Markups, Markdowns and Bankruptcy in the Banking Industry

Andrea Pasqualini*
University of Oxford

October 2021

Abstract

I develop a new structural approach for the joint estimation of markups on lending rates and markdowns on deposit rates for all US depository institutions between 1992 and 2019. Markups (markdowns) are wedges between the observed price for the output (input) good and the price that would realize if the bank was a price taker on that market. Gross markups have been generally decreasing over time with some procyclical variation, with an average value of 2. Gross markdowns do not display a trend, but feature strong countercyclical variation. The average gross markdown is 1.5. The Herfindahl-Hirschman Index and the Boone indicator are significantly different measures. I show that higher markups are associated with a lower bankruptcy probability, which is in contrast with previously known results. Instead, markdowns correlate positively with default probabilities. When considered jointly, markups and markdowns both correlate negatively with the probability of bankruptcy.

JEL Codes: E44, L11, L16, G21
Keywords: banks, depository institutions, markups, markdowns, bankruptcy

*Email: andrea.pasqualini@economics.ox.ac.uk. Website: https://andrea.pasqualini.io. I am grateful to Basile Grassi for his continued support during the project. I also thank Michael McMahon, Francesco Zanetti, Federica Romei, Sergio De Ferra, Tommaso Monacelli, Rustam Jamilov, Filippo De Marco, Nicolas Serrano-Velarde, Antonella Trigari, Giacomo Calzolari, Melika Liporace, participants to the ASSET 2021 conference and various seminar participants for helpful comments and suggestions. I gratefully acknowledge financial support from the *Fondazione Romeo ed Enrica Invernizzi*. Any error is my own.

1 Introduction

The Great Recession highlighted the importance of avoiding bankruptcies in the banking system. Academics and policy makers focused on assessing and improving financial stability, both at the aggregate and at the bank level. Many papers focused on the relationship between the probability of bankruptcy and market power banks have. Particularly on this topic, academics have not reached a clear consensus.

On the other hand, recent literature in Economics has brought the importance of market power front and center, especially in Macroeconomics. The literature acknowledges market power on output markets and recently started to investigate market power on input markets. Firms may exert monopsonistic power, which lowers the observed price for their input goods relative to the perfect competition benchmark. This conduct contributes to lowering marginal costs and, therefore, profit rates. In the banking literature, competition plays many roles and there is little agreement about the desirable state. Existing contributions highlight the importance of competition both on lending and deposit markets.

In this paper, I tackle these issues with a novel approach. I use the production approach to the estimation of markups and markdowns found in De Loecker and Warzynski (2012) and Morlacco (2020) to banking data. Given balance sheet and income statement data, the production approach allows for identification of markups and markdowns. Throughout this text, the markup is the wedge between the price for the output good one observes and the price one would observe were the seller a price taker. Conversely, the markdown is the wedge between the price for the input good we observe and the price that would realize if the buyer was a price taker. The strength of this methodology lies in the data requirement. While markups are typically thought of as arising from the elasticity of demand for the output good and markdowns from the elasticity of supply for the input good, the production approach does not require demand and supply data. Originally based on Hall (1988), the production approach consists of a structural model of firm behavior. The first-order condition to a cost-minimization problem links the unobservable markup to expenditure shares, which can be observed directly, and production elasticities, which can be estimated. Morlacco (2020) expands this methodology in order to allow for buyer's market power, therefore introducing markdowns. I jointly estimate both wedges for loan and deposit markets. For identification, I assume that banks hire labor on perfectly competitive markets and that both labor and deposits are not subject to adjustment costs.[1]

[1]The assumptions of no adjustment costs are credible given that I use yearly balance sheet data for the main results. At the yearly frequency, banks are unlikely to face significant frictions in hiring labor or taking deposits. On the other hand, while the assumption of no monopsonistic power goes against several papers documenting buyers' power on labor markets, it may be reasonable to believe that banks do not require highly specialized workers, such as tellers and administrative staff. They can therefore hire labor from a pool where banks compete with other industries.

I find that markups on lending rates are trending downwards between 1992 and 2019, while markdowns on deposit rates have substantially increased after the Great Recession. The average markup on lending rates is roughly 2, while the average markdown on deposit rates is roughly 1.5. These figures mean that lending rates are roughly twice as much as their price-taking benchmark. Deposit rates are instead roughly two thirds of their price-taking counterpart. The yearly cross-sectional dispersion of markups remains somewhat stable over the sample period, while the dispersion of markdowns follows the level: increases in the average (or median) markup are associated with increases in the cross-sectional dispersion, and vice versa. Markups correlate positively with the Federal Funds rate. Markdowns instead increase at the zero lower bound, but seem acyclical before the great recession. I compute the correlations with certain observable bank characteristics. Bigger banks tend to charge more markups and less markdowns, except for bank at the top of the asset distribution, which feature both higher markups and higher markdowns. Banks with relatively higher leverage have higher markups, but lower markdowns. Interestingly, profitability seems to correlate with neither markups nor markdowns. Banks that have a higher share of loans among their earning assets charge higher markups, but lower markdowns. Finally, banks that pay a larger share of their income as dividends are associated with higher markups and markdowns.

I compare markups and markdowns to the Herfindahl-Hirschman Index (HHI) computed at the state level for both deposits and loans and to the Boone (2008) indicator. HHIs are usually taken as measures of market power by policy makers, both for banks and manufacturing firms. The Boone indicator relies on an output reallocation effect, such that harsher competition increases (in absolute value) the elasticity of profits to marginal costs. On one hand, the HHI on deposits correlates positively with markdowns on deposit rates, although weakly. Instead, the HHI on loans correlates negatively with markups on lending rates. On the other hand, the Boone indicator barely correlates at all with neither markups nor markdowns. These findings suggest that HHIs or the Boone indicator do not capture the same phenomena as markups and markdowns.

I compute the Z-score (Altman, 1968) and the O-score (Ohlson, 1980) as measures of default probability. These consists of the predicted values of two regressions, where the left-hand side is an indicator of bankruptcy, which equals one if the bank will default within the following year. The two scores differ for the right-hand side terms, which are generally balance sheet ratios. The scores can be interpreted as probabilities of bankruptcy that can be predicted (in sample) by reports of condition, such as balance sheets. The advantage of these measures is that they require no data other than balance sheets. I regress these probabilities of default on markups and markdowns. The goal of such regression is to assess the *correlation* between market power and financial stability. Interestingly, I find that markups correlate negatively with default probabilities. The correlation is small, although significant. I also find that markdowns are positively correlated with default probabilities, although the sign reverses once I contemporaneously account for markups. While I control for bank fixed effects, this result may be driven by a generalized increase in bank size that occurs across the whole cross section over time. Bigger banks rely more on markups than markdowns. At the same time, there has been considerable mergers or acquisitions activity throughout the sample period, increasing the size of the average bank.

This paper brings five contributions. Firstly, I provide evidence of large bank market power, both on loan and deposit markets. Several existing papers are concerned with the presence of bank market power and its macroeconomic impact. For example, Scharfstein and Sunderam (2016) and Drechsler et al. (2017) show evidence of market concentration in the loan and deposit markets respectively. Each of those papers link bank market power to the transmission of monetary policy through two distinct channels. Wang et al. (2021) obtain estimates of markups and markdowns in the banking industry similar to mine, although I employ quite a different methodology. They treat markups and markdowns as latent variables in a macroeconomic model with heterogeneous banks, and obtain estimates by using the Simulated Method of Moments. Conversely, I estimate the same quantities directly from microdata with a methodology adapted from practices in Industrial Organization.

The second contribution consists of estimating a loan production function, along with a clean identification argument that applies specifically to the financial intermediation sector. Ackerberg et al. (2015) illustrate the econometric methodology for manufacturing firms, building upon Olley and Pakes (1996) and Levinsohn and Petrin (2003). The contribution in all these papers consists of dealing with an endogeneity issue that affects the OLS estimates of production functions. Endogeneity arises because the unobserved productivity shock in the OLS specification contains information that may be known to the firm, but not the researcher. The papers cited above suggest using a complementary input good, such as intermediate materials, to instrument the unobserved component that causes the endogeneity issue. However, banks are not subject to productivity shocks. I reinterpret productivity shocks as loan repayment shocks. To see this, consider that the output of a bank is total outstanding loans, net of a loan loss allowance. Repayment shocks naturally lend themselves to be unforeseen shocks that affect net loans either positively or negatively. A positive repayment shock is a loan repayment that the bank thought was noncollectable and has therefore set aside as loan loss in its accounting. A negative repayment shock is a loan repayment that the bank thought would occur and has not considered adding to its loan loss allowance. A bank may have some private information about repayment shocks that a researcher may not possess. In this paper, I instrument the component the bank may have information on using loan loss provisions (that is, the period contribution to the total allowance fund). This is a variable banks may adjust based on what information they have regarding their outstanding loans. The choice of this variable as an instrument is further motivated by regulatory constraints: banks need to justify their loan loss provisions in the quarterly call reports, and such justifications must be related to a bank's assessment of the riskiness of each loan.

The third contribution deals with the estimation of markups (and markdowns) using the production approach. The production approach to the estimation of markups was first introduced by Hall (1988). De Loecker et al. (2020) later developed the full procedure, which relies on the estimation of the production function. More recently, De Loecker et al. (2020) and related papers from the same authors employed the methodology to document global trend in markups across various industries. I contribute to this literature by providing evidence that more narrowly focuses on banks. This is motivated by the fact that the concept of production function does not obviously relate to banks. Morlacco (2020) modified the technique in De Loecker and Warzynski (2012) in order to accommodate market power in input markets. While the aforementioned papers deal with manufacturing firms, this paper brings the same insights to the banking sector. I provide empirical evidence of markups and markdowns specifically about the banking sector.

The fourth contribution consists in showing that markups and markdowns are significantly different than HHIs and the Boone (2008) indicator. In particular, HHIs on either loan or deposit are subject to confounding the channels of market power: what increases the HHI on loans may very well increase the HHI on deposits. For example, consider the large amount of mergers and acquisitions (M&A) activity in the banking sector during the Great Recession. Many M&A deals were either encouraged or facilitated by Central Banks, both in the US and in Europe, for the sake of soundness of banks. These M&A deals have the likely effect of increasing the HHI, with market power playing little to no role.[2] Conversely, the Boone indicator implicitly assumes that firms only exert power on the market for outputs. This is due to the fact that firms are assumed to take their cost functions and, hence their marginal costs, as given. Instead, banks are now known to exert part of their power on deposit markets, by offering particularly low interest rates. This highlights the fact that marginal costs can be strategic variables and banks can effectively lower them.

Finally, the fifth strand of the literature I contribute to deals with the relationship between idiosyncratic financial stability and market power. This literature contains conflicting results, both theoretically and empirically. For example, Hellmann et al. (2000) argue that competition may be detrimental for financial stability. Banks realize profits in a scarcely competitive environment, which can be accumulated and may serve as buffer against adverse shocks. Conversely, Boyd and De Nicoló (2005) argue that competition may foster financial stability. Banks with more market power do realize more profits, but also induce higher loan interest rates. This may increase risk-taking attitudes of firms that apply for loans. Empirically, Beck et al. (2006) show

[2] If anything, a bank that *has to* undergo a M&A deal has very little market power, because the alternative would be bankruptcy.

that more concentrated banking systems are associated with more financially stable economies. Instead, Schaeck et al. (2009) use the H-statistic of Panzar and Rosse (1987) to find that more competitive environments are associated with more stable banks. Berger et al. (2009) find that banks with more market power have less risk exposure. While this literature tends to focus on lending markets, I disentangle the effect of market power on lending and deposit markets. I *jointly* estimate markups and markdowns. I show that higher markups are associated with lower bank default probabilities, while the opposite holds for markdowns. Once I account for both markups and markdowns, I find that they both correlate negatively with the probability of bankruptcy. This change of sign for markdowns provides further motivation to the approach: it is important to disentangle market power for output goods and market power for inputs. In particular, a positive relationship between market power and bankruptcy is more likely associated to power being exerted on loan markets. Conversely, a negative relationship is probably due to buyers power on deposit markets. Alternative methods that do not disentangle these two sources of market power may misinterpret the role of competition for the stability of the banking system.

