Lie to Me: Using Facial Expressions to Detect ESG Washing *

Yiwei Dou
Department of Accounting
New York University
yd18@stern.nyu.edu

Binglei Duan
School of Accountancy
Central University of Finance and Economics
duanbinglei@cufe.edu.cn

Danqi Hu
Guanghua School of Management
Peking University
danqi.hu@gsm.pku.edu.cn

Changtuo (Alto) Luo
Guanghua School of Management
Peking University
luocht@stu.pku.edu.cn

December 2024

* We thank Beth Blankespoor and seminar participants at Peking University and Tsinghua University for very helpful comments. Any errors are our sole responsibility.

Lie to Me: Using Facial Expressions to Detect ESG Washing

ABSTRACT

We examine to what extent CEOs' facial expression reveals their commitment sincerity and thus facilitates the detection of ESG washing. Analyzing videos of bank CEOs' ESG commitment speech made available by the United Nations Principles for Responsible Banking program, we construct deception scores for 36 banks across 22 countries, representing a significant portion of total global bank assets. We find borrowers of banks that have higher deception scores in their commitment videos perform worse on various ESG outcomes including negative ESG incidents, ESG ratings, and emission intensity. The results are mainly driven by deception cues in the visual dimension, especially in the eye area, rather than text and audio dimensions. We also find the deception score to be more powerful when video disclosure is longer, and the facial recognition has lower blurriness. The results are robust to controlling for video persuasiveness scores and available bank ESG ratings. Overall, our evidence indicates the usefulness of video-based deception score in the detection of ESG washing.

Keywords: ESG, commitment, deception, video disclosure

I. INTRODUCTION

There are growing concerns that companies portray their environmental, social, and governance (ESG) activities and commitment in their ESG disclosure opportunistically. This practice, often referred to as "ESG washing," is especially worrying because it misleads ESG-focused shareholders and stakeholders who depend on company disclosures to evaluate companies' ESG practices and direct their capital. Existing research documents disconnect between disclosures with actual performance in the ESG space to show the prevalence of ESG washing (Basu et al., 2022; Reitmaier et al. 2024). However, this research cannot help market participants separate truly ESG committed companies and potential washers ex ante. A few studies measure ESG washing focusing on specific dimensions, on which it is easier to come up with benchmarks such as diversity (Baker et al. 2024). How to detect misrepresentation in broad ESG activities remains an open question.

In this paper, we take a different approach and explore the usefulness of video-based deception scores in detecting ESG washing. Our study draws on previous studies in psychology, video-analytics, and deception detection research, which have shown that there exist significant differences in visual features between truth tellers and liars. Building on this, our assumption is that truly committers and potential washers have some ideas of whether their ESG statements have been exaggerated (or manipulated), and that their (micro) expressions contain cues that can be used to separate them. Therefore, the advantage of our approach is that it is ex ante and does not require any historical or future ESG performance data to calculate. In addition, the methodology we propose is not restricted to a single ESG dimension and can be applied to any/all ESG dimensions of ESG activities that are of interest to investors or researchers.

To implement our approach, we need a representative sample of ESG disclosure in video, produced in a consistent manner. The United Nations (UN) Principles for Responsible Banking (PRB) program, one of the leading sustainable banking frameworks, represents more than half of the global banking industry. Most importantly, the PRB program encourages signatory banks' CEOs to produce a video disclosure upon signing up for the program. The PRB program provides detailed guidance on the content and shooting settings for the video disclosure and posts signatory banks videos in their YouTube channel to increase awareness of the program. Our sample period is from 2016 to 2022, including three years before and after the first wave of PRB commitments in 2019. The final sample includes 9,260 bank-borrower-year observations, corresponding to 2,563 bank-borrower lending relationships across 26 banks in 22 countries.[1]

With the ESG disclosure video at hand, we construct a video-based deception score in a few steps. First, we train a deception detection algorithm using the real-life deception detection video dataset (the Real-life Trial dataset) that contains 121 court trial videos of which 61 are deceptive and 60 are truthful. As the court trials involve multiple rounds of evidence collecting and truthful and deception labels are rigorously established, the real-life dataset, developed by Pérez-Rosas et al. (2015), first made lie detecting in real-life settings feasible and is widely used in automated deception detection literature. The training procedures include extracting numerous features from visual, audio and textual dimensions of the video and use the random forest approach to predict the given true/lie labels. The algorithm we trained using the court

[1] The video disclosure is not mandatory according the PRB requirements. Thus, one limitation of our study is that our analyses are restricted to observations with the available video disclosure to calculate the video-based deception score. Nonetheless, the total assets of PRB banks with available videos represent around 79% (81%) of the total assets of all PRB (listed) banks, indicating that our sample is a reasonable representation of the full PRB sample.

trial data achieves an AUC of 0.8424, in line with the state-of-the-art of automated deception detection research. Visual features contribute the most to the deception detection machine learning model, accounting for 83.89% of the total importance. This finding is consistent with existing psychological research, which identifies visual features, such as facial expressions, as the most significant behavioral cues in detecting deception (Ekman and Friesen 1969; Ekman 2009).

To validate the machine-empowered deception score, we compare the ESG performance of borrowers of banks that are of different degrees of deceptiveness in their commitment video disclosure. Following prior literature, we use borrowers' ESG performance as a proxy for banks' ESG alignment, because numerous studies show that banks integrate ESG in their lending activities by enhancing screening and monitoring of their portfolio companies (e.g., Wang 2023; Kim et al. 2023; Choy et al. 2024). To measure borrowers' ESG performance, we focus on negative ESG incidents, as they are realized outcome and thus more objective and less ambiguous than ESG ratings which are heavily affected by rating methodologies and exhibit significant disagreement across rating providers. Nonetheless, to supplement our main analyses, we also include ESG ratings and carbon emissions (CO2 Emissions) as additional relatively more controllable dimensions of the borrower ESG performance.

We find that borrowers of banks that have higher deception scores in their commitment videos exhibit a greater increase in the number of negative ESG incidents after the video disclosure. In terms of economic significance, a one-standard-deviation increase in the PRB bank's video-based deception scores is associated with a 5% (=0.504×0.099) greater increase in the number of negative incidents in the borrower firms from the pre-commitment-video period to the post-period. In addition to a continuous deception score, we also examine the deception score quartile groups and find that our results are mainly driven by the differences between the top two deception quartile groups and the bottom quartile groups. The dynamic effect analyses confirm that our results cannot be explained by differential pre-existing trends before the PRB commitment video disclosure. Furthermore, the documented patterns remain after we add bank-borrower pair fixed effects, suggesting that the results cannot be solely explained by changes in bank-borrower pairs, for example the initiation and termination of lending relationships. In other words, banks with higher deceptive video score are less likely to "walk the talk" and exert less efforts in disciplining their borrowers' ESG performance, consistent with them more likely being ESG washers.

We implement several cross-sectional analyses to further strengthen our inferences. First, longer videos may contain more useful lie-detecting cues compared to shorter videos. Consistently, we find our results are more pronounced in the subsample of longer videos. Second, we cut our sample based on the quality of facial area recognition and find our results are mainly driven by the subsample of videos with high facial area recognition quality. Third, we compare banks with high versus low past ESG failures and find our results mainly exist in the subsample of banks with low past ESG failures. This is consistent with banks with high past ESG failures either being more accustomed to lying about ESG commitments or more strategically managing their nonverbal cues in the video, leaving fewer deception cues and thus making our deception scores less useful.

To unpack the driving force of our video-based deception score, we decompose the comprehensive video-based score into the scores along the textual-, audio-, and visual dimensions. We find that our results are driven by the deception scores in the visual dimension rather than the text and audio dimensions, consistent with visual cues being more powerful in revealing deception than text and audio features. In particular, one would expect the spokesman of the video would be well prepared in terms of the textual transcripts, especially given our videos are very short with no Q&A opportunities at all. To further investigate the underlying mechanisms of visual features in deception detection, we trained six models, each using only one set of visual features as inputs. The results indicate that models trained with eye features demonstrate strong predictive power for borrowers' ESG performance, aligning with existing psychological research that highlights the role of eye movement as a key cue in detecting deception.

We perform several robustness checks of our results. First, we use two alternative ESG performance measures that are relatively more input-related: borrowers' ESG ratings and carbon emission intensities. Our results are robust to using these alternative measures. Second, we examine if our video-based deception scores capture distinct information beyond what is revealed by the most recent ESG ratings before video disclosure if available. Third, we consider the visual persuasiveness measure proposed by Hu and Ma (2024). While our measure is motivated to capture the truthfulness of the ESG statements in the video disclosure, the persuasiveness score is more about the emotions of the statements and the consequent impressions of the viewer. Our results remain after we control for the most recent ESG ratings and the video persuasiveness scores. Finally, our results are robust to using alterative machine learning models to build the deception-detection algorithms, and different ways of clustering standard errors.[2]

Our paper has the following contributions. First, it contributes to the literature on ESG washing. Previous research documents the prevalence of ESG washing practices, as the costs of providing exaggerated/misleading ESG disclosure are unclear. How to detect ESG washing ex ante, especially with no historical or reliable relevant ESG performance data available, is challenging. In this paper, we propose and validate a novel ex ante detection approach based on a machine-learning empowered deception score extracted from video disclosure.

Second, it contributes to the literature on video disclosure. Prior studies on video disclosure have mostly studied the informativeness of video disclosure in capital market settings such as entrepreneurs' pitch videos (Davila and Guasch 2022; Hu and Ma 2024), CEO interviews with respect to earnings announcements (Banker et al. 2023), and IPO roadshow videos (Blankespoor et al. 2017; Duan et al. 2024). We are the first to study the usefulness of ESG video disclosure in generating an ex-ante measure that helps separate truly committers and potential washers before subsequent realized outcomes become observable. In addition, most prior studies on video disclosure mainly focus on visual persuasiveness or investor perceptions of the visual cues presented, few studies explore the power of disentangling truthfulness based on video disclosure. We contribute to this endeavor by showing the usefulness of deception scores in an ESG video disclosure setting.

In the field of financial disclosure, Duan et al. (2024) apply the trained algorithm to IPO

[2] We also explore whether our results are different for different types of negative ESG incidents. First, the PRB program calls for the signatory banks to align their activities with UN Sustainable Development Goals, which cover all three ESG dimensions (i.e., "Environmental", "Social", and "Government" dimensions). Indeed, we find our results exist and are similar for negative incidents along all three dimensions.

roadshow video and find the video-based deception score predicts the IPO fraud. One major difference between our setting and theirs is that in the field of ESG, the costs to provide misleading disclosure seem small compared with IPO fraud, as the shareholder demanded audits and successful litigations against washers are rare during our sample period. As a result, lying about ESG commitments yields lower psychological burden and thus fewer deception cues (Ekman 2009). Ex ante it is unclear whether the machine-learning based deception score would be effective at detecting ESG washing and we provide initial empirical evidence on this front.

Third, the paper has policy implications. Our results show that bank CEOs' videos at the time of the disclosure, combined with AI deception detection technology, may reveal how truthful/deception they are with their ESG statements made in the video, and thus can be used to predict their subsequent ESG efforts in the post-disclosure period. This effective ex ante measure of ESG washing is particularly important, as it may help ESG-focused investors make better decisions of the ESG capital under their management. While our research does not analyze the full costs and benefits of requiring more ESG disclosure in the video format, the evidence suggests one potential benefit of the mandatory video disclosure. That is the better separation of ESG washers ex ante and better ESG capital allocation ex post.

II. THE SETTING AND ESG VIDEO DISCLOSURE DATA

The Principles for Responsible Banking (PRB) aims to encourage banks to align their strategy and business practices with the vision in the UN Sustainable Development Goals (SDGs) and the Paris Climate Agreement. Figure 1 shows the growth in the number of PRB signatory banks across various regions. By 2023, the PRB has over 330 signatory banks, representing more than half of the global banking industry, and has become the world's leading sustainable banking framework.[3]

The PRB requires their signatories to make public announcements at the time of signing the commitments. However, such ex ante textual disclosures tend to be brief and boilerplate, providing little information beyond the publicly announced commitment status. Therefore, the PRB encourages signatory bank CEOs to produce videos explaining why they sign the Principles and what the PRB means to their business. To help guide the content of the video disclosure, the PRB provides a list of exemplary questions that signatory CEOs can refer to. It includes topics covering motivations for joining the PRB program, views on the PRB principles, implications of the PRB for their business, etc. In addition, the PRB also provides guidelines on various aspects of video production, including the locations (i.e., where the videos should be filmed), the background of the videos, the CEO's posture, the set-up of the camera, microphone, and lighting, etc. These guidelines ensure that the videos are produced under a consistent standard and that the CEO's upper body is clearly visible and speech is audible. The list of exemplary questions and guidelines for video production can be found in online Appendix B.

