THESIS

FEELING THE SQUEEZE: LOOKING FOR LEMONS IN THE MARKET FOR USED BANKS

Submitted by

Gregory Totten

Department of Economics

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2024

Master's Committee:

Advisor: Stephan Weiler

Dawn Thilmany
Robert Kling

Copyright by Gregory Totten 2024

All Rights Reserved

ABSTRACT

FEELING THE SQUEEZE: LOOKING FOR LEMONS IN THE MARKET FOR USED BANKS

This study examines adverse selection and geographic information asymmetry costs stemming from branch network expansion on bank performance. Using a comprehensive dataset covering the period from 1994 to 2022 from the FDIC, this paper investigates how geographic expansion and resulting merger and acquisition (M&A) activity impact bank performance, as measured by Return on Assets (ROA), Return on Equity (ROE), and Non-Performing Loans. The findings reveal that while increasing bank size generally (as measured by total assets) yields benefits, likely due to operational efficiencies, these are accompanied by offsetting costs due to both local information losses resulting from increased geographic branch network size and adverse selection ("lemons") effects stemming from the M&A process. These costs are particularly acute during economic downturns, as evidenced by the interaction effects observed during the Great Recession. The results suggest that information asymmetries, resulting from geographic distance and a lemons effect, are significant in determining bank performance outcomes. This study contributes to the existing literature by reconciling previously mixed findings regarding the net benefits of bank branch network expansion, highlighting the importance of considering both the benefits and costs associated with these strategies.

ACKNOWLEDGEMENTS

I would like to thank my advisor Stephan Weiler, for his many years of guidance, support, and endless patience throughout this process. Additionally, I would also like to thank all of the wonderful faculty and staff in both the Economics Department, and the Agricultural Economics Department for providing the foundation for my career, and especially Dawn Thilmany and Robert Kling for their contributions of feedback and support as members of my committee. Finally, I would like to thank my friends and family, and especially my lovely wife, Mieke, for allowing me to spend so many weekends deeply engrossed in this research.

TABLE OF CONTENTS

ABSTRACT i
ACKNOWLEDGEMENTS ii
LIST OF TABLES iv
LIST OF FIGURES v
1 Introduction 1
1.1 Paper Structure 2
2 History of 20th Century Banking Regulation 3
2.1 Pre-1980 3
2.2 Deregulatory Period – 1980s and 1990s 4
2.3 Implications for the Banking System 4
3 Literature Review 5
3.1 Motivations for Expansion 5
3.2 Expansion Costs and Asymmetric Information 12
3.3 Benefits, Costs, and Banking Performance 18
3.4 Geography Matters 22
4 Theoretical Model 22
4.1 Motivation 26
5 Methodology 27
5.1 Data Sources 27
5.2 Data Description 27
5.3 Data Preparation 29
5.4 Model Specification 33
5.5 Strategy 36
6 Results 38
6.1 Descriptive Statistics 38
6.2 Main Results 38
6.3 Statistical Checks 45
7 Discussion 51
7.1 Interpretation of Results 51
7.2 Implications 54
7.3 Limitations and Further Analysis 54
8 Conclusion 55
8.1 Summary 55
9 Appendices 56
9.1 Outliers 56

LIST OF TABLES

1 Summary Statistics 39
2 Regression Results 41
3 Model Statistics 45
4 Variance Inflation Factor - Return on Assets(%) 47
5 Variance Inflation Factor - Return on Equity(%) 48
6 Variance Inflation Factor - Non-Current Assets(% of Total) 49
7 Results Interpretation - Overall 52
8 Results Interpretation - Recession 53
9 Summary Statistics of Dependent Variables 57
10 Summary Statistics of Dependent Variables with Outliers Removed 59
11 Regression Results comparing models with and without outliers 59

LIST OF FIGURES

1 Correlation Matrix 50
2 Outliers 57

1 Introduction

When one examines the past three decades of banking in the United States, consolidation of the industry has long been a dominant theme. In 1984, there were a total of 14,469 FDIC-insured commercial banks in the United States, with 40,853 Total Branches[1]. By 2022, the total number of FDIC-insured commercial banks had fallen to 4,136, while branches had increased to 69,905. This meant the average bank size went from 2.8 branches per bank in 1983, to 16.9 branches per bank in 2022 - a nearly sixfold increase.

The primary process by which this consolidation occurred was via mergers and acquisitions. That is, it was not only that some institutions grew organically (though some did), and the competition from that growth led to the destruction of existing institutions. Rather, much of the consolidation occurred through a Merger and Acquisition ("M&A") process, whereby institutions merged with or acquired[2] (often repeatedly) other institutions to grow their scale.

An important aspect of this process is that fundamentally the M&A process is a transaction, whereby a buyer determines a price they are willing to pay for the business, and the seller must determine a price which fairly compensates the owners and investors of the existing bank. In such a market the price is determined based on the Net Present Value of future returns of the business, discounted back to the current moment. In such a circumstance, a transaction will occur either when the buyer accrues excess benefits that they are able to achieve which the seller cannot, or when the seller is aware of additional costs or risks that the buyer is not. The latter is a classic example of the "lemons" problem, as first described by Akerlof (1970), which appears in many transactions of which an acquisition is one such example.

[1] FDIC BankFind Suite: Annual Historical Bank Data

[2] In the context of this paper, I use the words merger and acquisition interchangeably, as both are included in the business combinations data available from the FDIC

The existing literature, as described in more detail in Section 3 provides ample evidence of the existence of both excess benefits and countervailing costs which may end up accruing to the buyer. While a buyer may be able to estimate their benefits, many of the costs may exist in a form that is apparent to the seller, but which may be difficult for the buyer to ascertain. This dynamic leads to a classic information asymmetry The presence of sufficient benefits or a transaction to occur, either because the seller is able to extract excess value, or because their additional costs apparent only to the seller that cause them to reduce their own valuation.

This paper then applies a framework based on this transactional model (whereby buyer and seller must agree on a price) in the market for "used" banks, to identify the presence of both the benefits accruing to the buyer, as well as an information asymmetry in which there are "lemons" known only to the seller, to establish the effect of each on the acquiring bank's performance.

1.1 Paper Structure

The paper is organized in the following sections:

- Section 2 contains a brief history of regulatory changes impacting the US banking system over the century.
- Section 3 provides a literature review of the existing studies on M&A, highlighting the main theoretical perspectives, empirical methods, and findings.
- Section 4 presents the theoretical motivation for this study based on the existing literature.
- Section 5 includes the data and the econometric model used to test the hypotheses derived from theory.

- Section 6 reports and discusses the results of the estimation.
- Section 7 provides a discussion of the results, their implications, and limitations of the study, in the context of the existing literature and theory.
- Section 8 summarizes the major findings of the research, and remaining opportunities for continued investigation.
- Section 9 includes additional analysis on outliers present in the data.

2 History of 20th Century Banking Regulation

2.1 Pre-1980

Starting with the Civil War, through the passage of the McFadden Act of 1927 ("McFadden Act"), banks operated under either a national charter (national banks) where they were allowed a single building or under state charters (state banks) where various state laws dictated the number of branches that would be allowed. The McFadden Act changed these rules, allowing national banks to open branches in individual states but limiting the number of branches allowed to the number of branches allowed under a state charter – meaning that if a state did not allow branch banking, then national banks could not operate there, and therefore the extent to which a national bank can operate within a state is dictated by the state branching laws ("McFadden Act of 1927 | Federal Reserve History" 2013). Combined with later amendments, the McFadden Act effectively prevented interstate branching and gave states ultimate authority in managing their branch systems ("Going Interstate: A New Dawn For U.S. Banking | St. Louis Fed" 1994).

Additionally, the Bank Holding Company Act of 1956 prevented, at a national level, a "bank holding company (BHC) from acquiring an out-of-state bank subsidiary unless the home state of the acquired bank has a statute authorizing such an acquisition" ("Going Interstate: A New Dawn For U.S. Banking | St. Louis Fed" 1994).

2.2 Deregulatory Period – 1980s and 1990s

Beginning in the 1980s, individual states began relaxing policies on interstate branching, with 46 states allowing out-of-state BHCs to acquire in-state banks. However, these were on a state-by-state basis, and these restrictions varied in their approaches and rules. It was not until the passage of the Riegle-Neal Interstate Banking Act of 1994 ("Riegle-Neal") that national rules were created explicitly allowing for interstate branching at the federal level, allowing "well-managed" and "well-capitalized" banks to acquire banks in any state beginning in September 1995, and further allowing BHCs to merge branches in multiple states to create a single branch network beginning in June 1997("Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994 | Federal Reserve History" 1994).

With restrictions on interstate banking lifted via Riegle-Neal, one last major deregulatory act, the Gramm-Leach Bliley Act and Financial Integration ("GLBA") was passed in 1999. This act repealed parts of The Banking Act of 1933 ("Glass-Steagall Act"), which had placed several restrictions on deposit institutions, including the separation of commercial and investment banking activities. The removal of these restrictions reflected the last major act of the deregulatory banking period beginning in the 1980s and ending at the end of the 1990s ("Financial Services Modernization Act of 1999 (Gramm-Leach-Bliley) | Federal Reserve History" 2013).

2.3 Implications for the Banking System

As discussed in the Introduction, deregulation coincided with a major consolidation of the US banking system, with far fewer banks owning far more branches than in 1980 before this process began. Two results of this process were that the system of banks was inhabited by fewer banks that were both far larger, in terms of not only branches and assets, but also geographic scope, and the majority of which had undergone a significant number of acquisitions in their growth process. This paper examines the effects of both of those outcomes on bank performance.

3 Literature Review

While the 1980s and 1990s deregulatory trends provide the context of the changing policy environment that allowed for this consolidation in the banking sector, this consolidation would not have occurred absent the underlying economic motivations influencing banks to undergo this consolidation. This section provides a review of the literature examining the theoretical motivations and impacts of this consolidation cycle, as well as the empirical evidence examining the real-world outcomes resulting from it.

3.1 Motivations for Expansion

The primary motivating factor driving bank expansion, and with it M&A activity, is increasing shareholder returns, through either increased market, and therefore pricing, power, or from efficiency improvements (Allen N. Berger, Demsetz, and Strahan 1999). For the purposes of this paper, we will further break the latter into two categories - organizational efficiency and portfolio efficiency, both of which will be described in further detail below.

3.1.1 Market Pricing Power

The academic literature provides limited evidence for the thesis of increased market pricing power. This is likely because most studies suggest that a primary impact of the changes brought about by Riegle-Neal resulted in increased overall competition in the banking sector as banks were able to expand into greater numbers of markets.

One way to assess the presence of market pricing power is via the Herfindahl-Hirschman Index ("HHI"), a measure of market concentration used by the Department of Justice, with higher levels indicating a higher degree of market consolidation. Moore and Siems (1998) find that while nationwide banking HHI levels increased significantly from 1987 to 1997, from under 100 to 175, the values are quite low and represent a competitive overall market. Looking at the average across individual markets, they find little change in concentration over the same period. Additionally, from a practical standpoint, they find that while institutions

in heavily concentrated markets appear to have been able to leverage their market power into monopoly profits (as indicated by significantly higher ROA relative to those in less concentrated markets), by 1997, this advantage appears to have dissipated, with banks earning similar ROA regardless of market concentration. Rhoades (2000) find similar results, with HHI levels remaining relatively constant for banks in MSAs between 1980 and 1998 and a slight decrease in HHI levels for banks in non-MSA areas during the same period, and also note that overall bank profitability increased significantly for all banks during this period.

Examining a slightly different period, Giedeman (2004) finds that between 1994 and 2002, HHI levels and thus the overall market concentration of deposits decreased across the 743 cities studied. Neither a city's rural classification nor a state's laws regarding *de novo* branching had a significant impact on the change in market concentration.

Moving forward to the larger period of 2000 to 2020, Calzada, Fageda, and Martínez-Santos (2023) examine determinants of branch density at the Census Tract level, and find that tracts that have been exposed to bank merger activity generally have lower branch density - suggesting lower levels of bank competition. This finding on lower levels of competition in areas with merger activity is also supported by the negative correlation between branch density and HHI - indicating that areas with relatively lower levels of branch density have higher HHI levels, indicating lower levels of competition as well. They also test this in two periods - pre and post-Great Recession periods, and find similar effects, however, with a somewhat larger magnitude of effect in the post-Great Recession period.

Finally, Allen N. Berger et al. (1997) researched the impact of M&A activity on bank small business lending, with a dataset covering 6,000 M&As involving 10,000 banks from the late 1970s to the early 1990s, including data from both the Federal Reserve's Survey of the Terms of Bank Lending to Businesses and the FDIC June Call Reports. They find that there are two types of effects, with different impacts on small business lending in a market. They find that in the short run, banks often reduce the amount of small business lending performed

immediately after a merger, which could be related to a reluctance to lend due to reductions in soft information available to the institution. This perspective could also support the theory that management at the acquiring bank may be able to exert some market power evidenced by reduction in lending (directly or indirectly through more stringent criteria). Over time, however, they find banks tend to restart lending, offsetting any reductions through two main effects: a restructuring effect and increased competition from other firms in the market. The latter effect relates to what they deem the external effect, which is a reaction by other local banks to increase lending to compensate for the initial firm's reduction. The final net effect is that, overall in the merger-influence market, they find total lending increases. This process detailed in their analysis also again potentially indicates the competing dynamics of intended market power gains intended by the acquiring bank being offset by the overall increased market competition dynamics resulting from increased competition in the wake of Riegle-Neal.

