Persuading Investors: A Video-Based Study

Hu, Allen, and Song Ma. Working paper

解读者：屠雪永

2021.04.01

Outline

- Introduction
- Data and Setting
- Method: Processing **Video Data** with ML Algorithms
- Empirical Analysis
- **Experiment**: (Inaccurate) Beliefs vs. Taste?
- Conclusion

1. Introduction-- Background

- It is widely speculated that the delivery of a persuasion matters for the final outcome, example:
  - Sales agents achieve different results selling the same product using the same standard pitch;
  - Researchers of the same team convince peers to a different level when presenting the same paper using the same slides.

- Yet there is little evidence on how much and why the delivery features matter, especially in a real-world investment setting.

1. Introduction-- Content

- Topic: venture investment decisions after startup pitches.
- Our empirical method collects **full pitch videos** as data inputs and exploits machine learning (ML) algorithms to quantify features--facial expression, tone of voice, or diction of speech .
- We also collect the **investment decisions** made by the investors on each of the pitching startups and their long-term performance conditional on obtaining funding.

1. Introduction-- Question

- Startup teams that delivery well in their pitches , are more likely to obtain funding ?

Yes

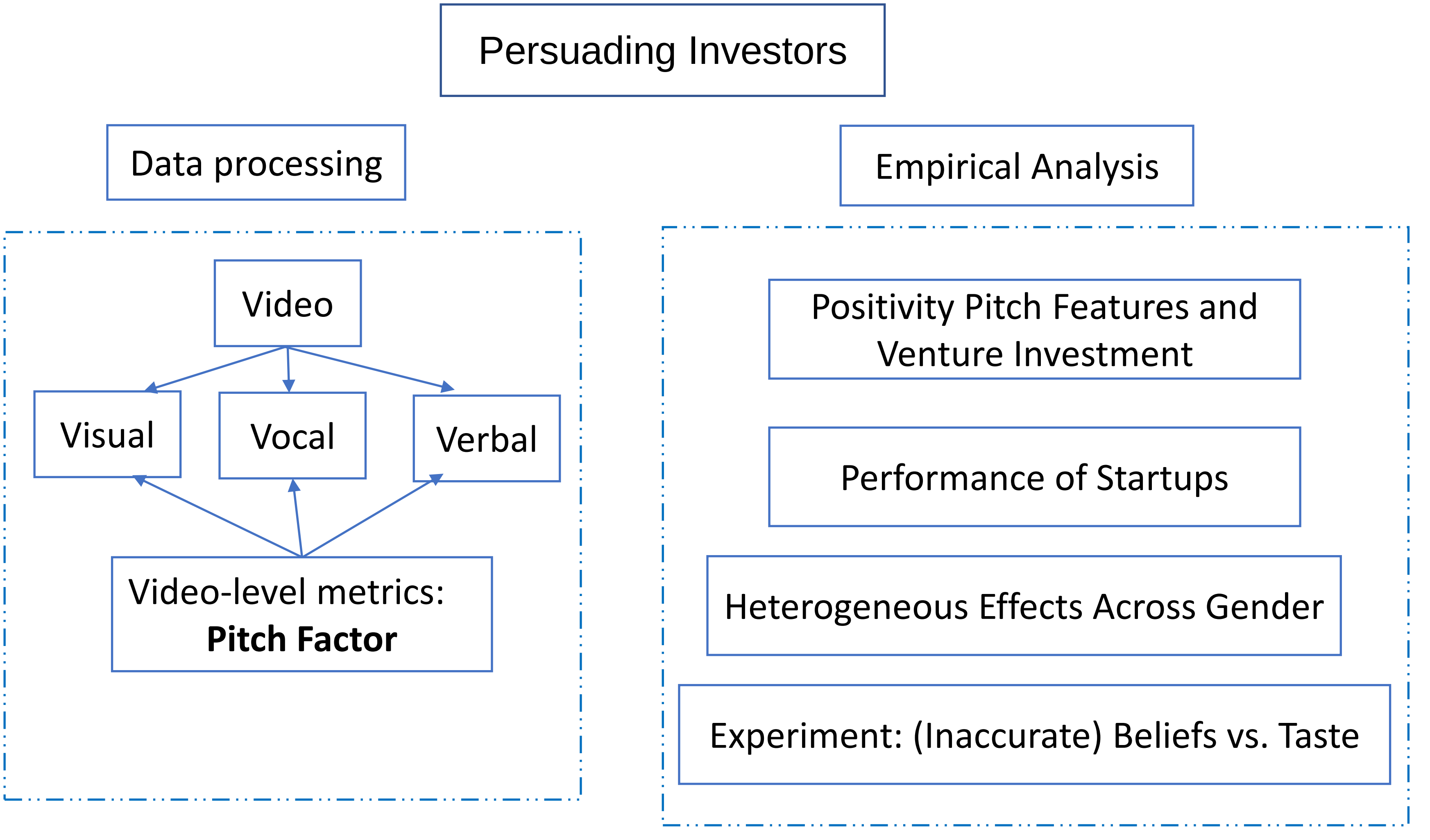
- Invested companies with higher levels of pitch positivity would likely perform better than those with poorer pitch features?