After the video is produced, the PRB posts these videos on its YouTube channel, where they can be viewed by anyone interested in the initiative.[4] The videos serve as a key communication asset to raise awareness, increase engagement, and inspire other banks to join the PRB. Signatory banks can also take their own videos and post them at other venues for internal and external communications about the PRB and their sustainability efforts.

We manually download videos of signatory CEOs from UNEP FI's YouTube account. Our initial download consists of 77 videos, covering 23% of the PRB banks. The mean (median)

[3] UNEP FI. "Principles for Responsible Banking" Accessed January 2024. https://www.unepfi.org/banking/banking-principles/.

[4] CEOs on the Principles: https://www.unepfi.org/videos-and-testimonials/.

video length is 101 (95) seconds, which is similar to the pitch videos in Hu and Ma (2024). The mean (median) number of words in the video transcripts is 204 (183). These videos often have low viewership and engagement, with few likes or dislikes.[5] The videos are timely accessible when signatory banks sign the PRB commitment. About 89.6% of the videos were posted on the YouTube channel within the month of signing or earlier, 97.4% of the videos were posted with no more than a 4-month delay, and only two videos had a maximum delay of 7 months.

III. MEASUREMENT AND SAMPLE

Video-Based Deception Score

We apply advanced machine learning techniques designed to handle high-dimensional data to construct a robust visual deception score for identifying deceptive behavior in PRB commitment videos. This methodology, which utilizes machine learning for deception detection in real-world video contexts, represents a growing area within the field of automated deception detection. Empirical evidence suggests that it surpasses human detection capabilities, achieving notably higher accuracy (Gogate et al. 2017; Wu et al. 2018; Ding et al. 2019). Following the methodology outlined by Duan et al. (2024), our video processing approach involves four steps, which are detailed in the subsequent section.

Step 1: Videos Pre-processing

To ensure the accuracy of our deception detection method and eliminate potential interference from non-CEO faces, we decompose the bank videos into individual clips and retain only those featuring the CEOs. This approach excludes any brief segments that might include other individuals' faces. Following Hu and Ma (2024), we sample images at ten frames per second to identify and compare human faces. Specifically, we feed our raw images into the cloud computing system and receive a host of face-related measures. The face-detection API

[5] The viewership and engagement characteristics of our PRB videos are in line with Hu and Ma (2024), whose pitch videos also have low views and few likes or dislikes, indicating that these videos are not marketing tools, and alleviating concerns that the people in the videos are intentionally acting out.

identifies faces in all images, while the face-comparison API compares two faces and determines whether they belong to the same individual, with an error rate of 0.001%.[6] Next, we compile and edit firm videos into several individual videos using a video editing package. This procedure excludes an average (median) duration of 9.07 seconds (5.96 seconds), which represents an average (median) of 10.54% (7.07%) of the duration of a bank video.

Step 2: Feature Extraction

Next, we construct a set of features extracted from textual, audio and visual dimensions. In visual dimension, we use established open-source Python packages for feature extraction. This approach ensures both replicability and transparency in our methodology. In line with previous studies (Morales et al. 2017), we extract 709 facial features per frame using OpenFace, a computer vision and machine learning toolkit for facial behavior analysis developed by the CMU MultiComp Lab (Baltrušaitis et al. 2018).[7] OpenFace categorizes the extracted visual features into four groups: gaze-related information, head and face location details, face shape characteristics, and facial Action Units (AUs). Each category contains between 35 and 348 features. The specific names and descriptions of the features in each category are provided in Table OA1 in Online Appendix A, while Figure OA1 in Online Appendix A offers a visual illustration of the OpenFace output. In the audio dimension, we extract audio files (.wav) from videos and use Librosa to compute 30 audio features, encompassing both prosodic and spectral characteristics, such as chroma STFT, spectral centroid, spectral bandwidth, spectral rolloff, RMS, and zero crossing rate. In the textual dimension, we first transcribe the audio into text and then extract syntactic features using the StanfordNLP package. This process yields 22 syntactic features, including tree depth, the number of root dependents, the number of unique

---

[6] The face-detection and face-comparison machine-learning algorithms are provided by Face++, a leading AI firm in China.

[7] OpenFace is an open-source facial behavior analysis toolkit that delivers state-of-the-art performance in facial landmark detection, head pose estimation, facial action unit recognition, and eye gaze estimation. For more detailed information, see https://github.com/TadasBaltrusaitis/OpenFace.

universal POS tags, the frequency of each part-of-speech (POS) tag, total word count, average word length, and a computed dependency distance measure.

To construct a feature matrix suitable for machine learning processing, we generate a feature vector representing three dimensions of each video using OpenMM, an open-source multimodal feature extraction tool.[8] Specifically, we apply 11 statistical functions to each of the 761 (709+22+30) features across all frames, including maximum, minimum, mean, median, standard deviation, variance, kurtosis, skewness, the 25th percentile, the 50th percentile, and the 75th percentile. These 11 statistical functionals allow us to condense the frame-level features (761 per frame * number of frames) into video-level features (761 * 11). In total, we obtain 8,371 video-level features.

Step 3: Model Construction and Evaluation

We train a machine learning model to distinguish between liars and truth-tellers, relying on a dataset with established labels indicating deceptive or truthful behavior for each video. For this purpose, we use the Real-life Trial dataset developed by Pérez-Rosas et al. (2015), a widely recognized benchmark in the automated deception detection literature. The dataset consists of 121 labeled videos depicting real instances of deception and truth during court trials. It includes recordings from 21 unique female and 35 unique male speakers, with ages ranging approximately from 16 to 60 years. Among the 121 videos, 61 are labeled as deceptive, while the remaining 60 are labeled as truthful.[9] Trial outcomes, such as guilty verdicts, non-guilty verdicts, and exonerations, combined with verified facts, are used to accurately label video clips as deceptive or truthful. In some instances, deceptive videos are derived from suspects denying involvement in a crime, while truthful clips feature the same suspects answering

[8] Detailed information can be found here: https://github.com/michellemorales/OpenMM/tree/openmm_v2.
[9] The videos in the RLT dataset are sourced from public multimedia platforms featuring trial hearing recordings, where truthful or deceptive behavior can be reliably observed and verified. The video selection process follows a rigorous protocol with strict guidelines. These criteria ensure that the defendant or witness in each video is clearly identifiable, their face remains visible for the majority of the clip duration, and the visual quality is sufficient to discern facial expressions accurately.

questions about verified facts. Witness testimonies corroborated by police investigations are labeled as truthful, whereas testimonies supporting guilty suspects are labeled as deceptive. We use the RLT dataset with its established labels to train our lying detection algorithm for three primary reasons. (1) The trial scenario, characterized by multiple rounds of evidence collection and investigative efforts by police and prosecutors, offers one of the most reliable real-world contexts for determining whether an individual is lying. (2) The creators of the RLT dataset invested substantial effort in selecting videos and assigning labels. The dataset's widespread citation in machine learning research on deception detection further attests to its credibility and the robustness of its labeling. (3) Prior psychological research on deception detection and its neuroanatomical basis suggests that certain facial features indicative of lying may be consistent across different settings, making the dataset relevant for training algorithms.

To develop the deception detection model, we first preprocess the videos in the RLT dataset and extract features following the procedures outlined in Step 2. We then train a prediction model to classify deception outcomes based on the extracted features. The model is implemented using the Random Forest classification algorithm in Python, which has demonstrated superior performance in previous studies utilizing similar feature extraction methods (Pérez-Rosas et al. 2015; Morales et al. 2017).[10] The model's performance is evaluated using 10-fold cross-validation, yielding the following metrics: AUC (0.8424), accuracy (0.7583), precision (0.7958), and recall (0.7631). AUC, representing the area under the Receiver Operating Characteristic (ROC) curve, provides a summary of the model's overall diagnostic accuracy, where values range from 0.500 (chance level) to 1.000 (perfect

[10] Random Forest is a type of ensemble learning. In contrast to conventional machine learning techniques such as SVM, which generate a single estimator, ensemble learning models can improve performance by aggregating the predictions of multiple models. We use the Random Forest model because Morales et al. (2017), Wu et al. (2018), and Khan et al. (2021) find that it has higher accuracy compared to the decision tree model. Similar to our study, these papers used automatically extracted features. We later find that our results are robust to using Gradient Boosted Decision Trees (GBDT) model.

prediction).[11] The performance of our model is consistent with results reported in the existing automated deception detection literature.

Step 4: Calculate Video-Based Deception Score

Using the trained predictive model from Step 3 and the input features of the PRB videos extracted in Step 2, we calculate visual deception scores for the CEOs featured in the videos. These CEOs' deception scores, *Deception Scores*, has a mean of 0.433 and a standard deviation of 0.150.

Measure of ESG performance

In line with previous studies, we use the ESG performance of borrowers to capture the ESG alignment of banks, as one significant aspect for banks to incorporate ESG factors is by enhancing their ESG involvement in lending activities, which can be achieved by carefully selecting and monitoring their borrowers (e.g., Choy et al. 2024; Wang 2023; Kim et al. 2023). In our main analyses, we first use realized ESG outcome to capture borrowers' realized ESG outcome. Specifically, we use negative ESG incidents from the RepRisk database, which screens over 150,000 public sources in 23 languages - such as print media, online media, social media, and regulatory filings - on a daily basis to identify any company or project associated with an ESG risk incident, covering over 250,000 public and private companies from around the world. These negative incidents represent realized outcomes, making them more objective and less ambiguous than ESG ratings, which are heavily influenced by rating methodologies and show significant disagreement across providers (Bams and van der Kroft 2022; Berg et al. 2022). Besides, the ESG issues identified by RepRisk align with the Ten Principles of the UN Global Compact and the 17 Sustainable Development Goals (SDGs), making them particularly relevant for examining banks' sincerity in their PRB commitments, as the PRB program is UN-

[11] The other performance metrics include accuracy, precision, and recall. Accuracy represents the overall classification accuracy across the sample. Precision is the proportion of instances classified as deception by the model that are actual cases of deception. Recall measures the proportion of actual deception cases that are correctly identified by the model.

led and aims to promote the SDGs.[12] Following Christensen et al. (2022), we focus on incidents classified as "severe" or "highly severe" and exclude those classified as "low severity" to avoid including cases where firms are unintentionally involved.[13] We construct *NegIncidents* as the number of negative ESG incidents for borrower firms in a given year.

To complement our main analyses of negative ESG incidents, we also examine borrowers' ESG performance on other dimensions. We collect data on borrowers combined ESG scores, ESG reporting scores, ESG strategy scores, and carbon emissions from Refinitiv ESG. These metrics are arguably more input-based and more reflective of firms' ESG efforts, and combined together, comprehensively assess firms' overall ESG disclosures, policies, and carbon emission practices. By including these additional analyses using more input/effort-based on ESG measures, we avoid failing to capture PRB banks' efforts in developing ESG policies (Christensen et al., 2022), as improving real ESG outcomes generally takes time.

Sample Construction

We start our sample period in 2016, which is the year after the announcement of the Paris Agreement, to mitigate the confounding effect of the Paris Agreement on our results (e.g., Mueller and Sfrappini 2022). We end our sample in 2022, so that we have three years before and after the first wave of PRB commitments in 2019. Our sample begins with firms that have lending relationships with PRB banks that have videos, sourced from the Thomson Reuters Dealscan database. Thomson Reuters Dealscan provides comprehensive loan contract data, including borrower identity and loan characteristics (Bharath et al. 2011). We focus on lead banks, which play a central role in establishing and maintaining borrower relationships and in information collection and monitoring (Sufi 2007). We also exclude financial industry

[12] See RepRisk's website for more introductions: https://www.RepRisk.com/research-insights/resources/methodology.
[13] The severity of incidents is assessed along three dimensions: (1) the consequences of the incident in terms of health and safety, (2) the extent of the impact, ranging from individual to a large group or population, and (3) whether the incident was caused by an accident, negligence, or intent.

borrowers (SIC=6000-6799).