3.1.2 Operational Efficiency

In defining operational efficiency, this research primarily focuses on two forms. The first is that which is most commonly associated with traditional benefits of scale, such as sharing back-office services or eliminating operational overlap. The second is related to benefits resulting from improved managerial practices typically resulting from a more efficiently managed firm acquiring a firm which is less so.

Several studies examining post-acquisition performance find benefits resulting from gains due to better cost efficiency.

Houston, James, and Ryngaert (2001) find that increasing profitably post-merger by focusing on cost savings, rather than through enhancing revenue, is generally a more productive M&A strategy. In their study of a series of 64 large bank mergers occurring during the period of 1985 to 1996, using a combination of data sources, the authors estimate the benefits accruing because of the merger, as well as examine the post-merger performance, including return on

assets, of the combined company. Additionally, in 41 of the transactions, they were able to look at management estimates of future cost savings and revenue enhancements stemming from the merger to determine the source of merger benefits. The authors find that while the actual performance of the combined company tends to fall short of management estimates, there is an overall benefit accruing to the combined company, with most of that benefit derived from cost savings. Additionally, the authors categorize transactions based on a focus of either geographic expansion of the combined merger company, or a focus on market-overlap, with the latter concentrated on cost savings accruing through efficiency gained by eliminating redundant operations. They find that the latter type – those focused on cost savings – tended to provide higher abnormal returns, stemming from more consistent cost savings, than those focused on geographic expansion.

Becher and Campbell II (2005) use a similar methodological approach, but with differing outcomes suggesting an expansive geographic approach leads to more favorable outcomes in the post Riegle-Neal era. They examine market reactions to transactions focused on cost savings vs those on geographic expansion (or diversification, as referred to in this context), and instead find evidence that the market rewards geographic diversification in M&A activity. In particular, the authors use an event study analysis to examine how merger announcements were digested by markets across pre and post-deregulation periods during the 1990s. Confirming the findings of Houston, James, and Ryngaert (2001), the authors again find significant benefits accruing to mergers centered on cost savings from high geographic overlap - indicating a focus on reducing operational redundancy - in the 1990s. Conversely, the authors find that mergers occurring in the post-deregulation period of the late 1990s failed to create value overall, and that in particular those with a high degree of geographic overlap actually saw significant losses in this period. The authors suggest that changes brought about by deregulation carried with them changes in market perception, that now saw benefits to geographic expansion as being more beneficial than potential cost savings.

Focusing more on managerial outcomes, Allen N. Berger and DeYoung (2001) examine over 7,000 banks during the period of 1993 to 1998 using FDIC Call Data. They find that parent banks can exercise some efficiency control over their affiliates, but that control dissipates with distance between the parent and affiliate. However, they do find that these effects tend to, on average, be relatively small, suggesting that some of these distance effects can be negated by the parent institution.

Stiroh and Strahan (2002), focusing on the regulatory impacts of competition in the US Banking Industry, also finds benefits accruing to better managed firms. The authors use FDIC call reports from 1976 to 1994 to examine bank performance, as measured by ROE, of banks. Specifically, they use what they call normalized return on equity, which is the ROE of a bank minus the ROE of other banks in the same state and year. They then use changes in bank performance in the panel data set to see how banks respond to changes in banking regulations. They find that bank performance is linked to market share growth, and that this link strengthens with deregulation, suggesting that increased competition further rewards better managed, and therefore better performing, banks. They also find that exit rates increased overall in markets after deregulation, and that higher performing banks exited at a higher rate as they were more attractive acquisition targets in a deregulated environment - again demonstrating the market preference for more well-managed institutions.

Evanoff and Ors (2008) examine the impact of deregulation on cost efficiency at the industry level, in the form of increased competition, resulting from the entry of new participants in the form of acquiring banks taking over existing competing local banks. The authors find that banks respond to this new entry by becoming more cost efficient, with improved accounting ratios, stemming from a more efficient cost structure. They also find some evidence of customer migration in the short run away from the acquired bank branches, and into existing banks, but that this migratory behavior only lasts approximately two years. Overall, they find that even though there may not be cost efficiency benefits to the acquiring bank, as

found by Allen N. Berger and DeYoung (2001), there are overall market efficiency gains from the increased competition.

3.1.3 Portfolio Efficiency

Efficiency can also be increased through portfolio diversification, following in the tradition of modern portfolio theory as originally proposed in Markowitz (1952). The basic premise is that by allocating assets across a variety of different types of assets, the overall risk of the portfolio can be reduced, while maintaining the same level of return. This is because to the extent that default risk is uncorrelated, the returns lost by a default on one loan can be offset by the returns gained by another loan. In the traditional sense, one can think about this in terms of a bank's overall portfolio of loans, with the interest earned from some borrowers more than offsetting the losses from others, generating an overall positive return over time. Similarly, one can apply this same logic to geographic diversification - to the extent regional economies are economically distinct, a relative downturn in one region, leading to either increased defaults in that region, or simply a decreased demand for loans, can be offset by increased demand in another region with an economic sector performing better in prevailing market conditions..

The portfolio effects of geographic diversification also have implications for lending risk and firm performance and valuation.

Emmons, Gilbert, and Yeager (2004) examine community banks who often have small loan portfolios that are more concentrated in both idiosyncratic and geographic terms. Using merger simulations, they assess the role that mergers would have in reducing risk through both channels and find that the greatest reduction is in idiosyncratic risk from increasing a bank's size. They do note that urban banks have more geographic risk than rural banks, likely due to the more heterogeneous nature of urban economies.

Deng and Elyasiani (2008) find that increased geographic diversification is associated with reduced levels of risk and increases to values of BHC value. They employ measures of geographic dispersion and distance between headquarters and branches to assess the benefits of dispersion. The results suggest that there is a significant benefit to geographic diversification in terms of risk and value as banks can spread localized geographic risks across more markets.

A series of papers, beginning with M. Goetz, Laeven, and Levine (2011), the field examines the impacts of increasing geographic portfolio diversity. M. Goetz, Laeven, and Levine (2011) employs a state-specific, time-series pattern around deregulation of banking restrictions, find that increasing geographic diversity instead reduces BHC valuations. They suggest the costs of increased complexity, and the ability of insiders to extract private rents due to less transparency to outside investors, reduce the overall values of these more geographically diverse companies. Martin Richard Goetz (2012) looks at the relationship between risk and geographic dispersion both in terms of a bank's own behavior, and that of its competitors. The author finds that as banks own geographic dispersion increases, so does their propensity to take on more risk, consistent with expectations under portfolio behavior as they have a more diversified risk portfolio. They also find, however, that competing banks in the area respond with a lower risk portfolio. The theoretical framework for this is that that the more geographically diversified bank, due to its larger geographic footprint and organization structure, is more necessarily reliant on hard information for lending decisions. This leads other banks in the area to specialize more and focus more on their competitive advantage in soft information, reducing their overall risk level. This is also consistent with Petersen and Rajan (2002) findings that distant firms are more likely to take on riskier small business borrowers. Finally, Martin R. Goetz, Laeven, and Levine (2016), again looking at patterns around state deregulation, finding that geographic expansion does, in fact, reduce overall portfolio risk for the BHC, though in this case results are inconclusive with respect to changes in loan quality.

Brealey, Cooper, and Kaplanis (2019) also examine the impact of merger activity on risk. Using a sample of US bank mergers spanning 1981 to 2014 the authors find that bank merger activity had no long term impact on equity and asset risk of the merging banks. The only significant change in risk found by the authors was instead that of systemic risk - which the authors find to generally occur with early acquisitions in the cycle for acquiring banks, with changes generally tapering off after the fifth acquisition.

Meslier et al. (2016) look at the tradeoffs inherent in geographic dispersion, weighing the benefits of reduced risk from more geographic exposure with the increased costs of being more thinly spread, looking at data from the period of 1994 to 2008. They find that overall geographic diversification benefits risk-adjusted returns, but with mixed results. They find that for banks with low levels of diversification, spreading across counties or states can benefit their risk adjusted returns, but only to a point, as eventually these returns become negative once geographic diversification overexpands. Generally, they find that for all small banks, with less than $1B in assets, there is a benefit from increasing geographic diversification, but that for medium and large banks above this level, no potential benefit still exists.

Jang (2019) looks at overall efficiency in the market for bank branches and finds a generally efficient market with large inefficient banks selling to smaller more efficient banks in order to raise capital. Using an event study approach the author finds significant effects on value for both parties to the transaction. The buyer in particular is able to generate these effects through improved operating performance and faster loan growth without a corresponding significant decrease in performance or loan rates for the seller.

3.2 Expansion Costs and Asymmetric Information

While a number of studies provide both a theoretical basis for a bank's idiosyncratic and geographic expansion, such a process does not come without costs.

Here we explore two particular types of costs accruing to expanding banks - one being an example of the "lemon" problem, and the other stemming from losses associated with geographic information asymmetries. These two costs are different in nature but both are related to the problem of asymmetric information.

3.2.1 The Lemon Problem

The idea of asymmetric information, and specifically that of adverse selection, is first formalized in Akerlof (1970). The first theoretical model proposed in the paper focuses on the used car market, defined as having four types of cars - new and used, each of which can be of either good or bad quality (the latter of which is referred to as a "lemon"). In this market good quality cars are of higher value than bad quality cars.

In the market for a new car neither party knows whether or not a car is, in fact, a lemon. In this context, we can reasonably assume that both parties have the same, or symmetric, information in that each car being sold has the same probability of being a lemon.

In the used car market, however, an asymmetry of information develops. In this context the seller of the car has had the opportunity to own the vehicle, spent time and money driving and repairing it, and therefore has been able to accumulate information about its quality that allows them to make a more accurate determination on the lemon status of the vehicle. The buyer, on the other hand, does not have all of this same information, and so must assume that each vehicle is of average quality. The asymmetry in this case is represented by this informational difference that has now arisen between the buyer and seller, as a result of the differing levels of information possessed by each party, and by the willingness of the seller, in particular, to disclose all information they possess.

3.2.2 Lemons in M&A

In contemplating the car market theorized by Akerlof (1970), it is straightforward to envision a similar dynamic emerging in the market for bank acquisitions. In this dynamic you again

have a buyer and seller, and the seller, looking to maximize the price received from the buyer, is not particularly motivated to disclose every potential pitfall they may see lurking as a future liability in the bank they are selling, or housed in its lending portfolio.

While just as a savvy used car buyer will take the vehicle they are purchasing to a quality mechanic for an inspection, an acquiring bank will partake in an extensive due diligence process to gather as much information about the asset they are purchasing as possible, in hopes of avoiding any lemons. However, in both instances there is a level of soft information possessed by the seller that may be impossible for the acquirer to discern. Perhaps the owner of a car knows that the vehicle only accelerates with a fraction of the "oomph" it did on the day they bought it - or that the starter seems to require just one more crank after the key is turned with each passing week. Similarly, the seller of a bank might know that the largest employer in the county is considering relocating, bringing hundreds of jobs (and the bank customer employees who possess them) with them. Or perhaps they know the CEO of the company that possesses the largest outstanding credit line for the bank just purchased a property in the Cayman Islands, and rumors around town involving the owner and his business keep involving acronyms like "IRS" and "DOJ". Such examples both relate to soft information (the former geographic, the latter related to the business), possessed by the seller which may be difficult, if not impossible, for the acquirer to easily discern, and which the seller has no motivation to disclose - lest it impact the sale.

3.2.3 Local and Geographic Asymmetric Information

Subsequent to Akerlof (1970), the idea of asymmetric information has been generalized to assess information density in losses in various other contexts, such as those of "local" or "geographic" asymmetric information.

In both of these contexts, one must re-frame the perspective with which we think of the root of the information asymmetry. While in the case of the "lemon" it is the seller who is intentionally withholding information from the buyer, in other contexts there may be a

similar lack of effective information transmission, but where the barrier constructed is not motivated by any incentives - and may be less immediately apparent to the subject. In the former case, it is simply that it might be extremely difficult (or costly) for the subject party to collect all relevant information to enable a fully informed decision; in short an information asymmetry because the information exists, however it is not available to the subject party.

A number of studies highlight the importance of geography in the context of information asymmetry and availability.

Weiler (2000) examines the role of first-movers in sparking future economic development, using the case of craft brewers in Lower Downtown Denver. The paper uses a game theory approach to demonstrate both the lack of incentive for businesses to initially move into economically distressed areas, given the potential inherent risks, and the potential overall welfare benefits that can emerge from the demonstrated successes of those first movers. A key component in this is that the lack of successful development does not provide enough quality information on probabilities of success for early businesses to be willing to make the initial investment decision. However, once those successes are realized and publicly discovered, this allows for future businesses to learn from that success - spurring a future development cycle.

