LIE TO ME: VIDEO-BASED DETECTION OF ESG WASHING

YIWEI DOU, NYU
BINGLEI DUAN, CUFE
DANQI HU, PKU
CHANGTUO LUO, PKU

ABFER 2025

光华管理学院
Guanghua School of Management

Motivation

- Increasing concerns about ESG washing among firms
  - Firms increasingly commit to ESG principles, but there are concerns about the sincerity of these commitments.
  - Studies indicate a mismatch between ESG commitments and subsequent actual performance (Kim and Yoon 2023).

- Challenges in detecting sincere ESG commitments ex ante (Dechow 2023; Briscoe-Tran 2024).
  - Ex post evaluations: Studies construct company-specific measures of ESG washing that benchmark the committed against actual performance (Marquis et al. 2016; Baker et al. 2024).
  - A viable ex ante approach to assess the ESG washing is still lacking.

Research Question

- Are video-based deception scores useful to detect ESG washing ex ante?
  - ESG washers will be identified ex ante before the ex post performance is realized
  - A timelier and more broad ESG washing indicator

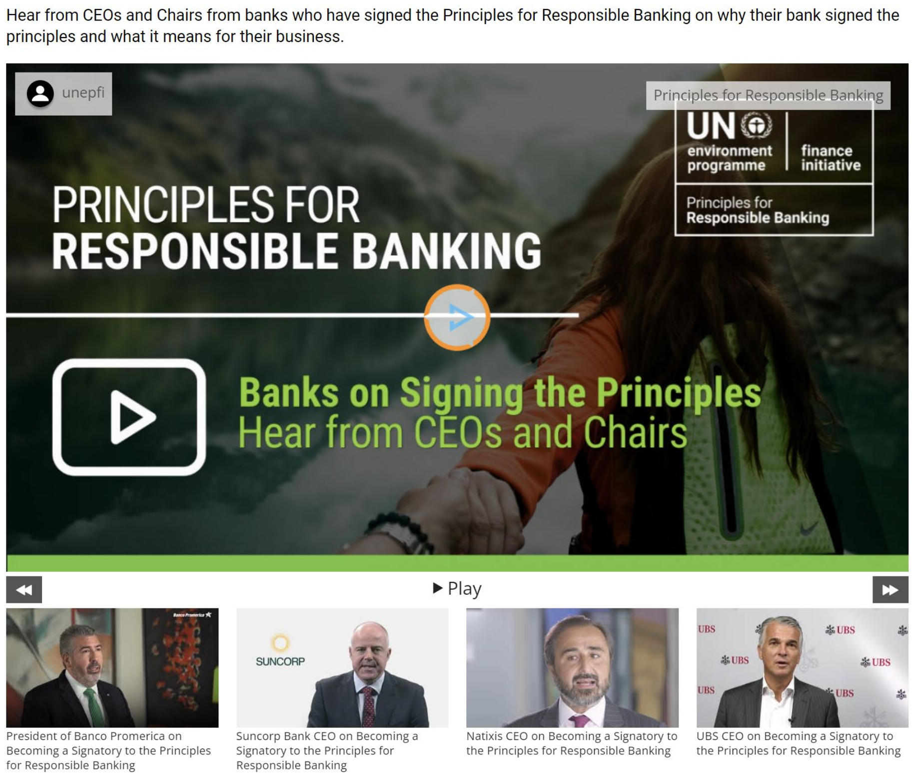
Preview of the Results

- Main Results
  - Banks with higher deception scores exhibit worse ESG performance during the post-commitment period
- The predicting power of the deception score is higher when:
  - The pressure to appear ESG-compliant is greater; CEOs do not have prior sales experience;
  - CEOs are more likely to have knowledge of their real ESG commitment levels
  - Video is longer or of higher quality
- Visual features, especially the eye area features, drive the detecting power

Setting

- Our setting: Commitment video disclosure of PRB banks
  - CEOs and Chairs from banks who have signed the Principles for Responsible Banking (PRB) talk about why their banks sign the principles and what it means for their business.

