Do Investors Overreact to Managerial Tones

Wei Wang
Indiana University

Research Question

1. Qualitative: Are investors overconfident on the public information?
2. Quantitative: If so, how much of market reaction is attributable to this overconfidence?

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Research Question

1. Qualitative: Are investors overconfident on the public information?
2. Quantitative: If so, how much of market reaction is attributable to this overconfidence?

Overconfidence

1. Overconfidence: An overestimation of the precision of one's information [**??**].
2. Example 1 (discrete time): An investor receives an unbiased and noisy signal s about the asset's fundamental value, v. The true distribution of s is $N(v, \sigma_s^2)$. Under the belief of an overconfident investor, s follows a normal distribution of $N(v, \sigma_o^2)$ and $\sigma_o^2 < \sigma_s^2$

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Overconfidence

1. Overconfidence: An overestimation of the precision of one's information [**??**].
2. Example 1 (discrete time): An investor receives an unbiased and noisy signal s about the asset's fundamental value, v. The true distribution of s is $N(v, \sigma_s^2)$. Under the belief of an overconfident investor, s follows a normal distribution of $N(v, \sigma_o^2)$ and $\sigma_o^2 < \sigma_s^2$

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Motivation

1. Overconfidence is an important psychological bias in behavioral asset pricing and macro-economics literature
   - Long-run reversal
   - Stock markets have bubbles that boom and burst
   - Market ”overreaction” to news and disclosure
   - Investors sentiments
2. Research on how investors form their subject beliefs is gaining more attention (challenges on rationality assumption)
   - use survey data for macro series
   - use analysts estimate for firms
3. What remains unanswered: Does overconfidence really exist for a representative investor? how to estimate the overconfidence?
   - There are many papers providing indirect evidence
   - Survey data are usually unavailable for equity investors.
   - Only a few firms have analysts.

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Motivation

1. Overconfidence is an important psychological bias in behavioral asset pricing and macro-economics literature
   - Long-run reversal
   - Stock markets have bubbles that boom and burst
   - Market ”overreaction” to news and disclosure
   - Investors sentiments
2. Research on how investors form their subject beliefs is gaining more attention (challenges on rationality assumption)
   - use survey data for macro series
   - use analysts estimate for firms
3. What remains unanswered: Does overconfidence really exist for a representative investor? how to estimate the overconfidence?
   - There are many papers providing indirect evidence
   - Survey data are usually unavailable for equity investors.
   - Only a few firms have analysts.

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Motivation

1. Overconfidence is an important psychological bias in behavioral asset pricing and macro-economics literature
   - Long-run reversal
   - Stock markets have bubbles that boom and burst
   - Market ”overreaction” to news and disclosure
   - Investors sentiments
2. Research on how investors form their subject beliefs is gaining more attention (challenges on rationality assumption)
   - use survey data for macro series
   - use analysts estimate for firms
3. What remains unanswered: Does overconfidence really exist for a representative investor? how to estimate the overconfidence?
   - There are many papers providing indirect evidence
   - Survey data are usually unavailable for equity investors.
   - Only a few firms have analysts.

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Main Results & Contribution

1. Results
   - Using LDA for topic analysis, I come up with a new way to estimate investors' overconfidence by comparing market reactions to different information
   - Using SP500 firms, I find no evidence of the existence of overconfidence
2. Contribution
   - There is no paper discussing how to quantify this overconfidence without survey data. This paper contributes to this blank
   - Provide novel evidence that may induce another round of debates.

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Main Results & Contribution

1. Results
   - Using LDA for topic analysis, I come up with a new way to estimate investors' overconfidence by comparing market reactions to different information
   - Using SP500 firms, I find no evidence of the existence of overconfidence
2. Contribution
   - There is no paper discussing how to quantify this overconfidence without survey data. This paper contributes to this blank
   - Provide novel evidence that may induce another round of debates.

