Shadows on Main Street: Local Financial Architecture and Asset Reallocation*

Taylor A. Begley† Amiyatosh Purnanandam‡ Virginia Traweek§

September 9, 2025

Abstract

We examine how shadow banks shape the reallocation of real assets at a local level. We first construct a novel county-level panel dataset of all financial institutions—not just banks—from Yellow Pages in Indiana, Michigan, and Ohio from 1920–2000. We document that shadow banks constitute a significant and persistent component of local financial architecture, representing over one-third of financial firms. Using changes in global corn prices as an exogenous shock to the economics of farming, we find that a higher shadow bank share in the local economy significantly enhances asset reallocation in response to economic shocks. In these counties, higher corn prices lead to greater farm consolidation, more efficient use of physical capital, and a larger increase in land values. The benefits come at the expense of higher volatility in land prices in these counties, consistent with an efficiency-volatility trade off. Our paper provides the first granular, long-run evidence that shadow banks play a key complementary role in facilitating asset reallocation.

Keywords: Banking, Shadow Banks, Allocative Efficiency, Measurement Error
JEL Classification: G21, G23, L1

*We are grateful to Paolo Pasquariello, Uday Rajan, and seminar and conference participants at the University of Michigan, the OCC-Treasury, Pontificia Universidad Catalica de Chile, the Fixed Income and Financial Institutions Conference, the Federal Reserve Monetary and Financial History Workshop, and the Lonestar Finance Conference. An earlier version of the paper was circulated under the title "Shadows on Main Street: Mapping America's Financial Architecture from the Yellow Pages of the 20th Century".

†Gatton College of Business and Economics, University of Kentucky; email: begley@uky.edu.
‡McCombs School of Business, University of Texas, Austin; email: amiyatosh@utexas.edu
§Neeley School of Business, Texas Christian University; email: v.traweek@tcu.edu

1 Introduction

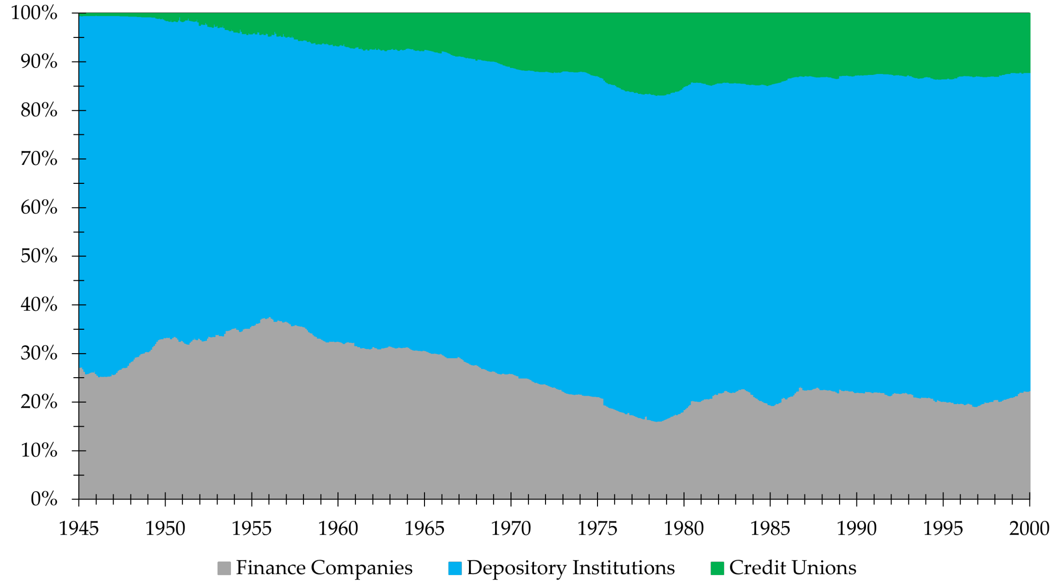
How does finance affect the allocation of productive assets across different agents in an economy? This is a foundational question for a vast literature in economics and finance at least dating back to Schumpeter (1911).Whether a study examines the effect of financial intermediaries on credit creation, compares bank- versus market-based financial systems on economic growth, or analyzes the impact of financial markets on asset valuation, the connection between finance and asset reallocation remains a key theme across the literature.[1] Yet, this vast literature has a critical gap: we know very little about how local shadow banks, such as unregulated finance companies, affect the allocative efficiency of productive assets. Our paper attempts to fill this gap in the literature.

In light of their prominent role in the global financial crisis, shadow banks are often viewed as a modern phenomenon, a recent innovation operating at the periphery of the traditional, regulated banking system. This view, however, overlooks a century of financial history. For much of the 20th century, a vibrant sector of unregulated finance companies operated on America's Main Streets, providing credit for everything from automobiles to farm equipment. What we miss by not studying these shadow banks is economically large: based on aggregate data, Figure 1 shows nonbank finance companies consistently provided between 16% and 38% of all consumer credit in the post-war era. Because researchers typically lack systematic data on nonbank financial institutions, this sector has unfortunately remained largely invisible in empirical research.

Lack of studies linking local shadow banks to asset allocation, and more broadly to other real and financial outcomes, has important implications for academic studies and policy initiatives. First, a vast literature studying the finance-growth nexus has often used data

[1] The literature is rich both on empirical and theoretical fronts (e.g., Leland and Pyle, 1977; Ramakrishnan and Thakor, 1984; King and Levine, 1993; Boot and Thakor, 1997; Jayaratne and Strahan, 1996; Rajan and Zingales, 1998; Subrahmanyam and Titman, 1999; Allen and Gale, 2000; Levine, 2002; Allen and Gale, 2004; Levine, 2005; Chen, Goldstein, and Jiang, 2007; Diamond and Rajan, 2000; Butler and Cornaggia, 2011; Rajan and Ramcharan, 2015, and several others).

Figure 1: The Supply of Consumer Credit in the U.S.
This figure presents the provisioning of consumer credit in the U.S. from 1945–2000 broken down across depository institutions, finance companies, and credit unions. The data are provided by the Federal Reserve Board's *G-19: Consumer Credit* files, which can be found at `https://www.federalreserve.gov/releases/g19/HIST/cc_hist_mh_levels.html`. The figure excludes consumer credit outstanding held by the federal government, nonprofit and educational institutions, and non-financial businesses.



from regulated institutions, such as banks, as a primary proxy for financial development (e.g., King and Levine, 1993; Jayaratne and Strahan, 1996; Levine, 2005, and several others). If shadow banks are an economically significant component of the financial system, then using banks alone to proxy for development can lead to systematic biases which can cloud our understanding of how finance truly impacts growth. Second, from a theoretical standpoint, banks and shadow banks are different types of intermediaries. Traditional banks are funded by sticky deposits and constrained by regulation, whereas shadow banks are primarily financed by market-based sourced of funds and operate with greater regulatory flexibility. Consequently, these institutions are likely to differ in their ability to write financial contracts, share risk, and allocate resources to different types of projects. Shadow banks may be better positioned to quickly react to price signals or technological changes and facilitate the efficient reallocation of capital when the need arises. As Dow and Gorton (1997) write "In a capitalist economy, prices serve to equilibrate supply and demand for goods and services, continually changing to reallocate resources to their most efficient uses." Whether and to what extent such a reallocation happens is an object of primitive interest in a wide range of theoretical models (e.g., Shleifer and Vishny, 1992; Dow and Gorton, 1997; Kiyotaki and Moore, 1997; Diamond and Rajan, 2000; Geanakoplos, 2010, and several others).

Data limitation, specifically lack of granular data on the composition of traditional and shadow banks at a local level, has been a primary reason for this gap in the literature. Our first contribution is to construct a novel, hand-collected dataset that provides the first granular, long-run map of the financial architecture for a large cross-section of the United States counties. We collect data on all financial service providers from the Yellow Pages directories for 171 counties across Indiana, Michigan, and Ohio, sampled at five-year intervals from 1920 to 2000. These directories were the most comprehensive source of information for local businesses in the pre-internet era.[2] These unique data allow us to identify and geocode over 17,000 branches from 5,700 unique financial institutions, distinguishing between regulated banks, thrifts, credit unions, and unregulated "finance companies," which we refer to as shadow banks. This massive data collection effort yields a unique panel that allows us to provide new facts on the level, mix, and evolution of financial architecture at a highly localized level for over 80 years.

Our empirical strategy exploits this rich dataset to test how local financial architecture impacts the reallocation of farmland. Farmland is a key productive asset whose financing, use, and value are inherently local, particularly during our sample. We use plausibly exogenous variation in global corn prices as a shock to the profitability of farming and consequently the need for reallocation. A rise in corn prices increases the incentive for capital investment such as replacing old tractors with newer and more efficient ones or installing a new irrigation systems as in the real option model of Dixit and Pindyck (1994). This need for capital can make consolidation of farmland among the most efficient farmers the optimal outcome. By

[2] "Perusal of telephone book 'yellow pages' of most cities discloses many competing credit grantors—commercial banks, finance companies, savings and loan associations, industrial banks and loan companies, credit unions, savings banks, and retailers" (National Commission on Consumer Finance, 1972, p.11).

offering contracts tied to changing market conditions and bringing in funds from market-based sources, shadow banks are well-positioned to facilitate such reallocations. For example, shadow banks are more likely to provide installment loans or lease financing for moveable assets like tractors and cashflows based loans for the purchase of raw materials like seeds and fertilizer than a traditional banks that face stricter regulation and are constrained in their sources of funds to local deposits. Therefore, shadow banks are likely to expand the set of contracts that can be written between the borrowers and the savers, changing the nature of resource allocation in the system (Allen and Gale, 2004).

Since we bring a novel dataset to the literature, the first part of our paper is descriptive in nature, presenting some new facts about U.S. financial architecture. Shadow banks have been an economically significant and persistent component of local finance for the entire 20th century, constituting, on average, over one-third of all financial institutions. This local financial architecture is remarkably persistent: a county's composition of banks and shadow banks is a strong predictor of its structure even 50 years later. When considering their interaction, we find that the presence of regulated banks and shadow banks is strongly positively correlated at the county level, suggesting they act as complements, not substitutes. This finding has important empirical implication for a wide range of studies: omitting shadow banks from empirical analysis can introduce significant non-classical measurement error bias and omitted variable bias depending on the application at hand.

We next study our main question of interest: How does the local mix of shadow and traditional banks shape asset reallocation following fundamental shocks? We employ a triple-difference research design, comparing the effect of shadow banks on outcomes in corn-producing versus non-corn-producing counties before and after corn price shocks. The key explanatory variable is the interaction of this shock with the lagged share of shadow banks in a county's financial system. This design allows us to isolate the role of shadow bank presence on the sensitivity of asset reallocation to a fundamental economic shock, while controlling for time-invariant county characteristics and aggregate macroeconomic trends. Specifically, our research design allows us to separate out the unconditional effect of corn price change on corn-producing farmlands as well as the effect of other macroeconomic or regulatory shocks on the agriculture sector.

We find that, in corn-producing counties with a higher initial share of shadow banks, a positive shock to corn prices leads to a significantly larger degree of farm consolidation. A one-standard-deviation increase in corn-price returns in a county at the 75th percentile of shadow bank presence results in a 9.2% larger reduction in the number of farms compared to a county at the 25th percentile. This consolidation is accompanied by a corresponding increase in average farm size. Our next tests examine whether this reallocation allows the owners to exploit economies of scale from larger farm size. We collect county-level data on number of tractors, a key capital input into farming, to test this conjecture. Following a corn price shock, high shadow bank counties experience a 4% increase in the number of tractors per farm. These results are consistent with farms leveraging economies of scale: farm consolidation allowing for a more efficient use of physical capital. We also investigate whether our effects are symmetrically distributed across rising and falling corn prices. We find the results are primarily concentrated within periods of rising corn prices, consistent with the idea that shadow banks facilitate reallocation when the need for capital investments is high.

Is the reallocation of assets efficient? Absent precise data on all the inputs and outputs at the farm level, our study cannot directly compute the profitability or efficiency of the farm's production function. However, we do observe the market value of land in these counties, which allows us to shed light on this question. Specifically, market value of land should increase in response to asset reallocation if assets move towards their efficient users. This is what we find: moving from 25th to the 75th percentile in its shadow bank ratio (an increase of 0.36) is associated with a 6.3% larger increase in land value per acre following a one-standard-deviation increase in corn price returns. We find no corresponding effect on crop yields, suggesting these efficiency gains stem primarily from cost savings and economies of scale.

While our results highlight the benefits of shadow banks in facilitating efficient reallocation, this channel may also introduce costs. Specifically, increased reallocation may be accompanied by higher asset price volatility as is the case with several models of financial development (Allen and Gale, 2000, 2004). We find some support for this trade off. Counties with higher presence of shadow banks, one average, show higher volatility in land prices.

In the final part of the paper, we provide direct guidance for future research by quantifying the bias from mismeasuring financial development. We show that omitting shadow banks creates a non-classical measurement error. Empirical studies using banks-per-capita in levels *overestimate* the true effect of financial development by 28% to 72%. In contrast, studies using the log of banks-per-capita *underestimate* the effect by 13% to 35%. This bias varies systematically across urban and rural counties, carrying significant implications for policy analyses that rely on these potentially flawed estimates. We also show that the presence of shadow banks can, in many circumstance, cause significant overestimates of banks on real outcomes because of omitted variable bias.

This paper makes several contributions. We contribute to the literature on financial intermediation and economic growth (e.g., Jayaratne and Strahan, 1996; Levine, 2005). Our work is related to Butler and Cornaggia (2011) who show that access to bank finance improves productivity of farming. Our results inform our understanding of U.S. financial development by creating a new dataset, showing that the shadow banking sector was a core component, not sideshow. Our results show that the common practice of using regulated banks as a proxy for the entire financial sector is potentially problematic and provide a quantitative framework to assess the resulting biases. Second, we contribute to a broad literature on asset allocation and the real effects of finance by providing new evidence on how intermediaries facilitate reallocation (e.g., Leland and Pyle, 1977; Ramakrishnan and Thakor, 1984; Boot and Thakor, 1997; Subrahmanyam and Titman, 1999; Chen, Goldstein, and Jiang, 2007). By documenting the distinct role of shadow banks, we also speak to the literature comparing bank-based and market-based financial systems (Allen and Gale, 2000; Levine, 2002). Finally, our work has implications for competition policy. By showing that standard, banks-only measures significantly understate the true level of local market competition, our findings are relevant for bank merger analysis and antitrust policy.

2 Data Collection & Sample Construction

Most existing studies in banking obtain measures of financial development from regulatory filings and databases maintained by banking supervisors such as the FDIC and Federal Reserve Banks. In light of banking regulators' mission and scope, these databases focus on regulated institutions such as commercial banks and, to a lesser extent, on institutions such as S&Ls. Even for the regulated banks, there is lack of data for a long time-series at a granular geographical level. For example, in the United States, one can obtain detailed branch-level location data for banks only after 1980s from the FDIC's bank branch statistics. Similarly, one can obtain data on national banks from the compilations of the Comptroller of the Currency. But these reports only include national banks, thereby omitting state banks, credit unions, S&Ls, and finance companies. Further, even for national banks, the Comptroller reports do not include the location or number of bank branches, which are important for understanding local access to financial services.

We take a novel approach for our data collection strategy. Instead of looking at regulatory databases and archives, we obtain information on financial institutions from the city and county directories, a source of data that has been overlooked by researchers. These directories were printed for major cities by telephone and telegraph companies starting in the mid-1800s, but with regularity for most cities starting around 1920. By the 1920s, a handful of major companies such as Polk's and Robinson's printed the majority of phone books for large cities, and, as such, the format of phone books quickly became standardized across the nation. Indeed, in all but the smallest cities, phone books followed roughly the same format from the early 1900s to the 2000s when phone book printing was largely replaced by the internet. A typical phone book contains multiple sections: the white pages (which include listings for phone numbers and addresses of individuals), the yellow pages (which include listings for businesses by category), reverse address look-up sections, and occasionally other sections such as important municipal phone numbers and local maps. Our data comes from the yellow pages section of these books.

In the pre-internet era, these directories were printed and widely distributed to local citizens and businesses every year. They were the primary mechanism via which consumers would find contact information for local merchants or other local individuals. Each book covers a specific geographic region — usually a specific city or metropolitan area. Many phone books also contain information about "rural routes," which include smaller municipalities outside the main municipal coverage area of the phone book. Other phone books cover the entire county or multiple cities and suburbs within the county. The fact that the majority of large cities are covered in our sample and that the overwhelming majority of directories include rural routes and adjacent suburbs means that our sample is comprehensive for the geographic areas and years in which we have coverage.

