Capital IQ
A Standard & Poor's Business

QUANTITATIVE RESEARCH / MARCH 2011

How much Alpha is in Preliminary data?

Illuminating the differences between Prelim and Final filings

Authors

Fei He, PhD
Quantitative Analyst
312-233-7150
fhe@capitaliq.com

James Osiol
Quantitative Associate
312-233-7128
josiol@capitaliq.com

Dave Pope, CFA
Director of Quantitative Research
617-530-8112
dpope@capitaliq.com

"In theory there is no difference between theory and practice. In practice there is." -Yogi Berra

Companies often report financials twice: first, through a preliminary press release and again in their official, i.e., final, SEC filings. In theory, there should be no difference between the numbers reported in a company's preliminary financial filings and their final filings with the SEC. In practice, often significant difference can occur between the preliminary and final filings. In this month's research report, we focus on these observed differences within the Capital IQ Point-In-Time database in order to ascertain the nature and exploitability of these differences. We find that:

- A simple event based trading strategy evaluating the differences between Preliminary and Final data for Diluted EPS and Net Income generates an annualized excess return of +22% for long positions and +24% for short positions
- Incorporating preliminary data into a simple Earnings Yield (EP) strategy improved the factor's monthly returns by ~0.2% (20 basis points) over a traditional final only EP factor
- Diluted EPS shows significant differences between Q4 and non-Q4 preliminaries and finals, i.e., Q1, Q2, and Q3. Additionally, we find that certain sectors have significantly more instances of higher preliminary observations.
- Unlike EPS, we find no discernable patterns in Net Income by sector or size when comparing preliminary and final filings.
- Repeat offenders, or companies that repeatedly report large differences between preliminary and final numbers, come in all shapes and sizes with underlying causes of their offenses being fairly independent across companies.
- We find larger companies, by market capitalization, and Utilities are more likely to be repeat offenders vis-à-vis smaller firms and/or other GICS sectors.

Acknowledgements:
The authors would like to thank R.G. Associates owner and 'The Analyst's Accounting Observer' publisher Jack Ciesielski

1 Initial Event Study Research

We start with a simple question. When there are differences between preliminary (prelims) and final filings is the market surprised? To explore this question we conducted an event study focusing on the relative magnitude of the difference between prelims and final filings.

We categorized the difference between prelim and final data into either "good news" or "bad news" based on the direction of the difference. Investors enter a long position if the final filing's data is higher than the preliminary data, i.e., "good news" is reported. Conversely, investors enter a short position after "bad news" is reported, or when the final data is lower than the prelim data. We defined this relative difference as the percentage change between the preliminary and final filing's data:

$$RelDiff = \frac{Data_{Prelim} - Data_{Final}}{ABS(Data_{Final})}$$

Using relative differences allows us to meaningfully compare differences across stock issues regardless of their initial magnitude. The ClariFI Event Study module was used to generate our results. Our events are classified as good or bad as follows:

Good news: $RelDiff < -20\%$;
Bad news: $RelDiff > 100\%$;

The thresholds for the good and bad news were chosen to make the number of long and short positions symmetric. Additionally, it reflects the fact that it is generally more difficult to short than long. Various thresholds were tested in both directions to ensure that our results were generally robust and not dependent on the parameters. The relative differences of -20% and 100% provides a more symmetric exposure to both long and short positions, e.g., number of events, returns, IRs.

Initially, we looked at Net Income events characterized by these news indicators. We hold positions for 20 business days to limit the number of potential interactions with other earnings events such as a press release for the next quarter. 20 days is chosen because it coincides well with the life cycle of the event in practice. The return starts slowly (0 to 5 days), gains momentum (5 to 15 days), and finally saturates (15 to 20 days) as seen in Figure 1. We find ancillary evidence for this short term hypothesis for earnings events in our Data Enhancement section. Specifically, we find that the gain from incorporating preliminary data into simple factor construction has a statistically significant short term impact on returns.

The results from these tests align with our hypothesis that Good (Bad) news results in positive (negative) returns, Figure 1. Unfortunately, the results are not particularly strong, and were significant only on the short side. However, we find that where the observed differences occur in both Net Income and Diluted EPS, the results were greatly improved.

Table 1 and Table 2 detail our findings for Net Income events exclusively. The long and short positions have annualized gains of 15% and 24% respectively. In addition to Information Ratio (IR), we also list the Sortino Ratio (SR) for reference. The SR is a risk-adjusted performance measure which only penalizes for downside risk/volatility. The observed difference between SR_{Long} and IR_{Long} suggests the return time series is slightly positively skewed with higher volatilities coming from positive returns.

Table 1: Statistics for Good News Events from Net Income

Russell 3000, 7/2002 to 11/2010

Days	Good NI	Std	Hit Ratio	t-Stat	p-Value	Ann. Gain	Ann. Std	SR	IR
0	0.000	-	-	-	-	-	-	-	-
5	0.000	0.06	0.49	0.04	0.97	0.7%	0.43	0.03	0.02
10	0.003	0.08	0.48	0.56	0.58	7.5%	0.40	0.30	0.19
15	0.011	0.10	0.51	1.53	0.13	17.7%	0.42	0.72	0.42
20	0.012	0.12	0.51	1.49	0.14	14.8%	0.41	0.58	0.36

Table 2: Statistics for Bad News Events from Net Income

Russell 3000, 7/2002 to 11/2010

Days	Bad NI	Std	Hit Ratio	t-Stat	p-Value	Ann. Gain	Ann. Std	SR	IR
0	0.000	-	-	-	-	-	-	-	-
5	(0.000)	0.11	0.50	0.04	0.97	(1.4%)	0.77	(0.02)	(0.02)
10	(0.010)	0.12	0.57	1.36	0.17	(24.9%)	0.59	(0.62)	(0.42)
15	(0.017)	0.13	0.59	2.24	0.03	(29.3%)	0.51	(0.89)	(0.57)
20	(0.019)	0.14	0.61	2.22	0.03	(24.2%)	0.50	(0.77)	(0.49)

Figure 1: Compounded Excess Returns for Net Income Event

Russell 3000, 7/2002 to 11/2010

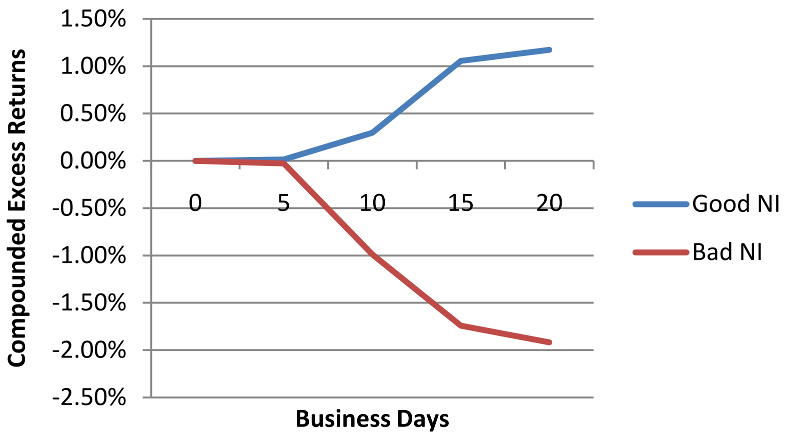


Figure 2: Compounded Excess Returns for DEPS Event

Russell 3000, 7/2002 to 11/2010

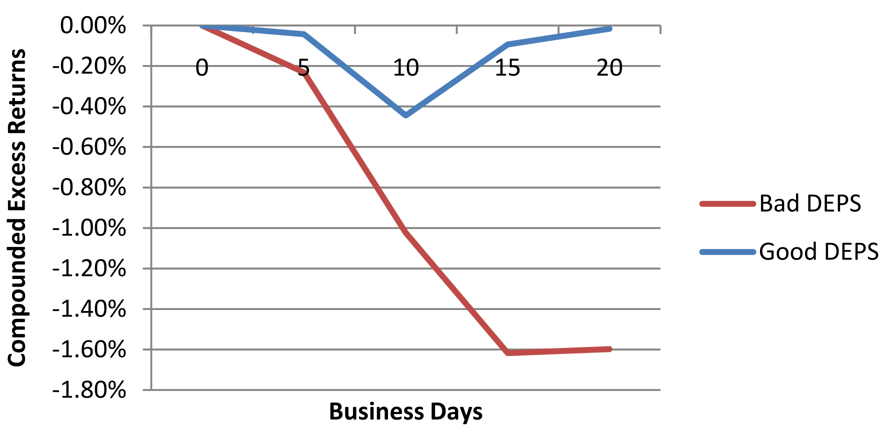


Table 4 similarly displays the detailed statistics for both long and short positions for Diluted EPS (DEPS). We follow the same conventions as with Net Income. We find that bad news results in negative returns of roughly the same magnitude as bad news Net Income events, however the relationship does not hold for good news.