Our loan-level sample consists of 18,272 loan facilities from 2016 to 2022. For each loan initiation, we assume the bank-borrower relationship persists throughout the loan's lifecycle, following Dou and Xu (2021). This results in 22,884 bank-borrower-year observations between 2016 and 2022. We include only firms that borrow at least one loan before and after the PRB program. We match the borrowers in DealScan with financial data from Worldscope using the link table from Beyhaghi et al. (2021). We also collect banks' accounting information from Bankscope. We exclude observations with missing borrower or bank control variables, or missing data on ESG negative incidents from the RepRisk, as well as those that are either singletons or separated by a fixed effect in the Poison regression. The final sample includes 9,260 bank-borrower-year observations, corresponding to 2,563 bank-borrower lending relationships.[14] The details of our sample selection procedure are reported in Panel A of Table 1.

[Insert Table 1 here]

IV. EMPIRICAL RESULTS

Main Results of Video-Based Deception Scores

Research Design and Descriptive Statistics

In our main test, we investigate whether video-based deception scores can explain the ESG performance of ex post lending relationships. To empirically examine this, we estimate the following regression:

$$NegIncidents_{ijt} = \beta_0 + \beta_1 Post_{it} \times Deception\ Scores_i + Controls + FEs + \varepsilon_{ijt} \tag{1}$$

where i denotes the bank, j the borrower, in a lending relationship (i.e., with an unmatured

[14] Our final sample consists of 36 PRB banks with videos across 22 different countries (i.e., Belgium, Brazil, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Malaysia, Mauritius, Netherlands, Norway, Poland, Portugal, South Africa, Spain, Sweden, Switzerland, United Kingdom, and United States). The borrowers in our sample are from 62 different countries.

loan contract), and *t* denotes the year. *NegIncidents* is the number of negative ESG incidents of borrower *j* who has a lending relationship to bank *i* in year *t*. *Post* is a dummy variable that equals one after bank *i* joins the PRB program (i.e., after 2019). *Deception Scores* represents the deception scores of bank *i'* video, capturing the likelihood of deception during a bank CEO's PRB ESG commitment disclosure video.

We control for several bank characteristics, including the bank size, capital adequacy, loan loss provisions, etc., and borrower characteristics, including the firm size, profitability, leverage, investment, etc. Recent studies suggest that the ESG-related regulations may affect the ESG performance of lending relationships (Wang 2023; Ivanov et al. 2024), so we control for *Country*×*Year* fixed effects and *Industry*×*Year* fixed effects. If PRB banks are genuinely committed to improving the ESG performance of their lending relationships—either by screening out borrowers with poor ESG performance or through monitoring, such as on-site inspections and private engagement—then their connected borrowers should exhibit fewer negative incidents post-PRB.

Panel B of Table 1 presents the summary statistics of the variables used in our main analysis. We winsorize all continuous variables at the 1% and 99% levels. The mean (median) number of negative incidents is 4 (2) in our sample. 59.4% of the observations are post-PRB. The average size of PRB banks is 20.723.

Deception Scores and the Detection of ESG Washing

Table 2 presents the main results comparing the ESG performance of borrowers of PRB banks with different levels of video-based deception scores. We first estimate Equation (1) without including firm-characteristic controls in column 1, and then add firm-characteristic controls, *Country*×*Year* fixed effects, and *Industry*×*Year* fixed effects in columns 2-4, respectively. The coefficient estimates on *Post×Deception Scores* are all positive and significant at the 1% level, suggesting that PRB banks with higher video-based deception scores exhibit poorer ESG outcomes, as evidenced by more negative incidents among their borrowers. In terms of economic magnitude, the estimate in column 4 indicates that a one-standard-deviation increase in the PRB bank's video-based deception scores is associated with a 5% (=0.504×0.099) higher number of negative incidents in the borrower firms during the post-commitment-video period. In column 5, we further include the *Bank×Borrower* fixed effects. We continue to find a positive and significant coefficient estimate on *Post×Deception Scores*, suggesting that the performance difference between PRB banks with different deception scores cannot be explained by screening efforts (i.e., changes of bank-borrower pairs) alone.

[Insert Table 2 here]

Dynamics of PRB Banks' Lending Relationships

One potential concern is that the results reported in Table 2 may simply reflect pre-existing divergent trends in ESG performance of PRB banks with different levels deception scores that have nothing to do with their changes in ESG practices during the post-commitment-video period. To address this, we examine whether the parallel trends assumption holds in the pre-PRB period. Specifically, we modify Equation (1) by replacing *Post* with six time indicators, each corresponding to the number of years relative to the PRB program, and interact them with Deception Scores. The six time indicators are *I(3yrs before PRB)*, *I(2yrs before PRB)*, *I(1yr before PRB)*, *I(1yr after PRB)*, *I(2yrs after PRB)*, and *I(3yrs after PRB)*. We use *I(PRB Launch Year)* as the reference group, which is omitted from the regression. Figure 2 shows the results. Regardless of whether *Bank×Borrower* fixed effects are included, the coefficient estimates on the two-way interactions between the time indicators and *Deception Scores* are all insignificant before the PRB launch year, and only become significant after the PRB launch year, and the significant effects remain relatively stable during the post-commitment period.

[Insert Figure 2 here]

Cross-Sectional Results

In this section, we explore the cross-sectional variations of the video-based deception scores' detecting power to corroborate our findings.

First, we examine the effect of the duration of the video disclosure. The longer the ESG commitment video is, the greater deception power it likely possesses. Because when a CEO speaks for a longer time, the chances that she leaves deception cues and that these deception cues are captured by our algorithm should both be higher. Therefore, we expect the power of the video-based deception score in detecting ESG washing to be higher for longer videos. To empirically examine this, we divide our sample into two subgroups based on the median video length in Panel A of Table 3. The coefficient estimates on *Post×Deception Scores* are more than three times larger in the subsample of longer videos than in the subsample of shorter videos. The differences in coefficients are significant at the 1% level.

[Insert Table 3 here]

Second, we examine the effect of video quality, proxied by the degree to which facial features can be accurately measured. Facial features are where most deception cues reside (Ekman and Friesen 1969; Ekman 2009), therefore, whether the facial features are captured accurately is crucial to the detection power of the video-based deception score. To measure the quality of face recognition in these CEO videos, we rely on the blurriness scores provided by Face++, which reflects the degree of clarity in the facial region of an image. The face blurriness is assessed by calculating the high-frequency information in the facial region of the image. High-frequency information refers to the finer details and textures that exhibit rapid changes in the image. Generally, blurred images contain less high-frequency information, whereas clearer images retain more of it. Higher levels of blur often result in the inability to extract clear facial features, which in turn affects the accuracy of facial recognition. We then divide our sample into two subgroups based on the median of blurriness scores. The results are presented in Panel B of Table 3. The coefficient estimates on *Post×Deception Scores* are positive and statistically significant in both specifications in the subsamples with high face recognition quality. In contrast, the coefficients are insignificant in the subsamples with low face recognition quality. The coefficient differences are significant at the 1% level.

Taken together, the patterns in the cross-sectional analyses suggest that the detecting power of the video-based deception score depends on the length and data quality of the video disclosure.

Finally, building on previous psychological literature, we investigate a situation in which individuals are potentially more strategic with their lies or more customed to lying about ESG. Specifically, we compare banks with high past ESG failures versus low past failures. Banks with high past ESG failures may become more customized to lying about their ESG commitments, thus experiencing lower psychological burden, and as a result, leaving fewer deception cues (Ekman 2009). In addition, according to the psychology literature, people tend to manage their verbal and nonverbal behavior and control their lying cues when they believe they are perceived as less truthful by their audience. If they can to some extent manage their deception cues, their lies may become harder to detect (Burgoon et al. 2021). When banks have a history of ESG governance failures, their ESG commitments become less credible to stakeholders, and the CEOs are more likely to engage in strategic activities to manage their verbal and nonverbal behavior in the commitment video, making our deception scores less useful. We measure a bank's past ESG governance failures as the total number of negative incidents of a bank's borrowers prior to the PRB program. We then divide the sample into two subgroups based on the median of past failures. Panel C of Table 3 reports the results. In line with our conjecture, the coefficient estimates on *Post×Deception Scores* are both positive and statistically significant in columns 2 and 4, where banks with fewer past failures (i.e., higher expectations of trustworthiness) are less likely to strategically control their lying cues during the videos and are not used to lying about ESG. In contrast, the video-based deception scores lose their effectiveness when CEOs are more likely to manage their lying cues and/or more customed to lying.

V. OTHER EXPLORATORY ANALYSES AND ROBUSTNESS CHECKS

Unpacking the Deception Scores

Visual, Audio, and Textual Features in Detecting Lies

In our main analyses, we focus on the deception scores trained using all video features from the visual, audio, and textual dimensions. In this subsection, we investigate which dimension of video features drive the detecting power of ESG washing. To do this, we construct the visual-based deception score (*Deception Scores_V*), the audio-based deception score (*Deception Scores_A*) and the text-based deception score (*Deception Scores_T*) by training the machine learning models using only the visual, audio, and textual features, respectively. Panel A of Table 4 shows the results. We first include the visual-based, audio-based and textual-based deception scores as independent variables separately. The coefficient estimates on the interaction terms are both positive and statistically significant when we use *Deception Scores_V* as the independent variable, while they are close to zero for both *Deception Scores_A* and *Deception Scores_T*. In columns 4 and 8, we include all deception scores in the model to assess whether the detecting power of the visual deception score is incremental to the audio-based and text-based scores. We continue to find that only the coefficient estimates on *Deception Scores_V* are positive and statistically significant, suggesting that compared to textual and audio features, the visual features are the most useful in detecting CEO's ESG washing in ESG commitment disclosure videos.

[Insert Table 4 here]

Different Categories of Visual Features

Since the visual parts of videos are the most informative about CEO deceptions, we next compare the detecting power of different categories of visual features. Following prior literature (Baltrušaitis et al. 2018), we categorize the visual features of videos into six groups: gaze, eyes, facial pose, facial landmarks (LMK), facial shape, and facial action units. Gaze refers to the eye gaze direction vector in world coordinates. The eye category comprises 56 eye landmarks, which capture the positions of the pupil, iris, and sclera. Facial pose describes the location and rotation of the head. Facial landmarks include the positions of 68 key points outlining the face, mouth, nose, eyes, and brows. Face shape is represented by parameters of a point distribution model that delineates both rigid and non-rigid facial shapes. Facial action units are a way to describe human facial expression, such as upper lid raiser, cheek puffer. Using the features extracted from each category, we build classifiers to predict deception. This approach simulates scenarios that evaluate the effectiveness of a deception detection model in predicting ESG performance based on different visual feature sets, thereby identifying which categories of visual features are most critical for accurate predictions. Panel B of Table 4 presents the results. The visual features of eye category are incremental to all other visual categories and drive the detecting power of the visual features. The underlying physiological mechanism for the importance of eye cues lies in the increased blood perfusion in the orbital muscles observed in individuals engaging in deceptive behavior (e.g., Tsiamyrtzis et al. 2007).

The findings also align with the broader deception detection literature, which consistently identifies eye movement cues as key features for detecting deception, including eye blinking, pupil size, and eye contact (Zuckerman et al. 1981; Hartwig et al. 2011).