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Road Map

1. A Toy Model
2. An Empirical Model for Estimation
3. Preliminary Evidence
4. Conclusion

A Toy Model

- Model setup
  1. There is one asset in the market
  2. The fundamental value is v, which is unobservable until being revealed
  3. There is one unit mass of representative investors in the market.
  4. Investors' prior belief on v is normally distributed, $N(\theta_0, \sigma_v^2)$
  5. Investors will receive an unbiased and noisy signal $s = v + \varepsilon$, where the true distribution of noise is $\varepsilon \sim N(0, \sigma_\varepsilon^2)$.
  6. Investors' subjective belief on the noise follows $\varepsilon^C \sim N(0, \sigma_C^2)$
- Sequence of Events
  1. At period 0, investors receive the noisy signal, form a posterior belief, and trade on that belief with an intermediate price P_1
  2. At period 1, the asset pays liquidating dividends

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

A Toy Model

- Model setup
  1. There is one asset in the market
  2. The fundamental value is v, which is unobservable until being revealed
  3. There is one unit mass of representative investors in the market.
  4. Investors' prior belief on v is normally distributed, $N(\theta_0, \sigma_v^2)$
  5. Investors will receive an unbiased and noisy signal $s = v + \varepsilon$, where the true distribution of noise is $\varepsilon \sim N(0, \sigma_\varepsilon^2)$.
  6. Investors' subjective belief on the noise follows $\varepsilon^C \sim N(0, \sigma_C^2)$
- Sequence of Events
  1. At period 0, investors receive the noisy signal, form a posterior belief, and trade on that belief with an intermediate price P_1
  2. At period 1, the asset pays liquidating dividends

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Model Solution

- Intermediate Price: $P_1 = \frac{1/\sigma_v^2}{1/\sigma_v^2 + 1/\sigma_C^2}\theta_0 + \frac{\lambda/\sigma_\varepsilon^2}{1/\sigma_v^2 + \lambda/\sigma_C^2} s$ where $\lambda \equiv \frac{\sigma_\varepsilon^2}{\sigma_C^2}$.
- Observation 1: As $\lambda \to \infty$, $P_1 \to s$. I use overconfidence price v^o to refer to the posterior belief on the value when investors think signals are completely precise.
- Observation 2: If $\lambda = 1$, P_1 equals to the posterior mean of an investor using rational Bayesian updating scheme. I use the v^* to the posterior belief on the value when investors are not overconfidence .
- Observation 3: We can express the intermediate price as a weighted average of P_1^* and s, i.e., $P_1 = (1-\chi)v^* + \chi v^o$ where $\chi = \frac{\lambda - 1}{\frac{\sigma_\varepsilon^2}{\sigma_v^2} + \lambda}$
- The market reaction to v^* or v^o depends on the level of overconfidence

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Model Solution

- Intermediate Price: $P_1 = \frac{1/\sigma_v^2}{1/\sigma_v^2 + 1/\sigma_C^2}\theta_0 + \frac{\lambda/\sigma_\varepsilon^2}{1/\sigma_v^2 + \lambda/\sigma_C^2} s$ where $\lambda \equiv \frac{\sigma_\varepsilon^2}{\sigma_C^2}$.
- Observation 1: As $\lambda \to \infty$, $P_1 \to s$. I use overconfidence price v^o to refer to the posterior belief on the value when investors think signals are completely precise.
- Observation 2: If $\lambda = 1$, P_1 equals to the posterior mean of an investor using rational Bayesian updating scheme. I use the v^* to the posterior belief on the value when investors are not overconfidence .
- Observation 3: We can express the intermediate price as a weighted average of P_1^* and s, i.e., $P_1 = (1-\chi)v^* + \chi v^o$ where $\chi = \frac{\lambda - 1}{\frac{\sigma_\varepsilon^2}{\sigma_V^2} + \lambda}$
- The market reaction to v^* or v^o depends on the level of overconfidence

Model Solution

- Intermediate Price: $P_1 = \frac{1/\sigma_v^2}{1/\sigma_v^2 + 1/\sigma_C^2}\theta_0 + \frac{\lambda/\sigma_\varepsilon^2}{1/\sigma_v^2 + \lambda/\sigma_C^2} s$ where $\lambda \equiv \frac{\sigma_\varepsilon^2}{\sigma_C^2}$.
- Observation 1: As $\lambda \to \infty$, $P_1 \to s$. I use overconfidence price v^o to refer to the posterior belief on the value when investors think signals are completely precise.
- Observation 2: If $\lambda = 1$, P_1 equals to the posterior mean of an investor using rational Bayesian updating scheme. I use the v^* to the posterior belief on the value when investors are not overconfidence .
- Observation 3: We can express the intermediate price as a weighted average of P_1^* and s, i.e., $P_1 = (1-\chi)v^* + \chi v^o$ where $\chi = \frac{\lambda - 1}{\frac{\sigma_\varepsilon^2}{\sigma_v^2} + \lambda}$
- The market reaction to v^* or v^o depends on the level of overconfidence