- Video-based deception detection to identify ES
  - When making ESG commitment, CEOs often already discussion and budgeting of whether their statement
  - Truth tellers differ from liars in verbal and nonverbal



Deception Detection

- Truth tellers differ from liars in verbal and nonverbal cues
  - Emotion and cognitive theories (e.g., Ekman 1985)
  - Self-Presentation theory (e.g., DePaulo 1992)

- Humans barely outperform chance
  - Overweighting of some cues based on their own lying experiences that are of no statistical significance
  - Lack of ability to identify micromomentary expressions as they are brief and fleeting

- The emerging of the Automatic Deception Detection (ADD) literature
  - machine learning that can analyze videos frame by frame, identify previously undiscernible features such as micro expressions, and weigh various features objectively

Automatically Calculate Deception Scores for the Bank CEO Video Data

Need to train a model by ourselves

- CEO video

Public Available

- Extract features

Open-Source Packages Available

- Deception Detection Model

No prior studies publish their deception classification model.

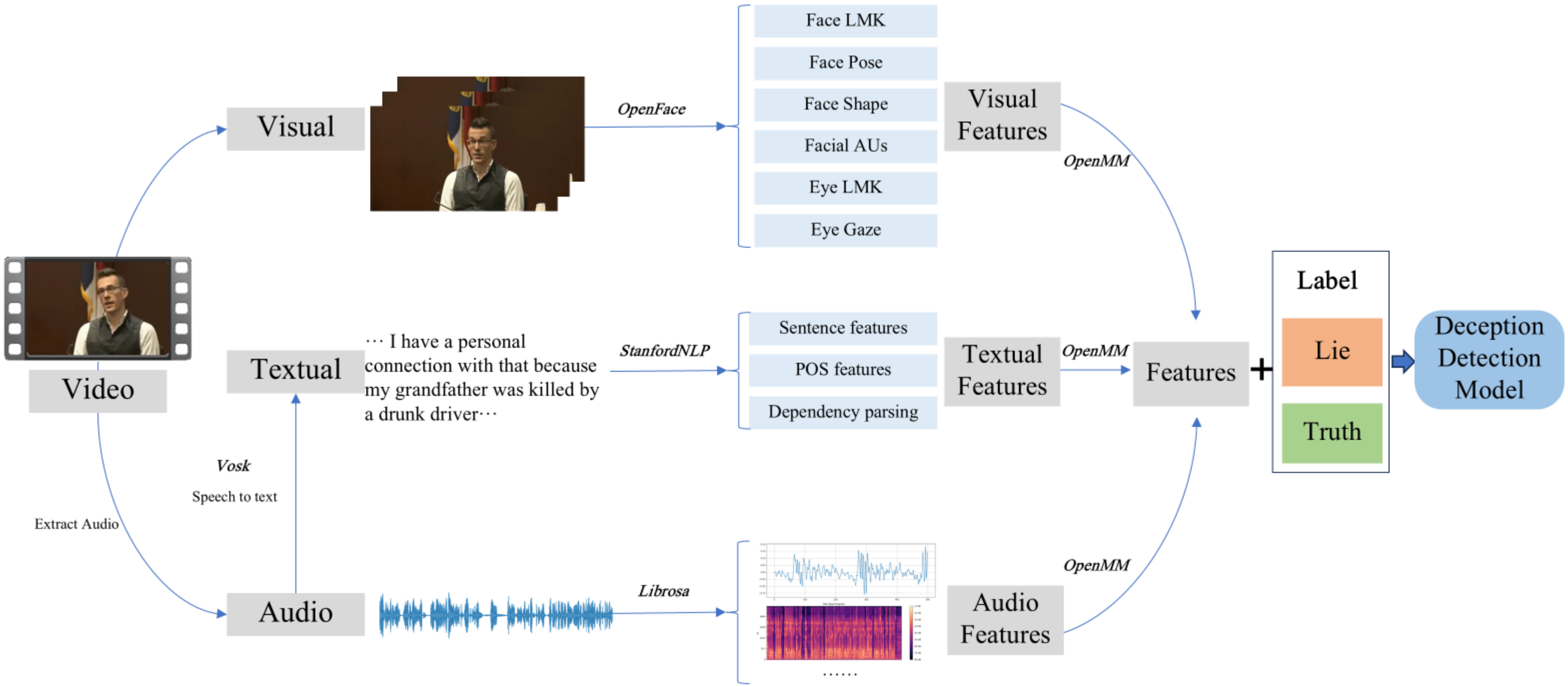
Deception Score for each CEO video

光华管理学院
Guanghua School of Management

Train the deception model

- Need a training data
  - With robust truthful/deceptive labels
  - Ideally in real-life context and for the CEOs
- No such data exists
  - The best alternative is Real-Life Trial (RLT) data
    - The RLT data contains 121 labeled individual videos of real deception and truth during public court trials (e.g., in a real-life scenario).
    - Other data are mock experiments or simulated lying experiments
  - We train a RF classifier on RLT data using various video features.