Robustness Tests

Other Measures of Borrower Firms' ESG Performance

In our main analyses, we capture borrower firms' ESG performance through the occurrence of negative ESG incidents, which are not subject to subjective assessments by ESG raters and reflect real ESG outcomes rather than mere the "cheaptalks" (Li and Wu 2020). To further corroborate the robustness of our findings, we also test other measures of borrower firms' ESG performance. Specifically, we obtain ESG combined scores, ESG reporting scores, ESG strategy scores, and carbon emissions intensity from the Refinitiv ESG database. Compared to negative incidents, these indicators are arguably more controllable by the borrower firms and reflect their different efforts on ESG issues. We then substitute these alternative measures for *NegIncidents* in Equation (1). Table 5 presents the results. Consistent with our baseline findings, PRB banks with higher video-based deception scores continue to exhibit poor ESG performance in their lending relationships, as evidenced by lower ESG combined scores, reporting scores, and strategy scores among their borrowers post-PRB program. Besides, the video-based deception scores also predict PRB banks' efforts to reduce the carbon footprints of their lending relationships. Overall, these results further reinforce our inference based on negative ESG incidents as a measure of ESG performance.

[Insert Table 5 here]

The Most Recent Available ESG Ratings of PRB Banks

We investigate whether our video-based deception scores have incremental detecting power over other available information sources, in particular, the most recent available ESG ratings of PRB banks. We, however, expect the most recent available ESG ratings (prior to joining the program) of these PRB banks to be less useful in detecting ex post ESG washing. Recent studies cast doubt on commercial ESG ratings, as they mainly reflect the disclosure rather than the true actions on ESG issues (Bams and Kroft 2022). In addition, they are subject to the rewrite of data (Berg et al. 2021) and are strategically influenced by the rated firms (Cornaggia and Cornaggia 2023) and data providers (Li et al. 2024). To compare the usefulness of available ESG ratings to our deception scores, we obtain the ESG combined scores from Refinitiv ESG. Only 47% of the PRB banks in our sample have available scores, so we include a dummy variable, *I(Missing_BankESGratings)*, to indicate the missing ESG scores from Refinitiv. Panel A of Table 6 reports the results. After controlling for the most recent available ESG ratings of these banks, our video-based deception scores are still significantly associated with the real ESG outcomes of their lending relationships. In comparison, the detecting power using commercial ESG ratings is close to zero.

Video-Based Measures of Persuasion

Next, we examine whether the detecting power of our deception scores can be explained by other video-based measures examined in previous research. In particular, Hu and Ma (2024) measure entrepreneurs' persuasiveness using a set of start-up pitch videos. They find that entrepreneurs with positive (e.g., passionate and warm) pitches have a higher funding probability but underperform after receiving funding. We do not expect that persuasion score to be particularly useful in detecting ESG washing as they mainly focus on the tone of the emotions conveyed by the video and are not designed to capture deception cues. Nonetheless, to measure persuasion, we follow Hu and Ma (2024) and employ principal component analysis (PCA) to extract the first principal component from four dimensions: visual emotion, audio emotion, text emotion, and facial beauty. Online Appendix C describes how we construct this persuasion measure in detail. We then include the persuasion measure (*Persuasiveness_PCA*) and its interaction terms with *Post* in our model. Panel B of Table 6 presents the results. The video-based deception scores continue to predict the ESG performance of PRB banks' lending relationships, even after controlling for the persuasion measure. Moreover, the persuasion measure contains little information about potential ESG washing, as it captures persuasiveness skills shown during the videos rather than lying cues.

Quartile Ranks of Deception Scores

We use continuous deception scores in our main analyses. In Panel C of Table 6, we also use quartile rank indicators of deception scores to explore the differences from the least deceptive to the most deceptive CEOs. The results suggest that the detecting power of potential ESG washings is primarily driven by the differences between the top two deception quartile and the bottom quartile. In terms of economic scale, borrower firms of PRB banks in the top deception quartile exhibit 25.4%-26.1% more negative ESG incidents in the post-PRB period, while borrower firms of banks in the second-top deception quartile have 18.4%-19.0% more incidents, both compared to the bottom quartile. The results indicate that our results hold not only with continuous deception scores but also with discrete deception groups.[15]

Alternative Specifications

We test the robustness of our findings to various alternative specifications. First, we construct our video-based deception scores using an alternative machine learning model - Gradient Boosted Decision Trees (GBDT). Online Appendix Table OA2 shows the results. The deception scores trained by the GBDT model exhibit very similar detecting power to our baseline deception scores. Second, we explore alternative clustering of standard errors in Online Appendix Table OA3. Our findings remain robust to different clustering schemes. Third, we examine whether our results are driven by specific dimensions of negative ESG incidents. RepRisk categorizes negative incidents into environmental, social, governance, and cross-

[15] The means of video-based deception scores in quartiles 1-4 are 0.249, 0.387, 0.502, and 0.592, respectively.

cutting issues.[16] The results reported in Online Appendix Table OA4 suggest that the video-based deception scores of PRB banks are associated with an increase in their borrowers' negative incidents across all four dimensions.

VI. CONCLUSION

Using banks' ESG commitment video disclosure, this paper proposes a video-based deception score as an ex ante measure of banks' potential ESG washing. To construct the score, we take advantage of the video processing technologies and advanced machine learning algorithms trained in widely adopted real-life court trial settings to incorporate features in all visual, audio, and textual dimensions to predict the level of deception of the video disclosure. To empirically assess the usefulness of the video-based score in detecting ESG washing, we examine the ESG performance of banks with different levels of deception during the post-video-disclosure period. Following prior literature, we measure banks' ESG performance by using their borrowers' ESG performance (Wang 2023; Kim et al. 2023; Choy et al. 2024). Across various ESG performance measures including outcome-based negative ESG incidence as well as relatively more input-based carbon emissions and ESG ratings, we find that banks with higher deception scores exhibit worse ESG performance during the post-disclosure period. In addition, we find that our results are more pronounced when the video disclosure is longer and only present when the facial recognition quality is high, suggesting that video quality and length is crucial for the video-based deception score to be useful. Comparing banks with high versus low past ESG failures, we find that our results only exist for banks with low past ESG failures, implying that high past failure banks are either more used to lying about ESG or more strategic with their nonverbal cues, resulting in a lower likelihood of leakage in deception cues and thus rendering our deception scores less useful.

[16] Cross-cutting issues refer to those spanning multiple dimensions of ESG.

To unpack the video-based deception score, we first separately examine the detection power of the visual, textual, and audio dimension of the video alone. The results show that only the deception score based on the visual dimension is powerful at detecting ESG washing. To further explore which area of the visual features explain our results, we train six distinct models, with each one solely using one set of visual features as inputs. The outcomes reveal that the model trained with eye features is the most powerful in predicting banks' ESG performance. This is consistent with psychology research that emphasizes the significance of eye movement as a primary indicator in unmasking deception.

Our results are robust to controlling for available bank ESG rating and other video-based measures in prior literature such as persuasion score and quartile groups of the deception scores instead of the continuous version used in the main analyses. All our results hold with or without bank-borrower pair fixed effects. This means that our results still exist even after we fix the lending relationship, suggesting that our results cannot be explained solely by the screening effects (i.e., changes in the bank-borrower pair such as initiation or termination of lending relationships) and could also be related to the monitoring effect. We nonetheless do not focus on the separation of these mechanisms of how banks affect borrowers' ESG performance, which has already been widely examined in prior literature. In addition, as theories provide no guidance on how bank ESG commitment sincerity affect these mechanisms differently, we leave these interesting questions of exploratory nature to future research.

REFERENCES

Baker, A. C., D. F. Larcker, C. G. McClure~~LURE~~, D. Saraph, and E. M. Watts. 2024~~2~~. Diversity washing. *Journal of Accounting Research* 62 (5): 1661-1709.

Baltrušaitis, T., Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. 2018. OpenFace 2.0: Facial Behavior Analysis Toolkit. *IEEE International Conference on Automatic Face and Gesture Recognition.*

Bams, D., and B. van der Kroft. 2022. Tilting the wrong firms? How inflated ESG ratings negate socially responsible investing under information asymmetries. MIT Center for Real Estate Research Paper, 22/12.

Banker, R. D., H. Ding, R. Huang, and X. Li. 2023. Market Reaction to CEOs' Dynamic Hemifacial Asymmetry of Expressions. *Management Science*, forthcoming.

Berg, F., J. F. Koelbel, and R. Rigobon. 2022. Aggregate confusion: The divergence of ESG ratings. *Review of Finance* 26(6): 1315-1344.

Berg, F., K. Fabisik, and Z. Sautner. 2021. Is history repeating itself? The (un) predictable past of ESG ratings. Working paper.

Beyhaghi, M., R. Dai, A. Saunders, and J. Wald. 2021. International lending: The role of lender's home country. *Journal of Money, Credit and Banking* 53(6): 1373-1416.

Beyhaghi, M., R. Dai, A. Saunders, and J. Wald. 2021. International lending: The role of lender's home country. *Journal of Money, Credit and Banking* 53(6): 1373-1416.

Bharath, S. T., S. Dahiya, A. Saunders, and A. Srinivasan. 2011. Lending relationships and loan contract terms. *The Review of Financial Studies* 24(4): 1141-1203.

Bharath, S. T., S. Dahiya, A. Saunders, and A. Srinivasan. 2011. Lending relationships and loan contract terms. *The Review of Financial Studies* 24(4): 1141-1203.

Blankespoor, E., B. E. Hendricks, and G. S. Miller. 2017. Perceptions and price: Evidence from CEO presentations at IPO roadshows. *Journal of Accounting Research* 55(2): 275-327.

Burgoon, J. K., V. Manusov, and L. K. Guerrero. 2021. Nonverbal communication. Routledge.

Choy, S., S. Jiang, S. Liao, and E. Wang, 2024. Public environmental enforcement and private lender monitoring: Evidence from environmental covenants. *Journal of Accounting and Economics* 77(2-3): 101621.

Christensen, D. M., G. Serafeim, and A. Sikochi. 2022. Why is corporate virtue in the eye of the beholder? The case of ESG ratings. *The Accounting Review* 97(1): 147-175.

Cohn, J. B., Z. Liu, and M. I. Wardlaw. 2022. Count (and count-like) data in finance. *Journal* *of Financial Economics* 146(2): 529-551.

Cornaggia, J., and K. Cornaggia. 2023. ESG ratings management. Working paper.

Dávila, A., and M. Guasch. 2022. Managers' body expansiveness, investor perceptions, and firm forecast errors and valuation. *Journal of Accounting Research* 60(2): 517-563.

Ding, M., A. Zhao, Z. Lu, T. Xiang, and J. R. Wen. 2019. Face-focused cross-stream network for deception detection in videos. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 7802–7811.

Dou Y, Z. Xu. 2021. Bank lending and corporate innovation: Evidence from SFAS 166/167. *Contemporary Accounting Research* 38(4): 3017-3052.

Duan, B., D. Hu, and H. Lu. 2024. Video-Based Deception Detection and Financial Fraud. Working paper.

Ekman P. *Telling lies: Clues to deceit in the marketplace, politics, and marriage* (revised edition). WW Norton & Company,1985/2009.

Ekman, P. 2009. *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage.* New York: Norton.

Ekman, P., and W. V. Friesen. 1969. Nonverbal leakage and clues to deception. *Psychiatry* 32 (1): 88–106.

Ekman, P., and W. V. Friesen. 1969. Nonverbal Leakage and Cues to Deception. *Psychiatry* 32(1): 88-105.

Gogate, M., A. Adeel, and A. Hussain. 2017. Deep learning driven multimodal fusion for automated deception detection. In: *2017 IEEE Symposium Series on Computational Intelligence*, 1–6.

Hartwig, M., and C. F. Bond Jr. 2011. Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin* 137(4): 643-659.

Hu, A., and S. Ma. 2024. Persuading Investors: A Video-Based Study. *Journal of Finance*, forthcoming.

Ivanov, I. T., M. S. Kruttli, and S. W. Watugala. 2024. Banking on carbon: Corporate lending and cap-and-trade policy. *The Review of Financial Studies* 37(5): 1640-1684.

Kim, S., and A. Yoon. 2023. Analyzing active fund managers' commitment to ESG: Evidence from the United Nations Principles for Responsible Investment. *Management Science* 69(2): 741-758.

Li, J., and D. Wu. 2020. Do corporate social responsibility engagements lead to real environmental, social, and governance impact?. *Management Science* 66(6): 2564-2588.

Li, X., Y. Lou, and L. Zhang. 2024. Do commercial ties influence ESG ratings? Evidence from Moody's and S&P. *Journal of Accounting Research* 62(5): 1901-1940.