No

- The economic mechanisms through which the non-content delivery features in persuasion influence investors' decisions ?

Belief and Preference

1. Introduction-- Framework



1. Introduction-- Contribution

- 1. The key conceptual contribution of this paper is the focus on the **delivery features** in interpersonal persuasions.
- 2.Our paper provides the first systematic exploration of **unstructured video data** in economic research.
- 3.Our paper presents a unique setting in which interpersonal persuasion is particularly important for **investment decisions**.

2. Data and Setting

- Data: Time period:2010 to 2019.
- 1. Entrepreneurs' pitch videos for accelerator applications: 1139 videos crawled form YouTube and Vimeo
  - These videos are typically one- to three-minute long, and they present the founder(s) introducing the team and describing the business idea.
- 2. Startups' company-level information: Crunchbase and PitchBook.
  - Startup-level variables include the year of founding, location, operating status, total funding round and amount, number of investors, and number of employees.
- 3.Team-level information:
  - compile a list of founders for each startup company using Crunchbase, PitchBook, and video content. Each startup team member, we use **LinkedIn** to extract the five most recent education experiences and the ten most recent work experiences.

2. Data and Setting

Table 1. Summary Statistics of Videos and Startups

Panel C: Summary Statistics of Teams

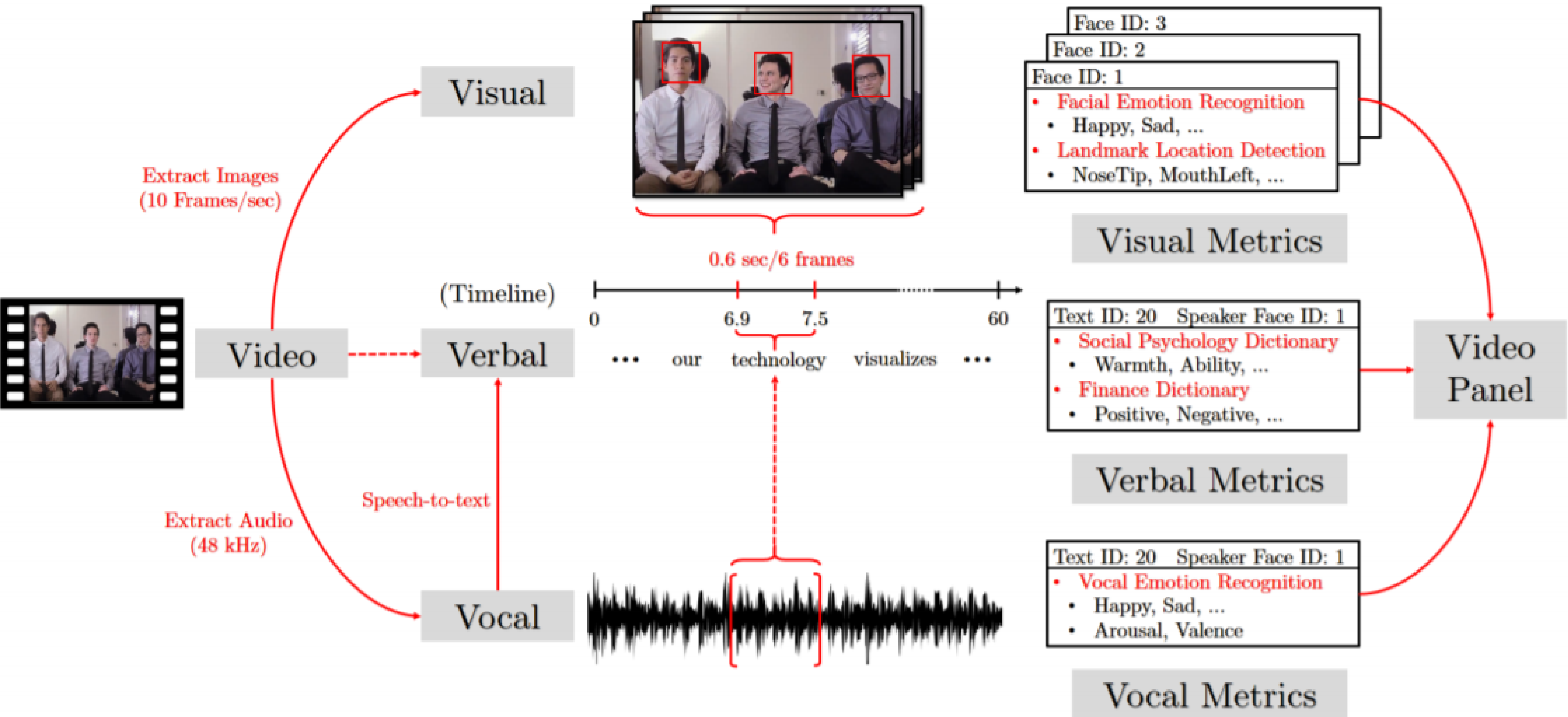
	N	Mean	STD	25%	50%	75%
Number of People	1,139	1.74	0.84	1.00	2.00	2.00
Single-Member	1,139	0.46	0.50	0.00	0.00	1.00
Multi-Member	1,139	0.54	0.50	0.00	1.00	1.00