Table 3 Statistics for the good DEPS (lower preliminary)

Russell 3000, 7/2002 to 11/2010

Days	Good DEPS	Std	Hit Ratio	t-Stat	p-Value	Ann. Gain	Ann. Std	SR	IR
0	0.000	-	-	-	-	-	-	-	-
5	(0.000)	0.06	0.49	0.13	0.89	(2.1%)	0.43	(0.08)	(0.05)
10	(0.004)	0.08	0.47	1.05	0.30	(11.2%)	0.41	(0.37)	(0.27)
15	(0.001)	0.10	0.50	0.18	0.85	(1.6%)	0.40	(0.06)	(0.04)
20	(0.000)	0.11	0.50	0.03	0.98	(0.2%)	0.40	(0.01)	(0.01)

Table 4 Statistics for the bad DEPS (higher preliminary)

Russell 3000, 7/2002 to 11/2010

Days	Bad DEPS	Std	Hit Ratio	t-Stat	p-Value	Ann. Gain	Ann. Std	SR	IR
0	0.0000	-	-	-	-	-	-	-	-
5	(0.002)	0.10	0.50	0.41	0.68	(11.7%)	0.71	(0.21)	(0.16)
10	(0.010)	0.11	0.56	1.64	0.10	(25.8%)	0.56	(0.69)	(0.46)
15	(0.016)	0.12	0.57	2.41	0.02	(27.2%)	0.49	(0.89)	(0.56)
20	(0.016)	0.14	0.57	2.06	0.04	(20.1%)	0.49	(0.64)	(0.41)

We find that when good (bad) differences for both items occur in unison, the results are improved. The compounded excess returns are shown in Figure 3 for both long (+22%) and short positions (-24.2%) for the full 20 days.

Table 5 and Table 6 presents detailed statistics for the respective long/short positions. The returns for long positions are improved, and are now statistically significant, along with almost all of our other performance statistics. Again, the observed difference between SR_{Long} and IR_{Long} suggests the return time series is positively skewed with more volatility coming from positive returns. Figure 3 provides a simple visual to illustrate impact to performance attributable to the event interactions. The majority of the observed improvement is realized in our long positions.

Table 5 Statistics for the good news interactions between NI and DEPS

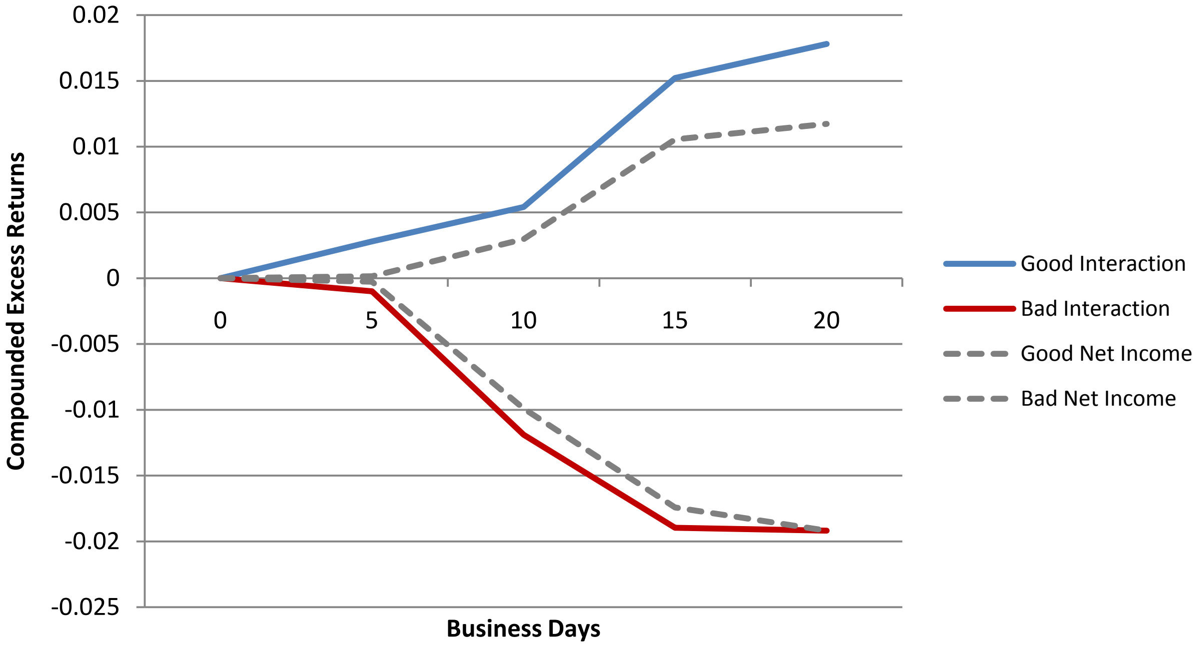
Russell 3000, 7/2002 to 11/2010

Days	Good Interac.	Std	Hit Ratio	t-Stat	p-Value	Ann. Gain	Ann. Std	SR	IR
0	0.000	-	-	-	-	-	-	-	-
5	0.003	0.06	0.50	0.65	0.52	14.1%	0.43	0.57	0.33
10	0.005	0.08	0.49	0.96	0.34	13.6%	0.40	0.56	0.34
15	0.015	0.10	0.52	2.10	0.04	25.6%	0.42	1.10	0.61
20	0.018	0.12	0.52	2.17	0.03	22.4%	0.41	0.93	0.54

Table 6 Statistics for the bad news interactions between NI and DEPS
Russell 3000, 7/2002 to 11/2010

Days	Bad Interac.	Std	Hit Ratio	t-Stat	p-Value	Ann. Gain	Ann. Std	SR	IR
0	0.0000	-	-	-	-	-	-	-	-
5	(0.001)	0.11	0.57	0.14	0.89	(5.0%)	0.78	(0.08)	(0.06)
10	(0.012)	0.12	0.64	1.59	0.11	(30.0%)	0.59	(0.74)	(0.51)
15	(0.019)	0.13	0.65	2.38	0.02	(31.9%)	0.52	(0.97)	(0.62)
20	(0.019)	0.14	0.66	2.17	0.03	(24.2%)	0.50	(0.77)	(0.49)

Figure 3: Improvement from Event Interactions
Russell 3000, 7/2002 to 11/2010



2 Factor Enhancement

In this section, we assess the performance impact of incorporating preliminary data into alpha factors as opposed to waiting until the data from the final filing becomes available. We are aware that most investors utilize preliminary data, but we wanted to quantify the incremental advantage to using the immediately available information.

To demonstrate, we define a simple factor EP:

$$EP = \frac{DEPS_{i,t=0}}{P_{i,t=0}}$$

where *DEPS* is the Diluted EPS, P is the monthly price for stock issue *i* and *t=0* to signify the most recent earnings and price information available at month end. We believe this factor should provide meaningful comparison between potential normal and enhanced factors due to its simplicity. With this comparison in mind, the following results should be viewed not for absolute performance but for incremental improvements.

Incorporating preliminary data into simple factor construction significantly improves short term performance.

Table 7: E/P IC Statistics, Before and After Preliminary Inclusion

Russell 3000, 7/2002 to 11/2010

Holding Period	Avg. IC Before	Avg. IC After	IC Diff	IC Diff T-Stat.	p-Value
1 Month	0.027	0.032	0.005	3.82	0.00
3 Months	0.031	0.033	0.002	1.32	0.19
6 Months	0.042	0.043	0.001	0.92	0.36
12 Months	0.032	0.032	0.000	0.26	0.79

Table 8: E/P Top-Bottom Statistics, Before and After Preliminary Inclusion

Russell 3000, 7/2002 to 11/2010

Holding Period	TB Spread Before	TB Spread After	Diff	t-Stat	p-Value
1 Month	(0.14%)	0.02%	0.17%	3.32	0.00
3 Months	(0.65%)	(0.60%)	0.05%	0.30	0.76
6 Months	(2.05%)	(2.11%)	(0.06%)	0.19	0.85
12 Months	(7.96%)	(8.39%)	(0.43%)	0.91	0.37

Table 7 and Table 8 compare the performance of the traditional and enhanced EP factors on an IC and return basis. Our traditional factor utilizes final filing's data, and our enhanced factor incorporates prelim data in addition to final filing data. The prelim data is incorporated at the monthly rebalance when the corresponding final has not yet been released. In this scenario, the preliminary represents the most recent data available for the given security.

Our subtle variations on an EP theme show distinct performance differences. The enhanced factor outperforms for both IC and return measures for short term holding periods. The performance improvement is particularly pronounced for a 1 month holding period. We believe this may be due to the fact that at longer holding periods final data would have become available, and there is already a slight lag inherent to this backtest in rebalancing at month-end.

Due to the difference in the 1 month holding period, we then examined the relative quintile difference at a 1 month horizon. As Table 9 demonstrates, we found statistically significant improvements in monthly return and IC. Furthermore, detailed quintile level analysis shows that the improvement to the spread is attributable to both the long and short side, each contributing approximately 0.1% a month to our spread. The inclusion of preliminary data factor produces a positive return improvement 64% of the time. We find the improvement is robust through time. Since 2002, the data enhancement provided a positive impact on average 1-Month return in 8 of the 9 years, see Figure 4.

