

# Media News and Cross Industry Information Diffusion

Li Guo

Singapore Management Univeristy

June 13, 2017

# Motivatiion

- **Cross Asset Return Predictability:**

- **Information Diffusion:**

- Hong and Stein (1999): Theory Model on Information Diffusion.
- Cohen and Frazzini (2008): Firm  $\leftrightarrow$  Firm.
- Hong and Valkavov (2007): Industry  $\rightarrow$  Stock Market.
- Menzly and Ozbas (2010): Industry  $\leftrightarrow$  Industry.
- Rapach et.al (2015): Core Industries + Technique Spillover Effect.

- **Liquidity:** Lo and MacKinlay (1990), Brennan and Swaminathan (1993), Badrinath and Noe (1995), Jegadeesh and Titman (1995)

- **Own-autocorrelation and cross-correlation:** Boudoukh and Whitela (1994)

- **Why Information Diffusion?**

Economic Linkage and Network Complexity

- **Research Question:**

**How News Travels across Different Industries?**

# Hypothesis

- **Assumptions of Information Diffusion (Hong and Valkavov, 2007):**
  - **Investors' Limited Attention:**
    - Fang and Peress (2009): # of News
    - Barber and Odeani (2008): Extreme trading volume/return
    - Jacobs and Weber (2012): Regional holidays
    - DellaVigna and Pollet (2009): Friday Dummy
    - ...
  - **News slowly diffuses across different industries (no formal test).**
    - Peress (2015): Media strike drives return predictability.
    - Hong (2009): Negative news travels slowly.

# Hypothesis I

- **Cross industry news contains valuable information of firm fundamentals.**
  - Dyck and Zingales (2002): Media selectively reduce the cost of acquiring and verifying information.
  - Tetlock et.al (2008): Negative words predict future earnings.
  - Bushee et.al (2010): Media serves as an information intermediary.
  - Peress (2015): Media increases the speed with which information diffuses across investors.
- **Predictions:**
  - Cross industry news predicts firm fundamentals (SUE).
  - Cross industry news is more valuable to those stocks with poor information environment.

## Hypothesis II

- **Cross industry news **slowly** diffuses.**
  - Hard interpretation of cross industry news (Information Intermediary)
- **Prediction:**
  - Cross industry news has long term effect on future stock returns.
  - Sophisticated investors are able to interpret cross industry news to the market. (Analysts/Institutional investors)
  - Cross industry news should have **longer** effects on stocks with **low analyst coverage** or **low institutional holdings** than others. (Not report yet)



## Data Sources

- **News Archive:** Thomson Reuters News Archive (From 01-Jan-1996 to 31-Dec-2014);
- **Active Institutional Fund Flow:** EPFR database.
- Earnings forecasts and analyst related data from the IBES detailed history;
- All other accounting variables come from COMPUSTAT, stock returns come from CRSP.
- **Sentiment Index:** <http://apps.olin.wustl.edu/faculty/zhou/>
- **Policy Uncertainty Index:** <http://www.policyuncertainty.com/>

## News Tone Measures

- News Tone (Tetlock et.al, 2008):

$$\text{Tone} = \frac{\# \text{ of negative word occurrences}}{\text{Total } \# \text{ of Words in the news}}. \quad (1)$$

- Firm Specific News:

$$\text{Firm News}_{i,t} = \frac{\sum_{d=1}^D \text{Tone}_{i,d}}{D}. \quad (2)$$

- Peer Industry News:

$$\text{Peer News}_{i,t} = \frac{\sum_{k=1}^K \text{FirmNews}_{k,t}}{K}, \text{ where } i \neq k. \quad (3)$$

- Cross Industry News:

$$\text{Cross Industry News}_{i,J,t} = \frac{\sum_{j=1}^J \text{Firm News}_{j,t}}{J}, \text{ with } i, j \text{ in different industries.} \quad (4)$$

