Measuring Bank Complexity Using XAI

Shengyu Huang* Majeed Simaan* Yi Tang†

shuang52@stevens.edu msimaan@stevens.edu ytang@fordham.edu

January 14, 2025

Abstract

Following the global financial crisis, banks' complexity and opacity have been scrutinized due to their impact on financial stability. Yet, measuring these attributes has posed challenges. We introduce a novel approach using Explainable AI (XAI) to quantify complexity and opacity, revealing a strong correlation at the firm and industry levels. We show that bank complexity exhibits a counter-cyclical pattern, rising before crises and declining during distress, with evidence of reduced trading activity in highly complex banks. Complexity is also associated with higher future returns and reduced volatility but increased systemic risk, providing important insights into bank structure and market stability.

*School of Business, Stevens Institute of Technology, 1 Castle Point Terrace, Babbio Center, Hoboken, NJ 07030, USA.

†Gabelli School of Business, Fordham University, 140 W 62nd St, New York, NY 10023, USA.

1 Introduction

Since the global financial crisis, regulators and the overall public have paid more attention to the banking industry, requiring increasing transparency to address the long-known problem of opacity (Bouvard et al., 2015). Regulators had initiated reforms designed to enhance the financial system's stability and resilience while promoting greater transparency (Baily et al., 2017). However, bank opacity is a subject rooted in both the historical conduct of financial institutions and the complexities of modern financial systems. At its core, banking is a business of confidentiality, handling sensitive customer information, which often extends into a bank's corporate dealings. Additionally, the competitive nature of the financial sector further intensifies such opaqueness. Banks tend to guard information on their strategies, risk models, and operations as valuable intellectual assets, fearing that transparency could compromise advantages to competitors, which is also shown theoretically in Moreno and Takalo (2016).[1]

Banks may be opaque in hiding information from investors, like potential mark-to-market loss from held-to-maturity (HTM) securities. For instance, the failure of Silicon Valley Bank (SVB) in early 2023 echoed the concerns that come with bank opacity again. As liquidity regulations were relaxed by 'Economic Growth, Regulatory Relief, and Consumer Protection Act' issued in 2018, SVB was no longer subject to the liquidity coverage ratio (Federal Reserve, 2023).[2] In particular, SVB held more HTM (also known as 'hide-till-maturity') securities that shielded the market values and hid unrealized losses from investors. The practice had made it difficult for investors and depositors to ascertain the true financial health of the institution (Granja, 2023).[3]

[1]Additionally, the risk of bank runs contributes to this lack of transparency, as the banking system is susceptible to customer panic, potentially leading to a bank run if the true extent of its risk exposure becomes publicly known (Diamond and Dybvig, 1983; Dang et al., 2017).

[2]Note that such regulation was initially established under the Dodd-Frank Act that mandated banks with assets ranging from \$100 to \$250 billion to hold sufficient high-quality liquid assets to cover expected net outflows during stress periods.

[3]In SVB's case, the bank had an HTM securities portfolio valued at \$91.3 billion at the end of 2022, yet it did not disclose the fair value of these securities on its balance sheet. The undisclosed market loss

The above begs at least two questions that our paper tries to address. First, how to measure the opacity and complexity that investors face in understanding the business nature of the banking industry? Second, how do active equity investors price both, given these arising challenges? Numerous academic studies empirically demonstrate the difficulty in understanding the true nature of bank activities by investors and professional analysts due to the inherent opaqueness of banks. As outlined by Morgan (2002), rating agencies encounter a greater disparity when rating banks than non-financial firms. Furthermore, bank holding companies (BHCs) tend to have lower market liquidity in terms of higher bid-ask spread, Amihud's (2002) illiquidity, and lower trading volumes compared to non-financial firms (Flannery et al., 2004; Blau et al., 2017). By examining the BHC stocks movements, other studies demonstrate that they usually have higher Hou and Moskowitz's (2005) price delay and lower stock price synchronicity, two measures of stock price efficiency (Blau et al., 2017; Abedifar et al., 2021). However, the reduced market liquidity and stock price efficiency in BHCs may not solely be attributed to the lack of transparency. The complexity inherent in banking also potentially serves as a driving factor.

Complexity is not easily defined (Cetorelli et al., 2014). It was solely attributed to the size of an institution and later generalized to include a wider range of factors (Cetorelli et al., 2014; Cetorelli and Goldberg, 2016). The current literature, therefore, considers three distinct complexity metrics based on organization, business, and geographic locations. At the same time, complexity also arises from a more interconnected economy and financial system, where the distress of one bank could have amplifying effects on other entities in the network. Additionally, the ambiguity that investors face in distilling such shocks amplifies these frictions (Caballero and Simsek, 2013).

Given the above challenges, recent literature has considered new technological innovations to come up with new measures of firm complexity. Employing a data-driven methodology via

on this HTM portfolio was over $15 billion, which amounted to over 90% of SVB's total equity. As these substantial unrealized losses were obscured, investors did not comprehend the true financial health condition of the bank until the sudden large deposit outflows, which directly led to its bankruptcy and a potential systemic banking crisis.

machine learning (ML) can efficiently capture and analyze the interaction of various factors simultaneously, rendering a more accurate and holistic measurement (Gu et al., 2020). For example, Loughran and McDonald (2020) proposes a text-based ML approach to construct a dictionary of complexity-related words from companies' 10-K reports. However, their ML analysis relies on simple linear technology that does not take into account the nonlinear interactions and higher-order mechanisms. This implies that their proposed dictionary-based complexity measure does not capture firm complexity to the fullest.[4]

Motivated by the approach of Bali et al. (2023), which employs various ML model specifications to represent heterogeneous investors with differing beliefs, we utilize ML models to emulate human investors and evaluate both the complexity and opacity of BHC entities through the lens of the trained ML models. Based on the proposed complexity and opacity measures, we investigate their impact on the trading activities of BHC stocks, their risk-adjusted returns, and systemic risk.

If we think about complexity in terms of ambiguity about stock returns, where investors share different views about the asset payoff, the literature on decision-making under uncertainty provides some foundational background. Such theoretical models suggest that firms exhibiting higher levels of uncertainty adversely affect some investors' trading activities, potentially leading to an equilibrium characterized by limited market participation (Cao et al., 2005; Eisfeldt et al., 2023). Such equilibrium implies that the return on the risk security diminishes with higher uncertainty among investors. In a similar intuition, Bali et al. (2023) show that stocks with higher ML-agents disagreement exhibit lower returns. Their argument can be easily understood from the point of investors' disagreement and limits to arbitrage (Miller, 1977). As investors face high uncertainty about the stock value, stock prices tend to be upward biased when less optimistic investors face greater short-sale constraints. On the other hand, Eisfeldt et al. (2023) contend that complex assets exhibit a higher risk-adjusted return. We investigate these hypotheses empirically at a later stage in our study.

[4]Additionally, the proposed dictionary is also subject to euphemisms obfuscation, as highlighted by Suslava (2021).

Our manuscript addresses two main questions. First, how can we utilize an ML data-driven approach to measure a firm's complexity and opacity levels? In addressing this, we are also interested in understanding the relationship or the difference between opacity and complexity. Second, the resulting measures reflect the complexity and opacity faced by investors in pricing banks. In this regard, what implications do these measures have for understanding investors' trading behaviors and the cross-sectional variation in stock returns?

In answering the above questions, we make several contributions to the literature. Our first contribution is the proposal of novel firm complexity and opacity measures that leverage recent advancements in ML and explainable artificial intelligence (XAI). We do so by utilizing the proposed methodology by Molnar et al. (2020) in quantifying ML complexity models and applying it to the firm level. The idea leverages what is known as "accumulated local effects" (ALE) functions (Apley and Zhu, 2020). In particular, these functions enable the evaluation of simplicity in mapping firm characteristics into contemporaneous stock returns. Using a pre-trained ML model, we proxy the sensitivity of the ML mapping function to each characteristic, allowing us to capture two aspects. In the first one, we gauge complexity in the sense of the shape of the mapping function, which is done via ALE.[5] Molnar et al. (2020) defines such measure as the "main effect complexity." The second aspect is how nonlinear the mapping function is. Thanks to the ALE function, we are able to decompose the bank's return into first-order approximations, whereas the residual constitutes the nonlinear exposure.[6] In line with Molnar et al. (2020), we refer to the latter as "interaction strength." Combined, these elements offer an innovative methodology for quantifying the complexity involved in mapping the characteristics of banks onto their stock returns.

To quantify opacity, we utilize the pre-trained ML model to measure the bank's opaqueness in line with the literature on bank transparency (see, e.g., Chen et al. (2022)). The

[5]For instance, if the mapping function is linear, then it can be explained by a single segment using piecewise regression, whereas a more complex nonlinear spline would require more than a single segment to depict the shape of the mapping technology.

[6]In the case of a linear model such ordinary least squares regression, the nonlinear exposure is zero, and the fitted value from the model can be explained using the ALE functions alone.

literature measures transparency using the coefficient of determination (R^2) in mapping a firm's characteristics into some response variable, such as loan performance. If the firm is more transparent (opaque), we expect higher (lower) R^2. In this context, we measure opacity as $1 - R^2$ based on the ML pre-trained model.

Both procedures allow us to gauge both complexity and opacity at the firm level using common characteristics studied by Gu et al. (2020), resulting in a rich firm-month panel. Utilizing such a panel, we conduct several descriptive and panel regression analyses to deepen our understanding of both measures. To the best of our knowledge, our study is the first to utilize XAI and nonlinear ML models for assessing bank complexity and opacity. Comparing the newly proposed measures, we find a high correlation between complexity and opacity, with an average correlation coefficient of 0.6 at the firm level, which increases to 0.9 at the industry level. This result illustrates an important aspect of the intertwined relationship between complexity and opacity. In particular, as the system grows in complexity, it becomes increasingly challenging for market participants to grasp its operations due to the increased opaqueness.

Our second main contribution is to evaluate the proposed complexity and opacity measures from investors' perceptions in terms of trading activity and stock return reaction. First, our findings reveal that increased complexity (opacity) leads to a decline in market participation in terms of a lower turnover ratio and dollar volume. This finding aligns well with the hypothesis raised from both Cao et al. (2005) and Eisfeldt et al. (2023). Second, our empirical analysis illustrates that this complexity and opacity are positively associated with future stock returns. These results align with the findings of Bali and Zhou (2016) and Bali et al. (2017b), who provide evidence at both portfolio and stock levels suggesting that equities with increased exposure to market-wide uncertainty tend to earn higher future risk-adjusted returns. However, these findings are inconsistent with those from Cao et al. (2005), Baltussen et al. (2018), and Ruan (2020), who show reduced stock (option) returns in the presence of uncertainties. Additionally, complexity and opacity are negatively (positively) related to future stock volatility (Sharpe ratio), which aligns with the study by Eisfeldt et al. (2023).

Third, our research also investigates the relationship between the two proposed measures and other observable macro variables at the aggregate level, offering critical insights into the relationship between complexity and financial stability. Our work empirically supports the theory that the complexities of modern financial markets can trigger systemic market failures (Battiston et al., 2016; Botta et al., 2022; Schwarcz, 2009) and calls for regulation on bank complexity and opacity level. Though we cautiously do not establish a causal relationship between complexity and financial stability, our analysis provides some descriptive insights. In particular, when we link our complexity index at the aggregate level with common macro indicators from Welch and Goyal (2008), the proposed measure exhibits a counter-cyclical pattern on the aggregated industry level, peaking on the eve of a financial crisis and then decreasing significantly during periods of financial distress; however, the relationship is weakened after the global financial crisis.

One potential explanation for the above macro results is asset bubbles. Specifically, when bubbles form in the financial market, the increasing number of noise traders makes it harder for rational investors to accurately grasp the market dynamics, resulting in a higher complexity index; however, when these bubbles burst, many noise traders opt out during the bear market, making it simpler for investors to comprehend the market mechanism. This speculation also provides some ground for the appropriateness of the "leaning against the wind" strategy.

Earlier literature (see, e.g., Goetz et al. (2016)) suggests that bank complexity – in terms of organizational and geographical – mitigates systemic risk, leading to stability during 2005–07 due to diversification benefits. However, these benefits are short-lived (Bakkar and Nyola, 2021). In light of this, we study the effect of complexity on systemic risk using Acharya et al.'s (2017) measure of systemic expected shortfall (SES) and marginal expected shortfall (MES). Our analysis reveals that during periods of market distress, complexity positively influences a BHC's MES level; in essence, the higher the complexity, the more susceptible a BHC is to the impacts of a financial crisis. Next, by conducting a case study on the 2007-2009 global financial crisis, we arrive at the same conclusion: complexity has a positive impact on a BHC's vulnerability and its overall contribution to systemic risk.

Finally, our proposed complexity and opacity measures shed important implications for the cross-section of stock returns. For instance, the proposed complexity measure can serve as an instrument for exploring existing anomalies - not only for the banks but also for non-financial firms. Indeed, our measure is not only bank-specific but can be generalized to a larger universe of stocks. This is especially relevant in the context of existing research which examines the impact of complexity on information processing and the subsequent effect on the efficiency with which asset prices incorporate all available information. From this view, higher complexity implies greater time to process information, leading to more pronounced return predictability (Cohen and Lou, 2012).[7] By highlighting how increased complexity extends the information processing timeline, our proposed measures are designed to provide a deeper comprehension of anomaly exploration, offering a valuable tool for future researchers.[8]

Our paper proceeds as follows. In Section 2, we conduct the literature review, whereas Section 3 outlines the methodology for quantifying ML model complexity utilizing a recent Explainable AI (XAI) tool based on the Accumulated Local Effect (ALE) function and introduces two distinct measures of model complexity. This is followed by a description of the data, its preprocessing, and the variables utilized in this study, as detailed in Section 4. The approach to forming the BHC group alongside two control groups is then demonstrated. Subsequently, the method to construct the complexity and opacity index is presented. In

[7]In Section IA.1.2 of the Internet Appendix, we build upon the "pseudo" methodology utilized in Cohen and Lou (2012) and deploy our proposed complexity measure to assess its return predictability in predictive panel regression. Consistent with the original paper, we find that the coefficient of the proposed counterfactual measure is positive and statistically significant.

[8]In Section IA.1.1 of the Internet Appendix, we provide basic cross-sectional analysis using portfolio formation. We put the proposed measure to the test to investigate one of the famous anomalies studied in the banking industry (Gandhi and Lustig, 2015). This analysis provides some further implications about the role of complexity in the cross-section of stock returns.

Section 5, an empirical investigation into the research questions developed earlier in this section is conducted, and the results are discussed. The paper concludes with a summary and discussion of future work in Section 6.

2 Literature Review

Our study first relates to measuring firms' complexity level, especially for the BHCs. From a regulatory perspective, they use an indicator-based measurement approach to quantify a financial institution's systemic importance, and one of the five categories is complexity. The more complex a bank is, the more likely it is to have a positively correlated systemic impact in the event of distress or failure. This means that highly complex banks tend to have higher costs and longer timeframes for resolution. The Financial Stability Board considers OTC derivatives notional value, Level 3 assets, and Held for trading and available for sale value (Basel Committee on Banking Supervision, 2013) to assess this asset complexity level. While in academia, the current literature proposes different approaches to quantifying firm and bank complexity. For instance, two distinct complexity metrics (geographic and business complexity) are proposed in Cetorelli et al. (2014). Business complexity is a normalized Herfindahl-Hirschman index (HHI) depending on the number of subsidiaries by business type relative to the total number of subsidiaries while geographical complexity is a normalized HHI depending on the number of subsidiaries by region relative to the total number of subsidiaries. Later, an additional aspect of complexity related to organizational structure is created, proxied with the count of non-bank and foreign subsidiaries that the BHC conglomerate owns. These three types of complexities, which primarily rely on entity, industry, and geographic location counts, are widely used in BHC complexity literature (Cetorelli and Goldberg, 2016; Correa and Goldberg, 2022; Carmassi and Herring, 2016; Goldberg and Meehl, 2020; Barth and Wihlborg, 2017).

Extending beyond literature that explores individual banks' complexity, Zhou (2009) centers on the systemic significance of banks within the broader banking network. The paper looks into the systemic relevance of financial institutions, challenging the traditional "too big to fail" (TBTF) notion through theoretical and empirical prospects, arguing that a financial institution's size should not be seen as a systemic importance proxy. Based on the measure 'probability that at least one bank becomes distressed' (PAO) introduced by Segoviano Basurto and Goodhart (2009), he introduces two innovative measures: a Systemic Importance Index (SII) gauging the expected count of bank failures within the system should a selected bank fail; and a Vulnerability Index (VI), the SII index's inverse, assessing the probability of a specific bank's failure given another failure within the system already exists. These metrics shed light on the TBTF effect and investigate big banks' complexity levels, particularly the top 100 banks identified by the Financial Stability Board as G-SIBs (Global Systemically Important Banks).

Our study relates not just to bank complexity but also to the broader topic of measuring complexity in firms. Researchers rely on 10-K file size or word count, Fog index, and number of segments of a firm to assess its complexity. The study by Loughran and McDonald (2020) introduces a novel text-based ML technique to gauge firm complexity through an analysis of 10-K annual reports. Initially, they develop a dictionary of complexity-indicative words derived from the 10-K reports, comprising 374 terms, such as "bankruptcies", "counterparties", "lawsuits", "leases", "swaps", and "worldwide." These terms cover the firm's complexity as perceived by investors estimating future cash flows or auditors preparing financial statements. Utilizing Lasso regression, they narrow down the most significant words from this preliminary list. The chosen dependent variables include audit fees, the absolute value of unexpected earnings, and stock return volatility post-filing date. Complexity-related words are then defined as those that simultaneously positively impact all three dependent variables. The ultimate complexity score for a firm is calculated as the sum of the word count of each identified word from this procedure relative to the total word count in the 10-K filing, presented as a percentage.