Scorsone and Weiler (2004) similarly examine the inherent problems in fostering start up and small business development in low information environments. Looking at the context of New Market business development, the authors argue that while new market initiatives often target entrepreneurship, this creates its own additional set of problems in which you have compounding effects of introducing new businesses, which, by their definition, have little to no track record of success, to an environment with uncertain probability of success, resulting in two relatively risky endeavors intersecting. When this occurs, then, it may be difficult to discern, in the event a business does not succeed, whether the culprit is inherent to the business, the location, or some combination of the two. What may occur, however, is that in the instance that a business does not result in what is deemed a “success”, it

may reinforce negative preconceptions on the geographic area, despite the fact that those preconceptions may be a result of the lack of information on the area, rather than anything specific to the area itself. When viewed through this lens, the findings of Low and Weiler (2012) that entrepreneurship is often most attractive in "areas with high employment risk and/or low returns", could potentially create its own negative bias on entrepreneurship, to the extent that areas with weak labor markets may also be correlated with weak economic markets, and therefore prospects.

Weiler, Hoag, and Fan (2006) formalize some of the ideas in Scorsone and Weiler (2004). Again, focusing on the geographic information asymmetry problems that often plague regional development, the authors look at the role that public research can play in exposing ("prospecting") information on the "market fringe" in a manner that can benefit both private entrepreneurs and the public good. The authors again focus on the problem inherent to entrepreneurs in new or unexplored markets, in that the lack of information on expected future returns to the entrepreneur is often too costly for the entrepreneur to obtain, thereby inhibiting their ability to make informed investment decisions. The authors argue that public research has an important role to play helping to unearth requisite information in a manner that is beneficial not only to potential entrepreneurs, but by providing a public externality that increases overall social welfare net gains. The authors apply a Bayesian updating framework whereby the agent (entrepreneur) is able to update expected future net present value returns with the benefit of public information to make better informed investment decisions. They provide two examples of these benefits. In the first, examining a Colorado Micromalting Project, the authors determine that private entrepreneurs would likely fail to develop such a project as the information gathering costs would outstrip low expected future returns, and so would be unlikely to invest. With the aid of a cost-effective form of gathering information (a university feasibility study), however, which demonstrated both high social returns, as well as an opportunity for returns to the private investor, the requisite information was determined and a substantial net societal gain was created. In short, as the studies

demonstrated private returns were able to spur private investor involvement, while also unlocking the demonstrated social gains. In another example, regarding a proposed Colorado meatpacking facility, researchers were able to demonstrate to potential investors, again in an information thin environment, that a proposed project was unlikely to be profitable, which resulted in deterring the investment. This saved the investors money since the information from the study dissuaded them from from engaging in a poor investment.

Bunten et al. (2015) examines more closely the role entrepreneurs play in contributing to economic growth via their own unearthing of economic information. The authors develop and empirically test a model of entrepreneurship to demonstrate the benefits of the information discovery process by entrepreneurs, and how it contributes to growth. The authors use establishment birth and deaths ("churn") to represent the information gathering process - as future entrepreneurs learn from both the successes and failures of their peers. In environments with more churn, then, you will have more information gathering, which can benefit future growth. The authors find this to be the case, with this process resulting in future establishment and employment growth.

Finally, in the context of banking, multiple studies have examined the role geographic distance has on lending decisions. This is particularly relevant, as soft information can play a significant role in lending decisions, especially when lending to small businesses who might not have an extensive history of verified financial and credit information.

Petersen and Rajan (2002) examine small business lending in the context of geographic dispersion. They find that even as distance to the borrower has increased among lenders, technological change around communications and the productivity of lenders have enabled the proliferation of credit to distant borrowers to continue, despite soft information losses. They find that distant banks are willing to take on more risky borrowers- that might not have traditionally been served through traditional lending relationships.

Building on this, Agarwal and Hauswald (2010) utilize a more detailed data set of loan applications by small firms to a large bank enabling them to more specifically examine the impact of distance and soft information on loan assessment. They find that proximity to the borrower enables lenders to make more loans than they would otherwise by utilizing additional soft information gained through those relationships. Relatedly, Nguyen (2019) finds that bank closings result in less access to credit for small businesses, finding that annual loan origination decreases significantly for businesses local to a branch that closes. The author finds that this impact tends to be limited geographically in scope, with effects dissipating outside of a six-mile scope of the closing bank.

Related to these distant changes is the real impact on borrowers that occurs as a result of these structural changes in the branch banking market. In particular Célerier and Matray (2019) finds that the overall expansion in the number of bank branches that occurred between 1994 and 2005 increased financial inclusion for low-income households due to greater to accessibility of banking services. They also find that financial inclusion, in the form of participating in banking services, increases household wealth accumulations for those participating, relative to their unbanked counterparts.

3.3 Benefits, Costs, and Banking Performance

Restating the potential benefits and costs thus far detailed as influencing bank performance we have the following:

- Potential Benefits
  - Market pricing power
  - Operational Efficiency
  - Portfolio Efficiency

- Potential Costs
  - Lemon Costs
  - Local Information Costs

Given the relative benefits and costs to both expansion and the M&A process often employed to facilitate that expansion there may be differing outcomes for banks undergoing these processes. At the most abstract, let's consider three potential outcomes to banking performance which might occur based on the presence of each:

1. Benefits Exist, Costs Do Not - As benefits would impact banking performance positively, and there are no costs to offset it, banking performance would necessarily be impacted positively;

2. Costs Exist, Benefits Do Not - As benefits would impact banking performance negatively, and there are no costs to offset it, banking performance would necessarily be impacted negatively;

3. Both Benefits and Costs Exist - As benefits would impact banking performance positively, and costs would impact it negatively, there would be an indeterminate impact on bank performance, as it could be impacted positively or negatively, depending on the relative magnitudes of each effect.

A common theme in much of the literature regarding both bank geography, and bank M&A patterns, is the presence of mixed results. Below are a number of results attempting to analyze the impact of M&A activity in relatively abstract terms (generally whether Mergers and Acquisitions are net beneficial), demonstrating these mixed results.

James and Wier (1987) examine the impact of market competition in their event study of 60 bank acquisitions occurring between 1972 and 1983. They analyze stock returns around

merger announcements to examine excess returns to the acquiring institution and find that who accrues the benefits is often related to market competition. In markets with many potential alternative target firms there is a benefit to the acquiring firm, but there is a negative impact to the acquiring firm if there are many potential bidders. This indicates that the bank M&A market is subject to effects from competition, as markets where there are potential substitute targets for the acquiring bank led to higher returns, while those with increased competition for purchasing lead to lower returns.

Agrawal, Jaffe, and Mandelker (1992) look at the long-term impacts on returns post-acquisition for acquiring banks. Using market data covering the period of January 1955 to December 1987, the authors find that firms that undergo an acquisition perform, on average, approximately 10% worse than other similar banks who had not acquired banks over a five-year period post-merger.

Carow and Heron (1998) look at the impact of deregulation on competition and returns on the stock prices of bank holding companies during the passage of Riegle-Neal. They find that large banks experienced a 2.2% increase in abnormal returns during the process of passing the legislation. These gains were particularly large when there was a large market of potential acquisitions targets with similar characteristics existing in states that were headquartered in states that previously outlawed interstate banking and were therefore previously closed off from the potential acquisition market. This latter finding also supports the findings of James and Wier (1987) with respect to the impacts of competition.

DeLong (2003) compares market reactions to merger announcements with long run performance. The author finds that while the market can anticipate the benefits accruing from bank merger activity from focusing geography, activities, and earning streams, it fails to anticipate other benefits including when the acquiring bank is relatively inefficient, and when it can reduce bankruptcy costs.

In examining the mixed results of many of these early studies, DeLong and Deyoung (2007) suggest the findings of poor to mixed performance may be a result of information density in the banking M&A market, and that the lack of positive results may be the result of lack of a information around the merger process, performance, and best practices. In particular, they focus on a data set of 216 M&As involving publicly traded companies during the period of 1987 to 1999, examining both market reactions (as is the focus in many of the prior studies), and longer term post-merger performance (similar to DeLong (2003)). They hypothesize that, over time as more merger activity takes place, there will be knowledge spillover effects whereby market participants are able to better execute and assess merger performance. They look at this both in the context of the institution partaking in the merger activity, as measured by long term performance, as well as the market reaction to that activity, as measured by short term abnormal stock market returns, and in both instances find evidence of learning from prior experience. The authors find that overall in the market outcomes for acquiring institutions are often associated with negative performance, and markets are not particularly good at assessing the outcomes of mergers when announced. However, with greater availability of merger information, over time both long run performance outcomes for banks, and the stock market anticipation of that performance, increase significantly.

Examining the period of 2005 to 2011, Christopoulos (2012) find that in the post deregulatory period of 2005 to 2011 abnormal stock market returns are positive overall for the combined bidding and target bank, however this takes the form of a 9.7% average appreciation for acquisition targets, while bidders lose approximately 0.45%. While this dynamic still results in a less than efficient market outcome for acquiring banks, it is indicative of an overall market reaction that views bank merger activity as favorable in the post-deregulatory period, in contrast to some earlier findings.

Similarly, Leledakis and Pyrgiotakis (2016) examine bank M&A in the era after the passage of the Dodd-Frank Wall Street Reform and Consumer Protection Act ("Dodd-Frank Act")

in 2010 and find large abnormal returns to target banks, insignificant returns to bidding banks, and significant positive returns to the combined bank. By focusing on this period, they focus primarily on smaller banks with combined assets of $10B or less post-merger (to avoid additional regulatory burdens).

3.4 Geography Matters

The primary motivating paper for this study is Petach, Weiler, and Totten (2024), which provides the primary framework for this analysis. The authors use a similar approach and specification to specifically examine the impact of geographic size of the branch network on bank performance, as well as spillover effects of highly dispersed banks on nearby banks during the financial crisis. The authors find some evidence that both high levels of geographic dispersion, as well as proximity to other banks with high levels of geographic dispersion, results in lower bank performance - an effect that is especially pronounced when interacting with the financial crises. In these models, multiple specifications are run on each of the performance measures (ROA, ROE, and NPL - as in this paper), with generally similar results.

It is important to note is that the highest levels of significance are often found when Distance is included as both a continuous variable and with a dummy variable to demarcate the highest levels of geographic dispersion. In particular, in these specifications, it is often the case that while the continuous distance variable often provides a positive performance impact, the high distance specification has an even larger offsetting negative impact - suggesting that there may be a tipping point where the portfolio benefits of geographic dispersion are outweighed by the costs stemming from geographic information asymmetry information losses.

4 Theoretical Model

The following model lays out a hypothetical transaction for the sale of a banking institution between Buyer B and Seller S, assuming the following conditions:

Based on discount rate i, number of time periods n, final time period T, and future profit π_t, the Net Present Value of a business, u, can be calculated as:

$$NPV_u = \sum_{n=0}^{T} \frac{\Pi_{un}}{(1+i)^n} \tag{1}$$

where in each future time period, n, profit is equal to revenue, R minus cost C

$$\Pi_{un} = R_{un} - C_{un} \tag{2}$$

Where $\frac{\partial \Pi}{\partial R} > 0$ and $\frac{\partial \Pi}{\partial C} < 0$.

Assuming each party has the same discount rate, and Π_u is constant over time for each party, Equation 1 can be factored out and rewritten as

$$NPV_u = \Pi_u \times \sum_{n=0}^{T} \frac{1}{(1+i)^n} \tag{3}$$

Defining discount factor D as

$$D = \sum_{n=0}^{T} \frac{1}{(1+i)^n} \tag{4}$$

Substituting this back into Equation 3 and solving for each party, u results in the following equations:

$$NPV_B = D \times \Pi_B \tag{5}$$

$$NPV_S = D \times \Pi_S \tag{6}$$

Additionally, the transaction will result in a cost, TC, born by each buyer and seller such that

$$TC = TC_S + TC_B \tag{7}$$

The Buyer will then accept any price, P, less than or equal to the net present value they place on the bank, minus their portion of the transaction costs

$$P \leq NPV_B - TC_B \tag{8}$$

The Seller will similarly accept any price, P, greater than or equal to the net present value they place on the bank, plus their selling costs

$$P \geq NPV_S + TC_S \tag{9}$$

Substituting Equation 8 into Equation 9, solving for the net present value to the buyer NPV_B, and substituting in Equation 7 results in the following conditions necessary for a transaction to occur

$$NPV_B \geq NPV_S + TC \tag{10}$$

Substituting in Equation 5 and Equation 6, and dividing by D results in

$$\Pi_B \geq \Pi_S + \frac{TC}{D} \tag{11}$$

Substituting Equation 4 back in for D we can rewrite the discount factor adjusted cost DAC as

$$DAC = \frac{TC}{D} = TC \times \sum_{n=0}^{n} (1+i)^n \tag{12}$$

Assuming i and t are both greater than 0 then $DAC > C$. Substituting Equation 12 back into Equation 11 results in the following necessary conditions for a transaction to occur:

$$\Pi_B \geq \Pi_S + DAC \tag{13}$$

Now let's envision three scenarios and what occurs in each based on the relative profit function ascribed by each party, Π_u, assuming $TC > 0$ for each party (implying $DAC > TC > 0$ as well).

1. The seller has higher expectations of future returns than the buyer, such that $\Pi_S > \Pi_B$. Since $DAC > 0$, $\Pi_S + DAC \geq \Pi_B$ and the transaction will not occur.

