Incorporating Equity Market Information into Supervisory Monitoring Models

John Krainer and Jose A. Lopez

Federal Reserve Bank of San Francisco
Economic Research Department
101 Market Street
San Francisco, CA 94105-1579

john.krainer@sf.frb.org
jose.a.lopez@sf.frb.org

Draft date: June 6, 2002

Abstract: We examine whether equity market variables, such as stock returns and equity-based default probabilities, are useful to bank supervisors for assessing the condition of bank holding companies. Using an event study framework, we find that equity market variables anticipate supervisory ratings changes by up to four quarters and that the improvements in forecast accuracy arising from conditioning on equity market information are statistically significant. We develop an off-site monitoring model that easily combines supervisory and equity market information, and we find that the model's forecasts also anticipate supervisory ratings changes by several quarters. While the inclusion of equity market variables in the model does not improve forecast accuracy by much relative to simply using supervisory variables, we argue that equity market information should still be useful for forecasting supervisory ratings and should be incorporated into supervisory monitoring models.

Incorporating Equity Market Information into Supervisory Monitoring Models

Abstract: We examine whether equity market variables, such as stock returns and equity-based default probabilities, are useful to bank supervisors for assessing the condition of bank holding companies. Using an event study framework, we find that equity market variables anticipate supervisory ratings changes by up to four quarters and that the improvements in forecast accuracy arising from conditioning on equity market information are statistically significant. We develop an off-site monitoring model that easily combines supervisory and equity market information, and we find that the model's forecasts also anticipate supervisory ratings changes by several quarters. While the inclusion of equity market variables in the model does not improve forecast accuracy by much relative to simply using supervisory variables, we argue that equity market information should still be useful for forecasting supervisory ratings and should be incorporated into supervisory monitoring models.

1 Introduction

Bank supervisors engage in extensive monitoring of banking organizations in order to enforce regulations and guard against systemic risk. In the United States, the various supervisory agencies conduct both on-site and off-site inspections of bank holding companies (BHCs) and their subsidiaries. Both of these monitoring methods have their uses and limitations. On-site supervisory visits produce a detailed picture of an institution's financial condition and risk profile. However, they absorb considerable resources from both supervisors and the supervised institutions. Importantly, most institutions typically receive about one on-site visit per year. Between these visits, there is potentially enough time for an institution's risk profile to change dramatically. Off-site monitoring typically consists of gathering and analyzing supervisory data through standard regulatory reporting forms on a quarterly basis. Off-site monitoring permits relatively more timely supervisory analysis and more efficient allocation of scarce supervisory resources. However, off-site monitoring relies heavily on the variables on the regulatory reporting forms, which may not capture certain changes in BHC risk profiles.[1]

To partly address these limitations in their monitoring technology, bank supervisors are considering the expanded use of market discipline and financial market information to supplement their efforts.[2] If markets are efficient and investors can assess a firm's true condition and riskiness, then financial market data could potentially provide valuable signals about changes in firm risk. Using market data could also be useful for validating supervisory assessments and actions. Although there seems to be agreement on the conceptual merits of incorporating financial market information into supervisory monitoring, no consensus on

---

[1]For most institutions, off-site monitoring consists of the analysis of quarterly reporting forms. However, large, complex, banking organizations often have a continuous, on-site supervisory presence. See DeFerrari and Palmer (2001) for a discussion of current supervisory policy.

[2]For example, the Basel Committee on Banking Supervision (2001) issued a proposal for a new international accord on commercial bank capital requirements that will rely explicitly on market discipline. The Committee stated that improved disclosure would enhance the role of market participants in the setting of appropriate levels of bank capital. This reasoning assumes that market participants will feel sufficiently at risk to make use of the additional disclosure.

how to do so has developed. To date, much of the research on this question has focused on subordinated debt, primarily because the concerns of debt holders are thought to be more closely aligned with those of the supervisors.[3]

In this paper, we focus on the potential contribution of equity market information to the supervisory monitoring of BHCs. We believe that the use of equity market information offers many important advantages over traditional bond market information. First, the equity market is thought to be much more efficient in processing information than the corporate bond markets.[4] Even though equity investors' preferences may not be as closely aligned with those of the supervisors (i.e., equity investors may prefer more risk than supervisors), equity prices contain potentially more information. Second, more BHCs have traded equity than have traded subordinated debt. In our sample, only 10% of the BHCs had outstanding subordinated debt at year-end 1999, and these firms accounted for 67% of total U.S. banking assets at that time. An even smaller percentage had issues with a satisfactory amount of liquidity to generate reliable prices. In contrast, about 20% of the BHCs had publicly traded equity, and these BHCs accounted for about 85% of total banking assets; see Table 1. We do not argue against using debt; supervisors could potentially use both when both are available. We simply note that if the aim is to harness market information for a large portion of the banking industry, supervisors cannot ignore equity markets. Third, a primary conclusion of the early studies of market discipline was that the perceived existence of "too-big-to-fail" policies weakened the incentives of bond market participants to monitor banks. While there is evidence that bond markets have changed following a retreat from these policies (see Flannery and Sorescu ,1996), the point remains that information in subordinated debt spreads will depend on how "tough" the regulatory regime is. Equity prices should be much less sensitive to this concern.

We examine the potential contribution of various equity market indicators of BHC performance for predicting supervisory BHC ratings, known as BOPEC ratings.[5] We consider

---

[3]Examples of research in this area include Bliss and Flannery (2001), DeYoung et.al. (2001), Evanoff and Wall (2000), Flannery and Sorescu (1996), Hancock and Kwast (2000) as well as Kwast et.al. (1999).

[4]For a further discussion of this point, see Saunders (2001).

[5]Note that in this paper we focus on supervisory ratings and not defaults, another key supervisory concern.

individual BHC stock returns and the "expected default frequency"^{TM} (or EDF^{TM}) measure generated by the financial software firm KMV.[6] We analyze these equity market measures for publicly traded BHCs from 1990 through 1999 by conducting event studies to determine whether these variables anticipate (or "lead") changes in supervisory ratings in a statistically significant way. We find that both of these variables detect BOPEC rating changes up to four quarters before their assignment. Conditioning on these equity market variables at up to four quarters before the event leads to statistically significant improvements in BOPEC rating forecasts. Hence, supervisors could reduce their probability of incorrectly assuming that a publicly traded BHC's next inspection will result in no change in BOPEC rating by conditioning on a single equity market variable.[7]

We also investigate whether BHC assessments based purely on supervisory information anticipate BOPEC changes. To generate such assessments at a quarterly frequency, we propose a BOPEC off-site monitoring (BOM) model that effectively extends the SEER modeling framework for bank supervisory ratings to BHC ratings.[8] Using variables from quarterly regulatory reports and supervisory databases, we estimate ordered logit models of BOPEC ratings for the purpose of generating one-quarter-ahead BOPEC forecasts for all BHCs in our sample. We find that fluctuations in these forecasts are capable of detecting BOPEC changes up to four quarters in advance.

We then incorporate equity market variables into the BOM model to test whether they improve its forecast performance. Theoretically, the combination of the supervisory and market information sets should improve the BOM model's forecasts. We find in practice that these forecasts also anticipate BOPEC changes by up to four quarters. However, the improvement in forecast accuracy is not statistically significant relative to the BOM forecasts

---

There exists an extensive literature on firm default dating back to Altman (1968) and bank default dating back to Meyer and Pifer (1970), Sinkey (1975) and Pettway and Sinkey (1980).

[6]Although an EDF value is a predicted default rate for a company's debt, it is derived primarily from equity market data. See Crosbie and Bohn (2002) for a description of KMV's methodology.

[7]See Flannery (1998, 2001) for discussions on the value of financial market variables to supervisory monitoring.

[8]See Cole, Cornyn, and Gunther (1995) for a discussion of the SEER model and its use in the off-site monitoring of banks.

based solely on supervisory variables.

In conclusion, our empirical results suggest that equity market information is useful for modeling BOPEC ratings changes. While these variables do not improve the forecasting performance of the BOM model dramatically within our sample period, we still recommend that equity market variables be incorporated into off-site supervisory monitoring, especially given the ease with which it can done using the BOM model. As noted by Flannery (2001), financial market information should complement supervisory information in that more accurate forecasts of BHC condition can be made using both sources of information than with either one alone.

The paper proceeds as follows. In Section 2, we provide a brief overview of the supervisory process for bank holding companies in the U.S. We also provide a brief survey of the academic literature on off-site monitoring models and the use of financial market information for supervisory monitoring. In Section 3, we describe the data used in the study. In Section 4, we conduct event studies that test whether equity market measures anticipate changes in BOPEC ratings. In Section 5, we present and estimate our proposed BOM model for BOPEC ratings, and Section 6 concludes.

2 Supervision of bank holding companies

The Federal Reserve is the supervisor of BHCs in the United States. Full-scope, on-site inspections are a key element of this supervisory process. These inspections are generally conducted on an annual basis, particularly for the case of large and complex BHCs.[9] Although limited and targeted inspections that may or may not be conducted on-site are also carried out, we focus here on full-scope, on-site inspections since they provide the most comprehensive supervisory assessments of BHCs.

At the conclusion of an inspection, the supervisors assign the institution a numerical

---

[9]A complex BHC has nonbank subsidiaries that extend credit or have debt outstanding to the general public. See DeFerrari and Palmer (2001) for an overview of the supervsiory process for large, complex banking organizations.

4

rating called a composite BOPEC rating that summarizes their opinion of the BHC's overall health and financial condition. The BOPEC acronym stands for the five key areas of supervisory concern: the condition of the BHC's <u>B</u>ank subsidiaries, <u>O</u>ther nonbank subsidiaries, <u>P</u>arent company, <u>E</u>arnings, and <u>C</u>apital adequacy. Like bond ratings given by the private rating agencies, BOPEC ratings are absolute ratings, not relative ratings.[10] BHCs with the best performance are assigned a BOPEC rating of 1, while those with the worst performance are given a BOPEC rating of 5. A rating of 1 or 2 indicates that the BHC is not considered to be of supervisory concern. Note that BOPEC ratings, as well as all other inspection materials, are highly confidential and are never made publicly available.

Between on-site inspections when private supervisory information cannot be gathered as readily, supervisors monitor BHCs using an off-site monitoring system based on quarterly regulatory reports filed by BHCs and their subsidiary banks. This off-site monitoring system is primarily based on two key information sources. The first source, known as the BHC Performance Report, is a detailed summary of their quarterly Y-9C regulatory reporting forms.[11] As of March 1999, the report summarized approximately 800 BHC variables across several years. From this report, certain variables are selected as key performance criteria, and if a BHC fails to meet these criteria in a given quarter, it is noted as an exception that requires further monitoring.

The second source of information for off-site BHC monitoring is the supervisory CAMELS ratings assigned to banks within a holding company. As with BOPEC ratings, CAMELS ratings are assigned after bank examinations. The acronym refers to the six key areas of concern: the bank's <u>C</u>apital adequacy, <u>A</u>sset quality, <u>M</u>anagement, <u>E</u>arnings, <u>L</u>iquidity, and <u>S</u>ensitivity to risk. The composite CAMELS rating also ranges in integer value from 1 to 5 in decreasing order (i.e., banks that perform best are assigned a rating of 1). Since the condition of a BHC is closely related to the condition of its subsidiary banks, the off-site BHC surveillance program includes monitoring recently assigned CAMELS ratings.

As with on-site BHC inspections, on-site bank examinations occur at approximately a

---

[10] For an international survey of supervisory bank rating systems, see Sahajwala and Van der Bergh (2000).

[11] For further information, see Board of Governors of the Federal Reserve System (2001).

yearly frequency, which is long enough for the gathered supervisory information to decay and become less representative of the bank's condition.[12] To address this issue, the Federal Reserve instituted an off-site monitoring system for banks, known as the System for Estimating Examiner Ratings (SEER), in 1993. The SEER system actually consists of two separate models that forecast bank failures over a two-year horizon as well as bank CAMELS ratings for the next quarter. The model that we are most interested in here is the latter, which is an ordered logit model with five categories corresponding to the five possible values of the CAMELS rating. The model is estimated every quarter in order to reflect the most recent relationship between selected financial ratios and the two most recent quarters of CAMELS ratings. Significant changes in a bank's CAMELS rating as forecasted by the SEER model could be sufficient to warrant closer monitoring of the bank. The off-site BHC surveillance program also explicitly monitors the SEER model's forecasted CAMELS ratings.

2.1   Literature review

Modern theories of financial intermediation, such as Diamond (1984), stress the role of banks as monitors and processors of information. One strand of empirical work that has flowed from this model of banking seeks to shed light on whether banks, as specialists in solving problems of asymmetric information between firms and investors, are themselves opaque to investors. To do so, several studies look at whether interest rates on bank liabilities are responsive to measures of default risk. Evidence supporting bank transparency has been mixed. Looking at new issues of large certificates of deposit (CDs), Hannan and Hanweck (1988) find that CD rates are sensitive to measures of bank risk such as the volatility of return on assets and capital ratios. Interestingly, it has been more difficult to establish a correlation between yields on debentures, which are typically junior and longer term than CDs, and the usual proxies for risk; see Gorton and Santomero (1990). Of course, this may not necessarily reflect the market's inability to evaluate a banking firm's assets, but rather the market's perception of a government guarantee on the bank's liabilities. In fact, Flannery and Sorescu (1996) demonstrate that these perceptions are empirically important. They find that following the

---

[12]See Cole and Gunther (1998) as well as Hirtle and Lopez (1999) for further discussion of this issue.

FDIC Improvement Act of 1991 and its stated retreat from too-big-to-fail, spreads on bank debt became more responsive to risk measures.

