“Let me get back to you” – A machine learning approach to measuring non-answers

Fabian Woebbeking
Leibniz Institute for Economic Research Halle (IWH),
Martin Luther University Halle-Wittenberg (MLU)

joint work with Andreas Barth and Sasan Mansouri

June 14, 2023

Motivation

- Information asymmetry is a key friction (in economics)
- Transparency through disclosures is essential for efficient markets (see e.g. SEC's plain English initiative.)

- **Questions & answers (Q&A)** are the most targeted form of information exchange
- Teach a computer to detect whether a question has been answered
- Empirical evidence for the economic relevance of non-answers
- Understanding beyond prediction (explainable AI)

Anatomy of an answer

Q

Answer

Non-Answer

Cause:
Intention to withold information
Inability to provide information

Symptoms:
Rejecting, e.g., "I won't answer that."
Blathering, i.e., "Beating around the bush"
Postponing, e.g., "Let me get back to you."
...

Anatomy of an answer cont'd

"Excuse me. Next. Boring, bonehead questions are not cool. Next?"
— Elon Musk, Tesla earnings call, May 2018

"Senator, [. . .] I can certainly have my team get back to you on any specifics there that I don't know, sitting here today."
— Mark Zuckerberg, US Senate hearing , April 2018

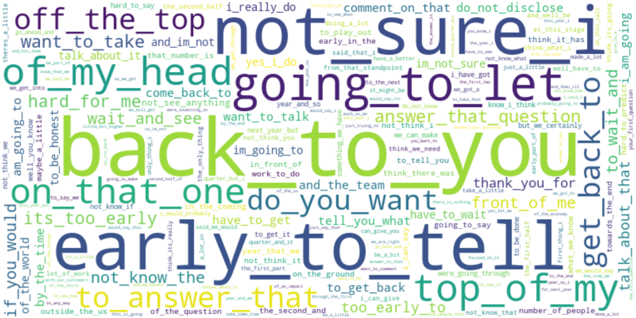
Research objectives

1. Develop a measure for non-answers
   - Supervised machine learning model
   - Training set with two classified symptoms: Rejecting and Blathering
   - Generate glossary with markers for non-answers

2. Show economic relevance of the measure
   - Empirical model to show the impact of not answering
   - Earnings conference calls
   - Large validation set with mapping to financial data

Results 1: the non-answer glossary

- 1,364 trigrams that indicate non-answers
- Not specific to the context of earnings conference calls
- A machine readable version is available at **econlinguistics.org**



Results 2: economic relevance

We show that the glossary has economic relevance

- Markets react to non-answers:
  - Lower stock returns after an earnings call
  - Higher uncertainty (option implied volatility)
- Within an earnings call, our non-answer metric is higher for managements' responses to:
  - Follow-up questions by the same analyst
  - More negatively toned questions
  - Forward looking questions
- Analysts are more likely to negatively adjust their EPS estimate when faced with non-answers

Overview

Multinomial Inverse Regression (MNIR)

- Supervised model developed by Taddy (2013)
- Requires labeled text examples, i.e. the training set
- A text is a combination of tokens (here: trigrams)

	"back_to_you"	"i_dont_know"	"gain_on_sale"	...
Text_1	1	0	0	...
Text_2	1	1	0	...
Text_3	0	0	1	...
...	...	...	...	...

- Here, text 1 and 2 might be labeled 'non-answer' while text 3 might be labeled 'answer'
- MNIR extracts tokens that frequently associate with a given label
- Technical details

Multinomial Inverse Regression (MNIR)

- A text is a combination of tokens (here: trigrams)
- In the universe of documents, $\mathcal{I}$, we have in total p distinct trigrams
- Document i consists of tokens counted in $\mathbf{x_i} = [x_{i1}, ..., x_{ip}]'$,
- and has observable attribute y_i (i.e. blathering or rejecting)
- A naive approach is to fit a linear regression model:

$$\mathbf{y} = \beta \mathbf{x}^\intercal + \epsilon,$$

where β contains each token's contribution to $\mathbf{y}$

$\Rightarrow$ Curse of dimensionality

Multinomial Inverse Regression (MNIR) cont'd

- Shrink dimensionality in pursuit of a parsimonious model $\rightarrow$ Lasso
- Multinomial Inverse Regression, a supervised generative model developed by Taddy (2013), i.e. a 'Gamma-Lasso' scheme
- Maximum A Posteriori (MAP) Estimation (Taddy, 2015)
- Choice of (hyper) prior parameters irrelevant (Taddy, 2015)