The rest of the paper is organized as follows. Section 2 details the production approach for the estimation of markups and markdowns, together with its application to the banking sector. Section 3 presents the data I use. Section 4 shows the empirical results on markups on lending rates and markdowns on deposit rates, together with correlations with observable bank characteristics and a comparison with the Herfindahl-Hirschman Indices and the Boone indicator. Section 5 shows details on the measures of default probabilities I use. Section 6 presents the results that relate market power on financial stability. Section 7 concludes.

2 The production approach for the estimation of markups and markdowns

The production approach for the estimation of markups and markdowns is detailed in De Loecker and Warzynski (2012) and Morlacco (2020). It relies on a simple structural model of firm behavior given a cost function and a production function. Before I delve into the details of the methodology, it is useful to clarify its use in the context of banking.

2.1 Conceptual framework

Production functions typically belong to the realm of manufacturing firms. They describe the transformation from input goods into output goods in a concise way. They are often referred to as "black boxes" because they do not describe how exactly such transformation takes place. Production functions are rarely encountered in the literature on banking: existing papers have focused on the roles of banks as intermediaries, exploring the economic mechanisms that justify such roles. I do not explicitly model any specific role of banks. Instead, I take those roles for granted and I model them with a production function. For banks, this function includes all the economic frictions that a bank addresses, such as informational asymmetries. In this sense a loan production function can be seen as a reduced-form characterization of the activity of the bank that does not ignore their economic role.

I assume that banks collect sources of financing, such as deposits and equity, and use them to provide loans. To do so, banks also need traditional input goods, such as labor and capital. This approach follows directly Sealey and Lindley (1977), who characterize the activity of a bank in terms of classical production theory. In their paper, the authors also describe the main difference between a bank and a manufacturing firm in terms of production. A manufacturing firm requires capital and labor to produce a physical good. Instead, banks *also* require sources of financing in order to supply a loan. To see this, consider a bank in a frictionless, simplified world that only uses deposits and labor and that such inputs are already being efficiently exhausted. Suppose that this bank can hire more labor, but cannot raise an additional unit of deposits. In this case, the bank cannot increase its outstanding loans, because the balance sheet constraint binds. On the other hand, if such bank can source additional deposits but cannot hire more labor, then it does not have the capacity to process more loans. Therefore, there is a degree of complementarity between sources of financing and physical input goods. Generally speaking, the production feasibility set of a bank is affected by the balance sheet constraint, the need for physical goods (e.g., premises, IT equipment, labor), regulatory constraints and the sources of risk, such as a creditor's default risk and the risk of bank runs.

In reality, banks also use sources of financing other than deposits, such as equity and, in some cases, corporate bonds. Some banks can also be seen as multi-output firms, because some buy financial assets and repackage them as securities to be either held or traded. Because I focus on depository institutions, most of earning assets in those banks are made of loans, as the summary statistics below show. Additionally, equity and deposits almost entirely describe the liabilities side of balance sheets. For these reasons, I focus on a single-product production function, where the output is loans, and I restrict my attention to deposits and equity as the only sources of finances.

2.2 The structural model

Building on De Loecker and Warzynski (2012) and Morlacco (2020), I assume that every bank i in period t solves the following static cost minimization problem, subject to a production function and for a given level of outstanding loans:

$$\min_{D_{it},E_{it},N_{it},K_{it}} r^D_{it} D_{it} + r^E_{it} E_{it} + w_{it} N_{it} + r^K_{it} K_{it} \tag{1}$$
$$\text{subject to } \; L_{it} = F\left(D_{it}, E_{it}, N_{it}, K_{it}\right),$$

where L_{it} are loans, D_{it} are deposits, E_{it} is equity, N_{it} is labor, K_{it} is capital, r^D_{it}, r^E_{it}, and r^K_{it} are the input interest rates paid on deposits, equity and capital respectively, w_{it} is the wage per efficient labor unit and $F(\cdot)$ is the loan production function. Given a level of outstanding loans L_{it}, the solution to this problem characterizes the optimal mix of physical input goods and financial assets to use in loan production. In order to identify the markup on lending rates, it is necessary to assume price-taking behavior and no adjustment costs for at least one input good or asset. Conversely, in order to identify the markdown on deposit rates, it is necessary to have an identified measure of markup on lending rates and, additionally, to assume that banks are not price-takers on deposit markets and that deposits are not subject to adjustment costs.

I assume that labor satisfies the required assumptions for identification of markups on lending rates. Two reasons justify this assumption. First, it is arguably the case that banks are not price-takers on the markets for deposits, while equity and capital may well be subject to adjustment costs. In particular, one objective of this paper is to identify a markdown on deposit rates. Second, banks compete with other industries for low-skilled workers and for administrative staff on one hand. On the other, banks may not perfectly compete with other industries for mid- or top-management workers. However, the more one climbs the job ladder within a bank and the more she is likely to be also paid with other forms of compensation, such as stock options, rather than wages. Because of this, the compensation over which banks may have some monopsonistic power would not appear in wages, but rather in other balance sheet items.

The first order condition for (1) with respect to labor is

$$w_{it} = \lambda_{it} \frac{\partial F}{\partial N_{it}},$$

where λ_{it} is the Lagrange multiplier associated to the production function and corresponds to the marginal cost of loan production. By multiplying each side by $N_{it}/(r^L_{it} L_{it})$, where r^L_{it} is the interest rate on loans a bank charges, and rearranging terms we obtain the following expression:

$$\underbrace{\frac{r^L_{it}}{\lambda_{it}}}_{\mu_{it}} = \underbrace{\left[\frac{\partial F}{\partial N_{it}} \cdot \frac{N_{it}}{L_{it}}\right]}_{\theta^N_{it}} \cdot \underbrace{\left[\frac{w_{it} N_{it}}{r^L_{it} L_{it}}\right]^{-1}}_{1/\alpha^N_{it}}. \tag{2}$$

The left-hand side is the interest rate on loans divided by the marginal cost of loan production, that is the gross markup on the lending rate, μ_{it}. The right-hand side is made of two components:

the first, θ_{it}^N, is the elasticity of loan production to labor and the second, $1/\alpha_{it}^N$, is the inverse expenditure share of labor relative to loan interest income. This expression has operational content. It implies that we can compute the unobservable markup given the inverse expenditure share, which is readily available in balance sheet data, and the production function elasticity, which can be estimated given a panel of banks.

Consider now the first-order condition to (1) with respect to deposits D_{it}. Repeating the steps taken above, the first-order condition is

$$\underbrace{\left[1 + \frac{\partial r_{it}^D}{\partial D_{it}} \cdot \frac{D_{it}}{r_{it}^D}\right]}_{\psi_{it}} \cdot \underbrace{\frac{r_{it}^L}{\lambda_{it}}}_{\mu_{it}} = \underbrace{\left[\frac{\partial F}{\partial D_{it}} \cdot \frac{D_{it}}{L_{it}}\right]}_{\theta_{it}^D} \cdot \underbrace{\left[\frac{r_{it}^D D_{it}}{r_{it}^L L_{it}}\right]^{-1}}_{1/\alpha_{it}^D}, \tag{3}$$

where the term ψ_{it} relates to the inverse supply elasticity of deposits. Morlacco (2020), who focuses on French manufacturing firms, interprets ψ_{it} as a *markdown*. Formally, the markdown is here defined as the wedge between the deposit rate banks pay relative to the interest rate that banks would pay if they were price-takers on deposit markets. In order to empirically recover the markdown component, one needs the same ingredients as before (i.e., the inverse expenditure share and the elasticity of the production function) and, additionally, a measure of markups. Assuming that Equation (2) identifies the markup, one can take the ratio between Equation (3) and (2) and identify the markdown:

$$\frac{\theta_{it}^D/\alpha_{it}^D}{\theta_{it}^N/\alpha_{it}^N} = \frac{\mu_{it}\psi_{it}}{\mu_{it}} = \psi_{it}. \tag{4}$$

To recap, the production approach to the estimation of markups and markdowns consists of three steps. The first is estimating the production function elasticity with respect to each input good. This exercise is standard in the empirical Industrial Organization literature. The second is computing the inverse expenditure shares of each input good, which is trivial given income statement variables. The third and final step is choosing two input goods such that one is subject to neither monopsonistic competition nor adjustment costs and the other is not subject to adjustment costs. The first-order condition with respect to the first good allows for identification of the markup on lending rates, while the ratio of first-order conditions identifies the markdown. While the second and third steps are trivial, the first one requires some attention and I therefore turn to it now.

2.3 Estimation of the production function and production elasticities

Consider the following net loan production function for bank i at period t, with lowercase variables being the logs of their uppercase variants:

$$l_{it} = f\left(d_{it}, e_{it}, n_{it}, k_{it}; \beta\right) + \tilde{\varepsilon}_{it}, \tag{5}$$

where β is the vector of production function parameters and $\tilde{\varepsilon}_{it}$ is normally referred to as an unobserved productivity term. In the context of financial assets as loans, it is not clear what productivity means. In this paper, I assume that net loans are subject to repayment shocks. When $\tilde{\varepsilon}_{it}$ is positive, the bank receives a repayment from a loan that was not expected to realize. Conversely, when $\tilde{\varepsilon}_{it}$ is negative, the bank does not receive a repayment from a loan that was instead expected to realize. Each bank has private information regarding the repayment shocks $\tilde{\varepsilon}_{it}$ while the researcher does not. Therefore, the OLS estimate of β will be subject to an endogeneity issue because of an omitted variable bias. The OLS estimate of β cannot be reliably interpreted as (the vector of) the loan elasticities to the input goods. This endogeneity issue is well documented in Olley and Pakes (1996), Levinsohn and Petrin (2003) and Ackerberg et al. (2015).

In this paper, I follow the approach of Ackerberg et al. (2015) in estimating the production function parameters. Suppose that the repayment shock $\tilde{\varepsilon}_{it}$ can be decomposed in two additive terms (in logs): a term that is known by bank i at time t, ω_{it}, and a term that is unknown to both the bank and the researcher, ε_{it}. Therefore, we can rewrite Equation (5) as

$$l_{it} = f\left(d_{it}, e_{it}, n_{it}, k_{it}; \beta\right) + \omega_{it} + \varepsilon_{it}.$$

To identify the production function parameters β, Ackerberg et al. (2015) propose a two-step GMM approach. Identification of β occurs at the second stage. Suppose that there exists an intermediate, complementary production good or asset that the bank chooses based also on the privately observed term ω_{it}. Such complementary good does not appear in the production function (5) because it is a value-added specification. In the context of manufacturing firms, such good can be materials, and is therefore denoted as m_{it}. In the context of banks, there is no such thing as materials. However, I write net loans in the value-added production function. A control variable for banks that appears in the balance sheets is the loan loss provisions, which reflects the fraction of repayments of gross loans the banks deems noncollectable. Although subject to some regulatory constraints, the determination of each period's loan loss provision is up to each bank, as each one is expected to have private information about its customers. Let

m_{it} be determined by the following demand function:

$$m_{it} = h\left(d_{it}, e_{it}, n_{it}, k_{it}, \omega_{it}\right). \tag{6}$$

This demand function arises from the optimization problem in (1) when the constraint is not the value-added production function but, rather, the gross production function, where intermediate goods or assets would appear. Assume that the function h is invertible with respect to ω_{it}, such that we can write

$$\omega_{it} = h^{-1}\left(d_{it}, e_{it}, n_{it}, k_{it}, m_{it}\right).$$

Plug this expression in Equation (5) to obtain the following:

$$\begin{aligned} l_{it} &= f\left(d_{it}, e_{it}, n_{it}, k_{it}; \beta\right) + h^{-1}\left(d_{it}, e_{it}, n_{it}, k_{it}, m_{it}\right) + \varepsilon_{it} \\ &= \Phi\left(d_{it}, e_{it}, n_{it}, k_{it}, m_{it}; \beta\right) + \varepsilon_{it}. \end{aligned} \tag{7}$$

Equation (7) constitutes the first step in the estimation procedure. It is estimated with OLS where the function $\Phi(\cdot)$ is approximated with a n-th order polynomial. Let Φ_{it} denote the predicted values of the regression.