In addition, our study also relates to the literature on measuring bank opacity level and its effect. In the study Chen et al. (2022), the notion of bank earnings informativeness is termed as the adjusted R-squared from a linear regression model, aiming to predict future loan write-offs from quantitative variables like loan loss provisions, earnings before loan loss provisions, and changes in non-performing loans. The paper studies the relationship between bank transparency and depositor behavior, particularly concerning uninsured deposit flows in US commercial banks from 1994 to 2019. Through the regression model, the authors illustrate that a higher level of bank transparency, represented by increased R-squared values, correlates with increased sensitivity of uninsured deposit flows to the banks' performance metrics. This exploration sheds significant light on the interplay between banking transparency and depositor behavior, underscoring the importance of transparency in banking regulations and its impact on the financial behavior of depositors.

An alternative approach involves utilizing the residuals from linear models as indicators of opacity. Researchers apply linear models to predict loan loss provisions (LLP) (Jiang et al., 2016; Zheng and Wu, 2023) or discretionary accruals (Hutton et al., 2009), and then use the absolute values of the residuals from these models as a metric for opacity. In this context, higher residual values correspond to increased levels of opacity. Intensification of competition reduces bank opacity levels as shown by Jiang et al. (2016), while Hutton et al. (2009) identify a greater risk of stock price crashes in opaque firms, and Zheng and Wu (2023) find a negative link between opacity and bank valuation during the 2007–2009 global financial crisis.

Blau et al. (2017) investigate the effect of bank opacity on stock price efficiency. They use three illiquidity measures (turnover ratio, bid-ask spread, and Amihud's (2002) illiquidity to represent the individual bank's opacity level. When the risk information about bank assets is relatively obscured, it can impede stock price efficiency, rendering them less indicative of the banks' actual value or inherent risks. Utilizing the measure of price delay from Hou and Moskowitz (2005) as a lens to scrutinize stock price inefficiency, the authors find evidence that opacity correlates positively with price delay. Compared to non-bank stocks, BHCs show a substantially higher delay. Within the BHC group, the regression results also demonstrate that three measures of opacity positively affect price delay.

Given that increased levels of BHC complexity and opacity could potentially escalate uncertainty, our study aligns with the literature examining the impact of uncertainty on stock performance and the risk-return trade-off. For instance, in the work of Baltussen et al. (2018), authors illustrate that stocks characterized by high-risk uncertainty, as measured by the volatility of implied volatility (vol-of-vol), consistently under-perform those with low-risk uncertainty by 8% annually[9]. This vol-of-vol effect is unique from 20 previously documented return predictors and withstands numerous robustness checks. They also explore the pricing mechanism behind the vol-of-vol effect and find empirical evidence favoring the limited participation theory proposed by Cao et al. (2005).

In Cao et al. (2005), authors illustrate theoretically that model uncertainty and heterogeneous uncertainty-averse investors can lead to limited market participation. In normal cases when the uncertainty level is low, there is an equilibrium with full market participation, while with the increase of model uncertainty, those investors experiencing higher levels of uncertainty (higher aversion level to uncertainty) opt out of the market, and this leads to a scenario of limited participation. In this state of limited participation, the investors who do participate tend to possess either more accurate information regarding outcomes or a reduced aversion to uncertainty, leading them to demand a smaller uncertainty premium.

Besides these papers discussing the relation between uncertainty and return premium, the recent work by Eisfeldt et al. (2023) examines the effect of asset complexity on realized risk-adjusted returns and market participation. The authors introduce a new theoretical model that generates lower equilibrium participation in markets with higher Sharpe ratios due to

[9]Related to the above literature, Ruan (2020) unveils a notable negative relation between equity option returns and the vol-of-vol after accounting for a wide range of existing options and stock characteristics. Furthermore, the univariate sort of option portfolios on vol-of-vol produces both statistically and economically significant negative risk-adjusted return of the high minus low return spread, and these alphas also exist after controlling for a range of control variables in the double sorting portfolio formation process.

different levels of idiosyncratic risk of complex assets raised by investors' different individual arbitrage models. They show that investors with more expertise, better models, and lower resulting idiosyncratic risk exposures realize higher Sharpe ratios, and their demand deters the entry of less sophisticated investors, causing market dislocations.

Finally, our work relates to the current literature on emulating economic agents with ML models. In particular, authors in Bali et al. (2023) propose a novel statistical model that captures differences in beliefs among heterogeneous investors, who are represented through distinct ML technology. Each investor (represented by a distinct set of model hyperparameters), forms return forecasts based on input data that are accessible to all investors. The level of disagreement is measured as the dispersion in forecasts across the different investor models. The authors find that their measure of disagreement proves to be a significantly stronger predictor of future returns than existing belief dispersion measures, such as analyst forecast dispersion. They document a robust negative cross-sectional relation between belief disagreement and future returns. They present that a long-short portfolio strategy that shorts stocks with high forecast disagreement while longs those with low disagreement can achieve a value-weighted annual alpha of 15%.

To the best of our knowledge, there is no existing literature on using nonlinear ML models to measure the complexity and opacity of BHCs at the same time. Our study holds the potential to bridge this existing gap significantly. It aims to validate the theoretical hypothesis and empirically explore the impact of bank complexity and opacity on BHC stock trading.

3 Methodology

Our study introduces a novel methodology to gauge the complexity of BHCs, drawing upon the emergent advancements in XAI within the domain of ML. Utilizing a trained ML model that maps different firm characteristics (predictors) with bank stock return, we provide two measures for bank complexity and opacity using ML to emulate equity investors. A more nonlinear and complex model structure suggests that the operational processes of the given BHC might be more challenging for human investors to comprehend, indicating a higher level of complexity. While it is challenging to disentangle complexity from opacity, we provide a measure of each inspired by the literature. We devote the next discussion to the ML complexity implementation.

3.1 Measuring Model Complexity

The proposed bank complexity measure relies on a measurement of ML complexity. To quantify the model complexity of different ML algorithms, we rely on two metrics proposed by Molnar et al. (2020), Main Effect Complexity (MEC) and Interaction Strength (IAS). Both ideas are built on ALE[10]. Before we implement these tools on our sample, we present a high-level numerical example using toy data based on the California Housing Dataset prepared by Pace and Barry (1997) to motivate the main idea of the ML complexity measure.

3.1.1 Functional Complexity

Let $f : \mathbb{R}^{N\times d} \to \mathbb{R}$ be an ML mapping function that transforms a data matrix $\mathbf{X}$ with dimensions $\mathbb{R}^{d\times N}$ into the output column vector $\hat{\mathbf{y}}$ of size $\mathbb{R}^{N\times 1}$. Here, d and N represent the number of features and number of observations of the training data, respectively. Based on Equation (3) from Molnar et al. (2020), for any given single observation (x) with d dimensions, the fitted value is then estimated as:

$$f(x) = f_0 + \sum_{j=1}^{d} \overbrace{f_{j,ALE}(x_j)}^{\text{MEC: How complex?}} + \overbrace{IA(x)}^{\text{IAS: Interaction strength?}} . \tag{3.1}$$

Equation (3.1) has three components. The first is an intercept f_0 that denotes the mean

[10]We refer readers to Apley and Zhu (2020) for more details.

of all fitted values $\hat{y}_i$, calculated as

$$f_0 = \sum_{i=1}^{N} \hat{y}_i. \tag{3.2}$$

The second component of Equation (3.1) corresponds to the first-order sensitivity of the ML mapping function f to each feature in the data. We refer to Equations (7) and (8) from Apley and Zhu (2020) for more details and introduce the ALE function at a high level. Specifically, x_j denotes the j^{th} feature value of the single observation (x) (and $x_{\setminus j}$ as a column vector to represent the other feature values for that observation. Similarly, X_j represents the whole j^{th} column and $\mathbf{X}_{\setminus j}$ is a matrix that contains all data except column j. For a function f that is differentiable, let $f^j\left(x_j, \mathbf{x}_{\setminus j}\right) \equiv \frac{\partial f\left(x_j, \mathbf{x}_{\setminus j}\right)}{\partial x_j}$ denote the partial derivative of $f(\mathbf{x})$ with respect to x_j. Then the uncentered ALE function g is calculated as:

$$g_{j,ALE}\left(x_j\right) = \int_{x_{\min,j}}^{x_j} \mathbb{E}\left[f^j\left(X_j, \mathbf{X}_{\setminus j}\right) \mid X_j = z_j\right] dz_j. \tag{3.3}$$

The (centered) ALE main effect of x_j, denoted by $f_{j,ALE}\left(x_j\right)$, is defined the same as $g_{j,ALE}\left(x_j\right)$ but centered and therefore the distribution of the first-order ALE function on j^{th} feature has a mean of zero.

$$\begin{aligned} f_{j,ALE}\left(x_j\right) &\equiv g_{j,ALE}\left(x_j\right) - \mathbb{E}\left[g_{j,ALE}\left(X_j\right)\right] \\ &= g_{j,ALE}\left(x_j\right) - \int p_j\left(z_j\right) g_{j,ALE}\left(z_j\right) dz_j \end{aligned} \tag{3.4}$$

In sum, the second component from Equation (3.1) (given by the ALE functions) quantifies the average effect on the model prediction as one feature changes while keeping other features fixed. In Figure 1, we visualize the ALE functions for four features from the California Housing Dataset, including MedInc, HouseAge, Latitude, and Longitude. We choose four models that include Lasso Regression, Support Vector Regression, Decision Tree Regression, and Random Forest Regression. The light blue histogram in the background is the distribution of values upon which the ALE is computed. Lasso Regression is the only model that has a linear ALE relationship, while Decision Tree and Random Forest exhibit a high degree of nonlinearity. The relationship between feature changes and their impact on model prediction is as expected. Specifically, areas with high median income tend to have a higher median housing value. The variable HouseAge captures the development of the neighborhood: a more developed and established neighborhood tends to have a higher median housing value. From the ALE plots examining Longitude and Latitude, it is evident that housing values are generally higher in Southern California (smaller Longitude), and properties near the coast (bigger Latitude) also typically have higher prices. In all cases, we use 30 bins to calculate the ALE approximation, and the smoothed lines are derived using a local regression with a span of 10%.

The third component from Equation (3.1), $IA(x)$, represents the residuals after controlling for the first-order ALE approximation. In line with Molnar et al. (2020), we refer to these residuals as the Interaction Strength that captures the higher-order effects of the model. A linear model, as it can be fully captured by the first-order ALE approximation, results in a zero IAS effect, while for other nonlinear models, this term should be non-zero. We provide further details on the IAS and MEC metrics in Sections 3.1.2 and 3.1.3, respectively.

3.1.2 Interaction Strength

Given on the decomposition from Equation (3.1), the ALE main effect model is defined as the sum of first-order effects, which consists of the first two parts of Equation 3.1:

$$f_{ALE,1st}(x) = f_0 + f_{1,ALE}\left(x_1\right) + \ldots + f_{d,ALE}\left(x_d\right). \tag{3.5}$$

IAS is the third part of the Equation (3.1) $IA(x)$, which is defined as the approximation error measured with loss L:

$$IAS = \frac{\mathbb{E}\left(L\left(f, f_{ALE1st}\right)\right)}{\mathbb{E}\left(L\left(f, f_0\right)\right)} \geq 0 \tag{3.6}$$

For a regression problem using mean squared error (MSE) as the loss function, the IAS equals 1 minus the R-squared, where the R-squared is the score of using first-order ALE approximation to represent the ML prediction output $\hat{y}$. If the chosen ML algorithm is linear, then the ALE function fully approximates the mapping function such that the residual part of $IA(x)$ equals 0, i.e., there are no interactions among different features given that the model is a linear map of the original feature space.

3.1.3 Main Effect Complexity

The second complexity measure provides a quantification on how complex the shape of the ALE function is. This is given by the Main Effect Complexity, which is a weighted average of the number of piece-wise line segments needed to approximate ALE functions:[11]

$$MEC = \sum_{j=1}^{d} \omega_j MEC_j. \tag{3.7}$$

with

$$\omega_j = \frac{V_j}{\sum_{j=1}^{d} V_j} \tag{3.8}$$

A couple of comments are in order. First, V_j denotes the variability of the ALE approximation of feature j. Since $f_{j,ALE}$ is centered around 0, V_j is calculated as the mean of the squared ALE approximated values for the j^{th} feature. Second, MEC_j is defined as the number of line segments needed to approximate the ALE function curve with piece-wise linear regression with a tolerance level of ϵ, which is usually set to be 0.05, namely to achieve an R-squared level of 95%. We use numerical examples to demonstrate the concept of MEC in Fig 2. Here, we choose Random Forest as the ML model, whereas the hyperparameters remain the same. The dots denote the values computed using ALE, whereas the red smoothed line is the approximated ALE function curve using a local regression with a span of 10%. What's new in the figure is the black line segments that use piece-wise linear regression to approximate

[11] See Algorithm 2 by Molnar et al. (2020) for further details.

the ALE values to achieve a minimum R-squared level of 95%. For instance, to approximate the ALE function for MedInc and Latitude, only one piece is needed to achieve an R-squared above 95%. On the other hand, to approximate the ALE values for HouseAge and Longitude, 3 pieces of line segments are needed. With all V_j and MEC_j calculated, one can derive the model MEC value.

Table 1 reports the MEC and IAS values for five different ML algorithms, including both linear and nonlinear models. We begin our discussion with MEC values first. We note that in all cases, we standardize all d features using Z-score normalization to ensure that the variability of ALE values from a single feature does not overly impact the overall model's MEC value. Two linear models (Elastic Net and Lasso) both have MEC values of 1, as one only needs one piece of line segment to approximate the ALE function of each feature. The MEC values for nonlinear models range from 1.4 to 1.8, and the Random Forest model is the most complex model in terms of the MEC measure. Both linear models also have zero interaction strength, as there are no residuals left after approximating the first-order ALE functions. As the IAS values for nonlinear models are non-zero, there is a noticeable interaction among the features.

Given the above summary of the XAI technique to measure the complexity of ML models, we will discuss how we link this to our proposed measure of bank complexity and opacity in Section 4.2.

4 Empirical Framework

In this section, we describe our study's data and empirical framework. In the first part, we describe our sample construction and data sources. Next, we discuss our proposed complexity and opacity measures based on the methodology described in the previous section.

4.1 Sample Construction and Data Source

We begin constructing our data sample using the Center for Research in Security Prices (CRSP) monthly data and applying several filters to select securities. Consistent with the empirical asset pricing literature, we choose securities with a Share Code (SHRCD) of 10 or 11 that are listed on the major exchanges of NYSE, NYSE, and NASDAQ, which corresponds to exchange codes (EXCHCD) of 1, 2, and 3. We also exclude any securities with trading status codes (TRTSCD) of 3, 4, or 5. After implementing these filters, we remove observations with missing return data, constituting approximately 0.77% of the data. In our next phase, we address potential duplicate data. Grouping by PERMCO and date, about 2% are identified as duplicates. Adjusting our grouping criteria to CUSIP and date, this rate is significantly reduced to less than 0.01%. For these duplicate entries, we consolidate them based on CUSIP and date and then compute the market cap weighted return and the month-end price.

Turning to the COMPUSTAT database, we gather fundamental company metrics like total assets and common/ordinary equity. Aware of potential lags in information transmission, we advance the regulatory data by six months. Duplicate entries in this set are also addressed by computing the average of our variables of interest grouping by CUSIP and quarter. We consider 94 firm characteristics by Gu et al. (2020) and Gu et al. (2021).[12] For brevity, we refer to the data as GKX94. Finally, we merge the CRSP and COMPUSTAT datasets by CUSIP and quarter and then with the GKX94 data by PERMNO and month. To address constraints related to liquidity and potential limits to arbitrage, we exclude stocks with a quarter-end price below 5 for the subsequent quarter, following the approach outlined by Li et al. (2014).

To identify BHCs in our sample, we adopt multiple methodologies. First, we source data from the Federal Reserve Bank of New York, [13] which provides us with names, entity types, PERMCO, and other pertinent details dating back to 1986. From this, we distinguish

[12]The monthly data is publicly available on the personal website of Dacheng Xiu

[13]See the website of Federal Reserve Bank of New York for more information.

BHCs with instance types explicitly labeled as "Bank Holding Company" or "Bank holding company." This approach yields a total of 1201 entities. Subsequent techniques pivot around the historical SIC code (SICCD) and the Header SIC code (HSICCD) from CRSP. Second, we identify banks that either have a specific SICCD of 6712 or codes commencing with the digits 60. This method leads to a total number of 1935 banks. Third, we target banks in CRSP with an HSICCD of 6712 or those starting with 60, resulting in the identification of 1793 BHCs. Last, we follow Gandhi and Lustig (2015)to isolate BHCs. This procedure emphasizes entities with a SICCD beginning with 60 or having an HSICCD of 6712, covering a total of 1933 BHCs[14]. As a confirmation, we examine the number of banks in our sample between 2000 and 2008. Overall, our final sample reveals 634 BHCs during this period, a number that is consistent with the literature (Gandhi and Lustig, 2015).

To better understand our findings, we consider a control group of non-banks. Adopting the approach from Flannery et al. (2004) and Blau et al. (2017), we pair BHC with non-bank entities annually. Specifically, we classify non-financial firms as those with SIC (Standard Industrial Classification) codes outside the 6000-6999 range. At the end of each year, we pair each BHC with a non-bank entity based on their year-end market capitalization and share price. Specifically, for each bank-year observation, a non-financial firm is matched if its market value is nearest to a particular BHC and its share price falls within 25% of that BHC's share price for the subsequent year. Additionally, we ensure that the matched non-financial entities are listed on the same exchange as their BHC counterparts.

4.2 Bank Complexity and Opacity Measure

Consistent with the methodology described in Section 3, we employ a machine-learning model to represent human investors. The model is trained to map the current period characteristics into stock returns. Our conjecture is that the more nonlinear and complex the mapping function is, the greater the complexity associated with processing BHC operations and,

[14] Further details on this approach are described in Gandhi and Lustig (2014).

hence, pricing their risk-taking, resulting in greater uncertainty. In the following, we provide further details on the final measurement of bank complexity as well as opacity.

