Extracting Market Implied Earnings From Equity Market Data

Carter Bryson **Meiyao Tysinger** **Jesse Meredith**

September 2014

Abstract

We propose a forecasting model that utilizes multiple financial market variables to predict firm earnings. The model is analogous to a measure of what the market expect the firm's earnings to be. We call this measure Market Implied Earnings (MIE), as it gives the earnings implied by financial market activity. MIE tends to predict actual earnings with an R-squared value of 0.33, relative to 0.21 for a simple Random Walk model. We anticipate that MIE could be used for a profitable trading strategy, particularly when combined with Post-Earnings-Announcement-Drift (PEAD).

We thank Dr. Michael Aguilar for helpful comments and critiques. We also thank the Economics department at University of North Carolina at Chapel Hill and the Herbert Brown Mayo Research Fellowship for sponsoring this endeavor.

1. Introduction

Recent studies have investigated the influence of earnings releases on equity prices, but little work has been done to address the true market's expectations of earnings prior to their release. Since cash flow news is a significant component in stock prices and stock returns, investors pay particular attention and effort on forecasting company's future earnings (Chen, Da and Zhao, 2013). The market then incorporates all available information future cash flows into the price. When the market undervalued or overvalued equities from either overly pessimistic or optimistic earnings projections, the price adjusts accordingly after the actual earnings announcement (Bernard and Thomas, 1989). While prior research has measured the surprise level of investors and measured the price changes after the event. This paper examines the linear correlation between market variables and earnings release, which enables us to attempt to forecast investor's expectation of earnings per share ex ante.

Most of the factors used in return prediction models tend to be exogenous to the equity market (Goyal and Welsh, 2008). However, these models tend to poorly predict the regressand (Goyal and Welsh, 2008). Though return models are often unsatisfying, trading strategies taking advantage of Post-Earnings-Announcement-Drift (PEAD) generate surprising returns, particularly given the simplicity of the strategy and that the Drift has been known for years (Ng, Rusticus, and Verdi, 2008). The trading strategies formulated around PEAD tend to involve purchasing the asset on the day of the earnings announcement (Ng, Rusticus, and Verdi, 2008)(Ke and Ramalingegowda, 2005). Should a factor model be developed that can predict the unusual earnings reports even one day before the release, it would allow investors to capture the day-of reaction to the report as well as the delayed reaction in PEAD.

Ferreira and Santa-Clara (2011) generated a return decomposition in which the returns for an equity over a period are a function of the change in the firm's earnings per share, the change in the P/E ratio, and the dividend yield, all over the same period as the returns. Since earnings are not released until after the period has ended, and all other factors are known in real time, the decomposition seems to imply that if one knows the returns, change in P/E ratio, and dividend yield for an equity for a firm for a period, it should be possible to solve for the earnings term and develop a measure of what the market expects earnings to be, which we call Market Implied Earnings (MIE). Though the decomposition was designed for predicticting returns, and is thus poorly suited to estimating MIE (see Appendix B), the insight persists. A regression comparing earnings reports to measures of market action, such as stock price and trading volume could approximately measure the MIE, provided that the forecasts of market actors are distributed reasonably close to the true value such that the errors cancel each other out. Such a regression resembles that attempted by Zhou, Ghysels, and Ball (2014), though theirs is optimized for predicting returns, rather than estimating MIE. In particular, their list of predictors only includes two measures of market activity, and the regression takes place in calendar time. To capture MIE, it is necessary to use event time in order to account for the differences in earnings report release dates. To do otherwise it to introduce the confounding effect of the release itself into the regression data. Working in event time also circumvents the issues typically involved in constructing regressions with data of multiple frequencies outlined in Zhou, Ghysels, and Ball (2014).