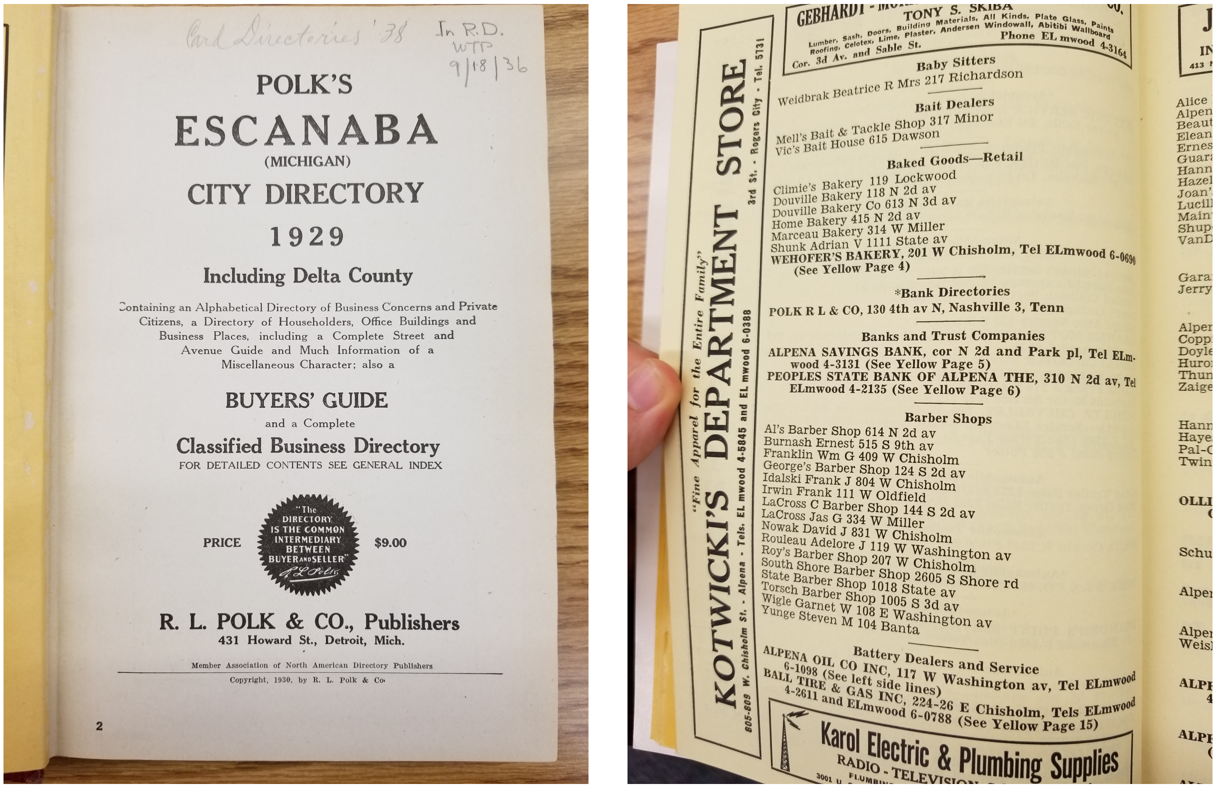
These directories have long been of interest to historians and genealogists who use the white pages to find the address for deceased individuals. As as result, city directories have been collected by genealogy libraries and state historical libraries to help facilitate these searches. Many state historical libraries offer their collections free of charge to the public on an open access basis. Other state historical libraries require advance notice in order to find books from their restricted shelving system. Many of these directories are either digitized or are available in print form from state historical libraries.

We obtain these data from all the various sources mentioned above for three states: Indiana, Michigan, and Ohio. We begin the exercise by obtaining all city directories and phone books from the Library of Michigan, the Indiana State Library, and the Ohio History Connection Library. We augment this collection with internet-based sources of phone books, most notably Ancestry.com, which is the largest repository of phone books in digital format in the county. Finally, we supplemented the collection of Detroit-area phone book data with the collection from the Royal Oak Public Library, which contains a more extensive collection of Detroit-area directories than the state Library of Michigan.

The phone book section of interest to us is the "yellow pages", which are, as the name suggests, often printed on yellow paper. This section lists all businesses in a specific town or county by the service they provide, with special advertising sections available for purchase by these businesses. The vast majority of directories contain specific sections that list businesses in rough categories such as "plumbing," "churches," or "banks." We obtain the precise location of all businesses that are present under the categories of banks, loans, and all such category that relate to financial services from these pages. Since there were no uniform listing requirement for businesses, we photograph specific entries across all books so as to maintain continuity. These entries include the following words: bank, collections, credit/credit unions/credit bureaus, finance companies, loans (and any subcategories listed after loans), savings/savings and loan, and trust/trust companies. These headings are fairly common across all books, although there are fewer headings as one looks further back in time. For instance, most directories from before 1930 feature entries only under the headings "banks" and "loan companies." Further, as we get closer to modern times, headings often include additional information that is of interest to us. For instance, in many cities, there are multiple categories of loans listed. It is not uncommon for a phone book to have "loans - auto," "loans - home/mortgage," and "loans - farm." We preserve this information in our data collection exercise. From these raw data, we can identify the branch networks of banks, credit unions, savings and loans (also called "building and loans"), and finance companies. The result is an extensive picture of financial services offered across 80 years, starting in 1920 and ending in 2000.

Figure 2(a) shows the cover page and one of the inner pages in our phone books. The cover page contains information related to the city/county and year that the book covered.

Figure 2: Example of Raw Phonebook Data
Panel (a) of this figure shows an example phone book cover page, which contains information regarding the type of directory, the publisher, publication date, and the location of the phone book and surrounding areas included in the book. Panel (b) shows an example of the raw data from the pages of a phonebook. We photograph each page containing entries that fall under the following categories: banks, collections, credit unions, credit bureaus, finance companies, loans, savings and loans, and trusts. Section 2 details the sample collection process.



(a) Sample Phone Book Cover Page **(b)** Sample Phone Book Data

It also includes smaller cities which are covered in the directory in addition to the larger municipality. Figure 2(b) shows a typical entry in the yellow pages. The type of business is listed in bold ("Banks and Trust Companies") with the respective companies listed below.

We photograph phone books at five-year intervals for every town or county in which a directory is available from approximately 1920 to 2000. We chose this range for a multiple reasons. First, 1920 is the approximate year in which phone book coverage appears to have become widespread, with the early 1900s and late 1800s not having sufficient coverage to make a panel. Second, there are gaps in coverage on an annual basis, so photographing at five-year intervals creates a more balanced panel. Third, the financial architecture is somewhat slow moving with relatively smaller changes from year to year. We stop in 2000 when the internet began to replace city directories.

Once we photograph or save digital copies of the relevant city directory pages, we transcribe the relevant entries into a spreadsheet. Each line of our spreadsheet contains the county, the state, the image number, the financial institution name, the address(es) of the listed branches, zip code, and the categories under which the financial institution is listed.

We then geocode every entry in our sample using the Google Maps API. Automatic geocoding is successful for most entries, but some instances are older bank branches where the address is no longer valid. For these, we select a zip code that is most likely to be associated with each particular entry and manually inspect records and enter county and zip code for entries that are clearly incorrect. Occasionally a bank branch will appear in two phone books for different cities. This usually occurs when cities have overlapping geographic coverage areas, and we remove these duplicates.

We also link branches and banks in the cross-section and across time. To do this, we follow a format similar to the modern call report format in that each parent bank has a unique identifier, and each branch bank has a unique identifier. We match branches based on address and name, and we match parent banks based on name and city or county. The result is a comprehensive panel of financial institution availability that allows us to examine financial access through time.

3 Empirical Analysis

In this section, we present several new facts about the nature of the local composition of supply of finance, which we refer to as "financial architecture." We show that the presence of nonbank financial institutions is (1) economically significant, (2) persistent over time, (3) very hetergeneous by county and time, and (4) are strongly positively correlated with bank

Table 1: Summary Statistics

This table presents sample summary statistics. *Total Financial Institutions (Br)* represents the total number of branches of all types of financial institutions in a county. *Total Financial Institutions (Firms)* represents the total number of unique financial institutions in a county. We follow the same nomenclature, representing statistics for the number of branches with *Br* in the parenthesis and statistics for the number for unique financial institutions with *Firm* in the parenthesis. *Reg'd Nonbanks* are regulated financial institutions that are not banks and include Savings/Building & Loans and Credit Unions. *Finance Cos* are unregulated institutions, i.e., finance companies. We present the raw numbers of these institution for every county-period pair as well as their scaled values for every 10,000 persons in the county. The scaled variables are represented with *Financial Institutions (Br)/10K* and so on. *Shadow Ratio* is the number of finance companies as a share of total financial institutions.

	mean	sd	min	p25	p50	p75	max	count
Population (000s)	149.59	280.57	2.74	33.93	62.41	131.99	2670.37	1,478
Total Financial Institutions (Br)	39.23	61.03	1.00	10.00	17.50	42.00	507.00	1,478
Total Financial Institutions (Firms)	24.45	33.93	1.00	8.00	14.00	27.00	320.00	1,478
Banks (Br)	17.02	26.80	0.00	4.00	8.00	18.00	292.00	1,478
Banks (Firms)	6.02	5.62	0.00	3.00	4.00	7.00	69.00	1,478
Reg'd Nonbanks (Br)	9.51	21.55	0.00	1.00	3.00	9.00	234.00	1,478
Reg'd Nonbanks (Firms)	7.14	16.97	0.00	1.00	3.00	7.00	226.00	1,478
Finance Cos (Br)	12.70	22.40	0.00	2.00	6.00	13.00	206.00	1,478
Finance Cos (Firms)	11.29	18.08	0.00	2.00	5.00	13.00	186.00	1,478
Financial Institutions (Br) / 10k	3.26	1.73	0.15	1.97	3.10	4.36	8.18	1,478
Financial Institutions (Firms) / 10k	2.39	1.25	0.12	1.52	2.27	3.12	6.61	1,478
Banks (Br) / 10k	1.56	1.19	0.04	0.67	1.23	2.18	5.56	1,478
Banks (Firms) / 10k	0.84	0.67	0.03	0.37	0.66	1.13	3.37	1,478
Reg'd Nonbanks (Br) / 10k	0.67	0.59	0.00	0.24	0.52	0.95	2.66	1,478
Reg'd Nonbanks (Firms) / 10k	0.55	0.47	0.00	0.21	0.47	0.80	2.28	1,478
Finance Companies (Br) / 10k	1.02	0.80	0.00	0.41	0.89	1.48	3.86	1,478
Finance Companies (Firms) / 10k	0.98	0.78	0.00	0.40	0.85	1.38	3.83	1,478
Shadow Ratio (Br)	0.33	0.22	0.00	0.14	0.33	0.50	1.00	1,478
Shadow Ratio (Firms)	0.39	0.22	0.00	0.24	0.41	0.56	1.00	1,478

presence.

3.1 Summary Statistics

In our main sample, we have an unbalanced panel of 171 unique counties from Indiana, Michigan, and Ohio over period from 1920 to 2000, where we measure time in five year increments. Table 1 presents our summary statistics. For each of our variables, we report their value when measured at the branch level (Br) and the institution level ($Firm$). The counties in our sample have an average (median) of about 39 (18) branches coming from nearly 24 (14) unique financial institutions.

The average county has 17 branches of regulated banks, 9.5 of regulated nonbanks, and

12.7 of shadow banks (finance companies). Therefore, regulated banks are the most common organizational form of the financial service providers in the U.S., but our data show that shadow banks also have a prominent presence. About 33% of all the branches of financial institutions come from shadow banks, as represented by the variable *Shadow Ratio* in the table.

Shadow banks cover a wide range of lending arrangements, but installment loans backing retail sales of consumer goods form an important part of their portfolio (National Commission on Consumer Finance, 1972). This lending spans the spectrum of consumer credit including home mortgages, farm lending, auto lending, loans for furniture and household goods, lending against future salary (payday lending), and other unsecured lending. Prominent names in our dataset are Household Finance Corp (eventually HSBC), Beneficial Finance Corp, and American General. Some of finance companies' growth in the early 20th century was propelled by financing the receivables of durable goods retailers (e.g., pianos, sewing machines, washing machines, etc.) and their aggressive supply of automobile credit, which banks were hesitant to supply (Nugent, 1939).[3] While there is much overlap in the types of customers served and products offered, finance companies have been usually willing to accept a higher average degree of risk than commercial banks (National Commission on Consumer Finance, 1972, p.115).

In terms of the number of unique firms in a county, which is commonly a measure of the competitiveness of the local market, we find that there is an average of 24 unique financial firms per county in our sample. The respective numbers for regulated banks, regulated nonbanks, and shadow banks are are 6, 7, and 11. Comparing the number of branches to the number of unique institutions, it is clear that banks typically operate multiple branches in a county (17 branches for 6 unique banks), whereas shadow banks are often a one-branch

[3] "Banks were extremely wary of loans secured by a commodity whose location could be shifted so readily and they looked askance at long-term credits for the purchase of such luxuries. The instalment finance company stepped into the breach between the consumer who wished to buy an automobile on partial payments and the dealer who wanted to make the sale but lacked the capital to finance it" (Nugent, 1939, p.80)

operation (12.70 branches for 11 unique firms). Therefore, the fraction of shadow banks in terms of unique institutions, is higher at 39%, compared to the corresponding ratio of 33% based on the number of branches.

3.2 Financial Architecture Over Time

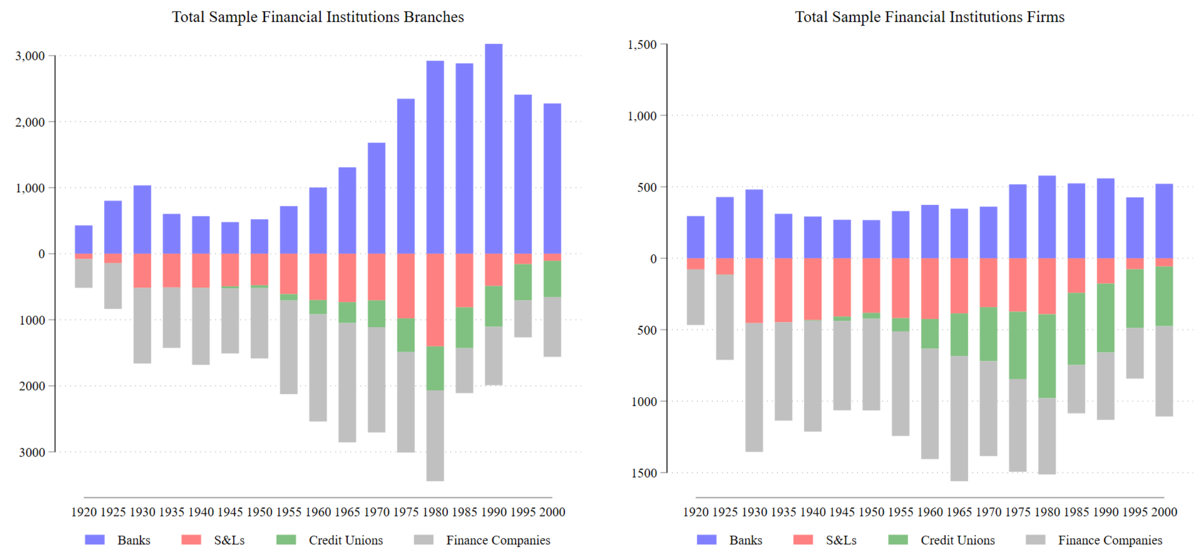
Figure 3 plots the counts of banks, S&Ls, credit unions, and finance companies throughout our sample.[4] The left panel shows the total number of branches of each category of institution for each year in the sample, and the right panel is the total number unique financial institutions (firms). In 1920, the first year of our sample, we have 429 bank branches, 80 building & loan branches, and 436 finance company branches in the sample. They come from 295, 78, and 389 unique banking, building & loan, and finance company firms. By 1980, we have 2,921 (578) branches (unique institutions) of banks, 1,403 (392) building & loans, 666 (587) credit unions, and 1,375 (534) finance companies, respectively. It is clear from these two figures that the financial architecture has changed considerably over time, but the shadow banks have remained an important part of the financial system in America over the past 100 years. It is also clear from the figure that the branches of overall financial institutions increased considerably between 1950 and 1980, after which it declined. The increase in the number of branches between 1950–1980 came from the growth in regulated institutions, both banks and nonbanks.

Figure 4 plots the *Shadow Ratio (branch)* (i.e., the number of branches of finance companies to the total number of branches of any type of financial institution), over the sample period for each of the three states in our sample.[5] Two striking patterns emerge. First, there is a strong time-specific element to the financial architecture: all three states experience an

[4]Our panel is not perfectly balanced because of data availability since we miss some city directories for some periods, but we find identical patterns when requiring a more balanced panel. Appendix Figure A.1 presents our sample coverage over time.