2. The buyer and seller have the same expectations of future returns, such that $\Pi_S = \Pi_B$. Since $DAC > 0$, $\Pi_S + DAC \geq \Pi_B$ and the transaction will not occur.

3. The buyer has higher expectations of future returns than the seller, such that $\Pi_S < \Pi_B$. Since $DAC > 0$, $\Pi_B \geq \Pi_S + DAC$, only when the price willing to be paid by the Buyer is greater than the transaction costs, TC, grossed up by the inverse of the discount factor, $1/D$, resulting in an amount greater than just the transaction cost itself.

Recalling Equation 2, a transaction only occurs if either a) the buyer expects they are able to generate either excess revenues or operate with lower costs than the seller, or b) the seller expects that they are unable to generate the same level of profits or will operate with a higher level of costs than the buyer. In either instance $R_S > R_B$, $TC_S < TC_B$, or both must hold true.

Given that only completed transactions are observable, and by substituting Equation 2 in for each Buyer and Seller in Equation 13, it then must hold that $R_B - TC_B \geq R_S - TC_S + DAC$, implying that the combined relative difference in revenue and cost benefits to the Buyer must be greater than the transaction Discount Factor Adjusted Costs. It then must be the case that either the buyer is able to generate excess profits through the detailed revenue effects, or the seller is able to observe asymmetric information costs that the buyer is unable to observe, generating a lemons effect. Section 3.3 details a number of such benefits and costs detailed in the literature review.

Applying this to the context of bank performance, it would then be the case that as banks are able to generate excess profit, acquiring banks should generally perform better than peers who have not undergone acquisitions. On the other hand, if there is a lemons effect, then acquiring banks should generally perform worse than peers who have not undergone acquisitions. It may also be the case that, due to the fact that the soft information being retained by the seller is often marginal in nature, that those costs may only become apparent when exposed to additional stressors, such as those stemming from the financial crisis.

4.1 Motivation

This paper contributes to the literature by applying this transactional model framework to examine whether historical banking transactions occurred as a result of a buyers ability to generate excess profits, or a seller being able to observe asymmetric information costs that the buyer is unable to observe, generating a lemons effect. In doing so, this paper also allows

for the existence of both benefits and costs accruing to acquiring banks, and examines the relative magnitudes of each.

5 Methodology

5.1 Data Sources

The data used in this analysis is from the FDIC BankFind Suite. Data is accessed using the FDIC Bankfind Suite: API for Data Miners and Developers. A description of all data available on the FDIC website is available at the FDIC Bank Data Guide.

5.2 Data Description

Data for the analysis covers the period of 1994 to 2020[3], which is the total number of years that are available for all institutions. Data is organized by the following categories:

- Institutions (CSV format) - Listing of FDIC-insured banking institution demographic and headquarter location data.
- Institutions Definitions (CSV format) - Definitions of fields in Institutions data.
- Institution API Definitions (YAML format) - for financial institutions.
- Locations (CSV format) - Listing of FDIC-insured banks branches and locations.
- Locations Definitions (CSV format) - Definitions of fields in Locations data.
- Location API Definitions (YAML format) - for financial institutions locations/branches.
- History Events Definitions (CSV format) - Definitions of fields in History data.
- History API Definitions (YAML format) - for details on structure change events.
- Summary API Definitions (YAML format) - for aggregate financial and structure data, subtotaled by year, regarding financial institutions.
- Failure API Definitions (YAML format) - for details on failed financial institutions.
- Summary of Deposits API Definitions (YAML format) - for details on sod institutions.
- Financial API Definitions (YAML format) - for financial information for financial institutions.
- Links to Create Common Financial Reports (.xlsx spreadsheet)

For this analysis we will primarily be using data from the *History Events*, *Summary of Deposits*, and *Financial* APIs, as these contain the relevant information for this analysis. These serve the following purposes:

[3]While the data utilized covers the period of 1994 to 2020, years preceding 2000 are only included for the purpose of developing the prior acquisition history.

- History Events - this API contains information on the opening and closing of branches, as well as the acquisition of branches by other banks. This is used to identify the opening and closing of branches, as well as the acquisition of branches by other banks.

- Summary of Deposits ("SOD") - this API contains information on the deposits held by each branch. This is an annual dataset, as of June 30th of each year, and broadly contains demographic and location information for the purposes of this analysis which are later combined with other datasets.

  - *Branch Demographics and Location Data* - this includes the branch's address, as well as other characteristics of the bank. Latitude and Longitude coordinates are provided in many pieces of the dataset, but after further analysis these were generally not complete, or always reliable. As such, we will use the `tidygeocoder`[4] package to geocode the addresses of each branch to obtain the latitude and longitude coordinates. This is used to calculate the distance of each branch to the bank's headquarters, as described below.
  - *Institution/Bank Holding Company Relationships* - includes both the institution (`CERT`) and BHC (`rsshcr`) relative to each branch, and which are used for both calculating distances, as well as identifying acquisitions. The three main units we are looking at in these relationships are:
    * Holding Companies. This is used for further identifying relationships between institutions, since some holdcos may own multiple institutions.
    * Institutions - this is the bank headquarters - the main branch of the banking unit. Individual branches are sub units of the main institution unit. Bank performance metrics (i.e., `roa`, `roe`, `npl`) are available at the institution level.

[4]Source: Cambon J, Hernangómez D, Belanger C, Possenriede D (2021). tidygeocoder: An R package for geocoding. Journal of Open Source Software, 6(65), 3544, https://doi.org/10.21105/joss.03544 (R package version 1.0.5)

These are uniquely identified by the `cert` variable, and are the primary unit of record in the dataset.

    * Branches - this is the individual branch of the bank. Branches are sub-units of the main institution unit. These are used for geographic analysis as they have geographic coordinates. These are also used for acquisitions. These are uniquely identified by the `uninumbr` variable, and variables associated with the branch end in `br`. It should be noted that `uninumbr` is not necessarily consistent across time, and branch locations can change `uninumbr` - generally through changes.

  - *Branch Deposits* - this includes the total deposits held by each branch. This is used for weighting of both the variables representing banks geographic distance and changes from mergers.

Conceptually the process can be broken down into a series of steps:

5.3 Data Preparation

In order to run the analysis, the data was prepared to create a number of datasets that were used in the analysis.

5.3.1 Pull Data from the SOD API

In order to access the API a package for R, `fdicapi` was created by the author and is available on Github.

All available data was pulled from the SOD API for the years 1994 to 2020. It was then broken up to create 3 datasets for "Branches", "Institutions", and "Holdcos"[5], which are then used in later parts of the analysis.

[5]Another term used to refer to Bank Holding Companies, or BHCs

5.3.2 Branch Distances

A branch distance dataset was constructed using branch level data available from the SOD using the Branch Office Deposits ("BOD") dataset. The dataset provides data for each branch of institutions covered by the FDIC for all years beginning in 1994, as of June 30th of each year. The data provided is primarily descriptive in nature, containing;

- *Geographic Data* including address and location data;
- *Regulatory Classification Data* such as the bank class, primary insurer, and chartering agent;
- *Relationship Data* such as unique identification data for each branch, and identification numbers for banks related owner institution, and top-level holding company; and
- *Financial Data* which is relatively limited in scope and related to assets held at the branch, such as branch deposits, as of June 30 of each year.

For the purpose of building the branch distance dataset, the primary data used from the BOD dataset is Geographic and Relationship in nature.

5.3.2.1 Branch Addresses

The first step in creating the branch distance dataset was to create a branch address dataset. Initially the plan had been to use Latitude and Longitude data provided in the FDIC dataset, however after further analysis it was determined that this data was not always complete, or always reliable.

In short, the data needed to be geocoded. In order to do this an address dataset was first created for each unique address in the branch dataset, and these were provided a unique `address_id`. Initially, it was planned to be done by collapsing data on `uninumbr`, which is the unique ID specified with a physical branch location, which does not vary with ownership,

according to the FDIC definitions file. However, across the dataset there are numerous banks with missing `uninumbr` values. Additionally, per Terrastrat, this rule is not always followed, and a bank may change ownership but not keep its `uninumbr`.

Once this unique address list was created the data was geocoded using the tidygeocoder package[6]. This was done in two steps, with a first pass using the Census Geocoder, provided by the United States Census Bureau, and a second pass using the Nominatim API with OpenStreetMap data.

The Census Geocoder was used first because it is free and has a higher rate limit than Nominatim. The API only accepts point addresses, however, and is limited when address data cannot be found. Nominatim is rate limited, and so a second pass was done afterward to backfill the remaining data. Nominatim allows lookups by City, Zip Code or County, so these were used to approximate locations when full address data was not available from the Census Bureau[7].

5.3.2.2 Average Distance

The next step was to calculate the average distance of each branch to the bank's headquarters. This was done by first identifying the bank headquarters, and then calculating the distance of each branch to the bank headquarters on an annual basis. This was done by calculating Haversine distances based on the latitude and longitude with the `geosphere` package in R. The distance is calculated in meters, so it is divided by 1000 to convert to kilometers.

[6]Source: inst/CITATION Cambon J, Hernangómez D, Belanger C, Possenriede D (2021). tidygeocoder: An R package for geocoding. Journal of Open Source Software, 6(65), 3544, https://doi.org/10.21105/joss.03544 (R package version 1.0.5)

[7]single address geocoding with Nominatim is possible but would take an extremely long time due to rate limiting

Next branch network sizes were calculated by taking the total number of branches for each institution (as calculated from the branch data) each year. The average distance to the main office was then calculated by taking the mean of this distance.

5.3.3 Merger Dataset

A dataset of merger history also needed to be created for each institution in order to perform the analysis.

This dataset was developed using the "History" API, which provides detail on Structure Change Events. This might be when a branch or institution is established or closed, but also includes when a branch is acquired or merged with another bank.

The first step to creating this dataset was to identify merger codes from the possible Change Codes available. Change codes 221 to 223, which relate to absorption, consolidation, or merger, all without assistance[8]. This was done to only include voluntary mergers and eliminate distressed mergers which might bias the sample.

Next related party mergers were eliminated from the dataset. These were determined when the two parties were both part of the same top level holding company.

Branch data was added by merging branch data from the distance dataset. Next the total number of mergers or acquisitions was calculated for each institution in each year at both the institution and branch level - i.e., the total number of other institutions acquired, as well as the total number of branches acquired as a result of those acquisitions.

Data was then checked and cleaned to ensure that the data was accurate and complete. The primary issue of missing branch data for a particular transaction, which generally occurred

[8]For purposes of this paper each type of combination is treated the same, with the "surviving" institution considered to be the acquiring institution

as a result of one party to the transaction not being covered by the FDIC, was resolved using branch counts from the FDIC merger dataset.

Because processing dates for the merger dataset begin in 1999, effective dates for merger activity needed to be filtered and checked. Effective years were filtered to only include after 1994 (as these date back as far as the 1960s), and acquisition totals were checked for consistency across time in order to ensure that the data was complete.

5.3.4 Financial Data

Financial data was compiled from the "Financials" API, all at the institution level. This includes data on the financial performance of each institution, including Return on Assets (ROA), Return on Equity (ROE), and Non-Performing Loans/Total Loans (NPL, and also referred to as Non-current Assets, or NPV, in this paper), as well as various other metrics that are used as independent variables in the regression analysis.

Data is provided on a quarterly basis, and as such, needed to be transformed to annual amounts in order to correspond with the annual nature of the remaining data. This was done on June 30^{th} fiscal year end in order to correspond with the timing of the SOD data. For stock variables this means using the value as of June 30^{th} of each year, while flow variables use the four quarter average of the preceding year.

5.3.5 Combining the Data

Finally, the various datasets were combined in order to create a single panel dataset for analysis.

5.4 Model Specification

To estimate the impact of bank consolidation on bank performance, the following model is estimated for the period $t = 2000 - 2020$:

$$Performance_{it} = \beta_0 + \beta_1 HighAcquisitionFlag_{it} + \beta_2 HighDistanceFlag_{it} +$$
$$\beta_3 HighAcquisitionxRecession_{it} + \beta_4 HighDistancexRecession_{it} + \beta_5 X_{it}^T + \epsilon_{ist}$$

In this model performance for bank i at time t is measured for each of the following:

- Return on Assets (ROA_{it})
- Return on Equity (ROE_{it})
- Percentage of Nonperforming Assets ($NPERFV_{it}$)

These in turn are a function of the following:

- High Acquisition Flag$_{it}$ - a dummy variable indicating whether the specific bank was 75th percentile or higher of banks in terms of the percentage of total bank branches owned by the bank were acquired in the previous 5 years. If a bank is 75th percentile or higher, then the dummy variable is equal to 1, otherwise, it is equal to 0. The five-year period post acquisition is consistent with the period referenced in Agrawal, Jaffe, and Mandelker (1992), in terms of post-acquisition effects.
- High Distance Flag$_{it}$ - a dummy variable indicating whether the specific bank was 75th percentile or higher of banks in terms of the average distance to bank headquarters for each branch in the bank network. If a bank is 75th percentile or higher, then the dummy variable is equal to 1, otherwise it is equal to 0.
- Recession - A recession dummy was included for the Great Recession, represented by the years 2008 to 2011. While the model uses year fixed effects, and therefore a Recession Dummy is not necessary, nor presented on its own in the results, it is included for testing interaction terms with each of our High Acquisition and High Distance flags,

in order to test whether institutions are particularly affected by these treatments in times of stress.