A related branch of the literature has sought to find whether banks are opaque to investors relative to comparable nonbank firms. Morgan (2000) provides some evidence that this is the case. He looks at ratings agency evaluations of bonds and finds that the agencies are more likely to disagree on a rating when the issuing firm is a bank. Flannery et al. (1999) turn to stock market data to address this question. Arguing that firms with opaque assets are likely to have large bid-ask spreads, they analyze the microstructure properties of a sample of small, publicly traded banks. Small banks are observed to have large spreads. Unlike Morgan, however, they find that bank spreads are not large relative to spreads of small firms in other industries, nor do other measures such as trading volume or analyst coverage indicate that banks are less transparent than nonbanks.

The branch of the literature most related to this paper concerns the interaction of market evaluations of bank condition with supervisory assessments. Berger and Davies (1998) use an event study framework to examine whether daily stock prices react to CAMELS rating changes. Even though these ratings are confidential, they find that BHC stock prices respond to these changes, implying that supervisory assessments provide valuable information that the equity market can detect. They conjecture that market participants view the supervisory rating as both a certification of financial statements and an indicator of future regulatory treatment.[13] Using data from the bond markets, DeYoung et al. (2001) find that supervisory information significantly affects contemporaneous and subsequent changes in the spreads on bank debentures. Specifically, they find that the private supervisory information component of bank CAMELS ratings impacts debenture spreads several months after the CAMELS assignment.

Our paper fits into the group of studies that looks at the ability of financial markets to anticipate (or "lead") supervisory ratings. Note that market participants are not actively

---

[13]Note that Allen et al. (2001) do not find evidence that supervisory ratings affect market assessments when the confidential ratings of banks' management (i.e., the "M" component of the CAMELS rating) became public knowledge. However, since these results are tied to a one-time change in charter status, it is not clear how general these results are.

trying to guess what BOPEC or CAMELS ratings will be, nor do they officially learn the outcomes of supervisory inspections. This literature tests whether there is some alignment between investor and supervisory preferences that would allow supervisors to use the more timely market data for their own purposes. Thus, supervisory ratings must be correlated with variables that investors care about, such as risk and expected return. This correlation could arise because supervisory actions could have an impact on the return distribution, or because the conditions that lead to good supervisory ratings are consistent with those that lead to changes in stock price.

Berger et al. (2000) pursue this idea when they examine the timeliness and accuracy of supervisory and market assessments of the condition of large BHCs. Their study is one of the few that uses both equity and bond market information. They find that market assessments appear to focus on different aspects of BHC performance than do supervisory assessments. Supervisory assessments, defined with respect to BOPEC ratings, are most closely related to bond market assessments, meaning that supervisory variables are shown to Granger-cause bond market variables, and vice versa. However, this type of interrelation is not apparent between supervisory assessments and equity market assessments; that is, they find very little Granger-causality from equity market variables to their supervisory assessment variables. Furthermore, they find that, after accounting for market assessments, supervisory variables do not contribute substantially to the modeling of future indicators of BHC performance, such as changes in nonperforming loans. Overall, their findings suggest that supervisors, bond market participants, and equity market participants produce complementary information on BHC performance.[14]

Policymakers have considered the possibility of explicitly linking policy to financial mar-

---

[14]Several recent studies have provided complementary evidence. Gunther et al. (2001) and Hall et al. (2001) find that equity-based market signals provide useful information that supplements supervisory assessments. Elmer and Fissel (2001) and Curry et al. (2001) further support this conclusion by finding that equity market variables add value to supervisory models of bank failure. Gilbert et al. (2001) find that default risk premia for jumbo CDs, as derived from financial statement variables, do not predict CAMELS downgrades as well as a traditional econometric model for a sample of small banks. Gropp et al. (2001) provide evidence in favor of using market information for the supervision of European banks.

ket signals. Since the interests of bank supervisors are supposedly most closely aligned with bank subordinated debt holders, most of the exploratory work on this issue has concentrated on the signal quality in subordinated debt spreads.[15] Evanoff and Wall (2000) model changes in the supervisory ratings of banks and BHCs with outstanding subordinated debt as a function of lagged subordinated debt spreads and regulatory capital ratios over the period from 1990 to 1999. They find that subordinated debt spreads do as well as or better than capital ratios at explaining supervisory ratings.

Our paper differs from the existing literature in several respects. First, our emphasis on equity market variables over the period from 1990 to 1999 allows us to consider a larger sample of institutions than much of the previous literature. Second, our focus on stock market activity up to a year before the assignment of supervisory ratings should provide clear insight into the timeliness of market signals. Third, we conduct explicit out-of-sample tests that examine the usefulness of equity market information to supervisors in a systematic and real-time manner.

3   The BOPEC sample

The core database for our analysis is the supervisory BOPEC ratings assigned to BHCs over the period from 1990 to 1999. We chose to analyze only BOPEC ratings assigned after on-site, full-scope inspections. This requirement reflects the concern that limited and targeted inspections produce a less comprehensive supervisory information set than a full inspection. Our sample of BOPEC ratings is further refined to include only inspections of top-tier BHCs with four quarters of available supervisory data and prior BOPEC ratings. We focus on top-tier BHCs since they are typically the legal entity within a banking group that issues publicly traded equity. We require each BHC to have filed four quarters of supervisory data and to have been inspected at least once before. This allows us to avoid issues regarding de

---

[15]One objection to this proposition is found in Bliss (2000). He shows that supervisory interests may diverge from bondholder interests in that both parties may not necssarily agree on the relative riskiness of different banks or bank portfolios.

novo BHCs and new BHCs arising from mergers.

Table 2 summarizes our sample with respect to asset size. The full sample contains 3,975 BOPEC ratings for 1,399 different BHCs.[16] Most of the BHCs in the sample are relatively small, with less than $1 billion in total assets. A roughly equal number of these inspections occurred in the first and second halves of the sample period. Note that publicly traded BHCs are generally larger than private BHCs, with a greater percentage having total assets ranging between $1 billion and $100 billion.

Tables 3A and 3B present the distribution of BOPEC ratings assigned in each year for all BHCs and for publicly traded BHCs, respectively. Clearly, the majority of the ratings fall in the upper two categories, which indicate that a BHC's financial condition and risk profile are of little supervisory concern. Although the distribution fluctuates over the sample, the percentage of ratings in the top two categories never falls below the 65% observed in 1991. The maximum percentage of 1 and 2-rated institutions at any point in time is 95.7% in 1997. Note that there are very few BOPEC 5 ratings in the sample. This outcome is probably due to both supervisors and bankers taking actions to try to prevent this outcome and to the fact that 5-rated institutions quickly improve their ratings, merge with a stronger institution, or simply close.

Our sample contains 1,602 BOPEC ratings for publicly traded BHCs, which represents approximately 40% of the full sample. These ratings correspond to 440 unique institutions, which implies a slightly higher ratio of BOPEC ratings per BHC than for the full sample (i.e., 2.8 for the full sample and 3.6 for the publicly traded sample). However, the ratings distribution for publicly traded BHCs is similar to that of the full sample.

Tables 4A and 4B present the patterns of changes in the BOPEC ratings in our sample. The most frequent outcome is no change in BOPEC rating, accounting for about 54% to 72% of the annual totals. The pattern of BOPEC upgrades and downgrades fluctuates over the sample. From 1990 through 1992, more downgrades occurred than upgrades, with the pattern reversing from 1993 through 1998. These patterns appear to follow the general

---

[16]The mean and median number of BOPEC rating assignments per BHC in the sample are 2.8 and 2, respectively. The minimum and maximum numbers are 1 and 11, respectively.

trends in the banking industry and the macroeconomy during the 1990s.

4   Equity market event studies

In this section, we conduct two event studies to examine whether equity market variables anticipate (or "lead") BOPEC assignments. These event studies examine whether BHC equity-based variables behave abnormally relative to a standard benchmark. Thus, the exercise in this section focuses on idiosyncratic variation in equity market variables that might be useful to regulators. Given the historical trends in the banking industry so clearly visible in the distribution of BOPEC ratings over time (i.e., tables 3A and 3B), it is also possible that systematic variation in equity market variables might be useful in predicting ratings outcomes. We explore this possibility directly in Section 5.

We define the *event date* to be the inspection exit date, which is the day that examiners conduct their final meeting with BHC management. The first event study is based on BHC stock returns and uses standard methods, as in Campbell et al. (1997). The second event study is based on the expected default frequencies (or EDF values) generated by KMV. To conduct this analysis, we follow the work of Delianedis and Geske (1998), who examine de-medianed default probabilities for publicly traded firms and use the sign test to study their behavior prior to actions by ratings agencies.[17]

There are a number of issues that arise in the process of testing whether equity market variables anticipate BOPEC rating assignments. First, as mentioned, BOPEC ratings are confidential and never revealed to the public. Hence, equity market variables cannot directly anticipate BOPEC ratings. Instead, we must assume that investors anticipate that changes in BHC performance, such as BHC condition and future profitability, would probably lead to changes in their future supervisory treatment and, hence, changes in BOPEC ratings. Detecting such a leading relationship between stock market variables and BOPEC ratings is evidence that market signals may be helpful to BHC supervisors. Failure to detect such a relationship may be due to BHCs being too opaque for there to be much agreement between

---

[17]Note that Delianedis and Geske (1998) generate an alternative set of default probabilities.

investors and supervisors or due to investor indifference regarding BOPEC ratings.

Second, the timing convention used in the event study is critical. If investors receive information about a change in BHC risk before the event window begins, then the event study will miss the stock market signal. To guard against this possibility, our event window begins one year before the BOPEC assignment.

Finally, equity market signals can only be interpreted as significant or abnormal with the aid of a model. Thus, as is the case with all event studies, the hypotheses tested are in fact joint hypotheses of whether market participants anticipate an event *and* whether the pricing model is correct.

4.1 Event studies using abnormal stock returns

In this section, we use four-week stock return data to investigate whether equity markets anticipate BOPEC changes.[18] For ease of exposition, we refer hereafter to four-week returns as *monthly* returns. The event window begins twelve months prior to the exit date and ends three months after. We use a two factor model of BHC stock returns,

$$R_{it} = \alpha + \beta_1 R_{mt} + \beta_2 f_t + \varepsilon_{it}, \tag{1}$$

where R_{it} is the monthly return for BHC i at time t, R_{mt} is the overall market return for that period, and f_t is the monthly change in the federal funds rate. The overall market return is the CRSP value-weighted index, which includes all stocks traded on the NYSE, AMEX, and NASDAQ. The interest rate factor is meant to control for interest rate shocks that impact BHCs in a systematic way. ε_{it} is assumed to be an independent normal error term.

The model is estimated using monthly returns for the five years prior to the beginning of the event window, and the residual vector $\widehat{\varepsilon}_i$ is retained for future use. Within the event window, abnormal returns are computed using the estimated coefficients from (1),

---

[18] Note that Berger and Davies (1998) conduct their BHC event study using daily stock returns. We chose a monthly frequency in order to match the monthly frequency used in generating EDF values and to more closely approximate the quarterly frequency of BHC off-site monitoring.

$$AR_{it} = R_{it} - \widehat{\alpha} - \widehat{\beta}_1 R_{mt} - \widehat{\beta}_2 f_t, \tag{2}$$

for t = -12,...,0,...,+3. If returns are jointly normally distributed, the distribution of the (16x1) abnormal return vector is $AR_{it} \backsim N(0, V_i)$. In general, the conditional covariance matrix V_i will not be diagonal; sampling error in the estimation of the parameters will generate autocorrelation in the errors for abnormal returns.[19] If we rewrite the regression system in equation (1) as

$$R_i = X\theta + \varepsilon_i, \tag{3}$$

where X is a three-column matrix of ones, the market returns, and the changes in interest rates within the estimation window, it can be shown that

$$V_i = I\left(\frac{\widehat{\varepsilon}_i'\widehat{\varepsilon}_i}{L-3}\right) + X_i^*(X_i'X_i)^{-1}X_i^{*\prime}\left(\frac{\widehat{\varepsilon}_i'\widehat{\varepsilon}_i}{L-3}\right), \tag{4}$$

where I is a (16x16) identity matrix, L is the length of the estimation window, and X^* is a three-column matrix of ones, the market returns, and the interest rate changes within the event window. The distribution of the cumulative abnormal return $CAR_{it} = \sum_{j=-12}^{t} AR_{ij}$ is given by $CAR_{it} \backsim N(0, \gamma_t'V_i\gamma_t)$, where γ_t is a vector of dimension equal to the event window with ones in the first t elements and zeroes elsewhere. The standardized CAR_{it} is denoted as $SCAR_{it} = CAR_{it}/\sqrt{\gamma_t'V_i\gamma_t}$.

As is standard in event study analysis, we examine whether the average CARs across BHCs at points in event time are statistically different from zero. To guard against contamination of the event windows from prior inspections, we restrict the analysis to BOPEC assignments separated by at least one calendar year. It is still possible that the results of previous BOPEC assignments contaminate the parameter estimates of the two factor model, but this problem is mitigated by our use of a long estimation window. We also restricted our sample to BOPEC assignments where a five-year estimation window was available. These restrictions result in a reduction from the original sample size of 1,602 BOPEC ratings of

---

[19]This discussion follows Campbell, Lo, and MacKinlay (1997).

publicly traded BHCs to just 813 ratings. The sample consists of 139 BOPEC upgrades, 85 downgrades, and 589 no-changes. Note that 50 of the upgrades are over the BOPEC-3 threshold and that 39 of the downgrades are below the BOPEC-2 threshold.

The results presented in Table 5 and graphed in Figures 1 and 2 show clearly that equity market participants anticipate BOPEC changes well in advance of their actual assignments. For upgrades, the average CARs increase as we approach the event date and are over 13% by the end of the event window. These CARs are significantly positive (i.e., $SCAR_{it} \geq 1.96$) as early as 12 months prior to the inspection. The average CARs for inspections with no change in BOPEC ratings are generally positive, yet quite small and never statistically significant. By the end of the event window, the average CAR for no change inspections is approximately -1%, while the average CAR for downgrades -15% . Importantly, the average downgrade CAR is statistically significant as early as ten months prior to the inspection. These results confirm the hypothesis that equity markets are capable of anticipating BOPEC changes. Additionally, on average, the market does not appear to send false signals of approaching changes; there is no significant market response in advance of inspections where nothing happens.