4.2.1 Firm Level Complexity and Opacity Index

The ML mapping function we define decomposes the current period return using a set of firm characteristics. We can view such a task as a filtering problem where the conditional expectation of the current return of stock i at month t is a function of current observable characteristics, $X_{i,t}$, such that

$$\mathbb{E}_t(r_{i,t}|X_{i,t}) = f(X_{i,t}). \tag{4.1}$$

We choose a set of characteristics, by following Gu et al. (2020)'s recommendation on feature importance. To derive this feature importance for a given predictor, they perform sensitivity analysis where they calculate the reduction in R^2 by assigning all zero values to that specific predictor within every training sample for each model, then average this to obtain a single importance measure for each feature. According to Gu et al. (2020), the most crucial firm-level characteristics are grouped into four categories: contemporary price movements, liquidity factors, risk indicators, valuation ratios, and fundamental signals.[15] There are in total of 20 firm characteristics in our final set of predictors used for our ML model. We report these variables in Table 2.

For the ML mapping technology, we choose the Random Forest algorithm given its robust capabilities in forecasting in line with Bali et al. (2023). For implementation, we apply a rolling window of 5 years, namely 60 observations, to derive the firm-level complexity metrics from Section 3 for bank i and time t. In all cases, we tune the hyperparameters in the Random Forest regression model to mitigate overfitting by following Bali et al. (2023)'s recommendation. Due to the problem of lack of data observations and features, we reduce the maximum tree depth from 6 to 4, and the number of trees in the ensemble from 2000 to 200, but remain the fraction of the features and of the sample to be taken as $\frac{\log_2 20}{20}$ and

[15] We refer readers to Section 2.3 in Gu et al. (2020) for more details.

0.1, respectively. Furthermore, a minimum of 36 data observations from the past five years is required to be incorporated into the training sample and to assess the model's complexity. After training the model for bank i and time t, we calculate the two model complexity scores, MEC and IAS, using the methodology from Section 3. From these three scores, we define the time t complexity score for bank i, $COM_{i,t}$, as the sum of the MEC and IAS scores:

$$COM_{i,t} = MEC_{i,t} + IAS_{i,t} \tag{4.2}$$

Additionally, we leverage the in-sample R^2 from the training data to measure the bank's opacity. Specifically, following the common approach in the literature (Chen et al., 2022), we measure opacity as a function of the goodness of fit. For instance, how transparent it is to map stock returns using current characteristics. The more transparent the process is, the higher the R^2 is. For this reason, we compute opacity as

$$OPA_{i,t} = 1 - R^2_{i,t} \tag{4.3}$$

5 Empirical Results

We divide this section into different parts. First, we examine the summary statistics of our data sample for the BHC group and the matched firms group. Then, we investigate the effect of complexity and opacity on investors' trading activities, market participation, and BHC's vulnerability and overall contribution to the systemic risk during the financial crisis. In the Internet Appendix, we utilize the newly proposed complexity to study the cross-section of stock returns. Specifically, we put the proposed measure into an important test to examine the bank size anomaly (Gandhi and Lustig, 2015).

5.1 Summary Statistics

As a first glance at the data, we provide summary statistics of both the BHC group and the matched group in Table 3. The table reports the sample mean value for different characteristics, covering Hou and Moskowitz's (2005) price delay, turnover ratio, illiquidity, book-to-market ratio, market beta, and idiosyncratic volatility.[16] Additionally, we report stock return, volatility, month-end price, and market cap. We represent these summary statistics based on ten size groups (deciles) to control for size effect. The last row of each panel reflects the equally weighted mean value of the entire sample. In all cases, we winsorize the chosen variables on a quarterly basis at the 5% level.

In line with the findings of Blau et al. (2017), we observe that, on average, BHC firms exhibit a higher Hou and Moskowitz's (2005) price delay compared to their matched group (0.38 vs. 0.22). Our analysis also confirms that BHC firms tend to have lower turnover (0.52 v.s. 1.32), higher illiquidity (0.83 v.s. 0.43), higher book-to-market ratio (0.82 v.s. 0.62), lower market beta (0.64 v.s. 1.02), and lower idiosyncratic volatility (3.93% v.s. 5.8 %). Additionally, we observe several trends across both categories. For instance, price delay and illiquidity decrease as firms increase in size, whereas turnover ratio and beta increase. Nevertheless, the book-to-market ratio displays an inverse pattern, where it decreases with size for the BHC firms but increases for the non-BHC firms. Lastly, since we run the matching algorithm based on the market cap and the price, the means of the monthly end price and market cap for both groups are close.

5.1.1 Overview of BHC Complexity and Opacity Measure

To ensure the robustness of our training sample and to evaluate the complexity and opacity of each bank at time t, we require a minimum of 36 data observations from the past five years for each bank. This requirement results in a notable reduction in the available training

[16] The idiosyncratic volatility is calculated by estimating the standard deviations of residuals from a daily CAPM model.

data. In total, we have 74,760 bank-month observations that include complexity and opacity scores. It is noteworthy that in the following empirical analysis, we first apply Z-score normalization to MEC, IAS, and in-sample R^2 to calculate the final complexity and opacity scores[17]. The number of BHCs significantly increases in 1996, which subsequently leads to a surge in the number of banks with complexity scores around 1999, as shown in Figure 3a. Furthermore, we aggregate the firm-level BHC complexity and opacity scores to compute the industry-level index. This aggregation is performed by assigning weights based on each BHC's total assets. Figure 3b presents the industry-level complexity and opacity scores for BHCs. To enhance the clarity of the data, we follow Jiang et al.'s (2019)'s recommendation and apply a moving average with a window size of 4 months to smooth the time series. As a result of this smoothing process, the correlation between Opacity and Complexity reached a substantial 0.9.

A couple of comments follow from Figure 3b. First, both indices display several pronounced declines, some of which coincide with recognized financial distress periods, highlighting the Asian financial crisis spanning June 1997 to December 1998, the 'dot-com bubble' burst from 2000 to 2002, and the recent global financial crisis between 2007 and 2008. During these periods, the BHC complexity and opacity indices drop significantly. Second, financial distress might not be the sole catalyst for a decline in the BHC complexity (opacity) Index. Notably, we observe a consistent, prolonged downward trend starting in 2016, a period devoid of any major financial crises. We suspect that this downward trend is caused by the BHC industry becoming more transparent. Following the 2008 global financial crisis, there was a strong global push for increased transparency in banking. BHCs had been facing stricter regulations regarding disclosing information on their lending practices, fee structures, risk management, and investment policies to avoid the opacity that contributed to the crisis.

[17] We note that such analysis is descriptive rather than predictive.

5.2 Relation between BHC Complexity and Trading Activeness

Next, we investigate how complexity/opacity is linked with trading activity. Both Cao et al. (2005) and Eisfeldt et al. (2023) contend that investor participation tends to decrease as asset complexity or uncertainty levels increase. This phenomenon may result in an equilibrium characterized by "limited market participation," where only investors with sufficient expertise or high confidence are willing to invest in such complex assets. To empirically test their hypothesis, we conduct a panel regression analysis. We estimate the following equations using pooled stock-month data for all BHCs:

$$\begin{aligned} turn_{i,t+1}\,(dolvol_{i,t+1}) =&\alpha + \beta_1\, COM_{i,t}(OPA_{i,t}) + \beta_2\, Market\_Distress_t \\ &+\beta_3\, COM_{i,t}\,(OPA_{i,t}) * Market\_Distress_{i,t} + \beta_4\, Size_{i,t} \\ &+\beta_5\, bm_{i,t} + \beta_6\, PRC_{i,t} + \beta_7\, illi, t + \beta_8\, beta_{i,t} + \epsilon_{i,t} \end{aligned} \tag{5.1}$$

The analysis covers the time period from January 1996 to December 2019, spanning a total of 23 years. Tables 5 and 6 summarize the regression results, where the dependent variable is the next period turnover ratio (turn) and dollar volume (dolvol), respectively, provided in the GKX94 dataset. The two variables of interest are complexity and opacity. Suspecting that complexity and opacity may have time-varying effects during normal times and market-distressed periods, we include an interaction term (Complexity/Opacity * Market Distress), where Market Distress is a binary indicator if the monthly-end VIX index is over the top 20% over the total data sample. Size is calculated as the logarithmic value of the market capitalization; BM is the book-to-market ratio; PRC is the monthly end price; 'ill' is Amihud's (2002) illiquidity[18]; and beta is the CAPM beta estimate for each firm during the year. We control for year and firm fixed effects. Standard errors are enclosed in parentheses for all reported values.

Tables 5 and 6 cover a total of 6 model results, respectively, as we choose two different

[18] Amihud's (2002) illiquidity is calculated as the ratio of the absolute value of the monthly return scaled by the monthly volume

sets of independent variables and three regression models. For the two sets of independent variables, we include either complexity or opacity and their interaction term with market distress, respectively. It is worth noting that with a correlation of approximately 0.6 between complexity and opacity, one should not use both measures simultaneously to avoid the problem of co-linearity. Within each set of independent variables, we conduct three distinct regression models, including pooled regression, regression incorporating fixed entity effects, and the regression model combining year and firm fixed effects.

Our variables of interest are complexity, opacity, and their interaction terms with the binary variable of market distress. The regression results show that the coefficients for the complexity and opacity variables are consistently negative. This indicates that an increase in asset complexity and uncertainty levels tend to reduce market participation, aligning with the theoretical predictions by Cao et al. (2005) and Eisfeldt et al. (2023). Furthermore, the coefficients of the interaction terms for the turnover ratio are negative and statistically significant, whereas those for the next period's dollar volume traded do not show such significance. This suggests that the impact of complexity and opacity on trading activities does not depend on market conditions.

5.3 Relation between BHC Complexity and Risk-Return Tradeoff

In the following section, we investigate the impact of complexity on the performance of BHCs, specifically focusing on their risk-return dynamics. As in the preceding section, we employ different panel regression models, with the dependent variables encompassing next period monthly returns (ret), monthly realized volatility (retvol), and Sharpe ratios(SR) — calculated as the monthly return divided by monthly volatility. The selection of independent variables mirrors that of the previous analysis:

$$
\begin{aligned}
ret_{i,t+1}\,(retvol_{i,t+1},\ SR_{i,t+1}) =& \alpha + \beta_1\, COM_{i,t}(OPA_{i,t}) + \beta_2\, Market\_Distress_t \\
&+\beta_3\, COM_{i,t}\,(OPA_{i,t}) * Market\_Distress_{i,t} + \beta_4\, Size_{i,t} \\
&+\beta_5\, bm_{i,t} + \beta_6\, PRC_{i,t} + \beta_7\, illi, t + \beta_8\, beta_{i,t} + \epsilon_{i,t}
\end{aligned}
\tag{5.2}
$$

Table 7 displays the findings on the relationship between BHCs' complexity/opacity and their monthly returns. During normal times, the coefficients for both complexity and opacity on future returns are positive, meaning that investors require a premium for holding stocks of more complex banks, compensating them for the additional risk or the effort needed to understand the complexity. The negative coefficient for the interaction term indicates that the relationship between complexity and next-period return becomes significantly more negative during market distress. Risk aversion tends to rise in such periods, leading investors to favor straightforward and transparent investments. Consequently, complex banks may become less preferred, and therefore, their lack of transparency might trigger more extensive sell-offs, driving down their returns.

Similarly, Table 8 presents the results when the dependent variable is the volatility of the monthly returns. During normal conditions, opacity and complexity are both associated with negative effects that increased level of these two factors will reduce the next period return volatility, even after accounting for the liquidity measure of Amihud's (2002) illiquidity. However, during market distressed periods, the impact of these factors on return volatility becomes positive, meaning that stocks with higher complexity are likely to experience increased volatility, as a result of potential sell-offs during market crashes.

Table 9 explores the connection between the complexity/opacity of BHCs and their risk-adjusted returns, represented by Sharpe ratios. Given that Sharpe ratios are calculated by dividing returns by volatility, it's anticipated that more complex banks would have higher Sharpe ratios under normal market conditions due to increased returns and lower volatility. Conversely, this trend is expected to invert in market downturns, with decreased returns and increased volatility. Our empirical analysis supports this hypothesis, which reveals positive coefficients for complexity and opacity and negative coefficients for the interaction terms.

In conclusion, contrary to Cao et al.'s (2005) assertion that limited market participation lowers the equity premium, our findings indicate that increased asset complexity will increase stock returns during normal times. Our explanation of the phenomenon is that agents in Cao et al.'s (2005) model settings do not demand liquidity premium, while investors in the real world do seek compensation for the illiquidity stemming from the high complexity level of assets. Furthermore, we find a negative relationship between asset complexity and stock variance. Therefore, our empirical analysis supports the theory in Eisfeldt et al. (2023), which suggests that the Sharpe ratio improves as the asset complexity level rises.

5.4 Relation between BHC Complexity and Systemic Risk

In the following discussion, we first link the proposed complexity index with common macroeconomic variables to conduct a descriptive analysis of the relationship between our measure of complexity and market regimes. In Welch and Goyal (2008), the authors investigate the influence of 14 macroeconomic variables on stock equity premiums. Given the significant role of banks in the financial industry, we explore the effects of our newly introduced BHC industry-level complexity score on these variables. Table 4 provides a summary of these macro variables. For each time period, we categorize the macro variables into three groups based on the BHC aggregate complexity score. Group 1 comprises those falling below the 25th percentile, Group 2 includes values ranging from the 25th to the 75th percentile, and Group 3 represents the top 25th percentile. The table displays the means of these variables within each respective group. We report the correlation between the BHC complexity score and the macro variables in the row labeled correlation. Furthermore, we conduct a one-way t-test and report the p-values. It is worth noting that we have adjusted the variables in accordance with the methodology outlined in Welch and Goyal (2008). The variables include logarithmic values of dividend-price ratio (dp), logarithmic values of dividend yield (dy), logarithmic values of Earnings-price ratio (ep), logarithmic values of Dividend-payout ratio (de), stock return variance (svar), book-to-market ratio (bm), net equity expansion (ntis), treasury bill rate (tbl), long-term yield (lty), long-term return (ltr), term spread (tms), default yield spread (dfy), default return spread (dfr), and inflation (infl).

As depicted in Figure 3b, the initial three significant drops in our complexity (opacity) indices align with known periods of financial distress. However, the fourth drop is linked to the bear market following the global financial crisis. Consequently, we split our analysis into three categories: 'Pre 2009', 'Post 2009', and 'Entire Data Duration'. The relationship between our complexity index and the macroeconomic factors varies across these three segments. Yet, several factors maintain the same association with our index throughout. The first set of variables includes dividend-price ratio, dividend yield, and dividend-payout ratio. All exhibit a negative correlation with the complexity index. Assuming firms, on average, keep dividend distributions steady regardless of good or bad years, this negative correlation suggests an increase in either price or net income when the complexity index goes up. Furthermore, stock return variance consistently relates to our index; during times of a high complexity index, the stock return variance tends to be low. Lastly, a negative relationship exists between the default yield spread and our complexity index.

In Figure 4, we plot the relationship between our complexity index and four chosen macroeconomic variables. The graph clearly illustrates a strong negative correlation between complexity and each of the four macroeconomic variables. Yet, this correlation weakens after the global financial crisis. We speculate that this change might be due to banks being mandated to enhance their transparency after the crisis. Therefore, we cautiously conclude that the complexity index tends to be high just before a financial crisis and subsequently drops during periods of financial distress. When bubbles form in the financial market, the increasing number of noise traders makes it harder for rational investors to accurately grasp the market dynamics, resulting in a higher complexity index. However, when these bubbles burst, many noise traders opt-out during the bear market, making it simpler for investors to comprehend the market mechanism.

After discovering that complex banks experience lower returns and higher volatility during financial crises, a natural follow-up question is how this complexity affects the systemic risk of BHCs. Goetz et al. (2016)) suggests that bank complexity - in terms of organizational and geographical - mitigates systemic risk, leading to stability during 2005–07 due to diversification benefits. However, these benefits are short-lived (Bakkar and Nyola, 2021). Using a cross-European dataset on bank internationalization, authors investigate whether complexity of the European banks affects systemic risk differently during normal times and distress times, they find out that though both organizational and geographical complexity in banks reduced systemic risk and increased stability during 2005-07, attributed to enhanced diversification benefits and minimized risks from asset similarity, during the acute crisis of 2008-11 and the post-crisis years of 2012-13, complexity contributed to systemic risk and instability. In summary, while bank complexity can lower systemic risk under normal conditions, its impact reverses during crises. In light of this, we study the effect of complexity on systemic risk using Acharya et al.'s (2017) measure of systemic expected shortfall (SES) and marginal expected shortfall (MES). Acharya et al. (2017) introduce an economic model focused on systemic risk; authors introduce SES and argue that the SES of each financial institution represents its likelihood of being undercapitalized at times when the entire system is experiencing under-capitalization, thereby measuring its contribution to systemic risk of the financial system. Furthermore, they find out that an elevated level of MES in a bank—referring to its losses in the tail of the system's loss distribution—results in an increased level of SES. We study the effect of our measure of complexity on the SES and MES to understand the role of complexity in the financial crisis.

To measure the MES for each BHC, we follow the Acharya et al.'s (2017) method. It is estimated at a standard risk level of $\alpha = 5\%$ using daily equity returns data from CRSP. For each year, we take the 5% worst days for the market returns and compute the equal-weighted average return on any given firm i for these days:

$$\text{MES}^i_{5\%} = \frac{1}{\#\text{days}} \sum_{t:\text{worst } 5\% R_t} R^i_t \tag{5.3}$$

Similar to (Bakkar and Nyola, 2021; Goetz et al., 2016), We hypothesize that during market normal times, the increased level of complexity will reduce BHCs' systemic risks through the diversification channel as allocating investments across different asset classes with imperfectly correlated returns mitigate idiosyncratic risks; however, during market distressed times an increase in a bank's complexity contributes to higher MES as complexity-averse investors are likely to divest from banks with elevated complexity during periods of heightened market-wide systematic risk. The resultant selling pressure significantly depresses the stock prices of complex banks, leading to larger losses during a financial crisis. Furthermore, building upon the findings of Acharya et al. (2017) indicating that SES increases with a bank's MES, we further posit that the amplification in the bank's complexity would also result in increased SES.

$$\begin{aligned} MES_{i,t} =& \alpha + \beta_1 \, COM_{i,t-1} + \beta_2 \, Market\_Distress_t + \beta_3 \, COM_{i,t-1} * Market\_Distress_{i,t} \\ &+ \beta_4 \, Size_{i,t-1} + \beta_5 \, bm_{i,t-1} + \beta_6 \, PRC_{i,t-1} + \beta_7 \, ill_{i,t-1} + \beta_8 \, beta_{i,t-1} + \epsilon_{i,t} \end{aligned} \tag{5.4}$$

Except for the variable Market Distress, all independent variables are lagged by one period. They are computed as the average of the given Bank Holding Company (BHC) over the previous year. Market Distress is a binary indicator that takes a value of one if the year t includes months classified as part of a recession by the NBER and zero otherwise. Our data sample includes three recession periods: August 1990 to March 1991, April 2001 to November 2001, and the well-known global financial crisis from January 2008 to June 2009.