Morales, M. R., S. Scherer, and R. Levitan. 2017. OpenMM: An Open-Source Multimodal Feature Extraction Tool. In: *Conference of the International Speech Communication Association (INTERSPEECH)*, 3354–3358.

Mueller, I., and E. Sfrappini. 2022. Climate change-related regulatory risks and bank lending. Working paper.

Pérez-Rosas, V., M. Abouelenien, R. Mihalcea, and M. Burzo. 2015. Deception detection using real-life trial data. In: *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, 59–66.

Reitmaier, C., W. Schultze, and J. Vollmer. 2024. Corporate responsibility and corporate misbehavior: Are CSR reporting firms indeed responsible?. *Review of Accounting Studies*: 1-69.

Sufi, A. 2007. Information asymmetry and financing arrangements: Evidence from syndicated loans. *The Journal of Finance* 62(2): 629-668.

Tsiamyrtzis, P., J. Dowdall, D. Shastri, I. T. Pavlidis, M. G. Frank, and P. Ekman. 2007. Imaging facial physiology for the detection of deceit. *International Journal of Computer Vision* 71(2): 197-214.

Wang, L. L. 2023. Transmission effects of ESG disclosure regulations through bank lending networks. *Journal of Accounting Research* 61(3): 935-978.

Wu, Z., B. Singh, L. Davis, and V. Subrahmanian. 2018. Deception detection in videos. *Proceedings of the AAAI Conference on Artificial Intelligence* 32 (1).

Zuckerman, M., B. M. DePaulo, and R. Rosenthal. 1981. *Verbal and nonverbal communication of deception.* In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 14, pp. 1–59). New York: Academic Press.

Figure 1. The growth of PRB banks around the world

This figure shows the growth of PRB signatories around the world from 2019 to 2023. The information is available on the PRB's website: https://www.unepfi.org/banking/prbsignatories/.

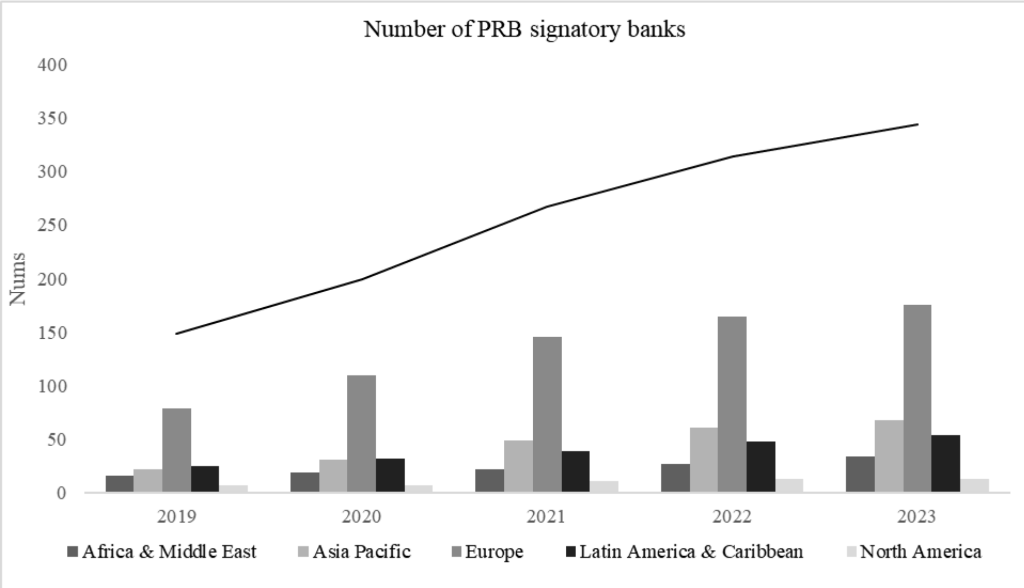
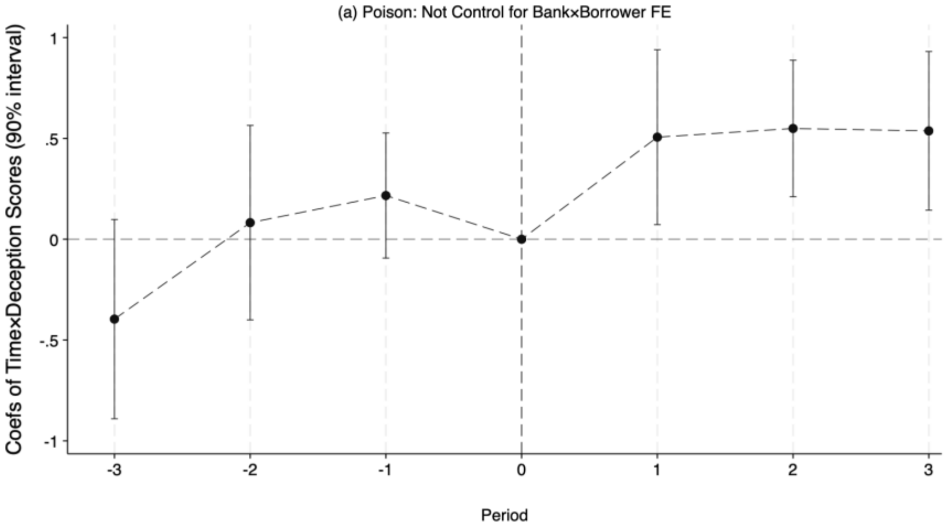


Figure 2. Dynamic effects of the video-based deception scores of PRB banks

This figure shows the Poison regression coefficients and the 10% confidence intervals. We include the full set of control variables that are consistent with Table 2. We include six time indicator variables to substitute *Post*: *I(3yrs before PRB)*, *I(2yrs before PRB)*, *I(1yr before PRB)*, *I(1yr after PRB)*, *I(2yrs after PRB)*, and *I(3yrs after PRB)*. We use *I(PRB Launch Year)* as the reference group and omit it from the regression.



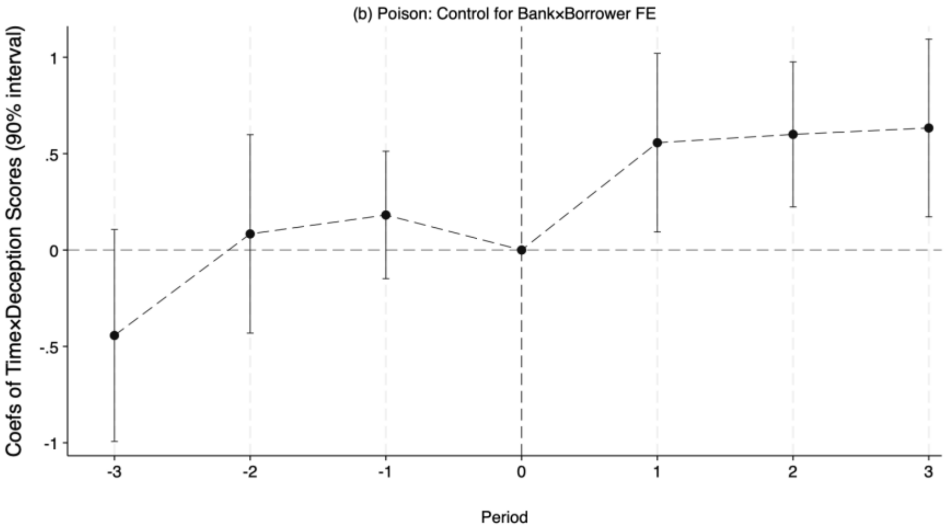


Table 1. Sample selection and descriptive statistics

This table summarizes the sample selection and descriptive statistics of our main analyses. Panel A shows the sample selection procedures. Panel B provides the descriptive statistics of the variables used in our main tests. The sample consists of all the loan contracts of PRB banks with videos in the period of 2016-2022. All variable definitions are listed in Appendix A.

Panel A. Sample selection

	Obs.
Initial Sample: loans data from DealScan between 2016 to 2022 (PRB banks with videos)	25,180
Less: Non-lead banks in the loan contracts	4,008
Less: Borrowers are financial firms (6000-6999)	2,900
Final loan-level observations for analysis	18,272
Initial Sample: Bank-Borrower-year level observations between 2016 to 2022 (PRB banks with videos)	22,884
Less: Borrowers that do not borrow any loans before or after the PRB program	565
Less: Observations with missing bank-level or borrower-level control variables	3,280
Less: Borrowers without ESG data	8,376
Less: Observations that are either singletons or separated by a fixed effect	1,403
Final Bank-Borrower-Year level observations for analysis	9,260
Bank-Borrower pairs in the final sample	2,563

Panel B. Descriptive statistics

	Obs.	Mean	SD	P25	Median	P75
NegIncidents	9,260	4.148	7.011	1.000	2.000	5.000
Deception Scores	9,260	0.498	0.099	0.412	0.530	0.588
Post	9,260	0.594	0.491	0.000	1.000	1.000
BankSize	9,260	20.723	0.823	20.385	21.004	21.134
LoanGr	9,260	2.332	5.486	-0.640	2.610	5.690
BankROE	9,260	7.053	4.748	4.480	8.060	11.060
Tier1	9,260	15.313	2.363	13.490	15.040	17.030
LoanRatio	9,260	39.989	13.206	32.000	37.670	50.220
NII	9,260	43.288	15.551	28.770	42.110	53.470
LLP	9,260	0.019	0.010	0.011	0.019	0.025
FirmSize	9,260	24.096	1.802	22.905	23.986	24.975
Lev	9,260	0.350	0.167	0.236	0.332	0.448
ROA	9,260	0.031	0.072	0.007	0.033	0.062
Current	9,260	0.331	0.176	0.203	0.306	0.429
InteCover	9,260	9.467	21.286	1.688	4.778	10.290
SGA	9,260	0.120	0.130	0.036	0.088	0.160
RD	9,260	0.011	0.020	0.000	0.001	0.015
CAPX	9,260	0.037	0.028	0.017	0.031	0.049

Table 2. The video-based deception scores and the real ESG outcomes

This table reports the results of using the video-based deception scores of PRB banks to evaluate the real ESG outcomes of their lending relationships during the post-video-disclosure period. The dependent variable is *NegIcidents*. *NegIcidents* is the number of negative ESG incidents of a bank's borrower. We run these regressions using the Poison model as suggested by Cohn et al. (2022). A set of firm characteristics and bank characteristics are controlled for. All variable definitions are listed in Appendix A. Standard errors, adjusted for heteroskedasticity and clustered by bank-borrower pair, are reported in parentheses. ***, **, and * indicate a two-tailed test significance level of less than 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)
Dep. var. =			*NegIncidents*		
Post×Deception Scores	0.546***	0.545***	0.517***	0.504***	0.573***
	(0.208)	(0.206)	(0.188)	(0.167)	(0.201)
BankSize	-0.011	-0.012	-0.017	-0.011	-0.010
	(0.020)	(0.020)	(0.018)	(0.016)	(0.018)
LoanGrowth	0.002	0.002	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
BankROE	0.003	0.003	0.003	0.003	0.002
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Tier1	0.010	0.009	0.010	0.010	0.011
	(0.011)	(0.011)	(0.010)	(0.009)	(0.010)
LoanRatio	0.002	0.003	0.002	0.003	0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
NII	0.000	0.000	0.000	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
LLP	-2.060	-2.160	-1.851	-1.955	-2.499
	(2.121)	(2.141)	(1.750)	(1.671)	(1.931)
FirmSize		0.181***	0.275***	0.231***	0.196**
		(0.061)	(0.071)	(0.072)	(0.078)
Lev		0.262	0.210	0.229	0.237
		(0.177)	(0.196)	(0.191)	(0.209)
ROA		0.767***	0.470**	0.509**	0.537**
		(0.178)	(0.209)	(0.220)	(0.226)
Current		-0.410**	-0.578**	-0.493*	-0.344
		(0.198)	(0.250)	(0.258)	(0.263)
InteCover		-0.001*	-0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.001)	(0.001)
SGA		-0.444	-0.454	-0.234	-0.152
		(0.337)	(0.352)	(0.367)	(0.412)
RD		7.969***	8.419***	6.837***	6.239**
		(2.503)	(2.486)	(2.567)	(2.586)
CAPX		0.906	-0.005	0.967	1.159
		(0.811)	(0.818)	(0.841)	(0.900)

Bank FE	Yes	Yes	Yes	Yes	-
Borrower FE	Yes	Yes	Yes	Yes	-
Bank×Borrower FE	-	-	-	-	Yes
Year FE	Yes	Yes	-	-	-
Country×Year FE	-	-	Yes	Yes	Yes
Industry×Year FE	-	-	-	Yes	Yes
N	9,260	9,260	9,260	9,260	8843
Adj. R^2	0.638	0.639	0.658	0.662	0.663

Table 3. Cross-sectional analyses: the detecting power of video-based deception scores

This table reports the cross-sectional results based on the video duration, face recognition quality, and banks'past failures. The dependent variable is *NegIcidents*. *NegIcidents* is the number of negative ESG incidents of a bank's borrower. We run these regressions using the Poison model as suggested by Cohn et al. (2022). A set of firm characteristics and bank characteristics are controlled for. All variable definitions are listed in Appendix A. Standard errors, adjusted for heteroskedasticity and clustered by bank-borrower pair, are reported in parentheses. ***, **, and * indicate a two-tailed test significance level of less than 1%, 5%, and 10%, respectively.