2. Methodology

We gathered data for the stocks in the S&P 500 from the FactSet database, which contained data going back to the year 1998. We then took linear regressions between the dollar change in actual earnings from quarter-to-quarter and a variety of market variables. Consistent with prior literature, we separately measure each variable for the day, week, month, and quarter prior to each earnings report (McAleer and Medeiros, 2008)(Zhou, Ghysels, and Ball, 2014). We used commercially available statistical software to develop those variables not found within the data (such as weekly returns, and changes in trading volume). However, missing values within the data set limited the number of variables which we could use within the model. We then performed standard statistical analysis in order to estimate the regression coefficients. Because both the estimated coefficients and the fit of the model vary from sector to sector, we computed the average of the sector coefficients and the average of the sector R^2's to give us estimates of the "market" coefficients and R^2. A similar procedure as used to estimate the Random-Walk model with our dataset, which we used as a comparison for our MIE model. The change in earnings predicted by both models over time is displayed in the Appendix.

Our work expands on that done previously in two different areas. First, we break from the convention of estimating earnings and instead attempt to estimate the market's consensus estimate of earnings. Second, our model also takes advantage of a greater variety of financial variables than prior work. Zhou, Ghysels, and Ball (2014) examine a broad array of variables proposed in the literature to be predictive of earnings. Only 2 of the 14 variables studied measure equity market activity, and thus could be tied to earnings predictions. Goyal and Welch (2008) include a broader array of market variables, but most of their regressors are exogenous to equity markets. Furthermore, the aforementioned variables are predominantly of monthly or quarterly frequency. Given that market actors are continually updating their forecasts as new information becomes available, monthly data are insufficient. Therefore, use use variables of daily and weekly frequency, as well as the conventional monthly and quarterly ones, allowing us a more granular look at market expectations (McAleer and Medeiros, 2008)(Zhou, Ghysels, and Ball, 2014).

<u>Regression Specification:</u>

$\Delta EPS_t = \square_0 + \square_1 \Delta EPS_{t-1} + \square_2 Returns(Day) + \square_3 Returns(Week) + \square_4 Returns(Month) + \square_5 Returns(Quarter) + \square_6 \%\Delta Volume(Day) + \square_7 \%\Delta Volume(Week) + \square_8 \%\Delta Volume(Month) + \square_9 \%\Delta Volume(Quarter) + u_t$

Where Day, Week, Month, and Quarter all refer to the time period preceding the earnings announcement. For example, Returns(Week) is the equity's return over the week preceding the earnings announcement.

3. Results

Our regression model for the entire S&P 500 predicts the change in EPS over a quarter from the following variables: the change in EPS in the prior quarter, equity returns for the day, week, month, and quarter prior to the earnings release, and the change in trading volume for the day, week, month, and quarter prior to the release.

Variable	Coefficient	p-value
Constant B0	-0.0802	0.3090
EPS Change (t-2,t-1)	-0.5108	0.0103
Day Returns	0.7253	0.5692
Week Returns	3.5425	0.4412
Month Returns	-0.1052	0.3042
Qtr Returns	1.6489	0.2408
Day Volume Change	-0.0805	0.3994
Week Volume Change	0.527	0.6132
Month Volume Change	-0.1004	0.4575
Qtr Volume Change	0.0162	0.2628

None of the regression coefficients are significant at the 10% level. The earnings change in the previous quarter is significant at the 11% level, with the next most robust results requiring 24% and 26% significance levels in order to accept. Given the difficulty in achieving significance when predicting earnings, these results are not surprising (Goral & Welch, 2008). In particular, the prevalence of null values within the data set severely limited our ability to run the regression. A future economist with more extensive data may attempt this exercise and find a significant relationship.

Our model specification for the Random Walk we used as a baseline appears in the chart below:

Variable	Coefficient	p-value
Regression constant	0.0159	0.527
EPS Change(t-2,t-1)	-0.4653	0.000

On average, our model for MIE predicts the movements of actual earnings with an R-squared value of 0.33. For comparison, we simulated a simple Random Walk model using the same data and achieved on average an R-squared value of 0.21. This improvement suggests that our MIE model has some use despite its lack of statistical significance.

The charts in Appendix A plot MIE, a simple Random Walk earnings prediction model, and realized earnings over time for the 10 sectors of the S&P 500.

