Does Alternative Data Improve Financial Forecasting? The Horizon Effect

Olivier Dessaint, Thierry Foucault, and Laurent Fresard

Discussion by: Fany Declerck

Toulouse School of Economics (TSM-R, CNRS)

Autorité des Marchés Financiers

Conseil Scientifique

19/10/2021

Synopsis: research questions

- Analysts' forecast informativeness
  - Impact of alternative data on short-term informativeness
  - Impact of alternative data on long-term informativeness
  - Multitasking is costly

First comments: why it matters (1/2)

- Impact of alternative data
- Relevance of forecasts produces by sell-side financial analysts

Synopsis: data

- I/B/E/S for sell-side equity analysts
  - US financial analysts
  - Period: from 1983 to 2017
  - h: earnings report date - forecast date *(in number of days)*
  - Horizon h: from one day up to 5 years
    - Except quaterly and semi-annual earnings forecasts
    - If well-defined fiscal period and well-defined announcement dates & analyst code
  - Size of the sample:
    - h < 1 year: 4,259,465 forecasts
    - h > 2 years: 1,260,796 forecasts (if h > 4 years: 102,431 forecasts)
    - > 65 million analyst-day-horizon R^2 observations
- StockTwits (2009): diff-in-diff analysis
  - Watchlists
  - Number of “hypothetical” messages (last 30 days)

First comments: why it matters (2/2)

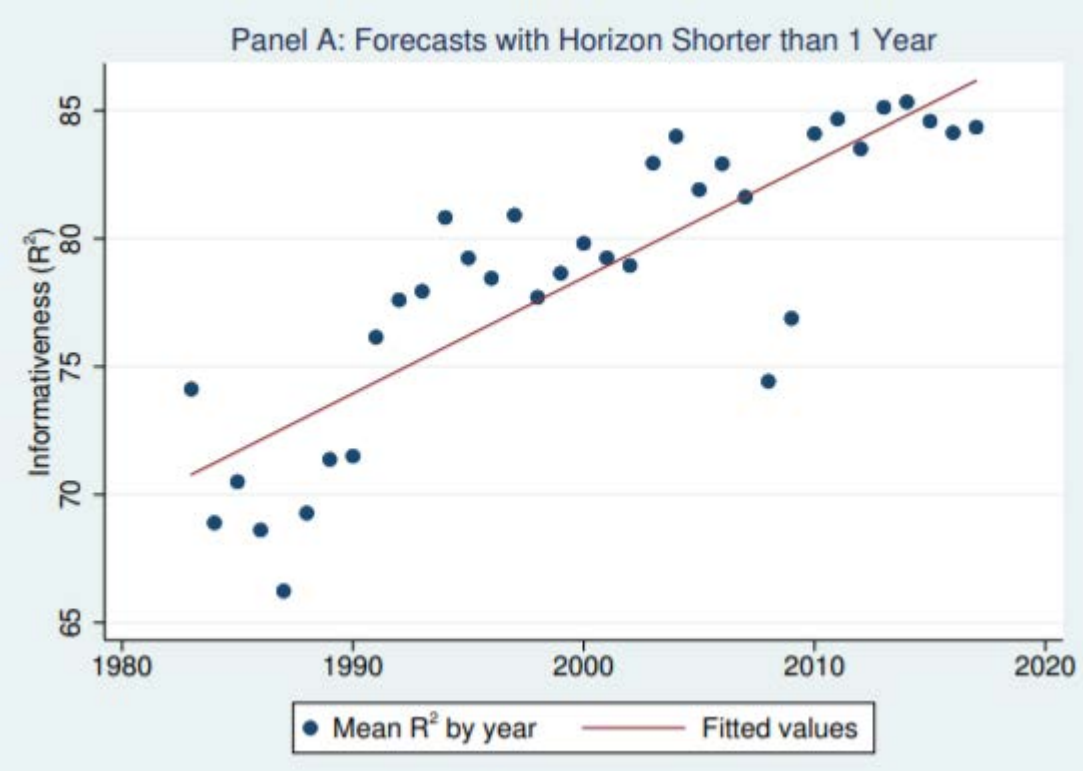
- Increase in the number of alternative data providers over time, particularly since the late 2000s
- Overall impact on order flow and stock prices is positive (Gao and Huang 2020, Zhu 2019, Grennan and Michaely, 2020)
- Impact is different for small-size firms and for large-size firms (Bai et al., 2016 and Farboodi et al. 2020)
- Social media is useful to financial analysts (Chi et al., 2021)