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Model Solution

- Intermediate Price: $P_1 = \frac{1/\sigma_v^2}{1/\sigma_v^2 + 1/\sigma_C^2}\theta_0 + \frac{\lambda/\sigma_\varepsilon^2}{1/\sigma_v^2 + \lambda/\sigma_C^2}s$ where $\lambda \equiv \frac{\sigma_\varepsilon^2}{\sigma_C^2}$.
- Observation 1: As $\lambda \to \infty$, $P_1 \to s$. I use overconfidence price v^o to refer to the posterior belief on the value when investors think signals are completely precise.
- Observation 2: If $\lambda = 1$, P_1 equals to the posterior mean of an investor using rational Bayesian updating scheme. I use the v^* to the posterior belief on the value when investors are not overconfidence .
- Observation 3: We can express the intermediate price as a weighted average of P_1^* and s, i.e., $P_1 = (1-\chi)v^* + \chi v^o$ where $\chi = \frac{\lambda - 1}{\frac{\sigma_\varepsilon^2}{\sigma_v^2} + \lambda}$
- The market reaction to v^* or v^o depends on the level of overconfidence

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Model Solution

- Intermediate Price: $P_1 = \frac{1/\sigma_v^2}{1/\sigma_v^2 + 1/\sigma_C^2}\theta_0 + \frac{\lambda/\sigma_\varepsilon^2}{1/\sigma_v^2 + \lambda/\sigma_C^2} s$ where $\lambda \equiv \frac{\sigma_\varepsilon^2}{\sigma_C^2}$.
- Observation 1: As $\lambda \to \infty$, $P_1 \to s$. I use overconfidence price v^o to refer to the posterior belief on the value when investors think signals are completely precise.
- Observation 2: If $\lambda = 1$, P_1 equals to the posterior mean of an investor using rational Bayesian updating scheme. I use the v^* to the posterior belief on the value when investors are not overconfidence .
- Observation 3: We can express the intermediate price as a weighted average of P_1^* and s, i.e., $P_1 = (1-\chi)v^* + \chi v^o$ where $\chi = \frac{\lambda - 1}{\frac{\sigma_\varepsilon^2}{\sigma_v^2} + \lambda}$
- The market reaction to v^* or v^o depends on the level of overconfidence

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Estimation Challenges

To quantify χ, there are many complexities of identification and calibration

1. (signal selection) In the real world, there are many signals about different dimensions of firms' fundamental values. But there are very few that are unbiased about fundamental values directly
2. Without good methods to calibrate σ_v and σ_ε, χ cannot be accurately identified.
3. Solution: starting from signals that can be measured accurately.

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Extension to Multifactor Models

Motivation: We want to incorporate factors that are easy to measure in the data

1. Multifactor fundamental value: $v = \sum_{i=1}^{N} w_i f_i$. where the prior belief on $f_i \sim N(\mu_i, \sigma_i^2)$
2. More signals: s_i, follows $s_i = a_i + b_i f_i + \varepsilon_i$, where $\varepsilon_i \sim N(0, \sigma_{\varepsilon,i}^2)$ is the noise term and independent of all other variables, including factors and other noises.
3. Investors: there are a unit measure of atomistic informed traders submitting orders
    - The information set is the same for all investors, $\mathscr{I} = \{(w_i)_{i=1}^{N}, (s_i)_{i=1}^{N}, (\mu_i)_{i=1}^{N}, (\sigma_i^2)_{i=1}^{N}, (\sigma_{\varepsilon,i}^2)_{i=1}^{N}\}$.
    - They are bounded-rational (or partially-overconfident), i.e., when they form posterior beliefs, they will use the weighted average: $E[v|\mathscr{I}] = (1-\chi)v^* + \chi v^o$

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Extension to Multifactor Models

Equilibrium:

1. The informed investor will submit orders x by maximizing the profits given the stock price p: $\max_x E[(v-p)x - \frac{m}{2}x^2|\mathscr{I}]$ where m captures the trading costs
2. There will be noisy trades, $u \sim N(0, \sigma_u)$
3. The market clearing requires $\int_i x_i di = u$

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Extension to Multifactor Models

Motivation: We want to incorporate factors that are easy to measure in the data

Theorem

If the representative investor has an overconfidence level χ, the equilibrium stock price will be

$$p = \chi \sum_{i=1}^{N} w_i \underbrace{\frac{s_i - a_i}{b_i}}_{\text{Original signal}} + (1-\chi) \sum_{i=1}^{N} w_i \underbrace{\frac{1}{1/\sigma_i^2 + b_i^2/\sigma_{\varepsilon,i}^2} \left[\frac{b_i^2}{\sigma_{\varepsilon,i}^2} \frac{s_i - a_i}{b_i} + \frac{1}{\sigma_i^2} \mu_i \right]}_{\text{Corrected signal}} + m \qquad (1)$$