Men-Only	1,139	0.49	0.50	0.00	0.00	1.00
Women-Only	1,139	0.27	0.45	0.00	0.00	1.00
Mixed Gender	1,139	0.24	0.43	0.00	0.00	0.00
Prior Senior Position	1,139	0.47	0.50	0.00	0.00	1.00
Prior Startup Experience	1,139	0.30	0.46	0.00	0.00	1.00
Elite University	1,139	0.06	0.24	0.00	0.00	0.00
Master's Degree	1,139	0.19	0.40	0.00	0.00	0.00
PhD Degree	1,139	0.03	0.17	0.00	0.00	0.00

3. Method: Processing Video Data with ML Algorithms

- **A. Video as Data: Key Properties**
- Video data are information intensive.
    - One second of a high-definition video, in terms of size, is equivalent to over 2,000 pages of text
- Video data are unstructured and high-dimensional, making them more complicated to process relative to other data formats
    - Consider a one-minute video with a resolution of 1280×720 (720p) and two 48 kHz audio channels.
- Video data have a low signal-to-noise ratio (SNR) and low economic interpretability
    - background noise, furniture in the video

3. Method: Processing Video Data with ML Algorithms

Figure 4. Data Structure and Processing Procedure



3. Method: Processing Video Data with ML Algorithms

C.1. Visual: ten images per second

- Identify human faces embedded in each video image using face detection Algorithms: **Face++ platform**

- First, the algorithm detects locations of **facial landmarks**:detect facial movements, such as smiles, eye blinks
- This information enters **emotion recognition algorithms**:
  - Positive : happiness
  - Negative: sadness, anger, fear, disgust
  - Neutral(exclude）
- We obtain a face **beauty** measure and some demographic characteristics of individuals, including gender and predicted age.