[5]The pattern for the ratio computed according to firm counts *Shadow Ratio (firm)* is similar, but shifted slightly upward since banks typically have more branches than nonbanks.

Figure 3: Financial Institutions Over Time
This figure presents the total number of unique branches (left panel) and institutions (right panel) in our sample. Section 2 details the sample collection process.



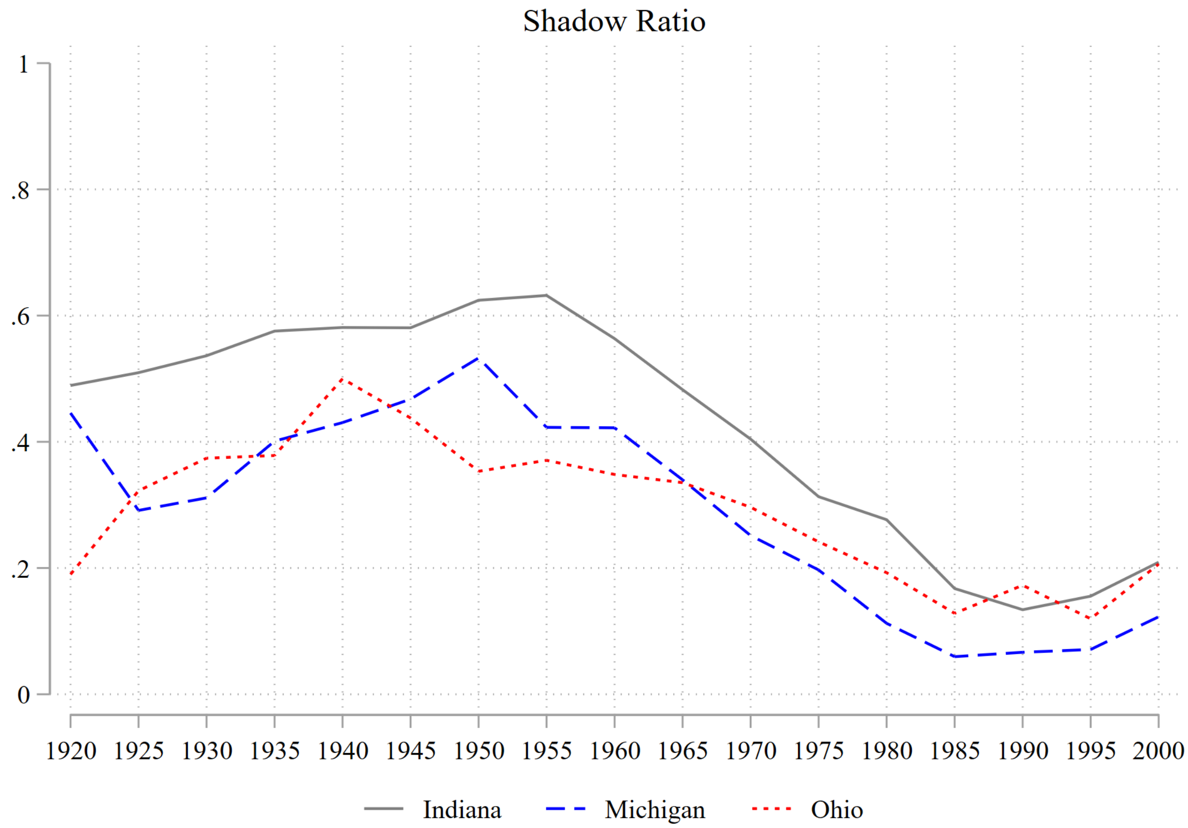
increase in the share of shadow banks in the initial three decades, and all of them show a decline thereafter before flattening out by the modern period late in the sample. Second, there is substantial and persistent cross-sectional heterogeneity at the state level. As we will show next, cross sectional differences are most pronounced at the local county level. Consequently, the effect of financial development on real outcomes can be different depending on the sample period or the geography used by the analysis.

3.3 Explaining Variation in Financial Architecture

What explains the variation in financial system architecture? While Figures 3 and 4 suggest strong time-series and state-specific components to the evolution of financial architecture, we now undertake a more systematic approach to examining the sources of this variation. We estimate the following regression model for observations in county c, state s, at time t with

Figure 4: Shadow Ratio Over Time

This figure presents the average county-level ratio of finance company branches to total financial institution branches (*Shadow Ratio*) for each state in our sample.



different sets of fixed effects. The adjusted R-squared (R_a^2) is the statistic of primary interest.

$$y_{ct} = \alpha + \beta \log(population_{ct}) + \Gamma(Fixed\ effects)_{sct} + \epsilon_{ct} \tag{1}$$

The dependent variable y_{ct} represents either the logged value of the number of branches of financial institutions in county c in year t, or the *Shadow Ratio*. We include $\log(population_{it})$ in all specification to account for baseline heterogeneity in the demand for banking services. We consider several combinations of fixed effects to shed light on which dimensions have the most explanatory power.

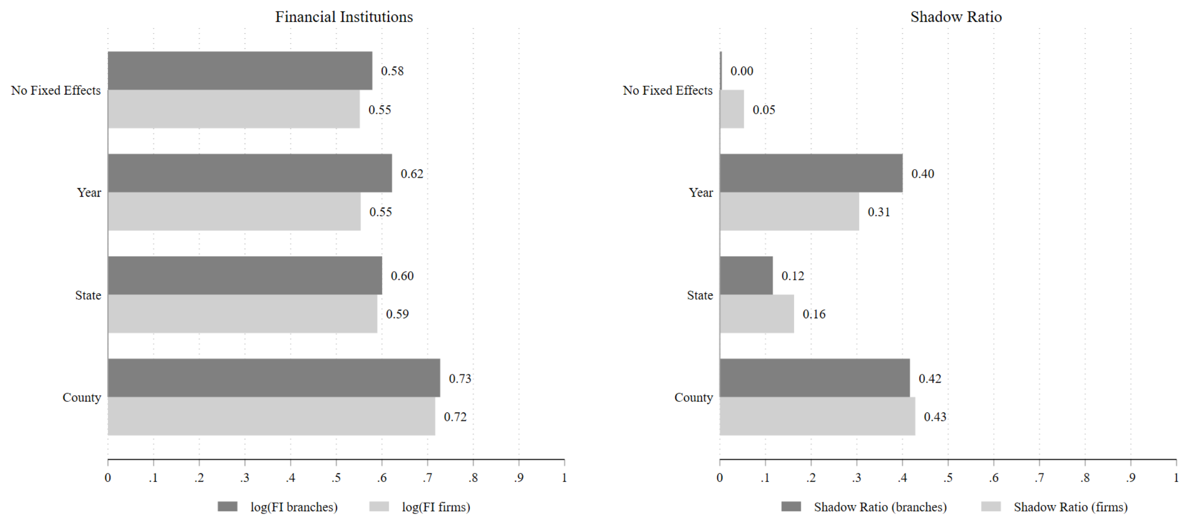
Figure 5 presents the results. Each bar in the figure presents the adjusted R^2 from the regression model (1), with the inclusion of the fixed effects that are noted on the y-axis. For example, without the inclusion of any fixed effect in the model, we obtain an adjusted R^2 of 58% for the number of branches and almost 0% for the *Shadow Ratio*. Said differently, the population of a county is a key predictor of the depth (i.e., number) of financial institutions, but not the mix of these institutions between shadow versus regulated institutions.

The left panel shows that county-level heterogeneity plays an important role in explaining the number of local financial institutions. Including state or year fixed effects yields only a modest increase in R_a^2 from 0.58 to 0.60 and 0.62. However, including county fixed effects yields a sizable increase in the R_a^2 to 73%.

The right panel studies the *Shadow Ratio* and shows that including year fixed effects alone significantly improves the model fit from 0% to 40%. Time-series variation, therefore, is a significant driver of the *Shadow Ratio*, as is shown visually in the earlier figures. The county-specific variation is also a substantially stronger predictor of the *Shadow Ratio* than the state-specific variation, increasing the R_a^2 from 12% for the model with only state fixed effects to 42% for the model with only county fixed effects. These results show that there is substantial county-specific variation in the financial architecture of the country. This, in turn, can have significant implications for empirical results that solely rely on state-level variation in banking deregulations to tease out the effect of financial development on real outcomes.

We further show the importance of this county-specific effect by providing an additional test highlighting the persistence of financial architecture. Here we examine the degree to which current financial depth (number of financial institutions, *FIs*) and *Shadow Ratio* predicts future financial architecture. Specifically, we regress the current measure of number of FIs (or the Shadow Ratio) on lagged number of FIs (Shadow Ratio) at a gap of up to 50 years. In these regressions, we control for time fixed effects, the lagged population, and the current population. Including the current population in the specification helps account for interim changes in local economic conditions, allowing us to isolate the persistence of the initial local conditions on future financial architecture after accounting for changes in the

Figure 5: Explaining Financial Architecture
This figure presents the adjusted R-squared from regressing the local financial architecture on the log of local population and various combinations of fixed effects (Equation 1). Financial architecture is measured as the log number of financial institutions in the left panel and the *Shadow Ratio* in the right panel. *Shadow Ratio* is the number of finance companies as a share of total financial institutions. The darker bars represents outcomes measured at the branch level, and the lighter bars represent outcomes measured at the firm level.



baseline demand for financial services.

Table 2 presents results. Columns (1)-(4) use the log of the number of financial institution branches as the dependent variable. At a lag of 20 years, we estimate an elasticity of 0.37 between the current and past financial depth, which gradually declines but remains at a significant and positive value of 0.14 even after 50 years. Columns (5)-(8) present the corresponding regression estimates for shadow bank ratio. The ratio is equally persistent. In other words, cross-sectional differences in financial depth as well as the mix of regulated versus non-regulated institutions have remained remarkably stable over a long period of time in the country. To the extent that finance matters for economic growth, these results suggest stickiness in financial architecture can be a crucial driver of differences in economic growth across different parts of the country, such as rural versus urban or rich versus poor counties.

Table 2: Persistence in Financial Architecture

This table present regressions of *log(Financial Institution Branches)* (columns 1-4) and *Shadow Ratio* (columns 5-8) on their lagged values, year fixed effects, and population. In columns (1)-(4), the main independent variable is *log(Financial Institution Branches)* lagged at 10, 20, 35, and 50 years, respectively. For each column, log(population) at the same lagged distance is included (denoted *LX.log(population)*) as well as contemporaneous population. Columns (5)-(8) are estimated similarly. *Shadow Ratio (Br)* is the number of finance company branches as a share of total financial institutions. All regressions include year fixed effects and present standard errors clustered at the county level.

	log(Financial Institution Branches)				Shadow Ratio			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
L10.log(FI Branches)	0.48***							
	(<0.01)							
L20.log(FI Branches)		0.37***						
		(<0.01)						
L35.log(FI Branches)			0.20***					
			(<0.01)					
L50.log(FI Branches)				0.14**				
				(0.04)				
L10.Shadow Ratio (Br)					0.48***			
					(<0.01)			
L20.Shadow Ratio (Br)						0.35***		
						(<0.01)		
L35.Shadow Ratio (Br)							0.21***	
							(<0.01)	
L50.Shadow Ratio (Br)								0.15***
								(<0.01)
log(population)	0.27**	0.37**	0.51***	0.58***	-0.01	-0.00	-0.03	-0.01
	(0.05)	(0.03)	(<0.01)	(<0.01)	(0.89)	(0.98)	(0.50)	(0.73)
L*X*.log(population)	0.14	0.10	0.13	0.10	0.01	0.01	0.04	0.01
	(0.33)	(0.57)	(0.45)	(0.58)	(0.75)	(0.84)	(0.42)	(0.78)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	982	781	543	323	982	781	543	323
Adjusted R-squared	0.70	0.64	0.61	0.59	0.61	0.56	0.51	0.25

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.4 The interplay between banks and nonbanks

How does the growth of the formal banking sector relate to the growth of nonbank institutions in the same area? It has been often argued that shadow banks fill in the gap left by regulated financial institutions as a result of relative differences in the technology, regulation, or cost structure.[6] In an early study, Nugent (1939) writes "The limited expansion of accommodation loans by banks was of the banks' own choosing. Although many country banks increased their consumer loan balances as the demand for such loans increased, many urban banks were convinced by cost-accounting studies of the unprofitableness of small personal loans and restricted them so far as possible." The willingness and ability of regulated banking institutions to expand in an area can directly influence the presence of shadow banks but the direction of the impact is unclear. Do they tend to be positively related to each other (more like complements), negatively relate to each other (more like substitutes), or are they independent?

Theoretically, banks and nonbanks could be substitutes, complements, or completely independent of one another. For example, an oft-cited narrative against tighter bank regulation regarding certain activities is that it drives those activities "into the shadows." On the other hand, increasing demand for financial services in an area can result in a simultaneous increase in both these types of institutions, especially to the extent that they may provide complementary services. Given the theoretical ambiguity, this is ultimately an empirical question. The answer to this question has significant implications for potential biases present in studies that only consider regulated banks as a proxy for financial development, as we discuss in detail in the next section. To examine this, we regress the number of banks on the contemporaneous number of nonbanks using the following regression equation:

$$\log(Nonbanks)_{cst} = \rho \log(Banks_{cst}) + \beta \log(population_{cst}) + \gamma_c + \tau_{st} + \epsilon_{cst} \tag{2}$$

The regressions control for population, county fixed effects γ_c, and state $\times$ year fixed effects τ_{st}, each of which we have shown are important drivers of local financial architecture. The coefficient of interest is ρ which is the elasticity of nonbanks to banks. We consider three variables for the measurement of nonbanks in these regressions: (a) all nonbank institutions

[6] For example, Buchak, Matvos, Piskorski, and Seru (2018) and Begley and Srinivasan (2022) show such that such differences influence the industrial organization of the modern mortgage market.

that includes regulated nonbanks such as S&Ls and shadow banks, (b) only the regulated nonbanks, and (c) only the shadow banks (finance companies). For each measure, we present estimates both for the number of branches and number of unique firms.

Table 3 presents the results. Column (1) shows a strong positive relationship between branches of all nonbank institutions and bank branches. The coefficient of 0.56 (p-value<0.01) indicates that a 10% increase in nonbank branches corresponds to approximately 5.6% increase in bank branches. Column (2) produces similar results when considering the number of unique firms as the measure of interest. In Columns (3)-(6), we use the same framework to separately examine regulated nonbanks including S&Ls and CUs (column 3 and 4) and shadow institutions (columns 5 and 6) as the dependent variable. These estimates show that the relationship is strong and similar for both types of (typically unmeasured) nonbanks. In Panel B, we perform the same analysis using Poisson regression. The coefficients represent the expected percentage change in the number of respective financial institutions in response to a unit change in log(Bank Branches). A benefit of Poisson regressions is they allow us to include observations with zero values, which have an economic meaning in our setting. These are the county-year observations where a certain type of institution is absent. Overall, the economic and statistical significance of both estimation strategies in Panels A and B provide very similar results: nonbanks tend to evolve alongside banks.

The preceding analysis shows that shadow banks act as complements to traditional banks, which leads to the question of their distinct economic role. We propose that this complementarity stems from fundamental differences in their design. Unlike highly-regulated deposit-funded banks, shadow banks are market-funded and contract-flexible. If these institutions can react to pricing signals or new productive technologies into funding more quickly, then counties with a higher shadow-bank share should reallocate productive assets more efficiently when fundamentals change. In the next section, we now take this implication to the data by asking whether counties with more shadow banks reallocate farmland more strongly when corn prices rise.

Table 3: Interplay Between Banks and Nonbanks
Panel A of this table presents estimates from regressions of the log number of nonbanks in a county on the log number of banks, with each specification controlling for *log(Population)*, county fixed effects, and state×year fixed effects. Columns (1)-(2) present estimates for all nonbanks; columns (3)-(4) for regulated nonbank; and (5)-(6) for shadow banks. The odd (even) columns present the estimates using branch (firm) counts. Panel B performs similar analysis using Poisson regression with counts of the respective financial institutions as the dependent variable. *Reg'd Nonbanks* are regulated financial institutions that are not banks and include S&Ls and Credit Unions. *Nonbanks* include the sum of regulated nonbanks and finance companies. All regressions present standard errors clustered at the county level.