- Sentiment-preserving dimension reduction for text data
- Unbiased estimator and
- Computationally superior to alternatives, such as classical Lasso

Q&A data

- Text data from earnings conference calls
  - Regular intervals (quarterly)
  - Relatively standardized format (Q&A)
  - Connected data (financial market, accounting, etc.)
- Collect all earnings calls from **S&P 500** companies

- The data (training) set in numbers:
  - Time range: 2002 to 2019
  - Number of companies: 650 (42)
  - Number of earnings calls: 25,675 (1,860)
  - Number of questions: 1,195,470 (48,197)

Training set classification (symptoms)

- For response j in earnings call of company i at time t, measure:
  - Rejecting (Gow et al., 2021)

$$y_{ijt}^{\text{Rejection}} = \begin{cases} 1 & \text{, if rejection phrase} \in \text{response } j \\ 0 & \text{, else.} \end{cases}$$

  - Blathering (Barth et al., 2021)

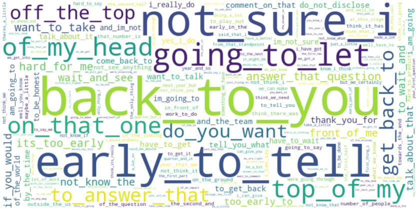
$$y_{it}^{\text{Blathering}} = 1 - \frac{\text{Finance glossary words}_{it}}{\text{Total words}_{it}}$$

- The blathering metric was developed for financial firms only
- Roughly **7%** of the earnings calls are used as training set

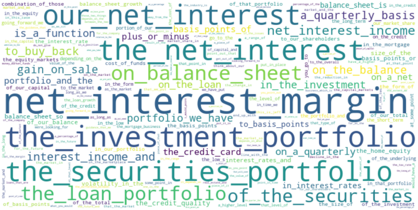
Overview

The glossary

- 1,364 tokens with positive factor load



- 1,351 tokens with negative factor load (**excluded from glossary**)



The non-answer score

- For the earnings call of company i in quarter t,
- Count glossary trigrams and total words,

$$NonAnswer_{it} = \frac{\text{Non-answer glossary tokens}_{it}}{\text{Total words}_{it}}.$$

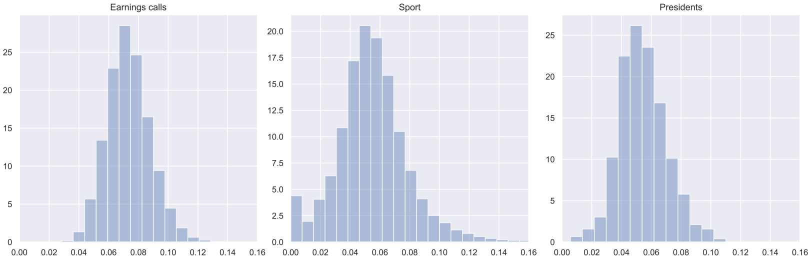
- Incorporate the information from the loadings

$$NonAnswer^{\phi}_{it} = \frac{\sum_{k=1}^{K} \phi_k \times \text{Non-answer glossary token}_{itk}}{\text{Total words}_{it}},$$

where ϕ_k is the loading associated with trigram $k \in \{1, 2, ..., K\}$.

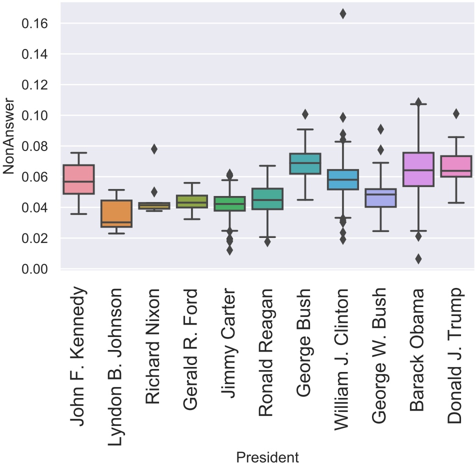
Alternative applications

- Non-answer glossary does not rely on a specific contextual domain
- Alternative sources of extensive Q&A type data, e.g.:
  - press conferences after major sport events, such as NBA basketball games and NFL football games
  - presidential interviews as collected by UCSB's Presidency Project.



