Discussion of:

Lazy Prices

by:

Lauren Cohen, Chris Malloy and Quoc Nguyen

Kent Daniel[†]

[†]Columbia Business School

RedRock Conference
8 September, 2016

Basic Idea: *DTP information*

- Information is not processed optimally.
- In particular, the representative agent is inattentive to difficult-to-process (DTP) information.
  - Sims (2003), Hirshleifer and Teoh (2003)
- Thus, this information is only partially incorporated into prices.
- This partial incorporation of information implies that (negative) difficult-to-process information should forecast future negative returns.

Empirical Methodology

- The authors process about 354,000 10-K's and 10-Q's, over 20 years
  - On average, they have these forms for about 4,000 firms
- They measure the "similarity" of the text to to the corresponding 10-K/10-Q from 1-year/4-quarters before.
- The authors conclude that "firms typically repeat what they most recently reported." However, when they change, it is bad news for returns over the next 12-18 months.
- Changers underperform non-changers by about 22%/year over the next 12-18 months; no future reversal of this performance.

Similarity Measures

- The paper proposes four similarity metrics:
  - `Sim_Cosine` and `Sim_Jaccard` are distance measures based exclusively on word counts.
  - `Sim_MinEdit` counts the minimum number of string-operations required to transform one document into another.
  - `Sim_Simple` counts the number of words resulting from a `diff` of the two documents.
- The metrics are fairly highly correlated.
- `Sim_Jaccard` appears to forecast future returns slightly better than the other three.
  - I don't know if there is a statistically significant difference.

Similarity Measures

- `Sim_Cosine`:

$$\texttt{Sim\_Cosine}_{1,2} = \frac{||D_1^{TF} \cdot D_2^{TF}||}{||D_1^{TF}|| \cdot ||D_2^{TF}||}$$

where D_i^{TF} is Term-Frequency vector with the number of counts of each word.

- `Sim_Jaccard`:

$$\texttt{Sim\_Jaccard}_{1,2} = \frac{|D_{S1} \cdot D_{S2}|}{|D_{S1} \mathrm{U} D_{S2}|}$$

where D_{Si} is now the set of words found in document i, and the norm ($|S|$) gives the number of elements in the set.

How Similar are 10-Ks from year-to-year?

Table I: *Summary Statistics*

	Count	Mean	SD	Min	Max
Document Size	353735	159873.7	159873.7	20357	5.24e+07
Document Size - 10K	90198	308633	282473	34660	2.43e+07
Document Size - 10Q	263537	114848.4	286663.9	18824	3.14e+07
Sentiment of Change	353735	-.0003371	.0011069	-.00409	.0048492
Uncertainty of Change	353735	.0007317	.0009165	0	.004885
Litigiousness of Change	353735	.0003252	.0009358	0	.0037628
Change CEO	353735	.0539817	.2259819	0	1
Change CFO	353735	.0238223	.1524956	0	1

- The biggest 10-Ks are $\sim$100 times bigger than the smallest.

How similar are 10-Ks from year-to-year?

Table II-A:

Panel A: Summary Statistics

	Count	Mean	SD	Min	Max
Sim_Cosine	349513	0.8582	0.2118	0.0004	.9999
Sim_Jaccard	349513	0.4234	0.1957	0.0001	.9950
Sim_MinEdit	349513	0.3846	0.1881	0.0000	.9993
Sim_Simple	332821	0.1247	0.1157	0.0000	.9966

- Note that the cosine measure is effectively weighted by the word count; the Jaccard measure is not.
  - The average Jaccard measure is not particularly high.

Similarity Measures — Examples

The authors propose the following example:

- D_A: We expect demand to increase.
- D_B: We expect worldwide demand to increase.
- D_C: We expect weakness in sales.

A and B are similar, based on the cosine and Jaccard measures, and B and C are not.

Similarity Measures — Examples

The authors propose the following example:

- D_A: We expect demand to increase.
- D_B: We expect worldwide demand to increase.
- D_C: We expect weakness in sales.

A and B are similar, based on the cosine and Jaccard measures, and B and C are not.

However, what about word changes that affect meaning?:

- D_D: We expect weakness in our competition.

Similarity Measures — Examples

The authors propose the following example:

- D_A: We expect demand to increase.
- D_B: We expect worldwide demand to increase.
- D_C: We expect weakness in sales.