Panel A: Linear Regression

	log(All Nonbanks)		log(Reg'd Nonbanks)		log(Shadow Banks)	
	(1) Branches	(2) Firms	(3) Branches	(4) Firms	(5) Branches	(6) Firms
log(Bank Branches)	0.56***		0.46***		0.54***	
	(<0.01)		(<0.01)		(<0.01)	
log(Bank Firms)		0.59***		0.41***		0.60***
		(<0.01)		(<0.01)		(<0.01)
log(population)	-0.00	0.10	0.03	0.08	0.19	0.25
	(0.99)	(0.57)	(0.87)	(0.70)	(0.25)	(0.14)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	1,403	1,403	1,257	1,257	1,320	1,320
Adjusted R-squared	0.84	0.81	0.84	0.81	0.80	0.79

Panel B: Poisson Regression

	All Nonbanks		Reg'd Nonbanks		Shadow Banks	
	(1) Branches	(2) Firms	(3) Branches	(4) Firms	(5) Branches	(6) Firms
log(Bank Branches)	0.56***		0.61***		0.54***	
	(<0.01)		(<0.01)		(<0.01)	
log(Bank Firms)		0.60***		0.57***		0.62***
		(<0.01)		(<0.01)		(<0.01)
log(population)	0.10	0.28	0.08	0.31	0.23	0.40**
	(0.53)	(0.10)	(0.73)	(0.36)	(0.17)	(0.02)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
State × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	1,446	1,446	1,429	1,429	1,437	1,437
Pseudo R-squared	0.84	0.79	0.80	0.75	0.76	0.72

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4 Shadow Banks and the Real Economy

A primary function of the financial system is to facilitate the efficient reallocation of assets to their highest-value users. As fundamentals in the economy change, financial institutions and markets facilitate such reallocation by aggregating information and overcoming frictions between borrowers and savers related to matching, transaction costs, and moral hazard. This foundational idea has been developed extensively in a variety of settings in finance and economics, with a large literature examining these issues in the context of financial intermediation (e.g., Leland and Pyle, 1977; Boot and Thakor, 1997; Ramakrishnan and Thakor, 1984; Subrahmanyam and Titman, 1999; Levine, 2002; Chen, Goldstein, and Jiang, 2007).

Shadow banks (finance companies in our setting) are hybrid institutions, sharing features of both a traditional banks and financial markets. While their local branch presence mirrors that of traditional banks, their liability structure is fundamentally different, as they raise capital from market-based sources rather than consumer deposits. They also face fewer regulatory constraints, which allows them to offer a more expansive set of lending contracts, broadening the contracting space available to borrowers and lenders. These features may position them to respond more quickly to changes in economic fundamentals to facilitate opportunities for asset reallocation.

4.1 Research Design

To empirically test how financial architecture affects asset reallocation, we require a setting with a locally financed asset that experiences plausibly exogenous shocks to its optimal ownership structure. We study farmland because its funding, use, and prices are inherently local. Specifically, we study corn-producing farmland and use fluctuations in the global price of corn as a shock that directly alters the profitability, economics, and thus the optimal allocation of farmland. As Dow and Gorton (1997) write "In a capitalist economy, prices serve to equilibrate supply and demand for goods and services, continually changing to reallocate resources to their most efficient uses."

Changes in corn prices naturally alters the optimal allocation of land. A rise in corn prices makes farming more profitable, increasing the optimal scale of capital investment in farm equipment. For example, farmers might find it optimal to invest in more mechanization by buying efficient tractors, replacing old tractors with new ones, or installing a new irrigation system when corn prices go up, consistent with real option investment models (Dixit and Pindyck, 1994). This can lead to the consolidation of land under the ownership of the most efficient producers who are best able to deploy this capital. As a result, we expect periods of rising corn prices to be associated with a decrease in the number of farms and an increase in their average size.

The use of corn price as a trigger for reallocation provides a distinct empirical advantage. Because corn prices are determined on a global market and are influenced by a host of factors such as wars or large scale weather events (Rajan and Ramcharan, 2015; Carlin and Mann, 2024) rather than local conditions, they provide a source of variation that is exogenous to a specific county's economic or financial environment. This plausibly exogenous shock to profitability allows us to estimate whether local financial architecture affects the transmission of these shocks to asset reallocation.

A key empirical challenge is to separate the effect of corn price changes from those of other aggregate economic variables that could independently drive asset reallocation. To address this, our research design uses non-corn-producing counties as a benchmark. This approach allows us to control for common macroeconomic shocks that affect the economics of farming in all counties irrespective of the crop they produce. We implement this strategy by dividing all counties in our dataset into two groups: those with heavy presence of corn farming (top quartile average corn production) and those that do not.[7]

To empirically examine these idea, we collect detailed, county-level U.S. Agricultural Census data and farmland value data (Barnard and Jones, 1987) to match to our dataset. This results in an unbalanced, 5-year interval panel of 143 counties during 1930-1985. Using

[7]In robustness tests, we use the percent of acreage in the county used for corn farming and show that our results consistent.

these data, we estimate the following regression model:

$$y_{ct} = \alpha_c + \alpha_t + \gamma(CornCnty_c \times CornPriceReturn_t \times ShadowRatio_{c,t-5}) + \Psi \mathbf{X}_{ct} + \epsilon_{ct} \quad (3)$$

y_{ct} represents an outcome in county c and year t (e.g., the number of farms). County fixed effects α_c absorb time-invariant county characteristics, and time fixed effects α_t absorb aggregate shocks. $CornCnty_c$ is an indicator for corn-producing counties; $CornPriceReturn_t$ is the global corn-price return over the past five years; $ShadowRatio_{c,t-5}$ measures the local five-year-lagged intensity of shadow banking. The vector $\mathbf{X}_{ct}$ includes the main effects not absorbed by the fixed effects and all pairwise interactions among these variables. To isolate the effects of real price changes from general price level movements, we also control for the interaction of inflation with corn price returns in corn-producing counties. The coefficient of interest, γ, is the difference-in-differences-in-differences estimate.

4.2 Shadow Banks and the Reallocation of Asset Ownership

We begin by examining whether corn price shocks differentially affect asset allocation according to local financial architecture (i.e., share of local financial institutions that are shadow banks). Table 4 presents estimates of equation (3), where we focus on two outcome variables: (a) number of farms in the county and (b) the average farm size in the county.

Column (1) presents our baseline result on the number of farms: a higher share of (five-year-lagged) shadow banks is associated with a greater reduction in the number of farms when corn prices rise. Column (2) confirms this finding is robust to controlling for a potential differential sensitivity to inflation across groups. In Column (3), we re-estimate the model on a "restricted" sample, requiring each county to have at least one shadow bank and one traditional bank. This test ensures our results are not driven purely by the extensive margin of financial development (i.e., the presence versus absence of shadow banks). The coefficient

Table 4: Farm Consolidation
This table presents estimates of regression (3), with log(Farms) (columns 1-3) and log(Average Farm Size) (columns 4-6) as the dependent variables. *Shadow Ratio* is the number of finance company branches as a share of total financial institutions, lagged five years. *(z)CornPriceReturn* and *(z)Inflation* are standardized five-year corn-price returns and inflation, *log(Farms)* is the log of the number of farms in the county, and *log(Avg Farm Size)* is the log of the average farm acreage in the county. All regressions include county and year fixed effects, and standard errors are clustered at the county level.

	log(Farms)			log(Avg Farm Size)		
	(1)	(2)	(3)	(4)	(5)	(6)
L5.ShadowRatio × (z)CornPriceReturn × CornCnty	-0.14***	-0.15***	-0.25***	0.06**	0.08**	0.10*
	(0.01)	(0.01)	(<0.01)	(0.02)	(0.02)	(0.05)
L5.ShadowRatio × (z)Inflation × CornCnty		0.07	0.13*		-0.09*	-0.09
		(0.28)	(0.08)		(0.09)	(0.17)
Sample	Full	Full	Restricted	Full	Full	Restricted
Nobs	988	988	698	976	976	694
Adjusted R-squared	0.94	0.94	0.93	0.95	0.95	0.94

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

of -0.25 is statistically significant and economically large. It implies that moving a county from the 25th to the 75th percentile in its shadow bank ratio (an increase of 0.36) is associated with a 9.2% larger reduction in the number of farms following a one-standard-deviation increase in corn price returns.

In columns (4)-(6), we repeat the analysis, now examining farm size. Across the specification, we find that a higher share of lagged shadow bank intensity corresponds to a greater increase in average farm size. Moving from the 25th to the 75th percentile in its shadow bank ratio is associated with a 3.5% larger increase in the average farm size following a one-standard-deviation increase in corn price returns. Taken together, these results provide evidence that the presence of shadow banks facilitates consolidation: larger farms held by fewer owners.

While these results show that reallocation is occurring, the key question is whether this reallocation is efficient. A complete answer requires granular data on farm-level inputs and outputs. Absent such data, we next examine two proxies for efficiency gains: capital deepening and changes in land values.

4.3 Shadow Banks and the Capital Deepening

A key motivation behind consolidation is to leverage economies of scale. This factor becomes increasingly important as rising corn prices raises the incentive for large capital investment. This mechanism is particularly relevant in our sample period, which saw rapid advances in farm mechanization with the available horsepower of tractors quadrupling between 1947 and 1977 (Olmstead and Rhode, 2000). To investigate this mechanism directly, we study county-level data on the number of tractors. While this is only one of several ways farmers can leverage scale economies, tractors are certainly a key form of physical capital. If consolidation improves efficiency, this should lead to higher capital intensity on the remaining farms. Specifically, we estimate regression 3 with tractors per farm and total county-level operating tractors.

Table 5 presents the results. Columns (1)-(3) show a positive and significant relationship between the shadow bank interaction term and the number of tractors per farm. The estimate of 0.18 in column (3) implies that moving from the 25th to the 75th percentile of shadow bank ratio is associated with an additional 0.063 tractors per farm (4% of the sample average) following a one-standard-deviation increase in corn price returns. Columns (4)-(6) show that this same interaction corresponds to a significant decrease in the total number of tractors in the county, with 4.8% larger reduction in tractors moving from the 25th to 75th percentile of shadow ratio. These results are consistent with a move towards greater mechanization by purchasing more efficient tractors in areas with more shadow banks: farmers are using fewer tractors in the aggregate, but they use more tractors on a per farm basis.

Combined with the earlier results, this evidence suggests that shadow banks facilitate a more efficient allocation of capital in response to economic shocks. Specifically, capital becomes more concentrated on the larger, surviving farms. In the next section, we examine whether these apparent efficiency gains are reflected in asset prices.

Table 5: Capital Deepening
This table presents estimates of regression (3), with Tractors per Farm (columns 1-3) and log(Tractors) (columns 4-6) as the dependent variables. *Shadow Ratio* is the number of finance company branches as a share of total financial institutions, lagged five years. *(z)CornPriceReturn* and *(z)Inflation* are standardized five-year corn-price returns and inflation, *Tractors/Farm* is the number of tractors divided by the number of farms in the county, and *log(Tractors)* is the log of the number of tractors in the county. All regressions include county and year fixed effects, and standard errors are clustered at the county level.

	Tractors/Farm			log(Tractors)		
	(1)	(2)	(3)	(4)	(5)	(6)
L5.ShadowRatio × (z)CornPriceReturn × CornCnty	0.08*	0.12**	0.18***	-0.10*	-0.08	-0.13**
	(0.09)	(0.02)	(0.01)	(0.09)	(0.21)	(0.05)
L5.ShadowRatio × (z)Inflation × CornCnty		-0.11	-0.18		-0.02	-0.01
		(0.36)	(0.24)		(0.84)	(0.94)
Sample	Full	Full	Restricted	Full	Full	Restricted
Nobs	785	785	558	785	785	558
Adjusted R-squared	0.95	0.95	0.94	0.89	0.89	0.88

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.4 Shadow Banks and Asset Values

The preceding evidence demonstrates that shadow banks facilitate both land consolidation and capital deepening. While we do not have precise information on all the inputs, outputs, costs, and revenues on a farm-by-farm basis, we do observe the price of land. We now test whether these reallocations are capitalized into asset values, which is a key indicator of efficient resource use. If the shift to fewer, larger, and more capital-intensive farms is value-creating, we expect the market price of farmland to increase accordingly.

Table 6 present estimates of our main regression model (3), with log(Land Value per Acre) as the dependent variable. Columns (1)-(3) show the coefficient on our triple-interaction term is positive and highly significant. The estimate in Column (3) implies that moving a county from the 25th to the 75th percentile in its shadow bank ratio (an increase of 0.36) is associated with a 6.3% larger increase in land value per acre following a one-standard-deviation increase in corn price returns.

Does this value come from greater land productivity or cost efficiencies? Columns (4)-(6) address this question by examining crop yields. We find no evidence that shadow bank-led

Table 6: Land Values and Yield

This table presents estimates of regression (3), with log(Land Value per Acre) (columns 1-3) and log(Yield) (columns 4-6) as the dependent variables. *Shadow Ratio* is the number of finance company branches as a share of total financial institutions, lagged five years. *(z)CornPriceReturn* and *(z)Inflation* are standardized five-year corn-price returns and inflation, *log(Land Value per Acre)* is the log of the average value of farmland in the county, and *log(Yield)* is the log of the average corn yield (bushels per acre) in the county. All regressions include county and year fixed effects, and standard errors are clustered at the county level.

	log(Land Value per Acre)			log(Yield)		
	(1)	(2)	(3)	(4)	(5)	(6)
L5.ShadowRatio × (z)CornPriceReturn × CornCnty	0.07**	0.09**	0.17***	-0.04	-0.04	-0.04
	(0.05)	(0.01)	(<0.01)	(0.31)	(0.41)	(0.52)
L5.ShadowRatio × (z)Inflation × CornCnty		-0.09*	-0.14**		-0.03	-0.02
		(0.09)	(0.05)		(0.70)	(0.78)
Sample	Full	Full	Restricted	Full	Full	Restricted
Nobs	988	988	698	973	973	688
Adjusted R-squared	0.99	0.99	0.99	0.91	0.91	0.91

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

consolidation corresponds to changes in physical output per acre, as the point estimates are economically small and statistically indistinguishable from zero. This null result, combined with the positive valuation effect, provides strong evidence that the efficiency gains are primarily channeled through cost savings from economies of scale.

To ensure our results are not sensitive to the binary classification of corn-producing counties, we perform a robustness test using a continuous measure of agricultural exposure. We replace the corn-county indicator with the five-year lagged percentage of county land in farms (*FarmingIntensity*) and re-estimate our main specifications. We present the results in Appendix Table A.1 and find that our results are qualitatively unchanged: the coefficients on the key triple-interaction term provide strong evidence of shadow banks facilitating asset reallocation, capital deepening, and higher land prices.

4.5 Shadow Banks and Asset Volatility

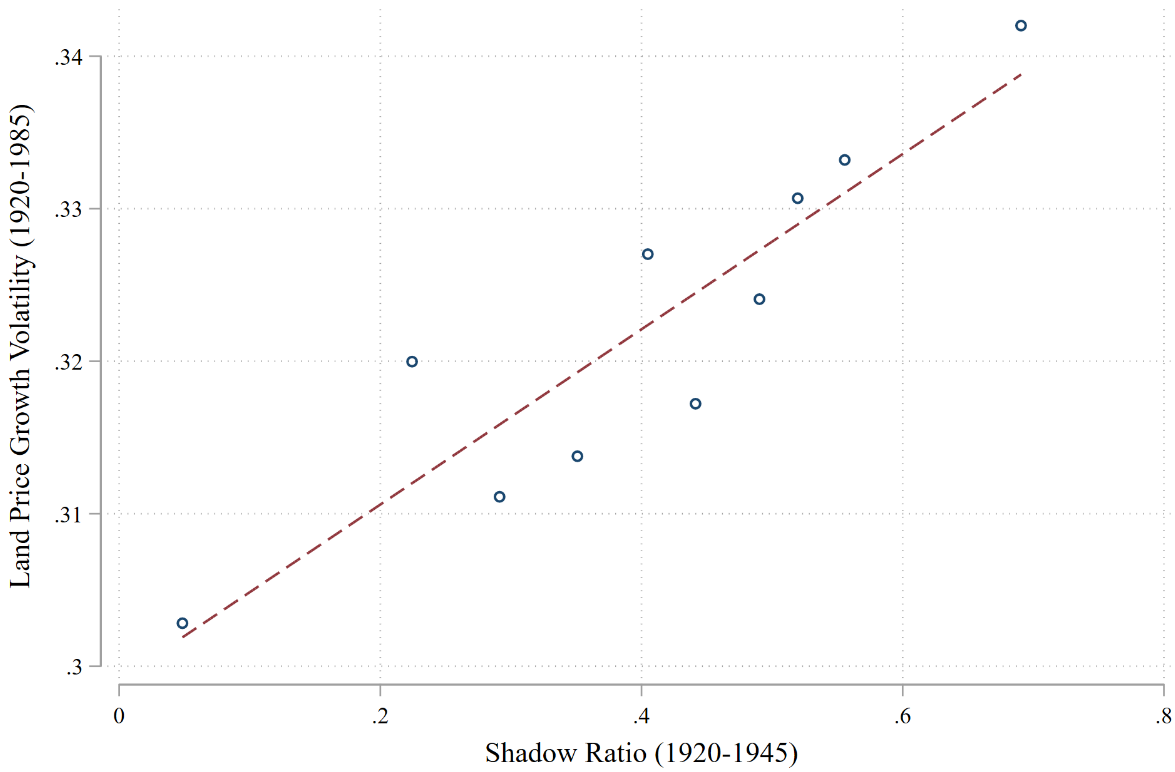
Our results so far establish a positive effect of shadow banks on the local economy. A potential cost of shadow banks is the increased volatility in real economic outcomes that they can create. There are at least two potential mechanisms for such an effect: increased financial leverage, and increased operating leverage. As shown earlier, shadow banks allow for efficient consolidation, which increases the operating leverage of farmlands. In addition, they are likely to increase the overall debt level in the system. Both these effects can lead to an increase in volatility. Such an efficiency-volatility trade-off is not unique to shadow banks; rather, this is an inherent feature of a number of financial systems. We test this conjecture by assessing the volatility of farmland in areas with higher shadow banks. For this analysis, our empirical design is cross-sectional in nature, since we need time-series data for every county to compute a measure of volatility.

