Indian Institute of Management Calcutta
Working Paper Series
WPS No. 898/ June 2023

Banking with Overreaction

Arpit Kumar Parija*
Indian Institute of Management Calcutta
PhD, IIM Calcutta
Email: arpitkp16@iimcal.ac.in

Vipul Mathur
Assistant Professor, IIM Calcutta
Email: vipul@iimcal.ac.in

* Corresponding author

Indian Institute of Management Calcutta, Joka, D.H. Road, Kolkata 700104

URL: https://www.iimcal.ac.in/faculty/publications/working-papers/

Banking with Overreaction*

Arpit Kumar Parija Vipul Mathur†

ABSTRACT

Using the data on loan loss provisions, we analyze expectations of US banks and find evidence of departure from rational expectations. In particular, we find that banks overreact to actual losses incurred in the recent past. In good times, the presence of overreaction leads to neglect of risks, resulting in a rise in credit growth. This subsequently results in higher non-performing loans and lower return on assets for banks when the risks get realized in future. Additionally, the shareholders fail to adequately recognise the risky lending behavior of such banks, and earn predictably lower returns in subsequent years.

JEL classification: G12, G21, G41

Keywords: Loan Loss Provision, Overreaction, Credit Growth, Equity Return

*We thank Viral Acharya, Raman Kumar, Sudipto Basu, Wei Wang, Justin Vitanza, Srinivasan Rangan, Gregor Matvos, Abhiman Das, Indrajit Roy, Sara Yasar, participants at EFA 2020, EFMA 2021, IGIDR Macroeconomics and Finance Conference 2021, Delhi Winter School 2021 (The Econometric Society), and Centre for Advanced Financial Research and Learning (CAFRAL) for helpful comments. We also thank four anonymous referees for their insightful and helpful suggestions. We owe special thanks to Claudia Ruiz Ortega for featuring our work in the Finance Blog series of the World Bank. All errors remain our own. The views expressed here are those of the authors and not necessarily those of the institutions they are associated with.
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
†Arpit Kumar Parija: Indian Institute of Management Calcutta (email: arpitkp16@iimcal.ac.in); Vipul Mathur: Indian Institute of Management Calcutta (email: vipul@iimcal.ac.in).

Introduction

In recent years, researchers in finance and economics have attempted to establish credit expansion as a leading indicator of impending financial crises (Schularick and Taylor, 2012; Jordà et al., 2013; Greenwood et al., 2022). Excessive credit growth has also been found to result in future underperformance of bank stocks (Baron and Xiong, 2017; Fahlenbrach et al., 2017). Why does credit expansion precede such poor outcomes? According to the long-standing argument by Minsky (1977) and Kindleberger (1978), excessive credit growth can be attributed to overoptimistic beliefs leading to underestimation of risks. Several studies pertaining to the non-bank credit markets have indeed established the role of such beliefs in driving credit outcomes, tracing them to non-rational or biased expectations of the investors (Greenwood and Hanson, 2013; Greenwood and Shleifer, 2014; Gennaioli et al., 2016; López-Salido et al., 2017; Bordalo et al., 2018, 2019). However, studies on banks' expectations have been elusive. This paper aims to address this gap by analyzing banks' expectations and exploring the implications of such expectations on bank credit and its outcomes.

Our study seeks to investigate the formation of expectations in the banking sector through the accounting category of loan loss provisions. Accounting standards require banks to make provisions in anticipation of future loan losses[1]. While provisions are primarily designed to reflect bank managers' expectations of future credit losses, it is widely recognised that managerial discretion makes provisions prone to manipulation. Correspondingly, we extract the "non-discretionary component" of loan loss provisions to arrive at a cleaner measure of expectations[2]. We juxtapose this component against actual loan losses to arrive at a measure of forecast errors. Crucially, we find that forecast errors are predictably related to loan losses. Such predictability of errors in expectations stands in stark contrast to rational expectations, which postulates that forecast errors be unpredictable from actual losses. In particular, our findings indicate that after periods of a declining (increasing) trend in loan losses, banks tend to consistently underestimate (overestimate) losses on loan portfolios

[1] Before the Financial Accounting Standards Board (FASB) introduced a new credit loss accounting standard in June 2016, banks were required to set provisions based on what is referred to as an "incurred loss model" (Acharya and Ryan, 2016; Jiménez et al., 2017). Under this accounting standard, provisions at any point in time were made only for loans that had turned bad based on the information available up to that time. Based on the information on bad loans, banks would typically arrive at some expectations on the magnitude of loan losses that could happen in the future.

[2] See Beatty and Liao (2014) for a review of loan loss provision models and Basu et al. (2020) for more recent approaches.

for the immediate future. Further, using the diagnostic expectations framework of Bordalo et al. (2020), we establish that such systematic variation in the forecast error can be explained by overreaction in expectations. Periods of reduced (increased) actual losses on the loan portfolio tend to increase optimism (pessimism) and, therefore, likely lead the bank manager to underestimate (overestimate) future losses. Overreaction to current economic conditions delays provisioning in good times whereas it causes excessive provisioning in downturns, thus exacerbating the cyclicality of loan loss provisions and credit.

Using a panel of publicly listed U.S. banks from 1988:1-2019:4, we find the presence of overreaction in expectations both at the panel and bank-by-bank level. Interestingly, our estimates on the degree of overreaction are comparable to those reported in the credit and financial markets literature. In addition, we find that the extent of overreaction varies with the loan type. Heterogeneous loan categories, which require greater due diligence at a loan-by-loan level (Real Estate, Commercial and Industrial loans), tend to exhibit weaker overreaction compared to the homogeneous loan categories assessed at a portfolio level (Credit Card and Consumer loans). Since our analysis of banks' expectations is based on the non-discretionary component of loan loss provisions, our findings remain robust to concerns about managers' manipulation of loan loss provisions. Nonetheless, to ensure that our results are not driven by any strategic accrual manipulation, we consider several competing hypotheses that could account for the predictability of forecast errors and confirm that the observed predictability is likely due to overreaction in expectations rather than any rational response on behalf of the bank manager.

We then explore how overreaction in expectations impacts bank credit and related outcomes. By utilizing the predictable component of forecast error, we infer the overreaction component in loss provisions. This approach is both intuitive and micro-founded, drawing on the theory of diagnostic expectations (Bordalo et al., 2018, 2020). We then compare the overreaction component with the Federal Reserve's Survey of Senior Loan Officers on Bank Lending Practice and find that the overreaction component in provisions is correlated to lending standards: overreaction in good times is concomitant with relaxed lending standards and vice-versa. We explore the implication on both quantity and quality of credit. Our results indicate that a one standard deviation increment in overreaction over a three-year period is associated with 20.8 percentage points higher credit growth during the same period. Additionally, it also leads to a 0.91 percentage points increase in *NPA* and a 0.48 percentage points decline in *ROA* in the subsequent three-year period. In sum, overreaction leads to distorted risk assessment by banks and leads to a cyclical variation in credit quantity and quality. Next, we explore whether the distortion in risk-assessment by banks on account of overreaction gets appropriately discounted in the equity markets. For instance, if shareholders were to take due cognizance of the risky behavior of banks, they are likely to demand higher returns by immediately lowering stock prices. However, if they either share similar expectations as banks or fail to recognize the risks, it is likely to result in an overvaluation of bank stocks followed by an underperformance in future, when the ignored risks get realized. We find that a one standard deviation increase in the overreaction component of provisions predicts a 14.2 percentage points decline in the subsequent three-year-ahead returns, thereby pointing to a failure of the investors in recognising risks.

As our analysis builds on using provisions to study banks' expectations, we have paid particular attention to verifying our results' robustness and several econometric issues. First, as stated before, we use only the "non-discretionary" component of loan loss provisions as a measure of expectations. We estimate the non-discretionary component using multiple empirical specifications and compared their ability to detect provision manipulation using the procedure in Kothari et al. (2005) and Basu et al. (2020). Second, we perform several robustness tests to ensure our results on forecast error predictability are not driven by any rational response of banks. Further, when investigating the role of overreaction in driving credit growth and future performance, we substantiate our results with a variety of checks as employed elsewhere in the literature.

Related Work: In most modern accounts of the financial crisis, explanations for why banks take excessive risks during credit booms rely on rational incentives, such as those of institutional investors and boards of directors (Cheng et al., 2015; Hagendorff et al., 2021), coordinated risks to increase chances of bailout (Acharya and Yorulmazer, 2007), or reach for yield (Rajan, 2005). However, our study shows that non-rational expectations may be central to this behavior. In particular, the presence of overreaction in banks' expectations in good times may lead banks to neglect risks and expand credit, resulting in increased bad loans and lower returns on assets in the future. Our findings offer evidence of expectations channel - both at the panel and bank-by-bank level - that may complement incentive-based channels from the literature. To the best of our knowledge, our work is the first to construct a measure of non-rational expectations for banks, contributing to the vast literature that studies the role of sentiments in financial markets (López-Salido et al., 2017; Pflueger et al., 2020)[3]. In contrast to the market-based sentiment metrics, our measure is micro-founded along the lines of Bordalo et al. (2021) and allows us to evaluate quantitative differences from the rational counterpart. Consequently, our work contributes to the recent literature on behavioral theories of credit cycle (Greenwood and Hanson, 2013) in which extrapolative expectations of creditors play a central role in generating credit boom and bust. We extend this literature by empirically establishing that such extrapolative expectations are also pervasive in the banking sector.

While there are studies that have identified the role of supply-side factors in impacting bank credit, such as those mentioned by Rajan (1994) and Dell'Ariccia and Marquez (2006), our research highlights the role of banks' expectations. The literature on provisioning has largely focused on examining strategic accrual manipulation, while research on non-strategic concerns remains limited [4]. In a recent paper, Bischof et al. (2021) argue that managerial reporting incentives could be the reason for the delay in loss recognition ahead of the 2007–09 crisis. In contrast, our findings suggest that overreaction to the benign conditions during the pre-crisis period turned banks too optimistic and thereby, engage in riskier lending practices.

Our results on equity returns are closely related to Fahlenbrach et al. (2017) who show that as a bank registers higher credit growth, the investors tend to become overly optimistic. As a result, the stock becomes overvalued, leading to a fall in the realized return as the price corrects in future. Our work extends this analysis to the supply-side. We establish that higher credit growth is likely an outcome of overreaction in the expectations of banks: over-optimistic banks engage in excessive

[3] In a related context, Hribar et al. (2017) studies managerial sentiment and its effect on accrual estimates of banks using Duke CFO Global Business Outlook survey. They measure bank sentiments as the residuals obtained from regressing median survey responses on economic fundamentals. By construction, it fails to capture any systematic (under) overreaction, which we observe in our analysis.

[4] A few exceptions are Francis et al. (2005) which separates accruals into innate and discretionary components where innate accrual considers difficulties in estimating future outcomes due to uncertainty. In the same spirit, Hennes et al. (2008) studies the restatements of financial results that can arise due to irregularities (strategic) and unintentional errors (non-strategic).

risk-taking during good times. This momentarily drives up the credit growth and stock prices but at the expense of credit quality. As those risks materialize in future, the return on assets and equity fall.

The article layout is as follows. In section 1 and 2, we present the empirical framework to analyse banks' expectations and provide evidence of overreaction. Section 3 introduces a stylized model and provide theoretical underpinnings for our empirical results. In section 4, we measure the extent of overreaction in banks' expectations and show its implications for credit growth and financial performance of banks in subsequent periods. Section 5 concludes. In the appendix, we perform all details of robustness exercises and discuss why our results on banks' expectations are robust to alternative confounds.

1. Some Preliminary Evidence on Banks' Expectations

We begin this section with a discussion on our proposed measure of expectations for the banking sector that is derived from the data on loan loss provisions using recent empirical approaches. Using this measure, we then discuss the methodology to determine whether expectations are rational or not, in line with the strategy used by Coibion and Gorodnichenko (2015), Bordalo et al. (2018, 2020), and Kohlhas and Walther (2021)[5]. We conclude the section by noting that the expectations of banks about future loan losses display a departure from the rational expectations framework and that our findings stand robust to alternative considerations.

1.1. Measure of Expectations

Before the Financial Accounting Standards Board (FASB) introduced a new credit loss accounting standard in June 2016[6], banks were required to set provisions based on what is referred to as an "incurred loss model" (Acharya and Ryan, 2016; Jiménez et al., 2017). Under this accounting standard, provisions at any point in time were made only for loans that had turned bad[7] depending

[5] While the authors can directly use survey data on expectations, we adapt provisions for expected loan losses to analyze banks' expectations.

[6] The new accounting standard introduced "current expected credit losses methodology (CECL)" for estimating provisions for credit losses. The standard is effective from 2020 for SEC registrant banks, and for all others, it takes effect in 2023. We restrict our analysis of banks' expectations using loan loss provisions to correspond to the old accounting standard.

[7] There is a well defined criteria to classify a loan as bad loan based on considerations such as significant financial difficulty of the borrower or a delay in principal or interest payments Gebhardt, 2016.

on the information available up to that time. Based on the information on bad loans, banks would typically arrive at some expectations on the magnitude of loan losses that could happen in the future. Banks were allowed sufficient discretion to arrive at this assessment. This renders the eventual loan loss provisions to potentially deviate from an objective assessment of expected loan losses, and often prone to manipulation.

To evaluate such deviations, the literature has attempted to empirically decompose the total provisions into two components: non-discretionary and discretionary component (see Beaver and Engel (1996), Kiridaran et al. (2004), and Beatty and Liao (2014)). The non-discretionary component is supposed to capture the provisions made against anticipated credit risks or expected loan losses, conditional on the currently available information on bad loans. On the other hand, the discretionary component is meant to largely capture managerial discretion towards meeting other objectives such as the desire for income smoothing (Ahmed et al., 1999), signaling financial strength (Kiridaran et al., 2004), or managing capital constraints (Calomiris and Wilson, 2004) among others.

In contrast to the discretionary component, the non-discretionary component of provision is - by definition - robust to such alternative motives and is primarily driven by expectations on immediate future loan losses (see (Ahmed et al., 1999)). Arguably so, it serves as a natural and direct measure of expectations. Therefore, we base our measure for banks' expectations on the non-discretionary component of loan loss provisions. We rely on well-established empirical methods in the literature to deduce the non-discretionary component from gross loan loss provisions. We describe our empirical approach and the related robustness exercises in detail below.

While the literature on isolating the non-discretionary component from loan loss provisions recognises multitude of approaches, there are certain common grounds (see Beatty and Liao (2014) for an exhaustive review). The key idea in identifying the non-discretionary component relies on projecting loan loss provisions against a set of explanatory variables comprising bank-specific and macroeconomic factors. More recent approaches, such as Basu et al. (2020), refine on this commonly accepted set of variables. For our main text, we keep track of one model which maps into the preferred model of Basu et al. (2020). In Appendix B.1, we consider different variants of the model and find that our key results remain robust to these extensions[8].

Loan Loss Provision Regression Model ***(Model 1)***:

Our preferred model largely mirrors that of Basu et al. (2020)and includes measures of changes in non-performing loans (both contemporaneous and past change), loan growth (since provisions change with amount of credit) and bank size (to account for different levels of regulatory scrutiny applied to banks based on their assets). Following Basu et al. (2020), we also control for net loan charge-offs. To account for the asymmetric effect of net loan charge-offs, we also allow for separate coefficients to distinguish positive and negative changes in non-performing loans through a dummy variable[9]. Specifically, we estimate the following:

$$llp_{it} = \gamma_0 + \gamma_1 \Delta npa_{it} + \gamma_2 D\Delta npa_{it} * \Delta npa_{it} + \gamma_3 \Delta npa_{it-1} + \gamma_4 \Delta npa_{it-2} + \gamma_5 Size_{it-1} + \gamma_6 \Delta Loan_{it} + \gamma_7 nco_{it} + \delta_t + \zeta_i + \eta_{it}$$

where $D\Delta npa_{it}$ is a dummy variable equal to 1 if $\Delta npa_{it} < 0$ and zero otherwise.

$llp_{it} = \frac{Loan\ Loss\ Provision_{it}}{Loan_{it-1}}$, $nco_{it} = \frac{Net\ Charge-offs_{it}}{Loan_{it-1}}$, $npa_{it} = \frac{Loans\ past\ due\ 90\ days_{it} + Non-accrual\ Loans_{it}}{Loan_{it-1}}$, $Size_{it} = Log(Assets_{it-1})$, and $\Delta Loan_{it} = \frac{Loan_{it} - Loan_{it-1}}{Loan_{it-1}}$.

We deduce the non-discretionary component $(ndllp_{it})$ as the fitted values from above regression. In the following subsection, we propose the second part of our empirical exercise, which explores the systematic deviation of expected loss from actual loss.

8 In our preferred model, we have not considered any lead variables (such as NPA_{t+1}). It can be argued that banks do not observe $t+1$ NPA while forming expectations about future loan losses and lack perfect foresight. Nonetheless, we include future non-performing loans and show that our results remain robust. Also we have included bank fixed effects in the models to control for bank-specific traits unrelated to managerial discretion. Our results remain qualitatively unchanged if we exclude bank fixed effects. Our results also remain robust to the use of macroeconomic variables à la Beatty and Liao (2014) in place of time fixed effect. For purely supplementary purposes, we also conduct a horse-race across the all models, a-la the the criteria employed in Kothari et al. (2005) and Basu et al. (2020) and consistent with our selection rationale, we find our preferred model to be the superior of the lot (see Appendix B.2).

9 As mentioned in Basu et al. (2020), not incorporating the effects of net charge-offs makes non-performing loan change a poor measure of the true change in loan portfolio quality. For instance, an increase in net charge-offs can cause non-performing loans to mechanically fall even though there might be a deterioration in loan portfolio quality since the last period. The non-discretionary component of loan loss provision would then understate the expectations of banks about future loan losses.

1.2. Are Expectations Rational?

A substantial body of literature has explored a few popular class of expectations formation models[10]. Of these models, the full information rational expectations model (FIRE) is typically the standard specification in economic modeling. Agents in this setting know the true data generating process and its parameters while forming optimal forecasts. The test of such forecasts against the benchmark of FIRE boils down to assessing whether forecast errors can be predicted using information available at the time the forecast is made. A growing body of research that has tested forecasts in this manner has found that expectations often exhibit significant departure from rational expectations (Greenwood and Shleifer, 2014; Gennaioli et al., 2016; Bordalo et al., 2019, 2020, 2021). The literature has further proposed several theoretical frameworks to explain this departure. Of these frameworks, the diagnostic expectations model of Bordalo et al. (2018) has emerged as a leading alternative. An important feature of this methodology is that it is immune to Lucas critique. Accordingly, we rely on Bordalo et al. (2018)for testing the departures from FIRE.