Table 10 displays the results. Our results align with Bakkar and Nyola's (2021) findings that complexity will have opposite effects across market normal and distressed periods. During normal times, complexity will reduce the extreme losses; if the market experiences a shock in the following year, then the heightened level of complexity at the end of the preceding year becomes significant. An increase of one standard deviation in our complexity measure can elevate the daily MES by 0.4%, which translates to an extra annual loss of around 10%.

Next, we explore whether complexity mitigates extreme losses via the diversification channel though we do not intend to establish a causal relationship but rather a descriptive analysis. We study the relationship between BHC complexity and diversification over time. BHC level data are obtained through the FR Y9-C reports available at the WRDS database. We then use the link file[19] provided by New York Fed to link regulatory identification numbers (RSSD ID) to the permanent company number (PERMCO) used in CRSP. We consider two measures of diversification. The first index considers the asset allocation diversification, calculated as 1 minus the sum of the squared weights of 18 asset classes (Duarte and Eisenbach, 2021) held by a BHC[20]. A higher asset diversification score indicates that a BHC has greater diversification in its asset allocations. The second metric measures diversification through the ratio of non-interest income to total income, which is calculated as BHCK4079 / (BHCK4074 + BHCK4079).

Figure 5 presents two diversification indices, categorized by size and complexity, across different time periods. Group 1 consists of entities below the 25th percentile, Group 2 encompasses those between the 25th and 75th percentiles, and Group 3 includes the top 25th percentile each quarter, based on either complexity or total assets. When comparing banks based on their size, it is evident that larger banks have a more diversified asset portfolio and sources of income than smaller banks. Using our measure of complexity, the cutoff becomes less clear. From 2000 to 2010, more complex banks exhibited greater diversification, but this trend became less pronounced after 2012. We suspect this is due to the implementation of the Comprehensive Capital Analysis Review (CCAR) stress test as part of the Dodd-Frank

[19]For access to the exact file, please refer to the provided link

[20]The mapping process follows the detailed description by (Duarte and Eisenbach, 2021) and Clark et al. (2023) provides detailed R code in their Internet Appendix.

Act, which mandated an increase in transparency disclosures (García and Steele, 2022). BHCs with assets exceeding \$100B in total assets for 2011-2012 or \$50B from 2013-2016 were required to comply, though later relaxed by the 'Economic Growth, Regulatory Relief, and Consumer Protection Act' issued in 2018. In summary, it can be concluded that larger banks exhibited greater diversification and complexity before the CCAR implementation, although this correlation was weakened after 2012.

Furthermore, given that market recession periods comprise only a small fraction of the data, we have chosen to focus on a case study examining the impact of complexity on the global financial crisis. We are interested in two variables, realized SES (the realized returns during the crisis) and the MES during the entire crisis. We take the 5% worst days for the market returns during the financial crisis and compute the equal-weighted average return on any given firm i during these days. As noted in Fig 3 there is a significant drop of the complexity index during the global financial crisis. We suspect that bubbles existing in the financial market before the financial crisis increase the complexity level. In contrast, when these bubbles burst, many noise traders opt out during the bear market, leading to a major drop in our complexity index. Therefore, we modify our regression to accommodate these observations.

$$\begin{aligned} MES_i\,(SES_i) =& \alpha + \beta_1\,COM_i + \beta_2\,COM\_F1_i + \beta_3\,Size_i \\ &+ \beta_4\,bm_i + \beta_5\,PRC_i + \beta_6\,ill_i + \beta_6\,beta_i + \epsilon_i \end{aligned} \tag{5.5}$$

In calculating the independent variables for the regression, we take the mean values from the last quarter of 2007 to represent the pre-global financial crisis conditions. Meanwhile, the variable COM_F1 is defined as the mean complexity level during the financial crisis period. Through this regression, we would like to investigate how the complexity level before and during the financial crisis may affect a BHC's vulnerability and its overall contribution to systemic risk. Table 11 reports the regression results.

In our data sample, 220 BHCs survived the global financial crisis. Our results show that complexity level is a significant determinant of a BHC's vulnerability and its contribution to overall systemic risk. The higher the complexity level before the crisis, the more negative the realized return, thereby increasing both the MES and SES. On the other hand, we show that complexity levels before and during the crisis exhibit an opposite relation, possibly indicating that complexity is resolved or attenuated through investors' active trading during the crisis.

6 Conclusion

Our study presents a novel method to employ machine learning (ML) algorithms to measure the complexity and opacity of individual banks. The premise behind our proposed measure is as follows. First, we use nonlinear ML models to emulate human investors who map observable characteristics into stock returns. Second, utilizing recent advancements in explainable artificial intelligence (XAI), we gauge how complex the mapping process is, corresponding to our proposed complexity measure. This is achieved after proper model training and tuning, which also allows us to assess bank opacity using the model's in-sample R^2. Our empirical examination shows that our proposed measure of complexity and opacity are highly correlated. The increased complexity (opacity) is associated with reduced market participation but, surprisingly, increases stock returns, reduces stock volatility, and improves the stock Sharpe ratio. Furthermore, we investigate the impact of complexity on systemic risk and find that the increased level of complexity is positively associated with BHC's vulnerability and its overall contribution to systemic risk during the financial crisis.

The current study inspires several directions for future research. First, our work has solely utilized Random Forest regression for assessing BHC complexity, leaving room for incorporating additional ML models in future investigations. By employing various ML models with differing levels of nonlinearity, it is possible to emulate investors of varying expertise levels, such that the bank complexity is represented by the average of heterogeneous model complexities. Second, our study can be extended later to measure the complexity of non-BHC firms and more research questions related to comparing the complexity between banks and non-banks. Due to the opaqueness of the banking industry (Morgan, 2002), such analysis would provide a proper identification to test our measures and investigate the effectiveness of bank-specific policies/regulations. We leave this for future research.

References

Abedifar, P., Bouslah, K., Zheng, Y., 2021. Stock price synchronicity and price informativeness: Evidence from a regulatory change in the us banking industry. Finance Research Letters 40, 101678.

Acharya, V.V., Pedersen, L.H., Philippon, T., Richardson, M., 2017. Measuring systemic risk. The review of financial studies 30, 2–47.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of financial markets 5, 31–56.

Apley, D.W., Zhu, J., 2020. Visualizing the effects of predictor variables in black box supervised learning models. Journal of the Royal Statistical Society Series B: Statistical Methodology 82, 1059–1086.

Baily, M.N., Klein, A., Schardin, J., 2017. The impact of the dodd-frank act on financial stability and economic growth. RSF: The Russell Sage Foundation Journal of the Social Sciences 3, 20–47.

Bakkar, Y., Nyola, A.P., 2021. Internationalization, foreign complexity and systemic risk: Evidence from european banks. Journal of Financial Stability 55, 100892.

Bali, T.G., Brown, S.J., Murray, S., Tang, Y., 2017a. A lottery-demand-based explanation of the beta anomaly. Journal of Financial and Quantitative Analysis 52, 2369–2397.

Bali, T.G., Brown, S.J., Tang, Y., 2017b. Is economic uncertainty priced in the cross-section of stock returns? Journal of Financial Economics 126, 471–489.

Bali, T.G., Kelly, B.T., Mörke, M., Rahman, J., 2023. Machine Forecast Disagreement. Technical Report. National Bureau of Economic Research.

Bali, T.G., Zhou, H., 2016. Risk, uncertainty, and expected returns. Journal of Financial and Quantitative Analysis 51, 707–735.

Baltussen, G., Van Bekkum, S., Van Der Grient, B., 2018. Unknown unknowns: uncertainty about risk and stock returns. Journal of Financial and Quantitative Analysis 53, 1615–1651.

Barth, J.R., Wihlborg, C., 2017. Too big to fail: Measures, remedies, and consequences for efficiency and stability. Financial Markets, Institutions & Instruments 26, 175–245.

Basel Committee on Banking Supervision, 2013. Global systemically important banks: updated assessment methodology and the higher loss absorbency requirement.

Battiston, S., Farmer, J.D., Flache, A., Garlaschelli, D., Haldane, A.G., Heesterbeek, H., Hommes, C., Jaeger, C., May, R., Scheffer, M., 2016. Complexity theory and financial regulation. Science 351, 818–819.

Blau, B.M., Brough, T.J., Griffith, T.G., 2017. Bank opacity and the efficiency of stock prices. Journal of Banking & Finance 76, 32–47.

Botta, A., Caverzasi, E., Russo, A., 2022. When complexity meets finance: A contribution to the study of the macroeconomic effects of complex financial systems. Research Policy 51, 103990.

Bouvard, M., Chaigneau, P., Motta, A.D., 2015. Transparency in the financial system: Rollover risk and crises. The Journal of Finance 70, 1805–1837.

Caballero, R.J., Simsek, A., 2013. Fire sales in a model of complexity. The Journal of Finance 68, 2549–2587.

Cao, H.H., Wang, T., Zhang, H.H., 2005. Model uncertainty, limited market participation, and asset prices. The Review of Financial Studies 18, 1219–1251.

Carmassi, J., Herring, R., 2016. The corporate complexity of global systemically important banks. Journal of Financial Services Research 49, 175–201.

Cetorelli, N., Goldberg, L.S., 2016. Organizational complexity and balance sheet management in global banks. Technical Report. National Bureau of Economic Research.

Cetorelli, N., Goldberg, L.S., et al., 2014. Measures of global bank complexity. FRBNY Economic Policy Review 20, 107–126.

Chen, Q., Goldstein, I., Huang, Z., Vashishtha, R., 2022. Bank transparency and deposit flows. Journal of Financial Economics 146, 475–501.

Clark, B., Francis, B., Simaan, M., 2023. Pricing banks: Risk and return in an opaque industry. Available at SSRN 3558546 .

Cohen, L., Lou, D., 2012. Complicated firms. Journal of financial economics 104, 383–400.

Correa, R., Goldberg, L.S., 2022. Bank complexity, governance, and risk. Journal of Banking & Finance 134, 106013.

Dang, T.V., Gorton, G., Holmström, B., Ordonez, G., 2017. Banks as secret keepers. American Economic Review 107, 1005–1029.

Diamond, D.W., Dybvig, P.H., 1983. Bank runs, deposit insurance, and liquidity. Journal of political economy 91, 401–419.

Duarte, F., Eisenbach, T.M., 2021. Fire-sale spillovers and systemic risk. The Journal of Finance 76, 1251–1294.

Eisfeldt, A.L., Lustig, H., Zhang, L., 2023. Complex asset markets. The Journal of Finance 78, 2519–2562.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of financial economics 33, 3–56.

Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. Journal of financial economics 116, 1–22.

Federal Reserve, 2023. Review of the Federal Reserve's Supervision and Regulation of Silicon Valley Bank. Technical Report. Federal Reserve. Washington, D.C. URL: https://www.federalreserve.gov/publications/files/svb-review-20230428.pdf.

Flannery, M.J., Kwan, S.H., Nimalendran, M., 2004. Market evidence on the opaqueness of banking firms' assets. Journal of Financial Economics 71, 419–460.

Gandhi, P., Lustig, H., 2015. Size anomalies in us bank stock returns. The Journal of Finance 70, 733–768.

Gandhi, P., Lustig, H.N., 2014. On identifying commercial banks in crsp. Available at SSRN 2518562 .

García, R.E., Steele, S., 2022. Stress testing and bank business patterns: A regression discontinuity study. Journal of Banking & Finance 135, 105964.

Goetz, M.R., Laeven, L., Levine, R., 2016. Does the geographic expansion of banks reduce risk? Journal of Financial Economics 120, 346–362.

Goldberg, L.S., Meehl, A., 2020. Complexity in large us banks. Economic Policy Review 26.

Granja, J., 2023. Bank fragility and reclassification of securities into htm. University of Chicago, Becker Friedman Institute for Economics Working Paper .

Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. The Review of Financial Studies 33, 2223–2273.

Gu, S., Kelly, B., Xiu, D., 2021. Autoencoder asset pricing models. Journal of Econometrics 222, 429–450.

Hou, K., Moskowitz, T.J., 2005. Market frictions, price delay, and the cross-section of expected returns. The Review of Financial Studies 18, 981–1020.

Hutton, A.P., Marcus, A.J., Tehranian, H., 2009. Opaque financial reports, r2, and crash risk. Journal of financial Economics 94, 67–86.

Jiang, F., Lee, J., Martin, X., Zhou, G., 2019. Manager sentiment and stock returns. Journal of Financial Economics 132, 126–149.

Jiang, L., Levine, R., Lin, C., 2016. Competition and bank opacity. The Review of Financial Studies 29, 1911–1942.

Li, X., Sullivan, R.N., Garcia-Feijóo, L., 2014. The limits to arbitrage and the low-volatility anomaly. Financial Analysts Journal 70, 52–63.

Loughran, T., McDonald, B., 2020. Measuring firm complexity. Journal of Financial and Quantitative Analysis , 1–55.

Miller, E.M., 1977. Risk, uncertainty, and divergence of opinion. The Journal of finance 32, 1151–1168.

Molnar, C., Casalicchio, G., Bischl, B., 2020. Quantifying model complexity via functional decomposition for better post-hoc interpretability, in: Machine Learning and Knowledge Discovery in Databases: International Workshops of ECML PKDD 2019, Würzburg, Germany, September 16–20, 2019, Proceedings, Part I, Springer. pp. 193–204.

Moreno, D., Takalo, T., 2016. Optimal bank transparency. Journal of Money, Credit and Banking 48, 203–231.

Morgan, D.P., 2002. Rating banks: Risk and uncertainty in an opaque industry. American Economic Review 92, 874–888.

Newey, W.K., West, K.D., 1987. A simple, positive-definite, heteroscedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Pace, R.K., Barry, R., 1997. California housing data set from statlib. StatLib repository, http://lib.stat.cmu.edu/datasets/.

Ruan, X., 2020. Volatility-of-volatility and the cross-section of option returns. Journal of Financial Markets 48, 100492.

Schwarcz, S.L., 2009. Regulating complexity in financial markets. Wash. UL Rev. 87, 211.

Segoviano Basurto, M., Goodhart, C., 2009. Banking stability measures. IMF Working paper .

Suslava, K., 2021. “stiff business headwinds and uncharted economic waters”: The use of euphemisms in earnings conference calls. Management Science 67, 7184–7213.

Welch, I., Goyal, A., 2008. A comprehensive look at the empirical performance of equity premium prediction. The Review of Financial Studies 21, 1455–1508.

Zheng, Y., Wu, D., 2023. The impact of opacity on bank valuation during the global financial crisis: A channel analysis. International Review of Financial Analysis 87, 102580.

Zhou, C., 2009. Are banks too big to fail? measuring systemic importance of financial institutions. Measuring Systemic Importance of Financial Institutions (December 1, 2009) .

Figures

Figure 1: **Accumulated Local Effects**
The figure provides an empirical illustration of the Accumulated Local Effects by Apley and Zhu (2020). In all cases, a given model is trained using the same data and features. The data corresponds to the California Housing dataset available from scikit-learn first created by Pace and Barry (1997). Four models chosen here are Lasso Regression (Lasso), Support Vector Regression (SVR), Decision Tree Regressor (Decision Tree), and Random Forest Regressor (Random Forest). The response variable denotes the median value of houses within a block and the four chosen explanatory variables include Median of Income, House Age, Longitude, and Latitude. The light blue histogram in the background is the distribution of values the ALE was computed over.

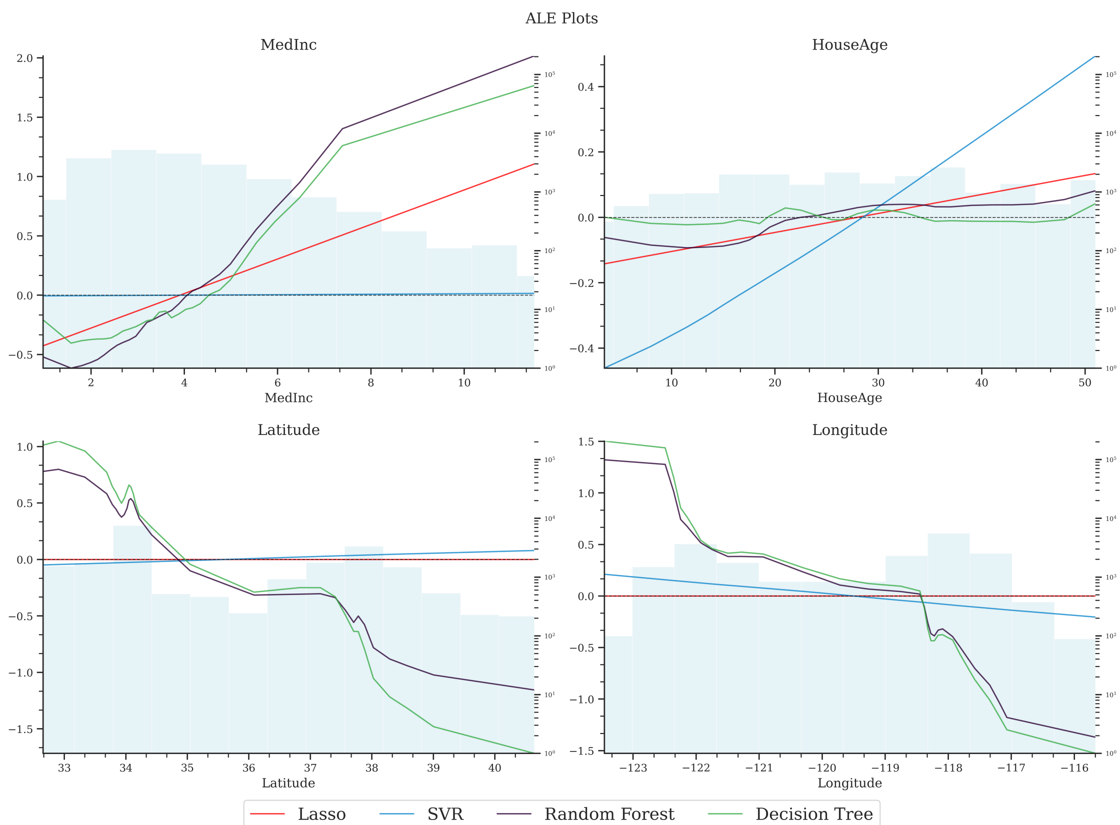


Figure 2: **Main Effect Complexity**

The figure provides an empirical illustration of the Accumulated Local Effects by Apley and Zhu (2020). The Red line denotes the first-order ALE approximation while the black line represents the curve fitted using a piece-wise linear regression to achieve a minimum level of 95% R-squared.