Figure 6 presents a binscatter plot that illustrates a clear positive relationship between a county's shadow bank ratio and its land price growth volatility. Overall, our evidence points to an efficiency-volatility trade-off: shadow banks improve efficiency by acting quickly on market signals and improving the reallocation of farmlands, but the benefit of improved allocation comes at the expense of higher volatility.

5 The Consequences of Mismeasuring Financial Architecture and Guidance for Researchers

As shown in the prior section, granular data on broader measures of financial architecture can yield new, more precise insights. Omitting nonbanks from the analysis can mask the effect of finance on real outcomes and potentially bias inferences in analysis that omit them. Our novel dataset on local financial architecture allows us to provide guidance to other researchers by quantifying the biases that arise from using incomplete, "banks-only" measures of financial development. While the detailed analysis is presented in Appendix A., we summarize the key takeaways here.

Figure 6: Shadow Ratio and Asset Volatility
This figure presents a binscatter plot of county-level land price growth volatility over 1920-1985 as a function of average Shadow Ratio during the early part of the sample (1920-1945). The regression coefficient is 0.056 (p-value=0.02). *Shadow Ratio* is the number of finance companies as a share of total financial institutions.



Local Competition is Mismeasured

Relying on banks-only data leads to a fundamental mischaracterization of local financial competition. Our findings in Table 3 earlier have shown that there is strong positive comovement between nonbanks and traditional banks: a 10% increase in bank branches is associated with a 5.6% increase in nonbank branches. Ignoring this relationship can lead to a substantial overstatement of market concentration. The average "banks-only" Herfindahl-Hirschman Index (HHI) in our data is 0.27, suggesting a moderately concentrated market. However, the "all financial intermediary" HHI including all institutions is only 0.09. Given the correlation between traditional banks and nonbanks is less than one, this also means that a rank ordering of competitiveness based on bank-only HHI can lead to substantial misclassification of relative competitiveness.

Omitting Nonbanks Creates Non-Classical Measurement Error Bias

Researcher often use banks presence to study the effect of financial development on real outcomes. They typically, almost always, assume measurement error to be classical, resulting in an attenuation bias in their estimated coefficients. However, because traditional bank and nonbank presence is correlated, omitting the latter creates a non-classical measurement error. Critically, the direction of the bias depends on the research design.

When measuring financial development in levels (e.g., branches per capita), the bias leads researchers to *overestimate* the true effect of finance on real outcomes by a staggering 28% to 72%. When measuring financial development in logarithms (e.g., log(banks), the bias flips, leading researchers to *underestimate* the true effect by 13% to 35%. Further, we show that this bias is not uniform, varying across urban vs. rural and higher-income vs. lower-income counties, meaning policy conclusions based on banks-only data may be inaccurate across geography and income levels.

Omitting Nonbanks Creates Omitted Variable Bias

For researchers specifically estimating the structural impact of regulated banks, omitting nonbanks creates a significant omitted variable bias (OVB) problem. Given the strong positive comovement between the two sectors, regressions that omit nonbanks incorrectly attribute the effects of the entire financial sector to banks alone. Depending on the nature of the correlation between nonbanks and the measured outcome, this can lead to a significant overstatement of the true, direct effect of banks.

A Practical Tool for Researchers

To help mitigate these issues, we use our novel data to provide a practical solution. Researchers with only bank-level data can use our regression calibration estimates to generate a more accurate, predicted measure of total financial development. For the full sample period, the relationship is estimated as:

$$E[\text{Total FIs per 10k}|\text{Banks per 10k}] \approx 1.44 + 1.17 \times (\text{Banks per 10k})$$

Appendix A.3.3. provides for discussion and detailed, period-specific estimates for urban and rural counties to allow for more precise corrections.

6 Conclusion

An accurate understanding of the link between finance and real outcomes like economic growth is of fundamental importance. Because banks play a key role in this link and their data are typically readily available, most empirical research examines the relationship between finance and growth with some measures of the development of the formal banking market alone. We collect detailed, novel, county-level data from phone books in Michigan, Indiana, and Ohio from 1920-2000 that includes regulated nonbank intermediaries, including savings/building & loans and credit unions, as well as unregulated finance companies, i.e., the shadow banks.

We show that nonbanks are ubiquitous across geography and time, constituting nearly 40% of financial institutions during the sample. In addition to showing strong time-series variation in shadow banks' share of institutions (Shadow Ratio), we find that county-level heterogeneity is substantial and persistent, with the local financial architecture having predictive power for future financial architecture for at least 50 years, despite massive changes in financial regulation and technology in the interim. We also find a strong positive correlation between the presence of banks and nonbanks — both regulated nonbanks and shadow banks — during our sample.

Our novel data opens the door for several interesting empirical studies: How do shadow banks affect credit creation? Do banks change their behavior when shadow banks are present? How does the presence of shadow banks affect the volatility of local asset prices? In our paper, we shed light on one of the fundamental questions of finance: how do shadow banks affect the allocation of local assets? Our key finding is that shadow banks act quickly on market signals and facilitate the reallocation of assets across users. The benefit comes at the expense of higher volatility of land prices in these areas, consistent with the classical efficiency-volatility trade-off of a financial system.

In the final part of the paper, we take seriously the challenge that empirical researchers face when attempting to link finance to real outcomes by simply using some version of access to bank finance as the measure of financial development. Using our novel data, we quantify the biases: they range from underestimating the true effect of finance by nearly 50% to overestimating the true effect by over 80%, depending on the empirical model the researcher uses. Finally, we provide a regression calibration exercise to help empirical researchers working with limited data overcome this problem.

References

Allen, Franklin, and Douglas Gale, 2000, *Comparing financial systems.* (MIT press).

Allen, Franklin, and Douglas Gale, 2004, Financial fragility, liquidity, and asset prices, *Journal of the European Economic Association* 2, 1015–1048.

Barnard, Charles Howard, and John Jones, 1987, *Farm real estate values in the United States by counties, 1850-1982.* (US Department of Agriculture, Economic Research Service).

Begley, Taylor A, and Kandarp Srinivasan, 2022, Small bank lending in the era of fintech and shadow banks: a sideshow?, *The Review of Financial Studies* 35, 4948–4984.

Boot, Arnoud WA, and Anjan V Thakor, 1997, Financial system architecture, *The Review of Financial Studies* 10, 693–733.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, 2018, Fintech, regulatory arbitrage, and the rise of shadow banks, *Journal of financial economics* 130, 453–483.

Butler, Alexander W, and Jess Cornaggia, 2011, Does access to external finance improve productivity? Evidence from a natural experiment, *Journal of Financial Economics* 99, 184–203.

Carlin, Bruce, and William Mann, 2024, An Experiment in Tight Monetary Policy: Revisiting the 1920–1921 Depression, *Journal of Financial and Quantitative Analysis* 59, 2299–2339.

Chen, Qi, Itay Goldstein, and Wei Jiang, 2007, Price informativeness and investment sensitivity to stock price, *The Review of Financial Studies* 20, 619–650.

Diamond, Douglas W, and Raghuram G Rajan, 2000, A theory of bank capital, *the Journal of Finance* 55, 2431–2465.

Dixit, Avinash K, and Robert S Pindyck, 1994, *Investment under uncertainty.* (Princeton university press).

Dow, James, and Gary Gorton, 1997, Stock market efficiency and economic efficiency: is there a connection?, *The Journal of Finance* 52, 1087–1129.

Geanakoplos, John, 2010, The leverage cycle, *NBER macroeconomics annual* 24, 1–66.

Jayaratne, Jith, and Philip E Strahan, 1996, The finance-growth nexus: Evidence from bank branch deregulation, *The Quarterly Journal of Economics* 111, 639–670.

King, Robert G., and Ross Levine, 1993, Finance and Growth: Schumpeter Might Be Right*, *The Quarterly Journal of Economics* 108, 717–737.

Kiyotaki, Nobuhiro, and John Moore, 1997, Credit cycles, *Journal of political economy* 105, 211–248.

Leland, Hayne E, and David H Pyle, 1977, Informational asymmetries, financial structure, and financial intermediation, *The journal of Finance* 32, 371–387.

Levine, Ross, 2002, Bank-based or market-based financial systems: which is better?, *Journal of financial intermediation* 11, 398–428.

Levine, Ross, 2005, Finance and growth: theory and evidence, *Handbook of economic growth* 1, 865–934.

National Commission on Consumer Finance, 1972, *Consumer Credit in the United States.* (United States Congress).

Nugent, Rolf, 1939, *Consumer Credit and Economic Stability.* (Russell Sage Foundation).

Olmstead, Alan L, and Paul W Rhode, 2000, The transformation of northern agriculture, 1910–1990, *Cambridge economic history of the United States* 3, 693–742.

Petersen, Mitchell A, and Raghuram G Rajan, 1995, The effect of credit market competition on lending relationships, *The Quarterly Journal of Economics* 110, 407–443.

Rajan, Raghuram, and Rodney Ramcharan, 2015, The anatomy of a credit crisis: The boom and bust in farm land prices in the United States in the 1920s, *American Economic Review* 105, 1439–1477.

Rajan, Raghuram G, and Luigi Zingales, 1998, Financial Dependence and Growth, *American Economic Review* 88, 559–586.

Ramakrishnan, Ram TS, and Anjan V Thakor, 1984, Information reliability and a theory of financial intermediation, *The Review of Economic Studies* 51, 415–432.

Schumpeter, Joseph A., 1911, *Theorie der wirtschaftlichen Entwicklung.* (Duncker & Humblot Leipzig).

Shleifer, Andrei, and Robert W Vishny, 1992, Liquidation values and debt capacity: A market equilibrium approach, *The journal of finance* 47, 1343–1366.

Subrahmanyam, Avanidhar, and Sheridan Titman, 1999, The going-public decision and the development of financial markets, *The Journal of Finance* 54, 1045–1082.

Appendix Figures and Tables

Figure A.1: Sample Coverage over Time

This figure presents the number of counties present in each year of our analysis.. We discuss Section 2 details the sample collection process.

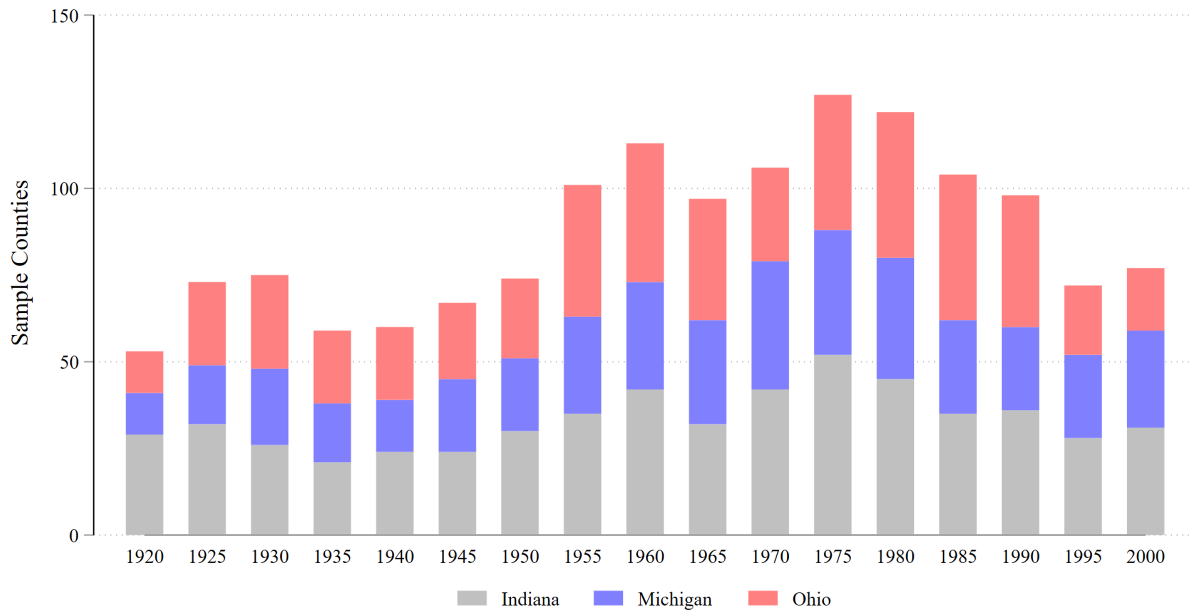


Table A.1: Robustness: Farm Intensity

This table presents estimates of regression (3) using lagged *Farming Intensity* as the measure of local focus on agriculture. *Shadow Ratio* is the number of finance company branches as a share of total financial institutions, lagged five years. *log(Farms)* is the log of the number of farms in the county, and *log(Avg Farm Size)* is the log of the average farm acreage in the county. *Tractors/Farm* is the number of tractors divided by the number of farms in the county, *log(Tractors)* is the log of the number of tractors in the county, *(z)CornPriceReturn* and *(z)Inflation* are standardized five-year corn-price returns and inflation, *log(Land Value per Acre)* is the log of the average value of farmland in the county, and *log(Yield)* is the log of the average corn yield (bushels per acre) in the county. *Farming Intensity* is the perent of land in the county devoted to farming, lagged five years. All regressions include county and year fixed effects, and standard errors are clustered at the county level.

	(1)	(2)	(3)	(4)	(5)	(6)
	log(Farms)	log(AvgFarmSize)	Trac/Farm	log(Tractors)	log(LandValue)	log(Yield)
L5.ShadowRatio × (z)CornPriceReturn × L5.FarmingIntensity	-0.63***	0.37***	0.35***	-0.33	0.40***	0.06
	(<0.01)	(<0.01)	(<0.01)	(0.11)	(<0.01)	(0.59)
L5.ShadowRatio × (z)Inflation × L5.FarmingIntensity	0.35*	-0.27	-0.64*	-0.79*	-0.30	0.22
	(0.08)	(0.13)	(0.08)	(0.08)	(0.20)	(0.23)
Sample	Restricted	Restricted	Restricted	Restricted	Restricted	Restricted
Nobs	698	694	558	558	698	688
Adjusted R-squared	0.95	0.95	0.95	0.90	0.99	0.91

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Regression Calibration for Financial Institution Firms

This table presents OLS estimates of the number of financial institution firms per 10,000 people on the number of banking firms per 10,000 people. Panels A, B, and C estimate the regressions for all counties, urban counties, and rural counties, respectively. Column (1) of each panel include observations from the entire sample period, while columns (2), (3), and (4) use data from the PreWar, PostWar, and Modern periods, respectively. We define counties as urban or rural based on whether their median population is above or below the sample median. *PreWar* is 1920-1945, *PostWar* is 1950-1980, and *Modern* is 1985-2000. All regressions present standard errors clustered at the county level.