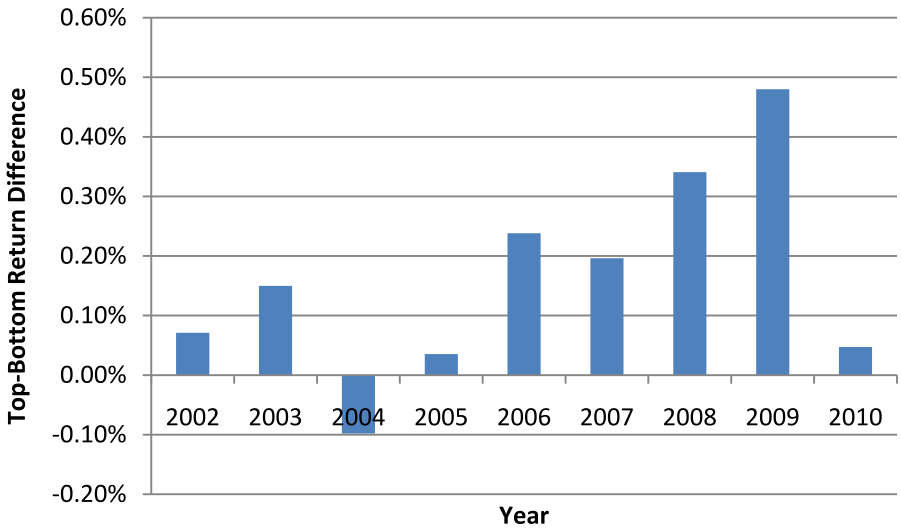
Table 9: E/P Improvement Drill Down for 1-Month Holding

Russell 3000, Aggregate Monthly Test of Differences from 7/2002 to 11/2010

Stat.	TB Diff	IC Diff	Q1 Diff	Q2 Diff	Q3 Diff	Q4 Diff	Q5 Diff
Mean	0.002	0.005	0.001	0.000	0.000	(0.000)	(0.001)
Std	0.005	0.013	0.003	0.002	0.002	0.002	0.003
t-Stat	3.32	3.82	3.26	0.66	0.40	1.84	2.47
p-Value	0.00	0.00	0.00	0.51	0.69	0.07	0.02
Hit Ratio	0.64	0.59	0.64	0.45	0.52	0.44	0.38

Figure 4: Data Enhancement, Average 1-Month Return Improvement by Years

Russell 3000, 7/2002 to 11/2010

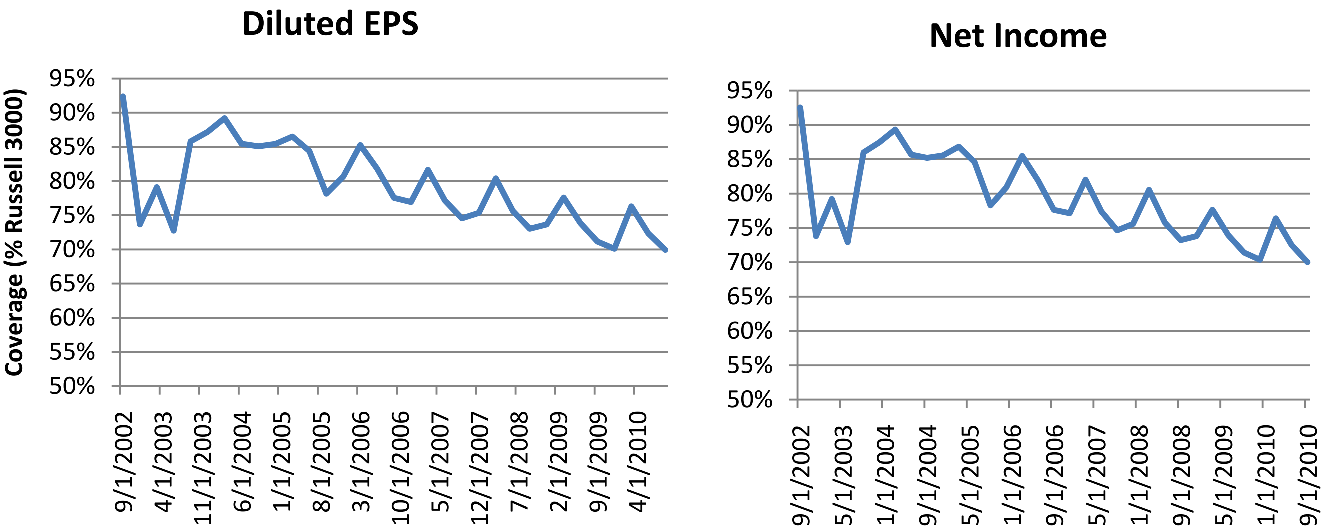


3 The Characteristics of Preliminary Data

Effective August 23, 2004, the Securities and Exchange Commission requires all public companies to file an 8-K within five business days of their earnings release (Securities and Exchange Commission, 2004). Prior to 2002, Capital IQ had been collecting preliminary data from official 8-K filings and other press releases. Unfortunately, due to the lack of formal requirements, the historical coverage for these documents is fairly sparse. Consequently, we focus on the period post Sarbanes-Oxley where, in tandem with new SEC rules in 2004, we see drastically improved reporting practices resulting in greater coverage.

The scope of our research has been limited to two data items, Net Income and Diluted Net EPS. These metrics are consistently reported in preliminary releases, leading to sufficient coverage. We show the Capital IQ Point-In-Time database coverage over time for the Russell 3000 in Figure 5. We compared the data counts to the total number of possible companies in our universe for each quarter. Overall, the preliminary data coverage for both items is very similar fluctuating between 70% and 90%. We do not expect 100% coverage because the database does not capture a preliminary value if the company files what would be considered the preliminary press release on the same day as their final filing. Companies are generally reducing the number of filing days between prelim and final. This contributes to the downward trend in coverage as more companies file their press release/8K on the same day as their final filling.

Figure 5: Diluted EPS and Net Income Coverage
Quarterly coverage of Russell 3000, 7/2002 to 11/2010



4 Prelim v. Final Comparison

The quality of the prelim data is naturally of concern to investors. Our research was performed using the Capital IQ PIT dataset which captures the all filings for its global dataset, i.e., both press releases/8-Ks and final filings. Given certain variations we found in companies and sectors, we then established a set of standards which would eliminate outliers and serve as the basis for our work. The following set of standards was applied in order to provide the most accurate representation of the data.

- **Date/Period Alignment through Fiscal Calendar Changes:** We aligned the prelim with the final filings to the same fiscal period and reporting period end, to adjust for possible misalignment surrounding changes in the fiscal calendar. This may seem self evident, but it is critical to ensure one is comparing preliminary and final filings for the same fiscal and calendar period.

- **Removal Trivial Differences:** We excluded trivial differences by aligning the number of significant figures between prelim and final. For example, if the Net Income from a prelim is $15.14 MM, but $15.135 MM from a final, we will round the final to the second digit and consider this case a trivial difference.

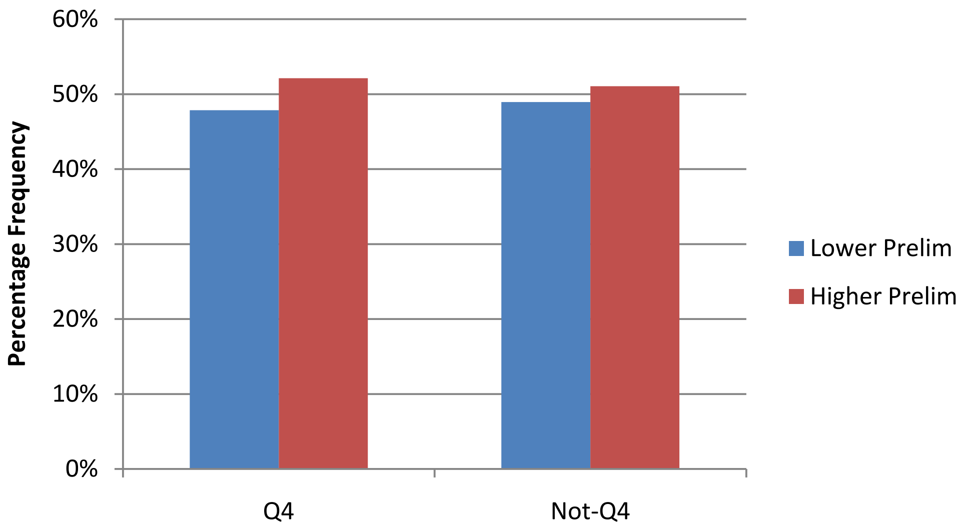
To start we compared preliminary and final filings data, absent any explicit partitions for the period 7/2002 to 11/2010. The following tables and figures outline the raw frequency and percentage of observed differences and their directions. As seen in Figures 6 and 7, Net Income and Diluted EPS both tend to have higher preliminary observations overall. Our results are consistent across fiscal quarters. Throughout the entire time period, approximately 5% of our observations show significant differences between the prelim and final for Net Income and Diluted EPS respectively.

We excluded differences between the preliminary and final that were less than 1 cent for Diluted EPS[1]. This exclusion acts as a noise filter, providing cleaner results and removes differences attributable to rounding errors of tangential data items. After applying this filter, we find about 5% observed differences greater than or equal to 1 cent for Diluted EPS, as show in Table 10.

Given the vastly different practices and regulations imposed on companies for Q4 filings, the most obvious being that the Q4 final report must be audited, we break our observations into Fiscal Quarter 4 and Non Fiscal Quarter 4, i.e., Quarters 1, 2, and 3. .

Figure 6: Diluted EPS Difference Comparison

Differences over entire time horizon from 7/2002 to 11/2010, broken out by Quarter 4 and Non Quarter 4



[1] The one-cent difference is calculated based on non-split-adjusted Diluted EPS data.