Empirical Models

1. Let's move from price to return level. Our regression model will be:

$$r_t = \alpha + \sum_{i=1}^{N} \tilde{w}_i[\tilde{s}_i + k\tilde{s}_i^*] + X_t\Phi + \varepsilon_t \tag{2}$$

2. $k = \frac{1-\chi}{\chi}$
3. $\tilde{s}_i = \frac{s_i - a_i}{b_i}$, $\tilde{s}_i^* = \frac{1}{1/\sigma_i^2 + b_i^2/\sigma_{\varepsilon,i}^2} \frac{b_i^2}{\sigma_{\varepsilon,i}^2} \frac{s_i - a_i}{b_i} + \frac{1}{1/\sigma_i^2 + b_i^2/\sigma_{\varepsilon,i}^2} \frac{1}{\sigma_i^2} \mu_i$.

More Comments

1. Our model considers the possibility that different dimensions may have heterogeneous effects on fundamentals by estimating w_i
2. By decomposing signals into different dimensions, we are able to calibrate the signal precision.
3. The regression is conceptually straightforward

Variable Construction

1. What are the signals?
   - I use the textual information from corporate annual reports to measure managerial tones on different topics
2. What are the fundamental measures?
   - For each topic identified, I construct a measure of it using accounting variables
3. How to execute it?
   - For topic analysis, I use an unsupervised technique called LDA
   - Divide documents into sentences
   - Use LDA to categorize each sentence into a topic
   - Calculate the managerial tones by counting the positive and negative words

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Variable Construction

1. What are the signals?
   - I use the textual information from corporate annual reports to measure managerial tones on different topics
2. What are the fundamental measures?
   - For each topic identified, I construct a measure of it using accounting variables
3. How to execute it?
   - For topic analysis, I use an unsupervised technique called LDA
   - Divide documents into sentences
   - Use LDA to categorize each sentence into a topic
   - Calculate the managerial tones by counting the positive and negative words

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Variable Construction

1. What are the signals?
   - I use the textual information from corporate annual reports to measure managerial tones on different topics
2. What are the fundamental measures?
   - For each topic identified, I construct a measure of it using accounting variables
3. How to execute it?
   - For topic analysis, I use an unsupervised technique called LDA
   - Divide documents into sentences
   - Use LDA to categorize each sentence into a topic
   - Calculate the managerial tones by counting the positive and negative words

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

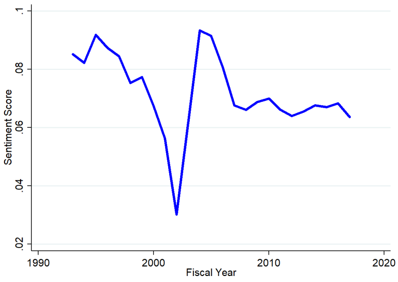
Variable Construction

1. I focus on SP500 index firms
2. Corporate accounting information: Compustat
3. Stock prices: CRSP
4. Coporate annual reports: Loughran and McDonald Stage-One 10-X Parse files updated to 2018.
5. Sentiment words: LM dictionary on positive vs negative words
6. The final sample contains 12,028 documents from 976 unique firms
7. 2,340 unique sentiment words and 112,961 neutral words

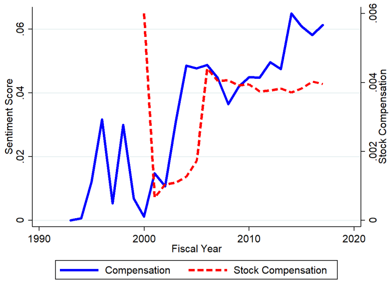
LDA Results

- Topic 1 (General): FINANCIAL COMPANY STATEMENTS REPORT FORM CONSOLIDATED
- Topic 2 (Compensation): COMPANY **SHARES COMMON** PLAN **SHARE OPTIONS** BASED NOTES **COMPENSATION PLANS** COURT DATE **EMPLOYEES**
- Topic 3 (Sales): **PRODUCTS** COMPANY **SERVICES BUSINESS PRODUCT CUSTOMERS** INCLUDING OPERATIONS **MARKET** SYSTEMS **OPERATING**
- Topic 4 (Energy): **GAS** COSTS COMPANY **ENERGY NATURAL** OPERATIONS **OIL** APPROXIMATELY **PRODUCTION COST**
- Topic 5 (Debts): COMPANY **CREDIT** FINANCIAL **INTEREST RATE RISK** MARKET **DEBT TERM SECURITIES**
- Topic 6 (Profits): **INCOME TAX NET ASSETS** COMPANY FINANCIAL RELATED CONSOLIDATED STATEMENTS **EXPENSE**