A change over the threshold between BOPEC ratings 2 and 3 is of particular concern to supervisors. Note that crossings of this regulatory threshold are relatively rare events (about 11% of the sample) and, thus, the abnormal returns are measured with considerable error. Nevertheless, for downgrades moving below the threshold, the equity market appears to signal a downgrade at nine months prior, approximately the same time as all downgrades. However, it is interesting to note that these ostensibly serious downgrades are not preceded by large declines in stock prices. Over the event window, the average CAR for downgrades below the threshold is -11.0%, compared to an average of -14.6% for all downgrades. The market does not deliver the same weak response for inspections resulting in upgrades moving above the threshold. For the sample of 50 upgrades above the threshold, the average CAR is +21.2% at the event date and +26.3% by the end of the event window, relative to +10.1% and +12.4%, respectively, for all upgrades. The equity market appears to anticipate BOPEC upgrades moving above the threshold by seven months. While the SCARs are significantly

14

different from zero for both upgrades and downgrades, we cannot reject the null hypothesis that the average CARs for threshold changes and for all BOPEC changes are equal.

Given that SCARs are shown to anticipate BOPEC changes, supervisors might be able to use them as an additional monitoring tool, if their implied forecasts were accurate. To evaluate this forecast accuracy, we test whether the distribution of BOPEC rating outcomes conditional on the SCARs at various periods prior to the assignments is statistically different from the unconditional distribution of outcomes. In the full sample of publicly traded BHCs, the unconditional probabilities of an upgrade, no change, or downgrade are 19.5%, 66.4%, and 14.2%, respectively; see Table 4B. Given a suitably strong signal in the form of a "large" SCAR, are the conditional probabilities statistically different? We use the Pearson goodness-of-fit statistic to test the null hypothesis that the conditional probabilities are not different from the unconditional ones and that the conditioning information is thus not informative.

Table 6 presents the forecast accuracy results for SCARs at one, three, six, nine, and twelve months prior to BOPEC assignments. Since the SCARs are assumed to be normally distributed, a reasonable method for defining a strong signal is whether its value is in the 5% tails; i.e., $|SCAR| > 1.96$. The results are decomposed for BOPEC upgrades, downgrades and no changes. Not surprisingly, BOPEC no changes are forecasted quite well; i.e., about 72% of SCARs in the interval [-1.96, 1.96] at each of the five forecast horizons are followed by no change in the BOPEC ratings. Focusing on changes, BOPEC downgrades are forecast more accurately. At twelve months prior, 7 of 26 (or 26.9%) of the large, negative SCARs are followed by a BOPEC downgrade. This percentage increases to a high of 42.4% at six months prior. The results are weaker for BOPEC upgrades; peak performance is at six months prior with 32.3% of large, positive SCARs leading to BOPEC upgrades. At all five horizons, the Pearson statistic is statistically significant, and we reject the null hypothesis that the SCARs are uninformative for the purpose of forecasting BOPEC ratings.

In summary, our analysis of abnormal stock returns indicates that equity markets are capable of anticipating BOPEC upgrades and downgrades by up to four quarters. This result supports the view that SCARs could be of use to supervisors for off-site monitoring. In fact, when using SCARs to forecast BOPEC assignments, 75% accuracy is achieved at up

to twelve months prior to assignment, and accuracy increases to 88% at three months prior to an inspection.

4.2   Event studies using EDF measures from KMV

As mentioned, a drawback to using stock returns to forecast BOPEC ratings is the somewhat ambiguous relationship between risk and required return when a firm is close to insolvency. For this reason, it has been argued that using subordinated debt prices should present less of a signal extraction problem. However, as per Merton (1974), it has long been understood that equity prices contain useful information for the pricing of corporate debt. This "structural approach" to pricing credit risk can also be used to supply estimates of a firm's asset volatility. KMV, a financial software firm, uses a structural model to estimate asset volatilities at a monthly frequency. With these volatility estimates, KMV uses a default database and information about a firm's liability structure to infer its expected default frequency (EDF) every month. In our event study, we analyze the one-year EDF values from KMV for publicly traded BHCs.[20]

To better conform with the spirit of previous event studies, we need to control for systematic variation in the EDF values due to common economic factors. We adjust the EDF values in month t by subtracting the median EDF value for the BHCs that have the same contemporaneous BOPEC rating. For example, for a BHC with BOPEC-3 rating in month t, we adjust by subtracting the median EDF value of all BOPEC-3 rated BHCs at that point in time.

The sample of EDF values linked to BOPEC assignments is slightly different from both the full sample of BOPECs and the sample of BOPECs used in the stock return event study.[21] For this event study, our sample consists of 742 BOPEC assignments, of which 122

---

[20]KMV makes a distinction between financial and nonfinancial firms because of the differing natures of the firms' balance sheets and funding sources. However, in theory and in practice, KMV's financial EDF values are similar to their nonfinancial counterparts. See Gunther et al. (2001) for further analysis of EDF values and supervisory ratings.

[21]The main reason for the difference in number of institutions in the EDF sample is KMV's reliance on

are upgrades, 66 are downgrades, and 554 are no changes. Of the upgrades, 32 move above the BOPEC-3 threshold, and of the downgrades, 22 move below the BOPEC-2 threshold.

In Figure 3, we plot the medians of the appropriately de-medianed EDF values (in percentage units) for each month in the event window for BOPEC upgrades, downgrades, and no changes. As expected, the medians of the adjusted EDF values for BOPEC upgrades are negative since the corresponding EDF values improve relative to their BOPEC rating cohorts. Similarly, the median for BOPEC downgrades is positive. To investigate these relationships statistically, we conduct nonparametric sign tests on the de-medianed EDF values for each period of the event window. The sign tests are based on the null hypothesis that the distribution of the adjusted EDF values at a point in event time has a median of zero. No other distributional assumptions are imposed. This null hypothesis would imply that the EDF values do not provide useful information, analogous to the hypothesis in the previous section that the expected value of the CAR is zero. The test is conducted at the 5% significance level.

The results in Table 7 are broadly consistent with the SCARs results. For downgrades, adjusted EDF values are significantly higher than zero as early as twelve months prior to the event date. For upgrades, the median values are significant up to nine months before the event date. The results are mixed for inspections that result in no change in BOPEC rating. Through most of the event window, the medians of the adjusted EDF values are practically zero, but they become statistically significant at two months prior to the event date. The reason for this result is not clear.

Note that an important difference between this event study and the one in Section 4.1 is the nature of the statistical tests. The sign tests only take into account whether an adjusted EDF value is above or below the median, not its distance from the median. That is, unlike the cumulative abnormal stock returns in the previous section, it is not easy to estimate the economic significance of the changes in EDF values. One step towards evaluating the economic significance of the signal extracted from the adjusted EDF values is to conduct a forecast accuracy exercise, as we did with the SCARs. However, an immediate problem is

Compustat data. Prior to 1993, Compustat had relatively light coverage of publicly traded BHCs.

determining what constitutes a strong signal in the absence of a clear distribution for the adjusted EDFs. To proceed, we use the observed distribution and determine its quantiles for the 5% tails, which were -0.25 and +0.93.

Table 8 presents the results from this exercise. An adjusted EDF value of less than -0.25 should indicate an upgrade, and does so approximately 55% of the time for the five selected horizons. This forecast accuracy peaks at 63% six months prior to the inspection. Forecast accuracy for downgrades is not as good; an adjusted EDF value greater than 0.93 leads to a BOPEC downgrade only about 40% of the time. Overall, the forecast accuracy of the adjusted EDF values is approximately 75% at all horizons, which is similar to the SCAR results in Table 6. Once again, the Pearson statistics allow us to reject the null hypothesis that the distribution of outcomes when conditioning on the adjusted EDF values is not different from the unconditional distribution. Thus, adjusted EDF values are informative with respect to forecasting BOPEC ratings and could be of use for off-site monitoring by bank supervisors.

5   The BOPEC Off-site Monitoring Model

Given our empirical evidence that equity market variables anticipate BOPEC assignments, we consider directly whether these variables provide supervisors with information beyond what is available to them from supervisory data sources. We conduct this analysis within the context of an off-site monitoring model.

Our proposed BOPEC off-site monitoring (BOM) model is similar in structure to the SEER model for CAMELS ratings. The model is an ordered logit, since BOPEC ratings take only integer values from one through five. The model assumes that the BOPEC rating assigned to BHC i in quarter t, denoted BP_{it}^*, is an unobservable continuous variable based on supervisory variables available in quarter $t-2$. We would prefer to use supervisory variables from quarter $t-1$ for this exercise, but the supervisory data are not widely available for about 60 days after the end of the quarter. Since we conduct what amounts to a real-time forecasting exercise in Section 5.1, we err on the conservative side and use supervisory data

dated two quarters prior to the inspection.

The BP_{it}^* rating is modeled as

$$BP_{it}^* = (\beta + \gamma I_{it-1})x_{it-2} + \varepsilon_{it}, \tag{5}$$

where x_{it-2} is a $(k \times 1)$ vector of explanatory variables unique to BHC i observed two quarters prior to the BOPEC assignment and I_{it-1} is an indicator variable for whether the BHC was publicly traded in the prior quarter. The interaction term $I_{it-1}x_{it-2}$ allows us to control for possible differences between private and publicly traded BHCs. The error term ε_{it} has a standard logistic distribution.

Since BP_{it}^* is unobserved, we must model the observable BOPEC rating $BP_{it} \in \{1, 2, 3, 4, 5\}$. Thus, in addition to the β and γ parameter vectors, we must also estimate four cutpoints denoted α_j such that

$$
\begin{aligned}
BP_{it} &= 1 \text{ if } BP_{it}^* \in (-\infty, \alpha_1]; \\
&= 2 \text{ if } BP_{it}^* \in (\alpha_1, \alpha_2]; \\
&= 3 \text{ if } BP_{it}^* \in (\alpha_2, \alpha_3]; \\
&= 4 \text{ if } BP_{it}^* \in (\alpha_3, \alpha_4]; \\
&= 5 \text{ if } BP_{it}^* \in (\alpha_4, \infty).
\end{aligned}
\tag{6}
$$

The parameter vector $\theta = (\beta, \gamma, \alpha)$ is of dimension $((2k + 4) \times 1)$.

The density function for an assigned BOPEC rating is constructed by defining Y_{ijt} as an indicator variable equal to 1 if rating j is assigned to BHC i at time t. Since the ratings are ordered, the probability that BHC i is assigned BOPEC rating j is calculated as the difference between the cumulative probability of receiving rating j and the cumulative probability of receiving rating $j - 1$; i.e.,

$$\Pr(Y_{ijt} = 1) = \Lambda[\alpha_j - (\beta + \gamma I_{it-1})x_{it-2}] - \Lambda[\alpha_{j-1} - (\beta + \gamma I_{it-1})x_{it-2}], \tag{7}$$

where $\Lambda[x]$ is the cumulative logistic function. In an estimation sample with N ratings, the likelihood function is

$$L(\theta) = \prod_{i=1}^{N}\prod_{j=1}^{5}[\Pr(Y_{ijt} = 1)]^{Y_{ijt}}. \tag{8}$$

19

The choice of which supervisory variables to include in x_{it-2} is challenging. No simple behavioral models exist of how supervisors assign BOPEC ratings. As mentioned, there are more than 800 variables at the supervisors' disposal for this purpose. For this study, we selected a relatively small number of explanatory variables that are reasonable proxies for the five components of the BOPEC rating. We chose a parsimonious specification in order to avoid overfitting the data and to construct an off-site monitoring model that can generate reasonable out-of-sample forecasts. Additionally, we face the practical concern that the number of observations available in any subsample is quite a bit smaller than the full sample.

The ten explanatory variables used in this study are summarized in Tables 9A and 9B. The first variable is the natural log of total BHC assets, which is our control variable for firm size. The next four variables are used to capture the supervisory concerns regarding the BHC's bank subsidiaries, as summarized in the "B" component of the rating. The second variable is the CAMELS rating of the BHC's lead bank.[22] The third variable is the ratio of the BHC's nonperforming loans, nonaccrual loans, and other real estate owned to its total assets. This "problem loans" variable proxies for the health and performance of the BHC's loans that are not making their scheduled payments. Note that CAMELS ratings, nonperforming loans, and nonaccrual loans variables are all confidential supervisory information. The fourth variable is the ratio of the BHC's allowances (or provisions) for losses on loans and leases (ALLL) to its total loans, another proxy for the health and performance of the BHC's lending portfolio.

The fifth variable is an indicator of whether the BHC has a Section 20 subsidiary, which is a subsidiary that can engage in securities activities that commercial banks were not permitted to engage in before the Gramm-Leach-Bliley Act of 1999. This variable is a proxy for the types of nonbank activities the BHC is engaged in and thus speaks to the "O" component of the BOPEC rating. The sixth variable is the ratio of a BHC's trading assets, whether located in banking or nonbanking subsidiaries, to its total assets, also as a proxy for BHC

---

[22]Some BHCs identify a lead bank in their regulatory filings. If a BHC did not report which bank was its lead bank, we assigned the lead status to its largest bank. Note that supervisors typically calculate an asset-weighted average CAMELS rating as an estimate of the bank rating component.

nonbanking activities.[23]

The seventh variable is the so-called "double leverage" ratio between the BHC and its lead bank, which is the ratio of the lead bank's equity capital to that of the parent's equity capital. This variable provides a measure of the soundness of the parent BHC, indicating the extent to which the parent's equity capital can be used to buffer against damage to the lead bank's equity capital. We use this variable as a proxy for the condition of the parent BHC as summarized in the "P" component of the BOPEC rating. The eighth variable is the BHC's return on average assets (ROAA), defined as the ratio of the four-quarter average of the BHC's net income to the four-quarter average of its assets. This variable is used to proxy for the "E" component of the BOPEC rating. The ninth variable is the BHC's ratio of equity capital to its total assets. This variable is used to proxy for the "C" component of the BOPEC rating.[24]

The tenth and final variable is the BOPEC rating assigned at the BHC's previous full scope, on-site inspection. This variable should contain information on BHC condition, although it should decay over time as per the studies on CAMELS ratings by Cole and Gunther (1998) and Hirtle and Lopez (1999). In addition, this variable can be viewed as a summary for other supervisory information, such as BHC liability structure, that is not explicitly captured by the other nine variables.