- High Acquisition Flag $\times$ Recession$_{it}$ - a dummy variable indicating an interaction between the High Acquisition Flag and the Great Recession (2008-2011 for purposes of this analysis). This is equal to 1 if the bank is in the 75th percentile or higher of banks in terms of the total combined number of acquisitions in the previous 5 years, and the bank is in the Great Recession, otherwise it is equal to 0.

- High Distance Flag$\times$Recession$_{it}$ - a dummy variable indicating an interaction between the High Distance Flag and the Great Recession (2008-2011 for purposes of this analysis). This is equal to 1 if the bank is in the 75th percentile or higher of banks in terms of the average distance to bank headquarters for each branch in the bank network, and the bank is in the Great Recession, otherwise it is equal to 0.

- X_{it}^{T} - a vector of bank characteristics that are expected to influence the performance of the bank, as follow:

  - Specialization Group$_{it}$ - Type of lending specialized in by the bank.

  - Fee Share$_{it}$ - the percentage of total income derived from non-interest activities (i.e., $FeeShare = \frac{Non-InterestIncome}{TotalIncome}$).

  - Capital/Asset Ratio$_{it}$ - the ratio of tier 1 capital to total assets (i.e., $Capital/AssetRatio = \frac{Capital}{Assets}$).

  - Log Assets$_{it}$ - the total assets of the bank (in log form).

  - Log Distance$_{it}$ - average distance from the bank headquarters to each of its branches (in log form).

5.5 Strategy

5.5.1 Difference-in-Differences

This study employs a difference-in-differences (DiD) model to estimate the impact of bank consolidation on performance metrics. The DiD approach is particularly well-suited for this analysis as it allows us to isolate the effect of our key variables of interest (High Acquisition Flag and High Distance Flag) relative to "baseline" banks who do not exhibit these characteristics.

The DiD model compares the changes in outcomes between a treatment group (in our case, banks with high acquisition activity or high distance) and a control group (banks without these characteristics). The coefficient on each of these flags reflects the impact of this treatment relative to the control group.

In our specification, we include interaction terms between our main treatment variables and a recession indicator for the 2008 financial crisis. This allows us to also examine how the effects of two variables, high acquisition activity and high distance, on bank performance may differ during periods of economic stress, and whether these characteristics made them acutely susceptible to these stressors, relative to their peers.

Per Callaway and Sant'Anna (2021) and Sant'Anna and Zhao (2020), while difference-in-differences estimators are quite popular, generally, the canonical iteration involves two time periods, before and after the treatment is applied, and two groups, the treatment and the comparison group. Oftentimes, however, when working with real world data in economics, such a straightforward analysis is often not possible. Rather, treatment variables are often introduced at various time periods in the analysis, and there may be more than two groups, and this analysis is one such case. As such one should be cautious in the initial results presented in this paper until further robustness checks are done to ensure robustness of the evidence, as suggested by Callaway and Sant'Anna (2021) and Sant'Anna and Zhao (2020).

5.5.2 Panel Fixed Effects

Parameters are estimated using a two-way fixed effects model, using the plm (panel linear model) package in R which means that both bank-specific and time-specific effects are controlled for. This is because an individual banks performance is expected to be a product of both its own characteristics (e.g., management, geographic dispersion, merger history, etc) impacting the bank itself, as well as year fixed effects related to overall macroeconomic conditions impacting all banks in a given year (e.g., such as Federal Reserve interest rate policy, or the overall economic conditions of the then current environment, such as may be characterized by a recession). Additionally, one might want to include additional fixed effects for either the state or region which the majority of a bank's assets are located to reflect there may additionally be other regional economic conditions impacting a specific bank[9].

Since a two-way transformation is being specified, the "within" model is also included as time effects are being included in the model.

As a two way fixed effect model is being used annual year fixed effects are also included in the model, but not reported. This also means that the *Recession* dummy is not provided absent its interaction as it is implicitly included within the time fixed effects.

The sample period is set from 2000 to 2019, as this allows for a 5-year window to calculate the High Acquisition Flag, and ends prior to the COVID-19 pandemic which may have had a significant impact on bank performance.

[9]As part of this analysis, an iteration of the model was run with state fixed effects included, resulting in little explanatory power. These were included based on the branch headquarters, however, and so may be limited in their usefulness in analyzing banks operating in multiple states.

6 Results

6.1 Descriptive Statistics

Table 1 presents summary statistics for the three performance measures and key independent variables. The table displays overall mean and standard deviation for each variable, as well as statistics based on banks' acquisition history and the difference between the two groups.

The mean Return on Assets (ROA) is 0.8%, mean Return on Equity (ROE) is 7.1%, and mean Nonperforming Loans/Total Loans (NPERFV) is 1.3%. Of the total observations, 2.0% are for banks with a High Acquisition Flag, 25% for banks with a High Distance Flag, and 22% of all observations are during the great recession.

Banks with a history of acquisitions generally exhibit higher levels of ROA and ROE, but also slightly higher levels of Non-Performing Assets. This suggests potential benefits associated with acquisition history, although these effects may be confounded with other factors positively correlated with returns, such as total assets.

Acquisition history is also associated with larger bank size, both in terms of number of branches (average of 70.6 branches compared to 5.8) and overall distance (average distance of 140.9 miles compared to 31.3).

6.2 Main Results

Table 2 presents the results of the regression analysis for each of the three performance measures. As this is a difference-in-differences model, the effects can be interpreted as follows:

1. High Distance Flag: The effect of being a high distance bank relative to the control group on each performance measure, in percentage points.

Table 1: Summary Statistics

Characteristic	**Overall** N = 145,470	**No Acquisition History** N = 131,203	**Prior Acquisition History** N = 14,267	**Difference**
Return on Assets (%)	0.8 (3.8)	0.8 (3.9)	1.0 (3.0)	-0.26
Return on Equity (%)	7.1 (22.2)	7.0 (22.5)	8.0 (18.9)	-0.97
Non-Current Assets (% of Total)	1.3 (2.1)	1.3 (2.2)	1.3 (1.8)	-0.01
Total Mergers	0.2 (1.2)	0.0 (0.0)	2.0 (3.3)	-2.0
Total Branches Acquired	2.3 (40.5)	0.0 (0.0)	23.5 (127.5)	-23
Distance (miles)	42.1 (153.7)	31.3 (126.5)	140.9 (288.2)	-110
Total Branches	12.1 (117.9)	5.8 (25.6)	70.6 (363.0)	-65
Total Assets (Millions $)	2,259.1 (43,556.7)	734.5 (6,220.8)	16,280.1 (137,009.4)	-15,546
Deposits (Millions $)	1,621.8 (31,082.7)	530.9 (4,387.2)	11,654.1 (97,790.4)	-11,123
Net Loans and Leases (Millions $)	1,253.4 (20,524.3)	451.7 (3,698.2)	8,625.3 (64,104.2)	-8,174
Capital/Asset Ratio	0.1 (0.1)	0.1 (0.1)	0.1 (0.1)	0.01
Fee Share	0.1 (0.1)	0.1 (0.1)	0.2 (0.1)	-0.04
High Acquisition Flag	2,845	603	2,242	
High Distance Flag	36,375	26,921	9,454	
Recession Flag	31,924	28,834	3,090	
Specialization Group				0.49
International Specialization	94	35	59	
Agricultural Specialization	30,405	27,894	2,511	
Credit-Card Specialization	469	425	44	
Commercial Lending Specialization	73,403	63,983	9,420	
Mortgage Lending Specialization	14,230	13,528	702	
Consumer Lending Specialization	1,964	1,900	64	
Other Specialized Under 1 Billion	6,884	6,600	284	
All Other Under 1 Billion	16,782	16,023	759	
All Other Over 1 Billion	1,239	815	424	

[1] Mean (SD); n

[2] Welch Two Sample t-test; Standardized Mean Difference

2. High Acquisition Flag: The effect of being a high acquisition bank relative to the control group on each performance measure, in percentage points.

3. High Distance Flag x Recession: The additional effect of being a high distance bank during the recession, relative to the control during the same period, in percentage points.

4. High Acquisition Flag x Recession: The additional effect of being a high acquisition bank during the recession, relative to the control during the same period, in percentage points.

5. Log Assets: As Log Assets is natural log transformed there is a specific interpretation of the coefficient. Specifically, the coefficient can be interpreted as the direct magnitude impact (expressed in percentage points) on the response variable due to a 1% increase in assets[10].

No individual effect for the recession is reported, as this is implicitly included in the time fixed effects.

6.2.0.1 Log Assets

The coefficient on Log Assets is positive and statistically significant at the 99.9% level for both ROA and ROE, and negative and significant at the 99.9% level for Non-Current Assets. This suggests that larger banks tend to perform better in terms of profitability and asset quality.

[10]It should be noted that in the case of Return on Assets and Return on Equity are correlated with ln(assets), as Assets is the denominator in each calculation. Therefore, we are making the assumption that to the extent each of the ROA and ROE variables are correlated with Assets, it is with respect to the causal effects being tested in the model, and that if not for the scale effects these variables would be otherwise independent

Table 2: Regression Results

Characteristic	Return on Assets (%) Beta (SE)	Return on Equity (%) Beta (SE)	Non-Current Assets (% of Total) Beta (SE)
High Distance Flag	-0.19***	-1.1***	-0.10***
	(0.031)	(0.239)	(0.021)
High Acquisition Flag	-0.28***	-1.2**	0.03
	(0.057)	(0.448)	(0.039)
High Distance Flag * Recession Flag	-0.19***	-2.4***	0.28***
	(0.034)	(0.268)	(0.024)
Recession Flag * High Acquisition Flag	-0.46***	-8.3***	0.38***
	(0.119)	(0.931)	(0.082)
Log Assets	0.56***	5.5***	-0.19***
	(0.022)	(0.169)	(0.015)
Log Distance (miles)	-0.01*	-0.02	0.01***
	(0.002)	(0.017)	(0.001)
Capital/Asset Ratio	-11***	6.9***	-4.2***
	(0.159)	(1.24)	(0.109)
Fee Share	3.7***	30***	-1.6***
	(0.089)	(0.700)	(0.062)
Specialization Group			
International Specialization	—	—	—
Agricultural Specialization	0.75	5.4 (5.22)	-0.84
	(0.666)		(0.459)
Credit-Card Specialization	-6.7***	-19***	0.05
	(0.716)	(5.60)	(0.493)
Commercial Lending Specialization	0.84	5.8 (5.21)	-0.94*
	(0.665)		(0.458)
Mortgage Lending Specialization	0.84	5.2 (5.21)	-1.0*
	(0.666)		(0.458)
Consumer Lending Specialization	0.73	5.2 (5.23)	-0.88
	(0.668)		(0.460)
Other Specialized Under 1 Billion	0.26	3.6 (5.21)	-0.94*
	(0.666)		(0.458)
All Other Under 1 Billion	0.71	4.6 (5.21)	-0.81
	(0.665)		(0.458)
All Other Over 1 Billion	0.11	2.7 (5.15)	-1.0*
	(0.658)		(0.453)

[1] *p<0.05;* ***p<0.01;*** p<0.001

[2] SE = Standard Error

For ROA, the coefficient of 0.56 indicates that a 1% increase in assets is associated with a 0.0056 percentage point increase in ROA. Given that the mean ROA is 0.8%, this represents a 0.7000% increase relative to the mean for each 1% increase in assets. For a bank with average assets of $2,259.1 million, a 1% increase ($22.6 million) would be associated with an increase in ROA from 0.8% to about 0.8056%.

For ROE, the coefficient of 5.5 suggests that a 1% increase in assets is associated with a 0.0550 percentage point increase in ROE. With a mean ROE of 7.1, this represents a 0.7700% increase relative to the mean for each 1% increase in assets at the mean.

For Non-Current Assets, the coefficient of -0.19 indicates that a 1% increase in assets is associated with a 0.0019 percentage point decrease in the ratio of non-current assets to total assets. Given the mean of 1.3, this represents a 0.1500% decrease relative to the mean for each 1% increase in assets at the mean.

It should be noted that while these magnitude effects seem quite small, mean Assets is quite low relative to some of the largest banks in the sample, and so the effects are likely to be economically significant for larger banks. As an example, the 99th percentile observation is 9.50 times larger than the mean bank, and the highest observed asset level is 1,542 and so the effects scale effects can become quite large for the largest banks.

6.2.0.2 High Distance Banks

Banks flagged as high distance exhibit statistically significant effects on all three performance metrics at the 99.9% level. The High Distance Flag is associated with a 0.1900 percentage point decrease in ROA, a 1.10 percentage point decrease in ROE, and a 0.1000 percentage point decrease in Non-Current Assets.

Compared to the means (ROA: 0.8%, ROE: 7.1%, Non-Current Assets: 1.3%), these effects represent substantial changes:

- A 23.8% decrease in ROA relative to the mean
- A 15.5% decrease in ROE relative to the mean
- A 7.70% decrease in Non-Current Assets relative to the mean

The reduction in Non-Performing Loans indicates a benefit to the banks, though the magnitude of the effect is relatively small compared to the negative impacts on profitability.