The second step consists of a GMM estimation. From Equation (5), and given a value for β, we have that

$$\omega_{it}(\beta) = \Phi_{it} - f\left(d_{it}, e_{it}, n_{it}, k_{it}; \beta\right).$$

A sufficient condition for identification of β is that ω_{it} follows a Markov process at the bank level. For concreteness, I assume that the repayment shock ω_{it} follows an AR(1) process for each bank i:

$$\omega_{it} = \rho\omega_{it-1} + \xi_{it}, \tag{8}$$

where the innovation term ξ_{it} is not in the information set of each bank. The GMM moment condition requires that

$$\begin{aligned} 0 &= \mathbf{E}_{it}\left(\xi_{it}(\beta)\right) \\ 0 &= \mathbf{E}_{it}\left(\Phi_{it} - f\left(d_{it}, e_{it}, n_{it}, k_{it}; \beta\right) - \rho\left[\Phi_{it-1} - f\left(d_{it-1}, e_{it-1}, n_{it-1}, k_{it-1}; \beta\right)\right]\right). \end{aligned}$$

Note that the expectation is conditional on the information set of bank i at time t, which makes it operationally difficult to deal with. Following Ackerberg et al. (2015), I instrument the conditioning information set with a vector of variables that I assume not to correlate with ξ_{it}. In particular, the instrumented GMM condition is

$$\mathbf{E}\left(\left[\Phi_{it} - f\left(d_{it}, e_{it}, n_{it}, k_{it}; \beta\right) - \rho\left[\Phi_{it-1} - f\left(d_{it-1}, e_{it-1}, n_{it-1}, k_{it-1}; \beta\right)\right]\right] \otimes \begin{bmatrix} 1 \\ k_{it} \\ l_{it-1} \\ \Phi_{it-1} \end{bmatrix}\right) = 0. \quad (9)$$

Operationally, the variable l_{it} is the log of total net outstanding loans, d_{it} is the log of total domestic deposits, e_{it} is the log of total equity, n_{it} is the log-expenditure on labor and k_{it} is the log of premises and equipment. I specify the production function (5) to be Cobb-Douglas. I implement the second step in the estimation of the production function with a numerical root-finding routine. I set the initial condition for β to the OLS estimate of Equation (5).

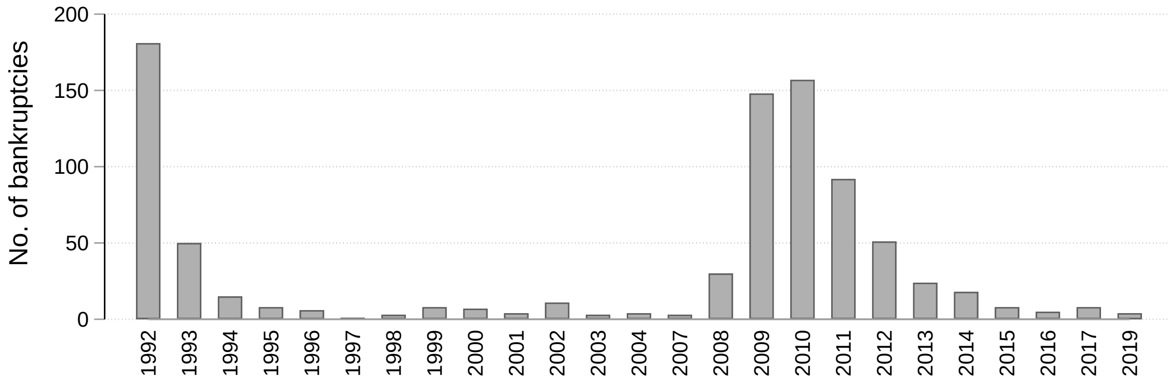
Assuming a Cobb-Douglas production function implies that the loan elasticities to each input good or asset are constants across time and banks. Hence, the cross-sectional and time-series features of markups on lending rates are driven by the features of the expenditure shares. The elasticities simply rescale the expenditure shares. This is easily seen in Equation (2).

3 Data and summary statistics

I use data from the Federal Deposit Insurance Corporation (FDIC). They maintain and provide the Statistics on Depository Institutions (SDI). These are balance sheet, income statement and demographic variables available at the quarterly frequency. They are compiled from the quarterly Call Reports, which are reports of condition and income. Each depository institution is required by law to fill the Call Reports. The structure of the filled form and the amount of detail in reported information depends on the amount of total assets and on whether banks have only domestic or domestic and foreign offices. The data are publicly available starting from 1992Q4. I use figures from 1992Q4 to 2019Q4. Income statement variables are cumulated within each fiscal year for each bank. When I take within year-bank differences to obtain the non-cumulated version of the variables, I observe significant seasonal variation at the year-bank level for all income statement variables. For example, cash dividends are typically registered in the income statement only at the end of the year. For this reason, I focus my attention to end-year observations, effectively using data at the yearly frequency. I estimate the production function also using quarterly data as a robustness check.

I use data from the Bank Failure and Assistance database, also provided by the FDIC, in order to obtain an indicator of bankruptcy for each bank. These data report the date of bankruptcy for every failed bank to date. Additionally, the FDIC reports the type of settlement after the bankruptcy: failure or assistance. In the former case, the financial institution is liquidated. In the latter, the FDIC provides guidance so that the bankrupt institution is acquired by another

Figure 1: Bankruptcies in sample period.



bank. From these data, I compile an indicator variable for each bank present in the SDI. The variable equals one if, in that year, the bank went bankrupt.

Table 1 shows summary statistics for the cross-sections of banks in 1992 and 2019. Their comparison provides an indication of changes in the US banking industry over 27 years. Each table groups statistics by percentile brackets of total assets. The number of depository institutions in the US went from 13,973 in 1992 to 5,186 in 2019. Roughly 48 percent of total banking system assets belonged to the top size percentile in 1992 and roughly 75 percent in 2019. The number of existing banks reduced over time. This is a first indicator that banking activities concentrated in bigger banks over time.

The composition of the balance sheets of banks has slightly changed over time. Net loans represented 51 to 58 percent of total assets in 1992 and 59 to 70 percent in 2019. Compared to held securities, net loans make for the majority of earning assets. Financing is primarily given by deposits, which backed roughly 73 to 88 percent of total assets in 1992 and 77 to 83 percent in 2019. Equity has become more important over time, backing roughly 7 to 9 percent of total assets in 1992 and 11 to 13 percent in 2019.

The primary source of income has always been interest from loans, ranging from roughly 55 to 64 percent of total income in 1992 and from 56 to 72 in 2019. Interest income from held securities decreased over time, from approximately 17 to 26 percent of total income in 1992 to 9 to 13 in 2019. Somewhat surprisingly, service charges on deposits are a relatively small source of income for banks, accounting for roughly 2-3 percent of total income both at the beginning and at the end of the sample. This reinforces the assumption that deposits are input goods. Although banks may compete over commission fees, they are not an important source of revenue for banks.

Figure 1 shows the number of bankruptcies in the sample period. There has been a number of bankruptcies at the beginning of the sample and right after the Great Recession. Of the 849 bankruptcies in the sample period, 640 resolved with other banks acquiring all deposits (insured

Table 1: Summary statistics at the beginning and at the end of the sample period.

	1992						2019					
Percentile bracket	[0, 75)	[75, 90)	[90, 95)	[95, 98)	[98, 99)	[99, 100]	[0, 75)	[75, 90)	[90, 95)	[95, 98)	[98, 99)	[99, 100]
No. of banks	10,479	2,096	699	419	140	140	3,889	778	259	156	52	52
Average total assets (bln USD)	0.089	0.370	0.900	2.448	6.258	28.050	0.200	0.958	2.537	8.014	25.297	269.025
Median total assets (bln USD)	0.075	0.344	0.843	2.196	6.136	16.237	0.160	0.897	2.318	6.810	24.757	117.848
Average income (bln USD)	0.008	0.031	0.077	0.206	0.547	2.602	0.011	0.052	0.138	0.411	1.290	14.148
Median income (bln USD)	0.006	0.029	0.071	0.179	0.498	1.431	0.008	0.037	0.093	0.264	0.816	9.000
Average expense (bln USD)	0.006	0.025	0.062	0.169	0.429	2.091	0.008	0.045	0.115	0.341	1.271	6.492
Median expense (bln USD)	0.005	0.023	0.057	0.144	0.398	1.162	0.006	0.032	0.077	0.212	0.739	4.041
Average NIM / assets (%)	4.227	3.935	3.863	3.674	3.607	3.576	3.460	3.423	3.325	3.305	3.293	3.144
Percentage of total system assets (%)	11.386	9.503	7.710	12.566	10.732	48.102	4.153	3.977	3.507	6.673	7.021	74.669
Average percentage relative to total assets in size category												
Net loans	51.821	57.297	59.242	57.662	57.984	58.907	63.186	70.388	69.810	70.024	68.123	59.389
Securities	31.508	28.643	26.603	26.641	24.572	21.420	19.057	16.392	16.666	17.806	17.074	21.885
Intangible capital	0.127	0.193	0.320	0.602	0.668	0.604	1.601	1.705	1.476	1.261	1.215	0.797
Physical capital	1.603	1.588	1.396	1.318	1.264	1.286	0.330	0.836	1.113	2.069	3.378	2.094
Deposits	88.194	87.272	84.978	80.650	75.862	73.357	83.179	82.696	81.198	78.660	76.812	77.338
Equity	9.559	8.221	7.415	7.403	7.284	6.889	13.096	11.973	11.729	12.634	13.037	11.191
Average percentage relative to income in size category												
Int. income from loans	60.655	63.799	64.486	62.516	59.799	55.750	71.548	72.198	69.929	69.366	69.500	56.787
Int. income from securities	26.681	24.112	22.168	23.580	18.777	17.344	10.717	8.755	9.392	10.962	9.026	12.850
Int. income from lease financing receivables	0.151	0.195	0.402	0.488	1.224	1.322	0.194	0.213	0.370	0.437	0.500	1.029
Income from charges on deposits	3.017	2.557	2.373	2.441	2.912	2.797	3.607	3.150	3.348	3.152	3.427	3.158
Average percentage relative to expenses in size category												
Int. expense on deposits	51.196	52.147	49.648	45.851	40.652	34.539	19.053	21.197	22.631	23.381	21.623	22.646
Wages and salaries	23.090	20.742	19.612	18.063	18.493	19.750	44.995	44.264	41.817	38.733	35.294	32.269
Cash dividends	6.343	6.293	5.163	5.344	4.856	5.923	16.964	16.728	16.882	28.960	27.209	36.589

Dollar figures are adjusted for inflation and expressed in terms of 2019 US dollars. The term "percentile bracket" refers to the cross-sectional distribution of total assets within each year.

and uninsured) and some assets, 126 resolved with other banks acquiring only insured deposits, 51 resulted in complete payouts, 17 resulted in other institutions paying insured deposits (without acquiring them) and 15 resulted in assisted transactions, where the FDIC managed transactions across banks such that the bankrupt institution's charter survives. These numbers highlight the relevance of acquisitions or assumptions relative to payouts or assisted transactions.

4 Markups and markdowns in the banking industry

In this section, I present the results on the estimation of markups and markdowns. I also describe their correlation with observable bank characteristics and I compare them with Herfindahl-Hirschman indices and with the Boone indicator.

4.1 Evolution of markups and markdowns over time

Following the procedure detailed in Section 2, I compute the markups and markdowns on lending and deposit rates respectively. Table 2 reports the estimates of the production function parameters, assuming that the production function is Cobb-Douglas. I report the results using OLS on Equation (5), together with the GMM results using the procedure described above. The estimates also represent the production function elasticities with respect to each input good *ceteris paribus*, because of the functional form and because all variables are expressed in logs. The standard errors on GMM estimates are bootstrapped, in line with the recommendation in Ackerberg et al. (2015).