We refer to the version of the BOM model in equation (5) as the *core* model. The *extended* BOM model is specified as

$$BP_{it}^* = (\beta + \gamma I_{t-1})x_{it-2} + \pi z_{it-1}I_{t-1} + \theta I_{t-1} + \varepsilon_{it}, \tag{9}$$

where z_{it-1} is an $(l \times 1)$ vector of equity market variables corresponding to publicly traded BHC i in quarter $t-1$ and I_{t-1} is again the indicator variable for publicly traded BHCs.[25]

---

[23]Note that the trading assets variable as currently reported first became available in the first quarter of 1995. Before then, we proxy for BHC trading assets using the self-reported replacement costs of interest rate and foreign exchange derivative contracts.

[24]A variety of capital measures have been used in previous studies; see Evanoff and Wall (2000) and Estrella et al. (2000). We use a simple measure to facilitate comparison over the entire ten-year period.

[25]Note that the z_{it-1} variables for privately owned BHCs are set equal to zero.

Note that we include a fixed effect θI_{t-1} for publicly traded BHCs in order to account for the possibility that an equity market variable takes on a specific value of zero. Since equity market variables are observed at a higher frequency than the supervisory variables, we can include equity market information at the end of quarter $t-1$ for forecasting purposes.

We use four equity market variables in our analysis. The first two variables are derived from the cumulative BHC return between event window quarters $t-4$ and $t-1$ (or equivalently, the event window months $t-12$ and $t-3$). Specifically, we decompose the monthly returns into idiosyncratic and systematic components by rewriting equation (1) as $R_{it} = SR_{it} + AR_{it}$, where $SR_{it} = \widehat{\alpha} + \widehat{\beta}_1 R_{mt} + \widehat{\beta}_2 f f_t$ and $AR_{it} = \widehat{\varepsilon}_{it}$. We define the cumulative systematic returns CSR_{it-1} and cumulative abnormal returns CAR_{it-1} as the sums of these two components over the first three quarters (i.e., first nine months) of an event window centered at quarter t. Hence, the cumulative return is expressed as $CR_{it-1} = CSR_{it-1} + CAR_{it-1}$. We standardize these returns using the estimated standard error of CAR_{it-1}, as defined previously; that is, $\dfrac{CR_{it-1}}{\sqrt{V(CAR_{it-1})}} = \dfrac{CSR_{it-1}}{\sqrt{V(CAR_{it-1})}} + \dfrac{CAR_{it-1}}{\sqrt{V(CAR_{it-1})}}$, or equivalently, $SCR_{it-1} = SCSR_{it-1} + SCAR_{it-1}$. We use both the $SCSR_{it-1}$ and $SCAR_{it-1}$ variables in the model. We chose to include the SCAR variable based on the empirical results from the previous section. We include the systematic SCSR variable to account for economic factors that affect all BHC stock returns.

The third equity market variable is the EDF values, as previously discussed, for the last month of quarter $t-1$. The fourth equity market variable is a simple proxy for the BHC's underlying asset volatility. We first estimate the BHC's quarterly equity volatility using stock return data over the previous two years and then combine the volatility estimate with balance sheet data to get an asset volatility, as per Ronn and Verma (1986). This measure should provide insight into the extent to which EDF values, which contain a more sophisticated measure of asset volatility, provide additional information.

For a variety of reasons, the equity market variables are not available for all publicly traded BHCs over the entire sample period. For example, we cannot generate reliable SCARs when a BHC does not have at least two years of stock return data with which to estimate the two-factor market model, and several EDF values are missing throughout the dataset,

particularly prior to 1995. To address this issue, we replaced these missing values with the variable's in-sample mean for the available observations. We also included fixed effects to account for this data adjustment.[26] This procedure does not affect the model's coefficient estimates for the variables with missing values and allows us to use the entire sample in our estimation.[27]

5.1 Full sample estimation results

Table 10A presents the estimation results for three BOM model specifications; the simple BOM model without interacted terms, the core BOM model, and the extended BOM model containing the four equity market variables. Note that the standard errors are adjusted for the fact that there are only 1,399 unique BHCs among the 3,975 observations.[28] A p-value of less than 5% indicates that the coefficient is significantly different from zero. Recall that the model coefficients for the supervisory variables corresponding to privately-owned BHCs are simply the reported β coefficients, while the coefficients corresponding to publicly traded BHCs are $(\beta + \gamma)$.

Our estimation results can be summarized as follows. The interacted variables are clearly important to the model. We can reject the null hypothesis that $\gamma = 0$ as in the simple BOM model at the 5% significance level using the likelihood ratio test. Focusing on the core BOM model results, most of the supervisory variables are significant and have the appropriate sign. Larger BHCs tend to have lower BOPEC ratings (i.e., better ratings), which could potentially be attributed to larger banks having more diversified asset portfolios and better management. However, if the BHC is publicly traded, this effect is reduced. The results show

---

[26]See Griliches (1986) for a full description of ways to cope with this type of data problem. Specifically, we use the "regression" adjustment (described on page 1488) that assumes the π parameters are constant across the subsamples of missing and non-missing variables.

[27]We tested the robustness of the results to this adjustment by estimating the extended BOM model on a sample that excluded observations missing any equity market variable. The results were qualitatively similar to those reported here.

[28]Since we have multiple BOPEC rating assignments for several BHCs, we must adjust the estiamted standard errors to relax the assumption that all the observations are independent; see Rogers (1993).

a strong link between the supervisory assessments of banks and their holding companies. An increase in a lead bank's CAMELS rating (i.e., a supervisory downgrade) causes the parent's BOPEC rating to increase as well. If the BHC is publicly traded, this effect is only slightly reduced. Our results indicate that an increase in the BHC's problem loans or provisions for loan losses cause the BOPEC rating to increase.

We find that the presence of a Section 20 subsidiary causes BOPEC ratings to increase, but that this effect is effectively zero for public BHCs. Since all the BHCs with Section 20 subsidiaries in the sample are public, this result suggests that BHCs engaging in securities underwriting and related activities are not likely to cause increased supervisory concern. In fact, Cornett *et al.* (2002) found that several BHC risk measures did not change significantly after their Section 20 activities commenced. Our results indicate that an increase in the BHC's equity capital ratio causes its BOPEC rating to decrease and improve its supervisory standing. For publicly traded BHCs, this effect is even stronger. Finally, several variables had no impact within the core BOM model. Note that the lagged BOPEC ratings variable is statistically significant and indicates that higher prior ratings should lead to higher current ratings. Specifically, our results indicate that a BHC's double leverage and ROAA do not materially impact its assigned BOPEC rating.

The estimation results for the extended BOM model are presented in the last two columns of Table 10A. The equity market variables are clearly significant and contribute to the model's overall fit. The null hypothesis that the equity market variables are not significant is rejected with a likelihood ratio statistic of 100.258, which is significant at the 1% level under a $\chi^2(8)$ distribution.[29] In addition, the coefficient estimates for the supervisory variables are practically unchanged from the core BOM model estimation results. The results indicate that a decrease in a BHC's three-quarter SCAR has an adverse impact on its BOPEC rating. Similarly, a decrease in the three-quarter SCSR portion of the BHC stock return also has an adverse impact on ratings. However, in contrast to the EDF event study, the results do not indicate that a change in a BHC's EDF value has an effect on its BOPEC rating. Our

[29]Note this likelihood ratio is distributed $\chi^2(8)$ due to the four equity market variables and the four fixed effects accounting for publicly-traded BHCs and publicly-traded BHCs with missing equity returns, EDF values or Ronn-Verma asset volatilities.

results also indicate that an increase in BHC asset volatility would not have an impact on its BOPEC ratings. In conclusion, the full-sample estimation results clearly indicate that equity market variables contribute to the modeling of BOPEC ratings and could be useful in off-site supervisory monitoring models.

As described in equations (5) and (9), we chose to model the BOPEC ratings themselves. However, an alternative model specification is to model the changes in BOPEC ratings from their previously assigned values.[30] That is, instead of modeling BP_{it}^*, we could model ΔBP_{it}^*, which would translate into possible ΔBP_{it} values of $\{\leq -2, -1, 0, 1, \geq 2\}$. Note that negative values correspond to downgrades and positive values correspond to upgrades; thus, since now lower numbers correspond to downgrades, we should expect the signs on the explanatory variables to be reversed. As this choice is effectively an empirical question, Table 10B presents the parameter estimates for the alternative BOM models. The results are qualitatively similar to those of Table 10A. Specifically, the significance of the parameter estimates for the three models of BOPEC ratings are almost identical to the three models for BOPEC changes. The main difference is for the lagged BOPEC ratings, which is positive for both specifications. This result suggests that higher (i.e., worse) lagged ratings lead to worse BOPEC ratings initially (in the level specification) and to improvements in BOPEC ratings over time (in the change specification). Taken together, the results suggest that there is persistence in BOPEC ratings, but also a tendency to improve over time. Given that the two sets of results are qualitatively similar, we focus on modeling the BOPEC ratings themselves in the rest of the paper.

5.2 The BOM model as a supervisory monitoring model

The BOM model presented above indicates that BOPEC ratings can be modeled as a function of lagged supervisory and equity market variables. However, in order to be a useful tool for supervisory monitoring, the model must be able to generate reasonably accurate forecasts of BOPEC ratings that supervisors could use to determine if and when additional supervisory actions are necessary. In this section, we show how to implement the BOM model

---

[30]We thank a referee for raising this issue.

as a supervisory model. We find that its forecasts are reasonably accurate and capable of anticipating future BOPEC changes.

5.2.1 Rolling estimation samples

To make the proposed BOM model operational, we need to permit it to adjust to current economic and banking conditions in the hope of producing more accurate forecasts. In the spirit of the SEER model, we address this concern by allowing the model parameters to be time-varying and by re-estimating the model every quarter based on a rolling estimation sample of four quarters. Specifically, let $S(t)$ denote the sample of BHCs assigned a BOPEC rating in the four-quarter period ending in period t.[31] For our sample, $S(0)$ encompasses the four quarters of 1989, which are used to generate forecasts for 1990:Q1, and $S(39)$ encompasses 1998:Q4 through 1999:Q3, which are used to generate forecasts for 1999:Q4. The average number of BOPEC ratings in $S(t)$ for our 40 subsample periods was about 400.

As before, each BHC in the $S(t)$ subsample has corresponding vectors of lagged explanatory variables $x_{iS(t)}$ of dimension $(k \times 1)$ and $z_{iS(t)}$ of dimension $(l \times 1)$. The rolling BOM model is specified as

$$BP^*_{iS(t)} = (\beta_{S(t)} + \gamma_{S(t)}I_{S(t)})x_{iS(t)} + \pi_{S(t)}z_{iS(t)}I_{S(t)} + \varepsilon_{iS(t)}, \tag{10}$$

where the supervisory variables are lagged by two quarters and the equity market variables are lagged by one quarter. Note that the fixed effect parameters are also used in the subsample regressions. We estimate the BOM model parameters using the BOPEC assignments in subsample $S(t)$ and use them to forecast the one-quarter-ahead BOPECs for all BHCs filing regulatory reports in quarter $t - 1$.

Table 11 summarizes the subsample estimation results for the core and extended BOM models by presenting the percentage of periods for which the variables are significant. Note

---

[31]We considered conducting this analysis using two-quarter rolling sample periods, as currently used in the SEER model. However, since fewer BOPEC ratings are assigned in a quarter, the estimation samples contain relatively few observations; i.e., less than 200 per sample periods. This relatively small sample size caused a concern in interpreting and using the estimated coefficients.

that the standard errors are again adjusted for possible multiple observations from a single BHC. However, the short, four-quarter estimation periods limit this concern. For the core model, lagged BOPEC rating, total assets, lead bank CAMELS rating, problem loans, ALLL and equity capital variables all have the expected signs and are statistically significant for more than 50% of the subsamples. This matches the full sample result. However, the results for the Section 20 subsidiary variable and for the interacted variables for publicly traded BHCs do not consistently match the full sample results, most likely because of the reduced number of publicly traded BHCs in a given subsample.

The results for the extended BOM model are qualitatively similar. The lagged BOPEC rating, total assets, lead bank CAMELS rating, problem loans, ALLL, and equity capital variables are significant for the majority of the subsamples. Once again, the results for the Section 20 subsidiary variable and the interacted variables do not consistently match the full sample results. In contrast to the full sample results, the coefficient estimates for the four equity market variables are not consistently significant across all subsample periods.

Why are the equity market variables statistically insignificant within certain subsamples and not in others? The probable answer is that, in a given four-quarter period, there may be simply too few publicly traded BHCs to capture the significance of the equity market variables. This explanation is supported by the fact that the publicly traded interaction terms do not come in as strong in the subsample results as they do in the full sample. The actual percentages of BOPEC assignments for public BHCs in the subsample periods range from 7% to 12% of the subsample total.

5.2.2 BOPEC forecast analysis

The subsample estimation results indicate that both the core and extended BOM models retain most of the descriptive characteristics of the full-sample estimation. Although these results are of independent interest, the purpose of conducting the subsample analysis was to generate BOPEC forecasts for potential use by supervisors in monitoring BHCs. In this section, we evaluate the accuracy of these forecasts using an event study framework. As with the EDF event study, we adjust the BOPEC forecasts by subtracting the median forecasts of

the similarly rated BHCs in that quarter. This procedure helps to remove systematic factors in the forecasts and isolates their idiosyncratic components. If a BHC is to be upgraded, then its adjusted forecast relative to its BOPEC peer group should be negative, and if a BHC is to be downgraded, then its adjusted forecast should be positive. To test these null hypotheses, we use the sign test at one to four quarters before the BOPEC assignments.