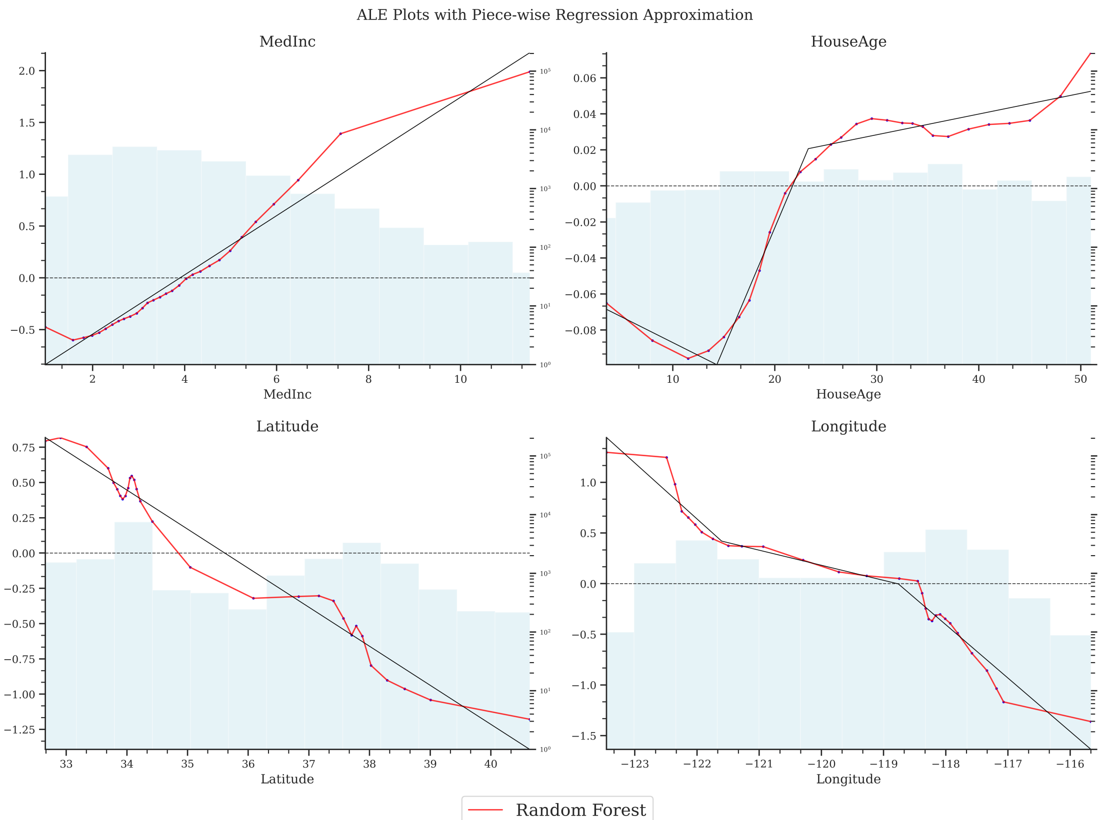
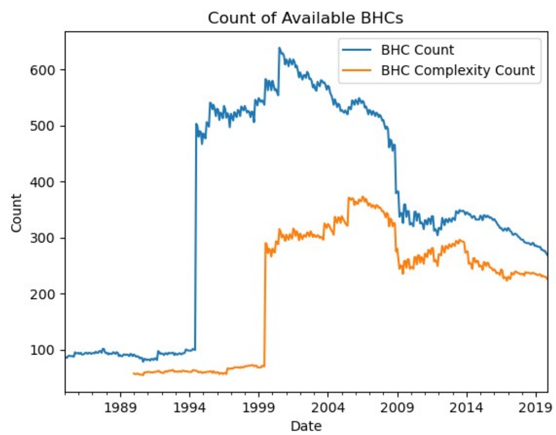


Figure 3: **BHC Complexity and Opacity Summary Statistics**
In the left figure, we present the number of banks and those with complexity scores over the whole time period. The number of BHCs significantly increased in 1996, which subsequently led to a surge in the number of banks with complexity scores around 1999. Furthermore, we aggregate the individual firm-level BHC complexity and opacity scores to compute the industry-level index. This aggregation is performed by assigning weights based on each BHC's total assets. The industry-level complexity and opacity scores for BHCs are shown in the right figure. To enhance the clarity of the data, we follow Jiang et al.'s (2019)'s recommendation and apply a moving average with a window size of 4 months to smooth the time series.

(a) BHC Counts



(b) Aggregated BHCs Complexity and Opacity Index

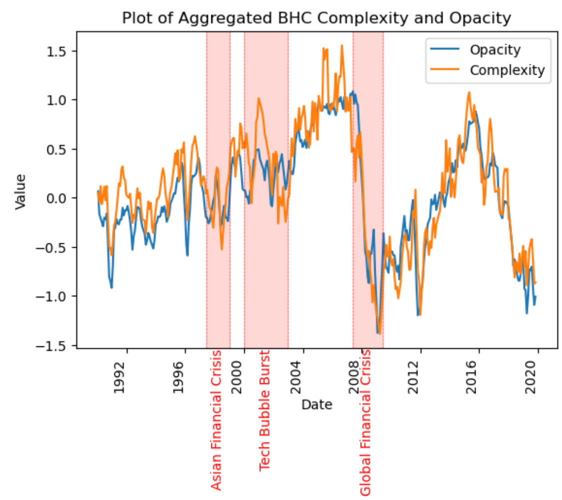
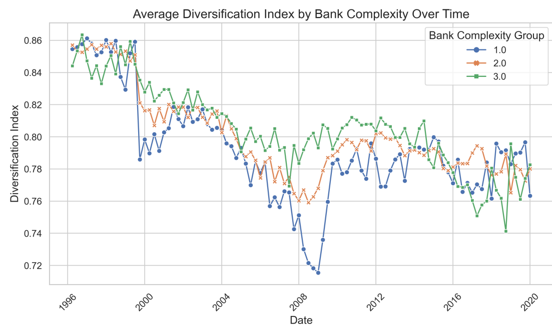


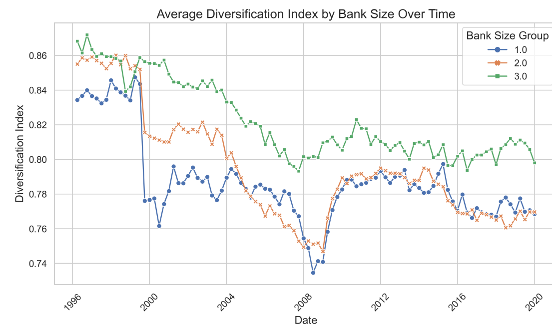
Figure 4: **Relationship between Complexity Index and Other Macroeconomic Variables**
We present the relationship between our complexity index and four macroeconomic variables over time. To enhance the clarity of the data, we follow Jiang et al.'s (2019) recommendation and apply a moving average with a window size of 4 months to smooth the time series. There is a strong negative correlation between complexity and each of these four macroeconomic variables. However, this correlation weakens after the global financial crisis. We cautiously conclude that the complexity index tends to be high just before a financial crisis and subsequently drops during periods of financial distress.

Figure 5: **Bank Diversification Indices Over Time**

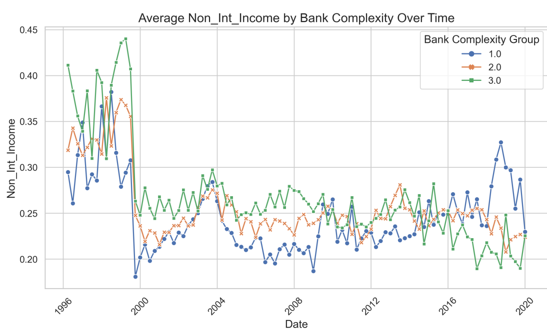
The figure illustrates the diversification trends of BHCs over time using two metrics. The first index is similar to the Herfindahl-Hirschman Index (HHI), calculated as 1 minus the sum of the squared weights of 18 asset classes held by a BHC. The second metric measures diversification through the ratio of non-interest income to total income. BHCs are categorized quarterly into three groups based on total assets or complexity, from the smallest/least complex (Group 1) to the largest/most complex (Group 3), to analyze diversification trends across different BHC profiles. The analysis begins in 1996, the year when BHCs were permitted to include securities within their asset portfolios.



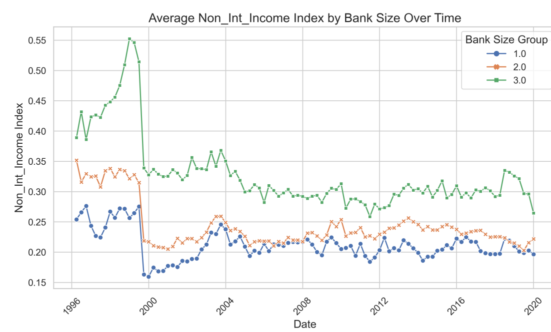
(a) Asset Holding Diversification by Complexity



(b) Asset Holding Diversification by Size



(c) Income Diversification by Complexity



(d) Income Diversification by Size

Tables

Table 1: **Measuring Complexity Statistics.**
This table reports the complexity results for five different machine learning algorithms. IAS denotes the interaction strength computed with respect to the first-order ALE approximation. MEC denotes the main effect complexity.

	IAS	MEC
Elastic Net	0.000	1.000
Lasso	0.000	1.000
SVR	0.277	1.370
Decision Tree	0.319	1.6598
Random Forest	0.229	1.7996

Table 2: **Summary of Variables Used in the ML Model**
According to Gu et al. (2020), the most crucial firm-level characteristics are grouped into four categories. The first category focuses on contemporary price movements, including short-term reversal (mom1m), stock momentum (mom12m), momentum shift (chmom), industry momentum (indmom), recent peak return (maxret), and extended-term reversal (mom36m). Following this, the second category relates to liquidity factors, featuring turnover and its volatility (turn, std turn), logarithmic market equity (mvel1), dollar volume (dolvol), Amihud's (2002) measure of illiquidity (ill), frequency of zero trading days (zerotrade), and the bid-ask spread (baspread). The third set comprises risk indicators, with total and idiosyncratic return volatility (retvol, idiovol), market beta (beta), and squared beta (betasq). The final set includes valuation ratios and fundamental signals such as earnings-to-price ratio (ep), sales-to-price ratio (sp), asset expansion (agr), and the count of recent earnings surges (nincr). Since our analysis corresponds to 'nowcasting' current stock returns rather than forecasting future returns, we replace the short-term reversal (mom1m) with 6-month momentum (mom6m) and remove the industry momentum (indmom).

No.	Acronym	Firm Characteristic	Paper's Author(s)	Year, Journal	Frequency
			Contemporary Price Movements		
1	mom6m	6-month momentum	Jegadeesh & Titman.	1993, JF	Monthly
2	mom12m	12-month momentum	Jegadeesh.	1990, JF	Monthly
3	chmom	Change in 6-month momentum	Gettleman& Marks	2006, WP	Monthly
4	maxret	Maximum daily return	Bali, Cakici & Whitelaw	2011, JFE	Monthly
5	mom36m	36-month momentum	Jegadeesh & Titman	1993, JF	Monthly
			Liquidity Factors		
6	turn	Share turnover	Datar, Naik & Radcliffe	1998, JFE	Monthly
7	std turn	Volatility of share turnover	Chordia, Subrahmanyam, &Anshuman.	2001, JFE	Monthly
8	mvel1	Logarithmic market equity (Size)	Banz	1981, JFE	Monthly
9	dolvol	Dollar trading volume	Chordia, Subrahmanyam & Anshuman.	2001, JFE	Monthly
10	ill	Illiquidity	Amihud	2002, JFM	Monthly
11	zerotrade	Zero trading days	Liu	2006, JFE	Monthly
12	baspread	Bid-ask spread	Amihud & Mendelson	1989, JF	Monthly
			Risk Indicators		
13	retvol	Return volatility	Ang, Hodrick, Xing & Zhang	2006, JF	Monthly
14	idiovol	Idiosyncratic return volatility	Ali, Hwang & Trombley	2003, JFE	Monthly
15	beta	market beta	Fama & MacBeth	1973, JPE	Monthly
16	betasq	Beta squared	Fama & MacBeth	1973, JPE	Monthly
			Valuation Ratios and Fundamental signals		
17	ep	Earnings-to-price ratio	Basu	1977, JF	Annual
18	sp	Sales-to-price ratio	Barbee, Mukherji, & Raines	1996, FAJ	Annual
19	agr	Asset growth	Cooper, Gulen & Schill	2008, JF	Monthly
20	nincr	Number of earnings increases	Barth, Elliott & Finn	1999, JAR	Quarterly

Table 3: **Summary Statistics.**
The table presents statistical summaries of the BHC data utilized in the analysis. The variables included in the table are as follows: price delay (Hou and Moskowitz, 2005), turn (ratio of monthly volume to shares outstanding), ill (Amihud's (2002) measure of illiquidity calculated as the ratio of the absolute value of the monthly return scaled by the monthly volume in millions), bm (Book-to-market equity), beta (CAPM 3 month rolling beta), idiovol (idiosyncratic volatility calculated by estimating the standard deviations of residuals from a daily CAPM model in percentage), baspread (Bid-ask spread rolling 3m in cents), ret (stock monthly raw return), Volatility (stock monthly volatility in percentage), PRC (monthly stock price) and MKTCAP (stock price times shares outstanding in millions).

Size_Group	pricedelay	turn	ill	bm	beta	idiovol	baspread	ret	volatility	PRC	MKTCAP
Panel A. Stock characteristics of banks											
(S)	0.69	0.29	2.03	1.04	0.43	4.33	3.31	0.90	7.23	14.19	61.97
(2)	0.64	0.28	1.74	0.95	0.41	4.11	2.83	1.00	7.18	16.70	79.75
(3)	0.54	0.28	1.52	0.89	0.43	4.03	2.72	1.03	7.23	18.44	103.66
(4)	0.45	0.32	1.08	0.87	0.47	3.88	2.55	0.97	7.41	19.86	124.82
(5)	0.39	0.36	0.86	0.83	0.55	3.87	2.59	1.08	7.48	21.31	197.54
(6)	0.31	0.44	0.56	0.80	0.65	4.00	2.62	1.05	7.89	21.70	252.18
(7)	0.24	0.52	0.32	0.76	0.73	3.85	2.69	1.06	8.05	23.94	379.61
(8)	0.18	0.62	0.14	0.71	0.83	3.80	2.80	1.13	8.26	26.62	645.05
(9)	0.17	0.89	0.05	0.67	0.89	3.79	2.79	1.05	8.29	28.27	1340.33
(B)	0.16	1.21	0.01	0.67	0.97	3.62	2.61	1.07	8.39	39.61	6904.46
Mean	0.38	0.52	0.83	0.82	0.64	3.93	2.75	1.03	7.74	23.06	1008.94
Panel B. Stock characteristics of non-banks											
(S)	0.37	0.92	1.23	0.51	0.85	7.27	4.34	0.27	11.51	13.78	85.92
(2)	0.30	1.01	0.79	0.59	0.91	6.62	4.06	0.53	11.31	16.93	147.69
(3)	0.24	1.18	0.57	0.62	1.02	6.58	4.10	0.89	11.66	19.36	225.64
(4)	0.22	1.19	0.48	0.62	1.04	6.08	3.98	0.79	11.30	21.23	294.21
(5)	0.22	1.22	0.44	0.67	1.01	5.80	3.84	1.04	11.03	21.86	354.52
(6)	0.20	1.29	0.33	0.70	1.09	5.75	3.80	0.88	11.42	23.68	460.50
(7)	0.20	1.39	0.24	0.68	1.09	5.55	3.69	1.00	10.93	25.11	627.36
(8)	0.13	1.60	0.13	0.62	1.14	5.31	3.54	0.79	10.70	27.61	1017.75
(9)	0.12	1.83	0.06	0.59	1.15	5.07	3.38	0.87	10.57	29.88	1870.31
(B)	0.15	1.55	0.00	0.57	0.90	3.93	2.68	0.90	8.97	39.55	9422.26
Mean	0.22	1.32	0.43	0.62	1.02	5.80	3.74	0.80	10.94	23.90	1450.62