We construct the forecast error as the difference between actual loss and expected loss, where expected loss is constructed as per the discussion in Section 1.1. For actual loss, we rely on the accounting category of net-charge offs which is simply the dollar amount representing the difference between gross charge-offs and any subsequent recoveries of a delinquent debt. We denote the net charge-offs and the non-discretionary component of loan loss provisions as nco and $ndllp$ respectively[11]. Resultantly,

$$Forecast\ Error_{iT+1,T} = Actual\ Loss_{iT+1} - E_T Actual\ Loss_{iT+1}$$

where $Actual\ Loss_{iT+1} = \frac{1}{4}\sum_{j=t+1}^{j=t+4} nco_{ij}$, and $E_T Actual\ Loss_{iT+1} = \frac{1}{4}\sum_{j=t-3}^{j=t} ndllp_{ij}$.

Under rational expectations, the constructed forecast errors on loan losses should be unpredictable from information available to banks when forecasts are made, such as past data on loan losses. Using

[10] For instance, adaptive expectations (Cagan, 1956); extrapolative expectations (Barberis et al., 2015; Hirshleifer et al., 2015); noisy information (Woodford, 2003).

[11] To capture the potential lag between the time that a bank recognizes future loan losses and the realization of actual losses, we follow Bordalo et al. (2018) and accordingly, the net charge-offs are averaged over quarter $t+1$ to $t+4$ and the expected losses are averaged over quarter $t-3$ to t. Besides, in a related context, **drechsler2021empty citation** mentions that it is necessary to adjust for seasonality due to the way loss provisions and other items are reported at the end of the year. We get qualitatively similar results if we compare the loan loss provisions with future net charge-offs without averaging as done in Hribar et al. (2017).

loan losses experienced by banks in the recent past (net charge-offs averaged over quarter $t-3$ to quarter t), the predictability of forecast errors is then assessed by estimating the following empirical specification,

$$Forecast\ Error_{iT+1,T} = \beta_0 + \beta_1 Actual\ Loss_{iT} + \zeta_i + \delta_t + \varepsilon_{iT+1,T} \tag{1}$$

where $Actual\ Loss_{iT} = \frac{1}{4}\sum_{j=t-3}^{j=t} nco_{ij}$. ζ_i and δ_t denotes bank fixed effect and time fixed effect respectively[12].

There is one more dimension to these tests: these tests can be conducted either by pooling all the banks in a panel setting, or at an individual bank level. As noted by Bordalo et al. (2020), the bank-by-bank specification has three main advantages. First, it does not impose the restriction of a common coefficient β_1. Second, it controls for bank-level heterogeneity in the extent of bias that can exist due to different priors. Third, it allows us to understand the interplay of overreaction under different persistence of the loan loss process at the bank level (Afrouzi et al. (2021)). On the flip side, the bank-by-bank analysis imposes a restriction on the number of unique bank-level observations, which can compromise the statistical power of the test. Our bank-by-bank regressions take the following form:

$$Forecast\ Error_{iT+1,T} = \beta_0^i + \beta_1^i Actual\ Loss_{iT} + \nu_{iT+1,T} \tag{2}$$

The above specification yields a distribution of individual coefficients. We can then take the median coefficient as indicative of whether the majority of banks over- or underreacts. For reliable estimation of β_1^i, we restrict the number of observations for each bank to at least 50.

Under the assumption of rational expectations, the forecast error should be unpredictable (or, $\beta_1 = 0$). On the other hand, a negative (or, positive) coefficient or $\beta_1 < 0$ (or, > 0) would indicate that banks tend to make systematic errors on the side of overreaction (or underreaction). The intuition behind the above is as follows: when loan losses decrease (increase), banks *overreact* to this positive (negative) development and become too optimistic (pessimistic) about the future. Consequently, expectations about future loan losses go below (above) the rational benchmark, and they

[12] We include bank fixed effect to account for any unobserved persistent differences such as differences among banks' prior expectations. For instance, some banks might be more optimistic than others. By using a fixed-effect model, we focus on the time-series dimension within banks. To avoid potential biases in computing t-statistics, we dually cluster standard errors both on bank and time.

keep lower (higher) provisions, which causes a predictable increase (decline) in the forecast error. Overreaction thus gives rise to a negative relationship between past loan losses and the forecast errors ($\beta_1 < 0$). By the same logic, $\beta_1 > 0$ indicates underreaction.

2. Data and Evidence of Non-rational Expectations

2.1. Sample

We use FR Y-9C reports for quarterly financial statements of bank holding companies[13] spanning the period 1987:1-2019:4[14]. Since the later part of our analysis involves stock returns, we obtain stock price information from CRSP. We use the link table from the regulatory identification numbers (RSSD ID) to CRSP's permanent company numbers (PERMCO) provided by the New York Fed to link the quarterly financial data from FR Y-9C with the CRSP data. Our final sample consists of an unbalanced panel of 71,696 bank-quarter observations containing only publicly traded bank holding companies[15]. We winsorize all bank-level variables at the top and bottom 1%-ile to limit the influence of outliers. In Appendix A, we discuss the construction of variables. We report the details of variables and their summary statistics in Table A.1 and Table A.2 respectively.

As is standard in the literature (Beatty and Liao, 2011, 2014; Bushman and Williams, 2015; Hribar et al., 2017; Basu et al., 2020), we scale both provisions and net charge-offs with total loan balances at the beginning of the quarter. Figure 1 shows the scatter plot of the forecast errors. The figure shows substantial heterogeneity across banks and time. There always exists banks on both sides of the zero line, suggesting the presence of both overestimation and underestimation of future loan losses at the bank level.

[13] Bank holding companies need to be included in this definition because banks that belong to a holding company are not traded themselves. Since our analysis in the later sections involve stock returns, we conduct our analysis on publicly traded bank holding companies. However, we repeat our analysis for forecast error predictability using bank call report data. Our inferences remain unchanged. Results are not reported for brevity.

[14] While the Financial Accounting Standards Board (FASB) proposed a new expected credit loss accounting standard in June 2016, the new standard is effective only from 2020 for SEC-registered banks and from 2023 for all others. Thus, we restrict our sample to 2019 to stay consistent with the accounting norms over our sample period.

[15] We also replicate our analyses that don't involve stock returns using the full sample from FR Y-9C reports. Our findings are robust.

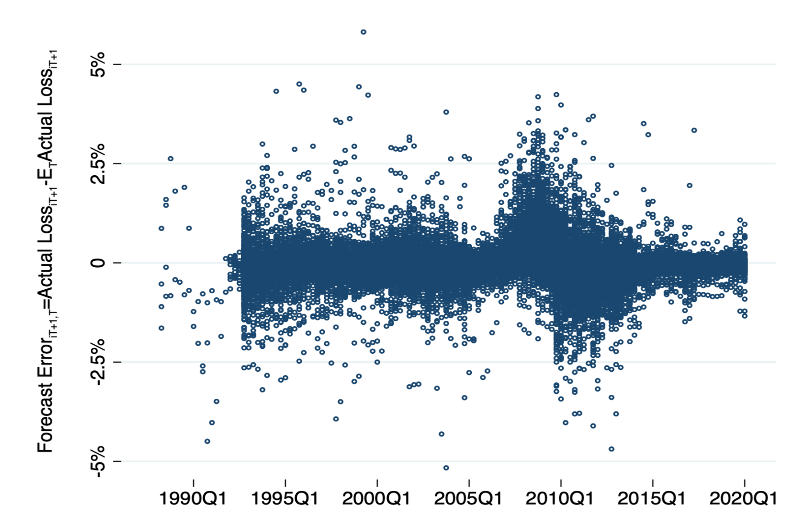


Figure 1

The figure shows scatter plot of forecast errors $(Actual\ Loss_{iT+1} - E_T Actual\ Loss_{iT+1})$. The forecast errors represent the difference between average loan loss realized over the next four quarters and expected loss averaged over the last four quarters as explained in Section 1.2.

This deviation of expectations from actual loan losses across banks and also over time could be completely random, which the banks fail to anticipate, or it could be systematic. Banks might be underestimating loan losses when they are optimistic and overestimating losses when they are pessimistic about the future. In Figure 2, we document some basic facts related to expected loan losses and actual losses before diving into the empirical analysis.

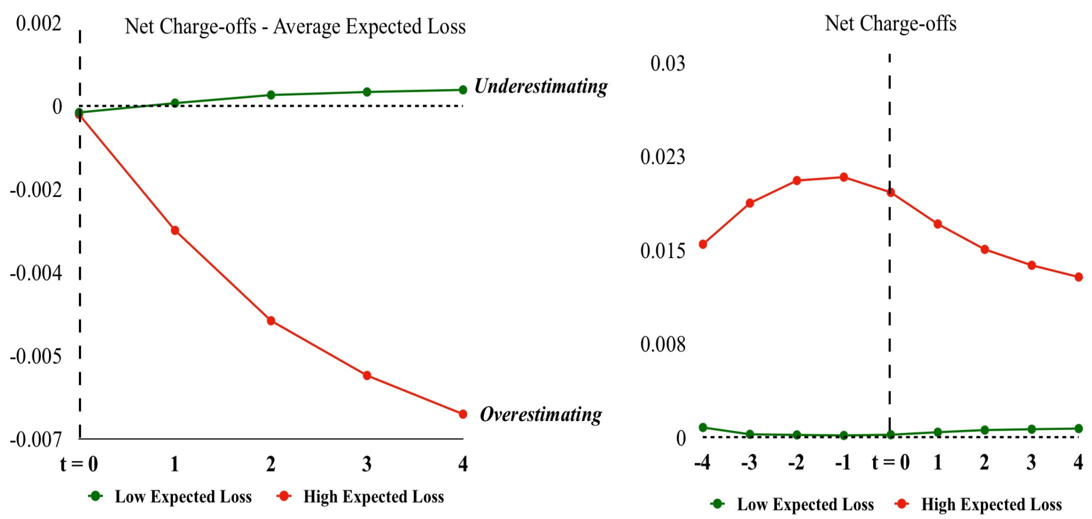


Figure 2

The left panel in Figure 2 shows the differences between net charge-offs (nco_{t+j}) from quarter

$t+1$ to $t+4$ and expected loss (the non-discretionary component of loan loss provision averaged over quarter $t-3$ to t). At $t=0$, we sort banks into groups based the amount of loans they are expecting to lose. *Low Expected Loss* and *High Expected Loss* in the figure corresponds to bottom 1 and top 1 percentile group respectively. Banks that are expecting higher level of loan losses, are actually overestimating; net charge-offs turn out to be lower than what they have expected. More importantly, these mistakes are not random. If we look past (before $t=0$), as shown in the right panel, banks in the *High Expected Loss* category have experienced increasing loan losses. Similarly, banks in the *Low Expected Loss* category have experienced a declining trend in loan losses in the past and have underestimated future loan losses. For these banks, net charge-offs turn out to be higher than what they have expected. This preliminary result suggests that banks tend to extrapolate the past too much while estimating future loan losses.

2.2. Evidence of Non-rational Expectations

As per our discussion in Section 1.2, we construct the forecast errors $(or,\ Forecast\ Error_{T+1,T})$ and plot them with past loan losses $(or,\ Actual\ Loss_T)$ in Figure 3. The figure suggests predictability between the two, in other words, a departure from rational expectation. When loan losses pile up, banks' expectations about future loan losses are too high. Banks keep *excess* provisions resulting in a predictable decline in the forecast error, implying that bad times are associated with *excessive* pessimism about the future. We observe such patterns during the 2001 recession and the recession following 2008 crisis. During good times, for instance, periods preceding the 2008 crisis, when loan losses are declining, expected losses are *too* low; the forecast errors predictably increase, implying that good times bring *excessive* optimism about the future.

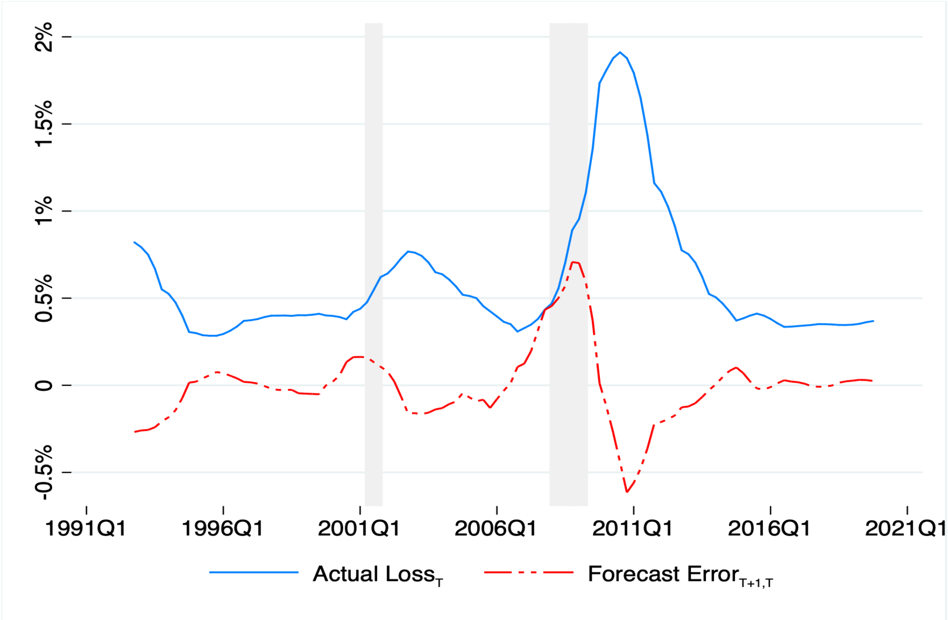


Figure 3

Predictability of Forecast Error: The figure plots the relationship between forecast errors in banks' expectations about future loan losses and past loan losses at the aggregate level. Similar to Basu et al. (2020), we scale these state variables with total loan book value of the previous period. Additionally, similar to Bordalo et al. (2018), we follow four quarter window for averaging these scaled variables. The blue line plots $Actual\ Loss_T$, defined as the average loan loss experienced by banks in the recent past, from quarter $t-3$ to t. The red lines plot the forecast error $(Forecast\ Error_{T+1,T})$ defined as the difference between $Actual\ Loss_{T+1}$ and $E_T Actual\ Loss_{T+1}$. $Actual\ Loss_{T+1}$ is the average loan loss realized over the next four quarters from quarter $t+1$ to $t+4$. $E_T Actual\ Loss_{T+1}$ is the non-discretionary component of loan loss provision $(ndllp)$ averaged over quarter $t-3$ to t. The figure documents a clear counter-cyclical relationship between the two variables suggesting predictability of the forecast error.

More formally, Table I reports the estimates from the econometric test of forecast error predictability as specified in Equation (1) and Equation (2). To avoid potential biases in computing t-statistics, we dually cluster standard errors both on bank and time. The key estimate of interest is the one that corresponds to the coefficient β_1 on past loan losses. If expectations were rational, our estimate of β_1 should be indifferent from zero. However, the value of our estimate is statistically different from zero - which indicates a systematic variation in expectations about future loan losses. Further, the sign of $\beta_1\ (< 0)$ reaffirms the cyclical variation of forecast error as depicted in Figure 3. The negative sign suggests an overreaction to current conditions in forming expectations about future loan losses.

Table I: Evidence on Predictability of Forecast Error

	β_1	**Median** β_1^i
$Actual\ Loss_{iT}$	-0.326^{***} (−16.07)	-0.338^{***} (0.016)
Bank F.E.	Yes	
Time F.E.	Yes	
Obs.	47,555	368
R^2	0.3151	
$Adj\ R^2$	0.2971	

Quarterly regression of errors in banks' expectations about future loan losses on past loan losses as specified in the regression model of Equation (1). Similar to Basu et al. (2020), we scale these state variables with the total loan book value of the previous period. Additionally, we follow a four-quarter window for averaging these scaled variables. The dependent variable $Forecast\ Error_{iT+1,T}$ is the difference between $Actual\ Loss_{iT+1}$ and $E_T Actual\ Loss_{iT+1}$ where $Actual\ Loss_{iT+1}$ is the average loan loss realized over the next four quarters from quarter $t+1$ to $t+4$. $E_T Actual\ Loss_{iT+1}$ is the non-discretionary component of loan loss provision $(ndllp)$ averaged over quarter $t-3$ to t. The first row reports β_1, the coefficient on $Actual\ Loss_{iT}$, defined as the average loan loss experienced by banks in the recent past, from quarter $t-3$ to t. Regressions are estimated using Correia (2016) multilevel panel fixed effect estimator. Standard errors are dually clustered both on bank and time; t-statistics is reported in parenthesis. In the last column, we report the median coefficient β_1^i from the bank by bank regression specification in Equation (2) and bootstrap standard errors in parenthesis. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

2.3. Alternative Explanations: Robustness Exercises

Notwithstanding our results on the predictability of forecast errors, there could still be potential concerns that might suggest that the predictability arises from features of loan loss provisions unrelated to overreaction in expectations. For instance, there might be concerns about whether the observed predictability is an outcome of some rational response. Our response is by way of counterfactual estimation. Suppose that the predictability in forecast error is being entirely driven by other potential mechanisms. Then, once such concerns are controlled for, the predictability should at the very least weaken, if not disappear. We do a series of such counterfactual exercises for each of the following concerns and find no statistical support for any of them. We only provide a sketch of our exercises below and keep the details in Appendix C.

Alternative Case 1: Level of Past Loan Loss Allowance: One potential concern is that current provisions are not only a measure of expected losses, but also a reflection of how wrong the bank was in its past provisions. For instance, if the provisions were too high in the past, then the current provisions are likely to be lower. This can create a model error that might mechanically drive the countercyclical relationship between forecast errors and past loan losses. We address this by controlling for banks' past loan loss allowance in the regression specification of Equation (1). In Table C.1, we document that the negative relationship between the forecast error and past loan losses remains significant even after controlling for the past allowance level.

The accounting item on loan loss allowance also opens up another possibility. Unlike loan loss provisions, which is an income-statement item, the allowance is a balance sheet item and is, arguably, less susceptible to quarter-by-quarter fluctuations. In the last column of Table C.1 we replicate our analysis of Equation (1), only this time we use allowance as a measure of expected loss to calculate forecast error. We find qualitatively similar results and in particular, $\beta_1 < 0$ continues to remain statistically significant.