Figure 4 plots the medians of the adjusted forecasts, and Tables 12A and 12B present the sign test results for both versions of the BOM model. The null hypothesis that the median of the adjusted forecasts is zero is not rejected for inspections that resulted in no change in BOPEC rating. This result mirrors the evidence presented in the other event studies. The more interesting question is whether the null hypothesis can be rejected for BOPEC upgrades and downgrades, and the empirical results reject these null hypotheses at the 1% significance level for both models for all four forecast horizons with the correct sign. These null hypotheses are also strongly rejected for upgrades and downgrades across the regulatory threshold between BOPEC-2 and BOPEC-3. Hence, the BOPEC forecasts are capable of anticipating BOPEC rating assignments up to four quarters prior.

Tables 13A and 13B present analysis of the distribution of BOPEC outcomes conditional on whether the BOPEC forecasts from the core and extended BOM models are more than one full rating away from the BHCs' current ratings. These cutoff points were selected to provide a comparison with the results in Tables 6 and 8. As before, the BOPEC no changes are the largest category and are forecasted well at all horizons; that is, the percentages of BOPEC no changes observed when the core and extended BOM forecasts "close" to the current rating (i.e., abs(forecast - current BOPEC) ≤ 1) are about 66%. Both sets of forecasts also accurately signal BOPEC changes at all four horizons. Downgrades are correctly signaled about 50% to 85% of the time, and upgrades about 50% to 90% of the time. Overall, both sets of adjusted BOM forecasts provide the correct signal about 66% of the time at all four horizons. In fact, as shown by the Pearson statistics reported in these tables, both sets of forecasts reject the null hypothesis that the conditional probabilities of BOPEC outcomes are equal to the unconditional probabilities. These results clearly indicate that the adjusted BOM forecasts are capable of detecting BOPEC assignments up to four quarters prior and

would thus be useful for supervisory monitoring purposes.

The sign test results indicate that both models forecast well, but an interesting question is whether the extended BOM model with its additional equity market variables performs better. The last column of Table 13B contains the Pearson statistics testing the null hypothesis that the extended BOM forecasts are not better than the core BOM forecasts. This null hypothesis is not rejected at any horizon, indicating that the additional equity market variables do not significantly improve the conditional forecasts based on the core BOM model. Hence, from the standpoint of forecast accuracy, these results suggest that supervisors may not need to incorporate equity market information into the BOM model to generate useful forecasts. However, given the low cost of doing so and the potential value of such information, we recommend that the extended BOM model be used for supervisory monitoring purposes.

6   Conclusion

This paper investigates two related issues regarding supervisory off-site monitoring of bank holding companies. First, we show that equity market variables anticipate BOPEC changes. Specifically, we find that cumulative abnormal stock returns and EDF values anticipate BOPEC changes by up to four quarters in advance. These results suggest that supervisors could derive useful monitoring information from these equity market signals.

Second, we propose a BOPEC off-site monitoring (BOM) model for summarizing supervisory information in a simple and practical way. As with the SEER model for bank CAMELS ratings, the BOM model could provide supervisors with a tool for detecting potentially significant BOPEC changes ahead of time. We also extend this model to include equity market variables and take advantage of the information that they are shown to contain. Our forecasting results indicate that the BOM forecasts do anticipate BOPEC assignments, but that including the equity market variables in the model does not improve forecast accuracy in a statistically significant way.

However, we argue that using the extended BOM model, which includes equity market

variables, should still have practical value for supervisors. Our argument is based on five points. First, equity market variables are useful in-sample and improve the BOM model's fit to the data. Second, including the equity market variables does not improve, but also does not degrade, the out-of-sample forecast performance of the core BOM forecasts. In fact, as shown in Table 13B, the extended BOM forecasts slightly improve the forecast accuracy for the BOPEC upgrades.

Third, equity market data is available on a more timely basis than standard supervisory data. The higher frequency equity data could potentially signal important BHC changes sooner than standard supervisory reporting or confirm supervisory concerns that might arise over the short term. Fourth, since the cost of incorporating equity market variables into supervisory off-site monitoring by using the BOM model is quite low, even small net improvements in forecast accuracy could still be of value. As noted by Stern (2001), the benefits of incorporating financial market data into the supervisory process "seem real enough and costs of increased use low enough" to warrant doing so. Finally, market assessments should provide an alternative opinion on BHC condition that may serve as a reasonable cross-check of supervisory evaluations, both for supervisors and for the public at large.

Note that our empirical results and conclusions are necessarily based on historical data. We have no guarantee that equity market signals, which were shown to be of some use given the current supervisory regime, would continue to be useful under a regime that explicitly incorporated market signals into supervisory policy. If a change in policy served to lessen the sense that equity market investors feel at risk, then the policy change could destroy their incentive to monitor and produce this information. This version of the Lucas critique must be acknowledged. However, as long as supervisors use equity market information solely for monitoring purposes and not for formal policy actions, the incentives to equity market investors should not be affected.

References

[1] Allen, L., Jagtiani, J., and Moser, M., 2001. "Further Evidence on the Information Content of Bank Examination Ratings: A Study of BHC-to-FHC Conversion Applications," *Journal of Financial Services Research*, 20, 213-232.

[2] Altman, E., 1968. "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *Journal of Finance*, 23, 589-609.

[3] Basel Committee on Banking Supervision, 2001. "The New Basel Accord." Consultative Paper, Bank for International Settlements. (http://www.bis.org/publ/bcbsca.htm)

[4] Berger, A. N. and Davies, S. M., 1998. "The Information Content of Bank Examinations," *Journal of Financial Services Research*, 14, 117-144.

[5] Berger, A.N., Davies, S.M., and Flannery, M.J., 2000. "Comparing Market and Supervisory Assessments of Bank Performance: Who Knows What When?," *Journal of Money, Credit and Banking*, 32, 641-667.

[6] Bliss, R., 2000. "The Pitfalls in Inferring Risk from Financial Market Data," Manuscript, Economic Research Department, *Federal Reserve Bank of Chicago*.

[7] Bliss, R. and Flannery, M., 2001. "Market Discipline in the Governance of U.S. Bank Holding Companies: Monitoring versus Influence," in Mishkin, R., ed. Prudential Supervision: Why Is It Important and What Are the Issues?, pp. 107-143. Chicago: University of Chicago Press.

[8] Board of Governors of the Federal Reserve System, 2001. A User's Guide for the Bank Holding Company Performance Report. (http://www.federalreserve.gov/boarddocs/supmanual/bhcpr/bhcpr_2000_access.pdf)

[9] Campbell, J., Lo, A., and MacKinlay, A., 1997. The Econometrics of Financial Markets. Princeton University Press.

[10] Cole, R.A., Cornyn, B.G. and Gunther, J.W., 1995. "FIMS: A New Monitoring System for Banking Institutions," *Federal Reserve Bulletin*, January, 1-15.

[11] Cole, R.A. and Gunther, J.W., 1998. "Predicting Bank Failures: A Comparison of On- and Off-Site Monitoring Systems," *Journal of Financial Services Research*, 13, 103-17.

[12] Cornett, M.M., Ors, E. and Tehranian, H., 2002. "Bank Performance Around the Introduction of a Section 20 Subsidiary," *Journal of Finance*, 57, 501-521.

[13] Crosbie, P.J. and Bohn, J.R., 2002. "Modeling Default Risk." Manuscript, KMV, LLC. (http://www.kmv.com/Knowledge_Base/public/crm/white/Modeling_Default_Risk.pdf)

[14] Curry, T.J., Elmer, P.J. and Fissel, G., 2001. "Regulator Use of Market-Related Data to Improve the Identification of Bank Financial Health," Manuscript, Federal Deposit Insurance Corporation.

[15] DeFerrari, L. and Palmer, D., 2001. "Supervision of Large Complex Banking Organizations." *Federal Reserve Bulletin*, February, 47-57.

[16] DeYoung, R., Flannery, M., Lang, M., and Sorescu, S., 2001. "The Information Content Bank Exam Ratings and Subordinated Debt Prices." *Journal of Money, Credit, and Banking*, 33, 900-925.

[17] Delianedis, G. and Geske, R., 1998. "Credit Risk and Risk Neutral Default Probabilities: Information about Ratings Migrations and Defaults," Finance Working Paper #19-98, The Anderson School at UCLA.

[18] Diamond, D.W., 1984. "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51, 393-414.

[19] Elmer, P.J. and Fissel, G., 2001. "Forecasting Bank Failure from Momentum Patterns in Stock Returns," Manuscript, Federal Deposit Insurance Corporation.

[20] Estrella, A., Park, S. and Peristiani, S., 2000. "Capital Ratios as Predictors of Bank Failure," *Federal Reserve Bank of New York Economic Policy Review*, July, 33-52.

[21] Evanoff, D., and Wall, L., 2000. "Sub-debt Yield Spreads as Bank Risk Measures," *Journal of Financial Services Research*, 20, 121-146.

[22] Flannery, M., 1998. "Using Market Information in Prudential Bank Supervision," *Journal of Money, Credit, and Banking,* August, Part I, 273-305.

[23] Flannery, M.J., 2001. "The Faces of 'Market Discipline'," *Journal of Financial Services Research*, 20, 107-121.

[24] Flannery, M., Kwan, S., and Nimalendran, M., 1999. "Market Evidence on the Opaqueness of Banking Firms' Assets," Working paper #99-11, Federal Reserve Bank of San Francisco.

[25] Flannery, M. and Sorescu, S., 1996. "Evidence of Bank Market Discipline in Subordinated Debenture Yields: 1983-1991," *Journal of Finance.* Vol LI, No. 4, 1347-1377.

[26] Gilbert, A., Meyer, A. and Vaughn, M., 2001. "Can Feedback from the Jumbo CD Market Improve Off-Site Surveillance of Small Banks?," Manuscript, Economic Research Department, Federal Reserve Bank of St. Louis.

[27] Gorton, G. and Santomero, A., 1990. "Market Discipline and Bank Subordinated Debt," *Journal of Money, Credit and Banking.* 22, 119-128.

[28] Griliches, Z., 1986. "Economic Data Issues," in Griliches, Z. and Intriligator, M.D., eds. Handbook of Econometrics, Volume III. Elsevier Science Publishers BV.

[29] Gropp, R., Vesala, J. and Vulpes, G., 2001. "Equity and Bond Market Signals as Leading Indicators of Bank Fragility," Manuscript, European Central Bank.

[30] Gunther, J.W., Levonian, M.E. and Moore, R.R., 2001. "Can the Stock Market Tell Bank Supervisors Anything They Don't Already Know?," *Federal Reserve Bank of Dallas Economic and Financial Review,* Second Quarter, 2-9.

[31] Hall, J.R., King, T.B., Meyer, A.P. and Vaughn, M.D., 2001. "What Can Bank Supervisors Learn from the Equity Markets? A Comparison of the Factors Affecting Market-Based Risk Measures and BOPEC Scores," Manuscript, Federal Reserve Bank of St. Louis.

[32] Hancock, D. and Kwast, M., 2000. "Using Subordinated Debt to Monitor Bank Holding Companies: Is it Feasible?," *Journal of Financial Services Research*, 20, 147-188.

[33] Hannan, T. and Hanweck, G., 1988. "Bank Insolvency Risk and the Market for Large Certificates of Deposit," *Journal of Money, Credit, and Banking.* 20(2), 203-211.

[34] Hirtle, B. and Lopez, J.A., 1999 "Supervisory Information and the Frequency of Bank Examinations," *Federal Reserve Bank of New York Economic Policy Review*, April, 1-20.

[35] Kwast, M., et al., 1999. "Using Subordinated Debt as an Instrument of Market Discipline: Report of the Study Group on Subordinated Notes and Debentures." Federal Reserve System Staff Study # 172, Board of Governors of the Federal Reserve System.

[36] Merton, R., 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance,* 29, 449-470.

[37] Meyer, P. and Pifer, H., 1970. "Prediction of Bank Failures," *Journal of Finance*, 25, 853-868.

[38] Morgan, D., 2000. "Rating Banks: Risk and Uncertainty in an Opaque Industry." Working paper #105, Federal Reserve Bank of New York.

[39] Pettway, R.H. and Sinkey, J.F., 1980. "Establishing On-Site Bank Examination Priorities: An Early-Warning System using Accounting and Market Information," *Journal of Finance*, 35, 137-150.

[40] Rogers, W.H., 1993. "Regression Standard Errors in Clustered Samples," *Stata Technical Bulletin*, 13, 19-23.

[41] Ronn, E. and Verma, A., 1986. "Pricing Risk-Adjusted Deposit Insurance: An Option-Based Model," *Journal of Finance*, 41, 871-895.

[42] Sinkey, J.E., 1975. "A Multivariate Statistical Analysis of the Characteristics of Problem Banks," *Journal of Finance*, 20, 21-36.

[43] Sahajwala, R. and Van der Bergh, P., 2000. "Supervisory Risk Assessment and Early Warning Systems," Basel Committee on Banking Supervision Working Paper No. 4, Bank for International Settlements.

[44] Saunders, A., 2001. "Comments on Evanoff and Wall / Hancock and Kwast," Journal of Financial Services Research, 20, 189-194.

[45] Stern, G.H., 2001. "Taking Market Data Seriously," *Federal Reserve Bank of Minneapolis Region*, September, 8-11

Table 1. Share of BHCs and BHC assets that are publicly traded

	1990	1999
Number of BHCs	1,291	1,402
Number of public BHCs	242	327
% public	18.8%	23.3%
Total BHC assets ($B)	$2,152.9	$4,609.2
Total public BHC assets ($B)	$1,579.2	$3,970.0
% public	73.4%	86.1%

Note: The definition of a bank holding company (BHC) used in this table is the definition used in constructing our dataset; i.e., a top-tier BHC with an identifiable lead bank and four quarters of available supervisory reporting data.

