skill clusters in Goldfarb et al. (2023) (See Appendix B for the detailed description). For banks in a holding company structure, we compute the AI job posting intensity at the holding company level.
Data Tech - AI Tech M&A	An indicator variable equal to one for a matched county pair in our sample if the treated county, defined as a county that experienced local newspaper closures, includes at least one bank that completed more than one AI technology acquisition in the preceding year (See Appendix B for the detailed description).
Data Tech - Analytics Web Technology	An indicator variable equal to one for a matched county pair in our sample if the treated county, defined as a county that experienced local news paper closures, includes at least one bank that employs analytics technologies on its website, zero otherwise (See Appendix B for the detailed description)

Appendix A: Variable Definitions, Contd.

Data Tech - Combined	An indicator variable equal to one for a matched county pair in our sample if at least one of the the following data technology variables - *Data Tech - AI Disclosure*, *Data Tech - AI Job Posting*, *Data Tech - AI Tech M&A*, *Data Tech - Analytics Web Technology*, and *Data Tech - Data Storage Budget* variable takes the value of one, zero otherwise.
Data Tech - Data Storage Budget	An indicator variable equal to one for a matched county pair in our sample if the treated county, defined as a county that experienced local newspaper closures includes at least one bank whose data storage budget scaled by the total assets is in the upper median of its distribution across banks with non-missing budget data over 2007-2019, zero otherwise (See Appendix B for the detailed description). For banks in a holding company structure, we compute the data storage budget at the holding company level.
Loan Growth	The ratio of the change in loan assets ($\Delta RCFD2122_t$) over the past year to loan assets in the preceding year ($RCFD2122_{t-1}$).
Loan Loss Provision	The ratio of the combined allowance for loan and lease losses (*RCFD3123*) and allocated transfer risk reserves (*RCFD3128*) to loan assets (*RCFD2122*). If RCFD series are missing, we replace with the RCON series.
Local Deposit Concentration	The herfindahl index of *Local Deposit Proportion* for a bank across all counties where the bank operates.
Local Deposit Proportion	The ratio of deposits collected by a bank in each county to the bank's total deposit amounts
Local SBL Lending Concentration	The herfindahl index of *Local SBL Lending Concentration* for a bank across all counties where the bank operates.
Local SBL Lending Proportion	The ratio of small business loans issued by a bank in each county to the bank's total small business loans. Small business loans are business and agriculture loans issued to borrowers with less than 1 million in revenues.
Log(Assets)	The natural logarithm of total assets (*RCFD2170*).

Appendix A: Variable Definitions, Contd.

Net Interest Margin	The ratio of net interest income (*RIAD4074*) to average earning assets. Earning assets are defined as the sum of interest-bearing balances (*RCFD0071*), loan assets (*RCFD2122*), total trading assets (*RCFD3545*), total held-to-maturity securities (*RCFD1754*), total amortized cost of available-for-sale securities (*RCFD1772*), federal funds sold (*RCONB987*), securities purchased under agreement to resell (*RCFDB989*). If RCFD series are missing, we replace with the RCON series.
Newspaper Closures	An indicator variable equal to one if a county experiences a local newspaper closure in prior years and zero otherwise.
NPL-to-Loans	The ratio of the combined past due 90 days loan assets (*RCFD1407*) and non-accrual loan assets (*RCFD1403*) to the total loan assets (*RCFD2122*)
Return-on-Assets	The ratio of the combined non-interest income (*RIAD4079*) and interest income (*RIAD4107*) net of non-interest and interest expense (*RIAD4130*) to the total assets (*RCFD2170*)
SBL M/S	The ratio of small business loans issued by the bank in each county to the total small business loans issued by all banks operating within that county. Small business loans are business and agriculture loans issued to borrowers with less than 1 million in revenues.
Small Business Loans	The ratio of the total small business loans to the total loan assets (*RCFD2122*). Small business loans are business and agriculture loans issued to borrowers with less than 1 million in revenues.

Appendix B: Methodology to Construct The Data-Driven Technology Measures

Appendix B.1: AI Job Posting

We use the BurningGlass data to measure AI job posting and categorize a job posting as AI-related if the job posting contains AI-related skills provided in Alekseeva et al. (2021) and machine learning-related skill clusters in Goldfarb et al. (2023). Specifically, Alekseeva et al. (2021) identify the following skills as AI-related:

Machine Learning, Computer Vision, Machine Vision, Deep Learning, Virtual Agents, Image Recognition, Natural Language Processing, Speech Recognition, Pattern Recognition, Object Recognition, Neural Networks, AI ChatBot, AI KIBIT, ANTLR, Apertium, Artificial Intelligence, Automatic Speech Recognition (ASR), Caffe Deep Learning Framework, Computational Linguistics, Decision Trees, Deeplearning4j, Distinguo, Google Cloud Machine Learning Platform, H2O (software), IBM Watson, IPSoft Amelia, Ithink, Lexalytics, Lexical Acquisition, Lexical Semantics, Machine Translation (MT), Madlib, Microsoft Cognitive Toolkit, MLPACK (C++ library), Mlpy, Modular Audio Recognition Framework (MARF), MoSes, MXNet, Natural Language Toolkit (NLTK), ND4J (software), Nearest Neighbor Algorithm, Object Tracking, OpenNLP, Semantic Driven Subtractive Clustering Method (SDSCM), Semi-Supervised Learning, Sentiment Classification, Supervised Learning (Machine Learning), TensorFlow, Text to Speech (TTS), Tokenization, Torch (Machine Learning), Vowpal, Wabbit, Text Mining, Unsupervised Learning, Image Processing, Mahout, Recommender Systems, Support Vector Machines (SVM), Random Forests, Latent Semantic Analysis, Sentiment Analysis / Opinion Mining, Latent Dirichlet Allocation, Predictive Models, Kernel Methods, Keras, Gradient boosting, OpenCV, Xgboost, Libsvm, Word2Vec, Chatbot, Machine Translation and Sentiment Classification

In addition, Goldfarb et al. (2023) categorize the following skill clusters as machine learning-related :

Machine Learning, Business Intelligence, Big Data, Data Mining, Data Science, Natural Language Processing

Appendix B.2: AI Tech M&A

We employ S&P 451 database of technology M&As to identify AI tech M&A transactions following a three-step procedure. First, we review the business descriptions of the target firms and develop the following seed list of AI terms:

analytics, predictive analytics, big data, ai, big-data, machine learning, cloud, data-based, ai-based, optimization, data management, data mining, market intelligence, business intelligence, BI, correlation, informatics, analysis software, data scientists, *data integration, data analysis, data and content management, natural language processing, database, performance analysis*

Second, we use the co-occurences of words in the business descriptions to identify additional AI terms and add to the AI dictionary list:

ai big data base, algorithmic, algorithms, analytics quantitative, analyzes, apache, automating, bi analytics, big data analytics, bioinformatics, blending reporting, bot, bots, budget forecasting, ci continuous, classification, clustering, cpm, data integration analytic, data management mdm, decisioning, deep learning, devops team, econometric, embedding, etl, evaluation, forecasting, fraud detection, geo location, geostatistical, gi mapping, hadoop, inference, informatica, informatics, information retrieval, jvm, kubernetes, labeling, likelihood, marketo, master data management, mathematical, ml, modal, model, modeling, modelling, nlg, nlp, nosql, optimize, parsing, predictions, predictive, predictive modeling, proprietary, qad, qlik, quantitative, relational, repositories, rpa, segmentation, semantic, source code, statistical, uat, unstructured data, validation, verisk, visualization, visualizations, vrealize

Third, we further include the AI-related skills provided in Alekseeva et al. (2021) to the AI dictionary. We categorize the tech M&A transaction as AI-related if the target firm's business description contains a term in the AI dictionary.