Alternative Case 2: Capital Management: Another possible concern could be that the negative relationship between the forecast error and past loan losses need not reflect overreaction, but regulatory capital management under the asymmetric cost of provisions. Recognizing that every unit of provision is one less unit available for regulatory capital, the cost of provisioning naturally varies across the business cycle. During periods of high credit demand - which are typically concomitant with low charge-offs - a bank under capital distress may rationally choose to delay the recognition of loan losses to maintain competitive levels of regulatory capital (see Calomiris and Mason (2003) and Calomiris and Wilson (2004)). Under such a situation, the shadow cost of provisions is higher and therefore, a bank is likely to underprovision. In contrast, during periods of low credit demand and high loan losses, as the bank's lending becomes less sensitive to its capital position, it may rationally choose to keep higher provisions. Together, this might drive the countercyclical relationship of the forecast error with past loan losses.

To address this concern, we introduce a variable to measure the slack in capital position. To the extent that concerns about capital management drive our results, the predictability of forecast errors should get attenuated as a bank's capital position improves. In Table C.2 we report the results of our test and find no statistical support for this channel.

Alternative Case 3: Regulatory Scrutiny: There could be yet another alternative explanation to the observed negative relationship between the forecast error and past loan losses, one that admits a role for variation in regulatory scrutiny. Typically, after periods of low charge-offs, regulatory scrutiny is likely to be relaxed. This may encourage banks to inflate earnings by underprovisioning. Likewise, in bad times, following the periods of high charge-offs, regulatory scrutiny may tighten, and banks may prefer to err on the side of overprovisioning.

To rule out the above alternative explanation, we derive a testable hypothesis by exploiting the variation in the extent of regulatory scrutiny across banks. Regulatory scrutiny can vary across the cross-section, with large bank holding companies being subjected to *extra* scrutiny, as they pose a greater degree of systemic risk[16]. In a recent paper, Hirtle et al. (2020) establishes that large banks, in each federal reserve district, receive a disproportional amount of supervisory attention. As a result, large banks are unlikely to be driven by earnings management of provisioning.

We introduce two dummy variables to indicate: (i) whether a bank can be considered large[17] as per Hirtle et al. (2020), and (ii) whether the banking sector is subjected to higher regulatory scrutiny in a given quarter. We interact these two with the actual losses.

To the extent that changes in regulatory scrutiny and earnings management, drive the predictability of the forecast error, the effect is likely to be weaker for large banks, especially following periods of heightened regulatory scrutiny. However, as reported in Table C.3, the coefficient for the interaction term turns out to be negative and statistically significant, therefore, offering no support for this channel.

Alternative Case 4: Variation in provisioning practices - 1990s Boom: Liu and Ryan (2006) argue that banks were keeping excess provisions during the 1990s boom period to smooth income and in a bid to obscure excessive loan loss allowance from regulators, they were also accelerating the charge-offs. Only after repeated regulatory interventions towards the end of 90s and early 2000s, the

[16] See section 1060.0 in FRB Supervision Manual.

[17] We also use an alternative definition where banks whose assets (averaged from quarter t - 3 to t) are more than the $95th$ percentile threshold in the corresponding quarter are considered *Large*. Our results remain robust to this alternative definition.

banks restrained their practice of overprovisioning and acceleration of charge-offs. For our sample, such manipulation may create a variation in the direction of overreaction and can cause an upward bias to our results.

To understand if the practices of the 1990s could have an overbearing influence on our results, we estimate Equation (1) separately for 1991Q1-2000Q4 and 2001Q1-2019Q4. This separation of time period is consistent with the analysis of Liu and Ryan (2006). In Table C.4, we report the results of these two subsamples. As expected, the estimated coefficients β_1 are higher in the 1990s - suggesting stronger predictability. However, the post-90s periods continue to exhibit strong predictability as well. This suggests that our findings continue to remain robust over different parts of the entire sample period.

Alternative Case 5: Error minimization: So far, the estimation of Equation (1) has been conducted by minimizing the mean of the squared (forecast) errors (mse). According to Basu and Markov (2004), banks may instead be setting the expected loss with the implicit objective of minimizing *absolute* forecast errors instead of *squared* forecast errors. And therefore, coefficients arrived by the mse method could be wrongly inferred as deviation from rational expectations[18]. To address this concern, we follow Basu and Markov (2004) and estimate Equation (1) by minimizing absolute error. In Table C.5, we report the results. Not only do we find qualitatively similar results with $\beta_1 < 0$, but the magnitudes of the coefficients also remain largely unchanged. Having addressed the above concerns, we now turn our attention to mapping the predictability of forecast errors to overreaction in expectations.

3. Diagnostic Expectations as a plausible framework

Having established the departure of banks' expectations about future loan losses from rational expectations ($\beta_1 \neq 0$), we now turn our attention to model the theoretical underpinnings of the finding. We propose a reduced form model to argue that the evidence of departure from rational expectations can be interpreted through - what the recent literature (see Bordalo et al. (2018, 2019, 2020, 2021)) has referred to as - "diagnostic expectations" framework. In this framework, the agent assigns disproportionate weight to current conditions when predicting the distribution of future outcomes.

[18] Except for the special case when the distribution of the underlying variable is symmetric.

Specifically, the agent "overreacts" on the current information set while forming expectations of the future, and, therefore, the current information set has a predictive or "diagnostic" utility for future forecasts.

To differentiate such expectations from rational expectations, the literature uses an operator E_T^θ, where E_T is the usual rational expectations operator for expectations formed at time T. In particular, for a variable $\omega \sim N(0, \sigma_\omega^2)$ that follows a standard AR(1) process (see Bordalo et al. (2018) for more details), such expectations are defined as:

$$E_T^\theta(\omega_{T+1}) = E_T(\omega_{T+1}) + \theta[E_T(\omega_{T+1}) - E_{T-1}(\omega_{T+1})] \tag{3}$$

where $\theta \in (-\infty, \infty)$ is the *diagnosticity* parameter that captures the excessive weightage given to current information. The case of $\theta = 0$ implies rational expectations, whereas > 0 implies overreaction, and < 0 implies underreaction.

Suppose that a representative bank forms expectations as per the diagnostic expectations framework. In addition, the bank lends an amount of L_T to a (risky) borrower for one period. At the time of maturity, $T+1$, the bank loses a fraction g_{T+1} of the original loan L_T. Therefore, at time $T+1$, the actual loss turns out to be $L_T g_{T+1}$. The accounting standards would require the bank to make provisions for such loss. Accordingly, the provisions at T would be made on the expected loss for $T+1$. Similar to our empirical analysis, we scale both the actual loss and expected loss for $T+1$ with total loans issued in T and denote the scaled variables as $Actual\ Loss_{T+1}$ and $E_T Actual\ Loss_{T+1}$, respectively. With this minimum setup, we can denote,

$$E_T Actual\ Loss_{T+1} = \begin{cases} E_T g_{T+1} & Rational\ Expectations \\ E_T^\theta g_{T+1} & Diagnostic\ Expectations \end{cases} \tag{4}$$

We make simplifying assumptions about the loan loss process: g_T; the fraction of the loan the bank stands to lose, follows an AR(1) process[19]:

$$g_T = \rho g_{T-1} + (1-\rho) g_0 + \varepsilon_T \tag{5}$$

where g_0 represents average loan loss fraction over the whole business cycle, $\rho \in (0,1)$ and ε_T

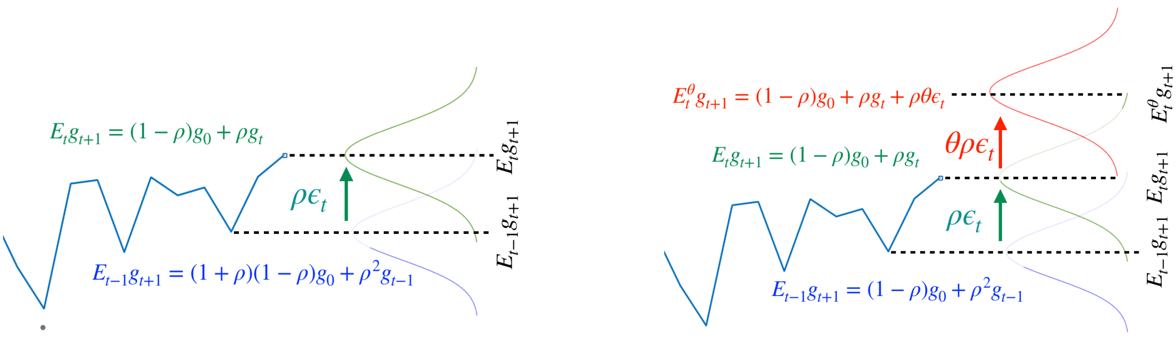
[19] This can be easily generalized to richer AR(n) processes without changing the intuition.

represents a normally distributed shock with zero mean and constant variance.

Then, in accordance with Equation (3), one can write

$$E^{\theta}_{T} g_{T+1} = E_T g_{T+1} + \theta[E_T(g_{T+1}) - E_{T-1}(g_{T+1})]$$
$$= E_T g_{T+1} + \theta\rho(g_T - E_{T-1}g_T) = E_T g_{T+1} + \theta\rho\varepsilon_T \quad (6)$$

Here, the term $\theta\rho(g_T - E_{T-1}g_T)$ denotes the "overreaction" component in loan loss provisions. For instance, upon an unexpected shock to current loan loss $(g_T \neq E_{T-1}g_T)$, the bank gets *predictably* (θ) swayed by this new information (ε_T) while forming its expectation of loan loss for the next period. In the case of an unfavourable shock to current period loan loss rate $(g_T > E_{T-1}g_T)$, a rational bank at time T would revise its expectation about g_{T+1} $(or,\ E_T g_{T+1} - E_{T-1}g_{T+1})$ upward by $\rho\varepsilon_T$. However, in the case of overreaction $(\theta > 0)$ the bank revises its expectation further upward $(or,\ E^{\theta}_{T} g_{T+1} - E_T g_{T+1})$ by an amount of $\theta\rho\varepsilon_T$ as illustrated Figure 4.



(a) Rational (b) Diagnostic

Figure 4

The forecast error constructed as the difference between actual and expected loss, can be written as

$$Actual\ Loss_{T+1} - E_T Actual\ Loss_{T+1}$$
$$= \begin{cases} (g_{T+1} - E_T g_{T+1}) \equiv \varepsilon_{T+1} & Rational\ Expectations \\ (g_{T+1} - E_T g_{T+1}) - \theta\rho(g_T - E_{T-1}g_T) \equiv \varepsilon_{T+1} - \theta\rho\varepsilon_T & Diagnostic\ Expectations \end{cases} \quad (7)$$

Under rational expectations, when $\theta = 0$, the forecast error is simply the random noise ε_{T+1} which, by definition, is orthogonal to all information available at time T. On the other hand, the forecast error under diagnostic expectations turns out to be $\varepsilon_{T+1} - \theta\rho\varepsilon_T$, and is predictably dependent upon the departure of current period loan loss from its expectations $(g_T - E_{T-1}g_T \equiv \varepsilon_T)$. For instance, in the case of an adverse shock to current period loan loss $(g_T > E_{T-1}g_T)$ - *and* in the particular case of overreaction $(\theta > 0)$ - the provisions for the future are set higher on account of observed higher loan loss, and this leads to a predictable decline in the forecast error. This drives an inverse relationship between observed loan loss and the forecast error. It is precisely this underlying mechanism that we propose, can help reconcile the evidence as depicted in Figure 3 and as established through the negative sign of β_1 in Table I.

Under our empirical specification of Equation (1), this simple setup also allows us to theoretically pin-down the coefficient β_1 and also lends itself to calibrating the overreaction parameter, (θ). Using Equation (7), we can write:

$$\beta_1 = \frac{-\rho\theta Var(\varepsilon_T)}{Var(g_T)} \tag{8}$$

Clearly, when $\theta = 0$ - as is the case under rational expectations - β_1 becomes identically equal to 0. More importantly, under overreaction in expectations, that is when $\theta > 0$, β_1 becomes unambiguously negative - which is in line with the results of Table I. Further, under the restriction that g_T follows an AR(1) process of Equation (5), we can use the estimate of β_1 to calibrate θ. Through the formulation of Equation (5), we estimate the persistence parameter $\rho \approx 0.5$[20]. We combine this with the estimates of β_1 of Table I, to produce estimates for θ in Table II (for more details, see Appendix D).

The overreaction parameter, as reported in Table II, is in the same ballpark as the estimates of (a) Bordalo et al. (2018), who use data on professional forecasts of credit spreads $(\theta = 0.91)$, (b) Bordalo et al. (2019), who use analyst expectations of US listed firms' long term earnings growth $(\theta = 0.91)$, (c) Pflueger et al. (2020), who use stock price-derived measures of risk perception $(\theta = 1)$, (d) d'Arienzo (2020), who use bond prices $(\theta = 1)$ and (e) Bordalo et al. (2021), who use firm-level forecasts from the IBES manager guidance database $(\theta = 1.069)$. This convergence in the estimated value of our overreaction parameter in the banking sector with the larger litera-

[20] Our estimate is similar to those reported in the finance literature, albeit in slightly allied contexts. For instance, Bordalo et al. (2018) estimate a persistence parameter of 0.65 for credit spread and Bordalo et al. (2019) estimate a value of 0.5 for earnings per share series.

ture in finance only serves to reaffirm the underlying mechanism of expectation formation. Having discussed the overreaction parameter in a panel-level setting, we now turn our attention to bank-by-bank estimation of the overreaction parameter.

Table II: Estimation of the Overreaction parameter θ

	Model 1
θ	0.87
	$(0.77, 0.98)$

The table provides estimation of the overreaction parameter (θ) along with the 95% confidence interval. As specified in Equation (1), we regress forecast error in banks' expectations about future loan losses, $Forecast\ Error_{iT+1,T}$ on actual losses, $Actual\ Loss_{iT}$ and estimate β_1. Persistence parameter of loan loss process, ρ is measured by regressing $Actual\ Loss_{iT+1}$ on $Actual\ Loss_{iT}$. Using β_1 and ρ, we compute the overreaction parameter θ as per Equation (8). For more details, see Appendix D.

3.1. Overreaction at bank-by-bank level

One obvious challenge here is the reduced number of observations per bank which may affect the statistical power of our results. For reliable estimation of θ^i, we restrict the number of observations for each bank to at least 50. Estimating θ^i follows a two-step process. First, we estimate the persistence parameter, ρ^i for each bank, by regressing $Actual\ Loss_{iT+1}$ on $Actual\ Loss_{iT}$ (see Appendix D). We then combine this with the estimates of β_1^i of Equation (2) using the expression for β_1^i in Equation (8). This allows us to estimate a bank specific overreaction parameter θ^i.

Table III: Estimation of the Overreaction parameter θ^i

θ^i	Mean	SD	Skewness	25th perc.	Median	75th perc.
Model 1	1.25	1.06	2.84	0.64	0.96	1.47
				(0.027)	(0.041)	(0.079)

The table provides estimation of the overreaction parameter (θ^i). As specified in Equation (2), we run bank by bank regression of forecast error in banks' expectations about future loan losses, $Forecast\ Error_{iT+1,T}$ on actual losses, $Actual\ Loss_{iT}$ and estimate β_1^i. Persistence parameter of loan loss process, ρ^i is measured by running bank by bank regression of $Actual\ Loss_{iT+1}$ on $Actual\ Loss_{iT}$. Using β_1^i and ρ^i, we compute the bank specific overreaction parameter θ^i as per Equation (8). Bootstrap standard errors are reported in parenthesis. For more details on estimation of θ_i, see Appendix D.

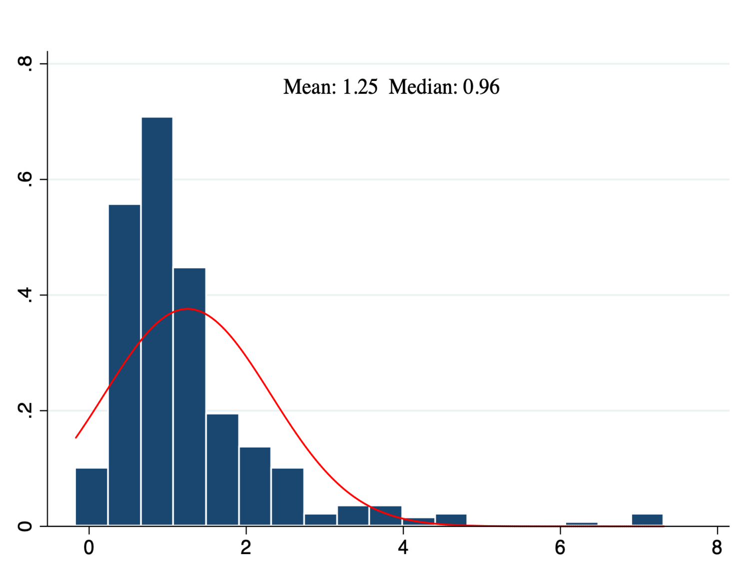


Figure 5

Overreaction Parameter θ^i: The figure plots the distribution of the overreaction parameter θ^i across banks. As specified in Equation (2), we regress forecast error in banks' expectations about future loan losses, $Forecast\ Error_{iT+1,T}$ on actual losses, $Actual\ Loss_{iT}$ at individual bank level and estimate β^i. Persistence parameter of loan loss process, ρ^i for each bank is measured by regressing $Actual\ Loss_{iT+1}$ on $Actual\ Loss_{iT}$. Using β_1^i and ρ^i, we compute a bank specific overreaction parameter θ^i as per Equation (8). For more details, see Appendix D. The median values of overreaction parameter is consistent with the corresponding panel-level estimate.

We report the descriptive statistics of θ^i in Table III after winsorizing at the 2% level to minimize the impact of outliers. Figure 5 plots the distribution of the same.

Interestingly, the bank-by-bank level estimation also permits us to explore the variability of overreaction parameter with the underlying persistence of the actual loan losses. Figure 6 presents a graphical representation of how the overreaction parameter θ^i varies with the persistence parameter ρ^i. Consistent with Bordalo et al. (2020) and Afrouzi et al. (2021)[21], we identify a strong negative association: the relative degree of overreaction parameter varies inversely with persistence. In other words, banks which face a more transitory loan loss process, seem to exhibit stronger overreaction towards current information about loan losses.