Training Process

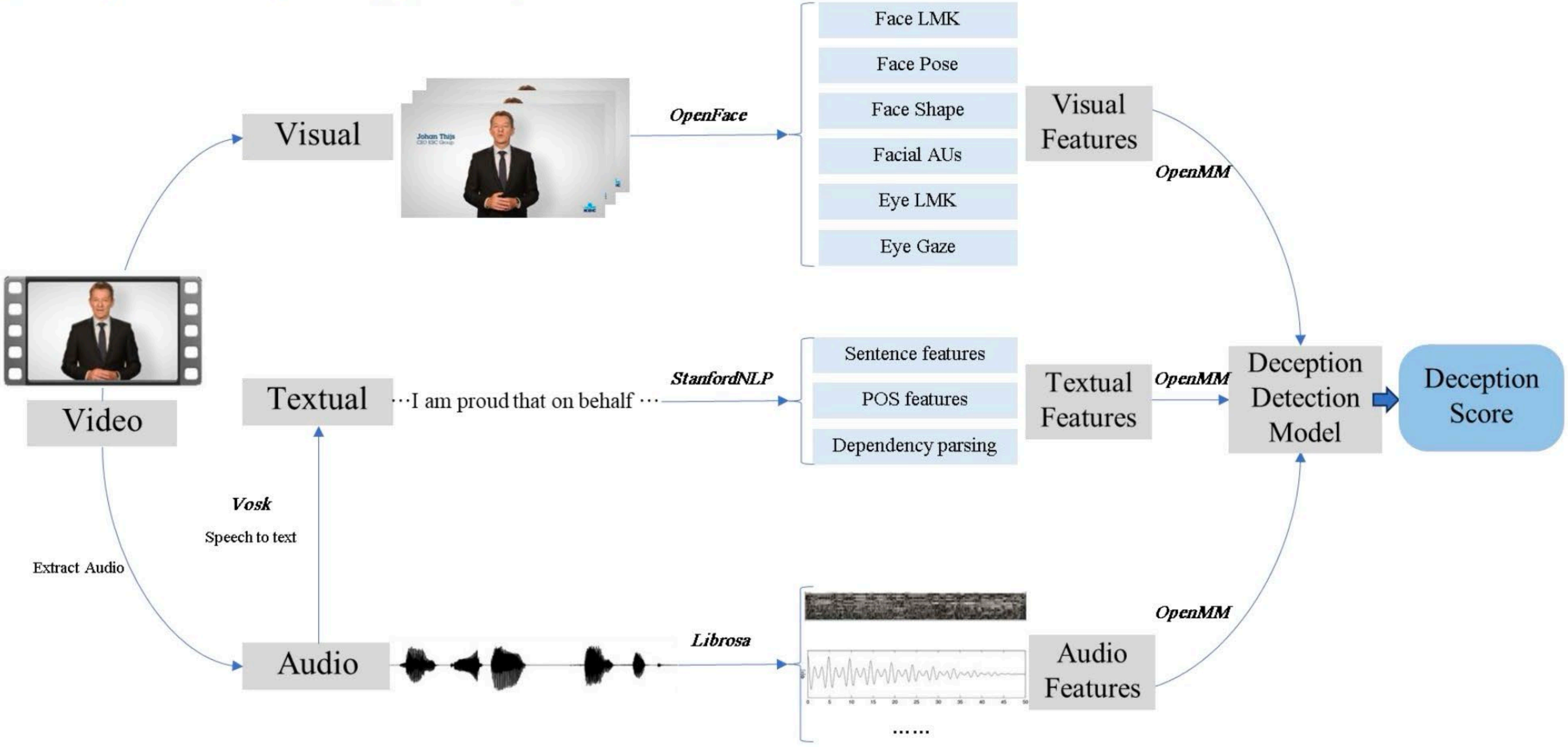


RLT dataset analysis and deception detection model training

Apple the Deception Model to CEO Videos

Figure OA1. Illustration of data structure and processing procedure

Figure OA1 presents the data processing pipeline using a real PRB video as an illustrative example.



Research Design

- In our main test, we investigate whether video-based deception scores can predict committed banks' future ESG alignment, reflected in their borrowers' ESG performance:

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

- Borrowers' Performance:
  - Number of negative ESG incidents from RepRisks; ESG ratings; Emissions
- We run the regressions using the Poison model as suggested by Cohn et al. (2022)
- Fixed effects: Bank FE; Borrower FE; Borrower Country×Year FE; Borrower Industry×Year FE