Presidential interviews

- Anecdotal evidence: particularly high non-answer scores during the Clinton–Lewinsky scandal.



Overview

Market reaction

- Investors participate in the Q&A session of earnings conference calls in order to reduce uncertainty about future performance
- The theoretical asset pricing literature suggests that higher uncertainty translates to larger risk premia (Andrei and Hasler, 2014)
- Not conveying information in earnings call leads to a negative stock market reaction (Hollander et al., 2010; Zhou, 2018)

- Faced with non-answers, we expect:
  - $\Rightarrow$ Lower cumulative abnormal stock returns
  - $\Rightarrow$ Higher expected volatility (implied volatility)

How do non-answers affect stock returns?

- Baseline model specification

$$CAR_{imt} = \alpha + \beta_1 \cdot NonAnswer_{imt} + \beta' \cdot \textbf{Language Controls} + \kappa' \cdot \textbf{Firm Controls} + ... + \epsilon_{imt} \quad (1)$$

- **Language controls**: $Negativity$ (Price et al., 2012); $Uncertainty$ (Dzielinski et al., 2021), $Numbers$ (Zhou, 2018), $Complexity$ (Loughran and McDonald, 2019)
- **Firm controls**: earnings surprise, i.e. reported earnings vs analysts consensus; BTM; Tobin's Q, total assets
- Fixed effects: QuarterYear, Industry, Firm, CEO

How do non-answers affect stock returns? Cont'd

Table: Management $NonAnswer$ and abnormal returns ($CAR_{0;1}$)

	$FF3-CAR_{0;1}$			$FF5-CAR_{0;1}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$NonAnswer$	-0.089***	-0.114***	-0.125***	-0.080***	-0.104***	-0.112***
	(-3.52)	(-4.27)	(-2.92)	(-3.30)	(-3.42)	(-2.77)
$Negativity$	-0.415***	-0.421***	-0.615***	-0.361***	-0.482***	-0.567***
	(-6.53)	(-6.53)	(-7.54)	(-5.98)	(-7.29)	(-7.21)
$Uncertainty$	-0.089	-0.038	-0.098	-0.071	-0.087	-0.103
	(-1.26)	(-0.52)	(-1.08)	(-1.08)	(-1.11)	(-1.17)
$Numbers$		-0.185**	-0.273***	-0.166**	-0.213**	-0.262***
		(-2.59)	(-2.86)	(-2.52)	(-2.56)	(-2.99)
$Complexity$		0.168	0.406***	0.183*	0.226*	0.375***
		(1.64)	(2.72)	(1.97)	(1.97)	(2.69)
Observations	21191	21035	20004	21191	21191	20004
R^2	0.045	0.048	0.132	0.044	0.080	0.130
FirmControls	Yes	Yes	Yes	Yes	Yes	Yes
QuarterYear FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Implied	No	Implied	Implied
Firm FE	No	No	Yes	No	Yes	Yes
CEO FE	No	No	Yes	No	No	Yes

Economic relevance

How do non-answers affect expected volatility?

- Investigate whether non-answers affect the implied volatility, i.e. markets expectation of future volatility (uncertainty)
- Option prices react instantaneous to information
- Collect the daily implied volatility $\sigma_{i,t}$ derived from liquid at-the-money options with 90 days to maturity from OptionMetrics
- Run baseline regression with dependent variable:

$$IV_{i,t}^{-1;1} = \ln\left(\frac{\sigma_{i,t+1}}{\sigma_{i,t-1}}\right).$$

$$\Delta IV_{i,t} = \frac{\sigma_{i,t+1} - \sigma_{i,t-1}}{2} - \frac{\sigma_{i,t-60} - \sigma_{i,t-1}}{59}.$$

How do non-answers affect expected volatility? Cont'd

Table: Management $NonAnswer$ and option implied volatility ($\Delta\ IV$ and $IV_{-1;1}$)