First, the OLS and GMM estimates are different, not only in terms of point estimates but also on standard errors. This comes from the fact that OLS estimates suffer an omitted variable bias. Second, the results show the predominant elasticity of loans with respect to deposits. A one percent increase in deposits translates into a roughly 0.54 percent increase in net loans. The elasticity of loans to equity is roughly 0.20, to labor is 0.28 and to physical capital is 0.03. For robustness, I also report the estimates using quarterly data. Quarterly income statement variables are cumulative at the bank-year level. For this reason, I take their first-differences within every bank-year pair. The results are qualitatively comparable. The results on markups and markdowns that follow rely on the GMM estimates using the yearly data.

Figure 2 shows the time series behavior of the average, median and interquartile range of markups on lending rates. The average net markup went from roughly 50 percent in 1992 to roughly 25 percent in 2019. Overall, markups have been trending downwards over time, with a temporary increase before the Great Recession.

Figure 3 shows instead the time series of the average, median and interquartile range of markdowns on deposit rates. The average markdown in 2019 is roughly ten times higher than

Table 2: Estimates of the production function parameters.

	Net loans			
	Yearly data		Quarterly data	
	OLS	GMM	OLS	GMM
Deposits	0.6376	0.7342	0.7062	0.6899
	(0.002)	(0.075)	(0.001)	(0.053)
Equity	0.1812	0.1464	0.2091	0.2386
	(0.002)	(0.054)	(0.001)	(0.041)
Labor	0.2082	0.1452	0.0843	0.0447
	(0.002)	(0.067)	(0.001)	(0.008)
Capital	0.0451	0.0425	0.0631	0.0769
	(0.001)	(0.013)	(0.000)	(0.010)

All variables are expressed in logs. The values in the OLS columns are used as initial condition for the GMM optimization. The columns for GMM estimates show the final results. The number in parentheses are standard errors. The standard errors for the GMM estimates have been bootstrapped.

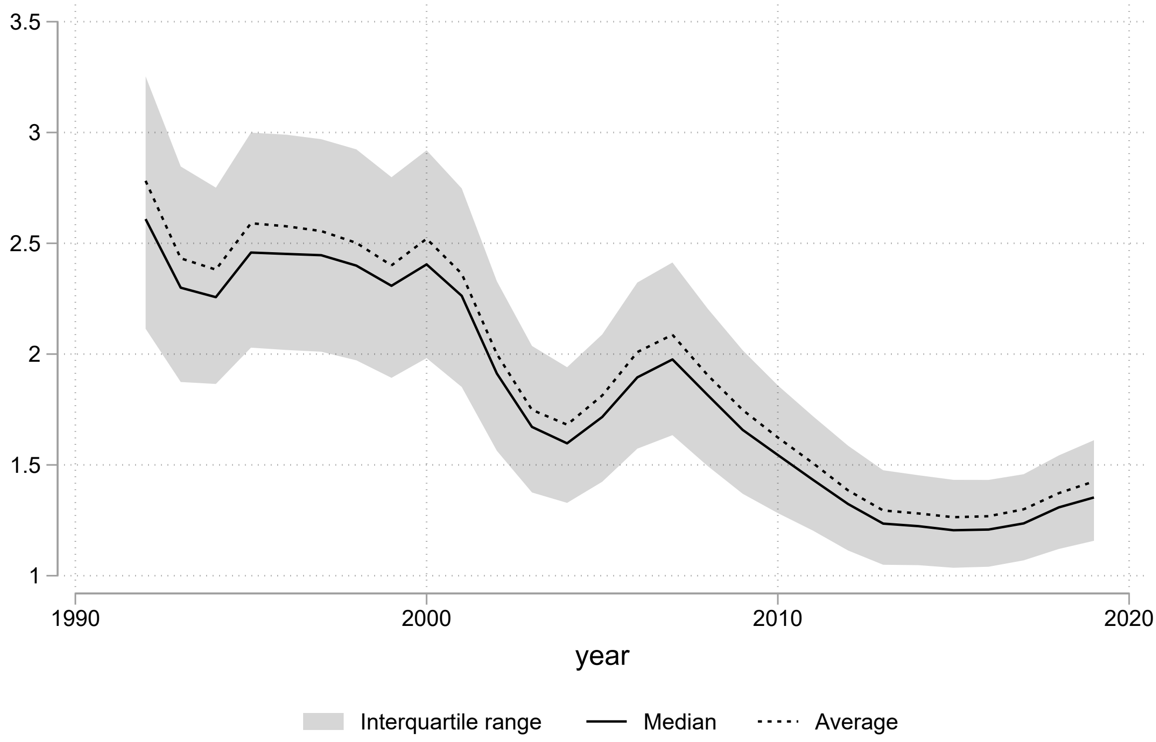
the average markdown in 1992. The increase predominantly occurred after the Great Recession. Importantly, the dispersion of markdowns has increased whenever the average increased, and vice versa.

As Equation (3) shows, there is a mechanical relationship between markdowns and markups. Keeping the expenditure share α_{it}^{D} and the elasticity θ_{it}^{D} constant, if the markup μ_{it} increases, then the markdown ψ_{it} has to decrease. However, the reported time series behavior suggest that there also is an economic interpretation to this relationship. We observe markups on lending rates to decrease with the Great Recession, while markdowns on deposit rates increase. We also observe that there has been an increase in the number of bankrupt banks during the Great Recession. Because policy makers were concerned with financial stability at the time, they had to impose more stringent rules about the risk-taking behavior of banks. This meant that banks faced harsher competition on lending rates and had to make up for the lost profitability in order to survive. Deposits have been a way for banks to sustain their profit streams, by paying deposit rates that are relatively lower than comparable interest rates (e.g., risk-free rates).

4.2 Markups, markdowns and bank characteristics

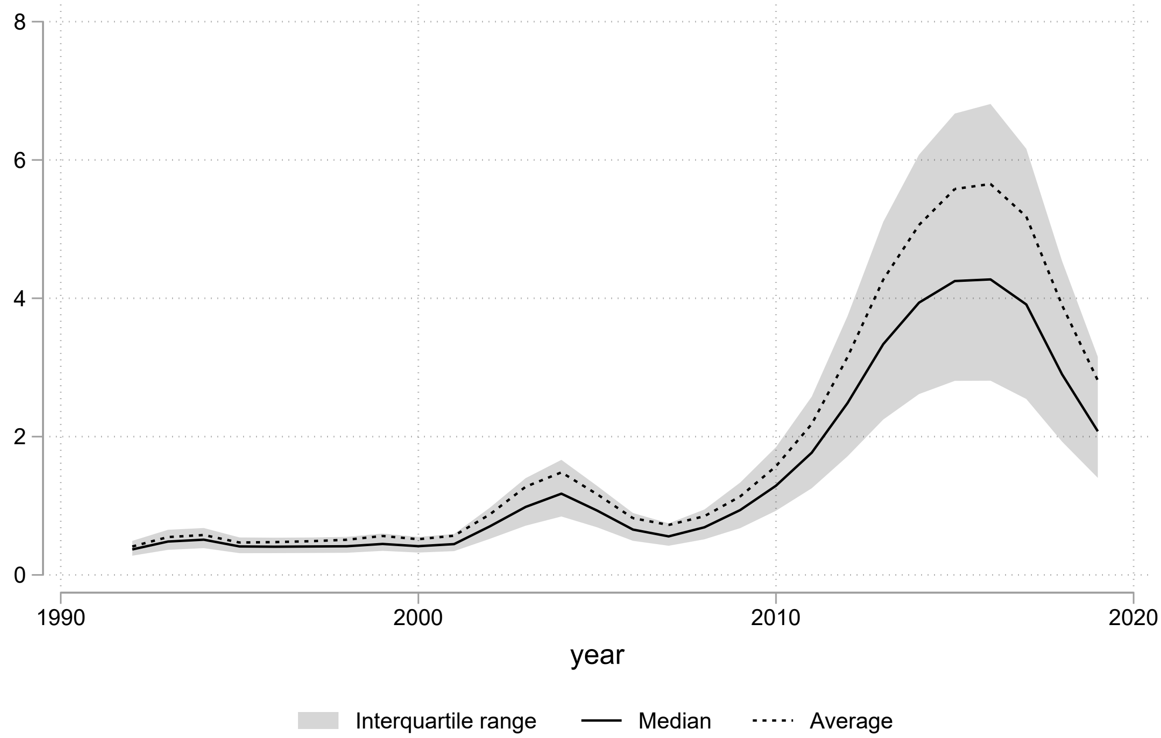
How do markups and markdowns correlate with observable bank characteristics? To answer this question, I regress them separately on a set of balance sheet and income statement variables. This exercise is useful to shed light on which banks are able to charge higher markups or markdowns. However, it does not help understand the determinants of market power in loan

Figure 2: Gross markups across years.



Each yearly cross-section of markups has been trimmed 1% top and 1% bottom to account for outliers.

Figure 3: Gross markdowns across years.



Each yearly cross-section of markdowns has been trimmed 1% top and 1% bottom to account for outliers.

and deposit markets, nor does it quantify causal relationships. The regression reads:

$$y_{it} = \varphi_0 + \varphi_1 \text{big}_{it-1} + \varphi_2 \text{biggest}_{it-1} + \varphi_3 \frac{\text{loans}_{it-1}}{\text{assets}_{it-1}} + \varphi_4 \frac{\text{deposits}_{it-1}}{\text{assets}_{it-1}} + \varphi_5 \frac{\text{liquid assets}_{it-1}}{\text{assets}_{it-1}} + \\ + \varphi_6 \frac{\text{NPL}_{it-1}}{\text{assets}_{it-1}} + \varphi_7 \frac{\text{NIM}_{it-1}}{\text{assets}_{it-1}} + \varphi_8 \frac{\text{cash div}_{it-1}}{\text{income}_{it-1}} + u_{it},$$

where y_{it} is either the log-markup or the log-markdown, big_{it} is an indicator variable that equals one when a bank's total assets are above the yearly median of total assets, biggest_{it} is an indicator that equals one when a bank's total assets are above the yearly 95$^{\text{th}}$ percentile of total assets, NPL_{it} is nonperforming loans and NIM_{it} is the net interest margin. The ratios of loans and deposits over assets give a sense of the composition of the balance sheet of each bank. Liquid assets refer to the sum of cash balances and securities held for short-term trading. The ratio of the net interest margin over assets is a measure of return on assets. Cash dividends over total income measures how keen a bank is to distribute profits with dividends. I always include year fixed effects, in order to control for various aggregate phenomena, such as recessions and regulatory changes. I investigate separately the results adding county, specialization and bank fixed effects. Considering these different fixed effects allows to have a sense of whether there is unobserved heterogeneity that is not captured by the regressors. In particular, specialization fixed effects control for different specialization in a bank's business, such as agricultural credit or household lending. Finally, I cluster standard errors at the bank level.

Table 3 shows the results. First, big banks charge higher markups and lower markdowns. The pattern however does not hold for the banks in the top 5% of the assets distribution, which also charge higher markdowns relative to small banks. Having relatively more loans among their assets or relatively more deposits as sources of financing are also correlated with higher markups but lower markdowns. The relationships between deposits and markdown could be due to the fact that depositors may prefer banks that offer more competitive deposit rates. Banks that have more liquid assets tend to have lower markups and higher markdowns. Interestingly, a higher share of NPLs is associated with less markups and markdowns. Finally, measures of profitability and propensity to pay cash dividends seems to be unrelated to markups and markdowns, which is puzzling. While markups and markdowns relate to economic profits, profitability as measured relates to accounting profits. Fixed costs may explain the difference.

The relationships described so far hold across different combinations of fixed effects. The results with county fixed effects are qualitatively the same as with specialization fixed effects, suggesting that geography and specialization capture roughly the same unobserved heterogeneity across banks. The columns with bank fixed effects are also roughly comparable, with the notable exception that the top 5% banks are associated with lower markdowns.