The interaction between the High Distance Flag and the Recession indicator reveals additional effects during economic downturns. During recessions, High-Distance banks experience an additional 0.1900 percentage point decrease in ROA, a 2.40 percentage point decrease in ROE, and a 0.2800 percentage point increase in Non-Current Assets, relative to control banks during the same period. These interaction effects suggest that the negative impacts of geographic dispersion may be amplified during economic stress periods.

6.2.0.3 High Acquisition Banks

Banks designated as High Acquisition are associated with generally lower performance relative to those not classified as High Acquisition. The High Acquisition Flag is statistically significant at the 99.9% level for ROA and ROE, though not for the Non-Performing Loan ratio.

The High Acquisition Flag is associated with:

- A 0.2800 percentage point decrease in ROA (35.0% of the mean)
- A 1.20 percentage point decrease in ROE (16.9% of the mean)
- A 0.0300 percentage point increase in Non-Current Assets (2.30% of the mean), though not statistically significant

These effects are substantial, particularly for ROA, indicating that the costs associated with recent acquisitions outweigh the potential benefits in the short to medium term.

The interaction between the High Acquisition Flag and the Recession indicator reveals even more pronounced effects during economic downturns. During recessions, High-Acquisition banks experience an additional 0.4600 percentage point decrease in ROA, a 8.30 percentage point decrease in ROE, and a 0.3800 percentage point increase in Non-Current Assets, relative to their control peers during the same period. The 8.30 percentage point reduction in ROE for High Acquisition banks during the recession represents more than the entire average ROE (7.1%), indicating a severe impact on profitability.

These results suggest that banks with high levels of recent acquisition activity are particularly vulnerable during economic downturns, experiencing substantially larger negative effects on their performance metrics compared to other banks.

6.2.1 Specialization

Credit-Card Specialization is the only specialization associated with a significant effect, with large negative effects on both ROA and ROE. There is a -6.7 percentage point impact on Return on Assets for banks with a Credit Card Specialization, and a -19 percentage point impact on Return on Equity. These banks are relatively rare, however, with only .3% (469) observations involving a bank desingated with the Credit-Card Specialization.

6.2.2 Relative Effects

While the largest banks are substantially larger than the mean asset size, most banks are substantially smaller. In fact, 95% of observations are under the mean asset level, and so the effects are likely to be quite small for the majority of banks.

Using a highly stylized model to put these relative magnitudes in context, these relative coefficients can be applied by calculating the relative level to the mean required to offset the relative impacts of the High Distance and High Acquisition Flags. For example, to offset the negative impact of the High Distance Flag, a bank would have to increase its asset base by 35%. If this requires the bank to acquire enough branches to also be classified as “High

Table 3: Model Statistics

Model	R-squared	Adjusted R-squared	F-Statistic	AIC
Return on Assets (%)	0.084	0.010	769 on 16 and 134,552 DF, p-value $<$ 00.001	659,137
Return on Equity (%)	0.024	-0.055	208 on 16 and 134,552 DF, p-value $<$ 00.001	1,257,792
Non-Current Assets (% of Total)	0.019	-0.060	166 on 16 and 134,552 DF, p-value $<$ 00.001	550,448

Acquisition" then the asset base would need to be increased by an additional 51%, or a total of 85% - a near doubling in size. Given that the threshold to qualify for the "High Acquisition Flag" is having acquired between 40% and 50% of the current level of branches in the last 5 years, it is quite possible that many banks would arrive at this statistical threshold in the process of acquiring enough branches to offset the negative impacts of the High Distance Flag. Even if they do not fully crossing the threshold level, there is likely still a relatively strong negative marginal effect on performance from acquisition levels that are near, if not exactly across, the threshold level. Again, this is a highly stylized example, and so these effects should not be taken as literal, however they may nonetheless signal relationships that are useful in interpreting the relative magnitudes of the coefficients.

6.3 Statistical Checks

6.3.1 Model Statistics

Table 3 presents the R-squared, F-statistic, and AIC for each of the regression models.

All three models have relatively low R-squared values, indicating that the models explain little of the variance in the data, with the Return on Assets having the highest percentage of the variance explained at 8.40%. The adjusted R-squared values are even lower with both ROE and Non-Performing Loans having negative adjusted R-squared values. This suggests that the models may not be capturing all the relevant factors that influence the dependent variables.

The F-statistics, however, are highly significant for all three models, indicating that while the models may not explain much of the variance, they do represent a significant improvement over the null model.

Finally, the Non-Performing Loans model has the lowest AIC value, suggesting that it is the best-fitting model of the three. However, all models have relatively high AIC values, and low R-Squared and Adjusted R-Squared values, indicating that there may be room for improvement in the model specification.

Overall, these statistics suggest that while the models are statistically significant, they are likely missing many relevant factors influencing bank performance, and therefor offer significant room for improvement with respect to explanatory power. The models do have some very high outlying values, particularly for ROA and ROE, and so Section 9 examines the impacts of these outliers on these statistics.

6.3.2 Variance Inflation Factor

One potential issue to check for is issues with multicollinearity. One way to check this is with a VIF check, however since this is a fixed effects model a standard VIF check cannot be used. Instead, regressions were rerun specifying the model as "pooling" and check for multicollinearity. The results of this check are presented in Table 4, Table 5, and Table 6.

The resulting GVIF values are generally between between 1.1 and 1.6, and Adjusted GVIF of between 1 and 1.3. These values generally reflect low to moderate multicollinearity, and so there does not appear to be any issues with multicollinearity.

6.3.3 Correlation Matrix

We can also create a correlation matrix for the key independent variables to check for multicollinearity. The results of this check are presented in Figure 1.

Table 4: Variance Inflation Factor - Return on Assets(%)

Characteristic	Beta (SE)	**GVIF**	**Adjusted GVIF**
High Distance Flag	-0.11*** (0.028)	1.6	1.3
High Acquisition Flag	0.12 (0.078)	1.3	1.1
High Distance Flag * Recession Flag	-0.14* (0.054)	1.6	1.3
Recession Flag * High Acquisition Flag	-0.74*** (0.177)	1.3	1.1
Specialization Group		1.5	1.0
International Specialization	—		
Agricultural Specialization	3.9*** (0.387)		
Credit-Card Specialization	4.5*** (0.417)		
Commercial Lending Specialization	3.1*** (0.385)		
Mortgage Lending Specialization	3.3*** (0.386)		
Consumer Lending Specialization	3.5*** (0.394)		
Other Specialized Under 1 Billion	3.8*** (0.388)		
All Other Under 1 Billion	3.5*** (0.387)		
All Other Over 1 Billion	2.3*** (0.395)		
Fee Share	6.5*** (0.083)	1.1	1.1
Capital/Asset Ratio	-1.1*** (0.151)	1.2	1.1
Log Assets	0.13*** (0.009)	1.8	1.3
Log Distance (miles)	0.00 (0.002)	1.5	1.2

[1] *p<0.05;* ***p<0.01;*** p<0.001

[2] SE = Standard Error, GVIF = Generalized Variance Inflation Factor

[3] GVIF^[1/(2*df)]

Table 5: Variance Inflation Factor - Return on Equity(%)

Characteristic	Beta (SE)	**GVIF**	**Adjusted GVIF**
High Distance Flag	0.23 (0.167)	1.6	1.3
High Acquisition Flag	-0.62 (0.461)	1.3	1.1
High Distance Flag * Recession Flag	-2.2*** (0.318)	1.6	1.3
Recession Flag * High Acquisition Flag	-9.2*** (1.04)	1.3	1.1
Specialization Group		1.5	1.0
International Specialization	—		
Agricultural Specialization	17*** (2.27)		
Credit-Card Specialization	16*** (2.45)		
Commercial Lending Specialization	11*** (2.26)		
Mortgage Lending Specialization	11*** (2.27)		
Consumer Lending Specialization	15*** (2.32)		
Other Specialized Under 1 Billion	12*** (2.28)		
All Other Under 1 Billion	14*** (2.27)		
All Other Over 1 Billion	8.0*** (2.32)		
Fee Share	24*** (0.486)	1.1	1.1
Capital/Asset Ratio	-8.7*** (0.887)	1.2	1.1
Log Assets	0.96*** (0.053)	1.8	1.3
Log Distance (miles)	0.04*** (0.010)	1.5	1.2

[1] *p<0.05;* ***p<0.01;*** p<0.001

[2] SE = Standard Error, GVIF = Generalized Variance Inflation Factor

[3] GVIF^[1/(2*df)]

Table 6: Variance Inflation Factor - Non-Current Assets(% of Total)

Characteristic	Beta (SE)	**GVIF**	**Adjusted GVIF**
High Distance Flag	-0.10*** (0.016)	1.6	1.3
High Acquisition Flag	-0.17*** (0.044)	1.3	1.1
High Distance Flag * Recession Flag	0.34*** (0.030)	1.6	1.3
Recession Flag * High Acquisition Flag	0.34*** (0.099)	1.3	1.1
Specialization Group		1.5	1.0
International Specialization	—		
Agricultural Specialization	0.21 (0.215)		
Credit-Card Specialization	1.1*** (0.232)		
Commercial Lending Specialization	0.90*** (0.214)		
Mortgage Lending Specialization	0.29 (0.214)		
Consumer Lending Specialization	0.36 (0.219)		
Other Specialized Under 1 Billion	0.14 (0.215)		
All Other Under 1 Billion	0.38 (0.215)		
All Other Over 1 Billion	0.26 (0.219)		
Fee Share	-0.34*** (0.046)	1.1	1.1
Capital/Asset Ratio	-2.3*** (0.084)	1.2	1.1
Log Assets	0.01 (0.005)	1.8	1.3
Log Distance (miles)	0.00** (0.001)	1.5	1.2

[1] *p<0.05;* ***p<0.01;*** p<0.001

[2] SE = Standard Error, GVIF = Generalized Variance Inflation Factor

[3] GVIF^[1/(2*df)]

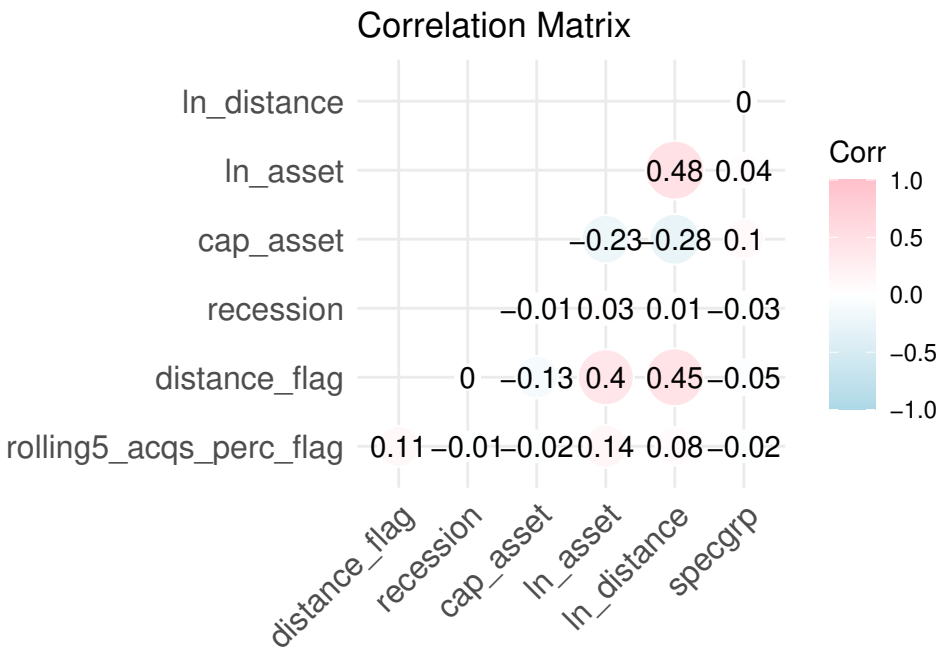

Figure 1: Correlation Matrix

Examining correlation between variables also does not show any issues with multicollinearity. The highest correlation is between the High Distance Flag and Log Assets, which is 0.4, and Log Asset and and Log Distance at .48. Neither of these are high enough to be of much concern.

6.3.4 Summary

While the models do not appear to be particularly well specified, the overall F statistics and individual variable significance suggest that the models are capturing some of the relevant effects. The VIF and correlation checks also suggest that multicollinearity is not an issue. However, the low R-squared values and high AIC values suggest that there is room for improvement in the model specification.

7 Discussion

7.1 Interpretation of Results

Recalling back to Section 4 we identified three separate vectors through which benefits and costs associated with bank branch expansion might be transmitted to bank performance:

- **Operational Efficiency Effects** - These include more traditional scale benefits accruing to the firm such as those derived from sharing administrative costs, reducing redundant operations, and improved management. The `Assets` variable, while primarily related to operational effects, is most representative of this concept in analyzing the results. Given that this is only associated with a benefit, and not a cost, we would expect this to have a **positive** impact on each performance variable.