Panel A: All Counties				
	(1)	(2)	(3)	(4)
Banks (Firms) / 10k	1.24***	1.48***	1.44***	1.16***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Constant	1.35***	1.16***	1.41***	1.16***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Sample Period	All	PreWar	PostWar	Modern
Nobs	1,478	387	618	473
Adjusted R-squared	0.44	0.51	0.38	0.51
Panel B: Urban Counties				
	(1)	(2)	(3)	(4)
Banks (Firms) / 10k	1.56***	1.49***	2.42***	1.99***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Constant	1.16***	1.11***	1.11***	0.72***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Sample Period	All	PreWar	PostWar	Modern
Nobs	732	236	288	208
Adjusted R-squared	0.36	0.45	0.33	0.42
Panel C: Rural Counties				
	(1)	(2)	(3)	(4)
Banks (Firms) / 10k	1.10***	1.40***	1.31***	1.08***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Constant	1.54***	1.31***	1.50***	1.27***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Sample Period	All	PreWar	PostWar	Modern
Nobs	746	151	330	265
Adjusted R-squared	0.35	0.43	0.30	0.43

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Analysis Appendix

A. The Consequences of Mismeasurement

In this section, we begin by highlighting the basic inaccuracy present when measuring competition in the supply of financial services with banks only. We then characterize the general nature of the measurement error bias that comes from omitting nonbanks from the analysis of financial development on real outcomes. Finally, we show how excluding nonbanks introduces significant omitted variable bias in estimates that seek to relate banks to real outcomes.

A.1. Mismeasuring Competition

Competition in the supply of credit is an object of fundamental importance both to policymakers and academics because of its implications for economic growth and efficiency, access to credit, inequality, outcomes in banking such as bank-borrower relationship, or the stability of the financial markets (e.g., Petersen and Rajan, 1995). Whether a regulator approves a bank merger or not often depends on the impact of that merger on banking competition at the local level. It is clear from our discussion so far that measures of competition computed solely from the presence of regulated banks can be misleading, with consequences for both policymakers and the academic literature.

We now explore the extent of this bias in the computation of competition. We begin by computing two aspects of competition for each county at each time period: the number of (a) bank branches per capita and (b) nonbank financial institution branches per capita present in the county. We plot these two measures against each other in Figure AA.1. The solid lines represent "competition isoquants" such that anywhere on the line has the same number of financial institution branches per capita. Thus, moving to the northeast in the figures represents more competition. The figure shows that the omission of nonbanks can lead to a nontrivial mischaracterization of local competition.

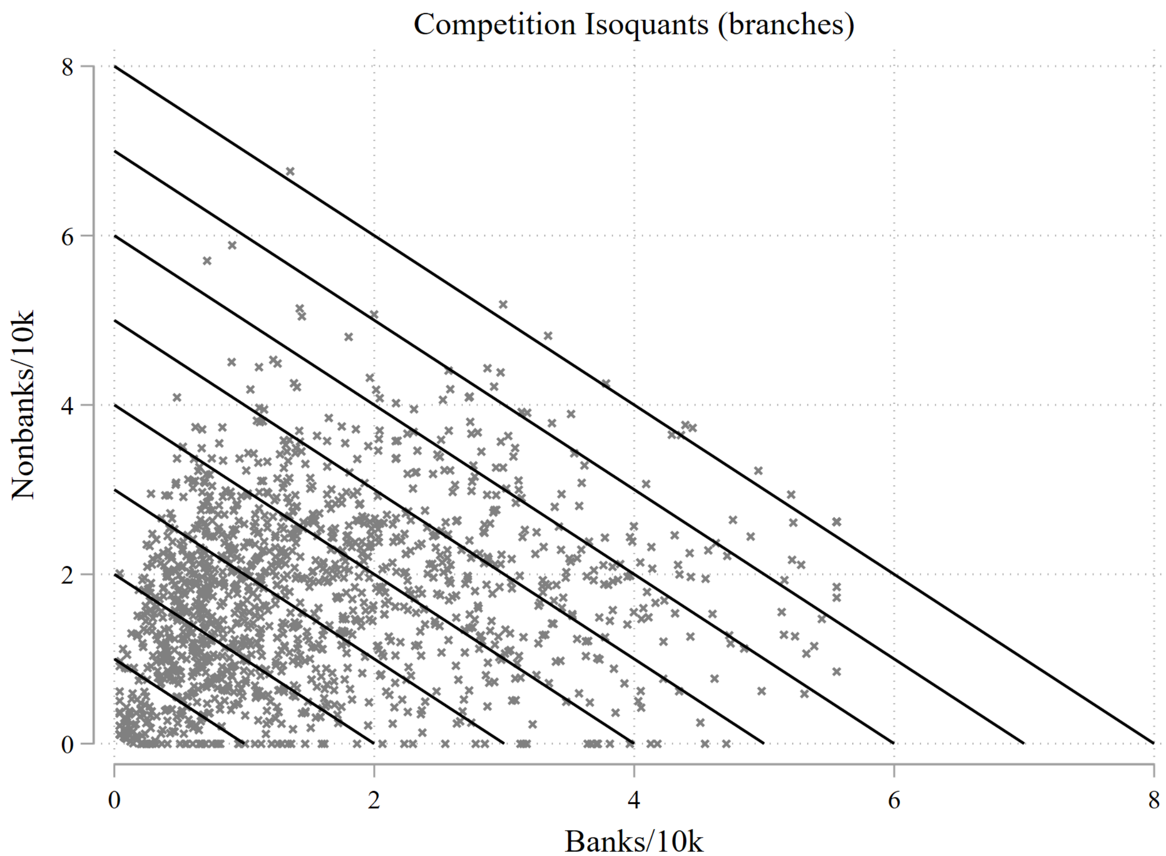
Consider Indiana in 1975. Clinton and Hamilton Counties each have about 1.6 banks per capita, which is about the 40th percentile of competitiveness by this measure. Measuring competition using total FIs paints a different picture: Clinton County has about 4.1 total FIs per capita (about the 60th percentile) while Hamilton County has about 2.5 total FIs per capita (about the 20th percentile). More broadly, if a regulator uses the median level of competition across counties as the cutoff rule for approving a merger deal, then about 10% of markets in our sample are classified as less competitive based on banks alone when they truly are more competitive, and about another 10% of markets are classified as highly competitive when the reverse is true. Such misclassifications can result in inaccurate policy decisions, which in turn can have welfare consequences.

A second common way of examining the competitiveness of an area is by computing the Herfindahl-Hirschman Index (HHI). For example, when the Department of Justice is considering the approval of bank mergers, they estimate the impact of the merger on the local deposit HHI based on banks and thrifts in a county.[8] Given that nonbanks provide several of the same products as banks, a "banks only" HHI will understate the true competitiveness of the area. While we do not have financial statement data on our firms, we can compute a local HHI measure based on the the number of branches financial firm has in a county. For example, a bank may have three branches while a finance company has one. In this case, our HHI measure will assume the bank has three-times the local market share as the finance company.

Figure AA.2 plots each county's HHI measured with banks only (x-axis) against the HHI measured using all local financial institutions. Along the diagonal represents counties where there are only banks, so the "bank only" HHI accurately captures the competitiveness of the local area. The further away from the diagonal represents a larger mismeasurement of local competition. The average bank-only HHI is 0.27, while the "all FIs" HHI is 0.09. The

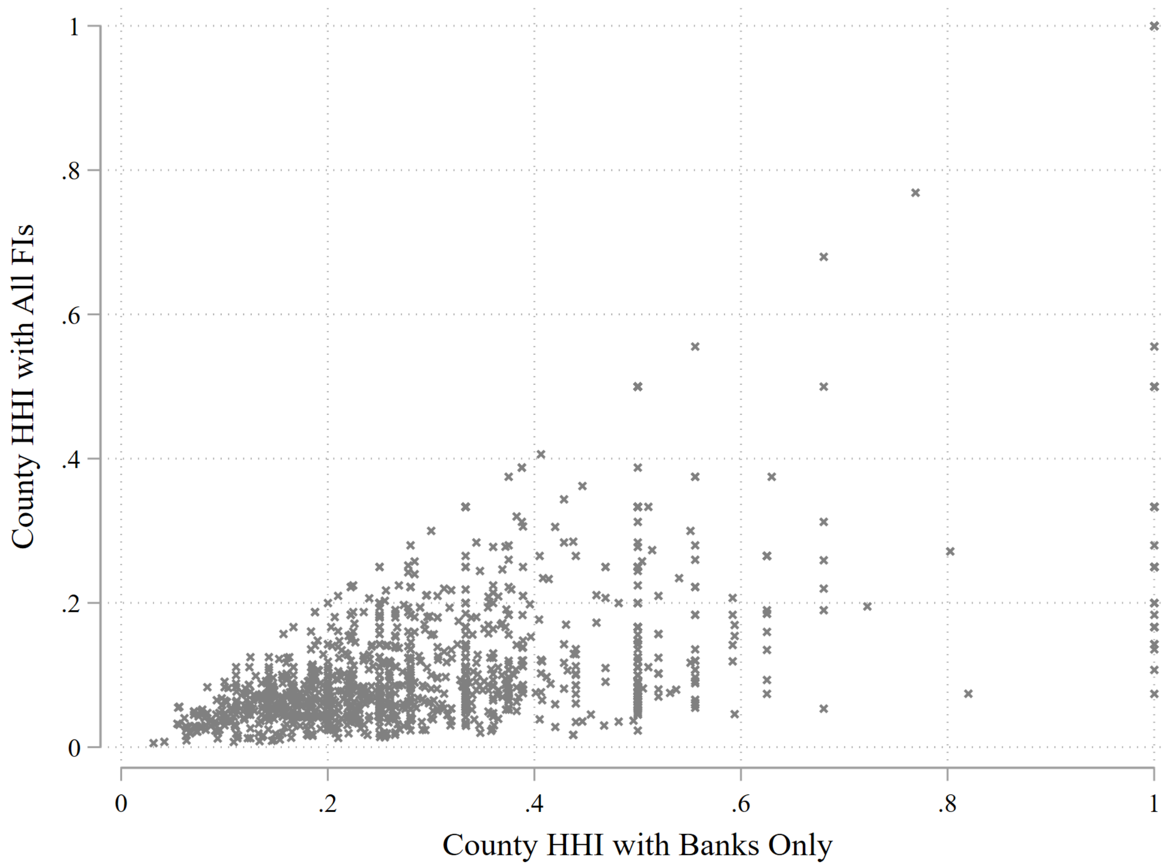
[8] For example, see https://www.justice.gov/atr/bank-merger-competitive-review-introduction-and-overview-1995.

Figure AA.1: Competition Isoquants
This figure plots the number of nonbanks per 10,000 people against banks per 10,000 people for each observation in the sample. The solid lines represents isoquants where, under some assumptions, competition in the supply of financial services is equal.



difference in the two measures range from 0.05 to 0.40 from the 10th to 90th percentile. While we recognize that giving equal weight to a bank branch and finance company branch may be more or less reasonable depending on the product or service, the point remains that ignoring the nonbank suppliers of credit leads to a mischaracterization of the local competitiveness of financial services.

Figure AA.2: Competition HHI
This figure plots each county's HHI using banks only on the x-axis and all financial institutions on the y-axis. To compute the HHI, we equally weight each financial instituion branch and aggregate to the firm level for each county. For example, bank with three local branches gets three-times the weight of a finance company with one local branch.



A.2. Measurement Error Bias

A.2.2. Characterizing the Bias from Mismeasuring Financial Architecture

In light of the results we have presented so far, ignoring shadow banks from a study that attempts to link financial development to economic outcomes can be problematic for at least three reasons: (a) shadow banks represent a large fraction of financial institutions at a local level, (b) the variation in shadow banks over time and across counties is economically meaningful, and (c) the evolution of banks and nonbanks are correlated. Critically, the correlation between banks and nonbanks means unmeasured nonbanks are not just a "noise" term in analyses that measures financial development with just the number of banks. We now quantify the extent of this bias using our data.

There is a vast empirical literature seeking to estimate the presence or magnitude of the effect on finance on real outcomes (e.g., economic growth, capital formation, etc.). Consider a true regression model linking economic outcomes (Y) to financial development (F) at country, state, or county level (i) is some form of the following:

$$Y_i = \alpha + \beta \times F_i + \epsilon_i \tag{4}$$

Since the true measure of financial development—which might include banks, credit unions, finance companies, peer-to-peer lenders, payday lenders, etc.—is not usually available, empirical researchers often estimate a model based on some parameterization of the number of banks B_i for whom data are more readily available. The result is that B is a measure of F with some error η, and the following is often estimated:

$$Y_i = \gamma + \beta \times B_i + \nu_i \tag{5}$$

$$B_i = F_i + \eta_i \tag{6}$$

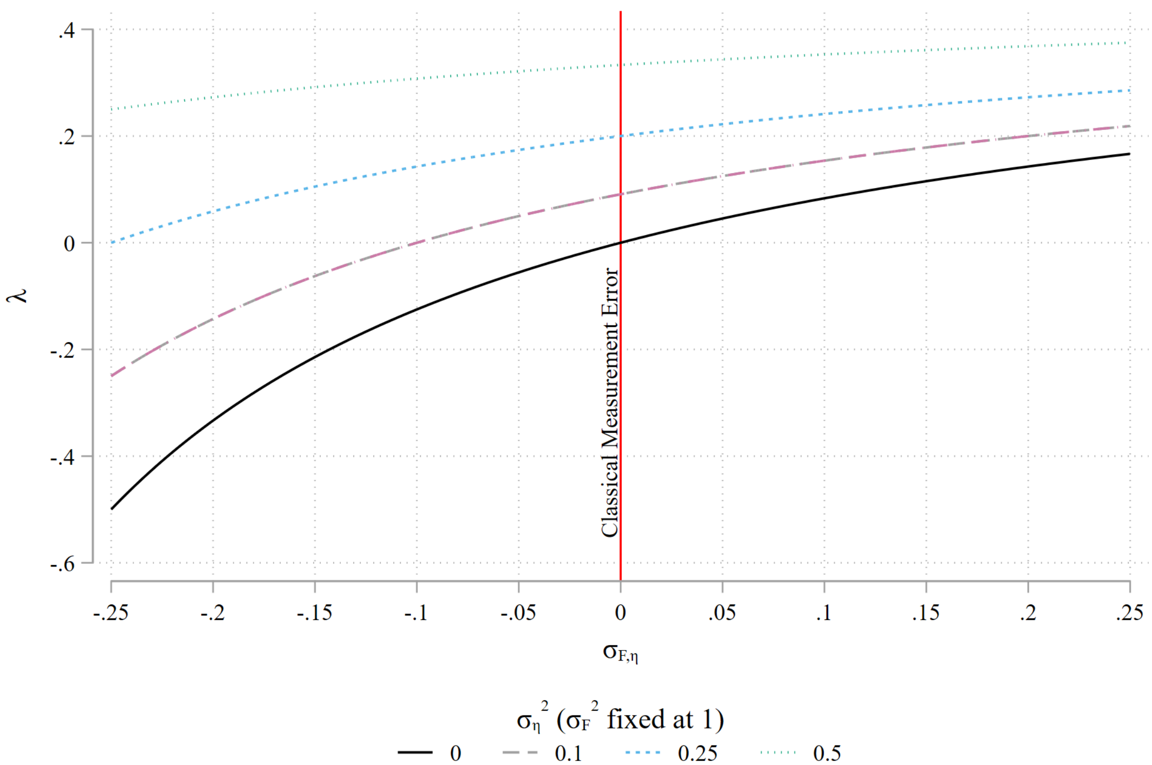
Given this measurement error, the relationship between the true coefficient β in (4) and the estimated coefficient $\hat{\beta}$ from the mismeasured regression (5) is given by the following equations:

$$\hat{\beta} = \beta \times (1 - \lambda) \tag{7}$$

$$\lambda = \frac{\sigma_\eta^2 + \sigma_{F,\eta}}{\sigma_\eta^2 + \sigma_F^2 + 2\sigma_{F,\eta}} \tag{8}$$

The bias λ depends on the relative importance of the variance in the measurement error (σ_η^2), the variance of the true measure (σ_F^2), and the covariance between the true measure and

Figure AA.3: Mismeasurement and the Extent of the Bias
This figure presents the value of the measurement error factor λ as a function of $\sigma_{F,\eta}$, which is the covariance of the true measure of financial development and the unmeasured portion of financing development. Each line plots this relationship for a given σ^2_η, which is the variance of the unmeasured portion η of financial development. The plot is normalized such that the variance of the true financial development σ^2_F is one.



the measurement error ($\sigma_{F,\eta}$). In our context, this will translate to the relative importance of the variance of nonbanks (i.e., $-\eta$ since $B_i = F_i + \eta_i$) compared to the variance in overall financial institutions, and the covariance between the two.

Figure AA.3 illustrates how the interaction of these components translates to bias in the empirical estimates in the mismeasured equation by plotting equation (8) for λ with σ^2_F normalized to one for various degrees of the variance of the error (σ^2_η) and the covariance between the true measure and the measurement error ($\sigma_{F,\eta}$).