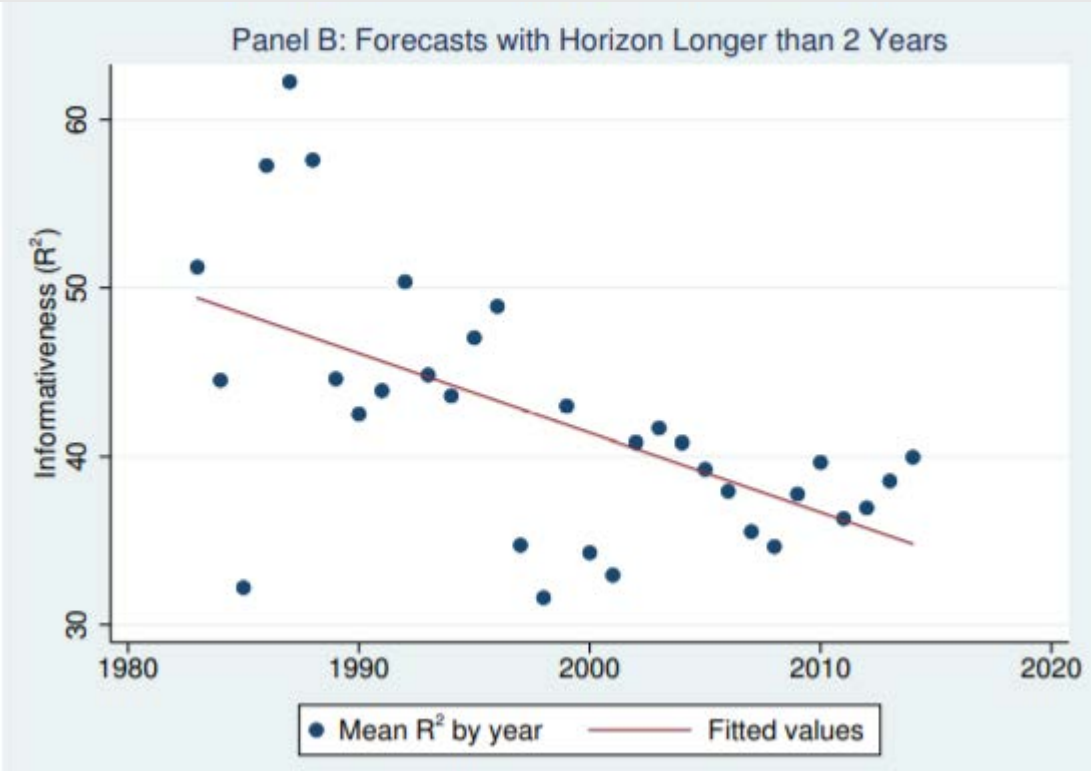
Impact on both short-term horizon and long-term horizon?
Determinants of analysts' forecasts?

Synopsis: contributions

- From a theoretical viewpoint
  - Greater exposure to alternative data increases R^2 for low values of h
  - Possibly decreases R^2 for high values of h
- From an empirical viewpoint
  - I/B/E/S
    - The relationship between R^2 and the time horizon h is downward-sloping
    - If h = one year, then R^2 increases by 10% (60% → 70%)
    - If h = five years, then R^2 decreases by 10% (40% → 30%)
    - The annual slope of the term structure has become steeper over time
  - StockTwits:
    - Significant improvement in the informativeness of analysts' short-term forecast (< 1 year)
    - Significant decline of long-term forecasts (> 2 years)
    - The slope is steeper if the financial analyst is posting messages on StockTwits
- Lot of robustness checks: trading volume, analysts who always cover the same firms, analysts and firms with non-missing long-term forecasts.