## Cross Industry News Signal (CIS)

- Step I: Predictive regression**

$$r_{i,t+1} = \alpha_i + \sum_{j=1}^N b_{i,j} \text{Industry News}_{j,t} + \epsilon_{i,t+1}, \text{ for } t = 1, \dots, T - 1, \quad (5)$$

- Adaptive Lasso Estimation:

$$\hat{b}_i^* = \operatorname{argmin} \|r_{i,t+1} - \alpha_i - \sum_{j=1}^N b_{i,j} \text{Industry News}_{j,t}\|^2 + \lambda_i \sum_{j=1}^N \hat{w}_i |b_{i,j}|, \quad (6)$$

- Choose inverse of OLS estimation as weighted parameter as Zou (2006).
- Use Cross Validation to select the optimal  $\lambda_i$ .
- Require at least 260 weekly observations for each firm and set the initial estimation window as 208.

- Step II: Out-of-sample Forecast**

$$\text{CIS}_{i,t} = \hat{r}_{i,T+1} = \alpha_i + \sum_{j=1}^N E_t[b_{i,j,t+1}] \text{Industry News}_{j,t}, \quad (7)$$

- Step III: Repeat step I and step II recursively to have a time series of CIS for each firm.**



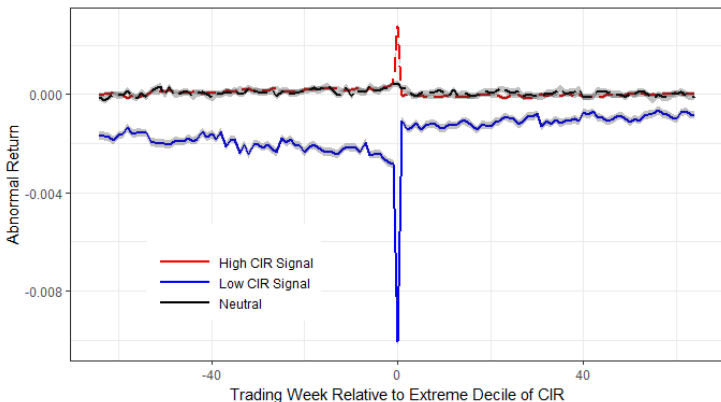
## Cross Industry News Signal (CIS) Formation

- Repeat Table 2 of Rapach, D., Strauss, J., Tu, J., & Zhou, G. (2015)

```
31 x 7 sparse Matrix of class "dgcmatrix"
```

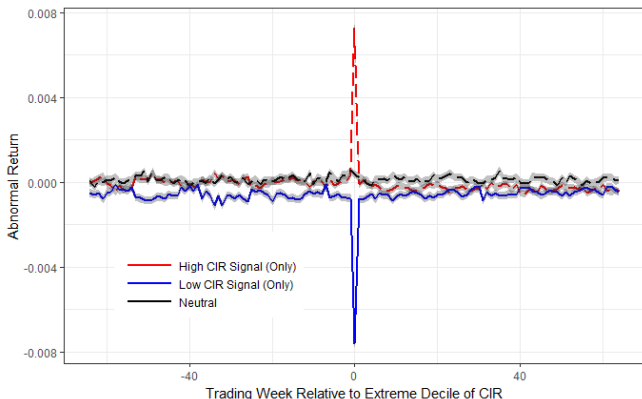
	food	beer	smoke	games	books	hshld	clths
intercept	0.01040179	0.009893651	0.01247286	0.008590713	0.008113343	0.008809311	0.009574341
food	.	0.157717279	.	.	.	.	.
beer	.	-0.029904254	.	.	.	.	.
smoke	.	.	.	.	.	.	.
games	.	.	.	.	.	.	.
books	.	.	.	0.175269375	.	0.019834329	0.030063468
hshld	.	.	.	.	.	.	-0.072644475
clths	.	0.011733393	.	.	.	0.099372463	0.006659671
hlth	.	.	.	.	.	.	.
chems	.	.	.	.	.	.	0.135371311
txtls	.	.	.	.	.	.	.
cnstr	.	.	.	.	.	.	0.041734600
steel	.	.	.	.	.	-0.031693673	-0.075104119
fabpr	0.01491884	.	.	.	.	.	.
elceq	.	.	.	.	.	.	-0.187419226
autos	.	.	.	.	.	.	.
carry	.	.	0.14399277	.	.	.	0.007594611
mines	.	.	.	.	.	.	.
coal	-0.02088105	-0.029619731	.	.	.	-0.032771990	-0.044677548
oil	-0.04990590	-0.002064895	-0.09567388	-0.049104047	-0.108354980	.	-0.121297797
util	0.10737464	.	0.23149128	.	0.016084642	.	0.132418643
telcm	.	.	.	.	.	.	-0.094311528
servs	.	.	-0.11193850	.	0.065461321	.	0.126879206
buseq	.	.	.	.	.	.	0.092677008
paper	.	.	-0.04333529	.	.	.	.
trans	.	.	.	.	.	.	.
whlsl	.	.	.	.	.	-0.042535392	.
rtail	0.01377803	.	.	.	.	.	0.098222013
meals	.	.	.	.	.	.	.
fin	.	.	0.01070637	0.107172036	0.156387160	0.066797468	0.086463627
other	.	.	.	.	.	.	.