Measure Construction: Validation



(1) General



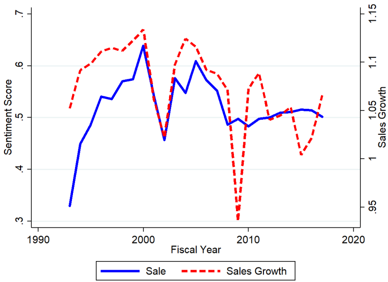
(2) Compensation

Figure: **Dynamics of Topic Sentiments**

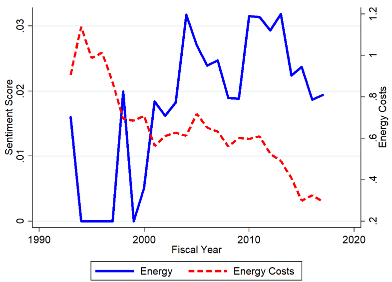
Note: Stock Compensation Expense is defined as STKCO/AT.

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Measure Construction: Validation



(3) Sales



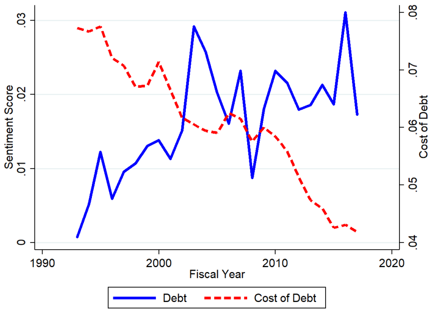
(4) Energy

Figure: **Dynamics of Topic Sentiments**

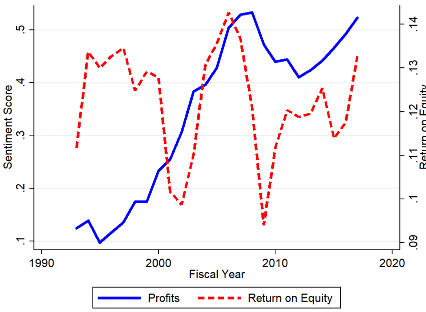
Note: Sales Growth is defined as SALE divided by lagged SALE. Projected Energy Costs is calculated in the following way. First, project the total operating cost, Compustat Item XOPR, on the global price of energy index using linear regressions and then scale it with book value of assets, COMPUSTAT item AT.

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Measure Construction: Validation



(5) Debts



(6) Profits

Figure: **Dynamics of Topic Sentiments**

Note: Cost of Debt is defined as Item XINT/(Item DLTT + Item DLC). Return on Equity is defined as IB/((BE+LAG(BE)/2))

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

Maximum Likelihood Estimation

We use five dimensions: 5 signals and 5 fundamentals.

	Compensation	Sales	Energy	Debts	Profits
Compensation*	0.0716	0.00942	0.0163	0.0162	-0.00173
Sales*	0.0260	0.0750	-0.00111	0.0198	-0.0143
Energy*	0.0162	0.0114	0.0356	0.00655	0.00348
Debts*	-0.0146	0.00450	0.00834	0.0901	-0.00656
Profits*	0.00368	-0.00939	0.00855	-0.00941	0.0752

Maximum Likelihood Estimation

Param	Estimate	90% Confidence Interval
Compensation (w_1)	-2.7989	[-17.914,12.211]
Sales (w_2)	-0.2332	[-1.099,0.737]
Energy (w_3)	0.1095	[0.017,0.217]
Debts (w_4)	-0.0191	[-1.928,2.134]
Profits (w_5)	1.0350	[0.282,2.091]
χ	0.0106	[0,0.021]
Size	-0.0629	[-0.118,-0.008]
Market-to-Book Ratio	-0.0577	[-0.169,0.046]
Leverage	0.2881	[-0.392,1.035]
General (s_i)	0.8908	[0.279,1.447]
Const	0.6748	[-0.599,1.803]
σ^2	17.2344	[16.493,18.066]

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

More for the future

1. Can this estimate generate some implications on real economy?
2. Can sentiments and overconfidence generate real impact?
3. How to shape this paper into an influential one?
4. Which direction to push?

Thank you very much for joining!