	ΔIV				$IV_{-1;1}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$NonAnswer$	0.024***	0.027***			0.101**	0.116**		
	(2.86)	(3.17)			(2.27)	(2.60)		
$NonAnswer^{\phi}$			0.017***	0.018***			0.067**	0.070***
			(3.51)	(3.67)			(2.57)	(2.66)
$Negativity$		0.018		0.017		0.255**		0.246**
		(0.95)		(0.86)		(2.65)		(2.57)
$Uncertainty$		0.037*		0.035*		0.222**		0.218**
		(1.93)		(1.87)		(2.40)		(2.37)
$Numbers$		0.039		0.037		0.195		0.183
		(1.57)		(1.47)		(1.61)		(1.51)
$Complexity$		0.036		0.037		-0.030		-0.028
		(1.16)		(1.18)		(-0.19)		(-0.18)
Observations	20108	20108	20108	20108	20113	20113	20113	20113
R^2	0.138	0.139	0.138	0.139	0.162	0.164	0.162	0.164
FirmControls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
QuarterYear FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Economic relevance

Overview

Information contained within an earnings call

- Additional information within the Q&A
  - Question order
  - Sentiment
  - Person (analyst) asking/answering
  - Contextual domain
- Managers avoid unfavourable or pessimistic analysts (Mayew, 2008; Cohen et al., 2020)
- Follow-up questions are often more critical (Clayman, 1993)

- Research questions:
  - How analysts respond to non-answers?
  - Which questions receive non-answers?

Which questions receive non-answers?

Table: $NonAnswer$ in response to follow-up questions.

	$NonAnswer$			$NonAnswer^{\phi}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$IsFollowUp_q$	0.0016***	0.0017***	0.0021***	0.0085***	0.0088***	0.0094***
	(10.66)	(11.63)	(14.13)	(15.95)	(16.54)	(17.49)
$Negativity_q$		0.0207***	0.0194***		0.0450***	0.0432***
		(10.17)	(9.57)		(5.54)	(5.32)
$ForwardSentiment_q$			0.0304***			0.0449***
			(25.72)			(9.86)
Observations	621696	621696	621696	621696	621696	621696
R^2	0.065	0.065	0.066	0.049	0.049	0.049
EarningsCall FE	Yes	Yes	Yes	Yes	Yes	Yes

How analysts respond to non-answers?

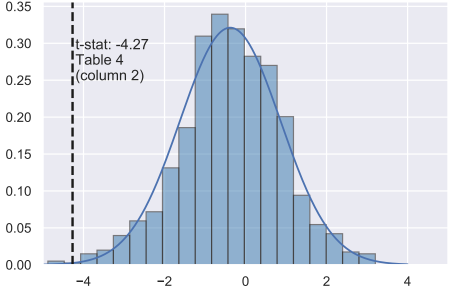
Table: Analysts EPS estimate around the call

	%Positive Revisions			
	(1)	(2)	(3)	(4)
$NonAnswer$	-0.240*	-0.167		
	(-1.71)	(-1.40)		
$NonAnswer^{\phi}$			-0.128**	-0.086*
			(-2.44)	(-1.94)
$Negativity$	-0.997***	-1.155***	-0.989***	-1.145***
	(-3.48)	(-4.47)	(-3.48)	(-4.50)
$Uncertainty$	0.098	-0.455	0.128	-0.435
	(0.31)	(-1.63)	(0.40)	(-1.56)
Observations	21129	20973	21129	20973
R^2	0.075	0.112	0.075	0.112
FirmControls	Yes	Yes	Yes	Yes
QuarterYear FE	Yes	Yes	Yes	Yes
Industry FE	No	Yes	No	Yes

Economic relevance

Robustness check: Can random glossaries produce similar results?

- Draw 1,000 times a number of 1,364 trigrams from the training set
- Repeat the baseline regression from Equation (1)
- t-statistics of β_1: (Dashed line in the figure corresponds to column (2) of the main (CAR) results table)



Conclusion

Identification of non-answers

- Using Multinomial Inverse Regression, we create a dictionary of markers for non-answers, such as ‘back to you’, ‘to be honest’, ‘a great question’, etc.

Economic relevance

- Investors place significant emphasis on information (not) conveyed in verbal disclosures (lower return, higher risk)
- Managers try to avoid to answer pessimistic analysts, as well as tougher/more critical questions

Future research

- Explore contextual domain
- Utilize more powerful (large-language) models

Thank you, for your attention!

Fabian Woebbeking
Leibniz Institute for Economic Research, Halle (IWH)
Kleine Maerkerstraße 8
D-06108 Halle (Saale)
fwg@iwh-halle.de

References I

Daniel Andrei and Michael Hasler. Investor Attention and Stock Market Volatility. *Review of Financial Studies*, 28(1):33–72, 09 2014.

Andreas Barth, Sasan Mansouri, Fabian Woebbeking, and Severin Zörgiebel. How to talk down your stock performance. Working paper, 2021.