3. Method: Processing Video Data with ML Algorithms

(a) Example of High-Positivity Visual Features

(b) Example of Low-Positivity Visual Features

3. Method: Processing Video Data with ML Algorithms

- **C.2. Verbal (Text).**
- we extract human speech from audio data using the **speech-to-text** conversion API provided by **Google Cloud**
- These transcripts include a list of **words**, time stamps (onsets,offsets, and durations) of these words, and punctuation
  - Loughran-McDonald Master Dictionary (LM):financial text analysis and provides text categories such as **negative and positive**, among others.
  - Nicolas,Bai, and Fiske(NBF,2019): social psychological traits such as **ability and warmth**

3. Method: Processing Video Data with ML Algorithms

(a) Example of High-Ability Pitch Script

Hi, I'm Vitali CEO of Fitness Lab. There are a lot of fitness apps in the world, but they all have the same problems. They offer their users random workout and diet plans. They use just do it marketing to push sales and they have low retention because people leave after three months. So we decided to fix it and make the best fitness app in the world and power it with artificial intelligence. We're going to offer our users highly personalized workout and diet plans by now. We have launched MVP. We have 50,000 installations and five thousand active users. We have received several design award including the one for the best interaction design by American Institute of graphic arts. So we have awesome design and very efficient technology. This is going to be more efficient than any living coach. We know that Y Combinator has a lot of connections with artificial intelligence businesses. That's why we're looking forward to your support. Thank you.

(b) Example of High-Warmth Pitch Script

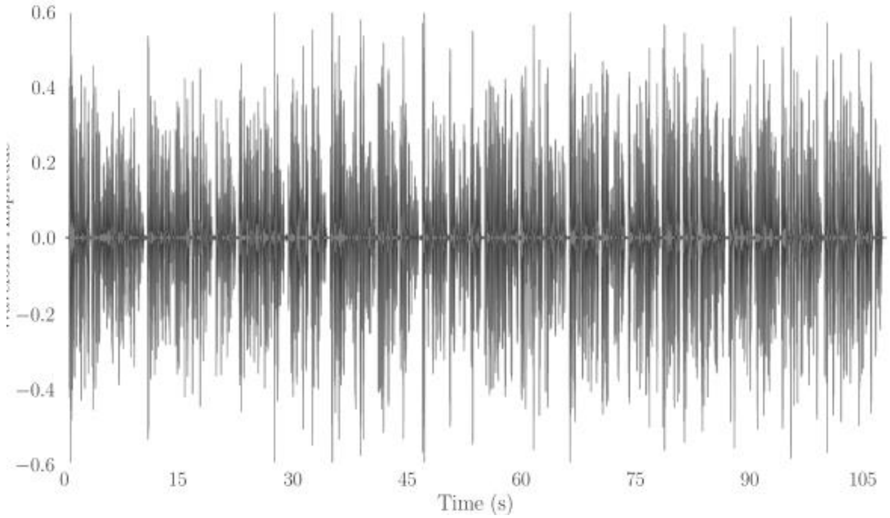
Hello, I'm Marcus and I'm Rebecca and together we're the proud founders of Fine Print Fighters LLC. We help expose small and misleading content in contracts. We help consumers make much more informed decisions during the purchasing process both pre and post purchase. We like to help the consumers gain back control of the purchasing process, and we like to create value well through our pleasing personalities as you can tell. Well, we look forward to working with angel pad, and we appreciate the opportunity in advance. Look forward to working with the staff and the rest of the constituents and hopefully be a good representation of what angel angel pad represents. So we thank you again in advance, and we look forward to speaking with you all and seeing you all soon. Thank you.