If the measurement error were classical ($\sigma_{F,\eta} = 0$) as is often implicitly assumed, then $\lambda = \frac{\sigma^2_\eta}{\sigma^2_\eta + \sigma^2_F}$, which constrains λ to be positive, bounded between zero and one. In that case, the consequence of mismeasurement is an attenuation bias in the estimated coefficient. This is shown in Figure AA.3 on the vertical line where $\sigma_{F,\eta} = 0$. Here, as the variance of the unmeasured financial depth increases, so does λ. This translates to larger attenuation bias, which biases the estimate from the mismeasured equation toward zero.

What if $\sigma_{F,\eta} \neq 0$? If there is nonzero correlation between the true degree of financial development F and its unmeasured portion η, then λ is no longer bounded between zero and one. As Figure AA.3 clearly shows, both the direction and magnitude of bias are unclear. In this case, using the number of banks rather than the true degree of financial development can either underestimate ($\lambda > 0$) or overestimate ($\lambda < 0$) the true effect of finance on the studied outcome. In the next subsection, we use our novel data to empirically examine the presence and nature of such biases during our sample period.

A.2.2. Quantifying of the Measurement Error Bias

Our results from the earlier Section 3.4 show a strong positive correlation between banks and nonbank institutions, so the data reveal that the measurement error is *not* classical in nature. Prior literature has ignored this source of bias. Our novel data allow us to more accurately observe the true value of F, and this allows us to directly estimate the key variances and covariance in equation (8) and therefore estimate the likely bias encountered in studies that estimates the relationship between economic outcomes and banks (equation 5) rather than with a more accurate measure of the supply of finance (equation 4) .

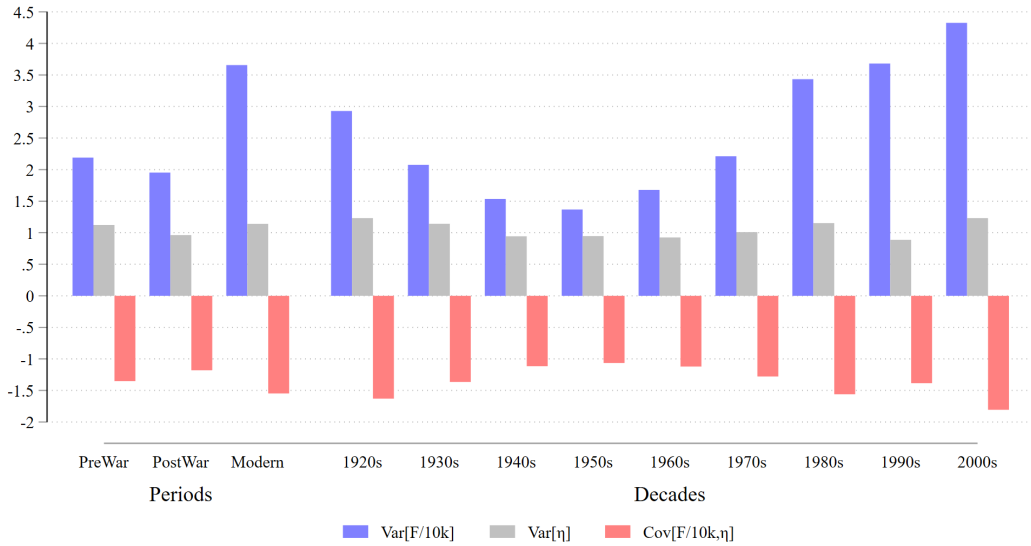
As shown earlier in the paper, there is significant time-series variation in the measures of F and B in our sample. Therefore, we quantify the bias based on a decade-by-decade estimation to ensure that our estimates of relevant variances and covariance is stable for the period of the estimation. The first measure of financial development we consider is per capita value of F and B, defined as number of branches of financial institutions or banks per 10,000 people. This is often referred to as "bank density" (e.g., Rajan and Ramcharan, 2015).

The estimates of relevant variances and covariance are shown in Figure AA.4a on a decade-by-decade basis as well as for three distinct periods: (a) pre-war period of 1920-1945, (b) post-war period of 1950-1980, and (c) the modern period of 1980-2000. We aggregate the data across these three periods based on the aggregate trend in the shadow bank ratio that we document in Figure 4. There are two key observations that emerge from this figure: (a) the magnitude of variance of the measurement error term (σ_η^2) is broadly of the same order of magnitude as the true measure (σ_F^2), and (b) the covariance term ($\sigma_{F,\eta}$) is negative and sizable. The first observation implies that the extent of bias can be large even if the measurement error were classical in nature. The second observation provides strong evidence that the measurement error is non-classical. This corresponds to the left side of Figure AA.3 and means that the quantification of the bias is practically relevant to determine both the direction and extent of the bias.

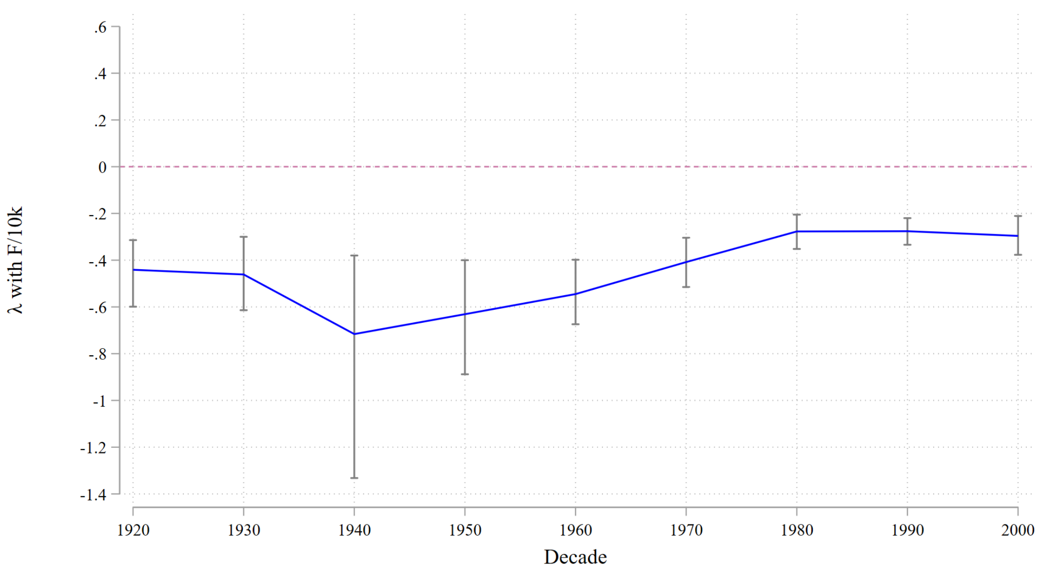
Using the correlation estimates, Figure AA.4b presents the decade-by-decade estimate of λ when finance is measured by the number of branches of financial institutions per 10,000 people, which we denote as $\lambda_{F/10k}$ (also shown in column (1) of Table AA.1). This figure shows that $\lambda_{F/10k}$ is negative for each decade in the sample, ranging from -0.72 in the 1940s to -0.28 during the Modern period. Plugging these values into equation (7), this indicates that the estimated coefficient $\hat{\beta}$ from the mismeasured regression (5) *overestimates* the true effect of finance β from the equation (4) by an economically large 28% to 72%.

We next consider an alternative parameterization banking depth by taking the log transform of the number of branches per 10,000 people. Figure AA.5a shows the components of the measurement error and shows that, relative to the variance in the true measure of financial development ($\log(F/10k)$), the variance of the unmeasured financial development (σ_η^2) is smaller and the covariance term ($\sigma_{F,\eta}$) still negative, but also smaller in magnitude. Figure AA.5b presents the decade-by-decade estimate of $\lambda_{log(F/10k)}$, showing a positive value that is declining over time. This example shows that seemingly innocuous changes in the parameterization — $F/10k$ to $\log(F/10k)$ — flips the sign of the bias, now leading to $\hat{\beta}$ from the mismeasured regression (5) to *underestimate* the true effect of finance β from the equation

Figure AA.4: Measurement Error Components: Finance Per Capita
This figure presents estimates of the measurement error components from equation (8) in Panel (a) and the bias factor λ in Panel (b) when the measure of financial development is the number of branches of financial institutions per 10,000 people ($F/10k$). The components are estimated within the specific time periods in the figure, and include the variance of the true measure of financial development $\sigma^2_{F/10k}$, the variance of the unmeasured financial development (nonbanks) σ^2_η, and their covariance $\sigma_{F/10k,\eta}$, and the measurement error bias factor is computed as $\lambda_{F/10k} = \frac{\sigma^2_\eta + \sigma_{F/10k,\eta}}{\sigma^2_\eta + \sigma^2_{F/10k} + 2\sigma_{F/10k,\eta}}$, with bootstrapped confidence intervals at the 10% level. *PreWar* is 1920-1945, *PostWar* is 1950-1980, and *Modern* is 1985-2000.



(a) Measurement Error Components ($F/10k$)



(b) Lambda ($F/10k$)

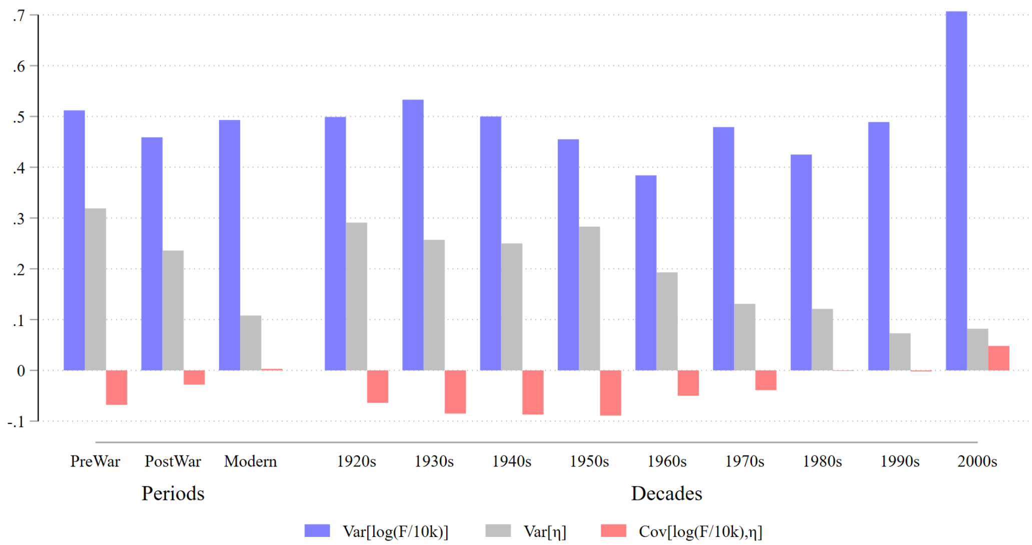
Table AA.1: Lambdas for Different Measures of Financial Development
This table presents the estimates of the measurement error bias factor λ for different measures of financial development rather than the true financial development which is typically unobserved. This factor is compute in equation (8): $\lambda_{F/10k} = \frac{\sigma_\eta^2 + \sigma_{F/10k,\eta}}{\sigma_\eta^2 + \sigma_{F/10k}^2 + 2\sigma_{F/10k,\eta}}$. Columns (1) and (2) use versions of banks per 10,000 people as the misumeasured true financial development, and columns (3) and (4) use all regulated financial institutions, which includes banks, S&Ls, and Credit Unions. *PreWar* is 1920-1945, *PostWar* is 1950-1980, and *Modern* is 1985-2000.

Measure:	Banks		All Regulated Financial Institutions	
	Banks/10k (1)	log(Banks/10k) (2)	RegdFIs/10k (3)	log(RegdFIs/10k) (4)
PreWar	-0.377	0.361	-0.219	0.174
PostWar	-0.390	0.326	-0.149	0.185
Modern	-0.240	0.183	-0.114	0.088
1920s	-0.441	0.343	-0.228	0.186
1930s	-0.461	0.277	-0.206	0.140
1940s	-0.716	0.283	-0.266	0.156
1950s	-0.631	0.347	-0.195	0.223
1960s	-0.545	0.300	-0.325	0.091
1970s	-0.408	0.172	-0.250	0.038
1980s	-0.277	0.220	-0.107	0.105
1990s	-0.276	0.128	-0.114	0.079
2000s	-0.296	0.147	-0.143	0.069

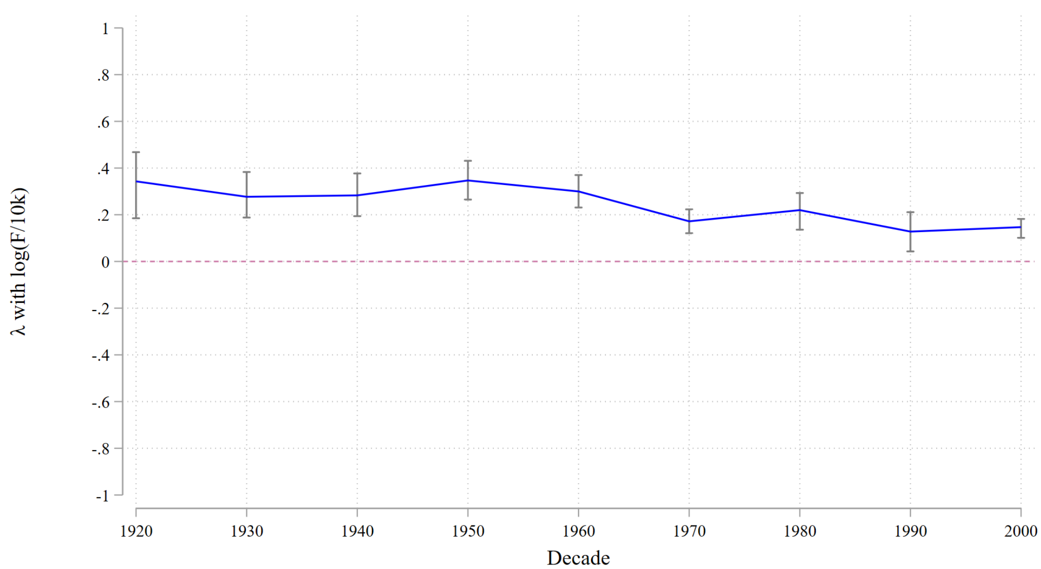
(4) by 13% to 35%. The figure also shows that during Modern period since the 1980s, the covariance term becomes relatively small which indicates closer to classical measurement error. However, quantifying the size of that attenuation bias requires our measurement of the σ_η^2.

These exercises have focused on using the number of respective branches as the measure of financial development. Alternatively, we compute the degree of bias using the number of unique financial firms as our measure. Appendix Figure AA.6 plots the results. We find qualitatively similar patterns, but the size of the bias is amplified for both parameterizations $F/10k$ and $\log(F/10k)$). This is driven by the aggregation of the firm level within county, which disproportionately affects banks, lowers σ_F^2, thus shrinking the denominator of λ as seen in equation (8).

Figure AA.5: Measurement Error Components: log(Finance Per Capita)
This figure presents estimates of the measurement error components from equation (8) in Panel (a) and the bias factor λ in Panel (b) when the measure of financial development is th log of the number of branches of financial institutions per 10,000 people ($\log(F/10k)$). The components are estimated within the specific time periods in the figure, and include the variance of the true measure of financial development $\sigma^2_{\log(F/10k)}$, the variance of the unmeasured financial development (nonbanks) σ^2_{η}, and their covariance $\sigma_{\log(F/10k),\eta}$, and the measurement error bias factor is computed as $\lambda_{\log(F/10k)} = \frac{\sigma^2_{\eta}+\sigma_{\log(F/10k),\eta}}{\sigma^2_{\eta}+\sigma^2_{\log(F/10k)}+2\sigma_{\log(F/10k)k,\eta}}$, with bootstrapped confidence intervals at the 10% level. *PreWar* is 1920-1945, *PostWar* is 1950-1980, and *Modern* is 1985-2000.