4. Discussion

Our results indicate that it may be possible to predict earnings using indicators of financial market activity to a greater extent than with a random walk model. While a simulated trading strategy is beyond the slope of this paper, the apparently link between earnings and equity market variables suggests that the MIE measure developed here could be used to make money (Ferreira and Santa-Clara, 2011). These possible gains are further compounded by the possibility of combining MIE forecasting with a trading strategy involving Post-Earnings-Announcement-Drift such as that described by Ng, Rusticus, and Verdi (2008).

It is notable that, throughout the sample period, instances in which the absolute change of earnings exceeds both MIE and the Random Walk (RW) model are far more numerous than those in which a model overestimated the size of the shift. In other words, it is during large earnings shifts that both models fail to capture the full magnitude of the event. Such a failure is to be expected of the backwards-looking RW model, but requires some explanation in the MIE model. The failure of MIE to track earnings at their most unusual could be attributed to inadequate information being available to the public, thus preventing market actors from fully predicting the change in earnings. This anomaly could also be due to earnings expectations having an adaptive component as well as a rational one.

The significance of our MIE model is limited at the outset by two factors: other models that predict earnings, and the data involved. The MIE model essentially extracts the results of the earnings models used by market actors from changes within the market. Thus, MIE by definition cannot be systemically more accurate than the models used by market actors. Given that forecasting earnings purely through extrapolation of past trends is common practice, the quality of the market's earnings estimators is in doubt (Ferreira and Santa-Clara, 2011).

Constraints on our data also limit the significance of our model. In addition to the difficulties outlined in Appendix B, some of our variables (particularly trading volume) do not have extensive historical data. Our earliest data is from 1998, 16 years ago. While this duration provides enough data to generate useful tools, it does not provide enough to allow for the tool to be fully robust. As time progresses, better earnings models and more extensive financial data should allow for more robust MIE models to be developed.

5 Conclusion

It is all but self-evident that financial market actors develop earnings forecasts and trade based on them. If market actors buy and sell based on their forecasts, then it should be possible to use measures of market activity, such as returns and trading volume, to estimate the market's aggregate forecast. While our results suggest that such estimation may be possible, the results are not currently significant. Given that significance proves elusive among earnings forecasters, this result is in line with the prior literature (Goyal & Welch, 2008). We anticipate that a MIE regression performed with a larger, more complete data set would yield results closer to significance.

6 References

Claus, J. and Thomas, J. 2001. “Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets.” *The Journal of Finance* 56 (5): 1629-1666, http://onlinelibrary.wiley.com/doi/10.1111/0022-1082.00384/abstract

Ferreira, Miguel A. and Santa-Clera, Pedro. 2011. “Forecasting stock market returns: The sum of the parts is more than the whole.” *Journal of Financial Economics*, 100 (3): 514-537 http://www.sciencedirect.com/science/article/pii/S0304405X11000365#bib5

Goyal, Amit & Welsh, Ivo. 2008. “A Comprehensive Look at The Empirical Performance of Equity Premium Prediction.” *The Review of Financial Studies* 21 (4): 1455-1508 http://rfs.oxfordjournals.org/content/21/4/1455.full

Ke, Bin and Ramalingegowda, Santhosh. 2005. “Do institutional investors exploit the post-earnings announcement drift?” *Journal of Accounting and Economics*, 39 (1):25-53 http://www.sciencedirect.com/science/article/pii/S0165410104000552