Appendix B.3: Analytics Website Technologies

We use the Builtwith data to identify the analytics website technologies that fall into its following subcategories: Audience Measurement, A/B Testing, Lead Generation, Marketing Automation, CRM, Customer Data Platform, Social Management, Application Performance

Appendix B.4 Data storage budget

We collect data storage budget for banks in our sample from the Aberdeen data. We extract the dollar amounts of data storage budgets reported under the storage category in its IT spend dataset. The storage category comprises of storage devices and the associated management software.

Appendix B.5: AI Disclosure

We measure AI disclosure from the business description of the annual report or presentation portion of the earnings conference call using the regex expressions related to analytics, AI, big data, cloud technologies, and machine learning provided in Chen and Srinivasan (2023) (See Table B1).

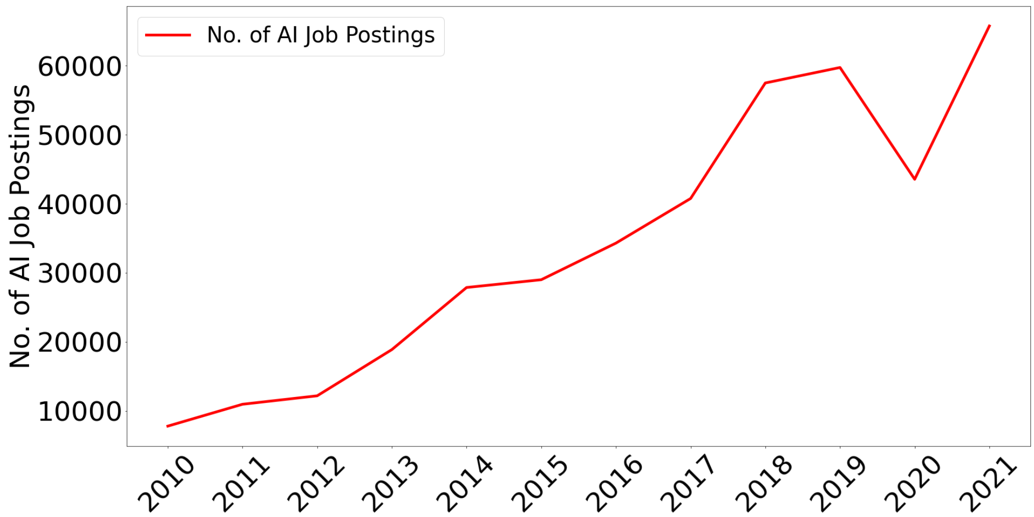
Table B.1: AI Disclosure Regex Expressions

<table>
<tr><th>Digital Term</th><th>Regex Expression</th></tr>
<tr><td>Analytics:</td><td></td></tr>
<tr><td>analytics</td><td>(\banalytics\b)</td></tr>
<tr><td>proprietary algorithm</td><td>(\bproprietary algorithm)</td></tr>
<tr><td>virtual reality</td><td>(\baugmented reality\b)|(\bvirtual realit)</td></tr>
<tr><td>AI:</td><td></td></tr>
<tr><td>artificial intelligence</td><td>(artificial ?[-]?intelligence) | (\bai ?[-]?tech) | (\bai ?[-]?related) | (\bconversational ai\b) | (\bevolutionary ai\b) | (\bevolutionary computing\b)</td></tr>
<tr><td>intelligence</td><td>(\bintelligent ?[-]?system)|(\bcomputer ?[-]?vision)</td></tr>
<tr><td>neural network</td><td>(\bneural ?[-]?network)</td></tr>
<tr><td>virtual assistant</td><td>(\bvirtual agent)|(\bvirtual ?[-]?assistant)</td></tr>
<tr><td>cognitive computing</td><td>(\bcognitive computing\b)</td></tr>
<tr><td>Big Data:</td><td></td></tr>
<tr><td>big data</td><td>(\bbig ?[-]?data)|(\bsmart ?[-]?data)</td></tr>
<tr><td>data science</td><td>(\bdata ?[-]?scien)</td></tr>
<tr><td>data mining</td><td>(\bdata ?[-]?mining)</td></tr>
<tr><td>data lake</td><td>(\bdata lake\b)</td></tr>
<tr><td>devops</td><td>(\bdevops\b)</td></tr>
<tr><td>digital twin</td><td>(\bdigital twin\b)</td></tr>
<tr><td>edge computing</td><td>(\bedge computing\b)</td></tr>
<tr><td>Cloud:</td><td></td></tr>
<tr><td>cloud platforms</td><td>(\bcloud ?[-]?platform) | (\bcloud ?[-]?based) | (\bcloud ?[-]?computing) | (\bcloud ?[-]?deployment)</td></tr>
<tr><td>cloud enablement</td><td>(\bcloud enablement\b) | (\bhybrid cloud\b)</td></tr>
<tr><td>virtual machines</td><td>(\bvirtual ?[-]?machine)</td></tr>
<tr><td>ML:</td><td></td></tr>
<tr><td>biometric</td><td>(\bbiometric)</td></tr>
<tr><td>deep learning</td><td>(\bdeep ?[-]?learning)</td></tr>
<tr><td>machine learning</td><td>(\bmachine ?[-]?learning)</td></tr>
<tr><td>NLP</td><td>(\bnatural ?[-]?language ?[-]?processing)</td></tr>
<tr><td>image recognition</td><td>(\bimage ?[-]?recognition)|(\bfacial ?[-]?recognition)</td></tr>
<tr><td>speech recognition</td><td>(\bspeech ?[-]?recognition)</td></tr>
</table>

Figure 1: AI Job Postings

This Figure presents the annual trend of the number of AI job postings (Panel A) and the median value of the number of AI job posting scaled by the total number of full time employees (Panel B) for banks in our sample. AI job postings are described in the Appendix A and B.

Panel A Number of AI Job Postings



Panel B Median Ratio of the Number of AI Job Posting to Total Employee Counts

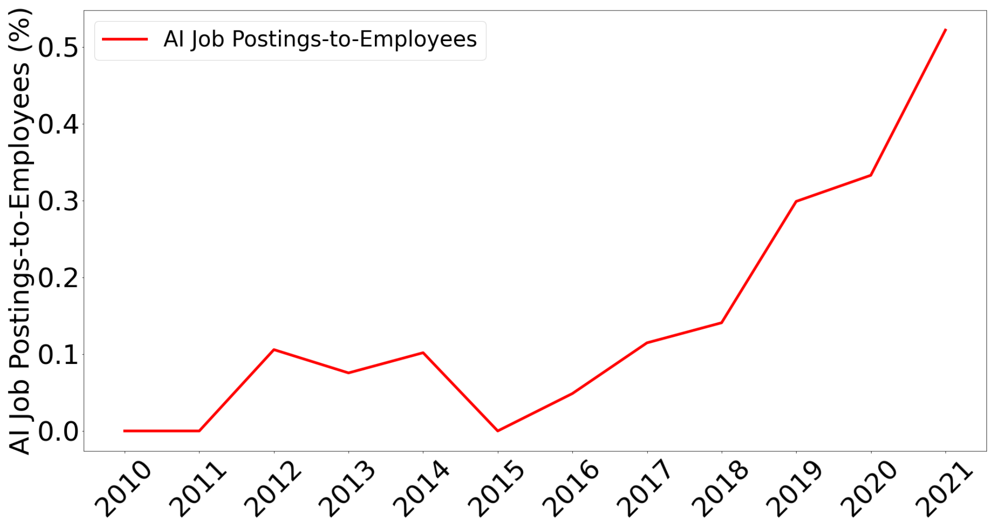


Figure 2: AI-related Technology M&As

This Figure depicts the annual trend of the number of AI related technology M&As for banks in our sample. AI technology M&As are described in the Appendix A and B.