Figure 7: Net Income Difference Comparison

Differences over entire time horizon from 7/2002 to 11/2010, broken out by Quarter 4 and Not Quarter 4

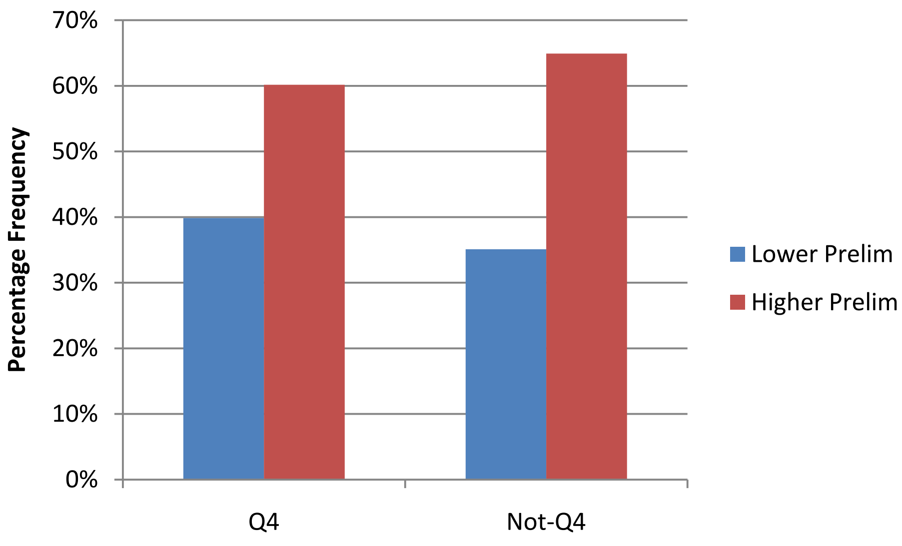


Table 10: Percentage of Large Differences in Diluted EPS and NI

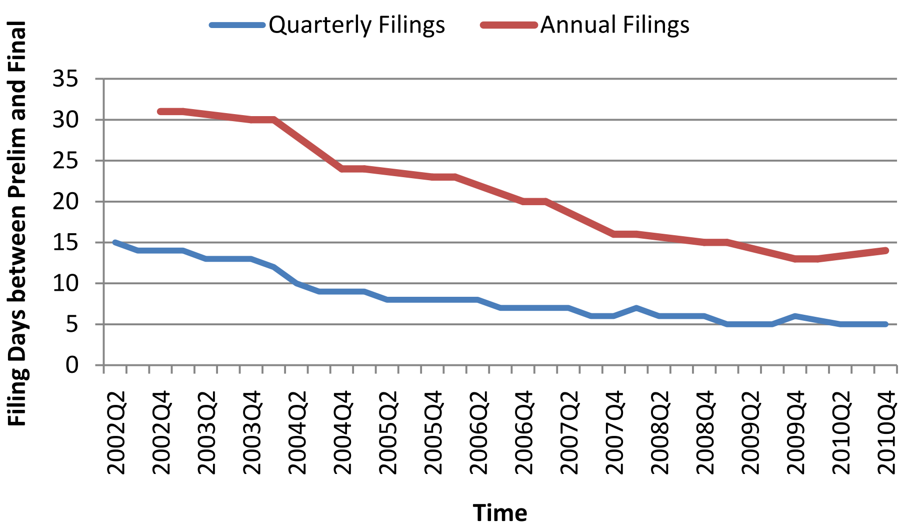
Percent differences over entire time horizon from 7/2002 to 11/2010

Difference	Diluted EPS[2]	Net Income
No Difference	94.98	94.56
With Difference	5.02	5.44

We noticed a very clear relationship between the number of days between filings through time. As seen in Figure 8, we find that the median date difference between prelim and final for annual reports has decreased from approximately 31 days in 2002, to approximately 15 days in 2010 in the Russell 3000. For quarterly reports, the median has decreased from approximately 15 days in 2002, to approximately 5 days in 2010.

Figure 8: Universe Median Days between Filings

Russell 3000, from 7/2002 to 11/2010



[2] For DEPS difference calculations, we excluded trivial difference less than one cent, which is likely to be introduced by rounding errors and has less significant economical meanings.

5 Prelim v. Final by Partitions

We then extended our analysis by examining differences between our preliminary and final filings based on different partitions of our universe in order to identify any systematic characterization/bias present in the differences between our aligned filings. Specifically, we analyzed differences by sector, size, and filing days.

The following analysis focuses solely on the subset of the universe where differences are present. We segmented our analysis by quarter and tested if our observed differences for non-Q4 are significantly different from Q4 (Table 11 and

Table 12). While we do not see a significant difference for Net Income across quarters, we find that there is significant difference across quarters for Diluted EPS. We do not identify an apparent difference in the relative frequency of higher prelim between Q1, Q2, and Q3. For consistency, we carry out subsequent analysis for both Q4 and non-Q4 for both items.

Table 11: ANOVA and Tukey Pairwise Multiple Comparison Test on Means Difference across Fiscal Quarters for NI

Russell 3000, Aggregated from 7/2002 to 11/2010

Overall ANOVA Statistics

Group Variable	DF	F Value	Pr > F
Fiscal Quarters	3	0.24	0.87

Pairwise Multiple Comparison Statistics with Controlled Type I Error

Fiscal Quarters	Means Difference	95% Confidence Limits		Significant at 0.05 Level
2 - 4	0.01	(0.06)	0.07	
2 - 3	0.02	(0.06)	0.10	
2 - 1	0.03	(0.06)	0.11	
4 - 3	0.01	(0.06)	0.07	
4 - 1	0.02	(0.05)	0.09	
3 - 1	0.01	(0.07)	0.10	

Table 12: ANOVA and Tukey Pairwise Multiple Comparison Test on Means Difference across Fiscal Quarters for DEPS

Russell 3000, Aggregated from 7/2002 to 11/2010

Overall ANOVA Statistics

Group Variable	DF	F Value	Pr > F
Fiscal Quarters	3	19.49	0.00

Pairwise Multiple Comparison Statistics with Controlled Type I Error

Fiscal Quarters	Means Difference	95% Confidence Limits		Significant at 0.05 Level
4 - 3	0.10	0.05	0.16	***
4 - 2	0.13	0.07	0.18	***
4 - 1	0.13	0.07	0.19	***
3 - 2	0.02	(0.04)	0.09	
3 - 1	0.03	(0.04)	0.10	
2 - 1	0.01	(0.06)	0.08	

BY SECTOR

When delineated along GICS sectors, we observed no obvious systematic tilts for the Net Income difference (Figure 9 and Figure 10). Each sector simply provides consistently higher prelims across quarters. Further ANOVA tests on the difference in means support these conclusions.

Figure 9: Net Income Difference Comparison by Sector, Q4

Russell 3000, Fiscal Quarter 4, GICS Sectors, Aggregated - 7/2002 to 11/2010

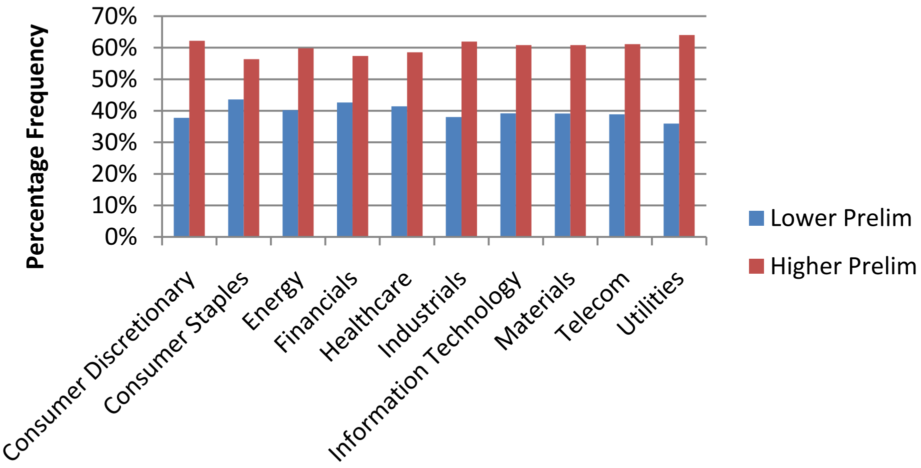
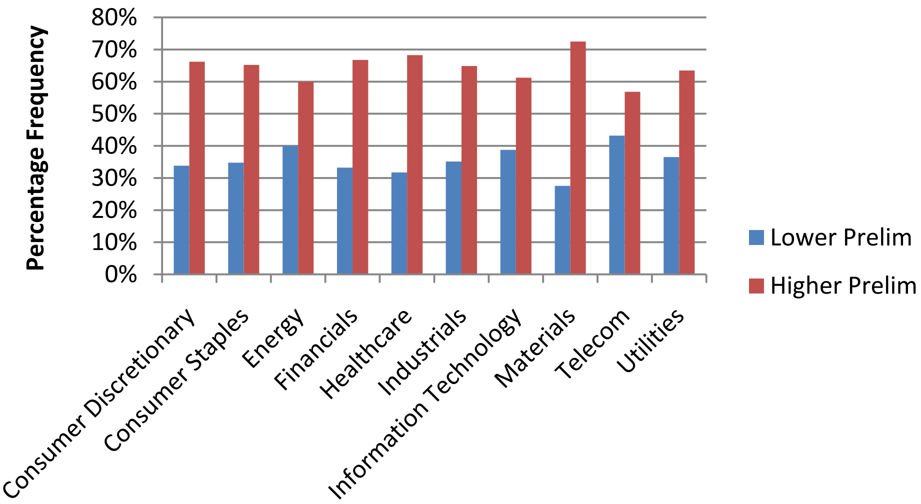


Figure 10: Net Income Difference Comparison by Sector, Not-Q4

Russell 3000, Not Fiscal Quarter 4, GICS Sectors, Aggregated - 7/2002 to 11/2010



In contrast, Diluted EPS is noticeably different between Q4 and non-Q4 data. Similar to Net Income, Q4 Diluted EPS showed consistently higher prelims for all sectors, Table 13. We believe that this relationship is being driven by differences in reporting requirements between Q4 and Non Q4. Since Q4 is audited, it is more likely that Q4 finals will come out less optimistic than prelims.