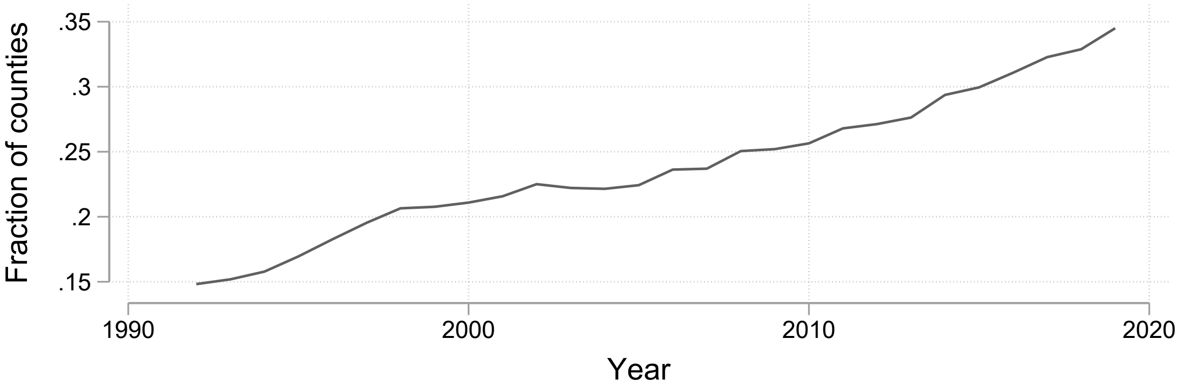
As mentioned above, Table 3 does not seek causal interpretation. It also describes some puzzling facts: for example, profitability does not seem to be related to either markups or mark-

Table 3: Correlations between markups, markdowns and observable bank characteristics.

	Log(Markup)			Log(Markdown)		
	(1)	(2)	(3)	(4)	(5)	(6)
Assets > median	0.08712***	0.08633***	0.12089***	−0.04895***	−0.05897***	−0.19187***
	(0.0042)	(0.0043)	(0.0042)	(0.0072)	(0.0082)	(0.0069)
Assets > 95th percentile	0.10435***	0.09538***	0.12615***	0.07982***	0.09233***	−0.14293***
	(0.0130)	(0.0127)	(0.0150)	(0.0214)	(0.0222)	(0.0254)
Loans / assets	0.28660***	0.33161***	0.36278***	−0.44147***	−0.62012***	−0.34552***
	(0.0197)	(0.0234)	(0.0142)	(0.0688)	(0.0801)	(0.0402)
Deposits / assets	0.08283**	0.10028**	0.35841***	−0.67676***	−0.58173***	−1.11312***
	(0.0288)	(0.0328)	(0.0269)	(0.0579)	(0.0848)	(0.0398)
Liquid assets / assets	−0.59679***	−0.71585***	−0.23153***	1.32427***	1.56952***	0.69079***
	(0.0342)	(0.0356)	(0.0243)	(0.0629)	(0.0672)	(0.0461)
NPL / assets	−1.47256***	−1.36833***	−1.06530***	−1.03225***	−1.35908***	−1.05619***
	(0.1525)	(0.1631)	(0.1102)	(0.2827)	(0.3225)	(0.1759)
NIM / assets	−0.86395	−1.35720	−0.24580	4.53811	6.10975	1.58971
	(0.6011)	(0.8887)	(0.1752)	(3.0545)	(3.9697)	(1.2930)
Cash dividends / income	0.00391	0.00727	0.00264	−0.00026	−0.00005	0.00025*
	(0.0052)	(0.0069)	(0.0033)	(0.0003)	(0.0003)	(0.0001)
Constant	0.42890***	0.41316***	0.08429***	0.55833***	0.51760***	1.09914***
	(0.0270)	(0.0299)	(0.0226)	(0.0603)	(0.0611)	(0.0343)
County FE	Yes	No	No	Yes	No	No
Specialization FE	No	Yes	No	No	Yes	No
Bank FE	No	No	Yes	No	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	0.50780	0.44318	0.78943	0.77890	0.74328	0.90451
Observations	208660	208668	207932	207953	207961	207221

Standard errors (in parentheses) are clustered at the bank level. Markups and markdowns have been trimmed one percent both at the bottom and at the top of each yearly cross-section. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 4: Fraction of counties that host only one bank.



downs. Rationalizing these results is left for future research.

4.3 Relationship with other measures of competition

How do markups and markdowns compare with existing measures of competition? To address this question, I compute the Herfindahl-Hirschman Index (HHI) and the Boone indicator. The HHI is a measure of concentration based on market shares. It requires bank-level observations and yields an aggregate number within a set of banks. Analytically, it is computed as

$$\mathrm{HHI}_{gt} = \sum_{i \in g} s_{it}^2, \tag{10}$$

for every period t, where g is a set of banks and s_{it} is the market share bank i has in period t within group g. In this formulation, the HHI ranges between $1/N_g$ and 1, where N_g is the number of banks in group g. When the HHI equals $1/N_g$, banks in group g have uniform market shares. This is usually associated with a highly competitive environment. Conversely, when the HHI equals one, there exists only one bank in group g that serves the entire market. This is usually thought of a highly monopolistic environment. As I have computed markups on lending rates and markdowns on deposit rates, I compute the shares s_{it} relative to both total net loans and total domestic deposits. I consider both US counties and states as delimiters that determine the sets g. However, there is a considerable number of counties in the US where only one bank operates. Figure 4 provides graphical evidence of this phenomenon. For all those counties, the HHI assumes it maximal value, one. For this reason, I present state-level evidence.

Figures 5 and 6 show the time series behavior of the average, median and interquartile range of the HHI computed on, respectively, loans and loans at the state level. These can be compared with Figures 2 and 3 respectively. The HHIs both have a slight upward trend over time. The cross-sectional variation in both indices is also slightly increasing over the years. The HHIs do not feature as much time series variation as the markups or markdowns. Markups tend to have a downward trend, while the HHI on loans trends upward. The HHIs being increasing over the sample period may be due to continued Mergers and Acquisitions (M&A) activity. In particular, M&A in the US banking industry are due to two main reasons. One is as part of recovery plans, often under the supervision or direction of the FDIC. The other is as part of deliberate deals for strategic reasons. The former reason increases the HHIs for mechanisms that are not related to market power but, rather, are due to financial stability concerns and to the intervention of the policy maker.

The graphical inspection is confirmed with Tables 4 and 5. The former shows the results from regressing the state median markup on the HHI on loans. The latter shows the results from regressing the state median markdown on the HHI on deposits. Particularly, regardless of whether I control for state or year fixed effects, markups correlate negatively with the HHI on loans. Conversely, markdowns correlate positively with the HHI on deposits. These results seem to suggest that the HHI on loans does not capture the same phenomenon as the markup on lending rates. The tables remain qualitatively unchanged if I compute the arithmetic average of markups or markdowns rather than the median.

The Boone (2008) indicator relies on measures of profitability and costs. Empirically, the Boone indicator is defined as the elasticity of profits to marginal costs, where both profits and marginal costs are relative to the respective maxima in the market. Such elasticity should be negative: higher marginal costs lead to less profits. However, the more negative the elasticity is, the harsher competition firms face. The reason behind this interpretation is that more efficient firms gain higher profits. An increase in competition makes less efficient firms exit the market and reallocates production to more efficient firms. This makes the elasticity of relative profits to relative marginal costs higher in absolute value.

Following Boone et al. (2005), I compute the Boone indicator using the following regression for every year t:

$$\log\left(\frac{\pi_{it}}{\pi_t}\right) = \kappa_{0,t} - \kappa_{1,t}\log\left(\frac{\mathrm{mc}_{it}}{\mathrm{mc}_t}\right) + \varsigma_{it}, \tag{11}$$

where π_{it} is the profit of bank i in period t and mc_{it} is the marginal cost. The terms π_t and mc_t are reference points within the cross-section. Ideally, they would correspond to the maximum. However, due to the presence of outliers and similarly to Boone et al. (2005), I use the 98$^{\text{th}}$ percentile. I use average variable costs as a proxy of marginal costs. The Boone indicator is the coefficient $\kappa_{1,t}$.

Figure 7 shows the OLS estimates of $\kappa_{1,t}$ in Equation (11). The regression is run once for every year-state pair using quarterly balance sheet data. Isolating each state gives a sense of cross-sectional variability of the Boone indicator in the US. The time series behavior of the

Figure 5: State-level HHI index on loans.

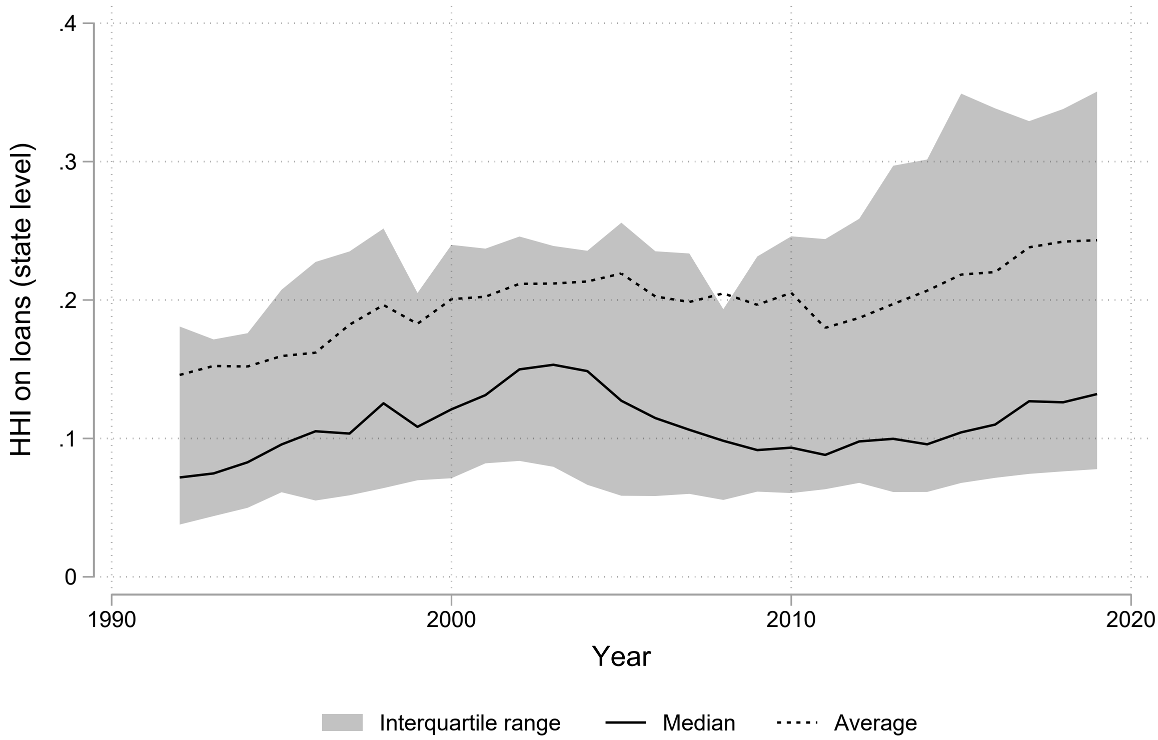


Figure 6: State-level HHI index on deposits.

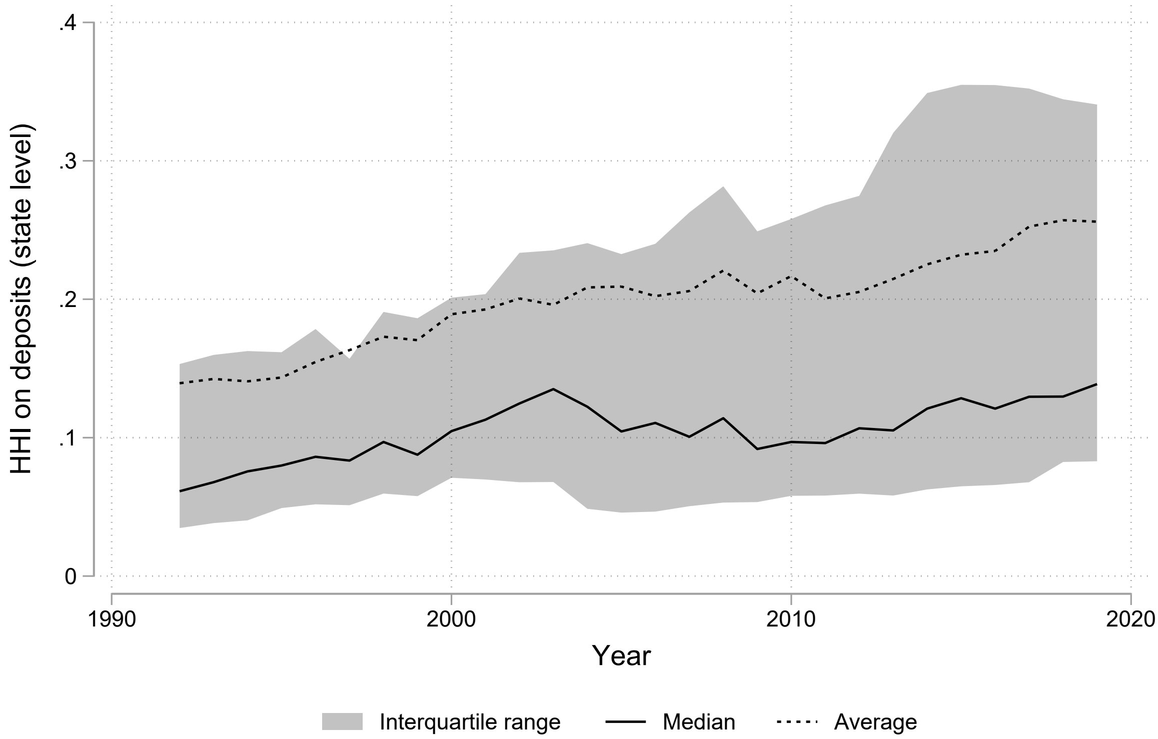


Table 4: Results from regressing state-median markups on the HHI on loans. All variables are in logs.