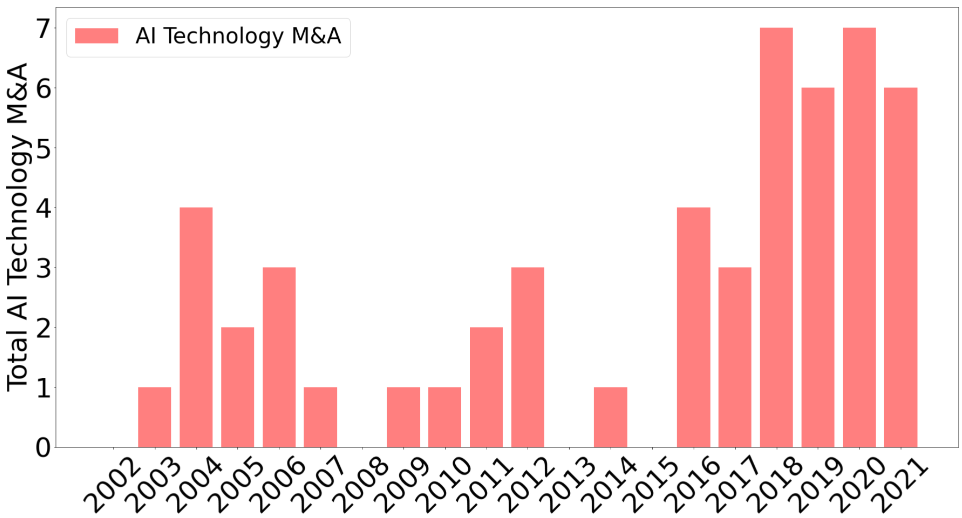
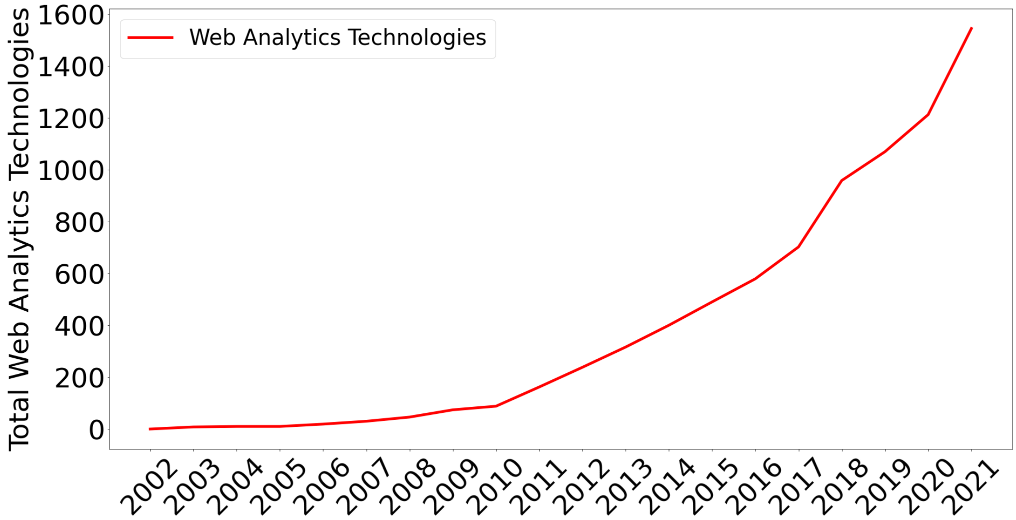


Figure 3: Analytics Web Technologies

This Figure displays the annual trend of the number of analytics web technologies (Panel A) and the median ratio of the number of analytics web technologies to total assets in billions (Panel B) in our sample. Analytics website technologies are described in the Appendix A and B.

Panel A Number of Web Analytics Technologies



Panel B Web Technologies-to-Assets

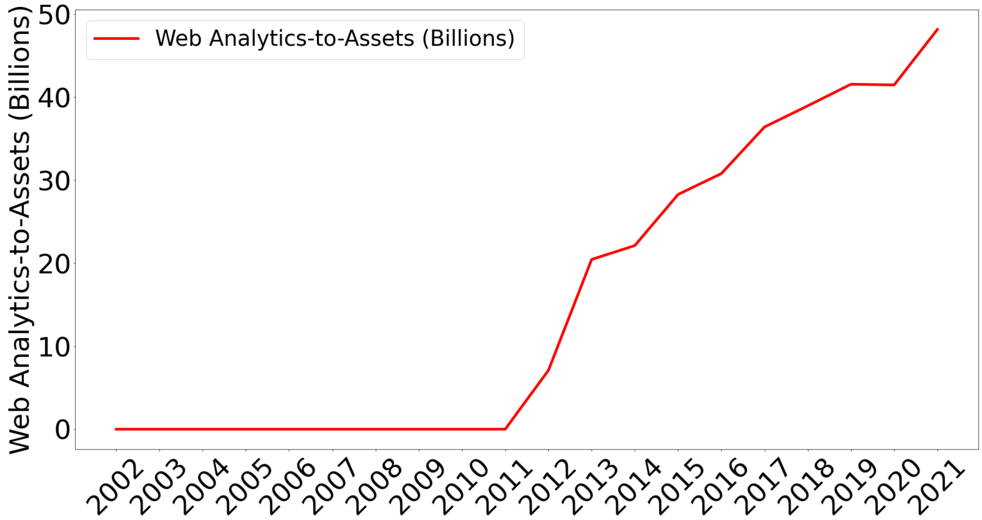


Figure 4: Data Storage Budget

This Figure shows the annual trend of the median value of the data storage budget to total assets for banks in our sample.The data storage budget is described in the Appendix B.

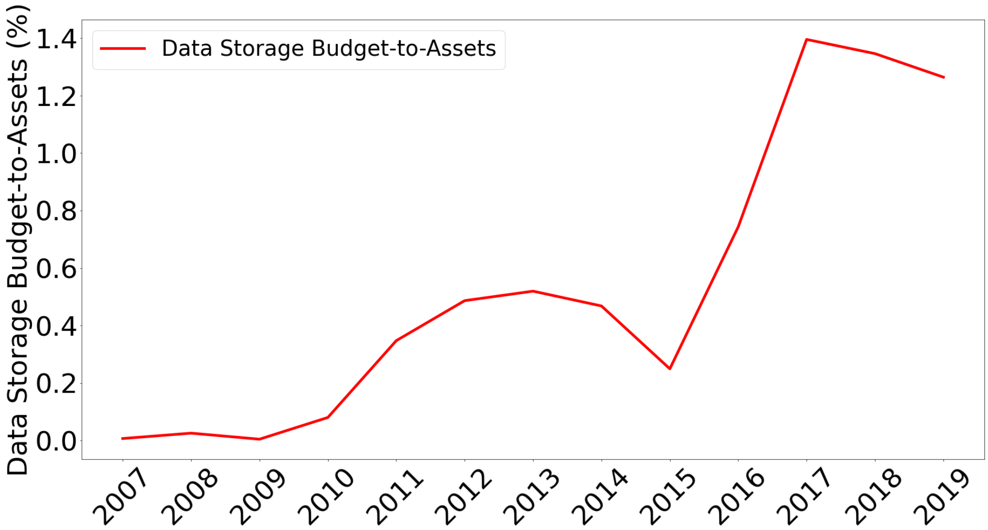
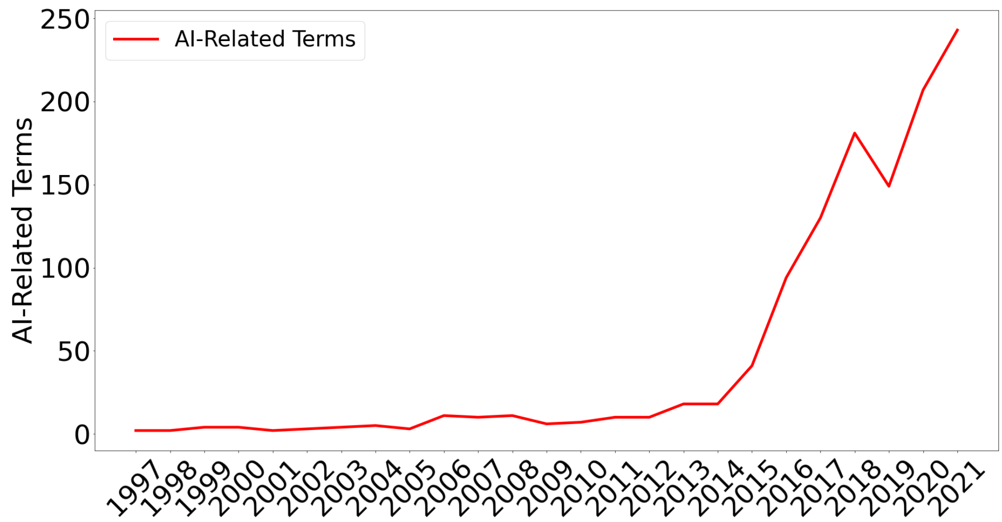