A and B are similar, based on the cosine and Jaccard measures, and B and C are not.

However, what about word changes that affect meaning?:

- D_D: We expect weakness in our competition.

or negation?

- D_E: We don't expect demand to increase

It would be nice to have some reassurance that a lack of similarity measures is really capturing changes in the meaning.

- (compare to human-scored 10-K's?).

Figure 1: NetApp Similarity over time

This figure plots the similarity score of NetApp, Inc. from 1996 to 2014.

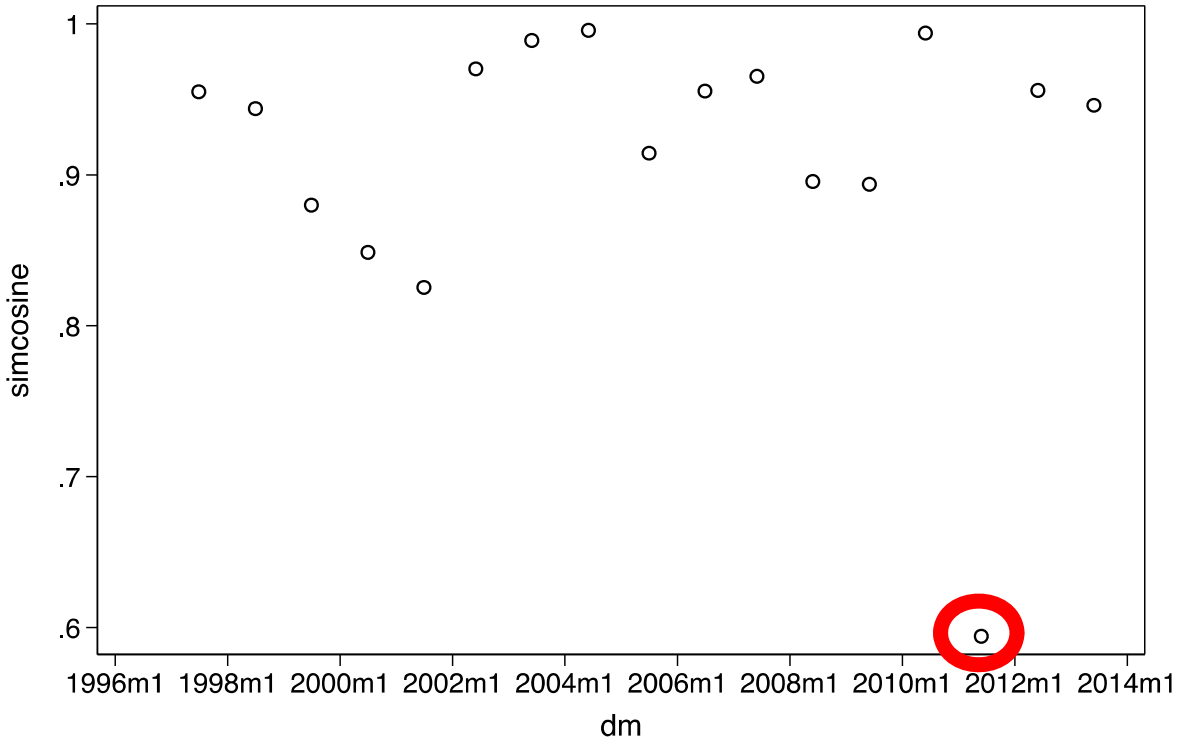
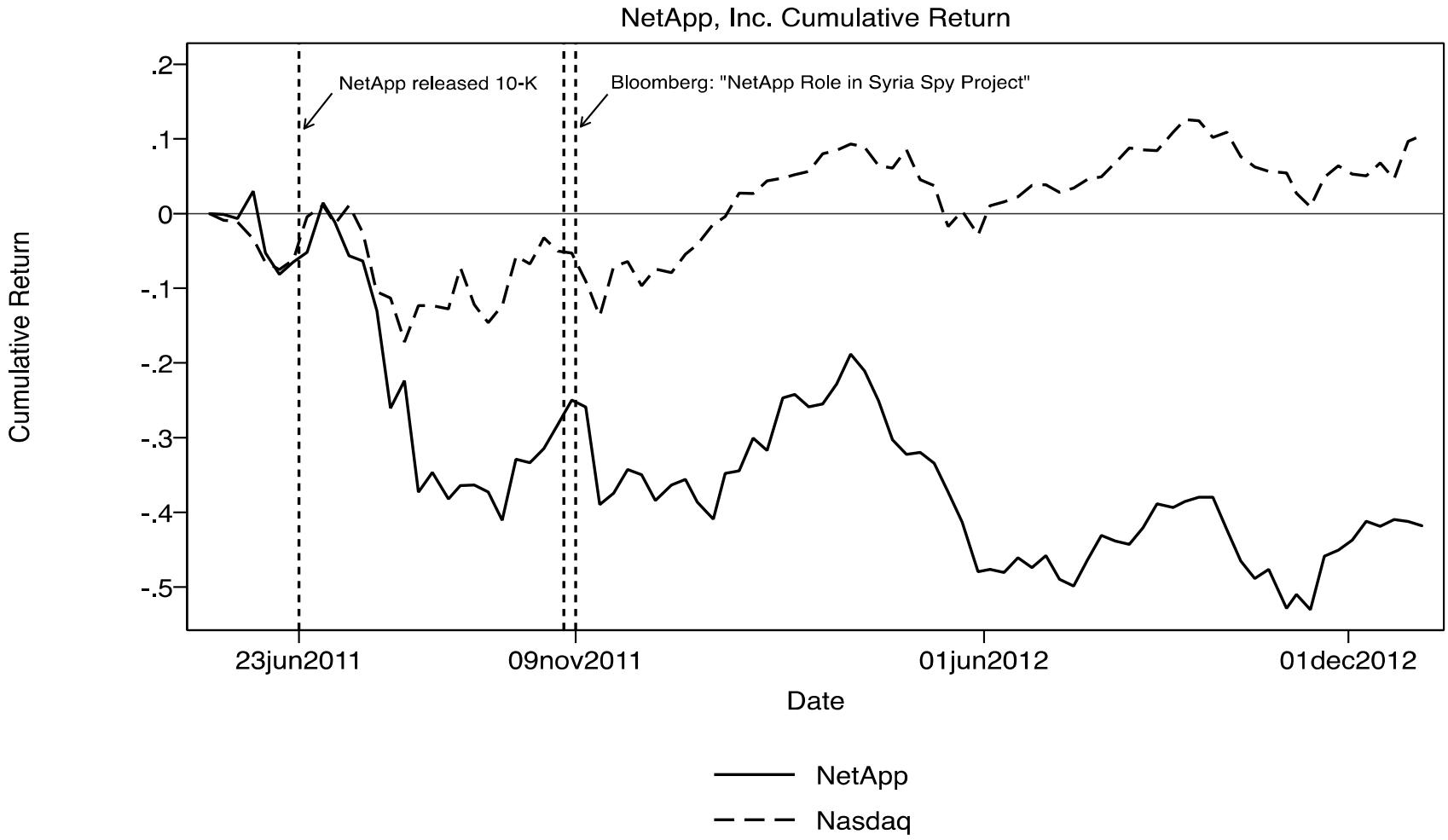