- **Geographic Effects** - This includes both benefits accruing from portfolio effects of diversification, and costs associated with geographic distance related information losses. Because this contains both benefits and costs, the **net** impact on each performance variable will provide information on the relative weights of each. In particular, if **benefits > costs** we expect a **positive** impact on performance, if **costs > benefits** we would expect a **negative** impact on performance, and if they have relatively equal impacts we might expect performance impacts to be **mixed** or **not significant**

- **M&A Effects** - As the M&A process in and of itself does not have any theoretical benefit (the benefits accruing are via the resulting Operational Efficiency and Geographic Portfolio Effects), we would only expect this to contain costs via a lemons effect, if it exists. Therefore, we would expect this to have a **negative** effect on each performance variable.

Additionally, recall that the following signs reflect relative **positive** impacts on bank performance;

Table 7: Results Interpretation - Overall

Variable	Concept	Expected	Actual	Return on Assets (%)	Return on Equity (%)	Non-Current Assets (% of Total)
High Distance	Geographic Information Geographic Portfolio	Cost Benefit	Cost Benefit	-0.19*** (0.031)	-1.1*** (0.239)	-0.10*** (0.021)
High Acquisition	Lemon	Cost		-0.28*** (0.057)	-1.2** (0.448)	0.03 (0.039)
Log Assets	Operational Efficiency, Portfolio Efficiency	Benefit		0.56*** (0.022)	5.5*** (0.169)	-0.19*** (0.015)
				Benefit = Positive	Benefit = Positive	Benefit = Negative

- ROA +
- ROE +
- Non Performing Assets -

Table 7 displays select explanatory variables their expected signs based on their theoretical inclusion as a related benefit or cost, and their actual signs. Areas shaded in green are those related to benefits, and those in red related to costs.

Given this context - both the `Assets` and `Acquisitions` variables perform as expected, with Log Assets having positive impacts on both ROA and ROE, and a Negative Impact on Non-Current Assets, and each significant at the 99.9% level. High levels of Acquisitions, while also containing the expected signs for each, are only significant at the 99.9% level for ROA, and 90% level for ROE, and were not statistically significant for Non-Current Assets.

In terms of Geographic Effects, the results were mixed with a negative impact on both ROA and ROE, and a positive impact on Non-Current Assets, each at the 99.9% level. It is quite possible that this indicates the presence of both costs stemming from Geographic Asymmetric Information - resulting in reductions in each ROA and ROE, while the geographic portfolio effects resulted in the lower overall default rates.

It should be noted that the impacts of the great recession were particularly acute on the High Acquisition Banks, as evidenced in Table 8, with Non-Current Assets increasing in magnitude 9 times (and becoming statistically significant at the 99.9% level), Average ROA

Table 8: Results Interpretation - Recession

Variable	Concept	Expected	Actual	Return on Assets (%)	Return on Equity (%)	Non-Current Assets (% of Total)
High Distance x Recession	Geographic Information	Cost	Cost	-0.19*** (0.034)	-2.4*** (0.268)	0.28*** (0.024)
	Geographic Portfolio	Benefit	Cost			
High Acquisition x Recession	Lemon	Cost		-0.46*** (0.119)	-8.3*** (0.931)	0.38*** (0.082)
				Benefit = Positive	Benefit = Positive	Benefit = Negative

nearly doubling, and average ROE increasing by over 7 times, to a -8.3 percentage point impact. The Great Recession impact on High Distance Banks was relatively less, with the magnitude of ROA approximately the same, ROE roughly doubling the impact, and, interestingly with Non-Current Assets having a negative effect, with a .28 percentage point increase due to the interaction term. This is likely because the Recession was a nationwide phenomenon, and therefore portfolio benefits were likely negated, while the Geographic Information Asymmetry costs were still realized.

While the lack of significance of the increase in Non-Current Assets during non crisis periods may signify a lack of "lemons", the large increase in magnitude during periods of stress may mitigate these doubts. In particular, the specific types of information that are often present in "soft" information, and not readily apparent through traditional due diligence, are also the type which are often marginal in nature and more often than not may fail to manifest in hard data, while they may represent a doubt to the originator. When confronted with a crisis though, those doubts may be more likely to become actualized - resulting in the much larger magnitude effect in the interaction term. As the famed investor Warren Buffet likes to say "you only find out who is swimming naked when the tide goes out" ("Chairman's Letter" 2001).

Finally, given the specifications the impact on increasing assets is not easily directly weighed against costs stemming from acquisitions and distance. It is possible that under the correct circumstances those benefits could outweigh the costs, and provide an overall net benefit.

7.2 Implications

Our findings support the existing literature and theoretical model regarding the expected presence of countervailing benefits and costs implicit in the process of expanding branch networks, and the M&A process that is often required to achieve that expansion. Prior literature tended to generate mixed results, likely due to a lack of specification allowing for adequate recognition of the full set of costs and benefits.

In particular, we find that scale, as indicated by bank `Assets`, provides benefits to institutions, most likely due to Operational Efficiency improvements. The M&A Process itself, as indicated through `High Acquisition` banks, has consistent costs, which are made especially apparent with the interaction with the Great Recession. The mixed results of the `High Distance` banks, also are indicative of the mixed nature of the variable, with Geographic Information Asymmetry Costs negatively impacting ROA and ROE, while the geographic portfolio benefits of a more distributed branch network have a positive impact on Non-Current Assets.

7.3 Limitations and Further Analysis

While this paper attempts to tease out the various costs and benefits, as evidenced by the above discussion and indications that much of the variance in performance is not fully explained, there is still room for additional analysis. In particular, determining a method for better identifying gains from management improvements, as well as methods for segmenting out portfolio effects from distance from their geographic information asymmetry costs present further areas of research opportunity.

Further robustness checks should be done to confirm the analysis. This includes testing differing specifications resulting from the determinations of the distance, and especially acquisition, flags. The models also appear to be under specified based on the low R-squared and Adjusted R-Squared Values, as well as the high AIC values, and as such there is room for

adding additionally relevant variables to the model. In particular, the addition of variables that continue to address the relevant effects of managerial improvement, as well as a more targeted approach to addressing the portfolio effects, may especially benefit the specification.

Finally, an opportunity exists to research spillover effects on other institutions, similar to the research performed in Petach, Weiler, and Totten (2024).

7.3.1 Difference-in-Differences

As noted previously in Section 5.5.1, this model deviates from the canonical DiD approach of two periods, two groups. Given this deviation, then, the question is whether the results are robust to these deviations. Both Callaway and Sant'Anna (2021) and Sant'Anna and Zhao (2020) suggest a number of additional robustness checks and alternative specifications that can be employed when employing DiD estimators over multiple treatment periods, as this paper examines. These additional checks and specifications should be performed in order to additionally validate the results.

8 Conclusion

8.1 Summary

Our findings support the existing literature and theoretical model regarding the expected presence of countervailing benefits and costs implicit in the process of expanding branch networks, and the M&A process that is often required to achieve that expansion. Prior literature tended to generate mixed results, likely due to a lack of specification allowing for adequate recognition of both costs and benefits.

In applying a transactional model, in which buyer and seller must agree on a price, this paper determines that a transaction is able to occur because both the buyer is able to generate excess value (increasing their NPV), and because of an information asymmetry that arises

where the seller has additional information about the banks portfolio that is not readily apparent to the buyer.

In particular, the analysis determines that scale generally, as indicated by bank `Assets`, provides benefits to institutions, most likely due to Operational Efficiency improvements. The M&A Process itself, as indicated through `High Acquisition` banks, has consistent costs, which are made especially during times of economic stress, as demonstrated by the interaction with the great recession. The mixed results of the `High Distance` banks, are indicative of the similarly mixed nature of the variable interpretation, with Geographic Information Asymmetry Costs negatively impacting ROA and ROE, while there is also an overall portfolio benefit demonstrated via the positive effects on Non-Current Assets.

9 Appendices

9.1 Outliers

When examining the model statistics in Section 6.3.1, it was noted that the R-squared values were quite low, and the AIC values were quite high. One potential reason for this is the presence of outliers in the data. In particular, there are some particularly large outliers in the dependent variables, as shown in Table 9. In fact, the scale of outliers is so large that the boxplots in Figure 2 needed to be presented in log scale. Looking at the values in Table 9 it is apparent some values are quite large, with Return on Equity having minimum and maximum values in the thousands, while the IQRs are in the range of 10s. This is likely to be at least part of the cause of the low R-squared values and high AIC values, given the extra variance inherent in these large values, and so it is these outliers were removed in order to examine impacts on model fit.

Outliers are removed by dropping values with a Z-score greater than 3, and the results are presented in .

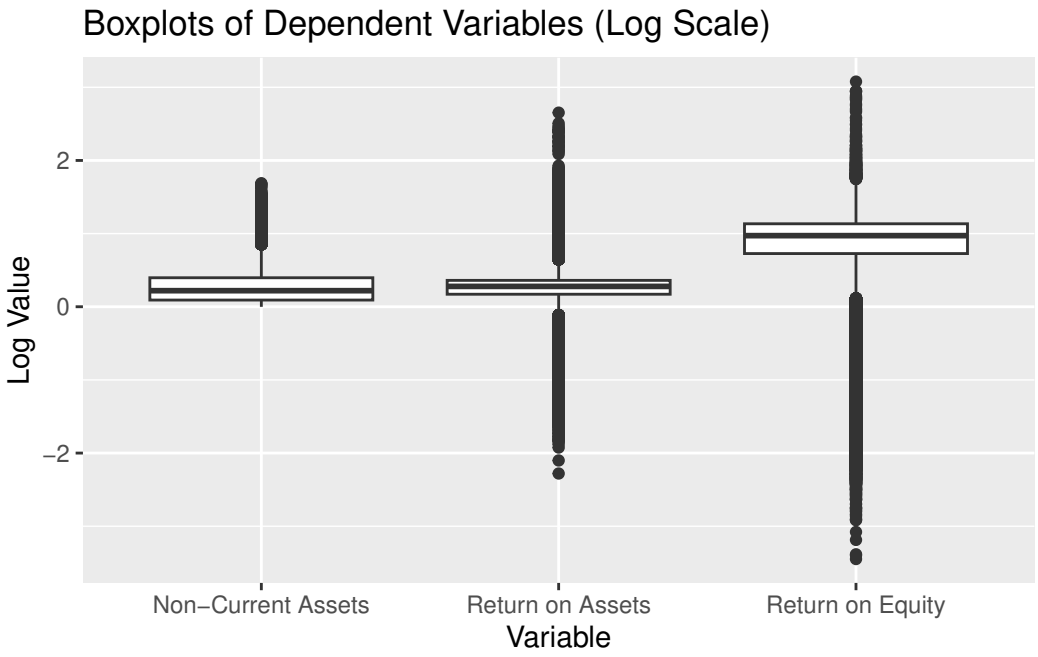

Figure 2: Outliers

Table 9: Summary Statistics of Dependent Variables

	Unique	Missing Pct.	Mean	SD	Min	Median	Max
Return on Assets (%)	9253	0	0.8	3.8	−189.4	0.9	449.3
Return on Equity (%)	30574	0	7.1	22.2	−2809.6	8.4	1195.6
Non-Current Assets (% of Total)	138204	0	1.3	2.1	0.0	0.7	47.8

Table 10 presents comparable statistics to Table 9. We find generally similar Median statistics, however, there are far more significant differences in terms of the other statistics, including Mean, Standard Deviation, Min, and Max. In particular, for Return on Equity we see a fairly large drop in Mean values, from 7.9 to 7.1, and Standard Deviation from 22.2 to 7.9. Min and Max values also drop significantly. While the changes are less large for Return on Assets and Non-Current Assets, there was a small drop in Mean Non-Current Assets from 1.3 to 1.1, as well as significant drops in Standard Deviation for both variables.

Table 11 contains regression results from both the initial regression models (including outliers) and the new results outliers removed. While significance levels remained generally similar, there was a general decreases in coefficient magnitude, which is unsurprising given the potentially large biases introduced by the extreme values.

For Log Assets we have generally similar ROA coefficients, but the marginal effect on ROE dropped from an increase of .055 percentage points to .036 percentage points for a 1% increase in assets, a decrease of around 35%. The marginal impact on Non-Current-Assets also declines to near zero, and is no longer statistically significant.

For the High Distance flag, magnitudes decreased fairly consistently in the 15 to 25% percent range for all variables, which is a lower marginal impact than Log Assets for both Return on Equity and Non-Current Assets. All three variables remained statistically significant at the 99.9% level.

For the High Acquisition Flag the marginal impact decreased more significantly for Return on Assets, by about 66%, and by around 25% for Return on Equity. Non-Current Assets remained non statistically significant.