3. Method: Processing Video Data with ML Algorithms

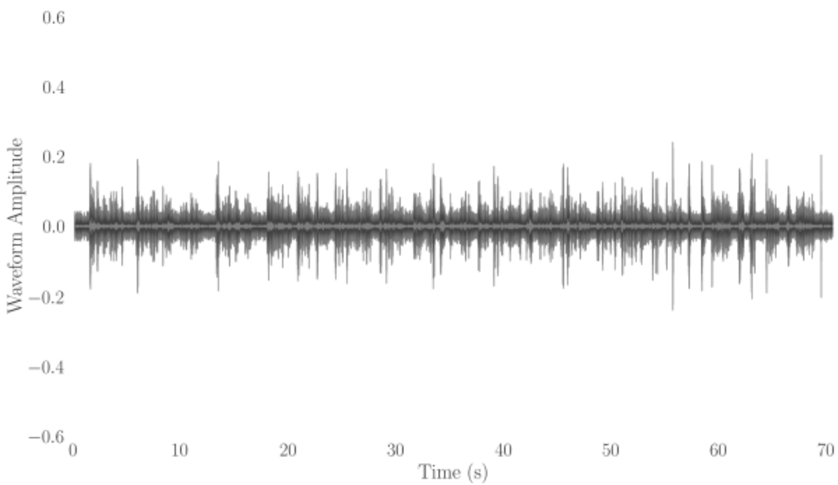
- **C.3. Vocal (Voice)**
- Audio data, essentially sound waves, code information in the audio’s dynamics and **auto-dependence structure**.
- Split each audio stream into segments by **words and sentences**
- PyAudioAnalysis:feature extraction, classification, segmentation, and application.(spectrogram, chromagram, and energy.)
- Vocal emotion analysis:
    - valance(positive affectivity) and arousal(the strength of such an attitude)
    - happy, sad, and neutral (speechemotionrecognition)

3. Method: Processing Video Data with ML Algorithms

(a) Visualized Example of High-Arousal Pitch



(b) Visaulized Example of Low-Arousal Pitch



3. Method: Processing Video Data with ML Algorithms

D. Measurement Aggregation

Table 2. Summary Statistics of Pitching Behavior Metrics

Panel A: Summary Statistics of Unstandardized Features

	N	Mean	STD	25%	50%	75%	Pitch Factor Loading	Uniqueness
Visual (Facial)								
Visual-Positive	1,139	0.17	0.16	0.05	0.12	0.25	0.08	0.29
Visual-Negative	1,139	0.15	0.14	0.06	0.11	0.20	−0.14	0.27
Visual-Beauty	1,139	0.58	0.08	0.54	0.59	0.64		
Vocal (Audio)								
Vocal-Positive	1,139	0.09	0.05	0.06	0.08	0.12	0.39	0.41
Vocal-Negative	1,139	0.01	0.01	0.01	0.01	0.02	−0.30	0.43
Vocal-Arousal	1,139	0.55	0.27	0.39	0.58	0.76	0.91	0.15
Vocal-Valence	1,139	0.44	0.22	0.31	0.46	0.59	0.88	0.18
Verbal (Text)								
Verbal-Positive	1,139	0.01	0.01	0.01	0.01	0.02	0.03	0.35
Verbal-Negative	1,139	0.01	0.01	0.00	0.01	0.01	−0.14	0.42
Verbal-Warmth	1,139	0.02	0.01	0.01	0.01	0.02	0.06	0.62
Verbal-Ability	1,139	0.03	0.02	0.02	0.03	0.05	0.06	0.56

- We compute the **average proportion** of time in the pitch that a team member shows certain facial or vocal emotion.
- Verbal are calculated as the **word counts** (in each category) scaled by the total number of words in each pitch

3. Method: Processing Video Data with ML Algorithms

D. Generating a “Pitch Factor”

- We create one variable to summarize“how well” entrepreneurs deliver their pitches.