Table 2. Asset size of the BHCs in the BOPEC sample

	1990-1999	1990-1994	1995-1999
Total inspections	3,975	1,986	1,989
Asset Size:			
Assets > $100B	50	13	37
$1B < assets < $100B	1,278	605	673
Assets<$1B	2,647	1,368	1,279
Inspections of publicly traded BHCs	1,602	762	840
Asset Size:			
Assets > $100B	48	13	35
$1B < assets < $100B	999	494	505
Assets<$1B	555	255	300

Note: The definition of a bank holding company used in this table is the definition used in constructing our dataset; i.e., a top-tier BHC with an identifiable lead bank and four quarters of available supervisory reporting data.

Table 3A. BOPEC ratings in the sample from 1990 through 1999

	BOPEC Rating						% of total				
	1	2	3	4	5	Total	1	2	3	4	5
1990	54	166	63	25	2	310	17.4	53.5	20.3	8.1	0.6
1991	58	163	82	30	6	339	17.1	48.1	24.2	8.8	1.8
1992	67	224	79	49	4	423	15.8	53.0	18.7	11.6	0.9
1993	109	253	62	26	2	452	24.1	56.0	13.7	5.8	0.4
1994	149	251	40	18	4	462	32.3	54.3	8.7	3.9	0.9
1995	173	251	36	19	3	482	35.9	52.1	7.5	3.9	0.6
1996	218	232	21	2	1	474	46.0	48.9	4.4	0.4	0.2
1997	203	222	18	1	0	444	45.7	50.0	4.1	0.2	0.0
1998	132	134	17	3	1	287	46.0	46.7	5.9	1.0	0.3
1999	124	156	17	5	0	302	41.1	51.7	5.6	1.7	0.0
Total	1287	2052	435	178	23	3,975	32.4	51.6	10.9	4.5	0.6

Table 3B.
BOPEC ratings for publicly traded BHCs in the sample from 1990 through 1999

	BOPEC Rating						% of total				
	1	2	3	4	5	Total	1	2	3	4	5
1990	25	72	18	8	1	124	20.2	58.1	14.5	6.5	0.8
1991	22	63	31	12	1	129	17.1	48.8	24.0	9.3	0.8
1992	36	72	28	19	2	157	22.9	45.9	17.8	12.1	1.3
1993	51	86	20	12	0	169	30.2	50.9	11.8	7.1	0.0
1994	67	91	17	7	1	183	36.6	49.7	9.3	3.8	0.5
1995	68	101	16	3	2	190	35.8	53.2	8.4	1.6	1.1
1996	87	91	6	0	0	184	47.3	49.5	3.3	0.0	0.0
1997	78	79	2	1	0	160	48.8	49.4	1.3	0.6	0.0
1998	73	66	7	2	0	148	49.3	44.6	4.7	1.4	0.0
1999	71	79	6	2	0	158	44.9	50.0	3.8	1.3	0.0
Total	578	800	151	66	7	1,602	36.1	49.9	9.4	4.1	0.4

Note: The BOPEC ratings are from full scope, on-site inspections from 1990 through 1999. The definition of a bank holding company used in this table is the definition used in constructing our dataset; i.e., a top-tier BHC with an identifiable lead bank and four quarters of available supervisory reporting data.

Table 4A. BOPEC rating changes in the sample from 1990 through 1999

	Change in BOPEC rating				% of total		
	Upgrade	No change	Downgrade	Total	Upgrade	No change	Downgrade
1990	45	176	89	310	14.5	56.8	28.7
1991	47	184	108	339	13.9	54.3	31.9
1992	85	244	94	423	20.1	57.7	22.2
1993	120	273	59	452	26.5	60.4	13.1
1994	117	296	49	462	25.3	64.1	10.6
1995	134	283	65	482	27.8	58.7	13.5
1996	121	318	35	474	25.5	67.1	7.4
1997	109	292	43	444	24.5	65.8	9.7
1998	43	206	38	287	15.0	71.8	13.2
1999	36	213	53	302	11.9	70.5	17.5
Total	857	2485	633	3,975	21.6	62.5	15.9

Table 4B.
BOPEC rating changes for publicly traded BHCs in the sample from 1990 through 1999

	Change in BOPEC rating				% of total		
	Upgrade	No change	Downgrade	Total	Upgrade	No change	Downgrade
1990	16	74	34	124	12.9	59.7	27.4
1991	8	80	41	129	6.2	60.0	31.8
1992	29	93	35	157	18.5	59.2	22.3
1993	53	105	11	169	31.4	62.1	6.5
1994	46	117	20	183	25.1	63.9	10.9
1995	51	117	22	190	26.8	61.6	11.6
1996	42	130	12	184	22.8	70.7	6.5
1997	34	113	13	160	21.3	70.6	8.1
1998	18	113	17	148	12.2	76.4	11.5
1999	15	121	22	158	9.5	76.6	13.9
Total	312	1063	227	1,602	19.5	66.4	14.2

Note: The BOPEC ratings are from full scope, on-site inspections from 1990 through 1999. The change in BOPEC ratings is relative to previously assigned rating. The definition of a bank holding company used in this table is the definition used in constructing our dataset; i.e., a top-tier BHC with an identifiable lead bank and four quarters of available supervisory reporting data.

Table 5. BHC event study based on CARs and SCARs

Month	BOPEC upgrades CARs (%)	BOPEC upgrades SCARs	BOPEC no changes CARs (%)	BOPEC no changes SCARs	BOPEC downgrades CARs (%)	BOPEC downgrades SCARs
-12	2.04	2.323*	-0.08	-0.294	-0.39	-0.461
-11	2.67	2.143*	-0.35	-0.827	-1.54	-1.272
-10	4.15	2.692*	-0.15	-0.288	-4.43	-2.952*
-9	3.52	1.959	0.58	0.976	-7.01	-4.006*
-8	5.02	2.470*	0.53	0.786	-9.83	-4.964*
-7	6.02	2.681*	0.97	1.297	-10.71	-4.866*
-6	6.67	2.725*	0.58	0.718	-10.49	-4.398*
-5	5.93	2.247*	0.61	0.699	-10.90	-4.247*
-4	5.98	2.116*	0.43	0.459	-12.36	-4.500*
-3	7.32	2.434*	0.44	0.436	-15.28	-5.242*
-2	8.04	2.531*	0.15	0.144	-14.79	-4.786*
-1	8.22	2.458*	0.59	0.534	-14.51	-4.430*
0	10.13	2.884*	0.38	0.328	-13.19	-3.852*
1	10.45	2.845*	-0.11	-0.092	-15.19	-4.233*
2	12.12	3.163*	-0.37	-0.292	-16.68	-4.467*
3	13.43	3.374*	-0.92	-0.695	-14.58	-3.743*
# obs.		139		589		85

Month	BOPEC upgrades above threshold CARs	BOPEC upgrades above threshold SCARs	BOPEC downgrades below threshold CARs	BOPEC downgrades below threshold SCARs
-12	2.10	1.201	-1.11	-0.869
-11	2.68	1.079	-1.40	-0.764
-10	4.71	1.534	-4.06	-1.797
-9	5.12	1.426	-6.83	-2.588*
-8	8.06	1.990	-11.85	-3.955*
-7	9.76	2.179*	-14.97	-4.494*
-6	12.08	2.474*	-13.78	-3.819*
-5	12.33	2.339*	-12.63	-3.247*
-4	11.65	2.057*	-13.46	-3.238*
-3	13.69	2.276*	-16.14	-3.662*
-2	16.44	2.577*	-15.60	-3.339*
-1	17.54	2.610*	-15.10	-3.046*
0	21.15	2.993*	-12.57	-2.421*
1	21.28.	2.877*	-14.94	-2.746*
2	25.14	3.258*	-16.24	-2.867*
3	26.28	3.277*	-11.02	-1.866
# obs.		50		39

Note: This table presents the average cumulative abnormal returns (CARs) and standardized CARs (SCARs) for the twelve months prior to and three months after the BOPEC assignments in our sample. An asterisk denotes statistical significance at the 5% level.

Figure 1. Event study CARs for BOPEC upgrades, downgrades and no changes

Average Cumulative Abnormal Returns

Date 0 is the exit date
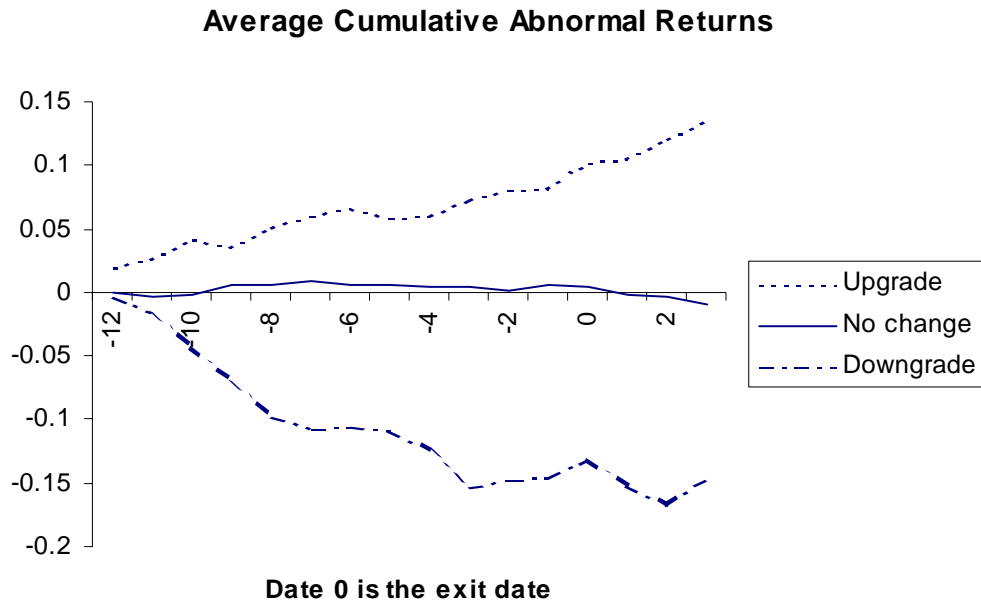
Figure 2. Event study SCARs for BOPEC upgrades, downgrades and no changes

Standardized Cumulative Abnormal Returns

Date 0 is the exit date

40
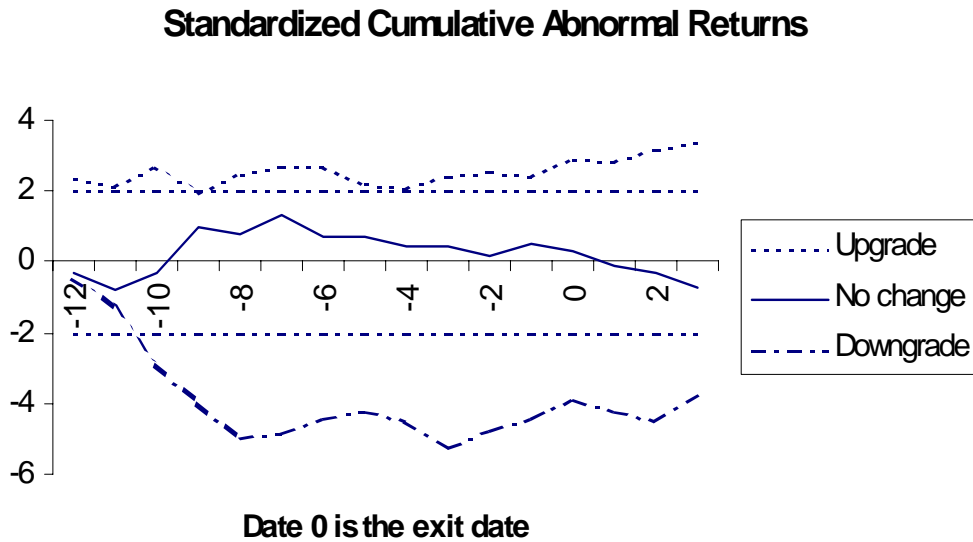
Table 6. Distribution of BOPEC outcomes after conditioning on the SCAR variable

SCAR @ -12 months	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	Pearson stat.
SCAR<-1.96	26	0.0	73.1	26.9		
|SCAR|≤1.96	749	18.0	72.0	10.0		
SCAR>+1.96	38	10.5	81.6	7.9	75.4	17.14*

SCAR @ -9 months	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	Pearson stat.
SCAR<-1.96	25	8.0	52.0	40.0		
|SCAR|≤1.96	756	17.1	73.4	9.5		
SCAR>+1.96	32	25.0	65.6	9.4	81.5	33.91*

SCAR @ -6 months	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	Pearson stat.
SCAR<-1.96	33	3.0	54.5	42.4		
|SCAR|≤1.96	749	17.1	73.8	9.1		
SCAR>+1.96	31	32.3	58.1	9.7	86.4	58.77*

SCAR @ -3 months	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	Pearson stat.
SCAR<-1.96	44	9.1	52.3	38.6		
|SCAR|≤1.96	733	17.5	73.7	8.9		
SCAR>+1.96	36	19.4	72.2	8.3	88.2	50.94*

SCAR @ -1 month	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	Pearson stat.
SCAR<-1.96	44	6.8	63.6	29.5		
|SCAR|≤1.96	739	17.9	72.6	9.5		
SCAR>+1.96	30	13.3	80.0	6.7	82.4	24.82*

Note: This table presents the forecast accuracy results based on conditioning on the SCAR variable at different horizons. The cells in bold indicate the outcome expected based on the signal observed. Note that the number of observations changes with the horizon due to data restrictions. The percentage columns represent the percentage of forecasts (i.e., number of observations in a row that were correct). The "total % correct" column reports the accuracy of all the forecasts. The Pearson goodness-of-fit statistic tests the null hypothesis that the distribution of BOPEC outcomes conditional on the SCAR variable is not different from the unconditional distribution and that the SCAR variable is not informative. The statistic is distributed χ_8^2. An asterisk denotes significance at the 5% level.