Figure 11: Diluted EPS Difference Comparison by Sector

Russell 3000, Fiscal Quarter 4, GICS Sectors, Aggregated from 7/2002 to 11/2010

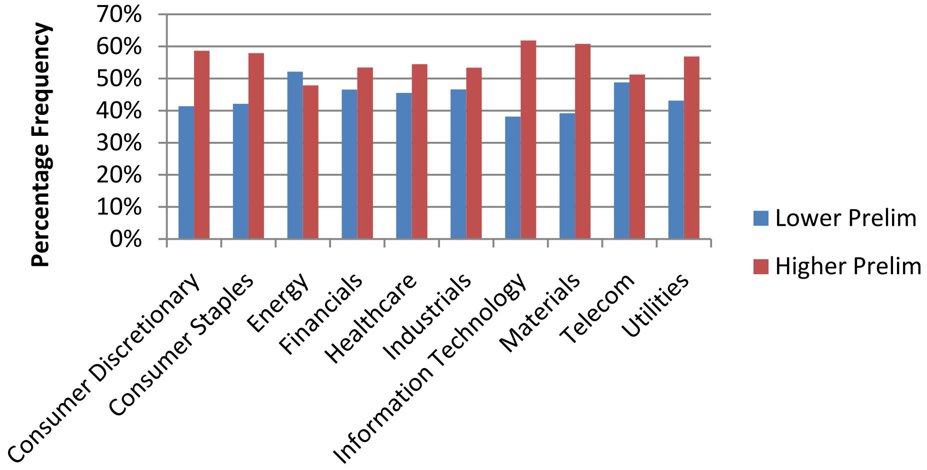


Figure 12: Diluted EPS Difference Comparison by Sector

Russell 3000, Not Fiscal Quarter 4, GICS Sectors, Aggregated from 7/2002 to 11/2010

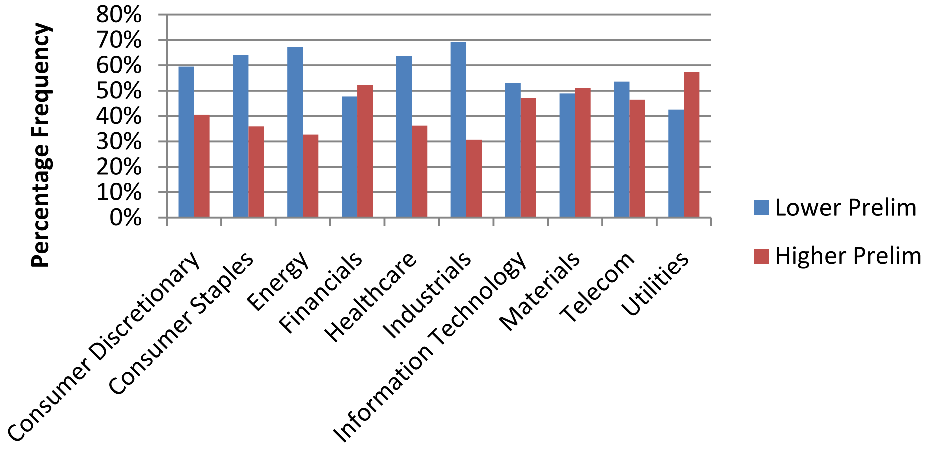


Table 13 ANOVA Test on Means Difference across Different Sectors for DEPS

Russell 3000,Q4, GICS Sectors, Aggregated -7/2002 to 11/2010

Overall ANOVA Statistics

Group Variable	DF	F Value	Pr > F
Sectors	9	1.32	0.22

Table 14 demonstrates that there are significant differences in DEPS across sectors. For the Tukey test, we only provide below the significant results due to the exhaustive nature of the pairwise comparisons. We find that Financials, Materials, and Information Technology have statistically significant higher percentage frequencies of higher prelims than all other sectors. The differences were particularly pronounced when compared to Energy, Industrials, Healthcare and Consumer Discretionary.

Table 14 ANOVA Test on Means Difference across Different Sectors for DEPS
Russell 3000, Non-Q4, GICS Sectors- 7/2002 to 11/2010
Overall ANOVA Statistics

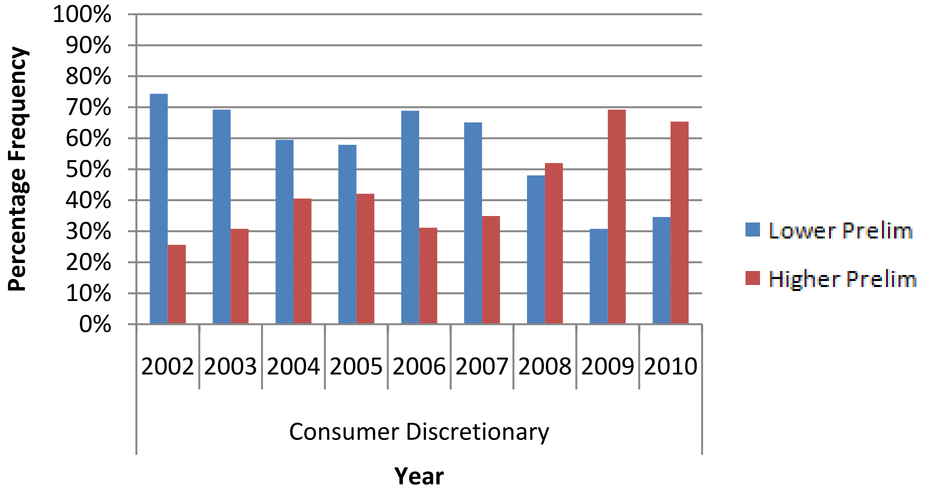
Group Variable	DF	F Value	Pr > F
Sectors	9	6.41	0.00

Pairwise Multiple Comparison Statistics with Controlled Type I Error

Fiscal Quarters	Means Difference	95% Confidence Limits		Significant at 0.05 Level
Financials – Cons. Disc.	0.12	0.02	0.23	***
Financials – Healthcare	0.15	0.03	0.27	***
Financials – Industrials	0.21	0.09	0.33	***
Financials – Energy	0.21	0.07	0.36	***
Materials – Industrials	0.21	0.03	0.38	***
Materials – Energy	0.21	0.02	0.40	***
Info. Tech. – Industrials	0.17	0.03	0.30	***
Info. Tech. – Energy	0.17	0.01	0.33	***

We then drilled down and looked at this relationship through time for the Consumer Discretionary sector with more lower-prelim instances. We see a fairly consistent trend, the lower prelim instances are skewed toward the early part of our sample. Figure 13 provides an example tracking the Consumer Discretionary sector.

Figure 13: Diluted EPS Difference Comparison for Consumer Discretionary
Russell 3000, Not Fiscal Quarter 4, Broken out by year - 7/2002 to 11/2010



By Size

When segmented by market capitalization, all of our capitalization groups showed higher prelims for Net Income (Figure 14 and Figure 15).

Figure 14: Net Income Difference Comparison by Market Capitalization
Russell 3000, Fiscal Quarter 4, Market Cap Deciles from 7/2002 to 11/2010

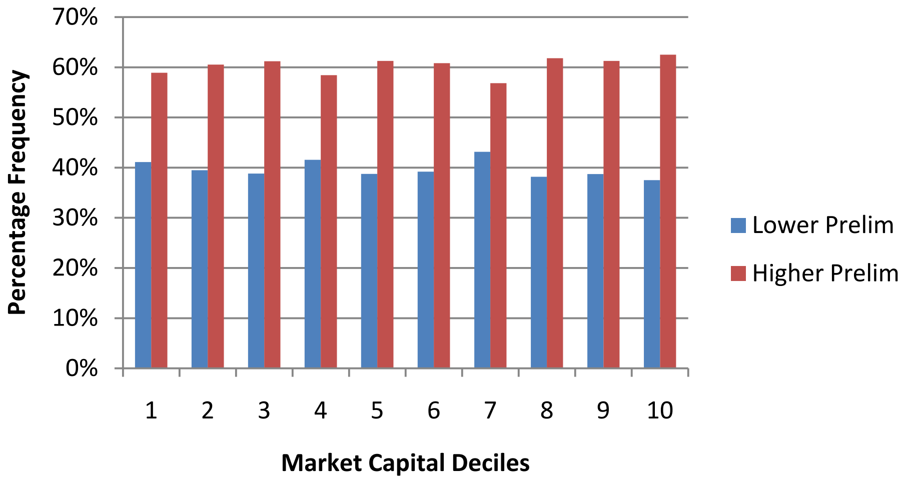
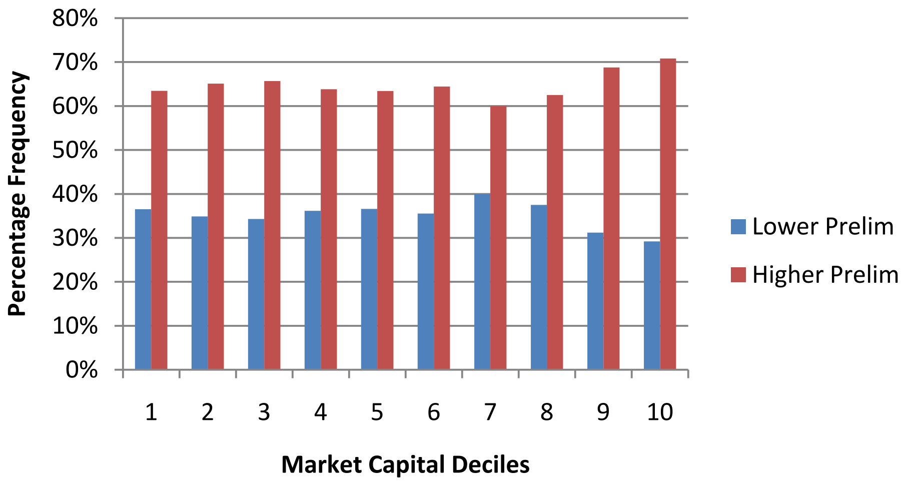


Figure 15: Net Income Difference Comparison by Market Capitalization
Russell 3000, Not Fiscal Quarter 4, Market Cap Deciles from 7/2002 to 11/2010



We found that small-cap companies are more likely to report higher Diluted EPS prelims relative to their finals, Figure 16 and
Figure 17**.** To formally test for statistical significance, we performed simple non-parametric sign tests for prelim final differences in small cap and other cap securities (Table 15). At the 5% level, the results are significant for small cap and inconclusive for other cap securities, confirming our hypothesis.