Robert Bloomfield. Discussion of "annual report readability, current earnings, and earnings persistence". *Journal of Accounting and Economics*, 45(2-3):248–252, 2008.

Brian J Bushee, Ian D Gow, and Daniel J Taylor. Linguistic complexity in firm disclosures: Obfuscation or information? *Journal of Accounting Research*, 56(1): 85–121, 2018.

Steven Clayman. Reformulating the question: A device for answering/not answering questions in news interviews and press conferences. *Text - Interdisciplinary Journal for the Study of Discourse*, 13, 01 1993. doi: 10.1515/text.1.1993.13.2.159.

Lauren Cohen, Dong Lou, and Christopher J. Malloy. Casting conference calls. *Management Science*, 66(11):5015–5039, 2020. doi: 10.1287/mnsc.2019.3423.

References II

Michal Dzielinski, Alexander F. Wagner, and Richard J. Zeckhauser. CEO clarity. *HKS Working Paper No. RWP17-017*, 2021.

Ian D Gow, David F Larcker, and Anastasia A Zakolyukina. Non-answers during conference calls. *Journal of Accounting Research*, forthcoming, 2021.

Wayne Guay, Delphine Samuels, and Daniel Taylor. Guiding through the fog: Financial statement complexity and voluntary disclosure. *Journal of Accounting and Economics*, 62(2-3):234–269, 2016.

Robert Gunning. *The technique of clear writing*. McGraw-Hill, New York, 1952.

Elaine Henry. Are investors influenced by how earnings press releases are written? *The Journal of Business Communication*, 45(4):363–407, 2008.

Stephan Hollander, Maarten Pronk, and Erik Roelofsen. Does silence speak? an empirical analysis of disclosure choices during conference calls. *Journal of Accounting Research*, 48(3):531–563, 2010.

Feng Li. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics*, 45(2-3):221–247, 2008.

References III

Tim Loughran and Bill McDonald. When is a Liability not a Liability? *Journal of Finance*, 66(1):35 – 65, 2011. ISSN 00221082.

Tim Loughran and Bill McDonald. Measuring readability in financial disclosures. *Journal of Finance*, 69(4):1643–1671, 2014.

Tim Loughran and Bimm McDonald. Measuring firm complexity. Working paper, 2019.

Sasan Mansouri. Does firm's silence drive media's attention away? *Available at SSRN: 3809792*, 2021.

William J Mayew. Evidence of management discrimination among analysts during earnings conference calls. *Journal of Accounting Research*, 46(3):627–659, 2008.

S. McKay Price, James S. Doran, David R. Peterson, and Barbara A. Bliss. Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking & Finance*, 36(4):992 – 1011, 2012.

Matt Taddy. Multinomial inverse regression for text analysis. *Journal of the American Statistical Association*, 108(503):755–770, Sep. 2013.

References IV

Matt Taddy. Distributed multinomial regression. *The Annals of Applied Statistics*, 9 (3):1394–1414, 2015.

Dexin Zhou. Do numbers speak louder than words? Technical report, 2018.

MNIR

- Logistic link $\mathbf{q_y}$ with parameter vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\Phi}$
- Discretization of metric y measure in m_y

$$\mathbf{x_y} \sim \mathrm{MN}(\mathbf{q_y}, m_y), \quad \text{with}$$

$$\mathbf{x_y} = \sum_{i:y_i=y} \mathbf{x_i},$$

$$m_y = \sum_{i:y_i=y} m_i,$$

$$q_{yj} = \frac{exp[\alpha_j + y\phi_j]}{\sum_{k=1}^{p} exp[\alpha_k + y\phi_k]}$$

$$\text{for } \; j = 1, ..., p, y \in \mathcal{Y} \; \text{ and } \; m_i = \sum_{j=1}^{p} x_{ij}.$$

Maximum a posteriori estimation

- For each ϕ_j, Taddy (2013) use a fat-tailed and sparsity-inducing independent Laplace priors instead of a shared λ (Lasso)
- Each Laplace rate parameter λ_j is left unknown with a gamma hyperprior $Ga(s, r)$
- **Choice of (hyper) prior leads to parsimonious model**
- Maximise a posteriori probability (MAP) using the algorithm in Taddy (2013) to fit the model
- Choice of hyper prior parameters irrelevant (Taddy, 2015)
- back