Panel A. Video duration and the detecting power

	(1)	(2)	(3)	(4)
Sample =	High Video Duration	Low Video Duration	High Video Duration	Low Video Duration
Dep. var. =	*NegIncidents*			
Post×Deception Scores	1.237***	0.365*	1.324**	0.375*
	(0.474)	(0.188)	(0.535)	(0.226)
Dif =	0.872***		0.950***	
Controls	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	-	-
Borrower FE	Yes	Yes	-	-
Bank×Borrower FE	-	-	Yes	Yes
Country×Year FE	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes
N	3591	5511	3476	5312
Adj. R^2	0.636	0.679	0.642	0.678

Panel B. The quality of face recognition and the detecting power

	(1)	(2)	(3)	(4)
Sample =	High Recognition Quality	Low Recognition Quality	High Recognition Quality	Low Recognition Quality
Dep. var. =	*NegIncidents*			
Post×Deception Scores	0.430**	0.238	0.544**	0.213
	(0.202)	(0.434)	(0.245)	(0.481)
Dif =	0.192***		0.331**	
Controls	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	-	-
Borrower FE	Yes	Yes	-	-
Bank×Borrower FE	-	-	Yes	Yes
Country×Year FE	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes

N	5742	3389	5505	3267
Adj. R^2	0.668	0.655	0.668	0.659

Panel C. Banks' past ESG failures and the detecting power

	(1)	(2)	(3)	(4)
Sample =	High Past Failures	Low Past Failures	High Past Failures	Low Past Failures
Dep. var. =		*NegIncidents*		
Post×Deception Scores	0.200	0.615***	0.269	0.537**
	(0.353)	(0.233)	(0.390)	(0.252)
Dif =	-0.414***		-0.339*	
Controls	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	-	-
Borrower FE	Yes	Yes	-	-
Bank×Borrower FE	-	-	Yes	Yes
Country×Year FE	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes
N	3559	4533	3496	4486
Adj. R^2	0.673	0.640	0.674	0.644

Table 4. The usefulness of different features in constructing the deception scores

This table reports the results of using only visual-, audio-, and textual-based features in the videos to train the deception scores of PRB banks. The dependent variable is *NegIcidents*. *NegIcidents* is the number of negative ESG incidents of borrower firms in the year. We run these regressions using the Poison model as suggested by Cohn et al. (2022). A set of firm characteristics and bank characteristics are controlled for, as in the Table 2. All variable definitions are listed in Appendix A. In Panel A, we compare the deception scores that are trained using only the visual, audio, and textual features, respectively. In Panel B, we compare the deception scores that are trained using specific categories of visual features in the videos. We divide the visual features of videos into six categories: gaze, eye, face pose, face LMK, face shape, and facial action units. Standard errors, adjusted for heteroskedasticity and clustered by bank-borrower pair, are reported in parentheses. ***, **, and * indicate a two-tailed test significance level of less than 1%, 5%, and 10%, respectively.

Panel A. Deception scores based on visual, audio, and textual features in the videos

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var. =				*NegIncidents*				
Post×Deception Scores_V	0.323**			0.337**	0.346**			0.363*
	(0.129)			(0.158)	(0.155)			(0.188)
Post×Deception Scores_A		0.252		-0.019		0.283		-0.000
		(0.193)		(0.235)		(0.229)		(0.274)
Post×Deception Scores_T			0.014	-0.046			-0.048	-0.110
			(0.149)	(0.148)			(0.178)	(0.177)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes				
Borrower FE	Yes	Yes	Yes	Yes				
Bank×Borrower FE					Yes	Yes	Yes	Yes
Country×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9,260	9,260	9,260	9,260	8843	8843	8843	8843
Adj. R^2	0.662	0.662	0.662	0.662	0.663	0.663	0.663	0.663

Panel B. Deception scores based on specific categories of visual features

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. var. =				*NegIncidents*			
Post×Deception Scores_Gaze	0.031						-0.003
	(0.066)						(0.126)
Post×Deception Scores_Eye		0.380***					0.711*
		(0.133)					(0.416)
Post×Deception Scores_Face Pose			0.118				-0.014
			(0.073)				(0.216)
Post×Deception Scores_Face LMK				0.188			-0.322
				(0.115)			(0.347)
Post×Deception Scores_Face Shape					-0.223*		-0.187
					(0.124)		(0.277)
Post×Deception Scores_Facial Action Unit						0.483**	0.039
						(0.200)	(0.404)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Borrower FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9260	9260	9260	9260	9260	9260	9260
Adj. R^2	0.662	0.662	0.662	0.662	0.662	0.662	0.663
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Dep. var. =				*NegIncidents*			
Post×Deception Scores_Gaze	0.059						0.050
	(0.080)						(0.153)

Post×Deception Scores_Eye		0.411**					0.811*
		(0.161)					(0.488)
Post×Deception Scores_Face Pose			0.145*				-0.049
			(0.087)				(0.256)
Post×Deception Scores_Face LMK				0.174			-0.383
				(0.140)			(0.410)
Post×Deception Scores_Face Shape					-0.314**		-0.285
					(0.155)		(0.328)
Post×Deception Scores_Facial Action Unit						0.609**	0.043
						(0.242)	(0.471)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank×Borrower FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	8843	8843	8843	8843	8843	8843	8843
Adj. R^2	0.663	0.663	0.663	0.663	0.663	0.663	0.663

Table 5. Other ESG performance of lending relationships

This table reports the results of using video-based deception scores of PRB banks to evaluate the other ESG performance of their ex post lending relationships. The dependent variables are *Borrower ESG Combined Ratings*, *Borrower ESG reporting Ratings*, *Borrower ESG strategy Ratings*, and *Borrower Co2 Intensity*. *Borrower Combined ESG Ratings* is the borrower firm's ESG combined score in the year, which captures the overall ESG performance of the borrower firms. *Borrower ESG reporting Ratings* is the borrower firm's ESG disclosure score in the year, which captures the borrower firms' ESG disclosure performance. *Borrower ESG strategy Ratings* is the borrower firm's ESG strategy score in the year, which reflects borrower firms' practices to communicate that it integrates the economic (financial), social and environmental dimensions into its day-to-day decision-making processes. *Borrower Co2 Intensity* is the borrower firms' total Co2 and Co2-equivalent emissions (in thousands of tons), scaled by sales (in millions) in the year. We run these regressions using the OLS model. A set of firm characteristics and bank characteristics are controlled for, as in the Table 2. All variable definitions are listed in Appendix A. Standard errors, adjusted for heteroskedasticity and clustered by bank-borrower pair, are reported in parentheses. ***, **, and * indicate a two-tailed test significance level of less than 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var. =	*Borrower Combined ESG Ratings*		*Borrower ESG reporting Ratings*		*Borrower ESG strategy Ratings*		*Borrower Co2 Intensity*	
Post×Deception Scores	-6.726***	-7.324***	-10.404**	-11.771***	-9.487**	-9.492**	0.487**	0.563**
	(2.571)	(2.835)	(4.130)	(4.563)	(3.824)	(4.205)	(0.238)	(0.268)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	-	Yes	-	Yes	-	Yes	-
Borrower FE	Yes	-	Yes	-	Yes	-	Yes	-
Bank×Borrower FE	-	Yes	-	Yes	-	Yes	-	Yes
Country×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	10795	10354	10795	10354	10795	10354	10795	10354
Adj. R^2	0.800	0.758	0.688	0.622	0.872	0.847	0.853	0.828

Table 6. Robustness tests

This table reports the results of using the quartile ranks of the video-based deception scores to replace the continuous deception scores. The dependent variable is *NegIcidents*. *NegIcidents* is the number of negative ESG incidents of borrower firms in the year. We run these regressions using the Poison model as suggested by Cohn et al. (2022). A set of firm characteristics and bank characteristics are controlled for, as in the Table 2. All variable definitions are listed in Appendix A. In Panel A, we compare the informativeness of our video-based deception scores with that of the most recent available ESG ratings of PRB banks in assessing the real ESG outcomes of their ex post lending relationships. In Panel B, we re-evaluate our video-based deception scores after controlling for the video-based persuasiveness scores suggested by Hu and Ma (2024). In Panel C, we use the quartile ranks of the video-based deception scores to substitute the continuous deception scores. Standard errors, adjusted for heteroskedasticity and clustered by bank-borrower pair, are reported in parentheses. ***, **, and * indicate a two-tailed test significance level of less than 1%, 5%, and 10%, respectively.

Panel A. Controlling for the most recent available ESG ratings of PRB banks

	(1)	(2)
Dep. var. =	*NegIncidents*	
Post×Deception Scores	0.574***	0.664***
	(0.180)	(0.218)
Post×Avail_BankESGratings	-0.002	-0.003
	(0.002)	(0.003)
Post×I(Missing_BankESGratings)	-0.122	-0.132
	(0.128)	(0.157)
Controls	Yes	Yes
Bank FE	Yes	-
Borrower FE	Yes	-
Bank×Borrower FE	-	Yes
Country×Year FE	Yes	Yes
Industry×Year FE	Yes	Yes
N	9,260	8843
Adj. R^2	0.662	0.663

Panel B. Controlling for the persuasiveness scores in the videos

	(1)	(2)
Dep. var. =	*NegIncidents*	
Post×Deception Scores	0.488***	0.542***
	(0.176)	(0.209)
Post×Persuasiveness_PCA	-0.003	-0.007
	(0.011)	(0.014)
Controls	Yes	Yes
Bank FE	Yes	-
Borrower FE	Yes	-

Bank×Borrower FE	-	Yes
Country×Year FE	Yes	Yes
Industry×Year FE	Yes	Yes
N	9,260	8843
Adj. R^2	0.662	0.663

Panel C. Quartile rank transformation of the video-based deception scores

	(1)	(2)
Dep. var. =	*NegIncidents*	
Post×I(Deception Scores: [25%, 50%])	0.130	0.109
	(0.081)	(0.098)
Post×I(Deception Scores: [50%, 75%])	0.174**	0.169*
	(0.079)	(0.094)
Post×I(Deception Scores: [75%, 100%])	0.226***	0.232**
	(0.084)	(0.100)
Controls	Yes	Yes
Bank FE	Yes	-
Borrower FE	Yes	-
Bank×Borrower FE	-	Yes
Country×Year FE	Yes	Yes
Industry×Year FE	Yes	Yes
N	9,260	8843
Adj. R^2	0.662	0.663