Table 15: Results of Sign test for Relative Difference, μ_0=0
Test of Directional Significance for Small and Large Cap Securities - 7/2002 to 11/2010

Size	Median	M	p-Value
Small Cap (<= 700 MM)	0	343	<.0001
Other Cap (> 700MM)	0	(95.5)	0.0614

Figure 16: Diluted EPS Difference Comparison by Market Capitalization
Russell 3000, Fiscal Quarter 4, Market Cap Deciles from 7/2002 to 11/2010

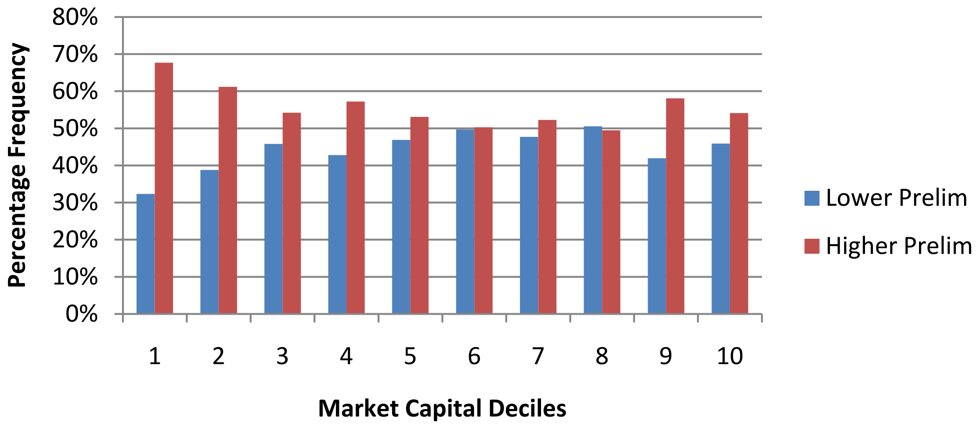
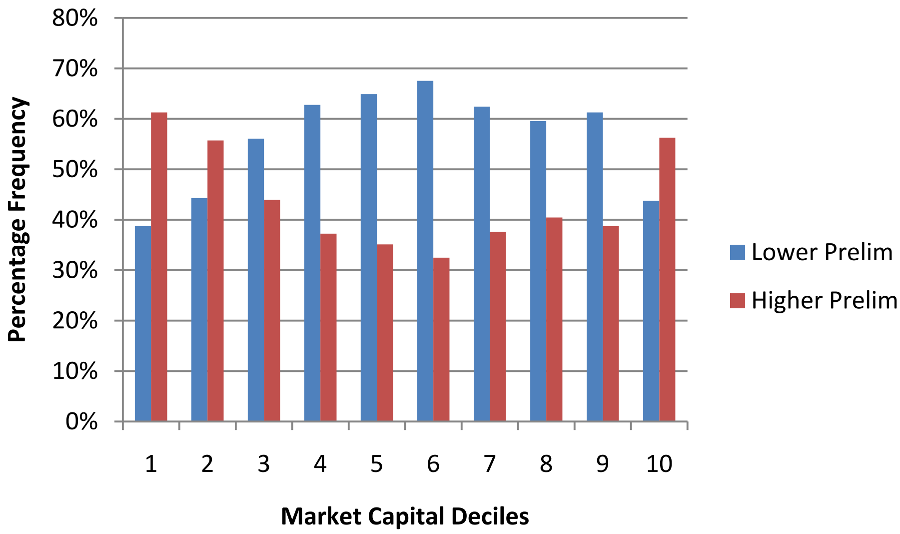


Figure 17: Diluted EPS Difference Comparison by Market Capitalization
Russell 3000, Not Fiscal Quarter 4, Market Cap Deciles from 7/2002 to 11/2010



We believe there are a couple plausible hypotheses to support the fact that small-cap companies report higher prelim Net Income. First, small-cap companies may have fewer resources at their disposal to dedicate to preliminary accounting, and tend to be overly optimistic. Second, certain small-cap companies may be trying to raise their profile by padding their preliminary earnings and providing a more accurate representation in the final with the full inclusion of balance sheet items.

6 Preliminary v. Final Repeat Offenders

We define repeat offenders as companies which display persistence in their reported differences between preliminary and final filings. Company level differences were tracked on a two year rolling basis. If a company has more than one recorded difference in the same direction during the trailing two years we categorized it as a "repeat offender". We believe this status may be a flag for potentially systematic aggressive/conservative accounting practices.

We were curious as to how pervasive this type of scenario was from 2002 to the present. Not surprisingly, the majority of companies are not repeat offenders. However, we do find that companies are twice as likely to be a "bad" repeat offender, i.e., higher prelim than final, than a good offender, i.e., lower prelim than final.

Table 16 Percentage Repeat Offenders in the Universe

Russell 3000 from 7/2002 to 11/2010

Repeat Offender Status	% of Universe
Non-Repeat Offender	87.0%
Repeat Offender Bad	9.2%
Repeat Offender Good	3.8%

We examined Higher Prelim (bad) and Lower Prelim (good) Repeat Offenders by sector and size. First, we looked at the percentage of repeat offenders in each sector and size bucket in Figure 18 and
Figure 19. We were interested if companies in a specific bucket more likely to be a repeat offender regardless of direction.

Who are the repeat offenders? Table 17 details companies that meet our predefined repeat offender requirements at some point from July 2002 to November 2010. It provides historical examples of repeat offenders and outlines a specific offense. Table 17 highlights the wide range of sizes, sectors, magnitudes of difference, and years where we see these occurrences.

Table 17: Net Income Historical Repeat Offender Examples

Russell 3000, 7/2002 to 11/2010, Net Income in Millions

Company	Lower/Higher Prelim	Count of Differences	Average Difference	Prelim	Final	Period	Sector
Netflix	Lower	4	.278	26.561	26.579	Q2 2008	Consumer Discretionary
MIPS Technologies	Lower	4	.565	(7.154)	(7.031)	Q1 2008	Information Technology
Toreador Resources	Lower	3	1.034	2.446	3.202	Q4 2005	Energy
NII Holdings	Lower	4	5.075	23.7	28.2	Q3 2004	Telecommunications
National Semiconductor	Lower	4	4.075	33.9	36.3	Q2 2009	Information Technology
Kodak	Higher	3	(36.667)	-146	-154	Q2 2005	Consumer Discretionary
Chemtura	Higher	6	(20.546)	2.507	.420	Q2 2006	Materials
Calpine Corp	Higher	3	(7.056)	22.332	15.019	Q3 2004	Utilities
Fifth Third Bancorp	Higher	3	(26.000)	292	286	Q1 2008	Financials
Zales	Higher	3	(1.295)	(26.447)	(27.363)	Q4 2006	Consumer Discretionary

Figure 18: Total Repeat Offenders by Sector

Russell 3000 from 7/2002 to 11/2010 for Net Income base of Total Sector Size

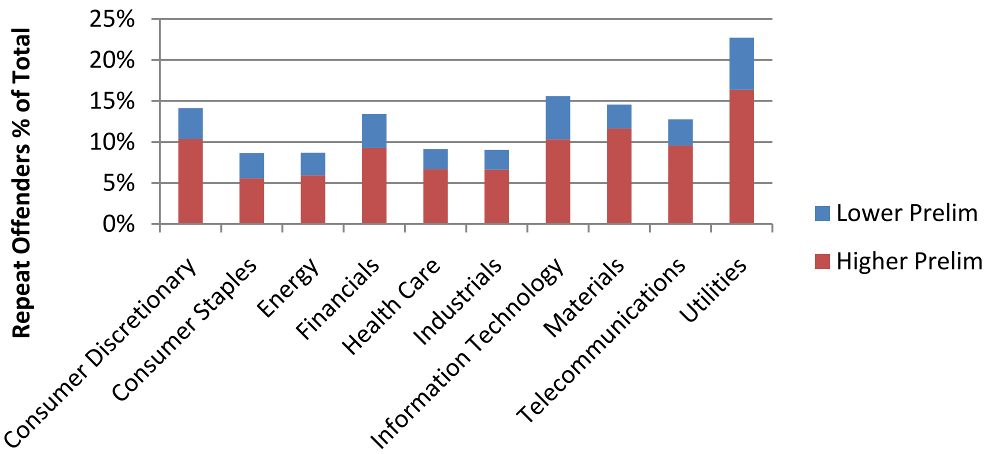
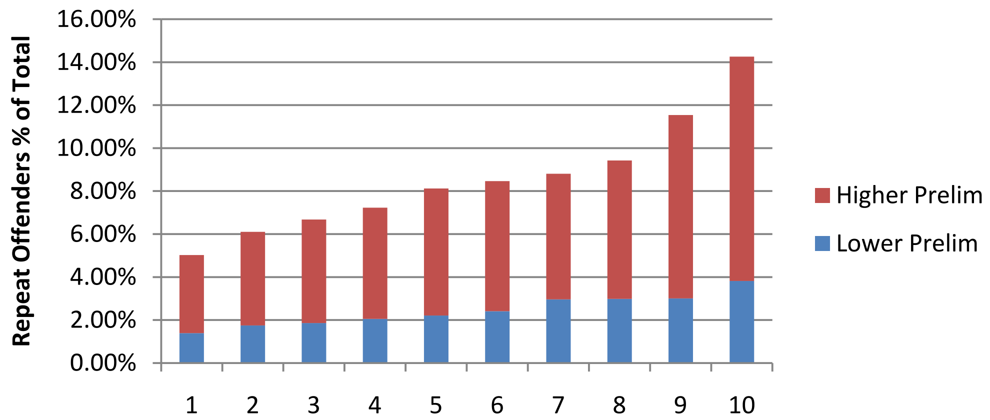