Table 4: **Relation between Macro Variables and BHC Complexity by Period**
The table provides a summary of the macro variables utilized in Welch and Goyal (2008). For each time period, we have categorized these macro variables into three groups based on the Complexity scores aggregated for BHCs. Group 1 comprises those falling below the 25th percentile, Group 2 includes values ranging from the 25th to the 75th percentile, and Group 3 represents the top 25th percentile. The table displays the means of these variables within each respective group. We report the correlation between the BHC Complexity score and these macro variables in the row labeled correlation. Furthermore, We conduct the one-way t-test and report the P-values. It's worth noting that we have adjusted the variables in accordance with the methodology outlined in Welch and Goyal (2008). The variables include logarithmic values of dividend-price ratio (dp), logarithmic values of dividend yield (dy), logarithmic values of earnings-price ratio (ep), logarithmic values of dividend-payout ratio (de), stock return variance (svar), book-to-market ratio (bm), net equity expansion (ntis), treasury bill rate (tbl), long-term yield (lty), long-term return (ltr), term spread (tms), default yield spread (dfy), default return spread (dfr), inflation (infl).

	dp	dy	ep	de	svar	b/m	ntis	tbl	lty	ltr	tms	dfy	dfr	infl
							Pre 2009							
Group 1	-4.00	-4.01	-3.31	-0.69	0.01	0.24	0.00	0.03	0.06	0.01	0.02	0.01	-0.00	0.00
Group 2	-4.18	-4.17	-3.26	-0.92	0.00	0.20	0.01	0.04	0.06	0.00	0.02	0.01	-0.00	0.00
Group 3	-4.09	-4.08	-2.97	-1.12	0.00	0.28	0.00	0.04	0.05	0.01	0.01	0.01	0.00	0.00
correlation	-0.29	-0.26	0.44	-0.67	-0.49	0.14	0.20	0.18	-0.08	-0.15	-0.27	-0.43	0.18	0.22
P-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.00	0.18	0.00	0.00	0.33	0.24
							Post 2009							
Group 1	-3.86	-3.84	-3.25	-0.60	0.00	0.33	-0.01	0.01	0.03	0.00	0.03	0.01	0.01	0.00
Group 2	-3.91	-3.90	-3.01	-0.90	0.00	0.32	-0.01	0.00	0.03	0.01	0.02	0.01	0.00	0.00
Group 3	-3.91	-3.90	-3.07	-0.84	0.00	0.31	-0.01	0.00	0.03	0.00	0.02	0.01	-0.00	0.00
correlation	-0.21	-0.24	0.18	-0.20	-0.31	-0.15	-0.00	-0.31	-0.39	-0.02	-0.04	-0.43	-0.11	-0.18
P-value	0.05	0.02	0.04	0.03	0.00	0.26	0.77	0.01	0.00	0.56	0.75	0.00	0.34	0.00
							Whole Data Period							
Group 1	-3.87	-3.86	-3.20	-0.67	0.00	0.32	-0.01	0.01	0.04	0.01	0.02	0.01	0.00	0.00
Group 2	-4.05	-4.05	-3.16	-0.89	0.00	0.25	0.01	0.02	0.05	0.01	0.02	0.01	-0.00	0.00
Group 3	-4.08	-4.07	-3.06	-1.02	0.00	0.27	0.00	0.03	0.05	0.00	0.02	0.01	0.00	0.00
correlation	-0.44	-0.43	0.17	-0.36	-0.29	-0.28	0.25	0.37	0.30	-0.05	-0.24	-0.44	-0.04	0.09
P-value	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.55	0.00	0.00	0.45	0.26

Table 5: **Panel regressions – Dollar Volume using Lagged Data.**

The table presents the outcomes of estimating the following equation using pooled data:

$dolvol_{i,t+1} = \alpha + \beta_1 COM_{i,t}(OPA_{i,t}) + \beta_2 Market\_Distress_{i,t} + \beta_3 COM_{i,t}(OPA_{i,t}) * Market\_Distress_{i,t} + \beta_4\, Size_{i,t} + \beta_5\, bm_{i,t} + \beta_6\, PRC_{i,t} + \beta_7\, ill_{i,t} + \beta_8\, beta_{i,t} + \epsilon_{i,t}$

The dependent variable is the dollar volume (dolvol) provided in the GKX94 dataset. The independent variables have a one-period lag. Complexity and Opacity are two variables of interest. 'Market Distress' is a binary indicator if the monthly-end VIX index is over the top 20% over the total data sample. 'Size' is calculated as the logarithmic value of the market cap. 'BM' is the book-to-market ratio. 'PRC' is the monthly end price. 'ill' is Amihud's (2002) illiquidity. 'beta' is the CAPM beta estimate for each firm during the year. We also control for fixed year effects. Standard errors are enclosed in parentheses for all reported values.

	Next Period Dollar Volume					
	M1	**M2**	**M3**	**M4**	**M5**	**M6**
Complexity	-0.0143***	-0.0067***	-0.0086***			
	(0.0020)	(0.0017)	(0.0016)			
Market_Distress × Complexity	0.0022	0.0130***	0.0052			
	(0.0056)	(0.0046)	(0.0045)			
Opacity				-0.0541***	-0.0364***	-0.0387***
				(0.0037)	(0.0033)	(0.0033)
Market_Distress × Opacity				0.0040	0.0135*	0.0063
				(0.0096)	(0.0080)	(0.0077)
Market_Distress	0.2182***	0.2147***	0.0430***	0.2094***	0.2109***	0.0437***
	(0.0091)	(0.0079)	(0.0089)	(0.0092)	(0.0080)	(0.0089)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects			✓			✓
Firm fixed Effects		✓	✓		✓	✓
No. Observations	69754	69754	69754	69754	69754	69754
R-squared	0.9053	0.5289	0.3069	0.9056	0.5298	0.3083

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 6: **Panel regressions – Turnover Ratio using Lagged Data.**
The table presents the outcomes of estimating the following equation using pooled data:
$turn_{i,t+1} = \alpha + \beta_1 COM_{i,t}(OPA_{i,t}) + \beta_2 Market\_Distress_{i,t} + \beta_3 COM_{i,t}(OPA_{i,t}) * Market\_Distress_{i,t} + \beta_4 Size_{i,t} + \beta_5 bm_{i,t} + \beta_6 PRC_{i,t} + \beta_7 ill_{i,t} + \beta_8 beta_{i,t} + \epsilon_{i,t}$
The dependent variable is the turnover ratio (turn) provided in the GKX94 dataset. The independent variables have a one-period lag. Complexity and Opacity are two variables of interest. 'Market Distress' is a binary indicator if the monthly-end VIX index is over the top 20% over the total data sample. 'Size' is calculated as the logarithmic value of the market cap. 'BM' is the book-to-market ratio. 'PRC' is the monthly end price. 'ill' is Amihud's (2002) illiquidity. 'beta' is the CAPM beta estimate for each firm during the year. We also control for fixed year effects. Standard errors are enclosed in parentheses for all reported values.

	Next Period Turnover Ratio					
	M1	**M2**	**M3**	**M4**	**M5**	**M6**
Complexity	-0.0322***	-0.0203***	-0.0206***			
	(0.0017)	(0.0014)	(0.0013)			
Market_Distress × Complexity	-0.0190***	-0.0115**	-0.0221***			
	(0.0057)	(0.0051)	(0.0049)			
Opacity				-0.0820***	-0.0626***	-0.0621***
				(0.0034)	(0.0033)	(0.0032)
Market_Distress × Opacity				-0.0242**	-0.0174*	-0.0292***
				(0.0116)	(0.0103)	(0.0098)
Market_Distress	0.2410***	0.2135***	0.0293***	0.2291***	0.2071***	0.0300***
	(0.0090)	(0.0074)	(0.0075)	(0.0089)	(0.0074)	(0.0074)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects			✓			✓
Firm fixed Effects		✓	✓		✓	✓
No. Observations	69754	69754	69754	69754	69754	69754
R-squared	0.3430	0.1374	0.0626	0.3485	0.1429	0.0683

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 7: **Panel regressions – Return using Lagged Data.**

The table presents the outcomes of estimating the following equation using pooled data:

$$ret_{i,t+1} = \alpha + \beta_1\, COM_{i,t}(OPA_{i,t}) + \beta_2\, Market\_Distress_{i,t} + \beta_3\, COM_{i,t}\,(OPA_{i,t}) * Market\_Distress_{i,t} + \beta_4\, Size_{i,t} + \beta_5\, bm_{i,t} + \beta_6\, PRC_{i,t} + \beta_7\, illi, t + \beta_8\, beta_{i,t} + \epsilon_{i,t}$$

The dependent variable is the next period stock return (ret). The independent variables have a one-period lag. Complexity and Opacity are two variables of interest. 'Market Distress' is a binary indicator if the monthly-end VIX index is over the top 20% over the total data sample. 'Size' is calculated as the logarithmic value of the market cap. 'BM' is the book-to-market ratio. 'PRC' is the monthly end price. 'turn' is the monthly turnover ratio. 'beta' is the CAPM beta estimate for each firm during the year. We also control for fixed year effects. Standard errors are enclosed in parentheses for all reported values.

	Next Period Return					
	M1	**M2**	**M3**	**M4**	**M5**	**M6**
Complexity	0.0007***	0.0011***	0.0006***			
	(0.0002)	(0.0002)	(0.0002)			
Market_Distress × Complexity	-0.0018***	-0.0019***	-0.0009			
	(0.0006)	(0.0006)	(0.0006)			
Opacity				0.0017***	0.0030***	0.0012***
				(0.0003)	(0.0004)	(0.0004)
Market_Distress × Opacity				-0.0035***	-0.0041***	-0.0020*
				(0.0010)	(0.0010)	(0.0010)
Market_Distress	-0.0014	-0.0059***	5.922e-05	-0.0014	-0.0060***	-0.0001
	(0.0010)	(0.0010)	(0.0012)	(0.0010)	(0.0010)	(0.0012)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects			✓			✓
Firm fixed Effects		✓	✓		✓	✓
No. Observations	69754	69754	69754	69754	69754	69754
R-squared	0.0024	0.0098	0.0113	0.0026	0.0103	0.0114

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 8: **Panel regressions – Return Volatility using Lagged Data.**
The table presents the outcomes of estimating the following equation using pooled data:
$retvol_{i,t+1} = \alpha + \beta_1 COM_{i,t}(OPA_{i,t}) + \beta_2 Market\_Distress_{i,t} + \beta_3 COM_{i,t}(OPA_{i,t}) * Market\_Distress_{i,t} + \beta_4\, Size_{i,t} + \beta_5\, bm_{i,t} + \beta_6\, PRC_{i,t} + \beta_7\, illi, t + \beta_8\, beta_{i,t} + \epsilon_{i,t}$
The dependent variable is the next period stock return volatility (retvol). The independent variables have a one-period lag. Complexity and Opacity are two variables of interest. 'Market Distress' is a binary indicator if the monthly-end VIX index is over the top 20% over the total data sample. 'Size' is calculated as the logarithmic value of the market cap. 'BM' is the book-to-market ratio. 'PRC' is the monthly end price. 'turn' is the monthly turnover ratio. 'beta' is the CAPM beta estimate for each firm during the year. We also control for fixed year effects. Standard errors are enclosed in parentheses for all reported values.

	Next Period Return Volatility					
	M1	**M2**	**M3**	**M4**	**M5**	**M6**
Complexity	-0.0006***	-0.0004***	-0.0002***			
	(0.0001)	(0.0001)	(0.0001)			
Market_Distress × Complexity	0.0008***	0.0007***	0.0003***			
	(0.0001)	(0.0001)	(0.0001)			
Opacity				-0.0015***	-0.0013***	-0.0005***
				(5.629e-05)	(6.026e-05)	(5.704e-05)
Market_Distress × Opacity				0.0012***	0.0011***	0.0004**
				(0.0002)	(0.0002)	(0.0002)
Market_Distress	0.0112***	0.0093***	0.0032***	0.0111***	0.0092***	0.0032***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects			✓			✓
Firm fixed Effects		✓	✓		✓	✓
No. Observations	69754	69754	69754	69754	69754	69754
R-squared	0.1790	0.1963	0.0491	0.1838	0.2001	0.0498

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 9: **Panel regressions – Sharpe Ratio using Lagged Data.**
The table presents the outcomes of estimating the following equation using pooled data:
$SR_{i,t+1} = \alpha + \beta_1\, COM_{i,t}(OPA_{i,t}) + \beta_2\, Market\_Distress_{i,t} + \beta_3\, COM_{i,t}\,(OPA_{i,t}) * Market\_Distress_{i,t} + \beta_4\, Size_{i,t} + \beta_5\, bm_{i,t} + \beta_6\, PRC_{i,t} + \beta_7\, illi,t + \beta_8\, beta_{i,t} + \epsilon_{i,t}$
The dependent variable is the next period stock Sharpe ratio (SR). The independent variables have a one-period lag. Complexity and Opacity are two variables of interest. 'Market Distress' is a binary indicator if the monthly-end VIX index is over the top 20% over the total data sample. 'Size' is calculated as the logarithmic value of the market cap. 'BM' is the book-to-market ratio. 'PRC' is the monthly end price. 'turn' is the monthly turnover ratio. 'beta' is the CAPM beta estimate for each firm during the year. We also control for fixed year effects. Standard errors are enclosed in parentheses for all reported values.

	Next Period Sharpe Ratio					
	M1	**M2**	**M3**	**M4**	**M5**	**M6**
Complexity	0.0447***	0.0589***	0.0299***			
	(0.0093)	(0.0099)	(0.0099)			
Market_Distress × Complexity	-0.0904***	-0.0887***	-0.0147			
	(0.0216)	(0.0216)	(0.0213)			
Opacity				0.1064***	0.1646***	0.0732***
				(0.0161)	(0.0182)	(0.0187)
Market_Distress × Opacity				-0.1611***	-0.1789***	-0.0544
				(0.0354)	(0.0356)	(0.0350)
Market_Distress	-0.0999***	-0.2745***	0.0131	-0.0965***	-0.2736***	0.0053
	(0.0369)	(0.0384)	(0.0463)	(0.0371)	(0.0386)	(0.0465)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects			✓			✓
Firm fixed Effects		✓	✓		✓	✓
No. Observations	69754	69754	69754	69754	69754	69754
R-squared	0.0044	0.0080	0.0084	0.0047	0.0086	0.0085

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 10: **Panel regressions – Marginal Expected Shortfall using Lagged Data.**
The table presents the outcomes of estimating the following equation using pooled data:
$MES_{i,t} = \alpha + \beta_1\, COM_{i,t-1} + \beta_2\, Market\_Distress_{i,t} + \beta_3\, COM_{i,t-1} * Market\_Distress_{i,t} + \beta_4\, Size_{i,t-1} + \beta_5\, bm_{i,t-1} + \beta_6\, PRC_{i,t-1} + \beta_8\, beta_{i,t-1} + \epsilon_{i,t}$
The dependent variable is the next period Marginal Expected Shortfall (MES). The independent variables have a one-period lag. Complexity and its interaction term with 'Market Distress' are two variables of interest. 'Market Distress' is a binary indicator if year t contains the recession months defined by NBER. 'Size' is calculated as the logarithmic value of the market cap. 'BM' is the book-to-market ratio. 'PRC' is the monthly end price. 'turn' is the monthly turnover ratio. 'beta' is the CAPM beta estimate for each firm during the year. We also control for fixed year effects. Standard errors are enclosed in parentheses for all reported values.

Next Period Marginal Expected Shortfall	**M1**	**M2**
Complexity	-0.0012***	-0.0011***
	(0.0001)	(0.0002)
Market_Distress × Complexity	0.0054***	0.0054***
	(0.0006)	(0.0007)
Market_Distress	0.0328***	0.0316***
	(0.0007)	(0.0007)
Controls	Yes	Yes
Firm fixed Effects		✓
No. Observations	5609	5609
R-squared	0.5361	0.5363

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 11: **A Case Study into the Global Financial Crisis**

This table presents the outcomes of estimating the following equation using pooled data:

$MES_i\ (SES_i) = \alpha + \beta_1\ COM_i + \beta_2\ COM\_F1_i + \beta_3\ Size_i + \beta_4\ bm_i + \beta_5\ PRC_i + \beta_6\ ill_i + \beta_7\ beta_i + \epsilon_i$

The dependent variables are the Marginal Expected Shortfall (MES) and the Realized Systemic Expected Shortfall (SES). For the independent variables, we use the average values from the last quarter of 2007 to represent conditions before the global financial crisis. The variable 'COM F1' represents the average complexity level during the crisis. 'COM' and 'COM F1' are the complexity measures of interest. 'Size' is calculated as the logarithm of market capitalization. 'BM' denotes the book-to-market ratio. 'PRC' represents the monthly ending price. An error in the original text was the mention of 'turn' which does not appear in the equation; it might be replaced or omitted as necessary. 'beta' reflects the CAPM beta estimates for each firm during the year. Standard errors are provided in parentheses for all reported coefficients

2007 - 2009 Financial Crisis Case Study	**M1**	**M2**
Dep. Variable	Realized SES	MES
Complexity	-0.0368**	0.0011**
	(0.015)	(0.001)
Complexity F1	0.1162***	-0.0022***
	(0.021)	(0.001)
Controls	Yes	Yes
No. Observations	220	220
R-squared	0.166	0.066

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Internet Appendix

IA.1 Cross-Section of Stock Returns

We put the proposed complexity and opacity measures into an important test to investigate how these characteristics are priced in the cross-section. Specifically, we conduct two empirical analyses. Our first one uses a conventional portfolio formation method, following the approach of Fama and French (1993). Building upon the work of Gandhi and Lustig (2015), we create portfolios based on bank size regarding total assets to reproduce the original size anomaly within BHCs and prove that it exists only in BHCs. Subsequently, we study the impact of various factors in addressing this anomaly by constructing bivariate portfolios. Our findings indicate that illiquidity is a potential solution to the anomaly, and the complexity measure we introduce also contributes to partially resolving it. In our second analysis, we innovate upon Cohen and Lou's (2012) approach and utilize our proposed complexity measure to assess its return predictability. Our empirical evidence demonstrates that investors require additional time to process information from stocks of high complexity. We find a strong return predictability group of low complexity firms to their more complex counterparts, which aligns with Cohen and Lou's (2012) findings.