Appendix A. Variable definition

Variable	Definition
NegIncidents	The number of negative ESG incidents of borrower firms in the year. Datasource: RepRisk.
Post	A dummy variable that takes the value of one after the bank joins the PRB program, and zero otherwise.
Deception Scores	The deception scores of PRB banks' videos, which captures the possibility of deception. We obtain the PRB banks' videos from UNEP FI's YouTube account, in which the CEOs from PRB banks talk about why their bank signs the principles and what it means for their business. Datasource: YouTube.
BankSize	The natural logarithm of bank's total assets. Datasource: Bankscope.
LoanGr	Bank's annual growth ratio of net loans. Datasource: Bankscope.
BankROE	Bank's return on equity, defined as the ratio of net income to total equity. Datasource: Bankscope.
Tier1	Bank's tier1 risk-based capital ratio. Datasource: Bankscope.
LoanRatio	Bank's ratio of net loans to total assets. Datasource: Bankscope.
NII	Bank's non-interest income over total income. Datasource: Bankscope.
LLP	Bank's loan loss provisions over net loans. Datasource: Bankscope.
FirmSize	The natural logarithm of borrower firm's total assets. Datasource: Worldscope.
Lev	Borrower firm's ratio of total debt to total assets. Datasource: Worldscope.
ROA	Borrower firm's return on assets, defined as the ratio of net income to total assets. Datasource: Worldscope.
Current	Borrower firm's ratio of current assets to total assets. Datasource: Worldscope.
InteCover	Borrower firm's ratio of earnings before interest and tax to the interest expense. Datasource: Worldscope.
SGA	Borrower firm's selling, general, and administrative expense scaled by total assets. Datasource: Worldscope.
RD	Borrower firm's research and development expense scaled by total assets. Datasource: Worldscope.
CAPX	Borrower firm's capital expenditures scaled by total assets. Datasource: Worldscope.
Deception Scores_V (A, T)	The visual- (audio-, texual-) based deception scores of PRB banks' videos, trained only on the visual (audio, textual) features of videos.
Deception Scores_Features Category	The deception scores of PRB banks' videos, trained by specific categories of visual features. We divide the visual features of videos into six categories: gaze, eye, face pose, face LMK, face shape, and facial action units.
Borrower ESG Ratings	Borrower firm's ESG ombined score in the year, which captures the overall ESG performance of the borrower firms. Datasource: Refinitiv ASSET4.
Borrower ESG reporting Ratings	Borrower firm's ESG disclosure score in the year, which captures the borrower firms' ESG disclosure performance. Datasource: Refinitiv ASSET4.

Borrower ESG strategy Ratings	Borrower firm's ESG strategy score in the year, which reflects borrower firms' practices to communicate that it integrates the economic (financial), social and environmental dimensions into its day-to-day decision-making processes. Datasource: Refinitiv ASSET4.
Borrower Co2 Intensity	Total CO2 and CO2-equivalent emissions (in thousands of tons), scaled by revenues (in millions) in the year. Datasource: Refinitiv ASSET4.
Avail_BankESGratings	The most recently available ESG ratings of PRB banks, zero if missing. Datasource: Refinitiv ASSET4.
I(Missing_BankESGratings)	A dummy variable that takes the value of one if the bank is not covered by Refinitiv ESG ratings, and zero otherwise. Datasource: Refinitiv ASSET4.
Persuasiveness_PCA	The factor with the highest eigenvalue using the Principal Component Method to estimate from visual, vocal, verbal emotions and visual beauty.
I(WithVideo)	A dummy variable that takes the value of one if the bank provides a video to PRB that is available on YouTube, and zero otherwise.
I(SLL)	A dummy variable that takes the value of one for sustainability-linked loan (SLL). We identify a loan facility to be an SLL based on the classifications of market segment and deal remark (e.g., "sustainability", "sustainable", "esg", and "green loans"). Datasource: Refinitiv DealScan.
LogMaturity	The natural logarithm of the number of months between loan start and end dates. Datasource: Refinitiv DealScan.
LogSpread	The natural logarithm of loan spread, defined as the sum of the spread over LIBOR plus the facility fee. Datasource: Refinitiv DealScan.
LogLoanSize	The natural logarithm of one plus the amounts (in million USD) of syndicated loans. Datasource: Refinitiv DealScan.
I(Collateral)	A dummy variable that takes the value of one if the loan is secured and zero otherwise. Datasource: Refinitiv DealScan.
LogLenderNum	The natural logarithm of one plus the total number of lenders involved in the loan. Datasource: Refinitiv DealScan.
I(Covenant)	A dummy variable that takes the value of one if the loan has any covenant and zero otherwise.
LnDeposit	The natural logarithm of total deposits. Datasource: Bankscope. Datasource: Refinitiv DealScan.
DepositRatio	Total deposits divided by total assets. Datasource: Bankscope.
I(Listed)	A dummy variable that takes the value of one if the bank is listed in year t, and zero otherwise. Datasource: Bankscope.
I(SBTi)	A dummy variable that takes the value of one if the bank has reported carbon-related targets to Science Based Targets initiative (SBTi), and zero otherwise. Datasource: Bankscope and SBTi.
I(Developed Country)	A dummy variable that takes the value of one for banks located in developed countries. Datasource: Bankscope and United Nations.
ClimateConcern	We measure the level of climate concerns from the Lloyd's Register Foundation (2020)'s 2019 World Risk Poll (e.g., Zhang, 2023). The survey asks how the interviewees perceive the threat of climate change: very serious, somewhat serious, or not a threat at all. We compute the country-level climate concerns as the total fraction who answer "very serious" or "somewhat serious".

I(PRB Founders)	A dummy variable that takes the value of one if the bank is one of the founding banks of the PRB program, and zero otherwise. Datasource: United Nations.

Online Appendix

Figure OA1. Illustration of Face Features Output using OpenFace

This figure illustrates the facial feature outputs generated by OpenFace. For presentation purposes, we display a subset of facial features in a graphical user interface (GUI) format. The complete set of extracted facial features is stored in CSV files during the data processing phase.

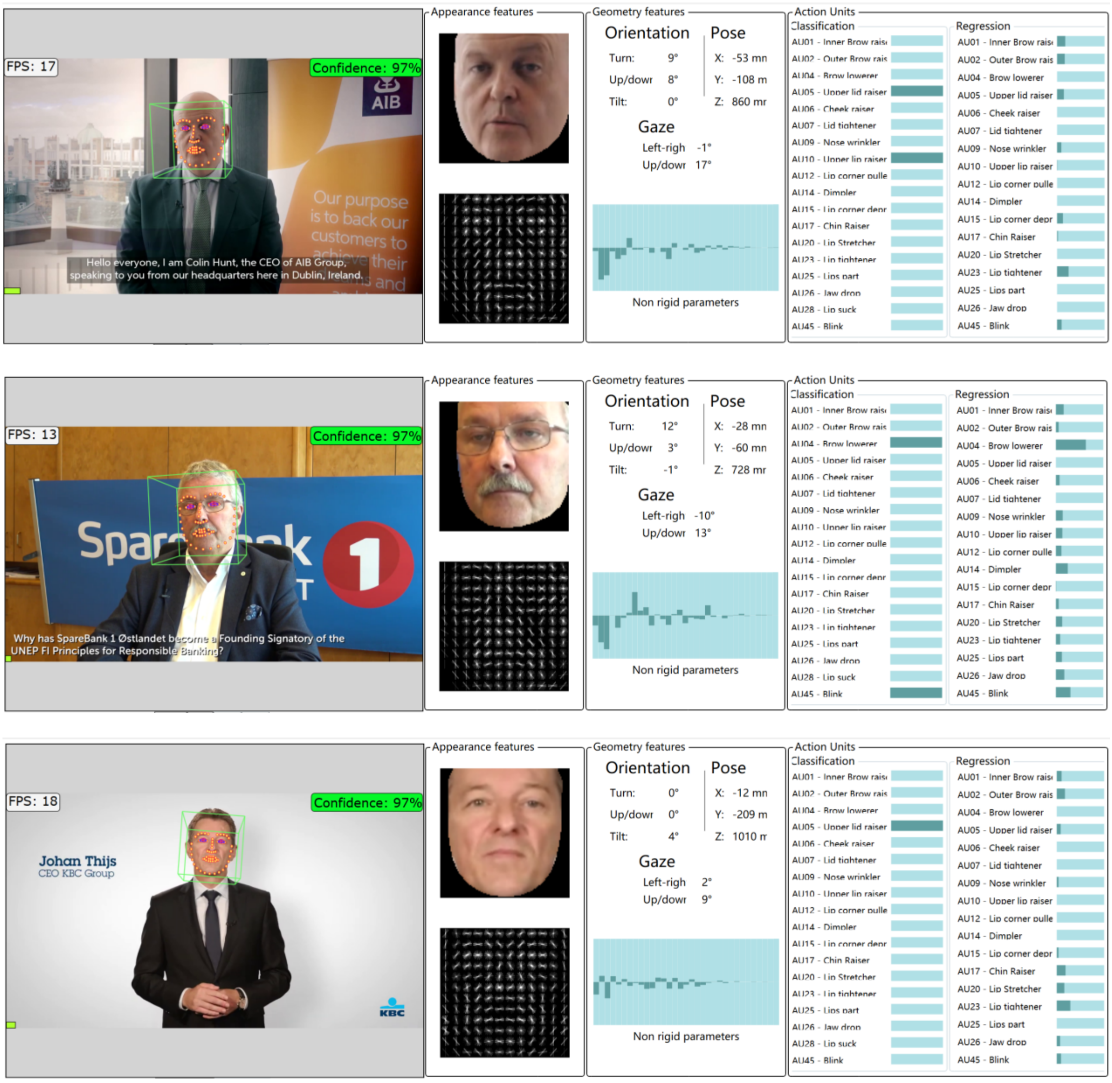
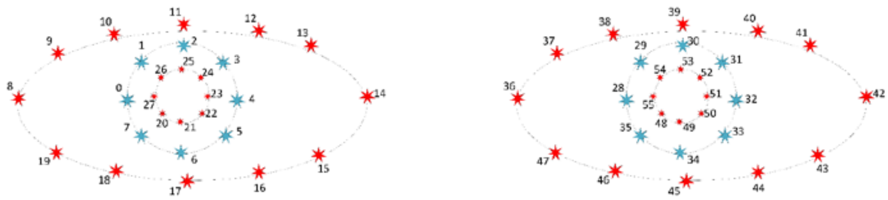
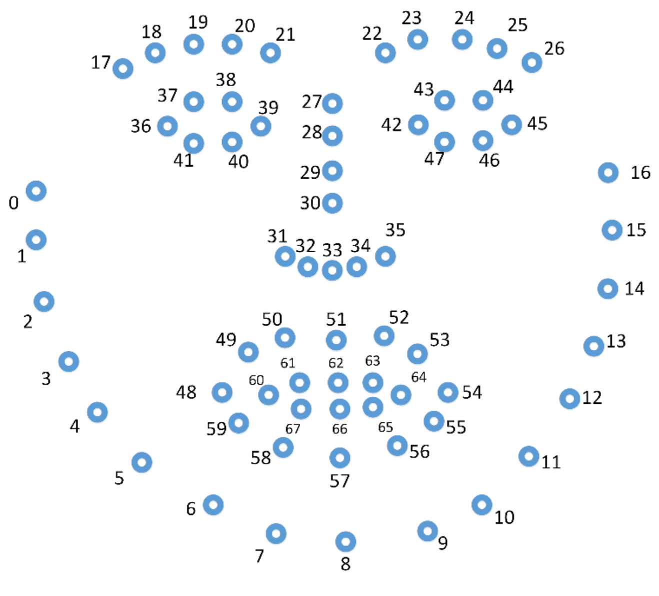


Table OA1. Visual Feature Description

This table provides the specific names and descriptions of the visual features extracted by OpenFace, organized by category.