Table 7. BHC event study based on adjusted EDF values

Month	BOPEC upgrades Median	BOPEC upgrades Sign test p-value (%)	BOPEC no changes Median	BOPEC no changes Sign test p-value (%)	BOPEC downgrades Median	BOPEC downgrades Sign test p-value (%)
-12	-0.0050	32.13	0.0000	13.25	0.1925	0.00*
-11	-0.0100	17.56	0.0000	48.34	0.1750	0.00*
-10	-0.0200	17.97	0.0000	62.76	0.2150	0.00*
-9	-0.0400	1.54*	0.0000	30.94	0.2125	0.00*
-8	-0.0425	0.27*	0.0000	100.0	0.2550	0.00*
-7	-0.0450	0.44*	0.0000	92.96	0.1775	0.00*
-6	-0.0550	0.59*	0.0000	56.64	0.2150	0.00*
-5	-0.0525	0.08*	-0.0100	12.91	0.2550	0.00*
-4	-0.0500	0.03*	-0.0250	14.78	0.3625	0.00*
-3	-0.0400	0.04*	0.0000	25.28	0.3800	0.00*
-2	-0.0500	0.34*	-0.0100	9.71	0.4650	0.00*
-1	-0.0475	0.21*	-0.0100	1.13*	0.3550	0.00*
0	-0.0475	0.02*	-0.0100	0.88*	0.3950	0.00*
1	-0.0600	0.22*	-0.0100	2.52*	0.3350	0.00*
2	-0.0500	0.04*	-0.0100	1.77*	0.2475	0.00*
3	-0.0500	0.00*	-0.0100	0.23*	0.2250	0.00*\
# obs.		122		554		66

Month	BOPEC upgrades above threshold Median	BOPEC upgrades above threshold Sign test p-value (%)	BOPEC downgrade below threshold Median	BOPEC downgrade below threshold Sign test p-value (%)
-12	-0.0200	42.78	0.2875	0.22*
-11	-0.0150	17.25	0.3075	0.04*
-10	-0.0275	29.23	0.5325	0.22*
-9	-0.1000	10.02	0.5000	0.02*
-8	-0.1700	4.94*	0.4875	0.04*
-7	-0.1375	7.48	0.6025	0.00*
-6	-0.1850	4.94*	0.6175	0.00*
-5	-0.1450	8.10	0.6400	0.01*
-4	-0.1675	2.14*	0.6875	0.00*
-3	-0.0850	3.07*	0.7525	0.00*
-2	-0.0950	4.94*	0.7650	0.00*
-1	-0.1550	10.02	0.7550	0.00*
0	-0.1550	7.48	0.6650	0.00*
1	-0.1025	2.14*	0.6950	0.00*
2	-0.1000	3.54*	0.5250	0.00*
3	-0.1150	0.63*	0.4800	0.00*
# obs.		32		22

Note: This table presents the medians of the adjusted EDF values arranged in event time from month -12 to month +3. The sign test p-values are the p-values for the null hypothesis that the medians are equal to zero. An asterisk denotes statistical significance at the 5% level.

Figure 3. Medians of adjusted EDF values in event time (Units: %)

Median-Adjusted EDF values

Median-Adjusted EDF values for Inspectioins Crossing the BOPEC 2-3 Threshold
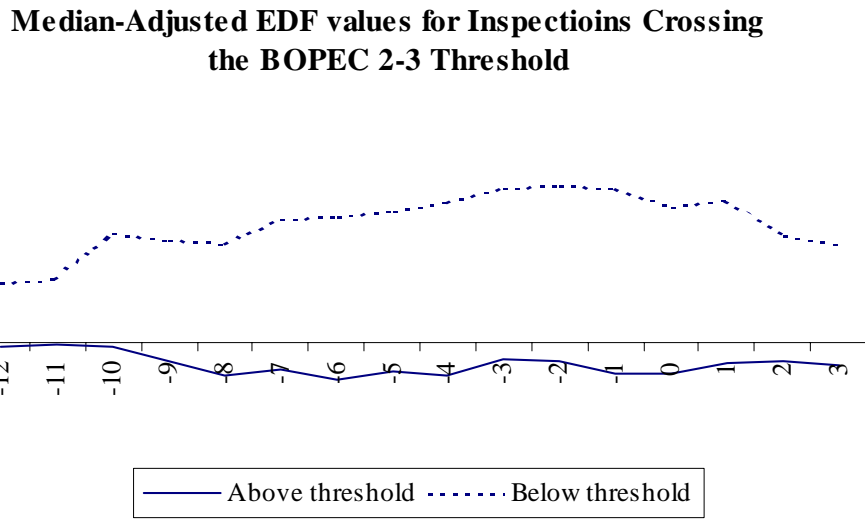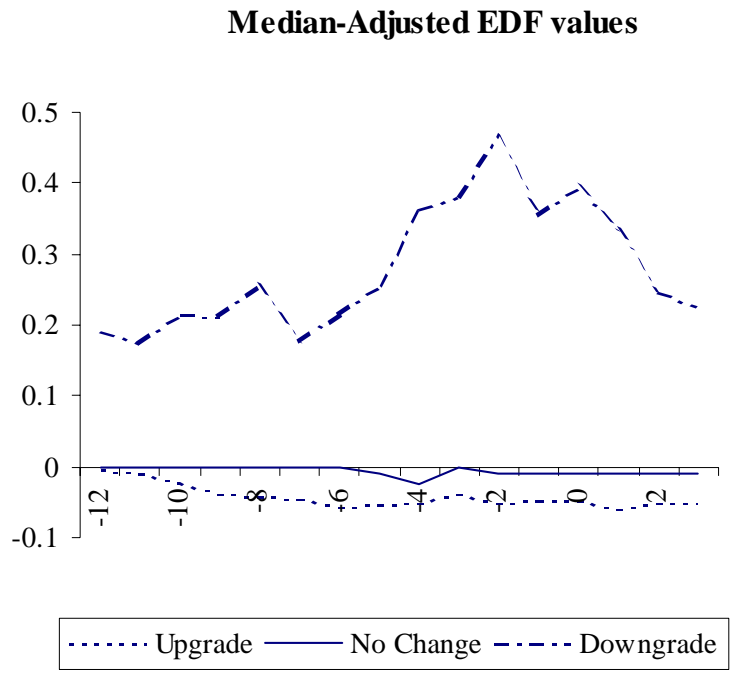
Table 8. Distribution of BOPEC outcomes after conditioning on the adjusted EDF values

Adj. EDF value @ -12 months	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	Pearson stat.
Adj. EDF<-0.25	25	56.0	40.0	4.0		
-.25≤Adj. EDF≤.93	676	14.6	78.1	7.2		
Adj. EDF >0.93	41	22.0	39.0	39.0	75.2	127.81[*]

Adj. EDF value @ -9 months	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	Pearson stat.
Adj. EDF<-0.25	33	51.5	39.4	9.1		
-.25≤Adj. EDF≤.93	667	14.5	78.3	7.2		
Adj. EDF >0.93	42	19.1	45.2	35.7	74.7	116.89[*]

Adj. EDF value @ -6 months	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	Pearson stat.
Adj. EDF<-0.25	24	62.5	33.3	4.2		
-.25≤Adj. EDF≤.93	678	14.6	77.7	7.7		
Adj. EDF >0.93	40	20.0	47.5	32.5	74.8	109.17[*]

Adj. EDF value @ -3 months	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	Pearson stat.
Adj. EDF<-0.25	30	56.7	40.0	3.3		
-.25≤Adj. EDF≤.93	665	14.6	78.6	6.8		
Adj. EDF >.93	47	17.0	40.4	42.6	75.4	161.46[*]

Adj. EDF value @ -1 month	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	Pearson stat.
Adj. EDF<-0.25	31	54.8	41.9	3.2		
-.25≤Adj. EDF≤.93	666	14.6	78.5	6.9		
Adj. EDF >0.93	45	17.8	40.0	42.2	75.3	152.46[*]

Note: This table presents the forecast accuracy results based on conditioning on the adjusted EDF values at different horizons. The cells in bold indicate the outcome expected based on the signal observed. Note that the number of observations changes with the horizon due to data restrictions. The % columns represent the percentage of forecasts (i.e., number of observations in a row that were correct). The "total percent correct" column reports the accuracy of all the forecasts. The Pearson goodness-of-fit statistic tests the null hypothesis that the distribution of BOPEC outcomes conditional on the adjusted EDF values is not different from the unconditional distribution and that the SCAR variable is not informative. The statistic is distributed χ_8^2. An asterisk denotes significance at the 5% level.

Table 9A.   Summary of BOM supervisory variables for all BHCs

Variable	Mean	Std. dev.	Median	25th percentile	75th percentile
log(total assets)	13.58	1.65	12.99	12.34	14.38
lead bank CAMELS	1.89	0.77	2.00	1.00	2.00
"problem" loans	1.85%	1.75%	1.41%	0.84%	2.29%
ALLL	0.37%	0.63%	0.19%	0.08%	0.40%
Sec. 20 indicator	0.04	0.18	0	0	0
trading assets	0.90%	36.80%	0.00%	0.00%	0.00%
double leverage	59.13%	96.96%	68.07%	8.97%	98.93%
ROAA	0.89%	1.25%	0.99%	0.69	1.24%
equity capital	8.35%	2.72%	8.04%	6.77%	9.57%
lagged BOPEC	1.94	0.82	2.00	1.00	2.00

Table 9B.
Summary of BOM supervisory and equity market variables for all publicly traded BHCs

Variable	Mean	Std. dev.	Median	25th percentile	75th percentile
log(total assets)	14.74	1.73	14.44	13.36	15.92
lead bank CAMELS	1.84	0.73	2	1	2
"problem" loans	1.81%	1.65%	1.40%	0.87%	2.19%
ALLL	0.38%	0.59%	0.20%	0.09%	0.43%
Sec. 20 indicator	0.08	0.03	0	0	0
trading assets	0.72%	3.60%	0.00%	0.00%	0.02%
double leverage	42.17%	43.62%	2.51%	17.01%	92.99%
ROAA	0.90%	0.90%	1.02%	0.75%	1.25%
equity capital	8.26%	2.48%	8.02%	6.30%	9.30%
lagged BOPEC	1.87	0.80	2	1	2
3-month SCSR	0.80	0.76	0.80	0.29	1.25
3-month SCAR	0.02	1.50	0.54	-0.72	0.73
EDF value	0.51%	1.21%	0.20%	0.09%	0.49%
asset vol.	0.05%	0.07%	0.04%	0.02%	0.06%

Table 10A. Estimation results for three versions of the BOM model using BOPEC levels

	simple BOM model		core BOM model		extended BOM model	
	Est. coeff.	p-value (%)	Est. coeff.	p-value (%)	Est. coeff.	p-value (%)
ln(assets)	-0.1724	0.0	-0.2689	0.0	-0.1962	0.3
public	.	.	0.1238	0.0	-0.110	89.5
CAMELS	1.2308	0.0	1.5174	0.0	1.5710	0.0
public	.	.	-0.6643	0.0	-0.7492	0.0
prob. loans	47.0437	0.0	49.9639	0.0	50.7289	0.0
public	.	.	-6.0334	46.0	-5.6624	48.4
ALLL	80.7028	0.0	87.0530	0.0	88.8192	0.0
public	.	.	-22.2707	45.8	-26.0223	36.9
Sec. 20	0.5209	3.5	2.4125	0.0	2.0732	0.0
public	.	.	-2.1484	0.0	-1.6463	0.8
trad. assets	-0.0094	41.6	-0.0077	24.8	-0.0048	47.9
public	.	.	-1.0442	66.2	-0.4188	86.6
dble. lev.	0.0387	65.1	0.0798	16.4	0.0877	6.4
public	.	.	-0.1558	36.6	-0.2542	14.5
ROAA	-0.5698	4.1	-0.3053	36.1	-0.2926	37.2
public	.	.	-0.7364	45.1	-0.5703	12.6
eq. capital	-25.9693	0.0	-25.3667	0.0	-24.9659	0.0
public	.	.	-2.0280	74.2	-6.2863	34.1
BOPEC lag	1.1627	0.0	0.9885	0.0	1.0338	0.0
public	.	.	0.4053	0.0	0.3643	2.8
3m. SCSR	.	.	.	.	-0.7006	0.0
3m. SCAR	.	.	.	.	-0.5298	0.0
EDF value	.	.	.	.	0.0180	85.3
asset vol.	.	.	.	.	-1.5560	15.4
# obs.	.	3,975	.	3,975	.	3,975
lkl. value	.	-2,535.887	.	-2,500.447	.	-2,450.318
χ^2 value	.	.	.	70.880	.	100.258

47

Table 10B.   Estimation results for three versions of the BOM model using BOPEC changes

	simple BOM model		core BOM model		extended BOM model	
	Est. coeff.	p-value (%)	Est. coeff.	p-value (%)	Est. coeff.	p-value (%)
ln(assets)	0.1630	0.0	0.2483	0.0	0.1893	0.4
public	.	.	-0.1188	2.4	-0.0148	86.0
CAMELS	-1.3148	0.0	-1.6278	0.0	-1.6830	0.0
public	.	.	0.7186	0.0	0.8140	0.0
prob. loans	-52.3892	0.0	-54.3290	0.0	-55.5081	0.0
public	.	.	4.9124	55.1	4.9682	55.1
ALLL	-99.5887	0.0	-101.2459	0.0	-103.3158	0.0
public	.	.	19.2034	52.4	24.6428	40.6
Sec. 20	-0.5110	3.3	-2.2252	0.0	-1.9629	0.0
public	.	.	1.9956	0.1	1.5835	0.8
trad. assets	0.0011	93.5	0.0033	59.4	0.0011	85.7
public	.	.	0.2619	90.9	-0.2901	90.0
dble. lev.	-0.0550	48.3	-0.0922	8.2	-0.0998	2.9
public	.	.	0.1621	32.9	0.2597	12.5
ROAA	0.4346	6.6	0.1919	36.5	0.1867	37.4
public	.	.	0.8206	0.3	0.6308	2.1
eq. capital	21.2641	0.0	21.6310	0.0	21.4845	0.0
public	.	.	-1.9218	77.0	1.3177	86.6
BOPEC lag	3.6807	0.0	3.8974	0.0	3.9655	0.0
public	.	.	-0.3410	2.1	-0.3253	4.4
3m. SCSR	.	.	.	.	0.7143	0.0
3m. SCAR	.	.	.	.	0.5426	0.0
EDF value	.	.	.	.	-0.0440	62.0
asset vol.	.	.	.	.	0.8894	48.8
# obs.	.	3,975	.	3,975	.	3,975
lkl. value	.	-2,566.047	.	-2,525.786	.	-2,475.349
χ^2 value	.	.	.	80.522	.	100.874