[21] To account for the pronounced overreaction for less persistent process, Afrouzi et al. (2021) suggests an alternative expectations formation model with costly information processing. Agents estimate the long-run mean of the process using a mix of recent observations and past data. Because of the cost of utilizing past information, recent observations have a disproportionate influence, resulting in overreaction. In our context, the costs of utilizing past information are unlikely to be of first order importance, because banks extensively use past information about the performance of loans and loan loss rate for many other purposes such as credit risk modelling, estimating regulatory capital, stress testing, etc.

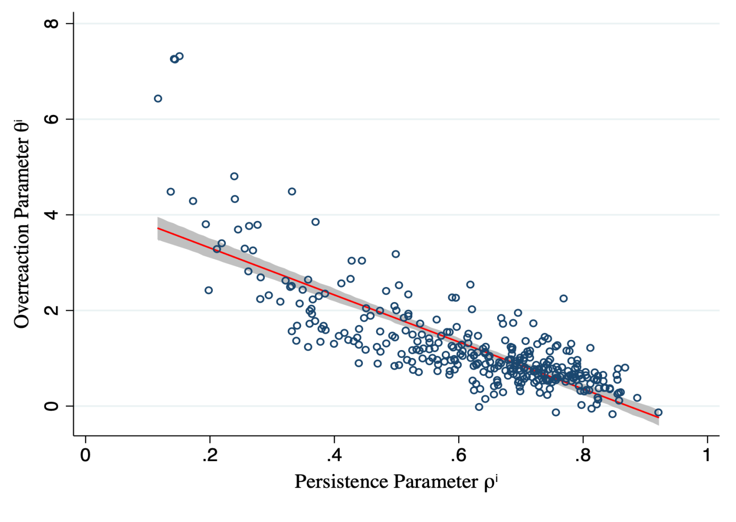


Figure 6

Overreaction and Persistence of Loan Loss Process: The figure plots bank-level overreaction parameter, θ^i against the persistence parameter of the loan loss process, ρ^i. The shaded region indicates 95% confidence interval. As specified in Equation (2), we regress forecast error in banks' expectations about future loan losses, $Forecast\ Error_{iT+1,T}$ on actual losses, $Actual\ Loss_{iT}$ at individual bank level and estimate β_1^i. Persistence parameter of loan loss process, ρ^i for each bank is measured by regressing $Actual\ Loss_{iT+1}$ on $Actual\ Loss_{iT}$. Using β_1^i and ρ^i, we compute a bank specific overreaction parameter θ^i as per Equation (8). The figure suggests that banks which face a more transitory loan loss shock, tend to exhibit higher overreaction.

3.2. Overreaction across Loan Categories

We also explore if there exist any heterogeneity in the overreaction parameter θ across loan categories. Liu and Ryan (1995, 2006) argue that banks' due diligence varies with the type of loan segment. Real estate loans and commercial and industrial loans are assessed on an individual loan basis with greater caution and careful deliberations. In contrast, loans related to credit card and consumer segments are primarily evaluated on a statistical basis, often, by relying on past information. Consequently, in line with the argument made in Gennaioli and Shleifer (2010) and Shleifer (2012), we expect overreaction to be more pronounced for the latter categories of loans.

The data permits us to analyze the forecast error predictability across four loan categories (i) real estate loans, (ii) commercial and industrial (CI) loans, (iii) credit card loans, and (iv) other consumer loans. We obtain disaggregated data on (i) net charge-offs from schedule HI-B, (ii) loans and loan loss allowance from schedule HI-C, and (iii) non-performing assets from schedule HC-N of FR Y-9C report. One challenge is that the number of observations reduces considerably because the disaggregated data on loan loss allowance is only available from 2013 onwards[22]. While banks are not required to report loan loss provisions at disaggregated level, we deduce it from the changes in loan loss allowance and net charge-offs. For each loan category, we first construct the expected loan loss component from the loan loss provisions and then construct the forecast error term as described in Section 1.1 and 1.2. Table IV shows the estimates of empirical specification in Equation (1) for each loan category.

Table IV: Predictability of Forecast Errors and Overreaction across Loan Categories

	Real Estate Loans	CI Loans	Credit Card Loans	Other Consumer Loans
$Actual\ Loss_{iT}$	-0.623^{***} (-9.02)	-0.571^{***} (-7.27)	-1.08^{***} (-8.66)	-0.855^{***} (-14.56)
θ	2.31 (1.77,2.85)	1.69 (1.21,2.18)	4.88 (3.68,6.06)	3.21 (2.74,3.68)
Bank F.E.	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes
Obs.	3,173	4,763	908	2,783
R^2	0.6199	0.4542	0.7534	0.7024
$Adj\ R^2$	0.5806	0.4117	0.7262	0.6711

Quarterly regression of errors in banks' expectations about future loan losses on past loan losses as specified in the regression model of Equation (1) across loan categories. For each category of loan, we scale these state variables with the respective loan book value of the previous period. Additionally, we follow a four-quarter window for averaging these scaled variables. We construct $Forecast\ Error_{iT+1,T}$ ($=Actual\ Loss_{iT+1}$ - $E_T Actual\ Loss_{iT+1}$) and regress it on $Actual\ Loss_{iT}$. $Actual\ Loss_{iT+1}$ is the average loan loss realized over the next four quarters from quarter $t+1$ to $t+4$. Expected loss ($E_T Actual\ Loss_{iT+1}$) is the non-discretionary component of loan loss provision ($ndllp$) averaged over quarter $t-3$ to t. The explanatory variable $Actual\ Loss_{iT}$ is defined as the average loan loss experienced by banks in the recent past, from quarter $t-3$ to t. Regressions are estimated using Correia (2016) multilevel panel fixed effect estimator. Standard errors are dually clustered both on bank and time; t-statistics is reported in parenthesis; *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. After estimating β_1 for each loan category, we measure the persistence parameter of loan loss process and compute the overreaction parameter θ as per Equation (8). As evident, heterogeneous loan categories (Real Estate, Commercial and Industrial loans) tend to exhibit weaker overreaction as compared to the homogeneous loan categories (Credit Card and Consumer loans).

[22] Earlier the banks of size above a certain threshold of $1 billion or more in total assets were required to report disaggregated data. This was revised to $5 billion later. Since schedule HI-C groups commercial loans with other categories, we obtain data on commercial and industrial loans from schedule HC-C.For more details, please refer to https://www.federalreserve.gov/apps/reportforms/reporthistory.aspx.

Consistent with our intuition, we find the overreaction parameter to be higher for Credit Card and other Consumer loans compared to Real Estate, and Commercial and Industrial loans. The observed heterogeneity in θ across loan categories thus provides further credence to our argument for the presence of overreaction or diagnostic belief in banks' expectations.

4. Overreaction, Credit Growth, and Equity Return

We now explore if the extent of overreaction in expectations has implications on credit growth, quality of loan portfolio, profitability and equity returns.

As noted in Equation (6), the diagnostic expectations contain an additional $\theta\rho\varepsilon$ term, which denotes the overreaction component. While we cannot directly observe the overreaction component from banks' expectations about loan losses, we can identify it through Equation (7). The forecast error at any point in time, t consists of two components: a stochastic component (ε_{t+1}), and the overreaction component $(-\theta\rho\varepsilon_t)$, which captures the predictability of the forecast error from past loan losses. When times are good $(\varepsilon_t < 0)$, overreaction can cause banks to become overly optimistic, resulting in biased expectations about future credit losses. As a result, banks set aside fewer provisions, leading to a predictable increase in the overreaction component of the forecast errors. Conversely, during bad times $(\varepsilon_t > 0)$, banks tend to be overly pessimistic, leading to excess provisions and a predictable decline in the overreaction component. Arguably, the predictable component of the forecast error for bank i at time t serves as a measure of the extent of overreaction.

To examine whether the extent of overreaction influences credit growth, return on assets and equity return, we split the sample in quartiles along the extent of overreaction component in the forecast errors, averaged over three years $(Overreaction_{it-3,t})$. Table V provides summary statistics of several bank characteristics by $0verreaction$ quartile. The first row of the table presents the median credit growth over a period of three years. We calculate the 3-year credit growth as the change in loans and leases from year $t-3$ to year t divided by total loans in year $t-3$. Credit growth increases as we move up the quartiles. Banks in the top quartile have a median credit growth that is around 26 percentage points higher compared to banks in the bottom quartile.

Table V: Summary Statistics by *Overreaction* quartile

	Quartiles based on $Overreaction_{it-3,\,t}$			
	1	2	3	4
3-year Credit Growth	0.186	0.304	0.378	0.447
A. *NPA (in %)*				
(a) Current	1.84	1.31	1.02	0.73
(b) 3 year ahead	1.41	1.19	1.09	0.93
Difference (b-a)	-0.43^{***}	-0.11^{***}	0.06^{***}	0.19^{***}
	(-26.40)	(-8.39)	(6.58)	(20.53)
B. *Return on Assets (in %)*				
(a) Current	0.52	0.59	0.63	0.66
(b) 3 year ahead	0.62	0.60	0.62	0.63
Difference (b-a)	0.09^{***}	0.007^{***}	-0.009^{***}	-0.03^{***}
	(23.69)	(2.77)	(-4.75)	(-17.90)
C. *Stock Return*				
(a) Prior 3-year	0.383	0.3403	0.333	0.325
(b) 3-year ahead	0.451	0.3406	0.312	0.264
Difference (b-a)	0.067^{***}	0.0004	-0.021^{***}	-0.063^{***}
	(9.78)	(0.064)	(-3.77)	(10.68)

The table shows medians for key bank characteristics by $Overreaction$ quartile. We add $Overreaction_{it}$ estimated quarterly using appropriate lag operator to measure the extent of overreaction in banks' expectations in the last three years or, $Overreaction_{it-3,\,t}$ and sort banks into quartiles based on it. 3-year Credit Growth is growth in loans from the quarter that ended three years ago until the end of the current quarter. NPA or Non-performing assets is the ratio of the sum of non-accrual loans and loans past due 90 days or more to total loans at the beginning of the quarter multiplied by 100. Return on assets (ROA) is defined as the ratio of net income to total assets multiplied by 100. The prior three-year return is the bank's stock return from the quarter that ended three years ago until the end of the current quarter. The 3-year ahead return is the bank's stock return from the current quarter to the quarter three years ahead. For the "Difference" row, we estimate the median value of every period's characteristics separately for each $Overreaction_{it-3,t}$ quartile and test the significance of the difference between future and current value by performing a t-test.

The next rows examine the quartile trends in banks' future performance. To measure performance, we use two metrics: (1) non-performing assets (NPA), which is the sum of loans past due by 90 days and non-accrual loans scaled by total loans at the beginning of the quarter, and (2) return on assets (ROA), which is calculated as net income divided by total assets. Within each quartile, we compare banks' current performance with their 3-year ahead performance[23]. Banks in higher $Overreaction_{it-3,t}$ quartiles (quartile 3 and 4) tend to perform worse in the future, as reflected by higher levels of non-performing loans and lower returns on assets. In contrast, banks in quartile 1 and 2 experience a drop in non-performing loans and an improvement in return on assets. These results suggest that overoptimistic beliefs driven by overreaction in good times can lead banks to engage in riskier lending practices, which eventually result in poor future performance. Further, the third row which tabulates the stock return over quartiles, seems to suggest that the banks in the higher $Overreaction_{it-3,t}$ quartile earn lower future returns for investors, relative to the banks belonging to the lower quartile. This suggests that the market perhaps fails to recognise the risk associated with banks with greater $Overreaction_{it-3,t}$, leading to a relative overvaluation of stocks vis-a-vis banks with lower $Overreaction_{it-3,t}$. These trends highlight the importance of overreaction channel in banks' expectations, particularly during good times when overreaction can lead to excessive credit expansion and future financial instability.

To further underscore the relevance of our measure, we make use of a qualitative data provided by the Federal Reserve Board's Senior Loan Officer Opinion Survey on Bank Lending Practices (SLOOS). This data reports whether the lending standards for a bank has been loosened or tightened over a quarter. Our contention is that banks whose expectations about the future improve (worsen) compared to the previous quarter - that is, $\Delta Overreaction_{it} > 0$ (or, < 0) - are likely to ease (tighten) the lending standards. We compute the net fraction of banks that report loosening their lending standards from the SLOOS as difference between the fraction of banks that have loosened their standards vis-a-vis those who have tightened them. On the other hand, we also compute the net fraction of banks that display an optimistic shift in the overreaction component as the difference between the fraction of banks with $\Delta Overreaction_{it} > 0$ and with $\Delta Overreaction_{it} < 0$. We juxtapose the two series in Figure 7.

[23] For each performance characteristic block, we report the current and future median values in row (a) and (b) respectively, and the last row reports the significance of the difference between future and current value. We estimate the median value of the characteristics every period separately for each $Overreaction_{it-3,t}$ quartile and test the significance of the difference between future and current value by performing a t-test. More information about how we construct these variables can be found in the Appendix A.

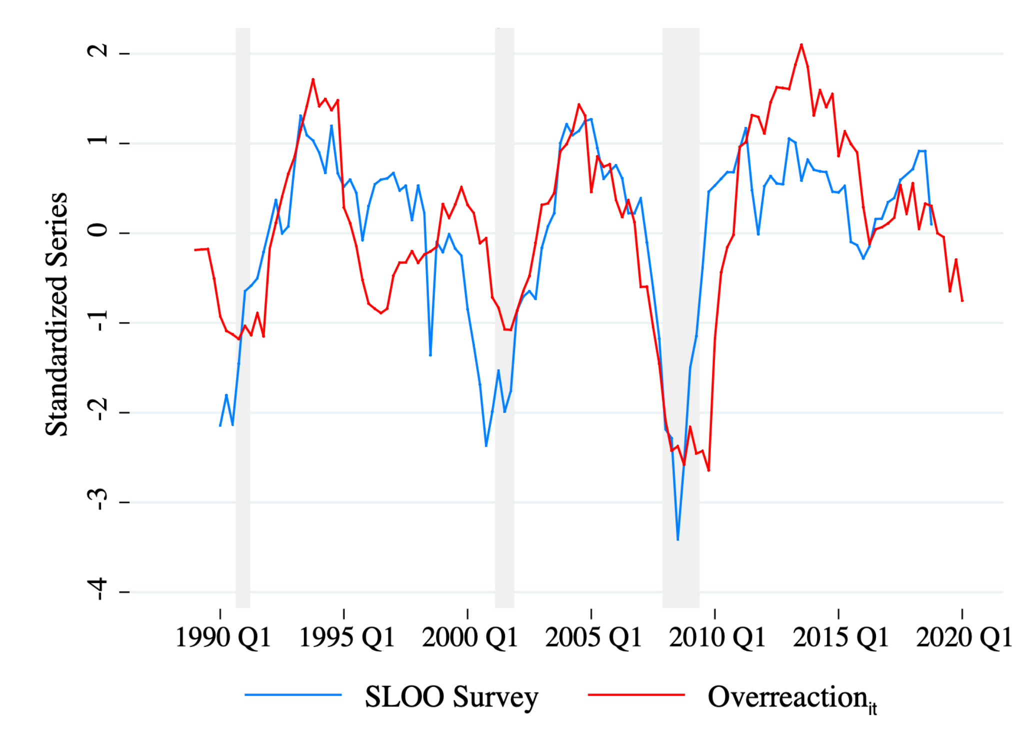


Figure 7

Fraction of Banks Loosening Credit Standard: The red line depicts the net percentage of banks with a rise in $Overreaction_{it}$, hence reasoned to ease credit standards. The blue curve depicts the net percentage of banks easing credit standards as reported by the Senior Loan Officer Opinion Survey on Bank Lending Practices.

Notably, the correlation between the two series — one obtained from a qualitative survey of commercial banks (blue line) and the other obtained from the expectations of banks (red line) - is strikingly high at 0.67 and statistically significant at the 5% level. This lends credence to our contention that overreaction in expectations is likely to play a central role in determining the quality and quantity of credit. In the following sub-sections, we establish these relationships formally.

4.1. Overreaction$_{it}$ and Credit Growth

To begin, we plot the time-series of median credit growth over the last three years, for two extremes of the level of overreaction during the same period in Figure 8. The figure shows that for the vast majority of sample years, banks in the highest quartile of $Overreaction_{it-3,t}$ (red curve) exhibited a significantly higher median credit growth than banks in the lowest quartile.

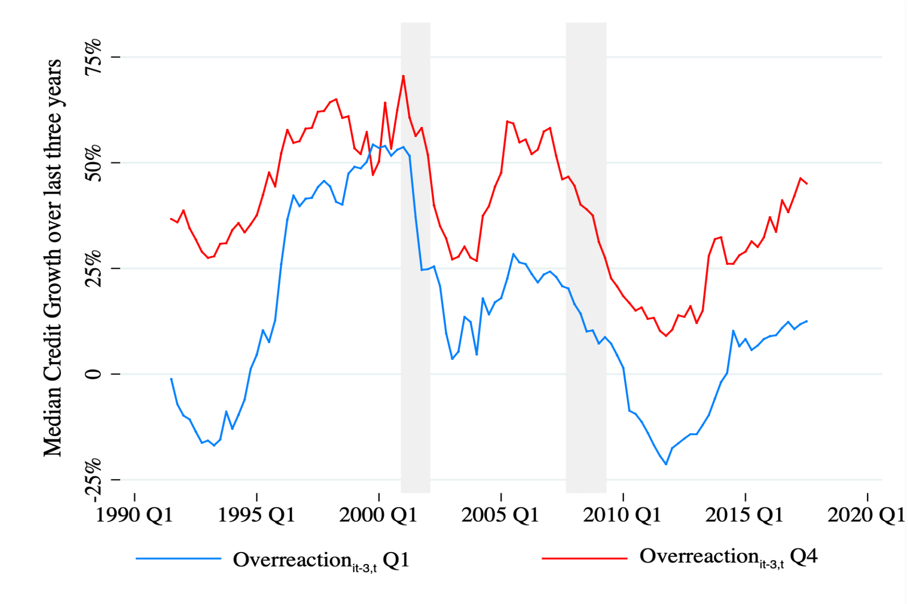


Figure 8

Median three-year credit growth for high versus low $Overreaction_{it-3,t}$ **banks:** The figure shows the time series of the median three-year credit growth for two groups of banks. Every quarter, we classify banks into quartiles based on $Overreaction_{it-3,t}$. The blue and the red line shows median credit growth for banks in the lowest and in the highest quartile respectively.