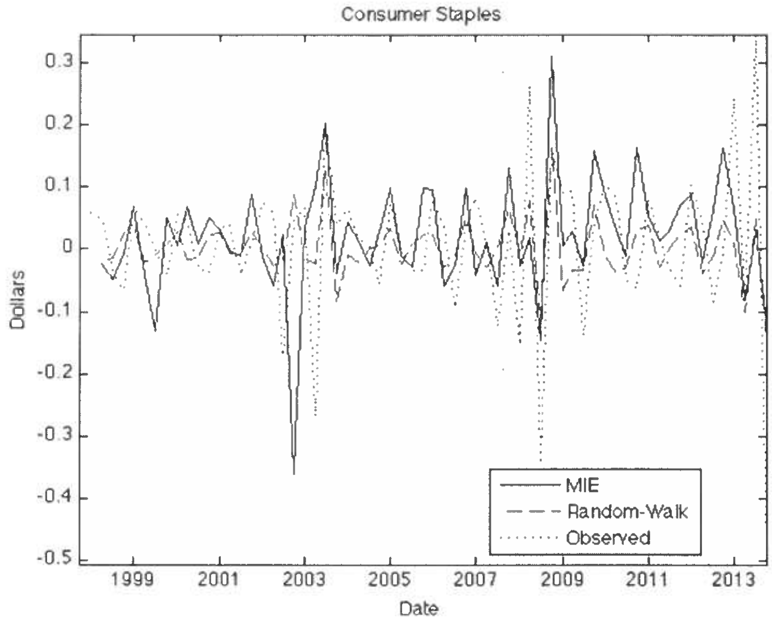
McAleer, Michael and Medeiros, Marcelo C. 2008 “A multiple regime smooth transition Heterogeneous Autoregressive model for long memory and asymmetries.” *Journal of Econometrics*. 147 (1): 104-119. http://www.sciencedirect.com/science/article/pii/S0304407608001280

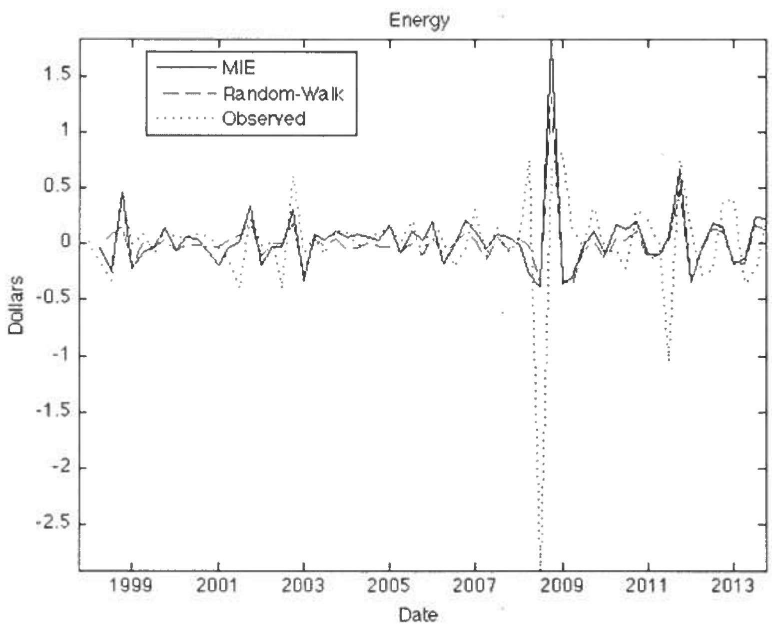
Ng, J., Rusticus, T. O. and Verdi, R. S. 2008. "Implications of Transaction Costs for the Post–Earnings Announcement Drift." *Journal of Accounting Research* 46 (3): 661-696.
http://onlinelibrary.wiley.com/doi/10.1111/j.1475-679X.2008.00290.x/full