- estimate the **factor analysis** using the principal component method

A. Baseline Result: Positivity Pitch Features and Venture Investment

$$I(Invested)_{ijt} = \alpha + \beta \cdot X_{it} + \delta_j + \varepsilon_{ijt}.$$

Table 3. Features in Pitch Delivery and Investment Decisions

Dependent Var: I(*Invested*)	Logit without Controls			Logit with Startup/Team Controls		
	Marginal Effect	S.E.	Pseudo R^2	Marginal Effect	S.E.	Pseudo R^2
Pitch Factor	0.030***	(0.007)	0.193	0.026***	(0.007)	0.253
Visual (Facial)						
Visual-Positive	0.015***	(0.005)	0.178	0.012**	(0.006)	0.240
Visual-Negative	−0.027***	(0.007)	0.187	−0.029***	(0.007)	0.253
Visual-Beauty	0.015**	(0.006)	0.178	0.015**	(0.007)	0.242
Vocal (Audio)						
Vocal-Positive	0.009**	(0.005)	0.174	0.011*	(0.006)	0.239
Vocal-Negative	−0.045***	(0.016)	0.183	−0.047***	(0.017)	0.248
Vocal-Arousal	0.023***	(0.009)	0.184	0.019**	(0.008)	0.245
Vocal-Valence	0.023***	(0.006)	0.185	0.020***	(0.007)	0.246
Verbal (Text)						
Verbal-Positive	−0.010	(0.009)	0.174	−0.011	(0.009)	0.239
Verbal-Negative	−0.026***	(0.007)	0.186	−0.022***	(0.008)	0.246
Verbal-Warmth	0.026***	(0.008)	0.190	0.028***	(0.008)	0.256
Verbal-Ability	−0.049***	(0.009)	0.243	−0.043***	(0.007)	0.298

A.1. Sample Selection of Videos.

Table 4. Sample Selection of Available Videos

	(1)	(2)	(3)	(4)	(5)	(6)
	Video Selected Out = 1					
Pitch Factor	0.006	0.015				
	(0.022)	(0.023)				
I(*Invested*)			−0.042	−0.044		
			(0.183)	(0.172)		
VC Invested					−0.011	−0.034
					(0.064)	(0.054)
Observations	527	527	527	527	527	527
Pseudo R^2	0.000	0.047	0.000	0.046	0.000	0.046
Startup/Team Controls		Y		Y		Y
Accelerator FE		Y		Y		Y

- 527 videos were uploaded between 2018 and July 2019.
- By the end of March 2020, 126 videos, or 23.9percent ,were selected out" (made private, unlisted, or completely removed)

A.2. The Value of the Video-Based Method

Table 5. Measure Construction—Full Video and Full Channels

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Var: I(*Invested*)					
	First Slice		Random Slice		Individual Channels	
Pitch Factor		0.026***		0.035**		0.064**
		(0.008)		(0.014)		(0.027)
Pitch Factor (First Slice)	0.015*	0.001				
	(0.008)	(0.008)				
Pitch Factor (Random Slice)			0.018***	−0.011		
			(0.005)	(0.013)		
Vocal Factor					0.023***	−0.040
					(0.007)	(0.028)
Visual Factor					0.025***	0.019***
					(0.006)	(0.006)
Verbal Factor					0.000	−0.009
					(0.008)	(0.009)
Observations	1,139	1,139	1,139	1,139	1,139	1,139
Pseudo R^2	0.241	0.253	0.243	0.254	0.263	0.269
Startup/Team Controls	Y	Y	Y	Y	Y	Y
Accelerator FE	Y	Y	Y	Y	Y	Y

- Slice: one second
- pitch factor provides a more comprehensive informational

C. Performance of Startups

Table 7. Long-Term Performance of Startups and Features in Pitches

	(1) Employment	(2) Raised VC	(3) VC Amount	(4) Startup Alive
Pitch Factor	−0.166**	−0.089***	−0.168*	−0.043**
	(0.050)	(0.018)	(0.086)	(0.021)
Observations	150	132	132	174
(Pseudo) R^2	0.267	0.257	0.306	0.290
Age Controls	Y	Y	Y	Y
Startup/Team Controls	Y	Y	Y	Y
Accelerator FE	Y	Y	Y	Y
Region FE	Y	Y	Y	Y

$$Performance = \alpha + \beta \cdot X + \delta_{FE} + \varepsilon.$$

Startups with a high Pitch Factor underperform in the long run.