While this reinforces the underlying interrelationship between overreaction and credit quantity, we investigate it formally with the following panel regression:

$$\Delta credit_{it-k,t} = \beta_0 + \beta^{cdt} * Overreaction_{it-k,t} + \delta_t + \zeta_i + \varepsilon_{it} \tag{9}$$

Here, $\Delta credit_{it-k,t}$ is the dependent variable that measures loan growth from quarter 4 in year $t-k$ to quarter 4 in year t. $Overreaction_{it-k,t}$ is the explanatory variable that measures a bank's overreaction during the last k years, where k takes values of 1, 2, or 3. We add $Overreaction_{it}$ estimated quarterly using appropriate lag operator to measure $Overreaction_{it-k,t}$. The coefficient β^{cdt} measures the association between *Overreaction* and credit growth. To control for any time-invariant unobserved differences in banks' lending opportunities, we include bank fixed effect ζ_i in the above regression. δ_t denotes time fixed effect.

We present the results in panel A of Table VI. The marginal effects in the table correspond to the change in credit growth for a one standard deviation increase in *Overreaction* over the last one, two, and three years, respectively. We use only non-overlapping credit growth to avoid potential biases in computing test statistics. That is, in regressing credit growth over the last one-, two-, or three-years, we drop the intervening observations from our data set, in effect estimating the regressions on an annual, biennial, or triennial data. Standard errors are dually clustered [24].

Table VI: *Overreaction*, Credit Growth, and Interest Income

Panel A	$\Delta credit_{it-k,t}$ = **Credit growth from year** $t-k$ **to** t		
	(1)	(2)	(3)
	$\Delta credit_{it-1,t}$	$\Delta credit_{it-2,t}$	$\Delta credit_{it-3,t}$
$Overreaction_{it-1,t}$	0.067^{***} (16.47)		
$Overreaction_{it-2,t}$		0.139^{***} (13.44)	
$Overreaction_{it-3,t}$			0.208^{***} (8.46)
R^2	0.3186	0.4288	0.4654
$Adj\ R^2$	0.2501	0.3176	0.3187
Observations	10,838	4,811	2,787
Panel B	$IntRate_{it+k,t}$ = **Interest Income from year** t **to** $t+k$		
	$IntRate_{it,t+1}$	$IntRate_{it,t+2}$	$IntRate_{it,t+3}$
$Overreaction_{it-1,t}$	-0.0009^{***} (-2.92)		
$Overreaction_{it-2,t}$		-0.0012^{**} (-2.86)	
$Overreaction_{it-3,t}$			-0.0013^{***} (-3.17)
R^2	0.8953	0.9094	0.9156
$Adj\ R^2$	0.8847	0.8916	0.8921
Observations	10,834	4,254	2,272
Bank Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes

The table reports results from the regression of credit growth and interest rate on *Overreaction*. In panel A, the dependent variables are bank's credit growth over the last $k \in (1,2,3)$ years. In panel B, the dependent variables are average interest income banks earn over the next one, two and three years. We regress the variables on $Overreaction_{it-k,t}$ that measures the extent of overreaction during the last $k \in (1,2,3)$ years as explained in Section 4. The explanatory variable is in standard deviation relative to its mean. Regressions are estimated using Correia (2016) multilevel panel fixed effect estimator. Standard errors are dually clustered on bank and time. t-statistics in parenthesis. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

[24] For example, we look at credit growth from 1988Q4 to 1989Q4, and so on for one year; from 1988Q4 to 1990Q4, and so on for two years and from 1988Q4 to 1991Q4, and so on for three years. See Appendix A.

Across all specifications, we find that the coefficient β^{cdt} is positive. A one standard deviation increase in $Overreaction_{it-1,\,t}$ is associated with 6.7 percentage points higher credit growth during that period. In columns (2) and (3), the effects of overreaction on credit growth during the last two and three years are reported. To interpret the magnitudes, a one standard deviation increase in $Overreaction_{it-2,\,t}$ and $Overreaction_{it-3,\,t}$ are associated with approximately 14 percentage points and 20 percentage points higher credit growth over the last two and three years respectively.

To be sure that the effect of overreaction on credit growth is due to increase in credit supply,rather credit demand, we make use of the methodology of Mian et al. (2017) who use the movement of interest rates for identification. Holding credit supply fixed, an increase in credit demand should lead to higher interest rates as credit demand rises. We scale interest income earned by banks in the future with total loans at the beginning of the period and average it over the next one to three years to measure the cost of credit. The results in Panel B of Table VI suggest that as $Overreaction$ increases, the cost of credit declines in the future. Thus, the positive association between overreaction and credit growth, followed by a decline in interest income in subsequent years, supports the argument that credit supply - and not demand - is at work here.

4.2. Do banks with higher $\mathbf{Overreaction_{it}}$ make bad loans?

Related to the previous section which documented whether overreaction impacts the quantity of credit extended, we also explore if overreaction impacts the quality of such credit. If overreaction cause banks' beliefs to improve in good times, underestimate credit risk and credit growth to rise, we should see an increase in $Overreaction_{it}$ to predict a decline in the quality of bank loan portfolio and a fall in profitability in the future.

We use two metrics: (i) *NPA* as the sum of non-accrual loans and loans past due 90+ days divided by total loans at the beginning of the period multiplied by 100, and (ii) *ROA*, as net income divided by total assets multiplied by 100. Specifically, we estimate the following setup:

$$\Delta X_{it,t+j} = \beta_0 + \beta^{performance} * Overreaction_{it-k,t} + \delta_t + \zeta_i + \varepsilon_{it} \tag{10}$$

where X denotes either the quality of bank loan portfolios or, the profitability of individual banks in the subsequent one to three years. In Table VII and VIII, we present the results where the dependent variables (X) are the evolution of *NPA* and *ROA* over the next one to three years respectively. The results demonstrate that banks report an increase in nonperforming loans and a decrease in the return on assets following an increase in $Overreaction$ (expressed in standard deviations with respect to mean).

Specifically, a one-standard-deviation (sd) increase in $Overreaction$ during the last one, two, or three years is associated with a rise in nonperforming loans in the subsequent one to three years. Panel A shows that banks with one sd higher overreaction during last year, $Overreaction_{it-1,t}$ report about 0.234pp rise in *NPA* in year $t+1$. The rise in *NPA* is also observed from year $t+1$ to $t+2$ as well as $t+2$ to $t+3$. Column (7) and (8) across all panels show the total change in *NPA* from year t to year $t+3$ with and without bank fixed effects. For one sd higher $Overreaction_{it-1,t}$, the total increase in *NPA* over the next three years is around 0.91pp, or about 57% relative to the sample mean. We observe similar patterns in panel B and C too, where we use overreaction averaged over the previous two and three years.

Table VIII reports results from the regression of changes in return on assets on $Overreaction_{it}$. The results show an immediate deterioration in the performance of banks following an increase in $Overreaction_{it}$. Column (1), panel A shows *ROA* decreases by 0.167pp when overreaction over the immediate previous year, $Overreaction_{it-1,t}$ increases by one standard deviation. The corresponding effects for $Overreaction_{it-2,t}$ and $Overreaction_{it-3,t}$ are 0.139pp and 0.110pp respectively. From year $t+1$ to $t+2$ and year $t+2$ to $t+3$, we also observe a drop in the *ROA* across all panels. Columns (7) and (8) show the total change in *ROA* from year t to year $t+3$ with and without bank fixed effects. Following a one standard deviation rise in $Overreaction_{it-1,t}$, the total drop in *ROA* over the next three years is around 0.4pp, or 78% of the sample mean. Similar trends are also observed in Panels B and C.

In summary, our findings suggest that banks tend to make riskier loans during periods of higher $Overreaction$, which leads to a rise in nonperforming loans and a decline in return on assets in the future. These results are consistent with the view that, in good times, overreaction can make banks too optimistic about the future, leading them to underestimate the risks of the loans they are making and ultimately resulting in a significant deterioration in the quality of loan portfolio and profitability.

Table VII: Relationship between *Overreaction* and Non-Performing Loans

	$NPA_{t+1} - NPA_t$		$NPA_{t+2} - NPA_{t+1}$		$NPA_{t+3} - NPA_{t+2}$		$NPA_{t+3} - NPA_t$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. *Overreaction over last one year*								
$Overreaction_{it-1,t}$	0.234*** (7.84)	0.370*** (11.52)	0.230*** (7.66)	0.301*** (8.15)	0.156*** (6.89)	0.155*** (5.57)	0.691*** (4.78)	0.909*** (7.13)
R^2	0.2160	0.2929	0.2155	0.2941	0.2217	0.2813	0.4811	0.5701
No. of Observations	10,908	10,831	9,968	9,908	9,078	8,981	2,866	2,630
B. *Overreaction over last two years*								
$Overreaction_{it-2,t}$	0.236*** (7.52)	0.350*** (9.71)	0.193*** (6.53)	0.245*** (6.61)	0.121*** (5.72)	0.130*** (4.39)	0.576*** (3.76)	0.831*** (5.50)
R^2	0.2216	0.3048	0.2281	0.2898	0.2275	0.2842	0.4687	0.5607
No. of Observations	9,660	9,596	8,816	8,716	8,007	7,927	2,655	2,442
C. *Overreaction over last three years*								
$Overreaction_{it-3,t}$	0.206*** (6.11)	0.310*** (6.79)	0.152*** (5.64)	0.217*** (6.19)	0.089*** (4.49)	0.112*** (3.42)	0.464*** (3.20)	0.690*** (4.01)
R^2	0.2358	0.3060	0.2306	0.2938	0.2356	0.2958	0.4610	0.5467
No. of Observations	8,552	8,456	7,783	7,703	7,099	7,020	2,452	2,272
Bank FE	N	Y	N	Y	N	Y	N	Y
Time FE	Y	Y	Y	Y	Y	Y	Y	Y

The table presents results from regressions of change in non-performing loans on *Overreaction*. Nonperforming loan is defined as the sum of non-accrual loans and loans past due 90+ days divided by total loans at the beginning of the period multiplied by 100. $Overreaction_{it-k,t}$ measures the extent of overreaction during the last $k \in (1, 2, 3)$ years as explained in Section 4. The explanatory variable is in standard deviation relative to its mean. Regressions are estimated using Correia (2016) multilevel panel fixed effect estimator. Standard errors are dually clustered on bank and year. t-statistics in parenthesis. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table VIII: Relationship between *Overreaction* and Profitability

	$ROA_{t+1} - ROA_t$		$ROA_{t+2} - ROA_{t+1}$		$ROA_{t+3} - ROA_{t+2}$		$ROA_{t+3} - ROA_t$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. *Overreaction over last one year*								
$Overreaction_{it-1,\,t}$	-0.167^{***}	-0.271^{***}	-0.09^{***}	-0.14^{***}	-0.05^{***}	-0.06^{**}	-0.338^{***}	-0.482^{***}
	(-9.48)	(-12.32)	(-5.10)	(-5.71)	(-2.82)	(-2.37)	(-4.06)	(-3.49)
R^2	0.1524	0.2211	0.1310	0.1815	0.1328	0.1752	0.4045	0.4858
No. of Observations	10,915	10,838	10,003	9,941	9,133	9,037	2,887	2,658
B. *Overreaction over last two years*								
$Overreaction_{it-2,\,t}$	-0.139^{***}	-0.228^{***}	-0.075^{***}	-0.104^{***}	-0.033^{**}	-0.027	-0.248^{**}	-0.376^{**}
	(-7.60)	(-8.78)	(-3.82)	(-3.72)	(-2.05)	(-1.19)	(-2.82)	(-2.59)
R^2	0.1523	0.2064	0.1368	0.1791	0.1331	0.1736	0.3821	0.4629
No. of Observations	9,665	9,601	8,847	8,750	8,059	7,980	2,675	2,460
C. *Overreaction over last three years*								
$Overreaction_{it-3,\,t}$	-0.110^{***}	-0.190^{**}	-0.054^{***}	-0.076^{**}	-0.018	-0.016	-0.188^{**}	-0.281^{*}
	(-5.43)	(-5.70)	(-2.90)	(-2.55)	(-1.34)	(-0.68)	(-2.47)	(-2.08)
R^2	0.1527	0.2003	0.1358	0.1790	0.1335	0.1728	0.3709	0.4416
No. of Observations	8,557	8,462	7,814	7,735	7,147	7,071	2,469	2,282
Bank FE	N	Y	N	Y	N	Y	N	Y
Time FE	Y	Y	Y	Y	Y	Y	Y	Y

The table presents results from the regressions of change in return on assets on *Overreaction*. Return on asset is defined as net income divided by total assets multiplied by 100. $Overreaction_{it-k,t}$ measures the extent of overreaction during the last $k \in (1, 2, 3)$ years as explained in Section 4. The explanatory variable is in standard deviation relative to its mean. Regressions are estimated using Correia (2016) multilevel panel fixed effect estimator. Standard errors are dually clustered on bank and year. t-statistics in parenthesis. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

4.3. Do Shareholders Anticipate the Poor Performance of Banks with higher $\mathbf{Overreaction_{it}}$?

Next, we examine whether shareholders recognize the risks banks take and anticipate the future poor performance. If so, the shareholders would demand higher average returns from holding such banks' stocks. On the other hand, when shareholders fail to recognize the risks, the future stock prices would correct downwards, resulting in lower returns.

First, we sort banks into groups based on their extent of overreaction during the last three years, or $Overreaction_{it-3,\,t}$. To avoid any look-ahead bias, we calculate percentile thresholds using only past information at each point in time. For $Overreaction_{it-3,\,t} \in (< 25\%)$, $Overreaction_{it-3,\,t}$ in that quarter must be less than the 25th percentile threshold of all previous observations. Further, we construct one-, two-, and three-year ahead returns by taking price returns, and adding in any dividend yield over the corresponding period. We attempt to predict returns by running a regression using a series of dummy variables for each percentile threshold[25]. Figure 9 plots predicted equity returns conditional on $Overreaction_{it-3,t}$. As shown in Figure 9, an increase in $Overreaction_{it-3,t}$ predicts a decline in equity return in the future. To establish this relationship formally, we estimate the regression of future returns on past $Overreaction_{it}$ using the following specification:

$$r_{it,t+j} = \beta_0 + \beta^{eqt} * Overreaction_{it-k,\,t} + \delta_t + \zeta_i + \varepsilon_{it} \tag{11}$$

where the dependent variable is the equity return from quarter 4 in year t to quarter 4 in year $t+j$, with j being 1, 2, or 3 years ahead. The variable of interest is $Overreaction$ during the last $k \in (1,2,3)$ years or $Overreaction_{it-k,\,t}$. We add $Overreaction_{it}$, estimated quarterly using appropriate lag operator to measure $Overreaction_{it-k,\,t}$[26]. The empirical specification includes both bank and time fixed effects; hence any effects we observe are due to within bank variations in overreaction after controlling for the general economic environment each year. To avoid potential biases in com-

[25] More specifically, we estimate the following empirical specification: $r_{it,t+j} = \beta_0 + \beta^{eqt} 1_{[Overreaction_{it-3,t} \in x]} + \delta_t + \zeta_i + \varepsilon_{it}$ with x representing the percentile threshold.

[26] We limit $Overreaction_{it-k,\,t}$ to the past three years as previous studies have shown that the second and third lags of credit growth have the greatest predictive power for subsequent equity returns. Baron and Xiong (2017) have shown that the greatest predictive power for subsequent equity returns comes from the second and third lags in the one-year change in bank credit to GDP. Schularick and Taylor (2012) also finds similar results for the greatest predictability of future financial crises with the second and third one-year lags of bank credit growth. This finding sheds light on the timing of distress, which generally seems to take place at a one- to three-year horizon after the neglect of risks.

puting test statistics, we consider only non-overlapping returns in our regression analysis[27]. That is, in regressing one-, two-, or three-year ahead returns, we drop the intervening observations from our data set, in effect estimating the regressions on an annual, biennial, or triennial data. Standard errors are dually clustered on bank and time.

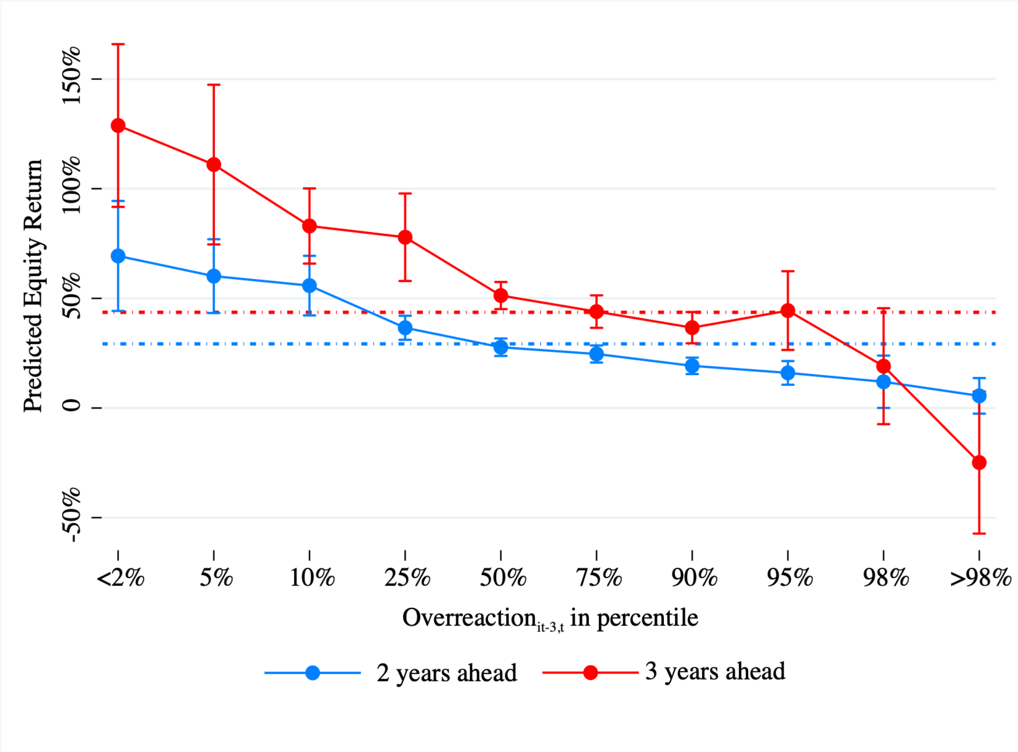


Figure 9

Predicted Equity Return: The plot shows predicted two- and three-year ahead equity returns with 95% confidence intervals conditional on overreaction during the last three years or $Overreaction_{it-3,t}$ belonging to a given percentile group. Average returns conditional on the thresholds are computed with non-overlapping returns by running a regression using a series of dummy variables for each percentile threshold. Standard errors are dually clustered. The dashed lines in blue and red shows average two- and three-year ahead returns respectively.