IA.1.1 Bank Size Anomaly

As an initial investigation, we replicate the original bank size anomaly by Gandhi and Lustig (2015). Consider the excess return of portfolio p during month m denoted as:

$$R_{p,m} - R_{f,m} = \alpha_p + \beta^{'}\mathbf{F_m} + \epsilon_{p,m} \tag{IA1}$$

where $\boldsymbol{F}_m = [mkt_m, smb_m, hml_m, rmd_m, cma_m, ltg_m, crd_m, liq_m]$ denotes a set of common factors. Specifically, mkt, smb, and hml correspond to the market, size, and value premiums, respectively, as introduced by Fama and French (1993). Similarly, rmd and cma denote the profitability and investment-related risk factors from Fama and French (2015) study. Additionally, we incorporate ltg and crd, representing two bond-related factors, namely the term and credit spreads, as utilized in Gandhi and Lustig (2015). Lastly, liq represents the Amihud's (2002) illiquidity factor return. The term ϵ accounts for the idiosyncratic component unique to each portfolio.

For our univariate portfolio analyses, we sort firms into deciles at the end of each quarter based on size (total assets). Next, we compute the loadings associated with each risk factor, leveraging α as a measure of risk-adjusted return. To accommodate the presence of heteroskedasticity and auto-correlation, we refer to Newey and West (1987) using three lags and adjust the standard errors accordingly. Our data sample dates between February 1986 and December 2019, yielding a total of 407 observations for each portfolio group.

The research by Gandhi and Lustig (2015) reveals a notable size anomaly within BHC firms after controlling the *smb* risk factor. Our investigation supports this observation, noting that this size anomaly is prevalent only in the BHC group and not in the matched group or the more comprehensive control group of non-financial firms. As seen in Table IA.1, a clear trend emerges when portfolios are grouped by size. There is a noticeable decrease in risk-adjusted return, represented by α, as size increases. To observe this size anomaly, we perform the factor regression on a long-short portfolio that invests $1 in the largest two deciles and shorts the smallest two deciles. We opt for the top and bottom two deciles to ensure stability in our results. Notably, for this long-short position, we find a monthly loss of 0.55% for BHC entities while this monthly loss is not statistically significant for the matched firms - as Panel B indicates. Furthermore, our research aligns with the original findings that, among BHC firms, the loadings associated with the risk factors *hml* and *smb* display an upward trend in conjunction with company size. However, this trend is not observed in the matched group. Similarly, the risk factor loadings associated with the two bond factors also demonstrate an ascending trend in tandem with size, but these factors exert comparatively less influence on the other two groups.

In Panel A from Table IA.2, we revisit the analysis we presented in Table IA.1 but introduce two more risk factors: profitability *rmw* and investment *cma*. We find that the inclusion of these two significantly largely affects the size anomaly in the BHC group. It significantly diminishes the risk-adjusted return of the long-short portfolio and makes it less statistically significant. The risk-adjusted return for the long-short position reduces to a monthly loss of 0.33%, which is about half of what we observed in IA.1. Indeed, profitability and investment risk factors play key roles in explaining the returns of bigger banks: larger banks have negative and significant loadings on these two factors. In Panel B, we add one more factor, the illiquidity risk factor. This further diminishes the risk-adjusted return of the long-short position in the BHC group. The monthly loss reduces to 0.24% and becomes statistically insignificant. As expected, small portfolios tend to have positive loadings on this illiquidity risk factor, while big portfolios' loadings on *liq* are relatively small and statistically insignificant.

IA.1.1.1 Bivariate Portfolio Analysis to Solve for Size Anomaly in BHCs

From the previous part, we have found that illiquidity plays an important role in explaining the size anomaly in the banking industry. Including the illiquidity risk factor largely reduces the magnitude of the risk-adjusted return for the long-short portfolio (Big-Small Portfolio) and makes it statistically insignificant. However, there is still an average monthly loss of 0.24%. To verify that if the remaining size anomaly α is caused by the fact that the liquidity factor is not solely designed for the banking industry, we next perform a bivariate portfolio analysis to assess the relation between stock excess returns and size (total assets) after controlling for Amihud's (2002) illiquidity measure and other firm characteristics discussed in Table IA.3 and Table IA.4, following Bali et al.'s (2017a) analysis. At the end of each month, all BHC stocks in the sample are sorted into decile groups based on an ascending order of the control variable. The sample date begins from January 1996 to December 2019 to ensure that there are more than 100 BHCs at the end of each month. First part of Table IA.3 illustrates the time-series average returns of assets-weighted portfolios formed by sorting initially on liquidity and then based on total assets. The last column labeled ILL Avg. presents the results for the average ill decile within the given size decile. In the second part, we show the mean return, risk-adjusted returns using different sets of factors (FF3 + two bond risk factors and FF 5 + two bond risk factors) for the long short position - investing \$1 in the largest decile and shorting the smallest decile within each ILL decile as well as the ILL Avg. group.

Table IA.3 shows that the size anomaly is no longer detected after controlling for illiquidity. Focusing on the last columns of ILL Avg where the portfolio is illiquidity-neutral, we can see that the risk-adjusted return for the (B)-(S) portfolio using either FF 3 + two bond risk factors or FF 5 + two bond risk factors, is small (-0.12% and -0.01%) and statistically insignificant (t-stats -0.68 and -0.07). Furthermore, the size anomaly is no longer detected in any single ILL decile, as the risk-adjusted returns shown in the table are not statistically distinguishable from 0. This confirms our hypothesis that illiquidity is a major factor behind the bank size anomaly.

Table IA.4 presents the time-series mean returns of the average control variable portfolio (e.g., ILL Avg.) within each decile of size for portfolios using that specific control variable as the first sorting variable. The columns labeled R, FF3, and FF5 also represent the mean returns and the risk-adjusted returns using FF3 + two bond risk factors and FF5 + two bond risk factors for the long-short portfolios. This is analogous to the last column of Table IA.3 labeled ILL Avg, and we also repeat the result in the first row of Table IA.4 for using ILL as the first sorting criterion. Notably, after controlling for some specific firm characteristics, the control variable average Big-Small portfolio's raw return is usually positive. This aligns with our previous summary statistics in Table 3 that big banks tend to have higher monthly raw returns than small banks (1.07% v.s. 0.90%). In the first three rows, we use different measures of liquidity risk, including Amihud's (2002) illiquidity, turnover ratio, and the bid-ask spread. Then, we test other different variables, including price delay, book-to-market ratio, 12-month momentum, market beta, and idiosyncratic risk related to the market and the return volatility. The results demonstrate that the size anomaly persists when using each control variable except for two liquidity measures (ILL and TURN) as the first sort variable.

The results of the bivariate portfolio analyses verify our previous hypotheses that illiquidity risk plays a crucial role in size anomaly in the banking industry. We have shown that liquidity factors can resolve the size anomaly in BHCs while it is still unclear to investors what is the main driver of illiquidity.

IA.1.1.2 Using Proposed Complexity Factor to Solve for Size Anomaly in BHCs

As the previous analysis sets the stage to evaluate the role of complexity in the cross-section, we next explore the capability of the complexity factor we propose in resolving the size anomaly. We form monthly bivariate portfolios first by individual banks' complexity ratings, followed by their total assets. Table IA.5 is similar to Table IA.3, where we replace illiquidity with our proposed complexity measure as the primary sorting criterion. Concentrating on the final columns of COM Avg, where the portfolio is neutral in terms of complexity, we observe that the risk-adjusted return of the (B)-(S) portfolio, when employing FF 5 + two bond risk factors, the risk-adjusted return is statistically insignificant. However, when applying FF 3 + two bond risk factors, the risk-adjusted return remains statistically significant. However, its magnitude diminishes from -0.55% (noted in the risk-adjusted return from the univariate portfolio analysis) to -0.45%. Though for the overall complexity-neutral portfolio, the monthly risk-adjusted return of -0.45% is still economically and statistically significant, it is evident that the size anomaly is mitigated for stocks classified within the high complexity groups (From COM 6 to COM 10) by the proposed measure. This suggests that the investor may have recognized and priced the complexity factor in these stocks. However, for stocks with low complexity, a significant size anomaly persists. For the two lowest complexity groups, the risk-adjusted alphas for (B) - (S) reach approximately 12% annually. This observation implies that for less complex banks, an unidentified factor, potentially priced by investors, remains. This factor could also have a strong correlation with the illiquidity factor.

We also run a panel regression on observations grouped by complexity as a robustness test. At the end of each month, we cluster the data by their complexity factor into three categories. Subsequently, we run a panel regression with a year-fixed effect for each of these groups. Table IA.6 shows the results. The regression results reveal that only in the group with the highest level of complexity is there a positive and statistically significant correlation. Furthermore, we conduct an independent bivariate sorting of the data into nine groups based on total assets and complexity and report the regression results in Table IA.7. The group comprising the smallest yet most complex banks exhibits the strongest positive correlation between complexity and next-period return. These results are consistent with earlier bivariate portfolio analysis, which indicates that within the high complexity group, investors do price the complexity risk factor ex-ante and demand a corresponding risk premium.

IA.1.2 Effect of Complexity on Information Process

We use our complexity measure to validate the hypothesis proposed in Cohen and Lou (2012) that investors' limited information processing capacity can cause delays in information revelation in asset prices. If investors need a longer time to process information from high-complexity firms, the previous month's returns from the low-complexity portfolio should have some predictive power. As we focus on BHCs, all firms in our sample should be subject to the same information shocks. The only difference is that more complicated firms may require longer information processing time and, therefore, cause delays in information revelation in asset prices. Being subject to the same industry shocks, the stock prices of those low-complexity firms should be updated first and capable of forecasting future price adjustments in response to the same information shocks of their peers in the high-complexity group.

We adopt the portfolio formation method to investigate the impact of complexity on investors' information processing and its following stock return effect. Firms are sorted into quintiles at the end of each month according to their complexity scores, and the monthly size-weighted average stock returns for the bottom quintile are calculated. To empirically test the hypothesis, we conduct a panel regression analysis of next-period stock returns for high-complexity firms on the previous month's portfolio returns of the low-complexity group. We estimate the following equations using pooled stock-month data for all BHCs in the top complexity quintile each month:

$$\begin{aligned} ret_{i,t+1} =& \alpha + \beta_1 ret\_lc_t + \beta_2 Size_{i,t} + \beta_3 bm_{i,t} + \beta_4 PRC_{i,t} \\ & + \beta_5 ill_{i,t} + \beta_6 beta_{i,t} + \beta_7 ret_{i,t} + \epsilon_{i,t} \end{aligned} \tag{IA2}$$

Table IA.8 reports the regression results. The dependent variable is the next period stock return (ret) for BHC stocks categorized in the high-complexity group. The independent variables have a one-period lag. 'ret_lc' is the variable of interest. It is calculated as the size-weighted average of stock returns from the previous month for all BHCs classified under the low complexity group. 'Size' is calculated as the logarithmic value of the market capitalization; 'BM' is the book-to-market ratio; PRC is the monthly end price; 'ill' is

Amihud's (2002) illiquidity; beta is the CAPM beta estimate for each firm during the year; and 'ret' is current period stock return. Our analysis comprises four regression models: the first pooled regression serves as a baseline without controlling for fixed effects, the second controls for fixed year effects, the third for fixed entity effects, and the fourth combines both year and firm fixed effects. Standard errors are enclosed in parentheses for all reported values. The regression results show that the coefficients for the return of the low-complexity portfolio from the previous month are positive for all models, which supports the hypothesis that the stock returns of the low-complexity group can predict the future returns of high-complexity stocks. Furthermore, we also observe the short-term reversal phenomenon as the coefficients of the previous month's returns are consistently negative.

Table IA.1: **Univariate Sorting for Size Portfolios**

The table presents the results of OLS regression analyses, examining the relationship between monthly value-weighted excess returns and size-sorted portfolios of U.S. bank holding companies (BHCs) in Panel A, their matched firms in Panel B, and a broader control group of all non-financial firms in Panel C. The regression models include three Fama and French (1993) stock risk factors and two bond risk factors. Specifically, Panel A uses the same econometric methodology employed by Gandhi and Lustig (2015) and replicates their findings of the size anomaly. The sample period consists of 407 months from February 1986 to December 2019. In all cases, the reported alphas, along with their corresponding standard errors, are on a monthly basis and expressed as percentages. The standard errors are adjusted for heteroskedasticity and autocorrelation using the Newey and West's (1987) method with three lags.

	(S)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(B)	(B)-(S)
					Panel A.BHC Size Portfolios Return Regression						
const	0.41**	0.33	0.44**	0.35**	0.41**	0.23	0.16	0.07	-0.06	-0.29	-0.55***
	(0.16)	(0.20)	(0.20)	(0.17)	(0.19)	(0.19)	(0.21)	(0.20)	(0.20)	(0.19)	(0.17)
Mkt-Rf	0.41***	0.51***	0.53***	0.58***	0.62***	0.78***	0.90***	0.97***	1.04***	1.51***	0.81***
	(0.04)	(0.06)	(0.07)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.07)	(0.06)
Smb	0.34***	0.31***	0.29***	0.36***	0.36***	0.50***	0.57***	0.62***	0.52***	-0.07	-0.10*
	(0.07)	(0.07)	(0.08)	(0.06)	(0.06)	(0.06)	(0.07)	(0.08)	(0.07)	(0.07)	(0.06)
Hml	0.29***	0.37***	0.35***	0.49***	0.49***	0.67***	0.73***	0.87***	0.89***	0.98***	0.61***
	(0.05)	(0.06)	(0.08)	(0.07)	(0.08)	(0.07)	(0.08)	(0.08)	(0.07)	(0.09)	(0.08)
Ltg	-0.06	-0.16*	-0.11	-0.03	-0.01	0.02	0.02	0.02	0.09	0.18*	0.25***
	(0.07)	(0.09)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.10)	(0.10)	(0.09)
Crd	-0.03	-0.12	-0.03	-0.06	0.04	0.40**	0.59***	0.74***	0.95***	0.19	0.64***
	(0.13)	(0.16)	(0.18)	(0.16)	(0.20)	(0.17)	(0.19)	(0.22)	(0.27)	(0.25)	(0.23)
R-squared	0.41	0.44	0.42	0.51	0.53	0.63	0.64	0.67	0.66	0.73	0.52
N	407	407	407	407	407	407	407	407	407	407	407
					Panel B.Matched Firms Size Portfolios Return Regression						
const	0.22	0.05	0.32*	0.23	0.53***	0.33**	0.22*	-0.06	-0.01	0.06	-0.11
	(0.21)	(0.16)	(0.17)	(0.16)	(0.15)	(0.15)	(0.13)	(0.11)	(0.12)	(0.15)	(0.17)
Mkt-Rf	0.66***	0.85***	0.80***	0.86***	0.87***	1.01***	0.98***	1.00***	1.10***	0.83***	0.22***
	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
Smb	0.74***	0.73***	0.90***	0.75***	0.84***	0.74***	0.78***	0.61***	0.44***	-0.09	-0.56***
	(0.09)	(0.07)	(0.05)	(0.11)	(0.05)	(0.06)	(0.08)	(0.06)	(0.05)	(0.06)	(0.07)
Hml	-0.22***	-0.09	-0.32***	-0.02	0.01	0.06	0.18***	0.09	0.20***	0.32***	0.41***
	(0.07)	(0.07)	(0.06)	(0.08)	(0.06)	(0.07)	(0.07)	(0.06)	(0.06)	(0.07)	(0.07)
Ltg	-0.07	-0.13*	0.01	0.05	0.09	0.07	0.01	0.13**	0.18***	-0.07	0.16**
	(0.08)	(0.07)	(0.07)	(0.07)	(0.06)	(0.08)	(0.07)	(0.05)	(0.06)	(0.06)	(0.07)
Crd	-0.27	-0.14	-0.08	-0.39***	-0.07	-0.18	-0.24	0.01	0.04	-0.04	0.20
	(0.19)	(0.21)	(0.17)	(0.15)	(0.17)	(0.18)	(0.20)	(0.12)	(0.19)	(0.15)	(0.16)
R-squared	0.47	0.67	0.72	0.72	0.75	0.75	0.79	0.80	0.82	0.65	0.27
N	407	407	407	407	407	407	407	407	407	407	407

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table IA.2: **Univariate Sorting for Size Portfolios Using FF 5 Factors and Illiquidity Risk Factor** The table presents the results of OLS regression analyses, examining the relationship between monthly value-weighted excess returns and size-sorted portfolios of U.S. bank holding companies (BHCs). The regression models include three Fama and French (1993) stock risk factors, two bond risk factors, profitability and investment factors from Fama and French (2015), and one illiquidity risk factor. The results are reported using the same econometric methodology employed by Gandhi and Lustig (2015). The sample period consists of 407 months, consistent with Table 2. In all cases, the reported alphas, along with their corresponding standard errors, are on a monthly basis and expressed as percentages. The standard errors are adjusted for heteroskedasticity and autocorrelation using the Newey and West (1986) method with three lags.