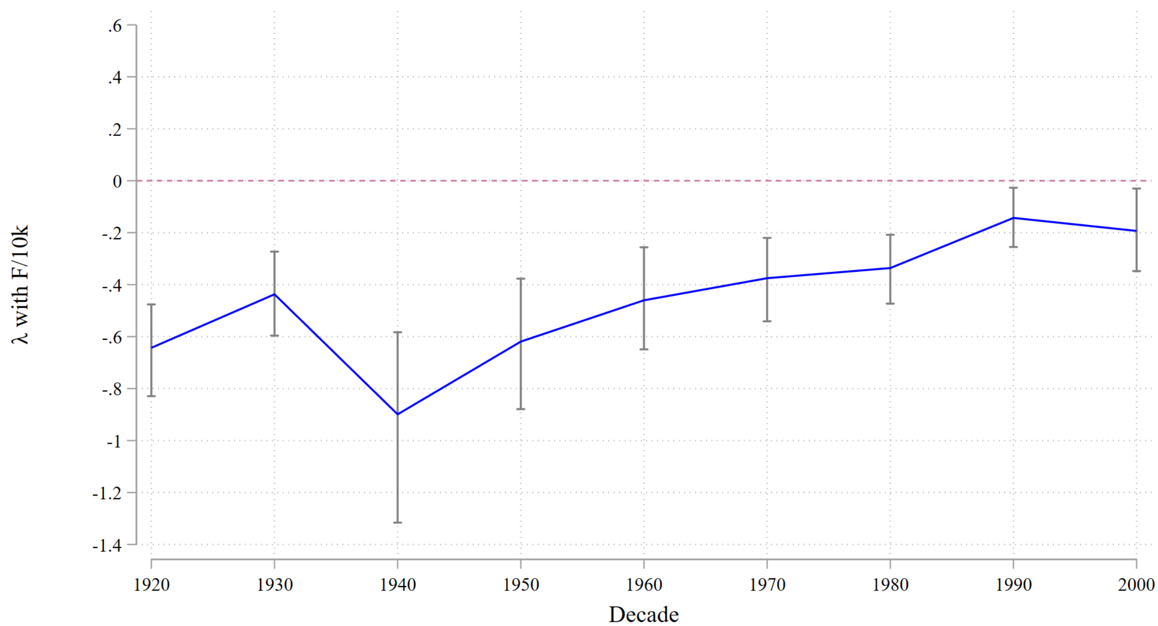


(a) Measurement Error Components ($\log(F/10k)$)

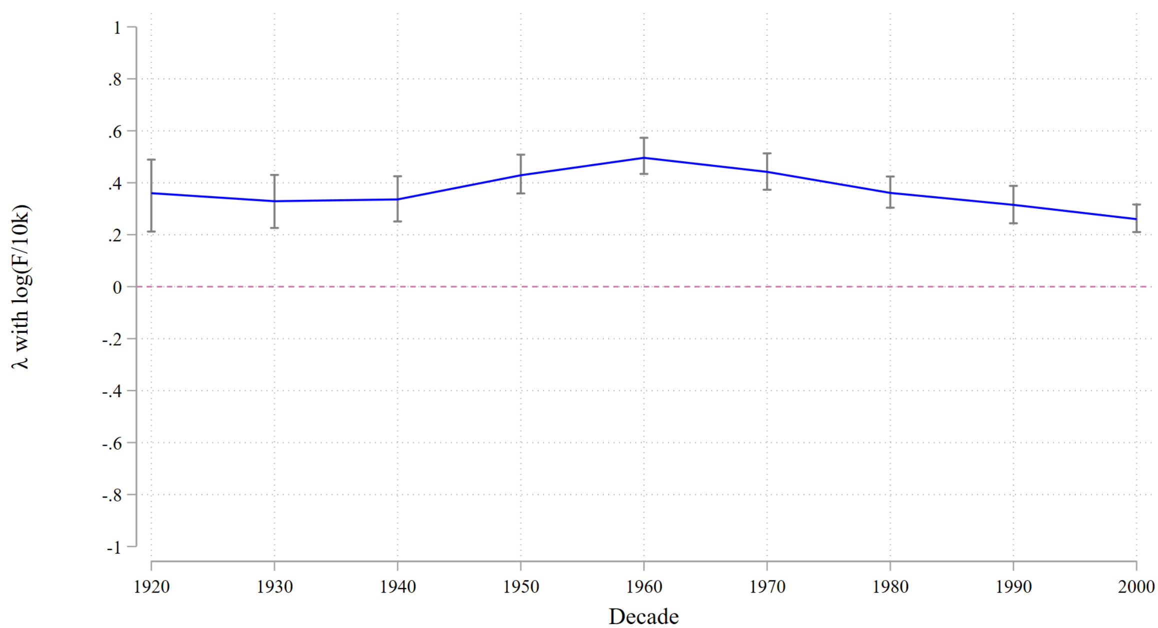


(b) Lambda ($\log(F/10k)$)

Figure AA.6: Measurement Error When Using Unique Financial Firms
This figure presents estimates of the bias factor λ from equation (8) over the sample period. Panels (a) and (b) presents the results when the measure of financial development is the number of firms of financial institutions per 10,000 people ($F/10k$) and $\log(F/10k)$, respectively. The components of λ are estimated within each decade, with bootstrapped confidence intervals at the 10% level.



(a) Lambda: $F/10k$



(b) Lambda: $\log(F/10k)$

A.2.2. Cross-Sectional Heterogeneity in Measurement Error Bias

The results so far have shown striking time series variation in financial architecture during the last hundred years and the consequent changes in the bias of empirical estimates using incomplete measures of financial development to study the role of finance on real outcomes. What about the cross-section? Given the interest of policymakers and academics alike in understanding how finance can affect particular areas and populations, it is important to understand whether the severity of the measurement error problems shown above examine have a differential impact across such groups.

We begin this analysis be splitting the counties in our sample based on urban and rural, which define based on whether the county's median sample population is above or below the overall sample median. In Panel A of Table AA.2, we present estimates of λ across urban and rural for each of the main time periods: PreWar, PostWar, and Modern. Across all time periods, $\lambda_{F/10k}$ is substantially more negative for urban than rural counties. The negative $\lambda_{F/10k}$ indicates that studies using *Banks/10k* as the measure of financial development — thus omitting nonbank financial institutions — will systematically overestimate the effect of finance on real outcomes, and this overestimate is substantially larger for urban as compared to rural areas. For example, the overestimation in the PostWar period for rural areas is 18% while it is 80% for urban areas. Panel A again highlights the crucial role that parameterization plays in the bias. When using $\log(Banks/10k)$, $\lambda_{\log(F/10k)} > 0$ indicates attenuation bias relative to the true effect, and the attenuation is larger for rural areas for most of the sample. For example, using $\log(Banks/10k)$ as the measure of financial development leads to an underestimate of the true effect in the PostWar period by 37% for rural areas compared to 26% for urban areas.

In Panel B of Table AA.2, we perform a similar exercise by splitting the counties into above- and below-median per capita income ("Rich" versus "Poor"). Since income data begin in 1960, we do not have data for any of the PreWar period. Here, we find that

Table AA.2: Lambdas Across Urban/Rural and High/Low Income
This table presents the estimates of the measurement error bias factor λ for urban versus rural areas over the periods in our sample for our two main measures of financial development in Panal A, and similarly for high income (Rich) versus low income (Poor) counties in Panel B. We define counties as urban or rural based on whether their median population is above or below the sample median. Rich/Poor are analogously defined, but the sample is limited to the 1960 onwards when county-level income data are available. *PreWar* is 1920-1945, *PostWar* is 1950-1980, and *Modern* is 1985-2000.

Panel A: Urban versus Rural			
	PreWar (1)	PostWar (2)	Modern (3)
Banks/10k – Urban	-0.510	-0.804	-0.496
Banks/10k – Rural	-0.233	-0.182	-0.163
log(Banks/10k) – Urban	0.301	0.264	0.174
log(Banks/10k) – Rural	0.463	0.373	0.154

Panel B: High versus Low Income			
	PreWar (1)	PostWar (2)	Modern (3)
Banks/10k – Rich	.	-0.700	-0.395
Banks/10k – Poor	.	-0.198	-0.141
log(Banks/10k) – Rich	.	0.196	0.113
log(Banks/10k) – Poor	.	0.374	0.282

the overestimates when using *Banks/10k* are much larger for Rich areas relative to poor ($\lambda_{F/10k}^{Rich} < \lambda_{F/10k}^{Poor} < 0$), and that the underestimates of the true effect of financial development is greater for relatively poorer areas ($\lambda_{\log(F/10k)}^{Poor} > \lambda_{\log(F/10k)}^{Rich} > 0$). Policy decisions based on mismeasured variables, therefore, can have significant distributional implications.

A.3. Omitted Variable Bias

The prior subsection shows that excluding nonbanks leads to measurement error bias when researchers try to link financial architecture to any generic outcome. Many times, researchers are more specifically trying to understand the structural relationship between banks and real outcomes. In this case, emprical researchers often estimate a form of the following regression:

$$y_i = \alpha + \psi_{short}(Banks_i) + \epsilon_i \tag{9}$$

However, such a regression excludes the presence of nonbanks, which leads to omitted variable bias. Ideally, the researcher would estimate the following:

$$y_i = \xi + \psi_{long}(Banks_i) + \gamma(Nonbanks_i) + \nu_i \tag{10}$$

The omitted bias can be written as follows:

$$\psi_{short} = \psi_{long} + \gamma \times \delta_{Nb\text{-}B} \tag{11}$$

ψ_{short} is the estimate from the "short" regression (9) that omits nonbanks; ψ_{long} represents the true structural relationship (10) the researcher would like to estimate; γ is the estimate on nonbanks in the long, structural regression (10); and $\delta_{Nb\text{-}B}$ represents that estimate of a regression of nonbanks on banks.

In light of the significant product overlap between banks and nonbanks, there is good reason to believe that nonbanks would also be related to most outcomes that banks affect ($\gamma > 0$). Further, as we have shown in Section 3.4, there is a strong positive relationship between bank and nonbank presence, which suggests that $\delta_{Nb\text{-}B} > 0$. In Panel A of Table AA.3, we regress county-level log(Nonbanks) on log(Banks) for each major time period: PreWar, PostWar, and Modern. We find that the coefficient estimates in the range of [0.90, 1.02] when only including state×year fixed effects (columns 1-3), and [0.47, 0.73] when controlling for local population (columns 4-6). These results show that the size of omitted variable bias is sizable, and increasing throughout our sample period.

For example, consider the auto loan market. If more banks is associated with lower auto loan rates or greater access, it is likely the case that more nonbanks could similarly drive rates down and access up. If it were the case that the effect of nonbanks is similar to banks ($\gamma \approx \psi_{long}$), then it follows that estimate the short regression (9) would overstate the true effect of banks by 47%–73%.

Table AA.3: Quantifying Omitted Variable Bias–Nonbanks on Banks
This table presents the estimates of the regression of the log of nonbanks (finance companies, savings and loans, credit unions) on the log of banks for various time periods. These estimates provide a necessary input to quantify the degree of omitted variable bias (equation 11). Specifically, it estimates the parameter $\delta_{Nb\text{-}B}$ for the omitted variable bias formula $\psi_{short} = \psi_{long} + \gamma \times \delta_{Nb\text{-}B}$. Panel A shows the estimates using all counties and Panel B divides the counties into Urban and Rural. We define counties as urban or rural based on whether their median population is above or below the sample median. *PreWar* is 1920-1945, *PostWar* is 1950-1980, and *Modern* is 1985-2000. All regressions present standard errors clustered at the county level.

Panel A: All Counties

	(1) PreWar	(2) PostWar	(3) Modern	(4) PreWar	(5) PostWar	(6) Modern
log(Bank branches)	0.90***	0.96***	1.02***	0.47***	0.55***	0.73***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)
log(population)				0.58***	0.50***	0.37***
				(<0.01)	(<0.01)	(<0.01)
State*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	364	598	441	364	598	441
Adjusted R-squared	0.66	0.64	0.73	0.76	0.72	0.76

Panel B: Urban versus Rural

	Urban			Rural		
	(1) PreWar	(2) PostWar	(3) Modern	(4) PreWar	(5) PostWar	(6) Modern
log(Bank branches)	0.53***	0.79***	0.93***	0.28**	0.20*	0.45***
	(<0.01)	(<0.01)	(<0.01)	(0.02)	(0.05)	(<0.01)
log(population)	0.51***	0.21*	0.12	0.33	0.62***	0.53***
	(<0.01)	(0.09)	(0.11)	(0.15)	(<0.01)	(<0.01)
State*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	227	283	203	137	315	238
Adjusted R-squared	0.74	0.70	0.78	0.54	0.29	0.40

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel B of Table AA.3 separately estimates the relationship between nonbanks and bank for urban counties (columns 1-3) and rural counties (columns 4-6). The estimates show that the pattern of increasingly larger relationships between nonbanks and banks over time hold for both urban and rural areas. However, the results show that the size of the relationship is 2–4 times larger for urban areas compared to rural areas. For example, consider again the case where nonbanks and bank have a similar effect on the outcome under study ($\gamma \approx \psi_{long}$). The estimated effect of banks on that outcome during the modern era while omitting nonbanks results in an overestimate of the true effect by 45% in rural areas and 93% in urban areas.

In sum, the omission of nonbanks from empirical research relating financial architecture to real outcomes can create significantly biased estimates. When using banks to measure local financial architecture, we have shown that estimates can have large measurement error bias, and the sign is contingent on the basic empirical choices such as whether or not to log transform the measure of local bank presence. This section also shows that trying to estimate the independent, structural relationship between banks and real outcomes is also fraught with challenges, with omitted variable bias leading to potentially significant overstatements of the true effect of banks.

A.3.3. Regression Calibration

Overall, a clear message emerges from this analysis. Earlier estimates of the effect of finance on outcomes such as economic growth that use only banks as the measure are often substantially biased and, depending on the parameterization, this bias can lead to an overestimate or underestimate of the true effect of finance. Moreover, we have shown substantial cross-sectional variation in this bias across different population density and affluence. Our analytical framework and empirical estimates provide guidance to empirical researchers on the sign and magnitude of bias as well as settings where the bias is likely to be higher.

We now provide a simple correction that can be used to estimate the true level of financial development when the researcher's data only include the number of banks, as is usually the case. We do so by providing a least squares estimate of $E[F|B]$ based on our data: the conditional expectation can be used in a regression calibration exercise to correct for the measurement error. Instead of using B, the researcher can use $E[F|B]$ to obtain a consistent estimate of the β in equation 4.

Table AA.4 provides regression estimates with the number of branches of financial institutions per 10,000 people (F/10k) as the dependent variable and bank branches per 10,000 people (B/10k) as the explanatory variable. Panel A provides full sample estimates in column (1) and PreWar, PostWar, and Modern subsamples in columns (2), (3), and (4). We also present the estimates based on whether the researcher is interested in urban counties (Panel B) or rural counties (Panel C). The relationship between banks and total institutions is larger in magnitude for urban areas than rural areas. These differences are consistent with the narrative that urban banks' lending is more based on hard information and collateral, which opened opportunities for finance companies, where as rural banks more willing to lend on soft information, as observed by Nugent (1939):

> "In general, banks were willing to lend only to the best risks among consumer applicants, but many circumstances influenced their credit policies. The most important determinant of the extent of consumer loans by banks appears, however, to have been the nature of the community. Liberal credit policies could readily be maintained in small stable communities where the financial worth, earning capacity, family history, and personal qualities of applicants were matters of common knowledge. In metropolitan cities, on the other hand, the difficulties of appraising credit worth on any basis other than the ownership of readily negotiable collateral led to the denial of bank credit to classes of consumers whose credit needs were supplied by banks in intimate communities."

Appendix Table A.2 present corresponding regression estimates when the researcher is considering the number of financial institution firms rather than branches. Depending on the specific application at hand, these estimates can be used as an approximate solution to address the mismeasurement problem.

Table AA.4: Regression Calibration for Financial Institution Branches

This table presents OLS estimates of the number of financial institution branches per 10,000 people on the number of banking branches per 10,000 people. Panels A, B, and C estimate the regressions for all counties, urban counties, and rural counties, respectively. Column (1) of each panel include observations from the entire sample period, while columns (2), (3), and (4) use data from the PreWar, PostWar, and Modern periods, respectively. We define counties as urban or rural based on whether their median population is above or below the sample median. *PreWar* is 1920-1945, *PostWar* is 1950-1980, and *Modern* is 1985-2000. All regressions present standard errors clustered at the county level.

Panel A: All Counties				
	(1)	(2)	(3)	(4)
Banks (Br) / 10k	1.17***	1.38***	1.39***	1.24***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Constant	1.44***	1.24***	1.39***	1.00***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Sample Period	All	PreWar	PostWar	Modern
Nobs	1,478	387	618	473
Adjusted R-squared	0.65	0.53	0.55	0.71
Panel B: Urban Counties				
	(1)	(2)	(3)	(4)
Banks (Br) / 10k	1.39***	1.51***	1.80***	1.50***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Constant	1.19***	1.08***	1.07***	0.63***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Sample Period	All	PreWar	PostWar	Modern
Nobs	732	236	288	208
Adjusted R-squared	0.69	0.48	0.68	0.77
Panel C: Rural Counties				
	(1)	(2)	(3)	(4)
Banks (Br) / 10k	1.06***	1.23***	1.18***	1.16***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Constant	1.64***	1.51***	1.61***	1.09***
	(<0.01)	(<0.01)	(<0.01)	(<0.01)
Sample Period	All	PreWar	PostWar	Modern
Nobs	746	151	330	265
Adjusted R-squared	0.61	0.48	0.45	0.69

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$