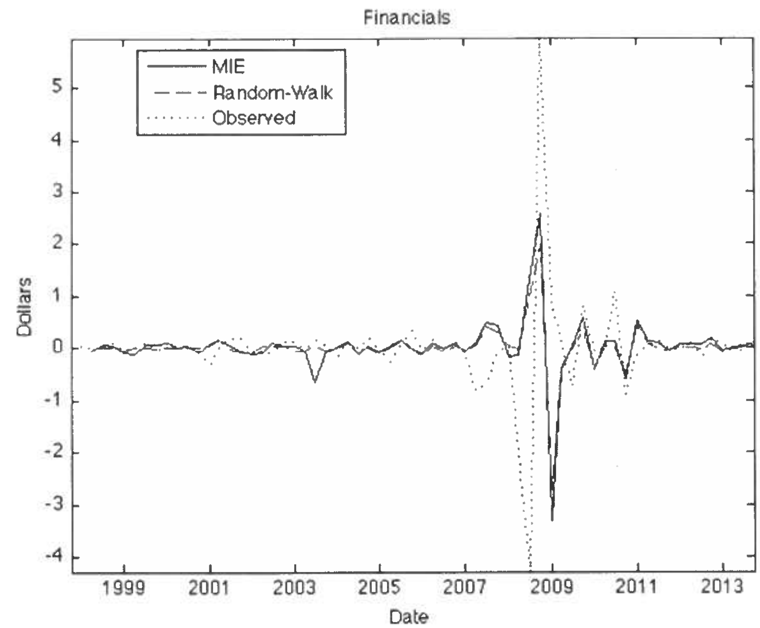
Zhou, Huan; Ghysels, Eric, and Ball, Ryan. 2014. "Beating Analysts' Earnings Forecasts with a MIDAS Touch."

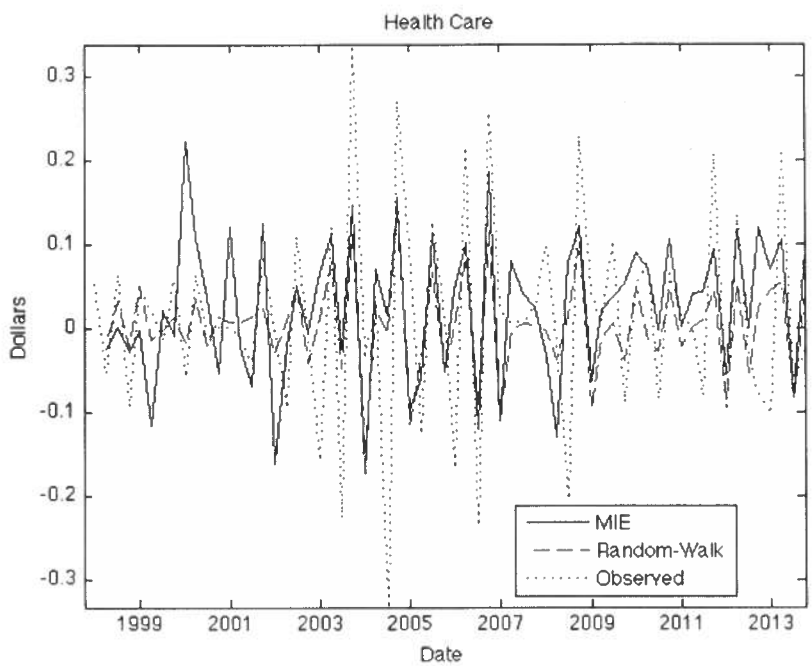
7 Appendices

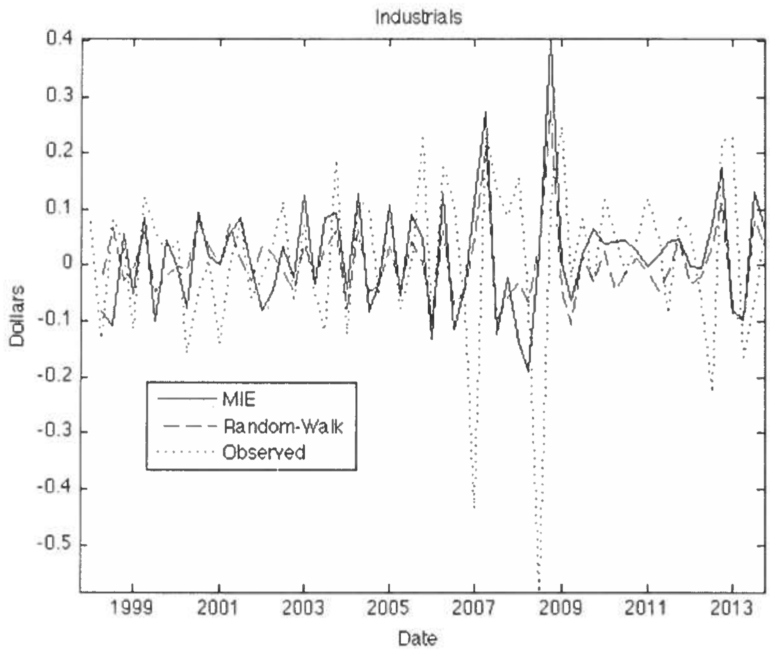
A)