Sample Construction

- PRB banks' CEO commitment videos from UNEP FI's YouTube channel:
  - Our initial download consists of 77 videos, covering 23% of the PRB banks.
  - The videos are timely accessible when signatory banks sign the PRB commitment.
- Bank-Borrow Pairs: Dealscan data from 2016 to 2022
  - Only consider the lead banks that play a central role in establishing and maintaining relationships with borrowers and in information collection and monitoring (Sufi 2007)
  - For each loan initiation, we assume the bank-borrower relationship persists throughout the loan's lifecycle, following Dou and Xu (2021)
  - The final sample includes 9,260 bank-borrower-year observations, corresponding to 2,039 bank-borrower lending relationships.

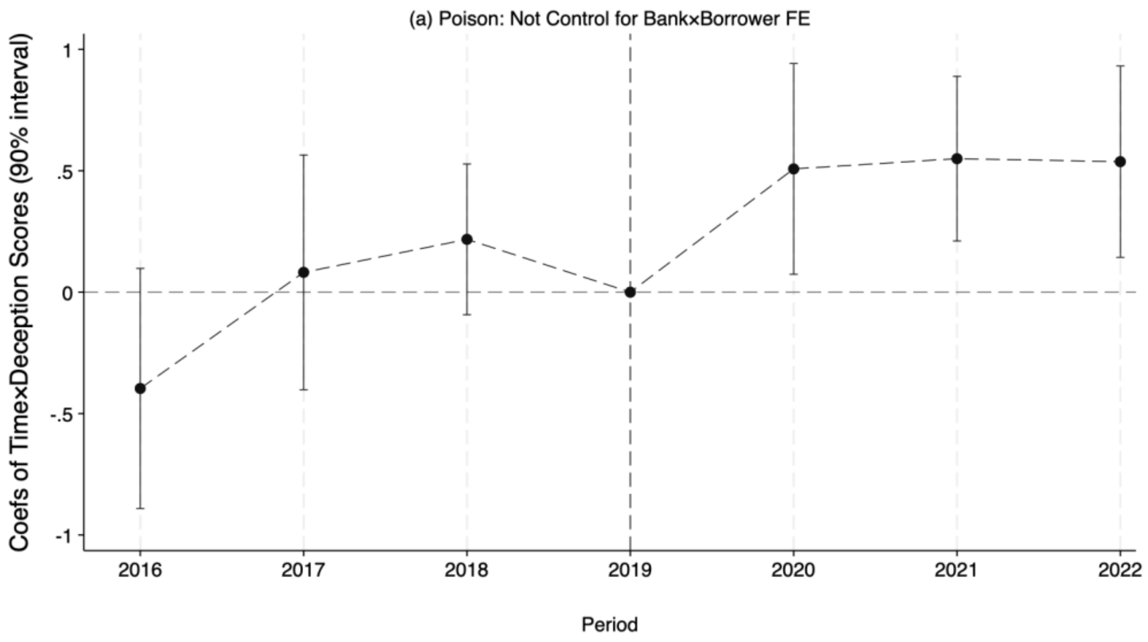
Deception Scores and the Detection of ESG Washing