Tables IX, X, and XI present the results of our regression analysis, where we estimate the effect of $Overreaction$ over the last 1, 2, and 3 years respectively by varying $Overreaction_{it-k,t}$ from last year $(k = 1)$ to further into the past $(till\ k = 3)$. Columns (1) to (3) in each table correspond to the one-year-ahead return, columns (4) to (6) correspond to the two-year-ahead return, and columns (7) to (9) correspond to the three-year-ahead return. To aid interpretation, we report the coefficients in

[27] For example, we look at equity return from 1988Q4 to 1989Q4, and so on for one year; from 1988Q4 to 1990Q4, and so on for two years and from 1988Q4 to 1991Q4, and so on for three years. See Appendix A.

terms of the effect of a one-standard deviation increase in $Overreaction$ on future equity returns. The results in column (1), (4) and (7) of Table IX show that a one-standard deviation improvement in overreaction during the last year, $Overreaction_{it-1,t}$ predicts a decline in subsequent one-, two-, and three-year returns by 7.5pp, 14.4pp, and 29.8pp, respectively. These effects remain statistically and economically significant when we use overreaction during the last two and three years, as shown in columns (1), (4), and (7) of Tables X and XI, respectively. Specifically, a one-standard deviation increase in $Overreaction_{it-2,t}$ predicts a decline in subsequent one-, two-, and three-year-ahead excess return by 8.3pp, 15.8pp, and 30.4pp, respectively. Similarly, for overreaction during the last three years, $Overreaction_{it-3,t}$, a one-standard deviation increase predicts a decline in subsequent one-, two-, and three-year-ahead return by 7.0pp, 14.8pp, and 27.1pp, respectively.

We further relate the return predictability of $Overreaction_{it}$ to that of credit growth, as the predictability of credit growth for equity returns is acknowledged by the literature (Baron and Xiong, 2017; Fahlenbrach et al., 2017) as a reflection of shareholders' failure to recognize the risks. We include $\Delta credit_{it-k,t}$ or credit growth over the past $k \in (1,2,3)$ years as an explanatory variable and find that it does not significantly affect the economic and statistical significance of $Overreaction$. Rather, $Overreaction_{it}$, in all cases, predicts a sharper decline in future returns than credit growth.

Taken together, the results in Section 4.1, 4.2 and 4.3 let us conclude that presence of overreaction in banks' expectations affects the quantity and quality of credit. In good times, banks become too optimistic and engage in relatively risky lending. Despite the neglect of risks by banks, shareholders do not seem to factor that in their expectations of future stock returns. When the risks are realized in future and banks start to under-perform, stock prices decline, generating lower returns.

Table IX: *Overreaction* and Equity Return

	1-year returns			2-year returns			3-year returns		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. *Overreaction during last one year*									
$Overreaction_{it-1,t}$	-0.075^{***} (−5.77)		-0.071^{***} (−5.41)	-0.144^{***} (−5.12)		-0.132^{***} (−5.13)	-0.298^{***} (−9.09)		-0.275^{***} (−8.57)
$\Delta credit_{it-1,t}$		-0.024^{***} (−4.78)	-0.014^{*} (−3.21)		-0.058^{***} (−3.33)	-0.034^{**} (−2.37)		-0.126^{***} (−3.20)	-0.061^{**} (−2.76)
R^2	0.4440	0.4098	0.4450	0.4827	0.4640	0.4852	0.5649	0.5028	0.5682
$Adj\ R^2$	0.3765	0.3427	0.3775	0.3650	0.3464	0.3678	0.4294	0.3520	0.4333
No. of Observations	8208	9951	8208	3485	4368	3485	1925	2471	1925
Bank FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

The table reports result from the regression model specified in Equation (11). The dependent variable is equity return which is regressed on $Overreaction_{it-1,t}$ and $\Delta credit_{it-1,t}$. $Overreaction_{it-1,t}$ measures the extent of overreaction during the last one year as explained in Section 4.$\Delta credit_{it-1,t}$ measures credit growth over last one year. Explanatory variables are in standard deviation units. Returns are non-overlapping at one-, two-, and three-year-ahead horizons. A coefficient of -0.075 means that a one standard deviation improvement in $Overreaction_{it-1,t}$ predicts a 7.5pp decline in subsequent return. Regressions are estimated using Correia (2016) multilevel panel fixed effect estimator. t-statistics in parentheses are computed from standard errors dually clustered on bank and time. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table X: *Overreaction* and Equity Return

	1-year returns			2-year returns			3-year returns		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
B. ***Overreaction during last two years***									
$Overreaction_{it-2,t}$	-0.083^{***} (-6.30)		-0.077^{***} (-6.01)	-0.158^{***} (-5.73)		-0.148^{***} (-5.48)	-0.304^{***} (-7.98)		-0.279^{**} (-7.35)
$\Delta credit_{it-2,t}$		-0.030^{**} (-3.83)	-0.017^{**} (-2.45)		-0.061^{**} (-2.76)	-0.028^{*} (-1.92)		-0.117^{***} (-3.82)	-0.06^{**} (-2.54)
R^2	0.4493	0.4265	0.4504	0.4922	0.4573	0.4936	0.5590	0.5282	0.5619
$Adj\ R^2$	0.3797	0.3598	0.3809	0.3780	0.3378	0.3794	0.4218	0.3857	0.4251
No. of Observations	7397	9239	7397	3229	3919	3229	1809	2149	1809
Bank FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

The table reports result from the regression model specified in Equation (11). The dependent variable is equity return which is regressed on $Overreaction_{it-2,t}$ and $\Delta credit_{it-2,t}$. $Overreaction_{it-2,t}$ measures the extent of overreaction during the last two years as explained in Section 4.$\Delta credit_{it-2,t}$ measures credit growth over last two years. Explanatory variables are in standard deviation units. Returns are non-overlapping at one-, two-, and three-year-ahead horizons. A coefficient of -0.083 means that a one standard deviation improvement in $Overreaction_{it-2,t}$ predicts a 8.3pp decline in subsequent returns. Regressions are estimated using Correia (2016) multilevel panel fixed effect estimator. t-statistics in parentheses are computed from standard errors dually clustered on bank and time. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table XI: *Overreaction* and Equity Return

	1-year returns			2-year returns			3-year returns		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
C. *Overreaction during last three years*									
$Overreaction_{it-3,\,t}$	-0.070^{***}		-0.065^{***}	-0.148^{***}		-0.142^{***}	-0.271^{***}		-0.257^{***}
	(-6.55)		(-5.97)	(-5.72)		(-5.42)	(-5.49)		(-5.10)
$\Delta credit_{it-3,\,t}$		-0.031^{***}	-0.014^{**}		-0.065^{***}	-0.014		-0.105^{**}	-0.036
		(-3.88)	(-2.20)		(-3.57)	(-1.10)		(-3.16)	(-1.17)
R^2	0.4351	0.4230	0.4359	0.4574	0.4522	0.4578	0.5502	0.5187	0.5512
$Adj\ R^2$	0.3638	0.3541	0.3646	0.3360	0.3354	0.3362	0.4093	0.3742	0.4102
No. of Observations	6625	8536	6625	2801	3705	2801	1703	2054	1703
Bank FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Time FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

The table reports result from the regression model specified in Equation (11). The dependent variable is equity return which is regressed on $Overreaction_{it-3,\,t}$ and $\Delta credit_{it-3,\,t}$. $Overreaction_{it-3,t}$ measures the extent of overreaction during the last three years as explained in Section 4.$\Delta credit_{it-3,\,t}$ measures credit growth over last three years. Explanatory variables are in standard deviation units. Returns are non-overlapping at one-, two-, and three-year-ahead horizons. A coefficient of -0.070 means that a one standard deviation improvement in $Overreaction_{it-3,\,t}$ predicts a 7.0pp decline in subsequent returns. Regressions are estimated using Correia (2016) multilevel panel fixed effect estimator. t-statistics in parentheses are computed from standard errors dually clustered on bank and time. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

5. Conclusion

Using the construct of expected loss from loan loss provisions, we study banks' expectations. We show that the expectations in the banking sector of the US display overreaction to current conditions, in the sense of expectations being too optimistic in good times and too pessimistic in bad times.

We trace the distortion in expectations to being caused by an overreaction to current conditions. We provide evidence supporting the relatively contentious view that emphasizes the role of irrational beliefs as an important driver of credit expansion. Overreaction to fundamentals leads banks to register excessive credit growth, which eventually leads to a deterioration in the quality of their loan portfolios and a decline in profitability. The shareholders fail to anticipate the future performance of such banks and when the ignored risks finally materialize, such banks' stocks tend to under-perform, generating a decline in returns.

Our measure of overreaction in expectations - which manifests itself through supply-side - proposes a mechanism that adds to the literature explaining why credit expansion is a leading indicator of financial crises ((Schularick and Taylor, 2012; Jordà et al., 2013)). It also helps explain the puzzle why excess credit growth leads to future under-performance of banking stocks ((Baron and Xiong, 2017; Fahlenbrach et al., 2017)). Relatedly, our findings also raise doubts over the proposals for relying on market-based measures as precursors to a banking sector crisis since market does not seem to price in the risks until it's too late. Future research could propose measures that factor in banks' and investors' expectations to complement the market-based measures.

References

Acharya, Viral V. and Stephen G. Ryan (2016). "Banks' Financial Reporting and Financial System Stability". *Journal of Accounting Research* 54.2, pp. 277–340.

Acharya, Viral V. and Tanju Yorulmazer (2007). "Too many to fail—An analysis of time-inconsistency in bank closure policies". *Journal of Financial Intermediation* 16.1, pp. 1–31.

Afrouzi, Hassan, Spencer Yongwook Kwon, Augustin Landier, Yueran Ma, and David Thesmar (2021). "Overreaction in Expectations: Evidence and Theory".

Ahmed, Anwer S., Carolyn Takeda, and Shawn Thomas (1999). "Bank loan loss provisions: a re-examination of capital management, earnings management and signaling effects". *Journal of Accounting and Economics* 28.1, pp. 1–25.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer (2015). "X-CAPM: An extrapolative capital asset pricing model". *Journal of financial economics* 115, pp. 1–24.

Baron, Matthew and Wei Xiong (2017). "Credit Expansion and Neglected Crash Risk". *Quarterly Journal of Economics* 2, pp. 1283–1288.

Basu, Sudipta and Stanimir Markov (2004). "Loss function assumptions in rational expectations tests on financial analysts' earnings forecasts". *Journal of Accounting and Economics* 38, pp. 171–203.

Basu, Sudipta, Justin Vitanza, and Wei Wang (2020). "Asymmetric loan loss provision models". *Journal of Accounting and Economics* 70.2, p. 101359.

Beatty, Anne and Scott Liao (2011). "Do delays in expected loss recognition affect banks' willingness to lend?" *Journal of Accounting and Economics* 52.1, pp. 1–20.

– (2014). "Financial Accounting in the Banking Industry: A Review of the Empirical Literature". *Journal of Accounting and Economics* 58, pp. 339–383.

Beaver, William H. and Ellen E. Engel (1996). "Discretionary behavior with respect to allowances for loan losses and the behavior of security prices". *Journal of Accounting and Economics*.

Bischof, Jannis, Christian Laux, and Christian Leuz (2021). "Accounting for financial stability: Bank disclosure and loss recognition in the financial crisis". *Journal of Financial Economics*.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer (Sept. 2020). "Overreaction in Macroeconomic Expectations". *American Economic Review* 110.9, pp. 2748–82.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer (2019). "Diagnostic expectations and stock returns". *Journal of Finance* 74, pp. 2839–2874.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer (2018). "Diagnostic Expectations and Credit Cycles". *Journal of Finance* 73.1, pp. 199–227.

Bordalo, Pedro, Nicola Gennaioli, Andrei Shleifer, and Stephen J. Terry (Dec. 2021). "Real Credit Cycles". *NBER Working paper*.

Bushman, Robert M. and Christopher D. Williams (June 2015). "Delayed Expected Loss Recognition and the Risk Profile of Banks". *Journal of Accounting Research* 53.3, pp. 511–553.

Cagan, Phillip (1956). "The monetary dynamics of hyperinflation". *Studies in the Quantity Theory of Money*.

Calomiris, Charles W and Joseph R. Mason (Dec. 2003). "Fundamentals, Panics, and Bank Distress During the Depression". *American Economic Review* 93.5.

Calomiris, Charles W and Berry Wilson (2004). "Bank Capital and Portfolio Management: The 1930s "Capital Crunch" and the Scramble to Shed Risk". *The Journal of Business* 77.

Cheng, Ing-Haw, Harrison Hong, and Jose A. Scheinkman (2015). "Yesterday's Heroes: Compensation and Risk at Financial Firms". *The Journal of Finance* 70.2, pp. 839–879.

Coibion, Olivier and Yuriy Gorodnichenko (Aug. 2015). "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts". *American Economic Review* 105.8, pp. 2644–78.

Correia, Sergio (2016). *Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator*. Tech. rep. Working Paper.

d'Arienzo, Daniele (2020). "Maturity increasing overreaction and bond market puzzles". *Working paper.*

Dell'Ariccia, Giovanni and Robert Marquez (2006). "Lending Booms and Lending Standards". *Journal of Finance* 61.5, pp. 2511–2546.

Fahlenbrach, Rüdiger, Robert Prilmeier, and René M Stulz (Nov. 2017). "Why Does Fast Loan Growth Predict Poor Performance for Banks?" *Review of Financial Studies* 31.3, pp. 1014–1063.

Francis, Jennifer, Ryan LaFond, Per Olsson, and Katherine Schipper (2005). "The market pricing of accruals quality". *Journal of Accounting and Economics* 39.2, pp. 295–327.

Gebhardt, Günther (May 2016). "Impairments of Greek Government Bonds under IAS 39 and IFRS 9: A Case Study". *Accounting in Europe* 13.2, pp. 169–196.

Gennaioli, Nicola, Yueran Ma, and Andrei Shleifer (2016). "Expectations and Investment".

Gennaioli, Nicola and Andrei Shleifer (2010). "What comes to mind". *The Quarterly journal of economics* 125.4, pp. 1399–1433.

Greenwood, Robin and Samuel G. Hanson (Apr. 2013). "Issuer Quality and Corporate Bond Returns". *Review of Financial Studies* 26.6, pp. 1483–1525.

Greenwood, Robin, Samuel G. Hanson, Andrei Shleifer, and Jakob AHM Sorensen (2022). "Predictable Financial Crises". *The Journal of Finance* 77.2, pp. 863–921.

Greenwood, Robin and Andrei Shleifer (2014). "Expectations of returns and expected returns". *Review of Financial Studies* 27.3, pp. 714–746.

Hagendorff, Jens, Anthony Saunders, Sascha Steffen, and Francesco Vallascas (2021). "The wolves of Wall Street? Managerial attributes and bank risk". *Journal of Financial Intermediation* 47, p. 100921.

Hennes, Karen M, Andrew J Leone, and Brian P Miller (2008). "The importance of distinguishing errors from irregularities in restatement research: The case of restatements and CEO/CFO turnover". *The Accounting Review* 83.6, pp. 1487–1519.

Hirshleifer, David, Jun Li, and Jianfeng Yu (2015). "Asset pricing in production economies with extrapolative expectations". *Journal of Monetary Economics* 76, pp. 87–106.

Hirtle, Beverly, Anna Kovner, and Matthew Plosser (2020). "The impact of supervision on bank performance". *The Journal of Finance* 75.5, pp. 2765–2808.

Hribar, Paul, Samuel J. Melessa, R. Christopher Small, and Jaron H. Wilde (2017). "Does managerial sentiment affect accrual estimates? Evidence from the banking industry". *Journal of Accounting and Economics* 63.1, pp. 26–50.

Jiménez, Gabriel, Steven Ongena, José-Luis Peydró, and Jesús Saurina (2017). "Macroprudential Policy, Countercyclical Bank Capital Buffers, and Credit Supply: Evidence from the Spanish Dynamic Provisioning Experiments". *Journal of Political Economy* 125.6, pp. 2126–2177.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor (2013). "When credit bites back". *Journal of Money, Credit and Banking* 45, pp. 3–28.

Kindleberger, Charles Poor (1978). *Manias, panics, and crashes: a history of financial crises*. Basic books.

Kiridaran, Kanagaretnam, Lobo Gerald J., and Mathieu Robert (2004). "Earnings Management to Reduce Earnings Variability: Evidence from Bank Loan Loss Provisions". *Review of Accounting and Finance* 3.3, pp. 128–148.

Kohlhas, Alexandre N. and Ansgar Walther (Sept. 2021). “Asymmetric Attention”. *American Economic Review* 111.9, pp. 2879–2925.

Kothari, S.P., Andrew J. Leone, and Charles E. Wasley (2005). “Performance matched discretionary accrual measures”. *Journal of Accounting and Economics* 39.1, pp. 163–197.

Liu, Chi-Chun and Stephen G. Ryan (1995). “The effect of bank loan portfolio composition on the market reaction to and anticipation of loan loss provisions”. *Journal of Accounting Research* 33.1, pp. 77–94.

– (2006). “Income Smoothing over the Business Cycle: Changes in Banks’ Coordinated Management of Provisions for Loan Losses and Loan Charge-Offs from the Pre-1990 Bust to the 1990s Boom”. *The Accounting Review* 81.2.

López-Salido, David, Jeremy C Stein, and Egon Zakrajšek (2017). “Credit-Market Sentiment and the Business Cycle.” *Quarterly Journal of Economics* 132.3, pp. 1373–1426.

Mian, Atif, Amir Sufi, and Emil Verner (May 2017). “Household Debt and Business Cycles Worldwide*”. *The Quarterly Journal of Economics* 132.4, pp. 1755–1817.

Minsky, Hyman P (Mar. 1977). “The Financial Instability Hypothesis: An Interpretation of Keynes and an Alternative to“Standard” Theory”. *Challenge* 20.1, pp. 20–27.

Pflueger, Carolin, Emil Siriwardane, and Adi Sunderam (Mar. 2020). “Financial Market Risk Perceptions and the Macroeconomy*”. *The Quarterly Journal of Economics* 135.3, pp. 1443–1491.