Visual feature category	**Visual feature name**	**Visual features**	**Description**	**No. of features**
Gaze-related information	Eye gaze direction vector in world coordinate	gaze_0_x, gaze_0_y, gaze_0_z; gaze_1_x, gaze_1_y, gaze_1_z	Eye 0 is the leftmost eye in the image, eye 1 is the rightmost eye in the image. Take eye 0 as an example. gaze_0_x, gaze_0_y, and gaze_0_z refer to eye gaze direction vector in world coordinates for eye landmark 0. Think of it as a ray going from the left eye in the image in the direction of the eye gaze.	6
	Eye gaze direction in radians in world coordinates averaged for both eyes and converted into more easy to use format than gaze vectors	gaze_angle_x, gaze_angle_y	If a person is looking left-right this will results in the change of gaze_angle_x (from positive to negative) and, if a person is looking up-down this will result in change of gaze_angle_y (from negative to positive), if a person is looking straight ahead both of the angles will be close to 0 (within measurement error).	2
	location of 2D eye region landmarks in pixels	eye_lmk_x_0, eye_lmk_x_1,... eye_lmk_x55; eye_lmk_y_1,... eye_lmk_y_55	There are a total 56 eye landmarks, leading to a total 112(56*2) features of 2D eye region landmark. The landmark index can be found below.	112
	location of 3D eye region landmarks in millimeters	eye_lmk_X_0, eye_lmk_X_1,... eye_lmk_X55; eye_lmk_Y_0,... eye_lmk_Z_55	There are a total 56 eye landmarks, leading to a total 168(56*3) features of 3D eye region landmark. The landmark index can be found below.	168
	 **Eye landmark index** ** The figure illustrates the eye landmark indices used by OpenFace. The landmarks are*			-

	plotted around the outline of each eye, with different indices corresponding to specific points on the eye's contour and within the eye region.			
	Total			**288**
Head and face location detail	The location of the head	pose_Tx, pose_Ty, pose_Tz	The location of the head with respect to camera in millimeters	3
	The rotation of head	pose_Rx, pose_Ry, pose_Rz	Rotation is in radians around X,Y,Z axes with the convention R = Rx * Ry * Rz, left-handed positive sign. This can be seen as pitch (Rx), yaw (Ry), and roll (Rz). The rotation is in world coordinates with camera being the origin.	3
	Face landmarks locations in 2D	x_0, x_1, ... x_66, x_67, y_0,...,y_67	Face location of 2D landmarks in pixels. There are a total 68 eye landmarks, leading to a total 136(68*2) features of 2D eye region landmark. The landmark index can be seen below.	136
	Face landmarks locations in 3D	X_0,...,X_67, Y_0,...,Y_67, Z_0,...,Z_67	Face location of 3D landmarks in millimetres. There are a total 68 eye landmarks, leading to a total 204(68*3) features of 3D eye region landmark. The landmark index can be seen below.	204
	(figure below)			-



Face landmark index

** The figure illustrates the facial landmark indices used by OpenFace. The landmarks are plotted around the outline of face, with different indices corresponding to specific points on the face.*

	Total			**346**
Face shape characteristics	Rigid face shape parameters	p_scale, p_rx, p_ry, p_rz, p_tx, p_ty	Parameters of a point distribution model (PDM) that describe the rigid face shape (location, scale and rotation)	6
	Non-rigid	p_0, p_1, ... p_33	Parameters of a point distribution model	34

	shape parameters		(PDM) that describe the non-rigid face shape (deformation due to expression and identity).	
	Total			**40**
Facial Action Units (AUs)	AU intensities	AU01_r, AU02_r, AU04_r, AU05_r, AU06_r, AU07_r, AU09_r, AU10_r, AU12_r, AU14_r, AU15_r, AU17_r, AU20_r, AU23_r, AU25_r, AU26_r, AU45_r	The intensity (from 0 to 5) of each facial AU. Facial Action Units (AUs) are a way to describe human facial expression, more details on Action Units can be found "https://www.cs.cmu.edu/~face/facs.htm"	17
	AU occurrences	AU01_c, AU02_c, AU04_c, AU05_c, AU06_c, AU07_c, AU09_c, AU10_c, AU12_c, AU14_c, AU15_c, AU17_c, AU20_c, AU23_c, AU25_c, AU26_c, AU28_c, AU45_c	The presence (0 absent, 1 present) of each facial AU. Facial Action Units (AUs) are a way to describe human facial expression, more details on Action Units can be found "https://www.cs.cmu.edu/~face/facs.htm"	18
	Total			**35**
Total				**709**

Table OA2. Deception scores trained by alternative model - Gradient Boosted Decision Trees (GBDT)

This table reports the results of using video-based deception scores trained by an alternative machine learning model (GBDT) to evaluate the real ESG outcomes of their ex post lending relationships. The dependent variable is *NegIcidents*. *NegIcidents* is the number of negative ESG incidents of borrower firms in the year. We run these regressions using the Poison model as suggested by Cohn et al. (2022). A set of firm characteristics and bank characteristics are controlled for, as in the Table 2. All variable definitions are listed in Appendix A. Standard errors, adjusted for heteroskedasticity and clustered by bank-borrower pair, are reported in parentheses. ***, **, and * indicate a two-tailed test significance level of less than 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)
Dep. var. =	*NegIncidents*			
Post×Deception Scores_GBDT	1.064***	1.246***		
	(0.353)	(0.423)		
Post×Deception Scores_GBDT_V			0.949***	1.092***
			(0.354)	(0.423)
Post×Deception Scores_GBDT_A			0.228	0.186
			(0.525)	(0.612)
Post×Deception Scores_GBDT_T			-0.098	-0.256
			(0.309)	(0.372)
Controls	Yes	Yes	Yes	Yes
Bank FE	Yes	-	Yes	-
Borrower FE	Yes	-	Yes	-
Bank×Borrower FE	-	Yes	-	Yes
Country×Year FE	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes
N	9,260	8843	9,260	8843
Adj. R^2	0.662	0.663	0.662	0.663

Table OA3. Alternative clustering of standard errors

This table reports the results of using alternative clustering of standard errors. The dependent variable is *NegIcidents*. *NegIcidents* is the number of negative ESG incidents of borrower firms in the year. We run these regressions using the Poison model as suggested by Cohn et al. (2022). A set of firm characteristics and bank characteristics are controlled for, as in the Table 2. All variable definitions are listed in Appendix A. Standard errors, adjusted for heteroskedasticity and clustered by bank-borrower pair, are reported in parentheses. ***, **, and * indicate a two-tailed test significance level of less than 1%, 5%, and 10%, respectively.

	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Clustering:	Bank, Borrower		Bank		Borrower		Industry, Country	
Dep. var. =	*NegIncidents*							
Post×Deception Scores	0.504***	0.573***	0.504***	0.573***	0.504***	0.573***	0.504***	0.573***
	(0.040)	(0.057)	(0.105)	(0.120)	(0.122)	(0.152)	(0.126)	(0.182)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	-	Yes	-	Yes	-	Yes	-
Borrower FE	Yes	-	Yes	-	Yes	-	Yes	-
Bank×Borrower FE	-	Yes	-	Yes	-	Yes	-	Yes
Country×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9,260	8843	9,260	8843	9,260	8843	9,260	8843
Adj. R^2	0.662	0.663	0.662	0.663	0.662	0.663	0.662	0.663

Table OA4. The dimensions of negative ESG incidents

This table reports the results of comparing the ESG negative incidents with different features. We run these regressions using the Poison model as suggested by Cohn et al. (2022). A set of firm characteristics and bank characteristics are controlled for, as in the Table 2. All variable definitions are listed in Appendix A. Standard errors, adjusted for heteroskedasticity and clustered by bank-borrower pair, are reported in parentheses. ***, **, and * indicate a two-tailed test significance level of less than 1%, 5%, and 10%, respectively.

Panel A. Not controlling for the bank-borrower pair fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var. =	*NegIncidents_Env*		*NegIncidents_Social*		*NegIncidents_Gov*		*NegIncidents_CrossCutting*	
Post×Deception Scores	0.486**	0.546**	0.573**	0.646**	0.540***	0.635**	0.457**	0.533**
	(0.214)	(0.263)	(0.244)	(0.298)	(0.203)	(0.250)	(0.181)	(0.215)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes		Yes		Yes		Yes	
Borrower FE	Yes		Yes		Yes		Yes	
Bank×Borrower FE		Yes		Yes		Yes		Yes
Country×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	6372	5913	7125	6751	8084	7627	8592	8201
Adj. R^2	0.705	0.704	0.655	0.655	0.555	0.552	0.663	0.664

Online Appendix B. PRB's Guide to Producing Commitment Videos

1. The exemplary questions provided by the PRB are as follows:

- #1 Why is [xx bank] involved in establishing these Principles for Responsible Banking?
- #2 Why is there a need for global Principles for Responsible Banking? What is different about them from existing frameworks? Why are they needed now?
- #3 Why do you see alignment with societal goals - as expressed in the Sustainable Development Goals and the Paris Climate Agreement - as important to strategically position your bank for future success? What value do these Principles bring to your bank, your shareholders and customers?
- #4 The Principles also call on banks to publicly set targets and report back on their progress. Why do you think that's an important feature of the Principles?
- #5 What changes in your bank do you see these Principles guiding or accelerating?
- #6 How do you see these Principles helping your bank to identify and seize emerging opportunities?

2. The guidelines for video production provided by the PRB are as follows:

Location/setting

- Maybe the CEO is sitting in a meeting room
- There should be something on the table or behind him/her that identifies it as your bank (e.g., logo, banner, etc.)
- The background should not be too distracting

Set-up

- We are looking for a tight close-up – head and shoulders – of your CEO in the frame
- Film landscape (horizontally) and place the camera level with the CEO
- Fix the camera on a tripod
- Use an external microphone (e.g., a lapel microphone) on the CEO
- If you can set the lighting, make it in front of the CEO, but to one side, not head on
- Have the CEO speak just to one side of the camera, i.e. at a hidden interviewer

Filming

- Test the focus and film and sound quality before conducting the whole interview
- Record both the questions and the answers

Let the film run on between question and answers so there is “white space” we can cut into to make the editing easier.

Online Appendix C. The Construction of Video-Based Persuasion Measure

In Online Appendix C, we describe the details of how we construct the video-based persuasion measure, following Hu and Ma (2024).

First, to capture visual emotion, we represent the PRB commitment videos as images sampled at ten frames per second. Using Face++, a face-detection machine learning algorithm, we identify human faces in these frames and generate a visual emotion measure. The *Face++* platform provides APIs through which we feed the raw images into the cloud computing system and receive a host of face-related measures constructed by *Face++*'s machine learning algorithms. Those measures include visual emotions, beauty, age, gender, etc. The *Face++* emotion recognition algorithm API classifies visual emotion into seven categories: happiness, neutral, sadness, surprise, anger, disgust, and fear. Specifically, for each frame, the API gives each category a predicted score between 0 to 1, indicating the likelihood that the frame's emotion belongs to that category. The scores of seven categories sum up to one. For each frame, the emotion category that has the highest predicted score from *Face++* is used to label the emotion of the frame. Following Hu and Ma (2024), we classify a frame as positive if its emotion label is "happiness", as negative if its emotion label is "sadness", "anger", "disgust", or "fear".[17] The visual positive tone during a PRB commitment videos is calculated as the number of positive frames scaled by the number of total frames, and the textual negative visual emotion is calculated as the number of negative frames scaled by the number of total frames.

[17] The "surprise" category is not classified as either positive or negative following prior literature (Curti and Kazinnik, 2023; Hu and Ma, 2024).

Second, as to audio emotion, we use the deep neural networks (CNNs) model trained and provided by Pinto et al. (2020) to classify emotions from audio files extracted from PRB commitment videos. The Pinto et al. (2020) model classifies the audio of each word into eight different emotion categories (neutral, calm, happy, sad, angry, fearful, disgust, surprise). In line with the face emotion classification, we classify “sad”, “anger”, “disgust”, and “fearful” as negative emotions, and “happy” as a positive emotion. The audio positive emotion is calculated as the number of words with positive audio emotion scaled by the number of words, and the audio negative emotion is calculated as the number of words with negative audio emotion scaled by the number of words.

Third, we construct textual emotion by extracting speech transcriptions and applying the Loughran and McDonald (2011) dictionary. Specifically, we use *Vosk*, a speech recognition toolkit, to transcribe the PRB commitment videos. The transcriptions include a list of words, timestamps (onsets and offsets), and punctuation. The textual positive tone during a PRB commitment videos is calculated as the number of positive words scaled by the number of words, and the textual negative tone is calculated as the number of negative words scaled by the number of words. To determine positive and negative words, we rely on the 2020 version of the Loughran and McDonald (2011) dictionary.

Fourth, we measure the CEO’s facial beauty using Face++’s face detection API. The API provides two predicted beauty scores for each detected face: a male beauty score and a female beauty score, both ranging from 0 to 100, indicating the perceived beauty level of the face from male and female perspectives, respectively. To calculate the CEO’s beauty score, we take the mean of the male and female beauty scores across all frames.