Figure 5: AI-related Disclosures

This Figure illustrate the annual trend of the the number of AI disclosures (Panel A) and the proportion of banks with an AI disclosure. AI disclosures are described in the Appendix A and B.

Panel A Number of Disclosed AI-Related Terms



Panel B Proportion of Banks with AI-Related Terms

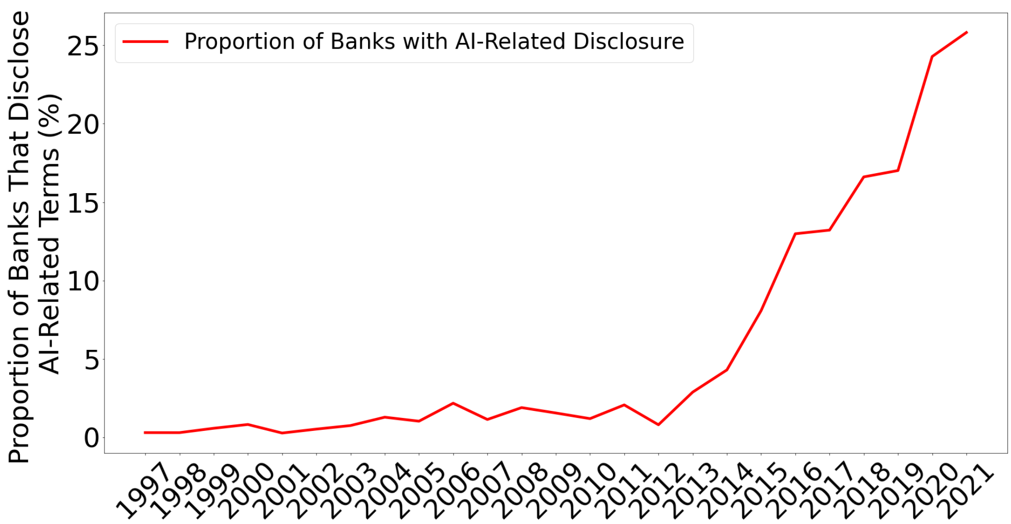


Figure 6: Local Newspaper Closures

This Figure presents the annual trend of the number of local news paper closures during our sample period.

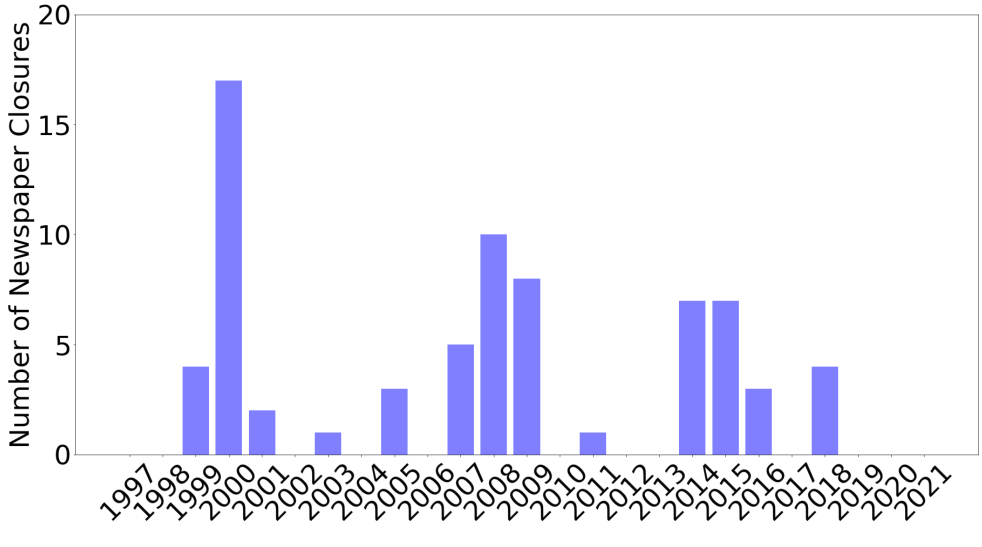


Table 1: Descriptive Statistics

This table reports the summary statistics of the main variables used in our sample. All variables are defined in the Appendix A. In panels B and C, we present the summary statiscs of banks in treatment and control counties, which are defined based on the newspaper closure variable in Appendix A.

Panel A: Full Sample						
Dependent Variable	Mean	SD	Median	25%	75%	N
Bank-Level Variables:						
Total Assets (Billions)	23.352	115.499	2.046	0.829	7.283	6938
Log(Assets)	1.049	1.715	0.716	-0.188	1.986	6938
Loan Loss Provision	0.014	0.007	0.013	0.01	0.016	6938
Net Interest Margin	0.041	0.013	0.039	0.034	0.047	6938
Return-on-Assets	0.014	0.012	0.014	0.005	0.022	6938
NPL-to-Loans	0.012	0.017	0.007	0.004	0.013	6938
Loan Growth	0.115	0.194	0.081	0.023	0.155	6938
Local SBL Lending Concentration	0.36	0.289	0.293	0.118	0.514	6938
Local Deposit Concentration	0.421	0.296	0.362	0.162	0.622	6938
Small Business Loans	0.04	0.042	0.027	0.011	0.054	6938
Small Business Loans (Millions)	163.236	584.325	40.526	15.678	103.362	6938
Bank-County-Level Variables:						
Small Business Loans (Millions)	9.508	24.664	3.61	1.128	9.706	35944
Small Business Loans Market Share (SBL M/S)	0.088	0.124	0.04	0.013	0.105	35944
Local Deposit Proportion	0.097	0.211	0.012	0.002	0.066	35944
Local SBL Lending Proportion	0.098	0.197	0.014	0.002	0.085	35944
Panel B: Banks in Treated County						
Bank-Level Variables:						
Total Assets (Billions)	39.662	151.355	2.6	1.025	13.193	2695
Log(Assets)	1.444	1.913	0.955	0.025	2.58	2695
Loan Loss Provision	0.015	0.007	0.013	0.011	0.017	2695
Net Interest Margin	0.043	0.014	0.041	0.035	0.049	2695
Return-on-Assets	0.017	0.012	0.018	0.009	0.024	2695
NPL-to-Loans	0.012	0.016	0.008	0.004	0.014	2695
Loan Growth	0.134	0.224	0.088	0.026	0.174	2695
Local SBL Lending Concentration	0.304	0.27	0.229	0.085	0.455	2695
Local Deposit Concentration	0.361	0.278	0.286	0.119	0.559	2695
Small Business Loans	0.04	0.041	0.027	0.011	0.055	2695
Small Business Loans (Millions)	284.551	853.799	56.913	20.816	177.089	2695
Bank-County-Level Variables:						
Small Business Loans (Millions)	16.164	46.205	5.191	1.62	14.348	5637
Small Business Loans Market Share (SBL M/S)	0.071	0.101	0.03	0.01	0.086	5637
Local Deposit Proportion	0.138	0.247	0.026	0.004	0.13	5637
Local SBL Lending Proportion	0.14	0.234	0.031	0.005	0.152	5637
Panel C: Banks in Control County						
Bank-Level Variables:						
Total Assets (Billions)	24.399	118.259	2.172	0.878	7.721	6606
Log(Assets)	1.104	1.72	0.776	-0.13	2.044	6606
Loan Loss Provision	0.014	0.007	0.013	0.01	0.016	6606
Net Interest Margin	0.041	0.013	0.039	0.034	0.047	6606
Return-on-Assets	0.014	0.012	0.014	0.005	0.022	6606
NPL-to-Loans	0.012	0.016	0.007	0.004	0.013	6606
Loan Growth	0.115	0.194	0.08	0.023	0.155	6606
Local SBL Lending Concentration	0.348	0.28	0.281	0.113	0.504	6606
Local Deposit Concentration	0.41	0.29	0.352	0.156	0.607	6606
Small Business Loans	0.04	0.042	0.027	0.011	0.053	6606
Small Business Loans (Millions)	170.049	597.972	42.294	16.583	107.807	6606
Bank-County-Level Variables:						
Small Business Loans (Millions)	8.27	17.739	3.378	1.065	9.028	30307
Small Business Loans Market Share (SBL M/S)	0.091	0.128	0.043	0.014	0.109	30307
Local Deposit Proportion	0.09	0.202	0.01	0.001	0.058	30307
Local SBL Lending Proportion	0.091	0.188	0.013	0.002	0.075	30307