Powell, David (2016). “Quantile regression with nonadditive fixed effects”. 1, pp. 1–33.

Rajan, Raghuram G. (May 1994). “Why Bank Credit Policies Fluctuate: A Theory and Some Evidence*”. *The Quarterly Journal of Economics* 109.2, pp. 399–441.

– (Nov. 2005). “Has Financial Development Made the World Riskier?” NBER Working Paper 11728.

Schularick, Moritz and Alan M. Taylor (Apr. 2012). “Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008”. *American Economic Review* 102.2.

Shleifer, Andrei (2012). “Psychologists at the gate: a review of Daniel Kahneman’s thinking, fast and slow”. *Journal of Economic Literature* 50.4, pp. 1080–91.

Woodford, Michael (2003). “Imperfect common knowledge and the effects of monetary policy”. *Information, and Expectations in Modern Macroeconomics*.

Appendix

A. Details on Data and Summary Statistics

The section presents additional information related to data sources and construction of variables beyond what is described in Section 1. We list all the variables and their sources in Table A.1. Summary statistics are given in Table A.2.

1. Credit Growth $(\Delta credit_{it-k,t})$*:* We use total loans and leases reported by bank holding companies as our measure of bank credit. Credit growth is constructed for the last one-, two, and three years from the change in credit. Mathematically it can be expressed as,

$$\Delta credit_{it-k,t} = \left(\frac{Credit_{it}}{Credit_{it-k}} - 1 \right)$$

where $k \in (1, 2, 3)$. Average credit growth for last one, two and three year period comes out to be 12.38%, 27.28% and 44.43% respectively. To avoid potential biases in computing test statistics, we consider only non-overlapping credit growth in our regression analysis in Section 4.1. For example, we look at credit growth from 1988Q4 to 1989Q4, and so on for one year; from 1988Q4 to 1990Q4, and so on for two year and from 1988Q4 to 1991Q4, and so on for three year.

2. Performance Measures: To compare banks' financial performance, we use *NPA*, and *ROA*. *NPA* reflects the quality of a bank's loan portfolio and is defined as the ratio of non-performing loans (sum of loans past due 90 days and non-accrual loans) to total loans at the beginning of the quarter. *ROA* is defined as the ratio of net income to total assets.

3. Equity Return $(r_{it,t+j})$*:* We obtain bank equity data from *CRSP* database. As is standard in the literature (Fahlenbrach et al., 2017), we use the link table provided by the *New-York Fed* which provides a mapping between the regulatory identification numbers (RSSD ID) and CRSP's permanent company numbers (PERMCO). We construct one-, two-, and three-year ahead returns by taking price returns, and adding in dividend yield over the period. The mean returns $(r_{it,t+j})$ for one, two, and three year periods are 13.81%, 29.24% and 43.63% respectively. To avoid potential biases in computing test statistics, we consider only non-overlapping returns in our regression analysis in Section 4.3. For example, we look at returns from 1988Q4 to 1989Q4, and so on for one year; from 1988Q4 to 1990Q4, and so on for two year and from 1988Q4 to 1991Q4, and so on for three year.

3. Control Variables: We also employ several financial and macroeconomic variables, which are known to affect the performance of bank loans. In addition to bank-specific variables such as asset, interest income, loan loss allowance, charge-offs, we consider several macroeconomic variables: Real GDP growth, Unemployment rate, and Case-Shiller real estate index. All macroeconomic variables are sourced from the data provided by *FRED Economic Research*.

Table A.1: Variables Descriptions

Variable	Source	Definition
llp_{it}	FR Y-9C	Loan loss provision scaled by one over total loans at the beginning of the quarter
$ndllp_{it}$	Estimated	Non-discretionary component of loan loss provision scaled by one over total loans at the beginning of the quarter
nco_{it}	FR Y-9C	Net charge-off recorded in quarter t scaled by one over total loans at the beginning of the quarter
$Loan_{it}$	FR Y-9C	Size of loan portfolio (in billion dollar) in quarter t
NPA_{it}	FR Y-9C	Loans past due for 90 days or more and Non-accrual loans scaled by one over total loan balance at the beginning of the quarter
$size$	FR Y-9C	Log of assets (in thousand dollar) at the beginning of the quarter
$\Delta credit_{it-k,t}$	FR Y-9C	Change in total loans and leases over the last k years
ROA_{it}	FR Y-9C	Return on assets defined as net income divided by total assets
$IntRate_{it,t+j}$	FR Y-9C	Average interest income earned by the bank over the next j years scaled by one over total loans at the beginning of the quarter
$r_{it,t+j}$	CRSP	j years ahead return on bank equity

Table A.2: Summary Statistics

	Obs.	Mean	SD	1st perc.	25th perc.	Median	75th perc.	99th perc.
llp_{it}	63,359	0.0036	0.0056	-0.0018	0.0007	0.0018	0.0038	0.0355
nco_{it}	63,361	0.0029	0.0051	-0.0010	0.0003	0.0011	0.0031	0.0301
$size_{it}$	66,100	14.22	1.66	11.57	13.01	13.88	15.12	19.52
$Loan_{it}$	49,910	6.87	21.1	0.10	0.42	11.06	3.60	157
NPA_{it}	63,271	0.0163	0.0229	0.0000	0.0049	0.0095	0.0187	0.1061
$\Delta credit_{it-1,t}$	59,703	0.1238	0.1816	-0.2248	0.0235	0.0903	0.1806	0.9372
$\Delta credit_{it-2,t}$	54,635	0.2728	0.3418	-0.3475	0.0715	0.2060	0.3957	1.7102
$\Delta credit_{it-3,t}$	49,949	0.4443	0.5221	-0.4104	0.1301	0.3344	0.6257	2.7066
$IntRate_{it,t+1}$	59,278	0.0650	0.0247	0.0272	0.0454	0.0641	0.0797	0.1597
$IntRate_{it,t+2}$	54,132	0.0650	0.0244	0.0274	0.0454	0.0641	0.0794	0.1596
$IntRate_{it,t+3}$	49,288	0.0649	0.0239	0.0276	0.0467	0.0639	0.0791	0.1565
ROA_{it}	66,100	0.0054	0.0062	-0.0236	0.0026	0.0052	0.0085	0.0282
$r_{it,t+1}$	43,367	0.1381	0.3358	-0.6273	-0.0639	0.1204	0.3154	1.2967
$r_{it,t+2}$	38,416	0.2924	0.5286	-0.7029	-0.0410	0.2394	0.5417	2.5413
$r_{it,t+3}$	33,993	0.4363	0.6911	-0.7200	-0.0119	0.3467	0.7219	3.565

The table reports summary statistics of all variables used in the analysis. Loan loss provisions and net charge-offs are scaled by one over total loans at the beginning of the quarter. $\Delta credit_{it-k,t}$ measures loan growth over the last $k \in (1,2,3)$ years. $r_{it,t+j}$ measures j year-ahead equity return. All bank level variables are winsorized at 1% top and bottom to mitigate the influence of outliers.

B. Analysis of Alternative Loan Loss Provisioning Models and Test for Forecast Error Predictability

B.1. Extensions to the main text model:

In this section, we employ alternative models to estimate the non-discretionary component of loan loss provision $(ndllp)$ as discussed in Section 1.1. Using $ndllp$, we measure expected loss, calculate the forecast error and then, test its predictability.

Model 2

Model 2 resembles our preferred model in the main text (call it Model 1) in all aspects except for the vector of macroeconomic variables. Instead of using time-fixed effects to control for prevailing macroeconomic conditions, we use direct measures of macroeconomic conditions from Beatty and Liao (2014).

$$
\begin{aligned}
llp_{it} = \gamma_0 + \gamma_1 \Delta npa_{it} + \gamma_2 D\Delta npa_{it} * \Delta npa_{it} + \gamma_3 \Delta npa_{it-1} + \gamma_4 \Delta npa_{it-2} + \gamma_5 Size_{it-1} + \gamma_6 \Delta Loan_{it} \\
+ \gamma_7 nco_{it} + \gamma_8 \Delta GDP_t + \gamma_9 \Delta UNEMP_t + \gamma_{10} \Delta CSRET_t + \zeta_i + \eta_{it}
\end{aligned}
$$

where ΔGDP_t, $\Delta UNEMP_t$, and $\Delta CSRET_t$ respectively denote real GDP growth rate, change in unemployment rate, and growth in Case-Shiller Real Estate Index over the quarter.

Model 3

In Model 1 and 2, we have not considered any lead variables (such as NPA_{t+1}). In this section, we control for future non-performing loans as in Beatty and Liao (2014). Except for an additional control of future non-performing loans, npa_{it+1}, the following Model 3 and 4 map into Model 1 and 2, respectively

$$
\begin{aligned}
llp_{it} = \gamma_0 + \gamma_1 \Delta npa_{it+1} + \gamma_2 \Delta npa_{it} + \gamma_3 D\Delta npa_{it} * \Delta npa_{it} + \gamma_4 \Delta npa_{it-1} + \gamma_5 \Delta npa_{it-2} + \gamma_6 Size_{it-1} \\
+ \gamma_7 \Delta Loan_{it} + \gamma_8 nco_{it} + \delta_t + \zeta_i + \eta_{it}
\end{aligned}
$$

Model 4

$$llp_{it} = \gamma_0 + \gamma_1 \Delta npa_{it+1} + \gamma_2 \Delta npa_{it} + \gamma_3 D\Delta npa_{it} * \Delta npa_{it} + \gamma_4 \Delta npa_{it-1} + \gamma_5 \Delta npa_{it-2} + \gamma_6 Size_{it-1}$$
$$+ \gamma_7 \Delta Loan_{it} + \gamma_8 nco_{it} + \gamma_9 \Delta GDP_t + \gamma_{10} \Delta UNEMP_t + \gamma_{11} \Delta CSRET_t + \zeta_i + \eta_{it}$$

We use the same averaging procedure as in Section 1.2 to construct the forecast error and test its predictability. Table B.1 reports the results. We do not observe any qualitative difference in the predictability of the forecast error from Table I.

Table B.1: Evidence on Predictability of Forecast Error

	β_1			Median β_1^i		
	Model 2	Model 3	Model 4	Model 2	Model 3	Model 4
$Actual\ Loss_{iT}$	-0.399^{***}	-0.325^{***}	-0.397^{***}	-0.308^{***}	-0.335^{***}	-0.312^{***}
	(-19.49)	(-16.24)	(-19.68)	(0.003)	(0.001)	(0.002)
Bank F.E.	Yes	Yes	Yes			
Time F.E.	Yes	Yes	Yes			
Obs.	47,908	47,511	47,853	369	367	369
R^2	0.3340	0.3225	0.3369			
$Adj\ R^2$	0.3166	0.3047	0.3196			

Quarterly regression of errors in banks' expectations about future loan losses on past loan losses as specified in the regression model of Equation (1). Similar to Basu et al. (2020), we scale these state variables with the total loan book value of the previous period. Additionally, we follow a four-quarter window for averaging these scaled variables. The dependent variable $Forecast\ Error_{iT+1,T}$ is the difference between $Actual\ Loss_{iT+1}$ and $E_T Actual\ Loss_{iT+1}$ where $Actual\ Loss_{iT+1}$ is the average loan loss realized over the next four quarters from quarter $t+1$ to $t+4$. $E_T Actual\ Loss_{iT+1}$ is the non-discretionary component of loan loss provision $(ndllp)$ averaged over quarter $t-3$ to t. The first row reports β_1, the coefficient on $Actual\ Loss_{iT}$, defined as the average loan loss experienced by banks in the recent past, from quarter $t-3$ to t. Regressions are estimated using Correia (2016) multilevel panel fixed effect estimator. Standard errors are dually clustered both on bank and time; t-statistics is reported in parenthesis. In the last three columns, we report the median coefficient β_1^i from the bank by bank regression specification in Equation (2) and bootstrap standard errors in parenthesis; *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. The estimates from Model 3 and 4 are comparable to the estimates from Model 1 and 2 as in Table I.

B.2. Assessing the Models

We compare the four models used for estimating the non-discretionary component $(ndllp)$ of loan loss provision using simulations like those in Basu et al. (2020) and Kothari et al. (2005). The simulations assess each model based on: (i) the specification to falsely reject the null of no abnormal accruals (Type 1 error) and (ii) the power of tests to detect earnings management (Type 2 error).

Specification tests: We estimate each model's propensity to falsely reject the null of no abnormal accruals, i.e., Type I error, using randomized trials. We first estimate the model using the full sample and randomly select 100 bank-quarter observations. We then assess the significance of the mean discretionary loan loss provision $(Residuals)$ using a t-test. In the top panel of Table B.2, we report the percentage of times out of 250 simulated samples the null hypotheses of zero discretionary accruals are rejected against a positive or negative alternative at the 5% significance level. If models are well-specified, there is a 95% probability that the rejection rate lies between 2.4% and 8.0% for 250 trials. We observe that the rejection rates lie within the bounds for all four models.

Power of tests: To assess each model's power to detect earnings management, we randomly draw 100 bank-quarters from the full sample and simulate earnings management in the selected bank-quarters by adding positive or negative abnormal provisions that are 1, 3, 5, or 10 basis points (bps) of lagged loans. The indicated seed level is added to total accruals before estimating the respective discretionary accrual model using all observations. The process is repeated 250 times, and we perform one-tailed t-tests for abnormal provision in the seeded observations at a significance level of 5%. We report the percentage of times out of 250 simulated samples the null hypotheses of zero discretionary accruals are rejected against a positive or negative alternative.

While all models perform reasonably well based on Type I error, we find that Model 1 has the greatest power in detecting seeded abnormal provisions. The findings are consistent with Basu et al. (2020) and provide our justification for using Model 1 as the preferred model for our analysis.

Table B.2: A Comparison of the Type I and Type II error rates of alternative Loan Loss Provision models

		Ha	Model 1	Model 2	Model 3	Model 4
Type I Error		> 0	4.8	3.6	2.8	3.2
		< 0	4.8	4.0	4.8	4.8
Type II Error	(+1 bps)	> 0	12	10.8	9.6	8.8
	(+3 bps)	> 0	36.4	38.0	32.0	36.0.2
	(+5 bps)	> 0	66.8	68.4	66.0	63.6
	(+10 bps)	> 0	96.4	97.6	94.0	95.6
	(-1 bps)	< 0	13.6	12.8	10.0	10.8
	(-3 bps)	< 0	40.8	40.4	36.4	35.6
	(-5 bps)	< 0	63.6	62.4	59.6	58.4
	(-10 bps)	< 0	92.0	91.6	90.8	92.0

The table compares Type I and Type II error rates of alternative models used for separating the discretionary LLP. Column 'Ha' mentions the alternative hypothesis (discretionary LLP being $>$, or < 0)against the null hypothesis of zero discretionary LLP.

Type I: We randomly draw 100 bank-quarter observations from the full sample and test the null hypothesis of zero discretionary LLP. The trial is repeated 250 times. The top panel reports the frequency (percentage of 250 samples) with which the null hypothesis of zero discretionary accrual is rejected at the 5% level. If models are well specified and the null hypothesis is true, the rejection rates for 250 trials should be between 2.4% and 8% with 95% probability.

Type II: We randomly select 100 bank-quarter observations from the full sample and add the indicated seed level (e.g. $\pm 1bps$, $\pm 3bps$, $\pm 5bps$, $\pm 10bps$ of lagged total loans) to total loan loss provisions. The models are then estimated using the full sample. We repeat the trial 250 times. The bottom panel reports the detection rates of positive and negative seeded abnormal loan loss provisions using a one-tailed t-test at the 5% significance level. Model 1 and 2 emerge as the superior of the four models.

C. Some Concerns on the predictability of Forecast Error and Overreaction

C.1. Alternative Explanations: Robustness Exercises

Notwithstanding our results on the predictability of forecast errors, there could still be potential concerns that might suggest that the predictability arises from features of loan loss provisions unrelated to overreaction in expectations. For instance, there might be concerns on how well-specified the regression models are and in particular, whether the observed predictability is an outcome of some rational response. Our response is by way of counterfactual estimation. Suppose that the predictability in forecast error is being entirely driven by other potential mechanisms that remain misspecified in the current setup. Then, once such concerns are controlled for, the predictability should at the very least weaken, if not disappear. We do a series of such counterfactual exercises for each for the following concerns and find no statistical support for any of them.

Alternative Case 1: Level of Past Loan Loss Allowance: One potential concern is that current provisions are not only a measure of expected losses, but also a reflection of how wrong the bank was in its past provisions. For instance, if the provisions were too high in the past, then the current provisions are likely to be lower. This can create a model error that might mechanically drive the countercyclical relationship between forecast errors and past loan losses. We address this by controlling for banks' past loan loss allowance in the regression specification of Equation (1). Allowance is averaged over the last four quarters, from quarter $t-4$ to $t-1$. First four columns of Table C.1 document that the negative relationship between the forecast error and past loan losses is significant even after controlling for past allowance level.

The accounting item on loan loss allowance also opens up another possibility. Unlike loan provisions, which is an income-statement item, the allowance is a balance sheet item and is, arguably, less susceptible to quarter-by-quarter fluctuations. We replicate our analysis of Equation (1), only this time we use allowance as a measure of expected loss. Specifically, we construct $E_T Actual\ Loss_{iT+1}$ as the average of loan loss allowance from quarter $t-3$ to t and calculate forecast error. We find qualitatively similar results. The last column of Table C.1 shows that $\beta_1 < 0$ continues to remain statistically significant.