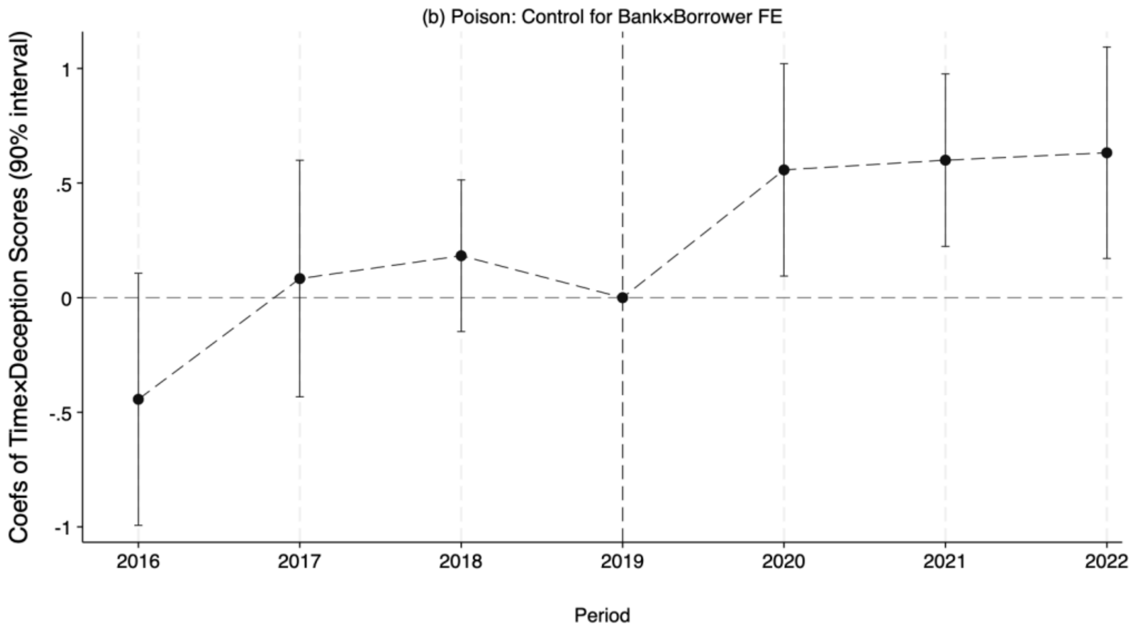
	(1)	(2)	(3)	(4)	(5)
Dep. var. =			*NegIncidents*		
Post×Deception Scores	0.544***	0.542***	0.516***	0.505***	0.573***
	(0.209)	(0.206)	(0.188)	(0.167)	(0.201)
Bank Controls	Yes	Yes	Yes	Yes	Yes
Borrower Controls	-	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	-
Borrower FE	Yes	Yes	Yes	Yes	-
Bank×Borrower FE	-	-	-	-	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	8,843
Adj. R^2	0.638	0.639	0.658	0.662	0.663

A one standard deviation increase in the deception score is associated with 5% more negative incidents

Dynamics of PRB Banks' Lending Relationships

- We include six time indicator variables to substitute *Post*: *I(year=2016)*, *I(year=2017)*, *I(year=2018)*, *I(year=2020)*, *I(year=2021)*, and *I(year=2022)*. We use *I(year=2019)* as the reference group and omit it from the regression.