Figure 4: NetApp Returns



NetApp Example: 2010-2011 10-Ks

2010:

In addition, selling our products to the U.S. government also subjects us to certain regulatory requirements. For example, in April 2009, we entered into a settlement agreement with the United States of America, acting through the United States Department of Justice (DOJ) and on behalf of the General Services Administration (the GSA), under which we paid the United States $128.0 million, plus interest of $0.7 million, related to a dispute regarding our discount practices and compliance with the price reduction clause provisions of GSA contracts between August 1997 and February 2005. The failure to comply with U.S. government regulatory requirements could subject us to fines and other penalties, which could have a material adverse effect on our revenues, operating results and financial position.

2011:

In addition, selling our products to the U.S. government, whether directly or indirectly, also subjects us to certain regulatory requirements. For example, in April 2009, we entered into a settlement agreement with the United States of America, acting through the United States Department of Justice (DOJ) and on behalf of the General Services Administration (the GSA) related to a dispute regarding our discount practices and compliance with the price reduction clause provisions of GSA contracts for certain specified prior years. Failure to comply with U.S. government regulatory requirements by us or our reseller partners could subject us to fines and other penalties, which could have a material adverse effect on our revenues, operating results and financial position.

NetApp Example: 2011-2012 10-Ks

2011:

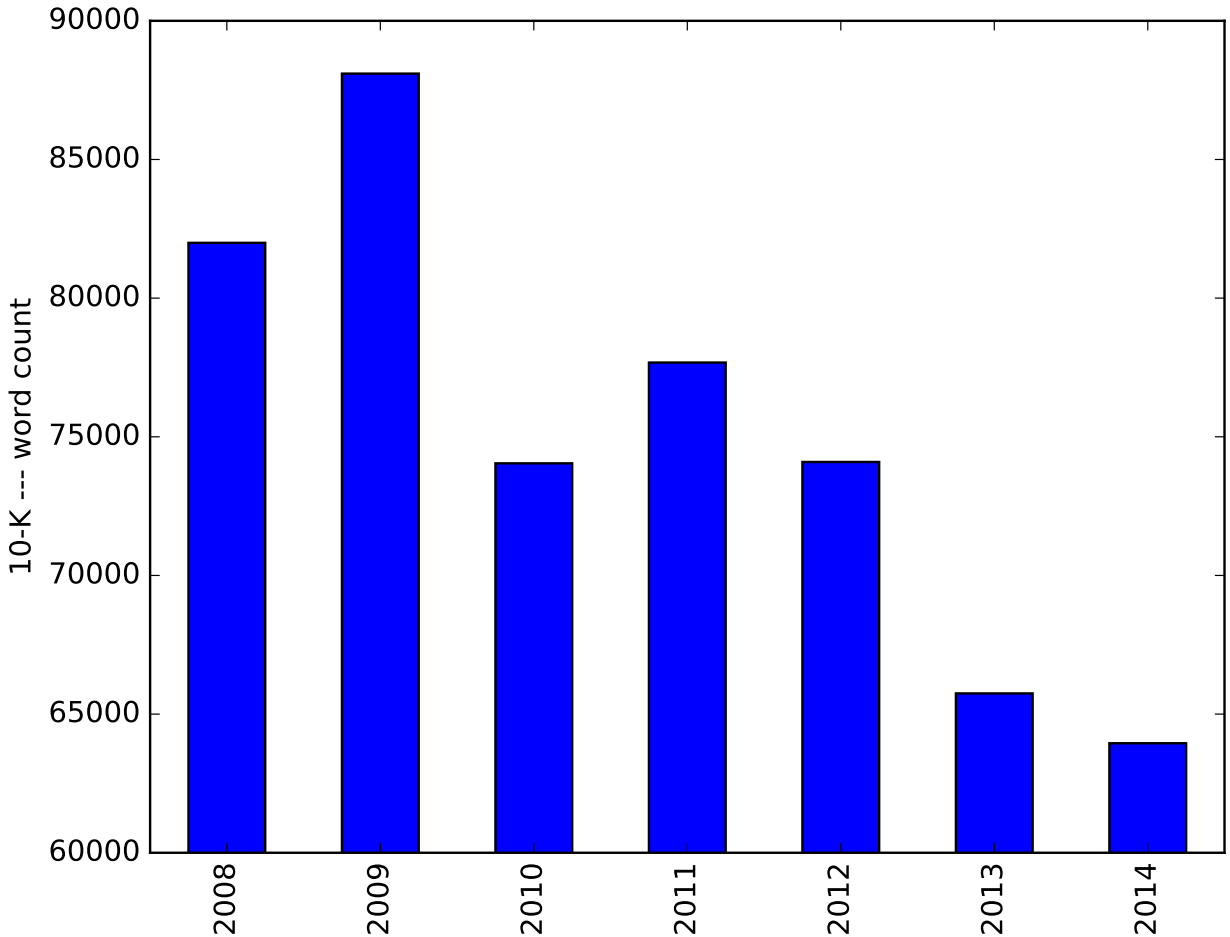
In addition, selling our products to the U.S. government, whether directly or indirectly, also subjects us to certain regulatory requirements. For example, in April 2009, we entered into a settlement agreement with the United States of America, acting through the United States Department of Justice (DOJ) and on behalf of the General Services Administration (the GSA) related to a dispute regarding our discount practices and compliance with the price reduction clause provisions of GSA contracts for certain specified prior years. Failure to comply with U.S. government regulatory requirements by us or our reseller partners could subject us to fines and other penalties, which could have a material adverse effect on our revenues, operating results and financial position.