Table 2: Local Newspaper Closures and Local Market Competition for Small Business Loans

This table reports the results of the analyses on the effects of local newspaper closures on the local market competition for small business loans. We restrict the sample period up to two years subsequent to each local newspaper closure. We also drop the years 2008 and 2009 to mitigate potential confounding effects of the global financial crisis. In Panel A, we measure a bank's local market concentration using *Local Deposit Proportion* and *Local Deposit Concentration.* In Panel B, we measure a bank's local market concentration using *Local SBL Lending Proportion* and *Local SBL Lending Concentration.* We include matched county group-year and bank-county fixed effects. Standard errors are clustered at the county, bank and year level. ***, **, and * indicate significance at the 0.01, 0.05, and 0.10 levels, respectively. All variables are defined in the Appendix A.

Panel A: Local Deposit-based Concentration

Dependent Variable	SBL M/S	SBL M/S
Concentration Variable	Local Deposit Proportion	Local Deposit Concentration
Newspaper Closure$_{i,t}$×Conc$_{i,t}$	0.019*	0.024**
	(0.011)	(0.011)
Newspaper Closure$_{i,t}$	0.003***	-0.000
	(0.001)	(0.002)
Conc$_{i,t}$	0.090***	-0.028
	(0.019)	(0.017)
Return-on-Assets$_{i,t}$	0.015	0.060
	(0.262)	(0.253)
Log(Assets)$_{i,t}$	0.030***	0.027***
	(0.006)	(0.006)
Loan Loss Provision$_{i,t}$	-0.094	-0.111
	(0.275)	(0.268)
Loan Growth$_{i,t,t-1}$	0.018**	0.017**
	(0.006)	(0.006)
Small Business Loans$_{i,t,t-1}$	0.965***	0.974***
	(0.113)	(0.113)
Matched County Group-Year FE	Yes	Yes
Bank-County FE	Yes	Yes
Observations	35,944	35,944
R^2	0.835	0.834

Panel B: Local SBL-based Concentration

Dependent Variable	SBL M/S	SBL M/S
Concentration Variable	Local SBL Lending Proportion	Local SBL Lending Concentration
Newspaper Closure$_{i,t}$×Conc$_{i,t}$	0.023**	0.032***
	(0.011)	(0.009)
Newspaper Closure$_{i,t}$	0.003***	-0.001
	(0.001)	(0.001)
Conc$_{i,t}$	0.130***	-0.043***
	(0.017)	(0.014)
Return-on-Assets$_{i,t}$	-0.012	0.061
	(0.264)	(0.250)
Log(Assets)$_{i,t}$	0.031***	0.026***
	(0.007)	(0.006)
Loan Loss Provision$_{i,t}$	-0.088	-0.123
	(0.277)	(0.268)
Loan Growth$_{i,t,t-1}$	0.018**	0.017**
	(0.007)	(0.006)
Small Business Loans$_{i,t,t-1}$	0.967***	0.973***
	(0.114)	(0.113)
Matched County Group-Year FE	Yes	Yes
Bank-County FE	Yes	Yes
Observations	35,944	35,944
R^2	0.837	0.834

Table 3: Local Market Competition for Small Business Loans - Before and After the Proliferation of Data-Driven Technology Infrastructure

This table reports the results of analyses on the effects of local newspaper closures on the local market competition for small business loans after segregating the sample into two distinct periods relative to the proliferation of data-driven technology infrastructure. We split the sample into the pre-cloud era before 2008 and the post-cloud era after 2009 which correspond to the widespread availability of cloud computing service. We restrict the sample period up to two years subsequent to each local newspaper closure. We also drop the years 2008 and 2009 to mitigate potential confounding effects of the global financial crisis. In Panel A, we measure a bank's local market concentration using *Local Deposit Proportion* and *Local Deposit Concentration.* In Panel B, we replace Newspaper Closure×Conc variable in the analysis in Panel A with separate interactions between Newspaper Closure and trend variables. The last two-year period preceding the local newspaper closure serves as a benchmark. In Panel C, we measure a bank's local market concentration using *Local SBL Lending Proportion* and *Local SBL Lending Concentration.* In Panel D, we replace Newspaper Closure×Conc variable in the analysis in Panel C with separate interactions between Newspaper Closure and trend variables. The last two-year period preceding the local newspaper closure serves as a benchmark. We include matched county group-year and bank-county fixed effects. Standard errors are clustered at the county, bank and year level. ***, **, and * indicate significance at the 0.01, 0.05, and 0.10 levels, respectively. All variables are defined in the Appendix A.

Panel A: Local Deposit Concentration

Sample	Post-cloud era	Pre-cloud era	Post-cloud era	Pre-cloud era
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local Deposit Proportion		Local Deposit Concentration	
Newspaper Closure$_{i,t}$×Conc$_{i,t}$	-0.010	0.033*	-0.005	0.042***
	(0.007)	(0.015)	(0.008)	(0.013)
Newspaper Closure$_{i,t}$	0.004	0.011*	0.004	0.004
	(0.003)	(0.005)	(0.003)	(0.005)
Conc$_{i,t}$	0.080***	0.108***	-0.037*	-0.025
	(0.019)	(0.021)	(0.020)	(0.022)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	18,586	17,358	18,586	17,358
R^2	0.883	0.895	0.883	0.894
Δ Newspaper Closure$_{i,t}$×Conc$_{i,t}$ βs	-0.042**		-0.047***	
	(0.016)		(0.016)	
Matched County Group-Year FE	Yes		Yes	
Bank-County FE	Yes		Yes	
Controls	Yes		Yes	
Observations	35,944		35,944	
R^2	0.889		0.889	