Cross-Sectional Results I

- Liars' deception cues are stronger and more prevalent when they are more motivated to be believed (Ekman 1985; DePaulo et al. 2003)

Panel A. Local E&S consciousness and the detecting power

	(1)	(2)	(3)	(4)
Sample =	High E&S consciousness	Low E&S consciousness	High E&S consciousness	Low E&S consciousness
Dep. var. =	*NegIncidents*			
Post×Deception Scores	0.644**	0.295	0.634*	0.299
	(0.286)	(0.350)	(0.333)	(0.423)
Dif =	0.358***		0.334**	
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	3,702	5,418	3,620	5,169
Adj. R^2	0.672	0.657	0.675	0.657

光华管理学院
Guanghua School of Management

Cross-Sectional Results II

- Vrij et al. (2010) suggest that expressive people such as salesmen tend to be good liars as they are more confident and comfortable, experience less cognitive stress, thus leave fewer deception cues

Panel B. Salesman experience and the detecting power

	(1)	(2)	(3)	(4)
Sample =	I (Salesman Experience)=1	I (Salesman Experience)=0	I (Salesman Experience)=1	I (Salesman Experience)=0
Dep. var. =	*NegIncidents*			
Post×Deception Scores	0.174	0.446**	0.125	0.469**
	(1.081)	(0.180)	(1.170)	(0.218)
Dif =	-0.272**		-0.344**	
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	3,373	5,743	3,278	5,505
Adj. R^2	0.652	0.670	0.656	0.670

光华管理学院
Guanghua School of Management

Cross-Sectional Results III

- For banks with higher information quality, the CEOs have better knowledge of their real ESG commitment levels, pronouncing the behavioral differences between ESG washers and others

Panel C. Internal information quality and the detecting power

	(1)	(2)	(3)	(4)
Sample =	High EAspeed	Low EAspeed	High EAspeed	Low EAspeed
Dep. var. =	*NegIncidents*			
Post×Deception Scores	0.632*	0.188	0.743*	0.293
	(0.349)	(0.253)	(0.411)	(0.294)
Dif =	0.444**		0.449**	
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	3,796	4,802	3,686	4,619
Adj. R^2	0.650	0.674	0.655	0.672

光华管理学院
Guanghua School of Management

Video Data Quality and the Detecting Power

- The longer the CEO video is, the greater detecting power it likely possesses.

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	1.251***	0.366*	1.338**	0.375*
	(0.481)	(0.189)	(0.542)	(0.226)
Dif =	0.885***		0.963***	
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	3,591	5,513	3,474	5,314
Adj. R^2	0.636	0.679	0.642	0.678

光华管理学院
Guanghua School of Management

Video Data Quality and the Detecting Power

- We measure the quality of face recognition in these CEO videos using the blurriness scores provided by Face++.

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**	0.241	0.543**	0.217
	(0.202)	(0.436)	(0.246)	(0.482)
Dif =	0.188**		0.326**	
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	5,744	3,389	5,507	3,267
Adj. R^2	0.668	0.655	0.668	0.659

Visual, Audio, and Textual Features in Detecting Lies

- Visual features are the main driving force of the usefulness of our deceptions score