2012:

Selling our products to the U.S. government, whether directly or through channel partners, also subjects us to certain regulatory and contractual requirements. Failure to comply with these arequirements by either us or our channel partners could subject us to investigations, fines, and other penalties, which could have a material adverse effect on our revenues, operating results and financial position. As an example, the United States Department of Justice (DOJ) and the General Services Administration (GSA) have in the past pursued claims against and financial settlements with IT vendors, including us and several of our competitors and channel partners, under the False Claims Act and other statutes related to pricing and discount practices and compliance with certain provisions of GSA contracts for sales to the federal government. The DOJ and GSA continue to pursue actively such claims. We are currently discussing contract compliance matters regarding sales made through a channel partner with the DOJ and GSA, and have produced documents and met with the DOJ and GSA on several occasions. If the DOJ determines to initiate an action against a channel partner and/or us, we would be subject to litigation, could be subjected to fines and penalties. We could also decide to pay the DOJ a settlement, either to avoid a potential action or in termination of an action. Violations of certain regulatory and contractual requirements could also result in us being suspended or debarred from future government contracting. Any of these outcomes could have a material adverse effect on our revenues, operating results and financial position.

NetApp Example: 10-K Word Counts

- 10-K word counts vary fairly dramatically from year-to-year.



Key Results – Table III.B

Panel B: Value Weighted

Sim_Cosine

	Q1	Q2	Q3	Q4	Q5	Q5 - Q1
Excess Return	0.0040	0.0044	0.0051	0.0079**	0.0078**	0.0038***
	(1.2095)	(1.3085)	(1.6391)	(2.5627)	(2.3629)	(2.7547)
3-Factor Alpha	-0.0018**	-0.0019**	-0.0007	0.0018**	0.0019*	0.0037***
	(-2.0280)	(-2.1017)	(-0.7910)	(1.9748)	(1.7411)	(2.7024)
5-Factor Alpha	-0.0013	-0.0021**	-0.0009	0.0021**	0.0021*	0.0034**
	(-1.4101)	(-2.2624)	(-1.0640)	(2.3542)	(1.9115)	(2.3996)

Sim_Jaccard

	Q1	Q2	Q3	Q4	Q5	Q5 - Q1
Excess Return	0.0015	0.0055	0.0072**	0.0072**	0.0076**	0.0061***
	(0.4459)	(1.6504)	(2.2260)	(2.3058)	(2.5168)	(3.9898)
3-Factor Alpha	-0.0046***	-0.0005	0.0012	0.0013	0.0018*	0.0063***
	(-4.8741)	(-0.4956)	(1.1990)	(1.3893)	(1.6714)	(4.4578)
5-Factor Alpha	-0.0044***	-0.0004	0.0014	0.0012	0.0019*	0.0063***
	(-4.5642)	(-0.3962)	(1.4451)	(1.2487)	(1.8656)	(4.4351)

Sim_MinEdit

	Q1	Q2	Q3	Q4	Q5	Q5 - Q1
Excess Return	0.0036	0.0043	0.0068**	0.0077**	0.0077***	0.0041**
	(1.0609)	(1.2900)	(2.0867)	(2.5586)	(2.6093)	(2.4051)
3-Factor Alpha	-0.0025***	-0.0018*	0.0007	0.0020**	0.0020*	0.0045***
	(-2.8874)	(-1.8498)	(0.7883)	(2.1000)	(1.8087)	(3.0695)
5-Factor Alpha	-0.0021**	-0.0016	0.0009	0.0020**	0.0012	0.0033**
	(-2.4416)	(-1.6325)	(1.1168)	(2.1022)	(1.0502)	(2.2778)

Sim_Simple

	Q1	Q2	Q3	Q4	Q5	Q5 - Q1
Excess Return	0.0024	0.0061*	0.0077**	0.0078**	0.0074**	0.0050***
	(0.6879)	(1.8821)	(2.4476)	(2.5284)	(2.4775)	(2.6924)
3-Factor Alpha	-0.0039***	0.0002	0.0018*	0.0019*	0.0019	0.0058***
	(-3.8893)	(0.1802)	(1.8704)	(1.8797)	(1.4452)	(3.5865)
5-Factor Alpha	-0.0036***	0.0005	0.0018*	0.0018*	0.0015	0.0051***
	(-3.4960)	(0.6607)	(1.7835)	(1.7139)	(1.1461)	(3.1419)