## CIS v.s. CIR I: Abnormal Return around CIR



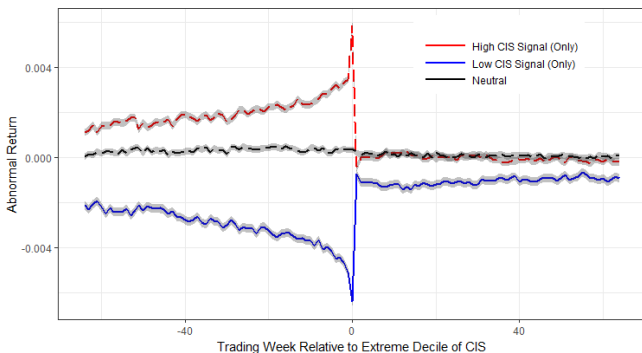
**Figure:** This figure plots firms' average abnormal return from 64 trading weeks preceding a top (bottom) decile of out-of-sample forecasted return based on CIR. All news stories come from Thomson Reuters between 1996 and 2014 inclusive. The out-of-sample period is 2000-2014. We also compute 95 confidence interval for both positive and negative signals represented by the gray area.

## CIS v.s. CIR II: Abnormal Return around CIR only



**Figure:** This figure plots firms' average abnormal return from 64 trading weeks preceding a top (bottom) decile of out-of-sample forecasted return based on unique CIR signals. All news stories come from Thomson Reuters between 1996 and 2014 inclusive. The out-of-sample period is 2000-2014. We also compute 95 confidence interval for both positive and negative signals represented by the gray area.

## CIS v.s. CIR III: Abnormal Return around CIS only



**Figure:** This figure plots firms' average abnormal return from 64 trading weeks preceding a top (bottom) decile of out-of-sample forecasted return based on unique CIS. All news stories come from Thomson Reuters between 1996 and 2014 inclusive. The out-of-sample period is 2000-2014. We also compute 95 confidence interval for both positive and negative signals represented by the gray area.