D. Heterogeneous Effects Across Gender

Table 8. Gender Differences in the Pitch-Investment Relation

	(1)	(2)	(3)	(4)
	Dependent Var: I(*Invested*)			
	Single-Gender Teams			Mixed-Gender Teams
	Men	Women	Pooled	Pooled
Pitch Factor (Men)	0.018**		0.018**	$\boxed{0.048^{*}}$
	(0.008)		(0.008)	(0.026)
Pitch Factor (Women)		$\boxed{0.170^{***}}$	0.077**	0.019
		(0.051)	(0.031)	(0.042)
p-value of Men vs. Women Test			$\boxed{0.079^{*}}$	0.661
Observations	559	310	869	270
Pseudo R2	0.194	0.334	0.217	0.653
Startup/Team Controls	Y	Y	Y	Y
Accelerator FE	Y	Y	Y	Y
Industry FE	Y	Y	Y	

Investment decisions on woman-only startups are significantly more sensitive to the performance in the pitch, women are ignored in the pitches when they co-present with a man

4. Experiment: (Inaccurate) Beliefs vs. Taste?

explore the economic mechanisms through which pitch delivery affects persuasion effectiveness: **taste-based models** and **inaccurate beliefs**.

$$I_{ij} = \mathbb{1}_{\{U_{ij} \geq \bar{U}\}}, \quad \text{where } U(\mu_{ij}, \sigma_{ij}, \theta_i) \equiv \gamma_\mu \mu_{ij} + \gamma_\sigma \sigma_{ij} + \kappa \theta_i. \tag{5}$$

$$\mu_{ij} = \lambda_\mu Q_i + \psi_\mu \theta_i, \tag{6a}$$

$$\sigma_{ij} = \lambda_\sigma Q_i + \psi_\sigma \theta_i. \tag{6b}$$

$$I(Invested)_{ijt} = \alpha + \underbrace{\beta}_{\kappa + \psi_\mu \gamma_\mu + \psi_\sigma \gamma_\sigma} \cdot X_{it} + \delta_j + \varepsilon_{ijt}. \tag{1}$$

Scenario	$\psi_{\mu,\sigma}$	κ	Beliefs Channel	Taste Channel	Decompose β in Eq. (1)
1	$\neq 0$	$= 0$	✓	✗	$\beta = \psi_\mu \gamma_\mu + \psi_\sigma \gamma_\sigma$
2	$= 0$	$\neq 0$	✗	✓	$\beta = \kappa$
3	$\neq 0$	$\neq 0$	✓	✓	$\beta = \kappa + \psi_\mu \gamma_\mu + \psi_\sigma \gamma_\sigma$

B. Experiment Design

- Master's students from the Yale School of Management :102 subjects to act as venture investors, and allocated 10 pitch videos
- The video pool consists of 62 videos
- After viewing each video i, the subject is asked to answer questions around three main themes:
  - (1) whether she/he would invest in company i, denoted as $I_{\boldsymbol{ij}}$;
  - (2) her/his expectation of the company's success probability, $\mu_{\boldsymbol{ij}}$, measured between 0 and 100%;
  - (3) her/his confidence level on her/his decision and expectation, $\sigma_{\boldsymbol{ij}}$, measured on a scale of 1 to 5.