	Log(Markups) — state-level medians			
	(1)	(2)	(3)	(4)
Log(HHI loans)	-0.0571^{***}	-0.1122^{***}	-0.0366^{***}	-0.0093
	(0.007)	(0.014)	(0.004)	(0.005)
Constant	0.4240^{***}	0.3072^{***}	0.4673^{***}	0.5252^{***}
	(0.017)	(0.030)	(0.008)	(0.010)
Observations	1559	1559	1559	1559
State FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Adjusted R^2	0.0370	0.1549	0.7784	0.9125

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Results from regressing state-median markdowns on the HHI on deposits. All variables are in logs.

	Log(Markdowns) — state-level medians			
	(1)	(2)	(3)	(4)
Log(HHI deposits)	0.2376^{***}	0.5349^{***}	0.1252^{***}	0.0290^{**}
	(0.021)	(0.040)	(0.007)	(0.009)
Constant	0.6465^{***}	1.2934^{***}	0.4019^{***}	0.1923^{***}
	(0.051)	(0.090)	(0.018)	(0.020)
Observations	1559	1559	1559	1559
State FE	No	Yes	No	Yes
Year FE	No	No	Yes	Yes
Adjusted R^2	0.0741	0.1612	0.8898	0.9658

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 7: State-level Boone indicator.

Boone indicator does not resemble the one of either HHI index shows above. In particular, the Boone indicator decreased with the Great Recession, but is otherwise relatively stable. The decrease coincides with a significant number of bankruptcies, as can be seen by comparing with Figure 1. Across states, the indicator features significant time-varying dispersion. Comparing the Boone indicator with the markups and markdowns in Figures 2 and 3 respectively, we see that these measures do not exhibit common time-series variation.

Table 6 shows the correlations between the Boone indicator and the measures of markups and markdowns. The main observation is that the Boone indicator does not significantly correlate with neither markups nor markdowns. The only exception occurs when I control for year fixed effects, where the Boone indicator correlates negatively with markups and positively with markdowns, although the size of the correlation is small.

Why are the HHIs and the Boone indicator different from markups and markdowns? The explanation lies in the assumptions behind the different measures. Neither the HHI nor the Boone indicator are vocal about the sources of market power. On one hand, the HHIs measure market share concentration within a set of banks, where the choice of the set is often guided by geography. There is no clear relationship between geographical concentration and the competitiveness of a bank, other than the suggestion that banks that are relatively near to each other may see customers easily switching between them. In particular, the HHI is subject to confounding

Table 6: Correlation between the Boone indicator and the measures of markup and markdown. All variables are the medians within each year-state pair.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log(Markup) — state-level medians				Log(Markdown) — state-level medians			
Boone ind.	0.0005	-0.0036^{**}	0.0050^{*}	0.0006	-0.0030	0.0083^{**}	-0.0095	0.0028
	(0.003)	(0.001)	(0.003)	(0.001)	(0.008)	(0.003)	(0.008)	(0.002)
Constant	0.5304^{***}	0.5284^{***}	0.5326^{***}	0.5304^{***}	0.1678^{***}	0.1733^{***}	0.1646^{***}	0.1706^{***}
	(0.008)	(0.004)	(0.007)	(0.002)	(0.024)	(0.009)	(0.023)	(0.004)
Observations	1503	1503	1503	1503	1503	1503	1503	1503
State FE	No	No	Yes	Yes	No	No	Yes	Yes
Year FE	No	Yes	No	Yes	No	Yes	No	Yes
Adjusted R^2	-0.0006	0.7560	0.1273	0.9114	-0.0006	0.8656	0.0653	0.9650

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

power on output and input markets. By using geography to delimit the sets within which the HHIs are computed, it is not clear that what drives the HHI on loans cannot be driving the HHI on deposits, and viceversa. On the other hand, the Boone indicator implicitly assumes that firms exert their power on the market for outputs. The Boone indicator implicitly assumes that firms take their cost functions as given, whereby input prices may be strategic variables for a firm that applies a markdown. Instead, markups and markdowns take an explicit stand on which channel market power is exerted through. The working hypothesis is that firms are price-takers on neither output nor input markets. In particular, markups and markdowns measure how much observed prices are different from their price-taking counterfactuals.

5 Measures of financial stability

I consider two alternative accounting-based measures of financial stability. They have been proposed by Altman (1968) and Ohlson (1980) respectively. The former is known as Z-score and the latter as O-score. They are the predicted values of reduced-form logit models of a future bankruptcy indicator on a set of balance sheet ratios. These measures rely only on the availability of balance sheet data. I can therefore compute them given the sample at hand. Hillegeist et al. (2004) show that these two measures of default probability are good out-of-sample predictors of a firm's bankruptcy. However, they also show that Merton (1974)'s *distance-to-default* outperforms the Z-score and the O-score in terms of out-of-sample predictive power. Because the computation of the distance-to-default requires stock market data, it would significantly reduce the number of observations and the scope of the results. Therefore

The Z-score and the O-score consist of the predicted values of two logit models. The dependent variable is an indicator, which equals one if a firm goes bankrupt in the two years ahead and zero otherwise. The independent variables are a set of balance sheet ratios. The two scores differ in the set of regressors. While Altman (1968) and Ohlson (1980) use logit models, I use linear probability models. This is driven by the scarcity of yearly bankruptcies in the data I use relative to the number of banks in each cross-section. Because of this, the maximum likelihood estimator for the logit parameters converges for neither the Z-score nor the O-score. Omitting the bank-year subscripts for notational convenience, the regression model I fit in order to compute the Z-score is

$$B = \delta_0 + \delta_1 \frac{WC}{TA} + \delta_2 \frac{RE}{TA} + \delta_3 \frac{EBT}{TA} + \delta_4 \frac{V_E}{TL} + \delta_5 \frac{S}{TA} + u, \tag{12}$$

where B is the bankruptcy indicator, TA is total assets, TL is total liabilities, WC is working capital, RE is retained earnings, EBT is earnings before taxes, V_E is market value of equity and S is sales. I approximate the market value of equity with Tier-1 capital. The regression model I use to obtain the O-score instead is

$$\begin{aligned} B = \zeta_0 + \zeta_1 \frac{TL}{TA} + \zeta_2 \frac{WC}{TA} + \zeta_3 \frac{CL}{CA} + \zeta_4 \frac{NI}{TA} + \zeta_5 \frac{FFO}{TL} + \\ + \zeta_6 INTWO + \zeta_7 OENEG + \zeta_8 CHIN + \nu, \end{aligned} \tag{13}$$

where CL is current liabilities, CA is current assets, NI is net income, FFO is pre-tax income plus depreciation and amortization, $INTWO$ is an indicator on whether cumulative net income over the previous two years is negative, $OENEG$ is an indicator on whether owners' equity is negative and $CHIN \equiv (NI_t - NI_{t-1})/(|NI_t| + |NI_{t-1}|)$ is the scaled change in net income. The Z-score and the O-score are the in-sample predicted values from the regression models in Equation (12) and (13) respectively.

Figures 8 and 9 show the time series behavior of the average, median and interquartile ranges of the Z-scores and the O-scores. Z-scores have historically trended upwards, while O-scores did not. The cross-sectional dispersion of Z-scores has remained substantially constant, while that of O-scores has spiked with Great Recession. It is also useful to compare these figures to the time series of bankruptcies in Figure 1. We observe that the O-scores pick up the spike of bankruptcies around the Great Recession better than the Z-scores.

6 Financial stability and market power

Similarly to Anginer et al. (2014), I regress measures financial stability on market power. The key difference here is the measure of market power. Rather than focusing on price-cost margins that relate to market power on loan markets, I disentangle a bank's market power as coming from two different markets: loan and deposit markets. In particular, I use the markup on lending rates as a measure of market power on loan markets. Conversely, I use the markdown on deposit rates as a measure of market power on deposit markets. The baseline regression model is

$$\text{Prob}\,(\text{bankruptcy}_{it}) = \gamma_0 + \gamma_i + \gamma_\mu \log(\mu_{it}) + \gamma_\psi \log(\psi_{it}) + \eta_{it}, \tag{14}$$

where μ_{it} is the markup on lending rates and ψ_{it} is the markdown on deposit rates. The left-hand side variable is either the Z-score or the O-score.

This work focuses on the *correlation* between market power and financial stability, rather than on causation. Both the right-hand side and the left-hand side variables in Equation (14) are computed from balance sheet data. One potential concern is that there will be spurious correlation. Because all data comes from balance sheet and income statement data, the regression may pick up mechanical correlation due to within-bank variation. However, all the variation in the measures of market power comes from income statement variables. Conversely, the left-hand

Table 7: Estimates of the coefficients for the model in Equation (14). The dependent variable is the log of the Z-score. All specifications include bank fixed effects.

	(1) Log(Z-score)	(2) Log(Z-score)	(3) Log(Z-score)	(4) Log(Z-score)	(5) Log(Z-score)	(6) Log(Z-score)
Log(markup)	-0.4091^{***}	-0.3505^{***}			-0.4191^{***}	-0.4744^{***}
	(0.004)	(0.006)			(0.008)	(0.010)
Log(markdown)			0.1186^{***}	0.1041^{***}	-0.0030	-0.1068^{***}
			(0.001)	(0.004)	(0.003)	(0.006)
Constant	-5.3837^{***}	-5.4136^{***}	-5.6179^{***}	-5.6149^{***}	-5.3799^{***}	-5.3308^{***}
	(0.002)	(0.003)	(0.000)	(0.001)	(0.005)	(0.006)
Observations	153604	153604	153070	153070	152249	152249
Adjusted R^2	0.6248	0.6560	0.5848	0.6187	0.6269	0.6658

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

side variables have been primarily obtained from balance sheet (not income statement) ratios. While income statement and balance sheet variables in levels are obviously correlated, adding bank fixed effects will capture their common variation within banks and leave variation across banks.

The main results are shown in Tables 7 and 8. The former uses the Z-score as dependent variable, while the latter uses the O-score. In all cases, markups on lending rates always correlate negatively with the scores. Markdowns on deposit rates correlate positively with both scores when considered alone, and negatively when holding markups constant. The inclusion of time fixed effects does not considerably change the point estimates. All correlations are highly significant. The magnitudes of the correlations differ, depending on whether one considers the Z-score or the O-score. A one percent increase in markups is associated with a decrease in the Z-score of 0.34 to 0.47 percent and a decrease in the O-score of 0.56 to 1.92 percent, depending on whether markdowns are held constant. Conversely, a one percent increase in markdowns is associated with an increase in the Z-score of roughly 0.11 percent and an increase in the O-score, or a decrease of 1–5 percentage points if markups are held constant.