Figure 19: Total Repeat Offenders by Market Cap

Russell 3000 from 7/2002 to 11/2010 for Net Income base of Total Market Cap Decile Size



We find that the Utilities and Information Technology sectors are significantly more likely to contain repeat offenders than Energy, Health Care, Industrials, and Consumer Staples, see Table 18. Furthermore, we find a likely large cap bias in the repeat offender space. Our top two deciles by market capitalization are significantly more likely to have higher prelim repeat offenders relative to many of their small counterparts and some midcaps are more likely than the very smallest names, Table 19. We hypothesize this may be attributable to large companies being more acquisitive and in a financial position that allows for more impactful discretionary moves more frequently than smaller names.

Table 18 ANOVA and Tukey Pairwise Comparison by Sector

Russell 3000 from 7/2002 to 11/2010 for Net Income base of Total Sector Size

Overall ANOVA Statistics

Group Variable	DF	F Value	Pr > F
Sectors	9	4.91	<.0001

Pairwise Multiple Comparison Statistics with Controlled Type 1 Error

Sectors	Means Difference	95% Confidence Limits		Significant at 0.05 Level
Utilities-Energy	0.15	0.02	0.27	***
Utilities-Health Care	0.14	0.03	0.25	***
Utilities-Industrials	0.14	0.04	0.25	***
Utiltiies-Consumer Staples	0.14	0.01	0.27	***
Info Tech-Health Care	0.07	0.01	0.12	***
Info Tech-Industrials	0,07	0.02	0.12	***

Table 19: ANOVA and Tukey Pairwise Comparison by Market Cap

Russell 3000 from 7/2002 to 11/2010 for Net Income base of Total Market Cap Decile Size
Overall ANOVA Statistics

Group Variable	DF	F Value	Pr > F
Market Cap Deciles	9	10.53	<.0001

- *Pairwise Multiple Comparison Statistics with Controlled Type 1 Error*
- *First Significant Comparison for each Significant Decile, all smaller deciles also significantly different*

Market Cap Deciles	Means Difference	95% Confidence Limits		Significant at 0.05 Level
10-8	0.05	0.005	0.096	***
9-5	0.04	0.001	0.073	***
8-1	0.03	0.002	0.067	***
7-1	0.04	0.010	0.070	***
6-1	0,04	0.008	0.065	***
5-1	0.03	0.004	0.059	***

Next we examined the directional repeat offenders relative to the total number of repeat offenders in each sector. This provides a slightly different view of the data allowing us to compare the likelihood of a repeat direction across sectors.

Figure 20: Directional Repeat Offenders by Sector
Russell 3000 from 7/2002 to 11/2010 for Net Income base of Total Repeat Offenders

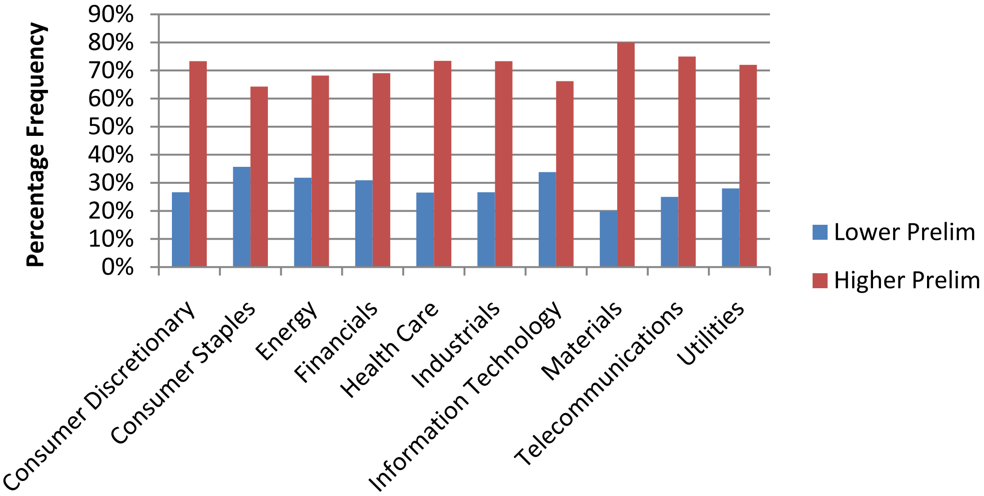
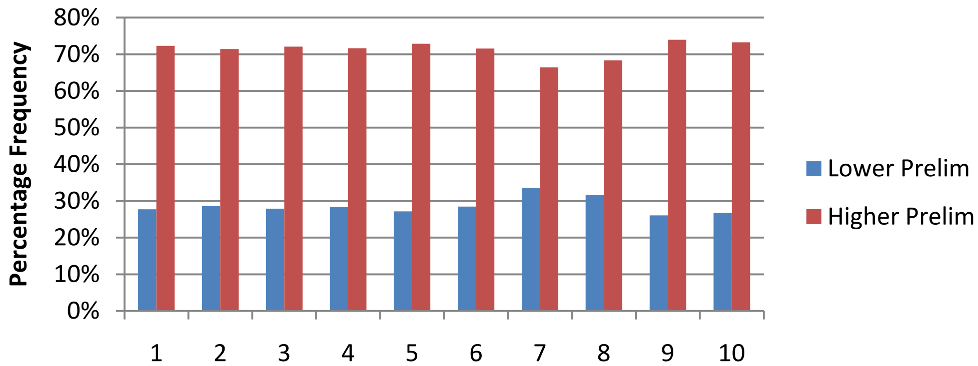


Figure 21: Directional Repeat Offenders by Market Cap
Russell 3000 from 7/2002 to 11/2010 for Net Income base of Total Repeat Offenders



We find no appreciable difference in the odds of being either higher or lower prelim across different sectors, Table 20. Similarly, there is no significant difference in the odds of being either higher or lower prelim across different market cap deciles, Table 21.

Table 20: ANOVA Repeat Offender Comparison by Sector

Russell 3000 from 7/2002 to 11/2010 for Net Income base of Sector Repeat Offenders

Overall ANOVA Statistics

Group Variable	DF	F Value	Pr > F
Sectors	9	.45	0.9072

Table 21: ANOVA Repeat Offender Comparison by Market Cap

Russell 3000 from 7/2002 to 11/2010 for Net Income base of Market Cap Decile Repeat Offenders

Overall ANOVA Statistics

Group Variable	DF	F Value	Pr > F
Market Cap Deciles	9	0.28	0.9793

7 Days to Final Filing Analysis

Finally, we analyzed what, if any, signals trends could we observed based on the number of days observed between prelim and final filings? We observed consistently higher prelim for all date differences of filings for Q4. This higher prelim relationship is present in both Q4 data items, Figure 22 and
Figure **23**. Higher preliminary filings are independent of filing time.

Figure 22: Net Income Difference Comparison by Filing Days

Russell 3000, Quarter 4, De-Median Filing Days between Prelim and Final from 7/2002 to 11/2010