C.1. Interaction Features and (Inaccurate) Beliefs

Table 9. Experiment Results: Pitch Factor and Investor Beliefs

	(1)	(2)	(3)	(4)	(5)
	P(*alive*|*invested*)		P(*success*|*invested*)		*alive*|*invested*
	μ	σ	μ	σ	Realized
Pitch Factor (θ)	0.020**	−0.020	0.016**	−0.030	−0.117**
	(0.009)	(0.027)	(0.007)	(0.028)	(0.053)
Observations	952	952	952	952	495
R^2	0.569	0.545	0.565	0.519	0.673
Startup/Team Controls	Y	Y	Y	Y	Y
Subject FE	Y	Y	Y	Y	Y

These results confirm the existence of a channel through beliefs.
actual survival probability negatively correlates with the Pitch Factor conditional on investment, in contrast to subjects' beliefs that Pitch Factor would positively predict survival. The miscalibration of beliefs has a magnitude of 0.137 (= 0.020 − (−0.117)).

C.2. Decomposing Inaccurate Beliefs and Preferences.

Table 10. Experiment Results: Inaccurate Beliefs, Tastes, and Investment

	(1)	(2)	(3)	(4)
	Dependent Var: I(*Invested*)			
Pitch Factor (θ)	0.125***			$\boxed{0.067^{***}}$
	(0.037)			$\boxed{(0.022)}$
$\mu(alive\|invested)$		2.309***		2.208***
		(0.120)		(0.132)
$\sigma(alive\|invested)$			−0.171***	−0.054**
			(0.041)	(0.026)
Observations	952	952	952	952
Pseudo R^2	0.157	0.423	0.135	0.436
Startup/Team Controls	Y	Y	Y	Y
Subject FE	Y	Y	Y	Y

$$I_{ij} = \underbrace{\kappa \cdot \theta_i}_{\text{Taste}} + \underbrace{\gamma_\mu \cdot \mu_{ij} + \gamma_\sigma \cdot \sigma_{ij}}_{\text{Beliefs}} + \delta_j + \varepsilon_{ij}.$$

There exists a taste/preference channel through which the pitch features affect investment decisions.

4. Conclusion

- Firstly, We find that non-content delivery features in persuasive interactions have statistically significant and economically sizable effects on investors’ decisions.
- Secondly, These features do not seem to help investors to make better investment decisions, a bias induced by those features, particularly through leading investors to form inaccurate beliefs.

5. Inspiration

- Many researches on investor sentiment can use video methods, mainly proper data acquisition are difficult.

B. Is Omitted Startup Quality Driving The Results?

$$I(Invested)_{ijt} = \alpha + \beta \cdot X_{it} + \delta_j + \varepsilon_{ijt}.$$

Table 6. Features in Pitches and Investment Decisions—Oster Test

$R^2_{max} = \min(2.2R^2_c, 1)$						
$\delta = 1$			$\delta = 2$			δ s.t. $\beta_{adj} = 0$
β_{adj}	Identified Set	Reject Null?	β_{adj}	Identified Set	Reject Null?	
0.023	[0.023,0.026]	Y	0.019	[0.019,0.026]	Y	6.157
$R^2_{max} = min(3R^2_c, 1)$						
$\delta = 1$			$\delta = 2$			δ s.t. $\beta_{adj} = 0$
β_{adj}	Identified Set	Reject Null?	β_{adj}	Identified Set	Reject Null?	
0.021	[0.021,0.026]	Y	0.014	[0.014,0.026]	Y	3.752
$\boxed{R^2_{max} = 1}$						
$\delta = 1$			$\boxed{\delta = 2}$			δ s.t. $\beta_{adj} = 0$
β_{adj}	Identified Set	Reject Null?	β_{adj}	Identified Set	Reject Null?	
0.012	[0.012,0.026]	Y	−0.007	[−0.007,0.026]	$\boxed{\text{N}}$	1.678

$$\beta_{adj} \approx \beta_c - \delta \frac{(\beta_u - \beta_c)(R^2_{max} - R^2_c)}{R^2_c - R^2_u}.$$

the null that $\beta = 0$.

Oster (2019) suggests a test for omitted variable bias that uses the information contained in the change in coefficient and the change in R2 when moving from uncontrolled to controlled regression