7 Conclusion

In this paper I use the production approach to the estimation of markups and markdowns to banking data. I compute markups on lending rates and markdowns on deposit rates. I find that markups have generally been trending downwards over the years, while markdowns have been increasing, especially after the Great Recession. I correlate these new measures to observable bank characteristics. I find that bigger banks tend to charge a higher markup on lending rates

Figure 8: Z-scores.

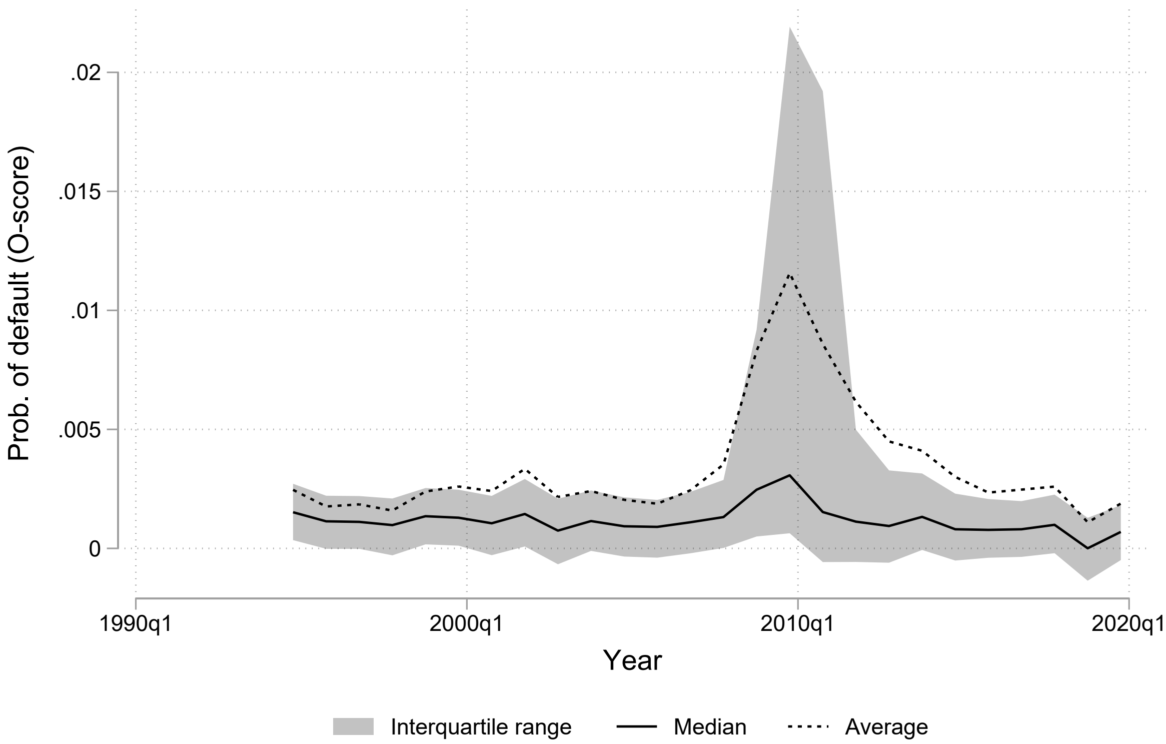
Table 8: Estimates of the coefficients for the model in Equation (14). The dependent variable is the log of the O-score. All specifications include bank fixed effects.

	(1) Log(O-score)	(2) Log(O-score)	(3) Log(O-score)	(4) Log(O-score)	(5) Log(O-score)	(6) Log(O-score)
Log(markup)	-0.5434^{***}	-1.0038^{***}			-1.8916^{***}	-1.8287^{***}
	(0.019)	(0.027)			(0.037)	(0.040)
Log(markdown)			0.0219^{***}	0.0994^{***}	-0.5445^{***}	-0.6840^{***}
			(0.006)	(0.016)	(0.013)	(0.024)
Constant	-5.9139^{***}	-5.6484^{***}	-6.2276^{***}	-6.2244^{***}	-5.1644^{***}	-5.2067^{***}
	(0.011)	(0.016)	(0.000)	(0.001)	(0.021)	(0.022)
Observations	141432	141432	141055	141055	140538	140538
Adjusted R^2	0.3127	0.3712	0.3010	0.3535	0.3372	0.3818

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 9: The average, median and interquartile range of the *O*-scores.



Some probabilities are negative because they are the predicted values of a linear probability model.

and a lower markdown on deposit rates. Both markups and markdowns correlate positively with bank profitability. These two findings together suggest that smaller, more local banks have driven their profitability through deposit rates, particularly after the Great Recession. I compare the measures with the Herfindahl-Hirschman Index (HHI), which is a widely used measure of market concentration. I find that the HHI on deposits correlates with markdowns on deposit rates, although imperfectly. On the other hand, markups on lending rates correlate negatively with the HHI on loans. Finally, I relate the measures of markups and markdowns with measures of default probability. Particularly, I estimate the Z-score and the O-score, which can be interpreted as bankruptcy probabilities that can be predicted, for each bank, by balance sheet data. I find that higher markups are associated with a lower probability of bankruptcy for banks. Conversely, markdowns are positively correlated with bankruptcy probability, but only if markups are not controlled for.

This paper contributes to three main branches in the banking literature. First, to the best of my knowledge, this is the first paper that estimates markups and markdowns with banking data. While the methodology is not new, I adapt it to better suit the intermediation approach to production in banking. Under the assumption that markups and markdowns are measures of market power in loan and deposit markets respectively, this paper provides a new measure of competition for the banking industry. The main novelty here is that I disentangle market power on output markets (loans) from market power on input markets (deposits). The coexistence of market power on both inputs is not new to existing literature in Macroeconomics and Industrial Organization. However, this paper brings this concept to the banking industry. Second, I provide evidence for the banking sector that markups and markdowns do not correlate with previously known measures of competition. I compare markups and markdowns to two HHIs, one on loans and the other on deposits, and to the Boone indicator. My results suggest that HHIs confound the channels of market power: the drivers of the HHI on loans may very well be the same drivers of the HHI on deposits. The Boone indicator, instead, captures Third, I revisit the correlation between market power and default probabilities. While most existing results find a positive relationship between market power and the probability of default for banks, I find that higher market power in lending markets is associated with lower bankruptcy probability, although the magnitude of the correlation is small. Instead, there is a positive relationship between markdowns and default probability, which vanishes once I control for markups. This finding suggests that previous studies confounded the channels of market power.

The main direction for future research consists in understanding the determinants of market power in the banking industry. While I disentangle markups and markdowns, I do not investigate what is determining them.

References

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification Properties of Recent Production Function Estimators. *Econometrica*, 83(6):2411–2451.

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 23(4):589–609.

Anginer, D., Demirguc-Kunt, A., and Zhu, M. (2014). How Does Competition Affect Bank Systemic Risk? *Journal of Financial Intermediation*, 23(1):1 – 26.

Beck, T., Demirgüç-Kunt, A., and Levine, R. (2006). Bank Concentration, Competition, and Crises: First Results. *Journal of Banking & Finance*, 30(5):1581 – 1603.

Berger, A. N., Klapper, L. F., and Turk-Ariss, R. (2009). Bank Competition and Financial Stability. *Journal of Financial Services Research*, 35:99–118.

Boone, J. (2008). A New Way to Measure Competition. *The Economic Journal*, 118(531):1245–1261.

Boone, J., Griffith, R., and Harrison, R. (2005). Measuring Competition. Advanced Institute of Management Research Paper No. 022.

Boyd, J. H. and De Nicoló, G. (2005). The Theory of Bank Risk Taking and Competition Revisited. *The Journal of Finance*, 60(3):1329–1343.

De Loecker, J., Eeckhout, J., and Unger, G. (2020). The Rise of Market Power and the Macroeconomic Implications. *The Quarterly Journal of Economics*, 135(2):561–644.

De Loecker, J. and Warzynski, F. (2012). Markups and Firm-Level Export Status. *American Economic Review*, 102(6):2437–71.

Drechsler, I., Savov, A., and Schnabl, P. (2017). The Deposits Channel of Monetary Policy. *The Quarterly Journal of Economics*, 132(4):1819–1876.

Hall, R. E. (1988). The Relation between Price and Marginal Cost in U.S. Industry. *Journal of Political Economy*, 96(5):921–947.

Hellmann, T. F., Murdock, K. C., and Stiglitz, J. E. (2000). Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough? *American Economic Review*, 90(1):147–165.

Hillegeist, S. A., Keating, E. K., Cram, D. P., and Lundstedt, K. G. (2004). Assessing the Probability of Bankruptcy. *Review of Accounting Studies*, 9:5–34.

Levinsohn, J. and Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *The Review of Economic Studies*, 70(2):317–341.

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *The Journal of Finance*, 29(2):449–470.

Morlacco, M. (2020). Market Power in Input Markets: Theory and Evidence from French Manufacturing. Working paper.

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1):109–131.

Olley, G. S. and Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64(6):1263–1297.

Panzar, J. C. and Rosse, J. N. (1987). Testing For "Monopoly" Equilibrium. *The Journal of Industrial Economics*, 35(4):443–456.

Schaeck, K., Cihak, M., and Wolfe, S. (2009). Are Competitive Banking Systems More Stable? *Journal of Money, Credit and Banking*, 41(4):711–734.

Scharfstein, D. and Sunderam, A. (2016). Market Power in Mortgage Lending and the Transmission of Monetary Policy. Manuscript, Harvard Business School.

Sealey, C. W. J. and Lindley, J. T. (1977). Inputs, Outputs, and a Theory of Production and Cost at Depository Financial Institutions. *The Journal of Finance*, 32(4):1251–1266.

Wang, Y., Whited, T. M., Wu, Y., and Xiao, K. (2021). Bank Market Power and Monetary Policy Transmission: Evidence from a Structural Estimation. R&R The Journal of Finance.

A Examples of production functions in banking

I illustrate the functional forms of the production function given some simple models. I start with a simple model of a bank in a frictionless world. This baseline is a simplified version of the model in Sealey and Lindley (1977). Then I expand the model by sequentially introducing two elements: a leverage constraint and risk in loan repayments. For simplicity, I assume that the former is imposed by regulation. In introducing these two elements, I abstract away from their microfoundations.

A.1 The baseline framework

A bank collects deposits D and requires labor N and capital K in order to provide loans L. The balance sheet constraint is

$$L \leq D, \tag{15}$$

where the inequality would be strict if the bank was choosing to hold part of deposits in cash reserves. To process lending requests and deposits, labor and capital are required. Therefore,

$$L \leq L(N, K) \tag{16}$$

$$D \leq D(N, K). \tag{17}$$

Combining the inequalities (15), (16) and (17), we obtain the following production function

$$L = \min\left\{L(N, K), D(N, K)\right\}, \tag{18}$$

where the equality follows from considering the efficient frontier of the production possibility set, given the usual monotonicity assumptions from classical production theory. In words, a bank requires both labor and deposits in order to "produce" loans. This result corresponds to the contribution of Sealey and Lindley (1977). In particular, there is some complementarity between classical inputs such as labor and capital and the need of sources of financing. This complementarity is what sets banks apart from manufacturing firms, from a production theory point of view.

A.2 Adding a leverage constraint

Let us introduce equity and suppose that the bank is subject to a leverage constraint. This would read

$$L \leq \lambda E, \tag{19}$$

where λ is the leverage ratio. Combining (15), (16), (17) and (19), the production function would become

$$L = \min\left\{L(N), D(N), \lambda E\right\}. \tag{20}$$

This functional form has an important implication. Consider a bank that maximizes profits subject to (20). The efficiency condition for such a problem implies that the leverage constraint hold with equality, regardless of whether equity is more expensive than deposits.

A.3 Adding risky loans

Now let us assume that loans are risky. Let ϵ be the fraction of loans that will not be repaid. Not that ϵ does not necessarily belong to the information set of the bank. Let total loans $L = L^n + L^d$, where L^n are net loans and L^d are loans that will not be repaid. By assumption, $L^d = \epsilon L$. Therefore, the new production function is

$$\begin{aligned} L^n &= (1-\epsilon)L \\ \frac{1}{1-\epsilon}L^n &= \min\left\{L(N), D(N), \lambda E\right\}. \end{aligned} \tag{21}$$

Note that this formulation captures the fact that the bank still needs all the necessary resources to "produce" all loans, even those that will not be repaid.