Short vs. long-term forecast informativeness by year (Fig. 4)





Data exposure and the slope of term structure (Tab. VI)

Slope 1-year horizon > 1-year horizon

$$R^2_{i,t,h} = \lambda_0 h^* + \lambda_1(\text{Data Exposure}_{i,t-1}) + \lambda_2(h^* \times \text{Data Exposure}_{i,t-1}) + ... + \omega_{i,t,h}. \quad (18)$$

Dep. variable:	Forecast informativeness (R^2)					
Data Exposure Proxy:	*#Watchlist*			*#Hypothetical Messages*		
OLS:	(1)	(2)	(3)	(4)	(5)	(6)
h^* × Data Exposure	-0.86***	-0.78***	-0.96***	-0.69***	-0.94***	-1.05***
	(-2.59)	(-3.06)	(-3.72)	(-2.75)	(-4.54)	(-5.03)
Data Exposure	0.13	-0.17	-0.35	0.34	-0.14	-0.32
	(0.50)	(-0.64)	(-1.29)	(1.42)	(-0.57)	(-1.30)
h^*	-16.66***			-16.62***		
	(-33.85)			(-32.13)		
Analyst FE	Yes			Yes		
Date FE	Yes			Yes		
Analyst FE (interacted)		Yes	Yes		Yes	Yes
Date FE (interacted)		Yes	Yes		Yes	Yes
Controls			Yes			Yes
N	30,959,281	30,105,556	27,860,429	30,959,281	30,105,556	27,860,429

Major comments

Comments / Suggestions: research question (1/2)

- Alternative data: sensor data tracking retailers' activity (satellite, geolocation); credit card data, social data…
- StockTwits is a social media BUT
  - OA#12: social media by **individuals** are alternative data (JP Morgan classification)
  - OA#8: Positive correlation between news (Capital IQ Key Developments: press release, news wires, regulatory fillings) and the number of messages on StockTwits
  - Integrated into Bloomberg.com or Reuters.com
  - StockTwit users are: traders, financial analysts, investors, etc.

Is StockTwits alternative data?
Isn't it digitization of previous network discussions?

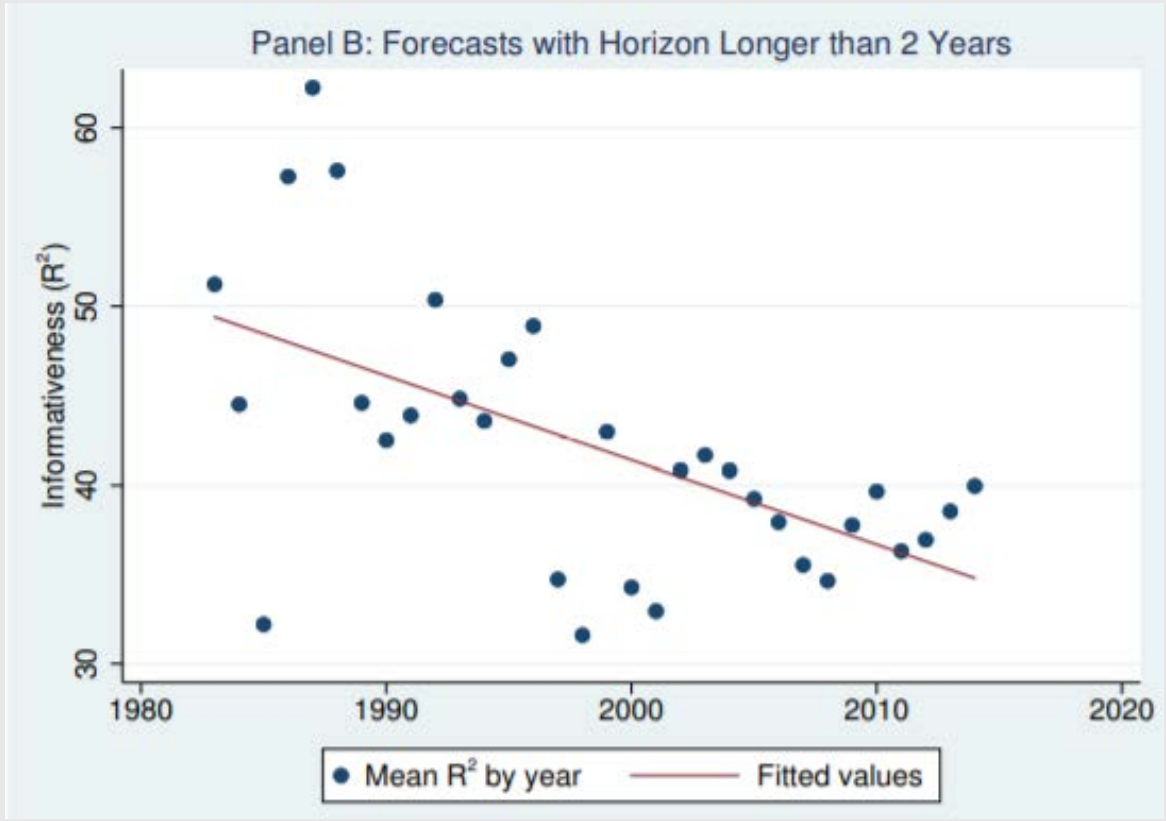
Comments / Suggestions: research question (2/2)