Panel B: Local Deposit Concentration - Parallel Trend Analysis

Sample	Post-cloud era	Pre-cloud era	Post-cloud era	Pre-cloud era
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local Deposit Proportion		Local Deposit Concentration	
Closure Event$_{i,t-2,t-5}$×Conc$_{i,t}$	0.002	-0.004	0.007	-0.001
	(0.010)	(0.010)	(0.014)	(0.009)
Closure Event$_{i,t,t+2}$×Conc$_{i,t}$	-0.008	0.030*	-0.001	0.042**
	(0.011)	(0.016)	(0.007)	(0.014)
Closure Event$_{i,t-2,t-5}$	0.001	0.001	0.000	0.000
	(0.005)	(0.003)	(0.007)	(0.003)
Closure Event$_{i,t,t+2}$	0.004	0.011*	0.004	0.005
	(0.005)	(0.006)	(0.006)	(0.006)
Conc$_{i,t}$	0.079***	0.109***	-0.038*	-0.025
	(0.019)	(0.021)	(0.020)	(0.022)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	18,586	17,358	18,586	17,358
R^2	0.883	0.895	0.883	0.894

Table 3: Local Market Competition for Small Business Loans - Before and After the Proliferation of Data-Driven Technology Infrastructure (*Continued*)

Panel C: Local SBL-based Concentration				
Sample	Post-cloud era	Pre-cloud era	Post-cloud era	Pre-cloud era
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local SBL Lending Proportion		Local SBL Lending Concentration	
Newspaper Closure$_{i,t}$×Conc$_{i,t}$	-0.002	0.036**	0.003	0.042***
	(0.010)	(0.015)	(0.009)	(0.013)
Newspaper Closure$_{i,t}$	0.004	0.010*	0.002	0.007
	(0.003)	(0.005)	(0.003)	(0.005)
Conc$_{i,t}$	0.135***	0.138***	-0.019	-0.041**
	(0.023)	(0.023)	(0.013)	(0.014)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	18,586	17,358	18,586	17,358
R^2	0.885	0.897	0.883	0.895
Δ Newspaper Closure$_{i,t}$×Conc$_{i,t}$ βs	-0.038**		-0.039**	
	(0.018)		(0.016)	
Matched County Group-Year FE	Yes		Yes	
Bank-County FE	Yes		Yes	
Controls	Yes		Yes	
Observations	35,944		35,944	
R^2	0.892		0.889	

Panel D: Local SBL-Based Concentration - Parallel Trend Analysis				
Sample	Post-cloud era	Pre-cloud era	Post-cloud era	Pre-cloud era
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local SBL Lending Proportion		Local SBL Lending Concentration	
Closure Event$_{i,t-2,t-5}$×Conc$_{i,t}$	-0.012	0.001	0.014	0.004
	(0.013)	(0.011)	(0.012)	(0.011)
Closure Event$_{i,t,t+2}$×Conc$_{i,t}$	-0.010	0.036*	0.012	0.043**
	(0.017)	(0.017)	(0.009)	(0.015)
Closure Event$_{i,t-2,t-5}$	0.002	0.001	-0.001	-0.001
	(0.005)	(0.003)	(0.006)	(0.003)
Closure Event$_{i,t,t+2}$	0.006	0.011*	0.002	0.007
	(0.005)	(0.005)	(0.005)	(0.005)
Conc$_{i,t}$	0.136***	0.138***	-0.020	-0.042**
	(0.023)	(0.023)	(0.014)	(0.014)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	18,586	17,358	18,586	17,358
R^2	0.885	0.897	0.883	0.895

Table 4: Data Driven Technologies and Local Market Competition for Small Business Loans - AI Job Posting

This table examines whether banks with data driven technologies diminish the effect of local newspaper closures on the local market competition for small business loans. We employ AI job posting to identify banks with the data driven technologies and use the variable of *Data Tech - AI Job Posting* to split our sample and perform the analyses. We restrict the sample period up to two years subsequent to each local newspaper closure. We also drop the years 2008 and 2009 to mitigate potential confounding effects of the global financial crisis. In Panel A, we measure a bank's local market concentration using *Local Deposit Proportion* and *Local Deposit Concentration.* In Panel B, we measure a bank's local market concentration using *Local SBL Lending Proportion* and *Local SBL Lending Concentration.* Standard errors are clustered at the county, bank and year level. ***, **, and * indicate significance at the 0.01, 0.05, and 0.10 levels, respectively. All variables are defined in the Appendix A.

Panel A: Local Deposit-based Concentration				
Sample	*Data Tech - AI Job Posting*=1	*Data Tech - AI Job Posting*=0	*Data Tech - AI Job Posting*=1	*Data Tech - AI Job Posting*=0
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local Deposit Proportion		Local Deposit Concentration	
Newspaper Closure$_{i,t}$×Conc$_{i,t}$	-0.003	2.245**	-0.002	0.890**
	(0.005)	(1.014)	(0.007)	(0.320)
Newspaper Closure$_{i,t}$	0.001	-0.038	0.001	-0.077**
	(0.002)	(0.024)	(0.003)	(0.033)
Conc$_{i,t}$	0.080***	-0.116	-0.037*	-0.085
	(0.019)	(0.265)	(0.020)	(0.110)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	17,395	1,191	17,395	1,191
R^2	0.883	0.911	0.883	0.911
Δ Newspaper Closure$_{i,t}$×Conc$_{i,t}$ βs	-2.248***		-0.892***	
	(0.304)		(0.165)	
Matched County Group-Year FE	Yes		Yes	
Bank-County FE	Yes		Yes	
Controls	Yes		Yes	
Observations	18,586		18,586	
R^2	0.896		0.895	

Panel B: Local SBL-based Concentration				
Sample	*Data Tech - AI Job Posting*=1	*Data Tech - AI Job Posting*=0	*Data Tech - AI Job Posting*=1	*Data Tech - AI Job Posting*=0
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local SBL Lending Proportion		Local SBL Lending Concentration	
Newspaper Closure$_{i,t}$×Conc$_{i,t}$	0.004	3.927**	0.005	0.526**
	(0.008)	(1.714)	(0.007)	(0.214)
Newspaper Closure$_{i,t}$	0.002	-0.019	-0.000	-0.049*
	(0.002)	(0.020)	(0.003)	(0.025)
Conc$_{i,t}$	0.124***	0.525***	-0.013	-0.203***
	(0.022)	(0.106)	(0.013)	(0.061)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	17,395	1,191	17,395	1,191
R^2	0.886	0.914	0.883	0.911
Δ Newspaper Closure$_{i,t}$×Conc$_{i,t}$ βs	-3.923***		-0.521***	
	(0.467)		(0.104)	
Matched County Group-Year FE	Yes		Yes	
Bank-County FE	Yes		Yes	
Controls	Yes		Yes	
Observations	18,586		18,586	
R^2	0.898		0.895	

Table 5: Data Driven Technologies and Local Market Competition for Small Business Loans - AI Tech M&A

This table examines whether banks with data driven technologies diminish the effect of local newspaper closures on the local market competition for small business loans. We employ AI tech M&As to identify banks with the data driven technologies and use the variable of *Data Tech - AI Tech M&A* to split our sample and perform the analyses. We restrict the sample period up to two years subsequent to each local newspaper closure. We also drop the years 2008 and 2009 to mitigate potential confounding effects of the global financial crisis. In Panel A, we measure a bank's local market concentration using *Local Deposit Proportion* and *Local Deposit Concentration.* In Panel B, we measure a bank's local market concentration using *Local SBL Lending Proportion* and *Local SBL Lending Concentration.* Standard errors are clustered at the county, bank and year level. ***, **, and * indicate significance at the 0.01, 0.05, and 0.10 levels, respectively. All variables are defined in the Appendix A.