Of note with respect to the changes in Log Assets and High Acquisition Flag is that the relatively large magnitude changes likely mean that outlier values were most likely correlated with these particular characteristics. This is especially interesting to note in the context of

Table 10: Summary Statistics of Dependent Variables with Outliers Removed

	Unique	Missing Pct.	Mean	SD	Min	Median	Max
Return on Assets (%)	8471	1	0.8	1.3	−10.6	0.9	12.1
Return on Equity (%)	29474	1	7.9	9.8	−59.3	8.4	72.9
Non-Current Assets (% of Total)	135266	2	1.1	1.3	0.0	0.6	7.7

Table 11: Regression Results comparing models with and without outliers

	Return on Assets (%)	Return on Equity (%)	Non-Current Assets (% of Total)	Return on Assets (%) No Outliers	Return on Equity (%) No Outliers	Non-Current Assets (% of Total) No Outliers
High Distance Flag	−0.193***	−1.075***	−0.104***	−0.169***	−0.845***	−0.081***
	(0.031)	(0.239)	(0.021)	(0.013)	(0.096)	(0.013)
High Distance Flag x Recession Flag	−0.189***	−2.363***	0.281***	−0.154***	−1.481***	0.246***
	(0.034)	(0.268)	(0.024)	(0.014)	(0.108)	(0.015)
High Acquisition Flag	−0.285***	−1.211**	0.028	−0.095***	−0.940***	−0.025
	(0.057)	(0.448)	(0.039)	(0.024)	(0.179)	(0.024)
Recession Flag x High Acquisition Flag	−0.459***	−8.256***	0.378***	−0.485***	−2.150***	0.219***
	(0.119)	(0.931)	(0.082)	(0.050)	(0.379)	(0.051)
Log Assets	0.559***	5.504***	−0.185***	0.531***	3.570***	0.003
	(0.022)	(0.169)	(0.015)	(0.009)	(0.068)	(0.009)
Log Distance (miles)	−0.005*	−0.018	0.006***	0.003**	0.005	0.004***
	(0.002)	(0.017)	(0.001)	(0.001)	(0.007)	(0.001)
Fee Share	3.685***	30.203***	−1.576***	2.996***	19.765***	−0.357***
	(0.089)	(0.700)	(0.062)	(0.042)	(0.323)	(0.042)
Capital/Asset Ratio	−11.182***	6.890***	−4.175***	−5.275***	−18.836***	−1.310***
	(0.159)	(1.243)	(0.109)	(0.076)	(0.507)	(0.068)
Num.Obs.	145 465	145 465	145 465	144 659	144 350	142 526
R2	0.084	0.024	0.019	0.131	0.082	0.009
R2 Adj.	0.010	−0.055	−0.060	0.060	0.007	−0.073
AIC	637 379.4	1 236 033.7	528 689.6	378 925.4	961 109.5	377 062.9
BIC	637 547.5	1 236 201.8	528 857.7	379 093.4	961 277.4	377 230.7
RMSE	2.16	16.94	1.49	0.90	6.75	0.91

+ p \num{< 0.1}, * p \num{< 0.05}, ** p \num{< 0.01}, *** p \num{< 0.001}

the High Acquisition Flag, as this a lower frequency characteristic, and also appears to be correlated with the highest levels of negative performance impacts - potentially indicating the presence of lemons.

Finally, by removing outliers model statistics generally improved, with R^2 and Adjusted R^2 improving across the ROA and ROE specifications, though declining slightly for Non-Current Assets. AIC and RMSE also decreased across the board. All of these decreases demonstrate that by removing outliers the models provide a better explanation of the remaining variance.

Agarwal, Sumit, and Robert Hauswald. 2010. "Distance and Private Information in Lending." *Review of Financial Studies* 23 (7): 2757–88. https://doi.org/10.1093/rfs/hhq001.

Agrawal, Anup, Jeffrey F. Jaffe, and Gershon N. Mandelker. 1992. "The Post-Merger Performance of Acquiring Firms: A Re-Examination of an Anomaly." *The Journal of Finance* 47 (4): 1605–21. https://doi.org/10.2307/2328956.

Akerlof, George A. 1970. "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics* 84 (3): 488–500. https://doi.org/10.2307/1879431.

Becher, David A., and Terry L. Campbell II. 2005. "Interstate Banking Deregulation and the Changing Nature of Bank Mergers." *Journal of Financial Research* 28 (1): 1–20. https://doi.org/10.1111/j.1475-6803.2005.00111.x.

Berger, Allen N, Rebecca S Demsetz, and Philip E Strahan. 1999. "The Consolidation of the Financial Services Industry: Causes, Consequences, and Implications for the Future." *Journal of Banking & Finance* 23 (2): 135–94. https://doi.org/10.1016/S0378-4266(98)00125-3.

Berger, Allen N., and Robert DeYoung. 2001. "The Effects of Geographic Expansion on Bank Efficiency." *Journal of Financial Services Research* 19 (2): 163–84. https://doi.org/10.1023/A:1011159405433.

Berger, Allen N., Anthony Saunders, Joseph M. Scalise, and Gregory F. Udell. 1997. "The Effects of Bank Mergers and Acquisitions on Small Business Lending." {{SSRN Scholarly Paper}}. Rochester, NY. https://doi.org/10.2139/ssrn.41964.

Brealey, Richard A, Ian A Cooper, and Evi Kaplanis. 2019. "The Effect of Mergers on US Bank Risk in the Short Run and in the Long Run." *Journal of Banking & Finance* 108 (November): 105660. https://doi.org/10.1016/j.jbankfin.2019.105660.

Bunten, Devin, Stephan Weiler, Eric Thompson, and Sammy Zahran. 2015. "ENTREPRENEURSHIP, INFORMATION, AND GROWTH." *Journal of Regional Science* 55 (4): 560–84. https://doi.org/10.1111/jors.12157.

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics*, Themed issue: Treatment effect 1, 225 (2): 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001.

Calzada, Joan, Xavier Fageda, and Fernando Martínez-Santos. 2023. "Mergers and Bank Branches: Two Decades of Evidence from the USA." *Empirical Economics* 64 (5): 2411–47. https://doi.org/10.1007/s00181-022-02307-4.

Carow, Kenneth A., and Randall A. Heron. 1998. "The Interstate Banking and Branching Efficiency Act of 1994: A Wealth Event for Acquisition Targets." *Journal of Banking & Finance* 22 (2): 175–96. https://doi.org/10.1016/S0378-4266(97)00052-6.

Célerier, Claire, and Adrien Matray. 2019. "Bank-Branch Supply, Financial Inclusion, and Wealth Accumulation." *The Review of Financial Studies* 32 (12): 4767–4809. https://doi.org/10.1093/rfs/hhz046.

"Chairman's Letter." 2001. https://www.berkshirehathaway.com/2001ar/2001letter.html.

Christopoulos, Apostolos G. 2012. "DO BANKING ACQUISITIONS CREATE VALUE IN AN ERA OF DEREGULATION AND INCREASING RISKS? RECENT FINDINGS IN THE US."

DeLong, Gayle. 2003. "Does Long-Term Performance of Mergers Match Market Expectations? Evidence from the US Banking Industry." *Financial Management* 32 (2): 5–25. https://doi.org/10.2307/3666334.

DeLong, Gayle, and Robert Deyoung. 2007. "Learning by Observing: Information Spillovers in the Execution and Valuation of Commercial Bank M&As." *The Journal of Finance* 62 (1): 181–216. https://doi.org/10.1111/j.1540-6261.2007.01205.x.

Deng, Saiying (Esther), and Elyas Elyasiani. 2008. "Geographic Diversification, Bank Holding Company Value, and Risk." *Journal of Money, Credit and Banking* 40 (6): 1217–38. https://doi.org/10.1111/j.1538-4616.2008.00154.x.

Emmons, William R., R. Alton Gilbert, and Timothy J. Yeager. 2004. "Reducing the Risk at Small Community Banks: Is It Size or Geographic Diversification That Matters?*."

Journal of Financial Services Research 25 (2): 259–81. https://doi.org/10.1023/B:FINA.0000020665.54596.f6.

Evanoff, Douglas D., and Evren Ors. 2008. "The Competitive Dynamics of Geographic Deregulation in Banking: Implications for Productive Efficiency." *Journal of Money, Credit and Banking* 40 (5): 897–928. https://doi.org/10.1111/j.1538-4616.2008.00141.x.

"Financial Services Modernization Act of 1999 (Gramm-Leach-Bliley) | Federal Reserve History." 2013. https://www.federalreservehistory.org/essays/gramm-leach-bliley-act.

Giedeman, Daniel C. 2004. "The Riegle-Neal Act and Local Banking Market Concentration." *International Advances in Economic Research* 10 (4): 328–36. https://doi.org/10.1007/BF02295145.

Goetz, Martin Richard. 2012. "Bank Diversification, Market Structure and Bank Risk Taking: Theory and Evidence from U.S. Commercial Banks." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2651161.

Goetz, Martin R., Luc Laeven, and Ross Levine. 2016. "Does the Geographic Expansion of Banks Reduce Risk?" *Journal of Financial Economics* 120 (2): 346–62. https://doi.org/10.1016/j.jfineco.2016.01.020.

Goetz, Martin, Luc Laeven, and Ross Levine. 2011. "The Valuation Effects of Geographic Diversification: Evidence from U.S. Banks." Working {Paper}. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w17660.

"Going Interstate: A New Dawn For U.S. Banking | St. Louis Fed." 1994. https://www.stlouisfed.org/publications/regional-economist/july-1994/going-interstate-a-new-dawn-for-us-banking.

Houston, Joel F, Christopher M James, and Michael D Ryngaert. 2001. "Where Do Merger Gains Come from? Bank Mergers from the Perspective of Insiders and Outsiders." *Journal of Financial Economics*, Complementary research methodologies: The InterPlay of theoretical, empirical and field-based research in finance, 60 (2): 285–331. https://doi.org/10.1016/S0304-405X(01)00046-0.

James, Christopher M., and Peggy Wier. 1987. "Returns to Acquirers and Competition in the Acquisition Market: The Case of Banking." *Journal of Political Economy* 95 (2): 355–70. https://doi.org/10.1086/261459.

Jang, Karen Y. 2019. "Corporate Assets and Enhancing Firm Value: Evidence from the Market for Bank Branches in the US." *Journal of Financial Services Research*, April. https://doi.org/10.1007/s10693-019-00317-5.

Leledakis, George N., and Emmanouil G. Pyrgiotakis. 2016. "As in the Post-Dodd-Frank Act Era : Do They Create Value ? By." In. https://www.semanticscholar.org/paper/As-in-the-post-Dodd-Frank-Act-era-%3A-Do-they-create-Leledakis-Pyrgiotakis/059443069e6f4db3216ec7b2bf4f3e70e674e1fd.

Low, Sarah A., and Stephan Weiler. 2012. "Employment Risk, Returns, and Entrepreneurship." *Economic Development Quarterly* 26 (3): 238–51. https://doi.org/10.1177/0891242412452445.

Markowitz, Harry. 1952. "Portfolio Selection." *The Journal of Finance* 7 (1): 77–91. https://doi.org/10.2307/2975974.

"McFadden Act of 1927 | Federal Reserve History." 2013. https://www.federalreservehistory.org/essays/mcfadden-act.

Meslier, Céline, Donald P. Morgan, Katherine Samolyk, and Amine Tarazi. 2016. "The Benefits and Costs of Geographic Diversification in Banking." *Journal of International Money and Finance* 69 (December): 287–317. https://doi.org/10.1016/j.jimonfin.2016.07.007.

Moore, Robert R., and Tom Siems. 1998. "Bank Mergers : Creating Value or Destroying Competition?, Third Quarter 1998," Financial Industry Issues, July. https://fraser.stlouisfed.org/title/financial-industry-issues-6374/bank-mergers-608584.

Nguyen, Hoai-Luu Q. 2019. "Are Credit Markets Still Local? Evidence from Bank Branch Closings." *American Economic Journal: Applied Economics* 11 (1): 1–32. https://www.jstor.org/stable/26565512.

Petach, Luke, Stephan Weiler, and Greg Totten. 2024. "Geography Matters: The Impact of Geographic Expansion on Bank Performance During the Great Recession." *Papers in Regional Science* 103 (1): 100009. https://doi.org/https://doi.org/10.1016/j.pirs.2024.100009.

Petersen, Mitchell A., and Raghuram G. Rajan. 2002. "Does Distance Still Matter? The Information Revolution in Small Business Lending." *The Journal of Finance* 57 (6): 2533–70. https://doi.org/10.1111/1540-6261.00505.

Rhoades, Stephen A. 2000. "Bank Mergers and Banking Structure in the United States, 1980-98." *Staff Studies*, Staff Studies,. https://ideas.repec.org//p/fip/fedgss/174.html.

"Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994 | Federal Reserve History." 1994. https://www.federalreservehistory.org/essays/riegle-neal-act-of-1994.

Sant'Anna, Pedro H. C., and Jun Zhao. 2020. "Doubly Robust Difference-in-Differences Estimators." *Journal of Econometrics* 219 (1): 101–22. https://doi.org/10.1016/j.jeconom.2020.06.003.

Scorsone, Eric, and Stephan Weiler. 2004. "New Markets as Informational Asymmetries." *Economic Development Quarterly* 18 (3): 303–13. https://doi.org/10.1177/0891242404265139.

Stiroh, Kevin J., and Philip E. Strahan. 2002. "Competitive Dynamics of Deregulation: Evidence from U.S. Banking." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.333621.

Weiler, Stephan. 2000. "Pioneers and Settlers in Lo-Do Denver: Private Risk and Public Benefits in Urban Redevelopment." *Urban Studies* 37 (1): 167–79. https://doi.org/10.1080/0042098002348.

Weiler, Stephan, Dana Hoag, and Chuen-mei Fan. 2006. "Prospecting for Economic Returns to Research: Adding Informational Value at the Market Fringe." *Journal of Regional Science* 46 (2): 289–311. https://doi.org/10.1111/j.0022-4146.2006.00442.x.