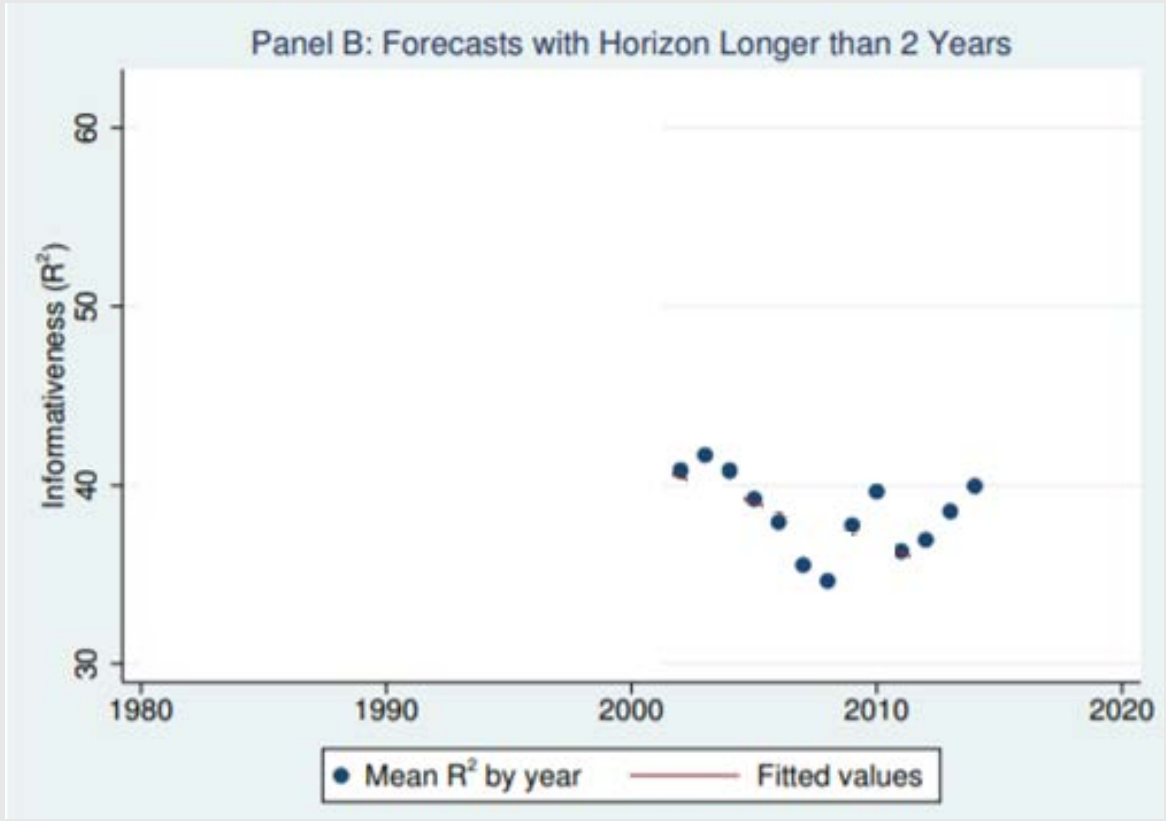
- Hypothesis: multitasking is costly
- If high effort level on short-term horizon, then less effort on long-term horizon