Panel A: Local Deposit-based Concentration

Sample	*Data Tech - AI Tech M&A=1*	*Data Tech - AI Tech M&A=0*	*Data Tech - AI Tech M&A=1*	*Data Tech - AI Tech M&A=0*
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local Deposit Proportion		Local Deposit Concentration	
Newspaper Closure$_{i,t}$×Conc$_{i,t}$	-0.003	0.011	-0.001	0.020***
	(0.005)	(0.008)	(0.007)	(0.006)
Newspaper Closure$_{i,t}$	0.002	-0.001	0.001	-0.004
	(0.002)	(0.003)	(0.002)	(0.003)
Conc$_{i,t}$	0.097***	0.064*	-0.015	-0.099***
	(0.022)	(0.033)	(0.022)	(0.033)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	12,167	17,042	12,167	17,042
R^2	0.867	0.855	0.867	0.856
Δ Newspaper Closure$_{i,t}$×Conc$_{i,t}$ βs	-0.014		-0.021*	
	(0.010)		(0.011)	
Matched County Group-Year FE	Yes		Yes	
Bank-County FE	Yes		Yes	
Controls	Yes		Yes	
Observations	29,209		29,209	
R^2	0.862		0.862	

Panel B: Local SBL-based Concentration

Sample	*Data Tech - AI Tech M&A=1*	*Data Tech - AI Tech M&A=0*	*Data Tech - AI Tech M&A=1*	*Data Tech - AI Tech M&A=0*
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local SBL Lending Proportion		Local SBL Lending Concentration	
Newspaper Closure$_{i,t}$×Conc$_{i,t}$	-0.012	0.023**	-0.002	0.032***
	(0.011)	(0.009)	(0.007)	(0.007)
Newspaper Closure$_{i,t}$	0.003	-0.002	0.001	-0.005*
	(0.002)	(0.003)	(0.002)	(0.003)
Conc$_{i,t}$	0.091***	0.159***	-0.006	-0.071***
	(0.019)	(0.029)	(0.010)	(0.023)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	12,167	17,042	12,167	17,042
R^2	0.869	0.858	0.867	0.856
Δ Newspaper Closure$_{i,t}$×Conc$_{i,t}$ βs	-0.035**		-0.034**	
	(0.014)		(0.012)	
Matched County Group-Year FE	Yes		Yes	
Bank-County FE	Yes		Yes	
Controls	Yes		Yes	
Observations	29,209		29,209	
R^2	0.864		0.862	

Table 6: Data Driven Technologies and Local Market Competition for Small Business Loans - Analytics Web Technology

This table examines whether banks with data driven technologies diminish the effect of local newspaper closures on the local market competition for small business loans. We employ analytics web technologes to identify banks with the data driven technologies and use the variable of *Data Tech - Analytics Web Technology* to split our sample and perform the analyses. We restrict the sample period up to two years subsequent to each local newspaper closure. We also drop the years 2008 and 2009 to mitigate potential confounding effects of the global financial crisis. In Panel A, we measure a bank's local market concentration using *Local Deposit Proportion* and *Local Deposit Concentration*. In Panel B, we measure a bank's local market concentration using *Local SBL Lending Proportion* and *Local SBL Lending Concentration*. Standard errors are clustered at the county, bank and year level. ***, **, and * indicate significance at the 0.01, 0.05, and 0.10 levels, respectively. All variables are defined in the Appendix A.

Panel A: Local Deposit-based Concentration

Sample	*Data Tech - Analytics Web Technology*=1	*Data Tech - Analytics Web Technology*=0	*Data Tech - Analytics Web Technology*=1	*Data Tech - Analytics Web Technology*=0
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local Deposit Proportion		Local Deposit Concentration	
Newspaper Closure$_{i,t}$×Conc$_{i,t}$	-0.005	0.042*	-0.003	0.049***
	(0.004)	(0.019)	(0.005)	(0.009)
Newspaper Closure$_{i,t}$	0.002*	0.011	0.002	0.006
	(0.001)	(0.020)	(0.001)	(0.017)
Conc$_{i,t}$	0.083***	0.083**	-0.008	-0.054
	(0.019)	(0.034)	(0.015)	(0.032)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	23,100	7,867	23,100	7,867
R^2	0.872	0.930	0.872	0.930
Δ Newspaper Closure$_{i,t}$×Conc$_{i,t}$ βs	-0.046**		-0.052***	
	(0.018)		(0.012)	
Matched County Group-Year FE	Yes		Yes	
Bank-County FE	Yes		Yes	
Controls	Yes		Yes	
Observations	30,967		30,967	
R^2	0.891		0.891	

Panel B: Local SBL-based Concentration

Sample	Web Analytics Bank	Rest of Sample	Web Analytics Bank	Rest of Sample
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local SBL Lending Proportion		Local SBL Lending Concentration	
Newspaper Closure$_{i,t}$×Conc$_{i,t}$	-0.007	0.053**	0.003	0.047***
	(0.010)	(0.020)	(0.005)	(0.010)
Newspaper Closure$_{i,t}$	0.002*	0.009	0.000	0.008
	(0.001)	(0.020)	(0.001)	(0.017)
Conc$_{i,t}$	0.124***	0.142***	-0.018	-0.036**
	(0.019)	(0.023)	(0.011)	(0.016)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	23,100	7,867	23,100	7,867
R^2	0.874	0.932	0.872	0.930
Δ Newspaper Closure$_{i,t}$×Conc$_{i,t}$ βs	-0.060**		-0.044***	
	(0.021)		(0.011)	
Matched County Group-Year FE	Yes		Yes	
Bank-County FE	Yes		Yes	
Controls	Yes		Yes	
Observations	30,967		30,967	
R^2	0.893		0.891	

Table 7: Data Driven Technologies and Local Market Competition for Small Business Loans - Data Storage Budget

This table examines whether banks with data driven technologies diminish the effect of local newspaper closures on the local market competition for small business loans. We employ data storage budget to identify banks with the data driven technologies and use the variable of *Data Tech - Data Storage Budget* to split our sample and perform the analyses. We restrict the sample period up to two years subsequent to each local newspaper closure. We also drop the years 2008 and 2009 to mitigate potential confounding effects of the global financial crisis. In Panel A, we measure a bank's local market concentration using *Local Deposit Proportion* and *Local Deposit Concentration*. In Panel B, we measure a bank's local market concentration using *Local SBL Lending Proportion* and *Local SBL Lending Concentration*. Standard errors are clustered at the county, bank and year level. ***, **, and * indicate significance at the 0.01, 0.05, and 0.10 levels, respectively. All variables are defined in the Appendix A.

Panel A: Local Deposit-based Concentration

Sample	*Data Tech - Data Storage Budget*=1	*Data Tech - Data Storage Budget*=0	*Data Tech - Data Storage Budget*=1	*Data Tech - Data Storage Budget*=0
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local Deposit Proportion		Local Deposit Concentration	
Newspaper Closure$_{i,t}$×Conc$_{i,t}$	-0.007	4.137**	-0.006	0.892
	(0.005)	(1.658)	(0.007)	(0.542)
Newspaper Closure$_{i,t}$	0.002	-0.033*	0.002	-0.063
	(0.002)	(0.018)	(0.002)	(0.049)
Conc$_{i,t}$	0.085***	0.107**	-0.033	-0.029
	(0.024)	(0.043)	(0.019)	(0.042)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	14,884	3,116	14,884	3,116
R^2	0.896	0.933	0.896	0.932
Δ Newspaper Closure$_{i,t}$×Conc$_{i,t}$ βs	-4.145**		-0.898	
	(1.658)		(0.540)	
Matched County Group-Year FE	Yes		Yes	
Bank-County FE	Yes		Yes	
Controls	Yes		Yes	
Observations	18,000		18,000	
R^2	0.903		0.903	