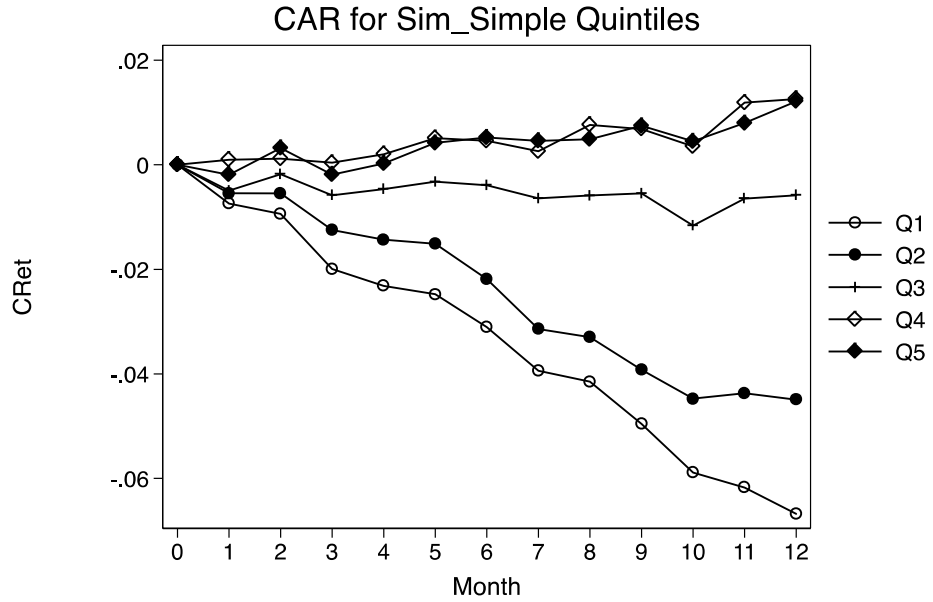
- The α is roughly consistent with a mkt loading of 1 ($\bar{r^e_m} \approx$ 67bps/mo.), and zero loadings on other factors. Is this right?
  - Any other result would be suspicious.

Asymmetry

- One thing that is striking about the return predictability is the strong asymmetry.
  - Changes in the 10-K's and 10-Q's reflect negative information, not positive.
- Why?
  - One possibility is that firm managers immediately and publically announce good news, but delay directly announcing potential bad news,
    - However, they shield themselves from the potential legal liabilities associated with failing to report risks.
  - This is consistent with the predictability associated with changes in the *Risks* section (See Table Vi.B)
- Again, it would be good to see more direct evidence on this.

Delayed information incorporation

"The returns [to the LS portfolio] continue to accrue out to 18 months ... implying that these return movements ... reflect true, fundamental changes to firms that only get gradually incorporated into asset prices over the 12-18 months after the reporting change." (p.



- I like this result (see, e.g., Daniel and Titman (2006)) but I'm not sure I trust it yet.

Delayed Fundamental Information Incorporation?

- The authors' basic story suggests that DTP information is incorporated slowly
- An additional (as yet untested) empirical implication of this hypothesis is that one be observe future negative returns around new information releases for Q1 firms, and not elsewhere.
  - For example, most of the negative returns associated with the Q1 portfolio should be around EADs.
    - Is this the case?
    - This effect should be especially true for small cap firms.

Additional Empirical Questions:

- This paper analyzes the "similarity" of the text to the corresponding 10-K/10-Q from 1-year/4-quarters before.
  - Is there a seasonality in the 10-Q format?
  - If not, why not use quarter-over-quarter change? This would presumably result in greater power.
- What do we know about the reaction to the filing of 8-Ks?

Conclusions

- Really great idea
- Apparently strong return predictability from the authors' similarity metric.
- The mechanism generating the predictability may be what the authors propose, but this is not yet clear.

References I

Daniel, Kent D., and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605–1643.

Hirshleifer, David, and Siew Hong Teoh, 2003, Limited attention, information disclosure, and financial reporting, *Journal of Accounting and Economics* 36, 337–386.

Sims, Christopher A., 2003, Implications of rational inattention, *Journal of Monetary Economics* 50, 665–690.