Is it still costly?

Comments / Suggestions: research question (2/2)

Impact of AI?





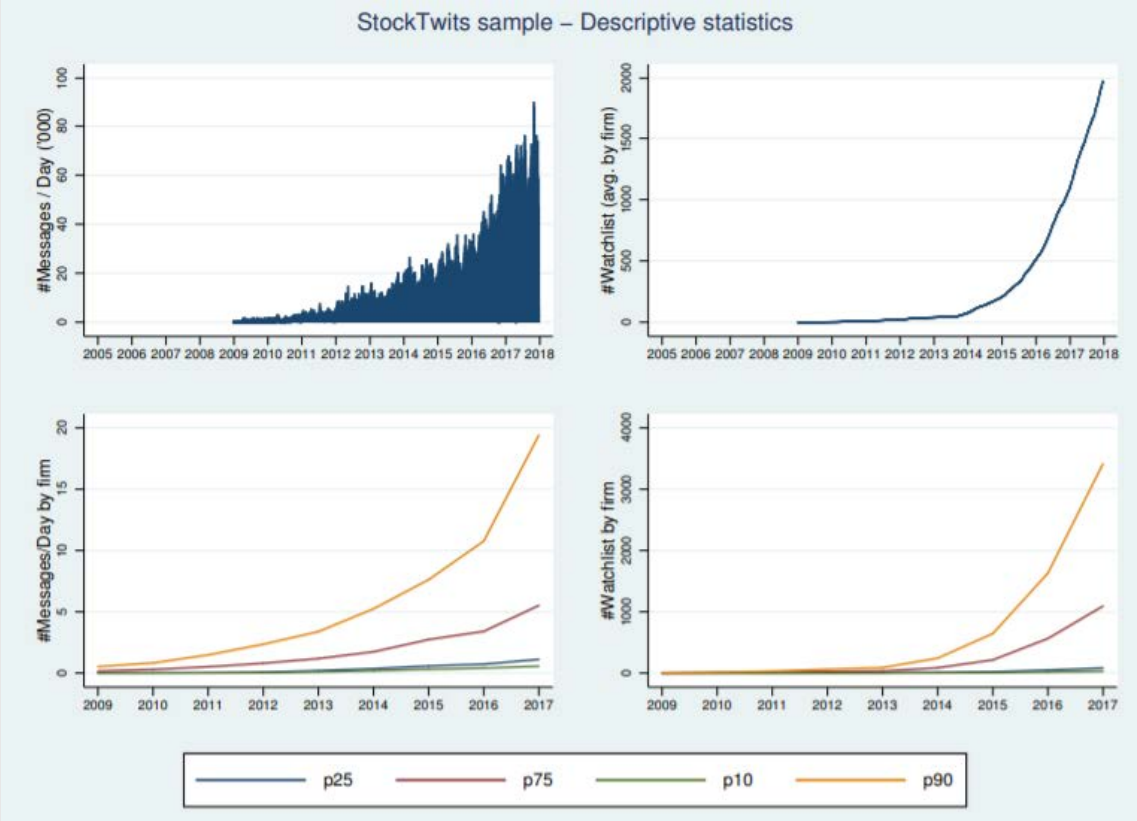
Comments / Suggestions: data (1/3)

- Foucault & Dugast (2021)
  - Data abundance
  - If messages are posted by traders and financial analysts
    - Noisy or not?
    - Reputation effect?