Panel B: Local SBL-based Concentration

Sample	*Data Tech - Data Storage Budget*=1	*Data Tech - Data Storage Budget*=0	*Data Tech - Data Storage Budget*=1	*Data Tech - Data Storage Budget*=0
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local SBL Lending Proportion		Local SBL Lending Concentration	
Newspaper Closure$_{i,t}$×Conc$_{i,t}$	-0.001	3.053***	0.002	0.716*
	(0.008)	(0.789)	(0.010)	(0.373)
Newspaper Closure$_{i,t}$	0.003	-0.025	0.001	-0.030
	(0.002)	(0.016)	(0.002)	(0.021)
Conc$_{i,t}$	0.141***	0.113**	-0.024	-0.040
	(0.025)	(0.041)	(0.015)	(0.034)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	14,884	3,116	14,884	3,116
R^2	0.899	0.934	0.896	0.932
Δ Newspaper Closure$_{i,t}$×Conc$_{i,t}$ βs	-3.054***		-0.714*	
	(0.788)		(0.372)	
Matched County Group-Year FE	Yes		Yes	
Bank-County FE	Yes		Yes	
Controls	Yes		Yes	
Observations	18,000		18,000	
R^2	0.905		0.903	

Table 8: Data Driven Technologies and Local Market Competition for Small Business Loans - AI Disclosure

This table examines whether banks with data driven technologies diminish the effect of local newspaper closures on the local market competition for small business loans. We employ AI disclosures to identify banks with the data driven technologies and use the variable of *Data Tech - AI Disclosure* to split our sample and perform the analyses. We restrict the sample period up to two years subsequent to each local newspaper closure. We also drop the years 2008 and 2009 to mitigate potential confounding effects of the global financial crisis. In Panel A, we measure a bank's local market concentration using *Local Deposit Proportion* and *Local Deposit Concentration.* In Panel B, we measure a bank's local market concentration using *Local SBL Lending Proportion* and *Local SBL Lending Concentration.* Standard errors are clustered at the county, bank and year level. ***, **, and * indicate significance at the 0.01, 0.05, and 0.10 levels, respectively. All variables are defined in the Appendix A.

Panel A: Local Deposit-based Concentration

Sample	*Data Tech - AI Disclosure*=1	*Data Tech - AI Disclosure*=0	*Data Tech - AI Disclosure*=1	*Data Tech - AI Disclosure*=0
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local Deposit Proportion		Local Deposit Concentration	
Newspaper Closure$_{i,t}\times$Conc$_{i,t}$	-0.004	0.036***	-0.001	0.040***
	(0.008)	(0.012)	(0.009)	(0.013)
Newspaper Closure$_{i,t}$	0.005***	0.002	0.005*	-0.003
	(0.002)	(0.004)	(0.002)	(0.004)
Conc$_{i,t}$	0.123***	0.080***	0.004	-0.044
	(0.026)	(0.025)	(0.020)	(0.028)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	12,537	23,407	12,537	23,407
R^2	0.879	0.849	0.878	0.849
Δ Newspaper Closure$_{i,t}\times$Conc$_{i,t}$ βs	-0.039***		-0.041**	
	(0.014)		(0.017)	
Matched County Group-Year FE	Yes		Yes	
Bank-County FE	Yes		Yes	
Controls	Yes		Yes	
Observations	35,944		35,944	
R^2	0.858		0.857	

P**Panel B: Local SBL-based Concentration**

Sample	*Data Tech - AI Disclosure*=1	*Data Tech - AI Disclosure*=0	*Data Tech - AI Disclosure*=1	*Data Tech - AI Disclosure*=0
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local SBL Lending Proportion		Local SBL Lending Concentration	
Newspaper Closure$_{i,t}\times$Conc$_{i,t}$	-0.002	0.035***	0.004	0.048***
	(0.015)	(0.012)	(0.009)	(0.010)
Newspaper Closure$_{i,t}$	0.005**	0.002	0.004*	-0.002
	(0.002)	(0.003)	(0.002)	(0.004)
Conc$_{i,t}$	0.131***	0.133***	-0.003	-0.062***
	(0.025)	(0.020)	(0.013)	(0.021)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	12,537	23,407	12,537	23,407
R^2	0.881	0.852	0.878	0.850
Δ Newspaper Closure$_{i,t}\times$Conc$_{i,t}$ βs	-0.037*		-0.044***	
	(0.019)		(0.014)	
Matched County Group-Year FE	Yes		Yes	
Bank-County FE	Yes		Yes	
Controls	Yes		Yes	
Observations	35,944		35,944	
R^2	0.860		0.858	

Table 9: Data Driven Technologies and Local Market Competition for Small Business Loans

This table examines whether banks with data driven technologies diminish the effect of local newspaper closures on the local market competition for small business loans. We use the variable of *Data Tech - Combined* to split our sample and perform the analyses. We restrict the sample period up to two years subsequent to each local newspaper closure. We also drop the years 2008 and 2009 to mitigate potential confounding effects of the global financial crisis. In Panel A, we measure a bank's local market concentration using *Local Deposit Proportion* and *Local Deposit Concentration*. In Panel B, we measure a bank's local market concentration using *Local SBL Lending Proportion* and *Local SBL Lending Concentration*. Standard errors are clustered at the county, bank and year level. ***, **, and * indicate significance at the 0.01, 0.05, and 0.10 levels, respectively. All variables are defined in the Appendix A.

Panel A: Local Deposit Concentration

Sample	*Data Tech - Combined*=1	*Data Tech - Combined*=0	*Data Tech - Combined*=1	*Data Tech - Combined*=0
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local Deposit Proportion		Local Deposit Concentration	
Newspaper Closure$_{i,t}$×Conc$_{i,t}$	-0.005	0.048**	-0.002	0.043**
	(0.004)	(0.019)	(0.005)	(0.016)
Newspaper Closure$_{i,t}$	0.002*	0.013	0.001	0.009
	(0.001)	(0.007)	(0.001)	(0.006)
Conc$_{i,t}$	0.085***	0.110***	-0.007	-0.082***
	(0.018)	(0.022)	(0.014)	(0.019)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	23,992	11,952	23,992	11,952
R^2	0.871	0.902	0.870	0.902
Δ Newspaper Closure$_{i,t}$×Conc$_{i,t}$ βs	-0.052**		-0.045**	
	(0.019)		(0.017)	
Matched County Group-Year FE	Yes		Yes	
Bank-County FE	Yes		Yes	
Controls	Yes		Yes	
Observations	35,944		35,944	
R^2	0.884		0.884	

Panel B: Local SBL-based Concentration

Sample	*Data Tech - Combined*=1	*Data Tech - Combined*=0	*Data Tech - Combined*=1	*Data Tech - Combined*=0
Dependent Variable	SBL M/S	SBL M/S	SBL M/S	SBL M/S
Concentration Variable	Local SBL Lending Proportion		Local SBL Lending Concentration	
Newspaper Closure$_{i,t}$×Conc$_{i,t}$	-0.008	0.050**	0.004	0.048**
	(0.010)	(0.018)	(0.006)	(0.016)
Newspaper Closure$_{i,t}$	0.002*	0.014*	0.000	0.010
	(0.001)	(0.007)	(0.001)	(0.006)
Conc$_{i,t}$	0.123***	0.164***	-0.017	-0.065***
	(0.018)	(0.022)	(0.011)	(0.017)
Matched County Group-Year FE	Yes	Yes	Yes	Yes
Bank-County FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	23,992	11,952	23,992	11,952
R^2	0.873	0.905	0.870	0.902
Δ Newspaper Closure$_{i,t}$×Conc$_{i,t}$ βs	-0.058***		-0.044**	
	(0.020)		(0.017)	
Matched County Group-Year FE	Yes		Yes	
Bank-County FE	Yes		Yes	
Controls	Yes		Yes	
Observations	35,944		35,944	
R^2	0.887		0.884	