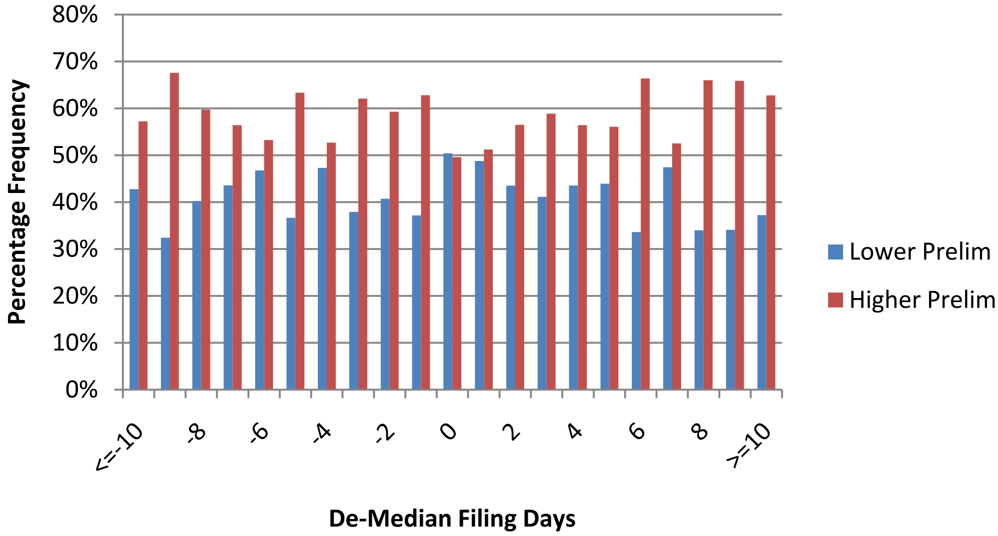
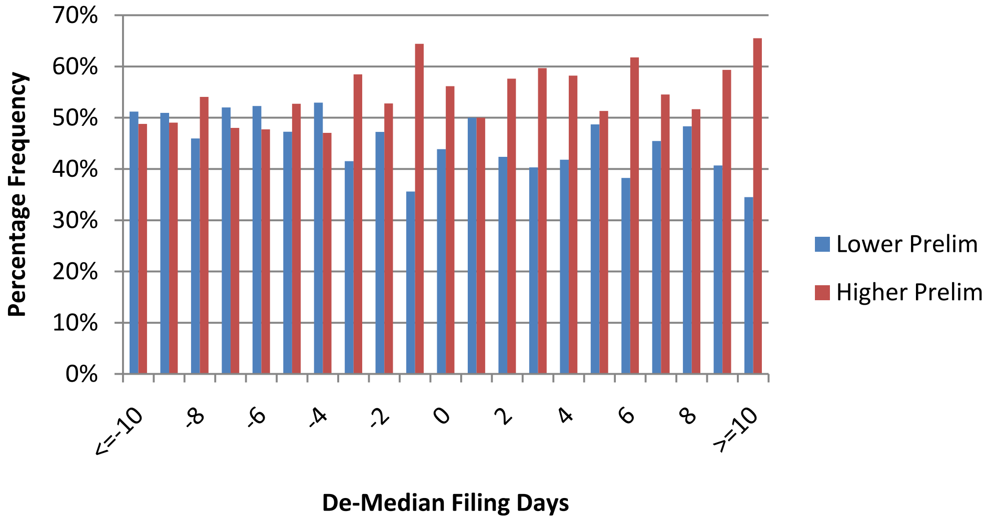


Figure 23: Diluted EPS Difference Comparison by Filing Days

Russell 3000, Quarter 4, De-Median Filing Days between Prelim and Final from 7/2002 to 11/2010



Unlike Q4 filings, Non-Q4 Net Income and Diluted EPS data have a strong inverse relationship between date difference and prelim final difference, Figure 24 and Figure 25. **As filing time increases, we find that it is more likely that the final number will be lower than the preliminary number.**

Figure 24: Net Income Difference Comparison by Filing Days

Russell 3000, Not Quarter 4, De-Median Filing Days between Prelim and Final from 7/2002 to 11/2010

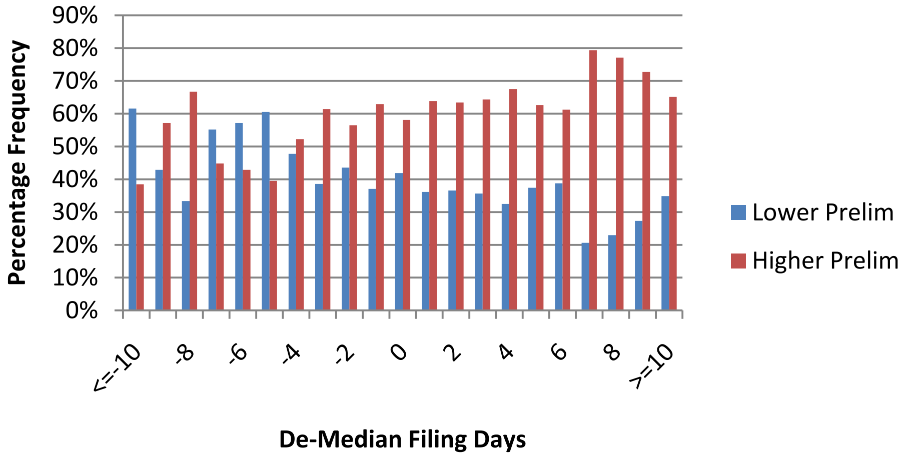
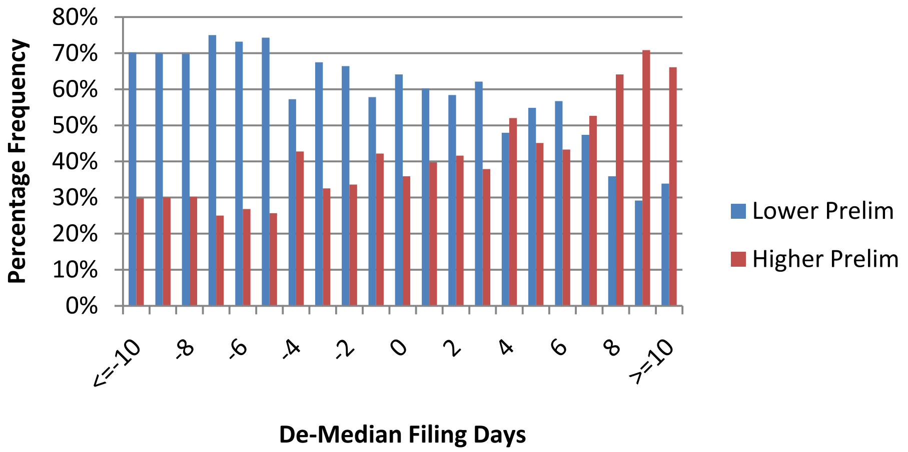


Figure 25: Diluted EPS Difference Comparison by Filing Days

Russell 3000, Not Quarter 4,De-Median Filing Days between Prelim and Final from 7/2002 to 11/2010



8 Conclusions

We leveraged the Capital IQ Point-in-Time database to explore differences between preliminary and final filings. The daily point-in-time format and breadth of coverage allowed us to find new and unique insights into company filings, market reactions, and difference characteristics.

The market indeed reacts both to differences between filings, and to the initial release of preliminary data. The spread between preliminary and final filings can be used as an alpha signal even in relatively simple implementations.

The incorporation of preliminary data in traditional alpha factors also improves individual factor performance, which we believe enables investors to differentiate and enhance their stock selection process.

There are systematic characteristics of both the differences and the offenders. While offenses are not common, companies are more likely to initially report higher than lower. We find higher preliminary reported numbers to be universally more likely across all sectors and sizes in the case of Net Income. The offenders themselves are most often utilities, information technology, and large companies.

REFERENCES

Securities and Exchange Commission, Additional Form 8-K Disclosure Requirements and Acceleration of Filing Date, 2004

OUR RECENT RESEARCH

February 2011: Industry Insights – Biotechnology: FDA Approval Catalyst Strategy
Biotechnology is a challenging sector for investors due to the binary nature of the product cycle. Indeed many biotechnology firms' futures rest upon the success of a single product. A critical stage in the product life-cycle is the FDA approval process. In this report we look at the exploitability of a strategy centered on FDA filings.

January 2011: US Stock Selection Models Introduction
In this report, we launch our four US Stock Selection models -- Value, Growth, Quality, and Price Momentum. Built using Capital IQ's robust data and analytics, these four models are the culmination of over two years of research and development. Each model is intended to be employed as the basis for a stand-alone stock selection strategy or integrated into an existing systematic process as an overlay or new component.

January 2011: Variations on Minimum Variance
Various explanations for why risk is mispriced have been offered; the most common one is that leverage restrictions incite some investors to chase volatility at the individual issue level. In this paper, we explore various methodologies for construction of minimum variance portfolios of US listed equities and analyze the features of these portfolios.

January 2011: Interesting and Influential Papers We Read in 2010
As researchers, we spend a large amount of time trying to generate new ideas. In order to discover and refine these ideas, we find ourselves in a continuous quest for innovative and interesting articles and papers from academics, analysts, and other researchers. There is such a large body of information out there that it can be difficult to wade through all the material to find what is truly of value and interest to us. To assist in sifting through all this information, our group recently took the time to find and discuss articles that recently struck us.

November 2010: Is your Bank Under Stress? Introducing our Dynamic Bank Model
Leveraging Capital IQ's Bank industry data, we have built a stock selection model that encompasses three themes -- Momentum, Value, and Balance Sheet Quality -- and includes a proprietary Markov-regime switching component which dynamically changes the model's weights depending on whether or not banks are in a "stressful" (or crisis) environment. This month, we will review how we built our model and its switching component.

October 2010: Getting the Most from Point-in-Time Data
In this paper, we will examine PIT data's origins, structure, variations, and proper use in implementations from Compustat and Capital IQ. Misusing PIT data, or applying it haphazardly, can discard valuable information and obscure otherwise clear signals.

October 2010: Another Brick in the Wall: The Historic Failure of Price Momentum
In 2009, investors witnessed the cataclysmic failure of Price Momentum strategies. Now that accounts of this failure have been on the books for some time, it is appropriate to place the events in a historical context and further analyze the fundamental relationships that affect this strategy. We look at a number of questions from practitioners interested in the strategy. Within a historical context, how pronounced has this recent failure been? When Price Momentum fails, what is the strategy's subsequent performance? And, what factors are concurrent or predictive of the performance of Price Momentum?

July 2010: Introducing Capital IQ's Fundamental US Equity Risk Model
In this paper we document the process of building and testing of our fundamental US Equity risk model across a number of short to medium term forecast horizons. The paper reviews typical risk model applications; discusses the relative merits of alternative forms of multifactor risk models; documents our data and methodology; 4 describes the chosen test metrics; and presents our results.