Comments / Suggestions: data (2/3)

Natural experiment

- Is StockTwits relevant to all financial analysts? Especially at the beginning?
- Is StockTwits relevant for all US stocks? Match or missmatch your sample?



Comments / Suggestions: data (3/3)

- More information about financial analysts
  - If on StockTwits (35% with perfect matching):
    - Are they junior or senior analysts?
    - Do they provide long-term forecasts?
    - Small stocks? Large stocks? Timing effect?
    - Did StockTwits change the number of firms they cover?
    - Did it change the firms that they cover?
  - If long-term forecasts:
    - Junior analysts? Senior analysts?
    - Small stocks? Large stocks?

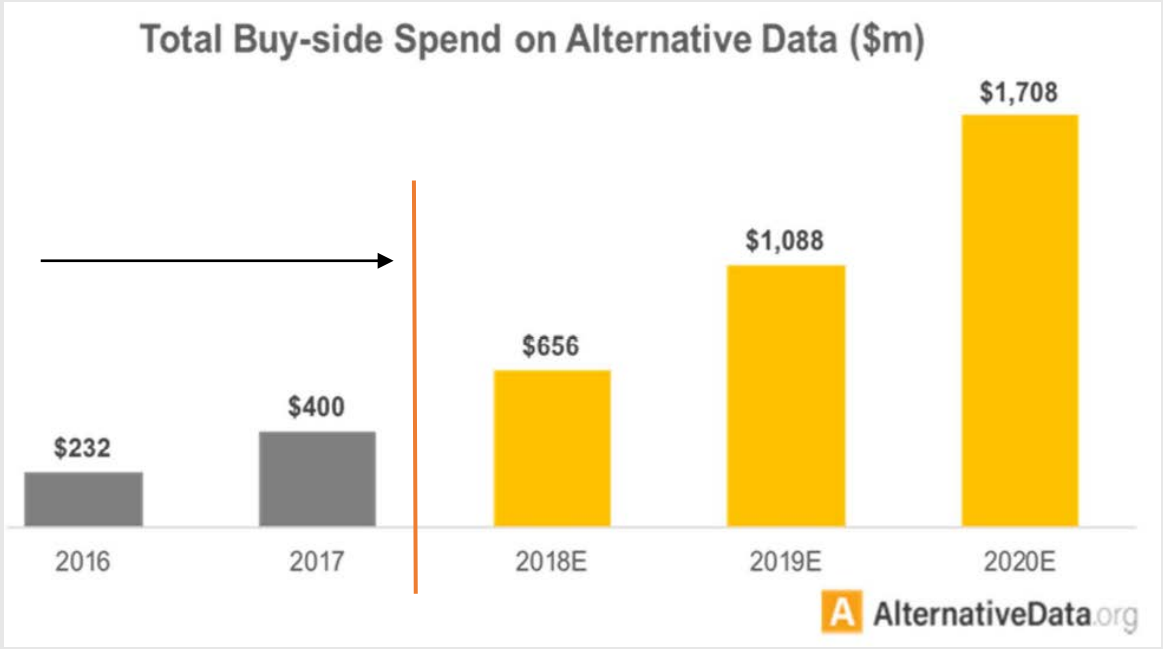
Comments / Suggestions: time period (1/2)

- Sample time period: sell-side from 1983 to 2017
  - → More data
  - → Change in trading strategies
    - HFT
    - Buy and hold: over 1 year on a given industrial sector

Do you capture alternative data impact and/or different trading strategy interests?
Over time, do you have less or more forecast if h > 2 years or if h > 5 years?

Comments / Suggestions: time period (2/2)

Sample time period: from 1983 to 2017



Do you plan to update your results?

Closing remarks

- Very interesting topic and well written paper
- Very promising research paper
- Really enjoyed reading it ☺