	(S)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(B)	(B)-(S)
Panel A.BHC Size Portfolios Return Regression with RMW and CMA											
const	0.32*	0.26	0.31	0.27	0.29	0.20	0.08	0.00	-0.14	0.05	-0.33*
	(0.17)	(0.20)	(0.20)	(0.18)	(0.19)	(0.18)	(0.19)	(0.19)	(0.19)	(0.20)	(0.18)
Mkt-Rf	0.44***	0.53***	0.56***	0.59***	0.65***	0.78***	0.91***	0.98***	1.05***	1.37***	0.73***
	(0.05)	(0.06)	(0.07)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)
Smb	0.38***	0.38***	0.39***	0.43***	0.46***	0.57***	0.68***	0.73***	0.62***	-0.13	-0.13**
	(0.06)	(0.07)	(0.07)	(0.06)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)	(0.09)	(0.06)
Hml	0.21***	0.36***	0.32***	0.48***	0.46***	0.72***	0.77***	0.92***	0.93***	1.38***	0.87***
	(0.07)	(0.08)	(0.09)	(0.09)	(0.09)	(0.08)	(0.09)	(0.10)	(0.10)	(0.13)	(0.12)
Ltg	-0.05	-0.14*	-0.08	-0.01	0.02	0.05	0.05	0.06	0.12	0.16*	0.24***
	(0.07)	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.09)	(0.09)	(0.08)
Crd	-0.04	-0.15	-0.07	-0.10	-0.01	0.36**	0.53***	0.68***	0.89***	0.16	0.62***
	(0.13)	(0.16)	(0.17)	(0.16)	(0.18)	(0.17)	(0.19)	(0.22)	(0.27)	(0.24)	(0.23)
Rmw	0.15*	0.21**	0.34***	0.24***	0.33***	0.22**	0.35***	0.34***	0.32***	-0.30***	-0.17**
	(0.08)	(0.09)	(0.09)	(0.09)	(0.09)	(0.10)	(0.10)	(0.11)	(0.10)	(0.10)	(0.08)
Cma	0.10	-0.09	-0.08	-0.10	-0.11	-0.22*	-0.26*	-0.29*	-0.24*	-0.81***	-0.53***
	(0.10)	(0.13)	(0.14)	(0.13)	(0.12)	(0.12)	(0.13)	(0.15)	(0.14)	(0.17)	(0.14)
R-squared	0.42	0.45	0.45	0.52	0.55	0.64	0.66	0.69	0.67	0.76	0.54
N	407	407	407	407	407	407	407	407	407	407	407
Panel B.BHC Size Portfolios Return Regression with RMW, CMA and Liq											
const	0.22	0.14	0.21	0.18	0.21	0.11	0.04	-0.01	-0.14	0.02	-0.24
	(0.16)	(0.19)	(0.19)	(0.18)	(0.19)	(0.17)	(0.19)	(0.19)	(0.19)	(0.20)	(0.17)
Mkt-Rf	0.60***	0.71***	0.72***	0.74***	0.78***	0.91***	0.97***	1.01***	1.06***	1.41***	0.58***
	(0.05)	(0.07)	(0.07)	(0.05)	(0.05)	(0.04)	(0.06)	(0.06)	(0.06)	(0.07)	(0.07)
Smb	0.24***	0.22***	0.26***	0.31***	0.35***	0.46***	0.62***	0.70***	0.61***	-0.16*	-0.01
	(0.06)	(0.07)	(0.08)	(0.06)	(0.06)	(0.07)	(0.07)	(0.08)	(0.08)	(0.10)	(0.07)
Hml	0.23***	0.39***	0.34***	0.50***	0.48***	0.73***	0.77***	0.92***	0.93***	1.39***	0.85***
	(0.07)	(0.08)	(0.09)	(0.08)	(0.09)	(0.08)	(0.09)	(0.10)	(0.10)	(0.13)	(0.11)
Ltg	-0.01	-0.10	-0.04	0.03	0.05	0.08	0.07	0.07	0.12	0.17*	0.20**
	(0.07)	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.09)	(0.09)	(0.09)
Crd	-0.03	-0.13	-0.06	-0.09	0.00	0.37**	0.53***	0.68***	0.89***	0.17	0.61**
	(0.15)	(0.14)	(0.17)	(0.14)	(0.18)	(0.16)	(0.18)	(0.22)	(0.27)	(0.23)	(0.26)
Rmw	0.16*	0.21**	0.34***	0.24***	0.33***	0.22**	0.35***	0.34***	0.32***	-0.30***	-0.17*
	(0.09)	(0.09)	(0.10)	(0.09)	(0.09)	(0.10)	(0.10)	(0.11)	(0.10)	(0.10)	(0.09)
Cma	-0.04	-0.26*	-0.22	-0.23*	-0.23*	-0.34***	-0.32**	-0.31**	-0.25*	-0.85***	-0.40***
	(0.09)	(0.13)	(0.13)	(0.12)	(0.12)	(0.12)	(0.14)	(0.14)	(0.13)	(0.17)	(0.13)
Liq	0.52***	0.61***	0.50***	0.46***	0.43***	0.43***	0.20**	0.08	0.02	0.14	-0.49***
	(0.08)	(0.09)	(0.10)	(0.09)	(0.08)	(0.09)	(0.09)	(0.11)	(0.12)	(0.13)	(0.10)
R-squared	0.48	0.52	0.50	0.56	0.58	0.67	0.66	0.69	0.67	0.76	0.57
N	407	407	407	407	407	407	407	407	407	407	407

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table IA.3: **Bivariate Portfolio Analyses - Control for ILL**

The table presents the findings derived from bivariate dependent-sort portfolio analyses on the relation between stock returns and size in terms of total assets, after controlling for Amihud's (2002) illiquidity. At the end of each month, all BHC stocks in the sample are sorted into decile groups based on an ascending order of the control variable. The sample date begins from January 1996 to December 2019 to ensure that there are more than 100 BHCs at the end of each month. The first part of the table illustrates the time-series average returns of assets-weighted portfolios, formed by sorting initially on liquidity and then based on total assets, using total assets as the weight. The last column labeled ILL Avg. presents the results for the average ill decile within the given size decile. In the second part, we show the mean return, risk-adjusted returns using different sets of factors (FF3 + two bond risk factors and FF 5 + two bond risk factors) for the long short position - investing $1 in the largest decile and shorting the smallest decile within each ILL decile as well as the ILL Avg. group. The numbers in parentheses are t-statistics, adjusted following Newey and West (1987) using three lags, testing the null hypothesis that the mean monthly or risk-adjusted return equals zero.

	ILL 1	ILL 2	ILL 3	ILL 4	ILL 5	ILL 6	ILL 7	ILL 8	ILL 9	ILL 10	ILL Avg
Small	0.86	1.08	1.03	0.78	0.87	0.68	1.11	1.08	0.81	0.87	0.92
Size 2	1.01	1.28	1.16	1.01	0.81	0.72	1.06	0.97	1.02	0.67	0.97
Size 3	0.78	1.20	0.88	1.02	1.28	1.00	0.87	1.40	0.94	1.31	1.07
Size 4	0.84	1.01	1.08	0.84	0.99	0.94	1.32	1.18	1.36	0.97	1.05
Size 5	0.92	1.22	1.03	1.05	1.00	0.84	0.94	1.11	0.79	0.61	0.95
Size 6	0.96	1.26	1.27	0.71	1.29	1.28	1.07	0.85	0.57	0.75	1.00
Size 7	0.71	0.94	0.95	0.87	1.31	1.23	1.48	1.25	0.95	0.84	1.05
Size 8	1.18	0.92	1.02	1.25	0.91	0.91	1.02	1.17	1.11	1.19	1.07
Size 9	1.14	1.20	1.02	1.06	0.68	0.90	1.26	0.87	1.37	0.72	1.02
Big	1.07	1.00	0.94	0.93	1.29	1.14	1.15	1.03	0.73	0.97	1.03
(B) - (S) Portfolios											
Mean Return	0.21	-0.09	-0.08	0.14	0.42	0.46	0.04	-0.04	-0.08	0.09	0.11
	(0.49)	(-0.23)	(-0.23)	(0.40)	(1.17)	(1.36)	(0.09)	(-0.12)	(-0.21)	(0.26)	(0.65)
FF3 + Bonds α	-0.26	-0.39	-0.27	-0.04	0.18	0.19	-0.14	-0.32	-0.19	0.06	**-0.12**
	(-0.77)	(-0.96)	(-0.75)	(-0.12)	(0.51)	(0.57)	(-0.41)	(-0.94)	(-0.51)	(0.17)	**(-0.68)**
FF5 + Bonds α	0.50	-0.32	-0.33	0.09	0.37	0.17	-0.03	-0.38	-0.26	0.07	**-0.01**
	(1.52)	(-0.80)	(-0.89)	(0.29)	(1.05)	(0.53)	(-0.09)	(-1.06)	(-0.68)	(0.18)	**(-0.07)**

Table IA.4: **Bivariate Portfolio Analyses - Other Control Variables**

The table presents the findings derived from bivariate dependent-sort portfolio analyses on the relation between stock returns and size in terms of total assets, after controlling for other factors besides Amihud's (2002) illiquidity as displayed in Table IA.3. This table shows the time-series mean returns of the average control variable portfolio (e.g. ILL Avg.) within each decile of size for portfolios using that specific control variable as the first sorting variable. The columns labeled R, FF3, and FF5 also represent the mean returns and the risk-adjusted returns using FF3 + two bond risk factors and FF 5 + two bond risk factors for the long-short portfolios. This is analogous to the last column of Table IA.3 labeled ILL Avg and we also repeat the result in the first row of the table. The numbers in parentheses are t-statistics, adjusted following Newey and West (1987) using three lags, testing the null hypothesis that the mean monthly or risk-adjusted return equals zero.

	Small Size 1	Size 2	Size 3	Size 4	Size 5	Size 6	Size 7	Size 8	Size 9	Big Size 10	Big-Small Size Portfolios R	FF3 α	FF5 α
ILL	0.92	0.97	1.07	1.05	0.95	1.00	1.05	1.07	1.02	1.03	0.11	−0.12	−0.01
											(0.65)	(−0.68)	(−0.07)
TURN	0.86	1.08	1.00	1.00	1.00	1.04	1.00	1.06	1.04	1.03	0.18	−0.11	−0.11
											(0.97)	(−0.83)	(−0.77)
Bid-Ask Spread	0.95	0.92	1.07	1.03	1.00	1.06	0.96	0.92	1.09	0.99	0.04	−0.50	−0.39
											(0.16)	(−2.59)	(−1.95)
Price Delay	0.90	1.05	1.06	1.03	1.05	1.04	0.94	1.13	0.93	0.96	0.06	−0.43	−0.27
											(0.24)	(−2.49)	(−1.36)
BM	0.80	0.94	1.05	1.11	0.95	1.14	0.99	1.03	0.98	1.17	0.36	−0.22	−0.17
											(1.23)	(−1.05)	(−0.53)
MOM12M	0.90	1.02	0.98	1.04	1.14	1.01	0.99	1.14	0.90	0.98	0.08	−0.50	−0.40
											(0.25)	(−2.76)	(−2.02)
BETA	0.88	1.08	1.09	1.00	0.97	1.10	1.20	1.06	0.88	0.86	−0.02	−0.26	−0.22
											(−0.10)	(−2.04)	(−1.61)
IDIOVOL	0.92	1.09	0.87	0.98	1.09	1.08	0.87	1.09	1.08	1.01	0.09	−0.54	−0.43
											(0.29)	(−2.75)	(−2.02)
RETVOL	0.94	1.06	1.10	0.95	1.04	1.11	0.90	1.01	0.98	0.90	−0.04	−0.61	−0.46
											(−0.14)	(−3.35)	(−2.41)

Table IA.5: **Bivariate Portfolio Analyses on Complexity**

The table presents the findings derived from bivariate dependent-sort portfolio analyses on the relation between stock returns and size in terms of total assets, after controlling for our newly proposed complexity measure. At the end of each month, all BHC stocks in the sample are sorted into decile groups based on an ascending order of the control variable. The sample date begins from January 1996 to December 2019 to ensure that there are more than 100 BHCs at the end of each month. The first part of the table illustrates the time-series average returns of assets-weighted portfolios, formed by sorting initially on complexity and then based on total assets, using total assets as the weight. The last column labeled COM Avg. presents the results for the average COM decile within the given size decile. In the second part, we show the mean return, risk-adjusted returns using different sets of factors (FF3 + two bond risk factors and FF 5 + two bond risk factors) for the long short position - investing $1 in the largest decile and shorting the smallest decile within each COM decile as well as the COM Avg. group. The numbers in parentheses are t-statistics, adjusted following Newey and West (1987) using three lags, testing the null hypothesis that the mean monthly or risk-adjusted return equals zero.

	COM 1	COM 2	COM 3	COM 4	COM 5	COM 6	COM 7	COM 8	COM 9	COM 10	COM Avg
Small	0.81	1.01	1.03	1.39	1.01	0.33	1.04	0.65	0.47	0.55	0.83
Size 2	1.20	1.35	1.10	1.29	1.06	1.27	0.94	0.92	0.68	0.92	1.07
Size 3	0.60	0.92	1.30	0.96	0.58	1.80	0.86	0.93	1.27	1.28	1.05
Size 4	1.05	0.92	0.69	0.81	1.45	1.11	1.23	0.95	1.27	1.22	1.07
Size 5	0.20	1.08	0.90	1.03	1.33	1.73	1.14	0.98	1.15	1.53	1.11
Size 6	1.60	0.63	0.88	0.20	1.06	0.98	0.92	1.53	1.41	0.96	1.02
Size 7	0.99	0.87	0.87	1.28	0.94	0.85	1.70	1.28	1.01	1.48	1.13
Size 8	0.03	1.00	1.27	0.97	1.34	1.16	0.78	0.88	1.20	0.91	0.95
Size 9	0.57	0.75	0.70	0.82	1.23	1.13	0.61	0.63	1.06	1.16	0.87
Big	0.36	0.45	1.15	0.76	0.82	1.20	1.28	1.17	0.61	0.80	0.86
(B) - (S) Portfolios											
Mean Return	-0.46	-0.55	0.12	-0.63	-0.19	0.86	0.25	0.52	0.14	0.25	0.03
	(-0.86)	(-0.90)	(0.23)	(-1.20)	(-0.30)	(1.63)	(0.50)	(1.03)	(0.32)	(0.53)	(0.09)
FF3 + Bonds α	-1.02	-1.23	-0.34	-1.03	-0.99	0.34	-0.07	0.05	0.10	-0.36	-0.45
	(-2.42)	(-2.39)	(-0.77)	(-2.53)	(-1.64)	(0.75)	(-0.16)	(0.12)	(0.24)	(-1.00)	(-2.13)
FF5 + Bonds α	-0.97	-0.72	-0.19	-0.78	-0.33	0.44	0.17	0.41	0.24	-0.15	-0.19
	(-2.24)	(-1.40)	(-0.43)	(-1.87)	(-0.55)	(0.89)	(0.37)	(0.99)	(0.51)	(-0.39)	(-0.80)

Table IA.6: **Panel regressions – Univariate Sort for Next Period Return on Complexity.**
The table presents the outcomes of estimating the following equation using pooled data:
$ret_{i,t+1} = \alpha + \beta_1 COM_{i,t} + \beta_2 Size_{i,t} + \beta_3 bm_{i,t} + \beta_4 PRC_{i,t} + \beta_5 illi,t + \beta_6 beta_{i,t} + \epsilon_{i,m}$
The dependent variable is the next period stock Sharpe ratio (SR). The independent variables have a one-period lag. Complexity and Opacity are two variables of interest. 'Size' is calculated as the logarithmic value of the market cap. 'BM' is the book-to-market ratio. 'PRC' is the monthly end price. 'ill' is Amihud's (2002) measure of illiquidity calculated as the ratio of the absolute value of the monthly return scaled by the monthly volume. 'beta' is the CAPM beta estimate for each firm during the year. We control for fixed year effects. Standard errors are enclosed in parentheses for all reported values.

	Next Period Return		
	Low Complexity	**Mid Complexity**	**High Complexity**
Complexity	0.0014	0.0008	0.0015***
	(0.0010)	(0.0012)	(0.0004)
Controls	Yes	Yes	Yes
Year Fixed Effects	✓	✓	✓
No. Observations	25032	24801	24927
R-squared	0.0008	0.0018	0.0017

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table IA.7: **Panel regressions – Double Sorting for Next Period Return.**
The table presents the outcomes of estimating the following equation using pooled data:
$ret_{i,t+1} = \alpha + \beta_1 COM_{i,t} + \beta_2 Size_{i,t} + \beta_3 bm_{i,t} + \beta_4 PRC_{i,t} + \beta_5 illi, t + \beta_6 beta_{i,t} + \epsilon_{i,m}$
The dependent variable is the next period stock return (ret). The independent variables have a one-period lag. Complexity is the variable of interest. 'Size' is calculated as the logarithmic value of the market cap. 'BM' is the book-to-market ratio. 'PRC' is the monthly end price. 'ill' is Amihud's (2002) measure of illiquidity calculated as the ratio of the absolute value of the monthly return scaled by the monthly volume. 'beta' is the CAPM beta estimate for each firm during the year. We control for fixed year effects. Standard errors are enclosed in parentheses for all reported values.

Complexity Coefficient β_1			
complexity \ size	(S)	(2)	(B)
(L)	0.0027**	0.0031*	-0.0019
(2)	-0.0015	0.0023	0.0013
(H)	0.0030***	0.0006	0.0011*

Table IA.8: **Panel regressions – Complexity and Return Predictability**

The table presents the outcomes of estimating the following equation using pooled data:

$ret_{i,t+1} = \alpha + \beta_1 ret\_lc_t + \beta_2 Size_{i,t} + \beta_3 bm_{i,t} + \beta_4 PRC_{i,t} + \beta_5 ill_{i,t} + \beta_6 beta_{i,t} + \beta_7 ret_{i,t} + \epsilon_{i,t}$

The dependent variable is the next period stock return (ret) for BHC stocks categorized in the high-complexity group. The independent variables have a one-period lag. 'ret_lc' is the variable of interest. It is calculated as the size-weighted average of stock returns from the previous month for all BHCs classified under the low complexity group. 'Size' is calculated as the logarithmic value of the market cap. 'BM' is the book-to-market ratio. 'PRC' is the monthly end price. 'ill' is Amihud's (2002) measure of illiquidity calculated as the ratio of the absolute value of the monthly return scaled by the monthly volume. 'beta' is the CAPM beta estimate for each firm during the year. 'ret' is the firm's previous month stock return. We control for fixed entity effects. Standard errors are enclosed in parentheses for all reported values.

	Next Period Return			
	M1	**M2**	**M3**	**M4**
ret_lc	0.1572***	0.1485***	0.1097***	0.1124***
	(0.0126)	(0.0127)	(0.0131)	(0.0132)
ret	-0.1242***	-0.1310***	-0.1328***	-0.1279***
	(0.0138)	(0.0138)	(0.0139)	(0.0140)
Controls	Yes	Yes	Yes	Yes
Year Fixed Effects			✓	✓
Firm fixed Effects		✓		✓
No. Observations	13286	13286	13286	13286
R-squared	0.0221	0.0274	0.0162	0.0258

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$