Table C.1: Bank Level Evidence on Predictability of Forecast Error

	β_1				
	Model 1	Model 2	Model 3	Model 4	Loan Loss Allowance
$Actual\ Loss_{iT}$	-0.369^{***} (-14.24)	-0.440^{***} (-16.85)	-0.370^{***} (-14.50)	-0.441^{***} (-17.17)	-0.325^{***} (-8.03)
Alw_{iT-1}	0.044^{***} (4.12)	0.041^{***} (3.99)	0.046^{***} (4.41)	0.044^{***} (4.31)	
Bank F.E.	Yes	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes	Yes
Obs.	40,986	41,386	40,993	41,347	46,176
R^2	0.3394	0.3534	0.3475	0.3572	0.6091
$Adj\ R^2$	0.3198	0.3344	0.3281	0.3382	0.5981

Quarterly regression of errors in banks' expectations about future loan losses on past loan losses as specified in the regression model of Equation (1). Similar to Basu et al. (2020), we scale these state variables with the total loan book value of the previous period. Additionally, we follow a four-quarter window for averaging these scaled variables. The dependent variable $Forecast\ Error_{iT+1,T}$ is the difference between $Actual\ Loss_{iT+1}$ and $E_T Actual\ Loss_{iT+1}$ where $Actual\ Loss_{iT+1}$ is the average loan loss realized over the next four quarters from quarter $t+1$ to $t+4$. $E_T Actual\ Loss_{iT+1}$ is the non-discretionary component of loan loss provision $(ndllp)$ averaged over quarter $t-3$ to t. Additionally, we control for past loan loss allowance by using Alw_{iT-1}, which denotes average loan loss allowance over the last four quarters, from quarter $t-4$ to $t-1$. For the last column, we run the same regression without controlling for past allowance Alw_{iT-1}, but construct $E_T Actual\ Loss_{iT+1}$ as the average of loan loss allowance from quarter $t-3$ to t. The first row reports β_1, the coefficient on $Actual\ Loss_{iT}$, defined as the average loan loss experienced by banks in the recent past, from quarter $t-3$ to t. Regressions are estimated using Correia (2016) multilevel panel fixed effect estimator. Standard errors are dually clustered both on bank and time; t-statistics is reported in parenthesis; *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Alternative Case 2: Capital Management: Another possible concern could be that the negative relationship between the forecast error and past loan losses need not reflect overreaction, but regulatory capital management under asymmetric cost of provisions. Recognizing that every unit of provision is one less unit available for regulatory capital, the cost of provisioning naturally varies across the business cycle. During periods of high credit demand - which are typically concomitant with low charge-offs - a bank under capital distress may rationally choose to delay the recognition of loan losses to maintain competitive levels of regulatory capital (see Calomiris and Mason (2003) and Calomiris and Wilson (2004)). Under such situation, the shadow cost of provision is higher and therefore, a bank is likely to underprovision. In contrast, during periods of low credit demand and high loan losses, as the bank's lending becomes less sensitive to its capital position, it may rationally choose to keep higher provisions. Together, this might drive the predictability of the forecast error.

Table C.2: Predictability of Forecast Error under Capital Constraint Management

	β_1			
	Model 1	Model 2	Model 3	Model 4
$Actual\ Loss_{iT}$	-0.211^{***} (−6.03)	-0.281^{***} (−7.90)	-0.213^{***} (−6.23)	-0.283^{***} (−8.19)
$Actual\ Loss_{iT} * ExRegCap_{iT-1}$	-2.28^{***} (−6.41)	-2.34^{***} (−6.56)	-2.25^{***} (−6.38)	-2.29^{***} (−6.53)
Bank F.E.	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes
Obs.	23,191	23,610	23,215	23,581
R^2	0.3555	0.3741	0.3639	0.3780
$Adj\ R^2$	0.3331	0.3527	0.3417	0.3567

Quarterly regression of errors in banks' expectations about future loan losses on past loan losses as specified in the regression model of Equation (1). Similar to Basu et al. (2020), we scale these state variables with the total loan book value of the previous period. We follow a four-quarter window for averaging these scaled variables. The dependent variable $Forecast\ Error_{iT+1,T}$ is the difference between $Actual\ Loss_{iT+1}$ and $E_T Actual\ Loss_{iT+1}$ where $Actual\ Loss_{iT+1}$ is the average loan loss realized over the next four quarters from quarter $t+1$ to $t+4$. $E_T Actual\ Loss_{iT+1}$ is the non-discretionary component of loan loss provision $(ndllp)$ averaged over quarter $t-3$ to t. The first row reports β_1, the coefficient on $Actual\ Loss_{iT}$, defined as the average loan loss experienced by banks in the recent past, from quarter $t-3$ to t. We augment the regression specification with $Actual\ Loss_{iT} * ExRegCap_{iT-1}$, where $ExRegCap_{iT-1}$ denotes excess tier 1 capital bank i holds over the regulatory minimum, averaged over the last four quarters from quarter $t-4$ to quarter $t-1$. Regressions are estimated using Correia (2016) multilevel panel fixed effect estimator. Standard errors are dually clustered both on bank and time; t-statistics is reported in parenthesis; *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

We exploit the variation in the regulatory capital position across banks with the hypothesis that the predictability of forecast error should be weaker for banks with a higher level of regulatory capital. We repeat our regression of forecast errors on past loan losses and additionally interact past loan losses with excess regulatory capital. We define $ExRegCap$ as the difference between the regulatory capital to risk-weighted asset ratio and minimum capital requirement, averaged over the last four quarters from quarter $t-4$ to quarter $t-1$. To the extent that concern about capital drives the provisioning behavior and predictability of forecast errors, the effect should get attenuated as a bank's capital position improves. And if that is the case, then one would expect a positive coefficient for the interaction term. However, as reported in Table C.2, the coefficient for the interaction term $Actual\ Loss * ExRegCap$ is negative, contrary to what was hypothesized.

Alternative Case 3: Regulatory Scrutiny: An additional explanation for the observed negative relationship between forecast errors and past loan losses could be the variation in regulatory scrutiny. Typically, during periods of low charge-offs, regulatory scrutiny tends to be relaxed, which may encourage banks to inflate earnings by underprovisioning. Conversely, after periods of high charge-offs, regulatory scrutiny may tighten, leading banks to err on the side of overprovisioning.

To rule out this explanation, we propose a testable hypothesis by leveraging the differences in regulatory scrutiny across banks. Regulatory scrutiny can vary across the banking sector, with larger bank holding companies facing *extra* scrutiny, as they pose a greater degree of systemic risk[28]. In a recent paper, Hirtle et al. (2020) establishes that large banks, in each federal reserve district, receive a disproportional amount of supervisory attention. As a result, large banks are unlikely to be driven by earnings management of provisioning.

We create two dummy variables. Our first dummy variable, *Large*, as per Hirtle et al. (2020), indicates whether a bank's asset (averaged from quarter $t-3$ to t) are in the top five or within 25% of the assets of the fifth largest bank in the federal reserve district. We also use an alternative definition, where banks, whose assets (averaged from quarter $t-3$ to t) are more than the $95th$ percentile threshold in the corresponding quarter are considered *Large*. Our second dummy variable, $Reg\ Scrutiny_T$ indicates the periods in which the banking sector has experienced higher than the median of average loan loss, thus is hypothesized to be subjected to higher regulatory scrutiny in the corresponding quarter. We then use Equation (1) and interact past loan losses with these dummy variables. Specifically, we estimate the following:

$$\begin{aligned} Forecast\ Error_{iT+1,T} &= \beta_0 + \beta_1 Actual\ Loss_{iT} + \beta_2 Actual\ Loss_{iT} * Reg\ Scrutiny_T \\ &+ \beta_3 Actual\ Loss_{iT} * Large_{iT} + \beta_4 Actual\ Loss_{iT} * Reg\ Scrutiny_T * Large_{iT} + \nu_{iT+1,T} \end{aligned}$$

If the predictability of forecast error is due to variation in regulatory scrutiny and earnings management, then compared to small banks, this effect is likely to be weaker for large banks. This is especially true for large banks during bad periods when regulatory scrutiny increases. In other words, the ex-ante expectation on the coefficients for the interaction term, $Actual\ Loss * Reg\ Scrutiny * Large$

[28] See section 1060.0 in FRB Supervision Manual.

is $\beta_4 > 0$. However, as reported in Table C.3, β_4 turns out to be negative.

Table C.3: Predictability of Forecast Error under Regulatory Constraint Management

	Hirtle et al. (2020)			
	Model 1	Model 2	Model 3	Model 4
$Actual\ Loss_{iT}$	-0.452^{***}	-0.530^{***}	-0.448^{***}	-0.524^{***}
	(-10.09)	(-11.77)	(-10.22)	(-11.89)
$Actual\ Loss^{i}_{T} * Reg\ Scrutiny_T$	0.143^{***}	0.146^{***}	0.140^{***}	0.141^{***}
	(3.08)	(3.12)	(3.08)	(3.10)
$Actual\ Loss_{iT} * Large_{iT}$	0.164^{***}	0.182^{***}	0.163^{***}	0.178^{***}
	(3.47)	(3.82)	(3.47)	(3.79)
$Actual\ Loss_{iT} * Reg\ Scrutiny_T * Large_{iT}$	-0.180^{***}	-0.187^{***}	-0.177^{***}	-0.183^{***}
	(-3.92)	(-4.03)	(-3.90)	(-4.00)
Obs.	41,560	41,961	41,567	41,924
R^2	0.3380	0.3530	0.3452	0.3557
$Adj\ R^2$	0.3185	0.3341	0.3259	0.3369
	>95th percentile			
$Actual\ Loss_{iT}$	-0.451^{***}	-0.527^{***}	-0.448^{***}	-0.524^{***}
	(-9.64)	(-11.28)	(-10.22)	(-11.89)
$Actual\ Loss_{iT} * Reg\ Scrutiny_T$	0.132^{***}	0.135^{***}	0.128^{***}	0.130^{***}
	(2.74)	(2.79)	(2.73)	(2.75)
$Actual\ Loss_{iT} * Large_{iT}$	0.123^{*}	0.147^{**}	0.123^{**}	0.145^{**}
	(1.96)	(2.33)	(2.00)	(2.34)
$Actual\ Loss_{iT} * Reg\ Scrutiny_T * Large_{iT}$	-0.112^{***}	-0.125^{***}	-0.111^{***}	-0.183^{***}
	(-2.16)	(-2.37)	(-2.15)	(-4.00)
Obs.	37,247	37,627	37,255	37,591
R^2	0.3348	0.3526	0.3422	0.3556
$Adj\ R^2$	0.3133	0.3318	0.3209	0.3349
Bank F.E.	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes

Quarterly regression of errors in banks' expectations about future loan losses on past loan losses as specified in Equation (1). Similar to Basu et al. (2020), we scale the state variables with total loan book value of previous period. We follow a four-quarter window for averaging these scaled variables. The dependent variable $Forecast\ Error_{iT+1,T}$ is the difference between $Actual\ Loss_{iT+1}$ and $E_T Actual\ Loss_{iT+1}$ where $Actual\ Loss_{iT+1}$ is the average loan loss realized over the next four quarters from quarter $t+1$ to $t+4$. $E_T Actual\ Loss^{i}_{T+1}$ is the non-discretionary component of loan loss provision $(ndllp)$ averaged over quarter $t-3$ to t. The first row reports β_1, the coefficient on $Actual\ Loss_{iT}$, defined as the average loan loss experienced by banks in the recent past, from quarter $t-3$ to t. We augment the regression specification with $Large_{iT}$ and $Reg\ Scrutiny_T$. In the first two columnns, $Large_{iT}$ indicates whether banks' assets are in the top five as well as those whose assets are within 25% of the assets of the fifth largest bank in the federal reserve district from quarter $t-3$ to t(Hirtle et al., 2020) and in the last two columns, it denotes whether the bank's asset size (averaged from quarter $t-3$ to t) is more than the $95th$ percentile threshold. $Reg\ Scrutiny_T$ denotes whether the banking sector has experienced higher than the average loan loss in a period. Regressions are estimated using Correia (2016) multilevel panel fixed effect estimator. Standard errors are dually clustered both on bank and time; t-statistics is reported in parenthesis; *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Alternative Case 4: Variation in provisioning practices: 1990s Boom: Liu and Ryan (2006) argue that banks were keeping excess provisions during the 1990s boom period to smooth income and in a bid to obscure excessive loan loss allowance from regulators, they were also accelerating the charge-offs. Only after repeated regulatory interventions towards the end of 90s and early 2000s, the banks restrained their practice of overprovisioning and acceleration of charge-offs. For our sample, such manipulation may create a variation in the direction of overreaction and can cause an upward bias to our results.

Table C.4: **Predictability of Forecast Error:Concerns for 1990s Boom and Overprovisioning**

	β_1			
	Model 1	Model 2	Model 3	Model 4
1991Q1-2000Q4				
$Actual\ Loss_{iT}$	-0.478^{***}	-0.552^{***}	-0.474^{***}	-0.546^{***}
	(-15.39)	(-18.03)	(-15.38)	(-17.96)
Obs.	18,293	18,234	18,275	18,223
R^2	0.4225	0.4393	0.4247	0.4383
$Adj\ R^2$	0.3953	0.4128	0.3976	0.4117
2001Q1-2019Q4				
$Actual\ Loss_{iT}$	-0.341^{***}	-0.416^{***}	-0.342^{***}	-0.415^{***}
	(-12.35)	(-14.97)	(-12.55)	(-15.24)
Obs.	25,353	25,792	25,378	25,763
R^2	0.3377	0.3558	0.3461	0.3597
$Adj\ R^2$	0.3156	0.3346	0.3243	0.3386
Bank F.E.	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes

Quarterly regression of errors in banks' expectations about future loan losses on past loan losses as specified in the regression model of Equation (1). We estimate the model separately for 1991Q1-2000Q4 and 2001Q1-2019Q4. Similar to Basu et al. (2020), we scale these state variables with the total loan book value of the previous period. Additionally, we follow a four-quarter window for averaging these scaled variables. The dependent variable $Forecast\ Error_{iT+1,T}$ is the difference between $Actual\ Loss_{iT+1}$ and $E_T Actual\ Loss_{iT+1}$ where $Actual\ Loss_{iT+1}$ is the average loan loss realized over the next four quarters from quarter $t+1$ to $t+4$. $E_T Actual\ Loss_{iT+1}$ is the non-discretionary component of loan loss provision $(ndllp)$ averaged over quarter $t-3$ to t. The first row reports β_1, the coefficient on $Actual\ Loss_{iT}$, defined as the average loan loss experienced by banks in the recent past, from quarter $t-3$ to t. Regressions are estimated using Correia (2016) multilevel panel fixed effect estimator. Standard errors are dually clustered both on bank and time; t-statistics is reported in parenthesis; *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

To understand if the practices of the 1990s could have an overbearing influence on our results, we estimate Equation (1) separately for 1991Q1-2000Q4 and 2001Q1-2019Q4. This separation of time period is consistent with the analysis of Liu and Ryan (2006). In Table C.4, we report the results of these two subsamples. As expected, the estimated coefficients β_1 are higher in the 1990s - suggesting stronger predictability. However, the post-90s periods continue to exhibit strong predictability as well. Our primary finding of $\beta_1 < 0$ is continues to hold for both the periods, thus alleviating concerns about provisioning practices during 1990s boom.

Alternative Case 5: Error minimization: So far, the estimation of Equation (1) has been conducted by minimizing the mean of the squared (forecast) errors (mse). According to Basu and Markov (2004), banks may instead be setting the expected loss with the implicit objective of minimizing the *absolute* forecast errors rather than minimizing *squared* forecast errors. And therefore, coefficients arrived by the mse method could be wrongly inferred as deviation from rational expectations[29].

Table C.5: Predictability of Forecast Error under minimization of mean absolute deviation

	β_1			
	Model 1	Model 2	Model 3	Model 4
$Actual\ Loss_{iT}$	-0.315^{***}	-0.409^{***}	-0.312^{***}	-0.399^{***}
	(-15.23)	(-25.04)	(-16.19)	(-24.95)
Obs.	47,583	47,947	47,578	47,902

Quarterly regression of errors in banks' expectations about future loan losses on past loan losses as specified in the regression model of Equation (1). Similar to Basu et al. (2020), we scale these state variables with the total loan book value of the previous period. Additionally, we follow a four-quarter window for averaging these scaled variables. The dependent variable $Forecast\ Error_{iT+1,T}$ is the difference between $Actual\ Loss_{iT+1}$ and $E_T Actual\ Loss_{iT+1}$ where $Actual\ Loss_{iT+1}$ is the average loan loss realized over the next four quarters from quarter $t+1$ to $t+4$. $E_T Actual\ Loss_{iT+1}$ is the non-discretionary component of loan loss provision $(ndllp)$ averaged over quarter $t-3$ to t. The first row reports β_1, the coefficient on $Actual\ Loss_{iT}$, defined as the average loan loss experienced by banks in the recent past, from quarter $t-3$ to t. We report parameter estimates and standard errors using a quantile regression estimator for panel data (Powell, 2016); t-statistics is reported in parenthesis; *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

To address this concern, we follow Basu and Markov (2004) and estimate Equation (1) by min-

[29] Except for the special case when the distribution of the underlying variable is symmetric.

imizing the absolute error. We use the quantile regression estimator from (Powell, 2016). As reported in Table C.5, the esitmates of median β_1 remain quantitatively similar.

D. Estimation of (ρ) and (θ)

Using Equation (7), we can write

$$\beta_1 = \frac{Cov\Big(Actual\ Loss_{T+1} - E_T Actual\ Loss_{T+1}, Actual\ Loss_T\Big)}{Var(Actual\ Loss_T)}$$

$$= \frac{Cov\Big(g_{T+1} - E_T^{\theta} g_{T+1}, g_T\Big)}{Var(g_T)} = \frac{-\rho\theta Var(\varepsilon_T)}{Var(g_T)}$$

Further, if g_T follows an AR(1) process of Equation (5), then $Var(g_T) = \frac{Var(\varepsilon_T)}{1-\rho^2}$. Substituting this in the above expression for β_1 yields, $\theta = -\frac{\beta_1}{\rho(1-\rho^2)}$.

As per the formulation in Equation (5), we estimate the persistence parameter, ρ, by using the following empirical specification,

$$Actual\ Loss_{iT+1} = c + \rho Actual\ Loss_{iT} + \varepsilon_{iT+1}$$

We also run the above regression bank-by-bank and estimate bank-specific persistence parameter ρ^i. Below we present the results from the two regressions for ρ.

Table D.1: Estimation of Persistence parameter ρ

	Panel	Median
ρ	0.564*** (26.85)	0.530*** (0.01)
Bank F.E.	Yes	
Time F.E.	Yes	
Obs.	54,133	878
R^2	0.6772	
$Adj\ R^2$	0.6693	

The table provides estimation of the persistence parameter (ρ) for loan loss process. The first column report results from the panel regression of $Actual\ Loss_{iT+1}$ on $Actual\ Loss_{iT}$ and t-statistics in the parenthesis. $Actual\ Loss_{iT+1}$ is the average loan loss realized over the next four quarters from quarter $t+1$ to $t+4$ and $Actual\ Loss_{iT}$ is defined as the average loan loss experienced by banks in the recent past, from quarter $t-3$ to t. The last column shows the median coefficient ρ in bank-by-bank regression of $Actual\ Loss_{iT+1}$ on $Actual\ Loss_{iT}$ and bootstrap standard error in the parenthesis.