Table 11. Summary of subsample parameter estimation results:
% of the subsample periods for which the coefficients are significant at the 5% level

	core BOM model	extended BOM model
ln(assets)	52.5%	30.0%
public	0.0%	12.5%
CAMELS	100.0%	100.0%
public	32.5%	45.0%
prob. loans	97.5%	97.5%
public	2.5%	5.0%
ALLL	45.0%	45.0%
public	7.5%	7.5%
Sec. 20	32.5%	30.0%
public	12.5%	15.0%
trad. assets	250%	35.0%
public	12.5%	27.5%
dble. lev.	12.5%	12.0%
public	5.0%	2.5%
ROAA	55.0%	55.0%
public	17.5%	15.0%
eq. capital	67.5%	65.0%
public	0.0%	5.0%
BOPEC lag	72.5%	80.0%
public	17.5%	10.0%
3m. SCSR	---	25.0%
3m. SCAR	---	42.5%
EDF value	---	15.0%
asset vol.	---	10.0%
any eq. var.	---	65.0%

Note: This table presents the percentage of the 40 subsample periods for which the estimated coefficients are significant at the 5% level.
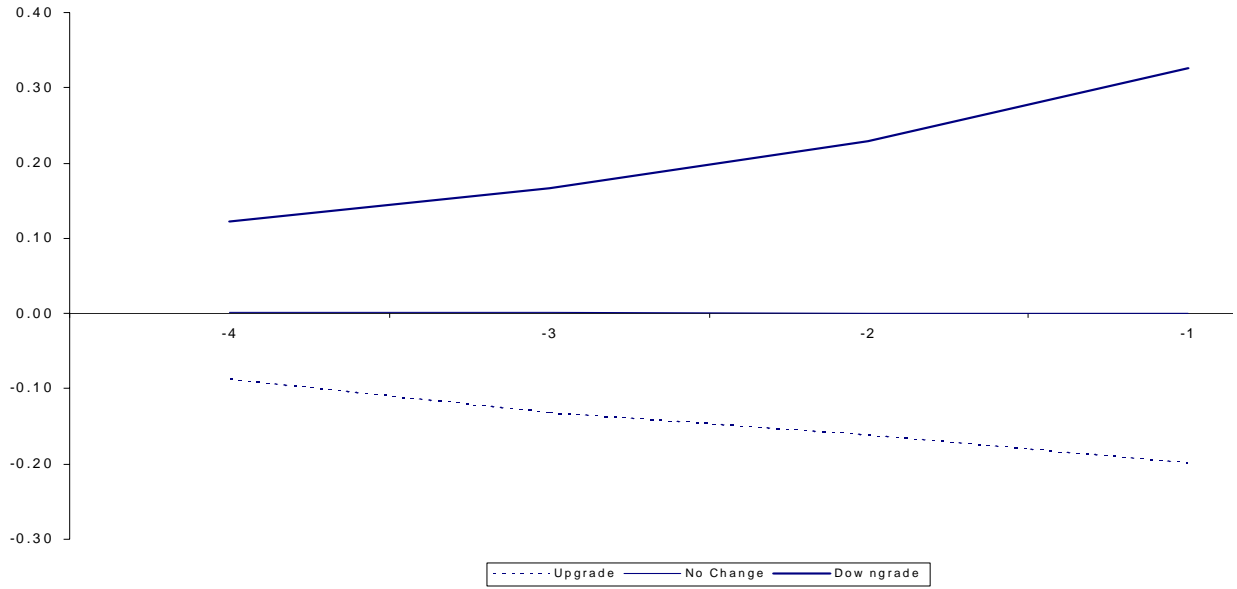
Figure 4A.   Median forecasts from the core BOM model
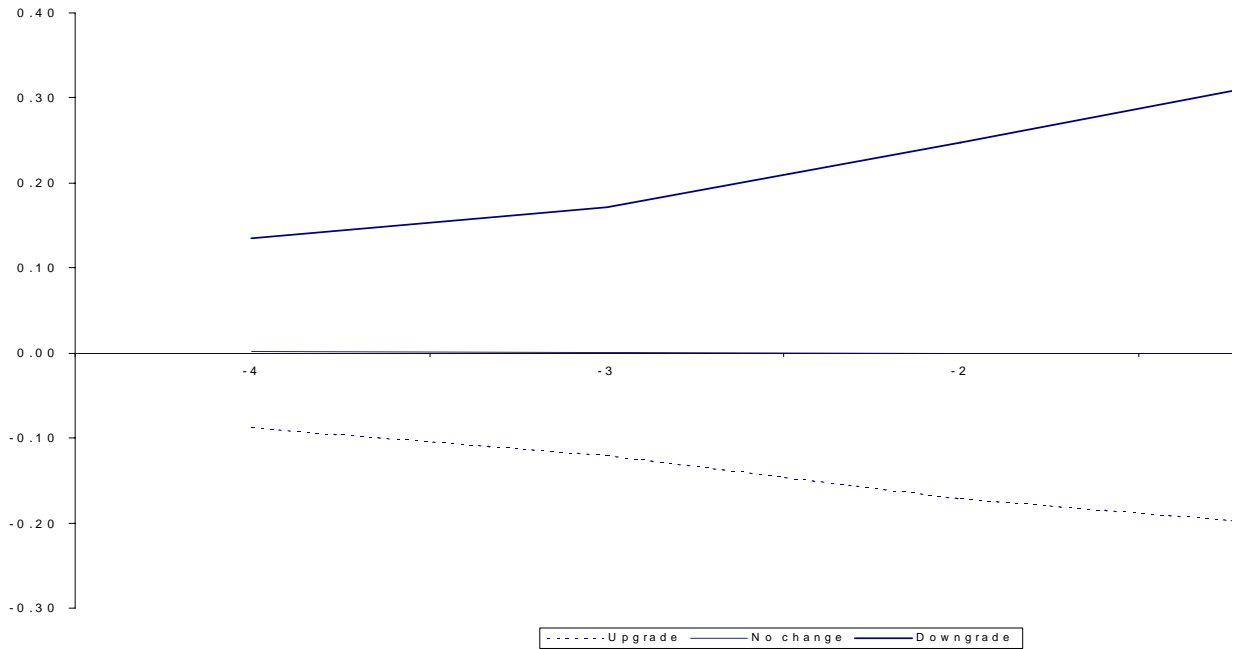

Figure 4B.   Median forecasts from the extended BOM model

Table 12A. BHC event study based on BOPEC forecasts from the core BOM model

Quarters prior	BOPEC upgrades			BOPEC no changes			BOPEC downgrades		
	Median	Sign test p-value (%)	# obs.	Median	Sign test p-value (%)	# obs.	Median	Sign test p-value (%)	# obs.
-4	-0.0866	0.0	746	0.0010	15.5	2,265	0.1222	0.0	486
-3	-0.1307	0.0	783	0.0014	16.8	2,345	0.1672	0.0	537
-2	-0.1613	0.0	805	0.0006	17.3	2,396	0.2300	0.0	564
-1	-0.1978	0.0	839	0.0000	60.8	2,434	0.3264	0.0	599

Quarters prior	BOPEC upgrades above the threshold			BOPEC downgrades below the threshold		
	Median	Sign test p-value (%)	# obs.	Median	Sign test p-value (%)	# obs.
-4	-0.0344	0.0	270	0.1492	0.0	191
-3	-0.0873	0.0	285	0.1973	0.0	209
-2	-0.1360	0.0	292	0.3154	0.0	214
-1	-0.1514	0.0	306	0.3987	0.0	225

Note: This table presents the medians of the adjusted BOPEC forecasts from the core BOM model arranged in event time from month -12 to month +3. The sign test p-values are the p-values for the null hypothesis that the medians are equal to zero.

Table 12B. BHC event study based on BOPEC forecasts from the extended BOM model

Quarters prior	BOPEC upgrades			BOPEC no changes			BOPEC downgrades		
	Median	Sign test p-value (%)	# obs.	Median	Sign test p-value (%)	# obs.	Median	Sign test p-value (%)	# obs.
-4	-0.0881	0.0	746	0.0021	21.8	2,265	0.1351	0.0	486
-3	-0.1198	0.0	783	0.0009	16.2	2,345	0.1765	0.0	537
-2	-0.1713	0.0	805	0.0001	19.3	2,396	0.2722	0.0	564
-1	-0.2044	0.0	839	0.0000	31.5	2,434	0.3736	0.0	599

Quarters prior	BOPEC upgrades above the threshold			BOPEC downgrades below the threshold		
	Median	Sign test p-value (%)	# obs.	Median	Sign test p-value (%)	# obs.
-4	-0.0600	0.0	270	0.1516	0.0	191
-3	-0.0750	0.0	285	0.1623	0.0	209
-2	-0.1225	0.0	292	0.2818	0.0	214
-1	-0.1452	0.0	306	0.3804	0.0	225

Note: This table presents the medians of the adjusted BOPEC forecasts from the extended BOM model arranged in event time from month -12 to month +3. The sign test p-values are the p-values for the null hypothesis that the medians are equal to zero.

Table 13A.
Distribution of BOPEC outcomes after conditioning on the adjusted BOPEC forecasts from the core BOM model

Adj. forecast -4 quarters	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	Pearson stat.
less than -1	136	48.5	46.3	5.1		
abs. value ≤1	3,232	20.8	66.3	12.9		
greater than 1	607	19.4	46.0	34.6	60.9	407.7*
Adj. forecast -3 quarters	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	Pearson statistic
less than -1	115	56.5	36.5	7.0		
abs. value ≤1	3,430	20.8	66.0	13.2		
greater than 1	430	18.4	41.4	40.2	63.0	488.3*
Adj. forecast -2 quarters	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	Pearson statistic
less than -1	92	77.2	17.4	5.4		
abs. value ≤1	3,561	20.5	66.1	13.5		
greater than 1	322	17.7	36.0	46.3	64.7	674.1*
Adj. forecast -1 quarters	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	Pearson statistic
less than -1	123	88.6	8.1	3.3		
abs. value ≤1	3,614	20.2	66.4	13.4		
greater than 1	238	7.6	31.5	60.9	66.8	1109.0*

Note: This table presents the forecast accuracy results based on conditioning on the adjusted BOPEC forecasts from the core BOM model at different horizons. The cells in bold indicate the outcome expected based on the signal observed. Note that the number of observations changes with the horizon due to data restrictions. The percentage columns represent the percentage of forecasts (i.e., number of observations in a row that were correct). The "total % correct" column reports the accuracy of all the forecasts. The Pearson goodness-of-fit statistic tests the null hypothesis that the distribution of BOPEC outcomes conditional on the adjusted BOPEC forecasts is not different from the unconditional distribution and, hence, that the SCAR variable is not informative. The statistic is distributed χ^2_8. An asterisk denotes significance at the 5% level.

Table 13B.
Distribution of BOPEC outcomes after conditioning on the adjusted BOPEC forecasts from the extended BOM model

Adj. forecast -4 quarters	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	P.stat.1	P.stat.2
less than -1	147	51.0	44.9	4.1			
abs. value ≤ 1	3,225	20.6	66.5	12.9			
greater than 1	603	19.7	45.4	34.8	61.1	442.5*	1.1
Adj. forecast -3 quarters	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	P.stat.1	P.stat.2
less than -1	121	55.4	38.0	6.6			
abs. value ≤ 1	3,419	20.8	65.9	13.3			
greater than 1	435	18.2	42.8	39.1	62.6	455.7*	0.9
Adj. forecast -2 quarters	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	P.stat.1	P.stat.2
less than -1	103	74.8	19.4	5.8			
abs. value ≤ 1	3,553	20.4	66.1	13.5			
greater than 1	319	17.6	36.4	46.1	64.7	671.8*	0.7
Adj. forecast -1 quarters	# obs.	% Upgrade	% No change	% Downgrade	Total % correct	P.stat.1	P.stat.2
less than -1	125	87.2	10.4	2.4			
abs. value ≤ 1	3,614	20.2	66.4	13.4			
greater than 1	236	7.6	30.5	61.9	66.8	1110.6*	1.6

Note: This table presents the forecast accuracy results based on conditioning on the adjusted BOPEC forecasts from the extended BOM model at different horizons. The cells in bold indicate the outcome expected based on the signal observed. Note that the number of observations changes with the horizon due to data restrictions. The % columns represent the percentage of forecasts (i.e., number of observations in a row that were correct). The "total % correct" column reports the accuracy of all the forecasts. The Pearson goodness-of-fit statistic denoted as P.stat.1 tests the null hypothesis that the distribution of BOPEC outcomes conditional on the adjusted BOPEC forecasts from the extended BOM model is not different from the unconditional distribution and that the forecasts are not informative. The Pearson goodness-of-fit statistic denoted as P.stat.2 tests the null hypothesis that the distribution of BOPEC outcomes conditional on the adjusted BOPEC forecasts from the extended BOM model is not different from the distribution conditional on the adjusted forecasts from the core model and that the forecasts from the extended model are not relatively informative. The statistics are distributed χ_8^2. An asterisk denotes significance at the 5% level.