## Cross Industry News and SUE

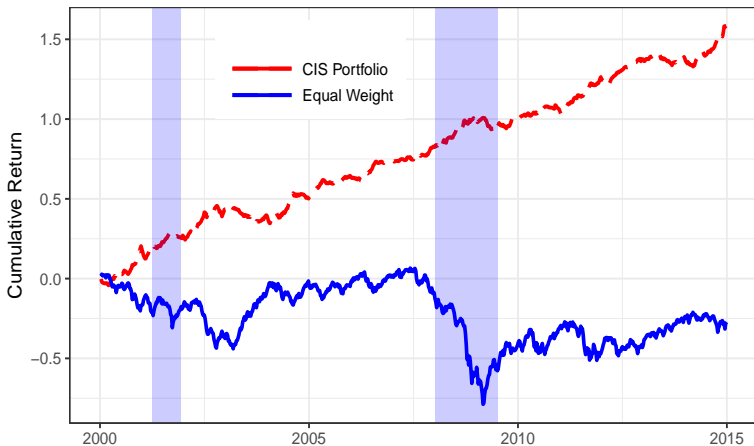
Adaptive Lasso:	No Controls		Tetlock 2008		All Controls	
	Coef	<i>P</i> – value	Coef	<i>P</i> – value	Coef	<i>P</i> – value
Food	<b>0.09</b>	< 0.01			0.01	0.28
Beer	<b>-0.06</b>	< 0.01				
Smoke	<b>-0.03</b>	0.02				
Games	<b>-0.03</b>	0.02				
Books	<b>-0.03</b>	0.01				
Hshld	-0.02	0.13				
Cltls	0.00	0.31				
Hlth	0.02	0.20				
Chems						
Txtls	<b>-0.07</b>	< 0.01				
Cnstr						
Steel						
FabPr	-0.01	0.36				
ElcEq	<b>-0.07</b>	< 0.01	<b>-0.01</b>	< 0.01	<b>-0.02</b>	< 0.01
Autos	<b>-0.11</b>	< 0.01	<b>-0.11</b>	< 0.01	<b>-0.08</b>	< 0.01
Carry						
Mines	<b>-0.03</b>	0.01				
Coal	<b>0.06</b>	< 0.01				
Oil	<b>0.12</b>	< 0.01	<b>0.01</b>	< 0.01	<b>0.02</b>	< 0.01
Util	<b>0.11</b>	< 0.01	<b>0.04</b>	< 0.01	<b>0.05</b>	< 0.01
Telcm	0.02	0.23				
Servs	<b>0.04</b>	0.03				
BusEq	<b>0.04</b>	0.02				
Paper	<b>0.02</b>	0.10				
Trans	<b>-0.07</b>	< 0.01	<b>-0.02</b>	< 0.01		
Whlsl	<b>-0.03</b>	0.07				
Rtail	<b>-0.05</b>	< 0.01				
Meals	<b>0.05</b>	< 0.01				
Fin	<b>0.04</b>	0.04				
Other	-0.01	0.16				

# Fama-MacBeth regressions of stock returns on CIS

	2000 - 2014			2000 - 2007			2008 - 2014		
<b>CIS</b>	<b>0.137***</b> (7.15)	<b>0.130***</b> (6.48)	<b>0.095***</b> (5.09)	<b>0.126***</b> (5.21)	<b>0.118***</b> (4.51)	<b>0.079***</b> (3.32)	<b>0.153***</b> (4.90)	<b>0.148***</b> (4.74)	<b>0.118***</b> (3.98)
Lagged Return	-0.036*** (-21.08)	-0.036*** (-22.12)	-0.042*** (-23.60)	-0.039*** (-17.62)	-0.039*** (-18.46)	-0.045*** (-19.27)	-0.032*** (-11.81)	-0.032*** (-12.43)	-0.038*** (-13.73)
Peer News		-0.001 (-0.60)	-0.002 (-1.53)		-0.001 (-0.56)	-0.001 (-1.14)		-0.001 (-0.28)	-0.002 (-1.02)
Firm News		-0.001** (-2.04)	-0.002*** (-5.02)		-0.000 (-0.68)	-0.001*** (-2.72)		-0.001*** (-2.63)	-0.002*** (-5.09)
# of Peer News		0.000 (0.01)	0.000 (0.24)		0.000 (0.06)	-0.000 (-0.54)		-0.000 (-0.07)	0.000 (1.11)
# of Firm News		0.001*** (5.17)	0.000 (0.76)		0.000* (1.66)	-0.000** (-2.25)		0.001*** (7.29)	0.001*** (6.05)
Size			0.001*** (8.39)			0.001*** (6.67)			0.001*** (5.13)
B/M			0.000*** (8.19)			0.000*** (6.10)			0.000*** (5.52)
Turnover			-0.000*** (-4.83)			-0.000*** (-3.84)			-0.000*** (-3.00)
Leverage			-0.003*** (-5.90)			-0.003*** (-4.58)			-0.002*** (-3.78)
Volatility			-0.013*** (-3.41)			-0.009* (-1.73)			-0.020*** (-3.44)
Intercept	-0.000 (-0.04)	-0.000 (-0.14)	-0.001** (-2.41)	-0.001 (-0.94)	-0.001 (-1.25)	-0.001*** (-2.95)	0.001 (0.99)	0.001 (0.82)	-0.000 (-0.64)
N	1,401,162	1,401,162	1,401,162	855,092	855,092	855,092	546,070	546,070	546,070
Average R <sup>2</sup> (%)	1.17	2.43	4.00	1.20	2.54	4.06	1.14	2.28	3.91