B Time-stability of production function estimates

In this Section I check whether the GMM estimates of the production function parameters are robust to the choice of the sample period. The results in Table 2 use the entire sample, from 1993 to 2019. Here I estimate again the production function using a rolling-window approach, both with annual and quarterly data. Figures 10 and 11 show the results, together with the point estimates from Table 2 and their associated confidence intervals. Overall, the figures show that the production function estimates are relatively stable over time, mostly within the confidence intervals. The results obtained with yearly data show a significant change in the estimated parameters around the Great Recession. This is likely due to the intense M&A activity during the Great Recession, where the balance sheets of certain banks suddenly got bigger. The significant changes around the Great Recession do not appear in the results that use quarterly data.

Figure 10: GMM estimates of the production function parameters using a rolling-window approach on yearly data. The horizontal black line denoted as "Point estimate" is the GMM estimate from Table 2.

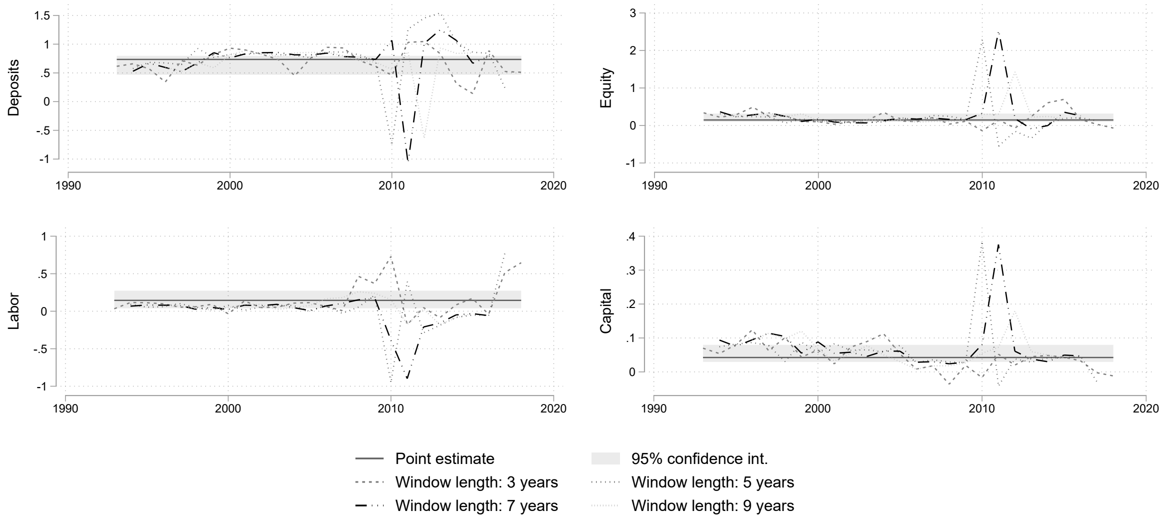
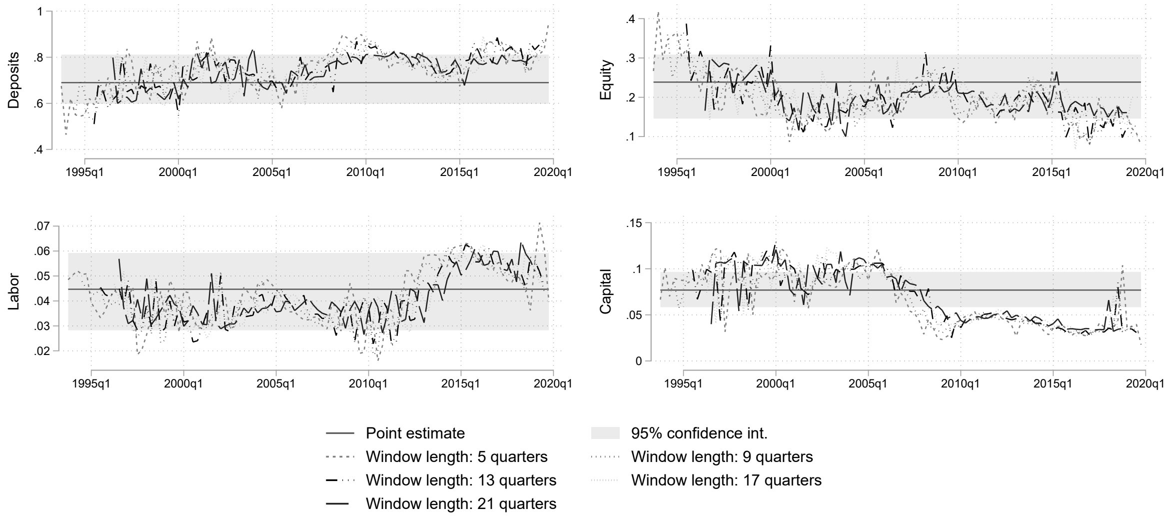


Figure 11: GMM estimates of the production function parameters using a rolling-window approach on quarterly data. The horizontal black line denoted as "Point estimate" is the GMM estimate from Table 2.



C Geographic maps of markups and markdowns

Where are the banks with higher markups or higher markdowns? I compute within-state averages of markups weighted by interest income on loans. Similarly, I compute within-state averages of markdowns weighted by interest expense on deposits. I do so by focusing on 1993 and 2019, separately. Figures 12 and 13 show the average state-level markups in 1993 and 2019 respectively. Figures 14 and 15 show the average state-level markdowns in 1993 and 2019 respectively.

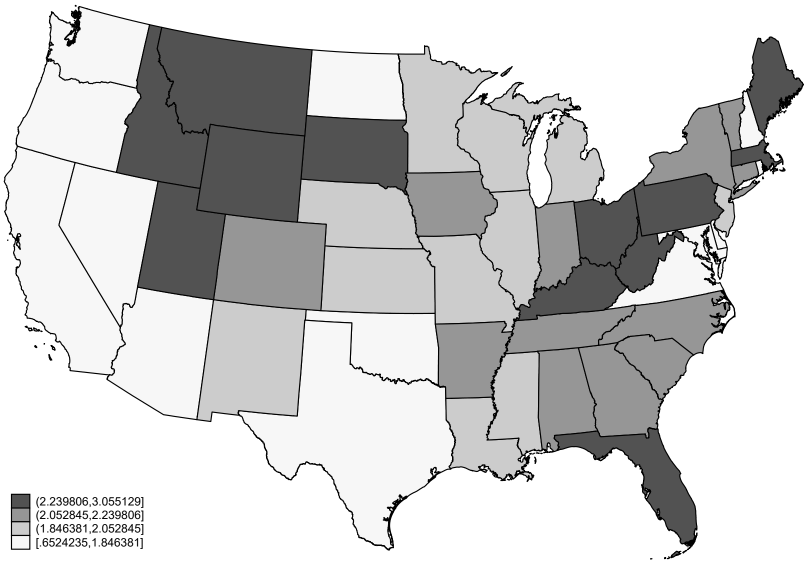


Figure 12: Average state-level markups in 1993.

Bank-level markups have been aggregated to state-level by computing the within-year, within-state average of markups, weighted by interest income from loans.

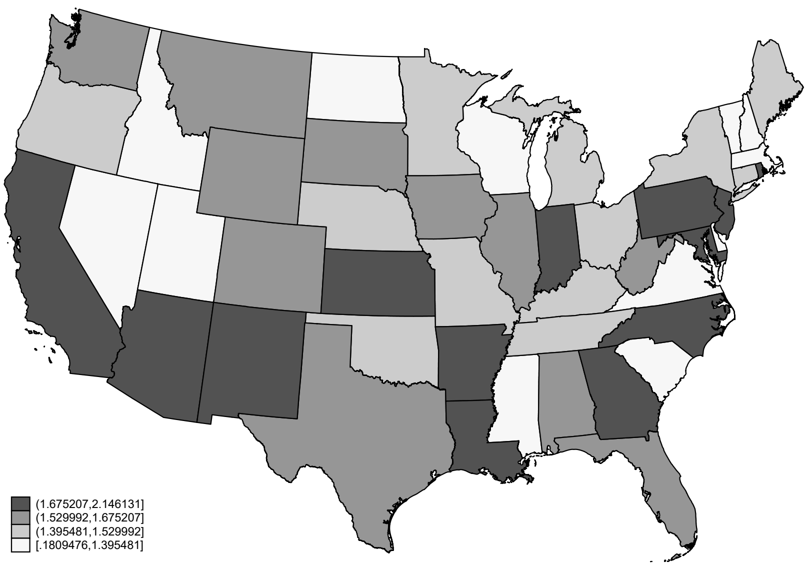


Figure 13: Average state-level markups in 2019.

Bank-level markups have been aggregated to state-level by computing the within-year, within-state average of markups, weighted by interest income from loans.

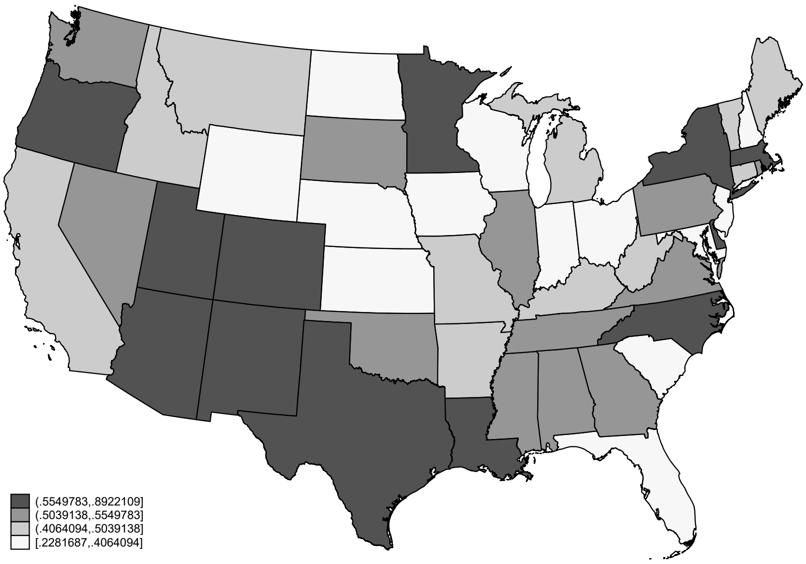


Figure 14: Average state-level markdowns in 1993.

Bank-level markdowns have been aggregated to the state-level by computing the within-year, within-state average of markdowns, weighted by interest expense on deposits.

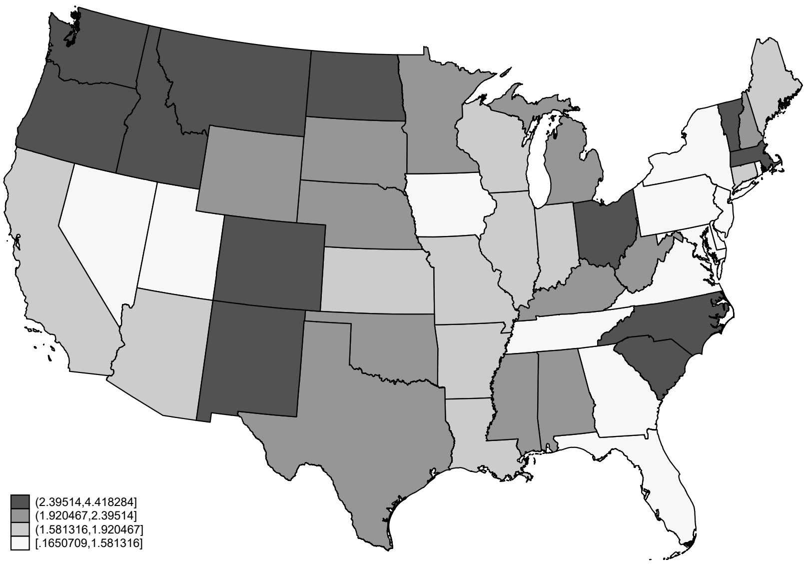


Figure 15: Average state-level markdowns in 2019.

Bank-level markdowns have been aggregated to the state-level by computing the within-year, within-state average of markdowns, weighted by interest expense on deposits.