Panel A. Randomly reshuffling visual, audio, and textual features

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var. =			*NegIncidents*			
Post × Deception_RandomVisual	0.027	0.048				
	(0.094)	(0.113)				
Post × Deception_RandomAudio			0.486***	0.542***		
			(0.167)	(0.200)		
Post × Deception_RandomTextual					0.475***	0.530***
					(0.170)	(0.204)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	-	Yes	-	Yes	-
Borrower FE	Yes	-	Yes	-	Yes	-
Bank×Borrower FE	-	Yes	-	Yes	-	Yes
Country×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Bootstrap p-value: *Post×Deception_Random ≥ Post×Deception Scores*	0.014	0.032	0.290	0.226	0.164	0.176
N	9,260	8,843	9,260	8,843	9,260	8,843
Adj. R^2	0.662	0.663	0.662	0.663	0.662	0.663

Different Categories of Visual Features

- We categorize the visual features of videos into six groups: facial landmark (LMK), face pose, face shape, facial AUs, eye landmark (LMK), and eye gaze.

Panel B. Randomly reshuffling different categories of visual features

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var. =			*NegIncidents*			
Post ×Deception_RandomEyeGaze	0.559***					
	(0.180)					
Post ×Deception_RandomEyeLMK		0.149				
		(0.172)				
Post ×Deception_RandomFacePose			0.508***			
			(0.169)			
Post ×Deception_RandomFaceLMK				0.540***		
				(0.173)		
Post ×Deception_RandomFaceShape					0.570***	
					(0.174)	
Post ×Deception_RandomFacialAU						0.484***
						(0.168)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes
Borrower FE	Yes	Yes	Yes	Yes	Yes	Yes
Country×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Bootstrap p-value: *Post×Deception_Random ≥ Post×Deception Scores*	0.938	0.008	0.660	0.654	0.948	0.128
N	9,260	9,260	9,260	9,260	9,260	9,260
Adj. R^2	0.662	0.662	0.661	0.662	0.662	0.662

光华管理学院
Guanghua School of Management

Robustness Tests

Panel A. Controlling for banks' ESG ratings, video-based persuasion, lying words and hemifacial asymmetry of expressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var. =				*NegIncidents*				
Post×Deception Scores	0.575***	0.664***	0.481***	0.535***	0.539***	0.621***	0.521***	0.598***
	(0.180)	(0.218)	(0.174)	(0.207)	(0.172)	(0.207)	(0.170)	(0.203)
Post×Avail_BankESGratings	-0.002	-0.003						
	(0.002)	(0.003)						
Post×I(Missing_BankESGratings)	-0.123	-0.133						
	(0.128)	(0.157)						
Post×Persuasiveness_PCA			-0.006	-0.010				
			(0.013)	(0.015)				
Post×LieWords					-0.756	-1.049		
					(0.857)	(1.001)		
Post×HFAsy							-0.153	-0.255
							(0.203)	(0.250)
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	8,843	9,260	8,843	9,260	8,843	9,260	8,843
Adj. R^2	0.662	0.663	0.662	0.663	0.662	0.663	0.662	0.663

Alternative Specifications

- First, we construct our video-based deception scores using an alternative machine learning model - gradient boosted decision trees (GBDT).
- Second, we use quantile groups of deception scores than the continuous score
- Finally, we examine whether our results are driven by specific dimensions of negative ESG incidents. RepRisk categorizes negative incidents into environmental, social, governance, and cross-cutting issues.
  - The video-based deception scores of PRB banks are associated with an increase in their borrowers' negative incidents across all four dimensions

Conclusion

- In this paper, we propose a video-based deception score as an ex ante measure of banks' potential ESG washing using banks' ESG commitment video disclosure.

- Contribution and Implications:
  - It helps ESG-focused investors make better allocation of the ESG capital under their management
  - Beyond communicating ESG commitment, more firms are adopting videos to convey important corporate information (e.g., The Times 2024) and more analysts use video presentations to achieve greater and faster reach to their investor customers. Our method can help users better assess truthfulness of video disclosure and inform their decision making.

Thank you

Danqi.hu@gsm.pku.edu.cn