## Long-short Portfolio

- Log Cumulative Returns of Cross-Industry-News (CIS) portfolio



## Sensitivity of News Based Trading Returns to Forecast Horizons

Week after News	Cross Industry News			
	Raw Return (%)	$T_{Raw}$	$\alpha$ (%)	$T_{\alpha}$
2	11.49	5.65	13.16	6.63
3	9.40	4.59	11.28	5.60
4	10.77	5.31	12.25	6.18
5	13.01	5.81	14.84	7.11
6	10.14	4.97	12.43	6.28
7	10.48	4.96	12.69	6.08
8	11.97	5.46	13.65	6.33
9	13.77	6.63	15.79	7.79
<b>10</b>	9.86	4.48	<b>10.85</b>	4.99



## Sensitivity of News Based Trading Returns to Forecast Horizons

Week after News	Firm Specific News			
	Raw Return (%)	$T_{Raw}$	$\alpha$ (%)	$T_{\alpha}$
2	2.96	1.70	<b>3.12</b>	<b>1.85</b>
3	2.99	1.63	3.19	1.94
4	3.82	2.18	3.87	2.30
5	4.58	2.67	5.13	3.11
6	1.91	1.09	2.25	1.32
7	1.43	0.84	1.86	1.14
8	1.18	1.49	1.38	1.68
9	3.82	2.28	3.97	2.44
<b>10</b>	1.69	1.05	<b>2.05</b>	<b>1.31</b>

## Robustness Test & Further Discussion

### • Robustness Test

- CIS is more valuable to **small** stocks, **illiquid** stocks, **high volatility** stocks, **low analyst coverage** stocks and **high analyst dispersion** stocks.
- Cross industry news is more valuable during a **high uncertainty period** proxied by VIX and news dispersion, while it is not sensitive to policy uncertainty.
- CIS is more valuable during a **high sentiment** period than the low sentiment periods.

### • Channels of News Traveling

- Cross Industry News Tone affect **analyst forecast** revision and improves their forecast accuracy.
- Cross Industry News Tone affect active **institutional fund flows**.
- CIS is incorporated into **firm specific news tone** by a delayed timer.

## Conclusions

- Media News Travels Slowly across Different Industries.
  - Cross industry news contains **valuable information** about firm fundamentals.
  - A long-short portfolio based on CIS generates an annualized risk adjusted return **10.85% after 10 weeks** of the signal.
  - Cross industry news is more valuable to stocks with **small size, illiquidity, high volatility, low analyst coverage and high analyst dispersion**.
  - Cross industry news is more valuable during a **high uncertainty** period and **high sentiment** period.
  - **Analyst forecasts, institutional fund flows and media news** might be 3 channels that interpret cross industry news to the market.