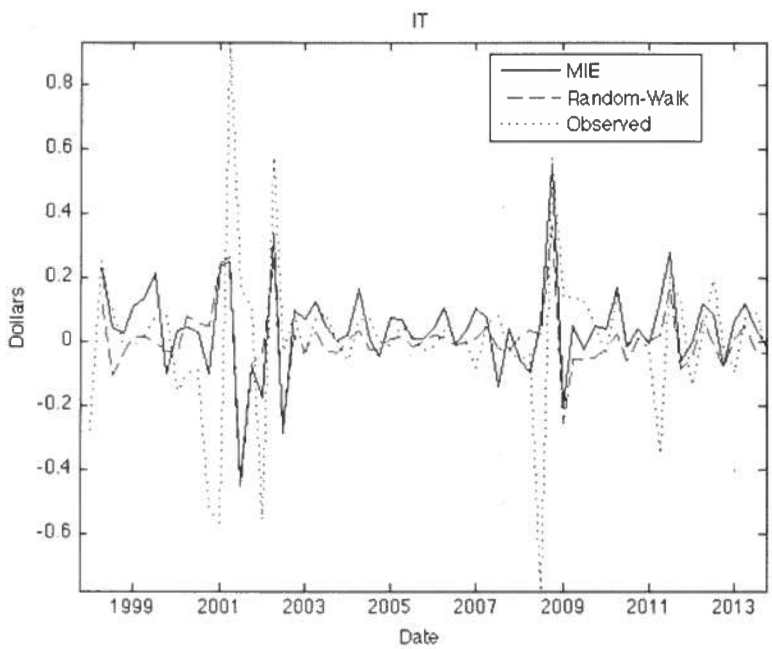


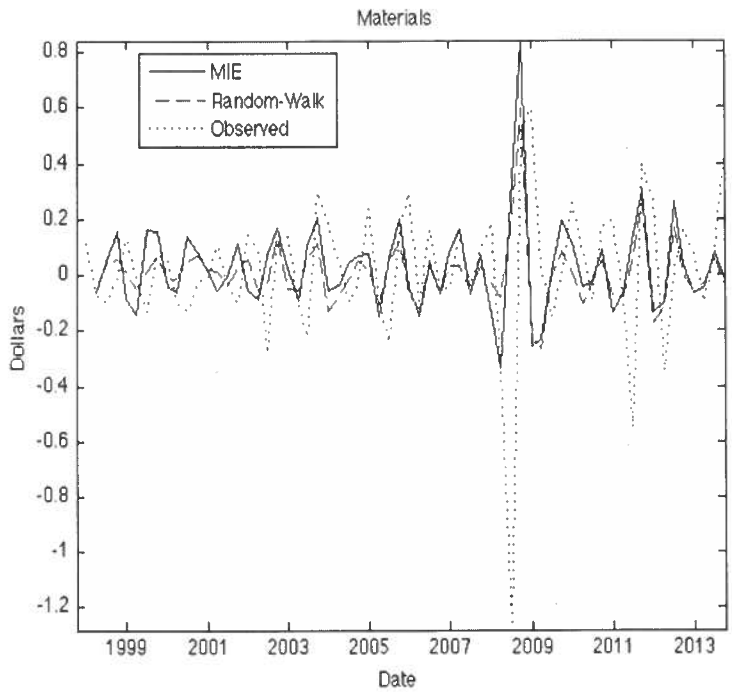


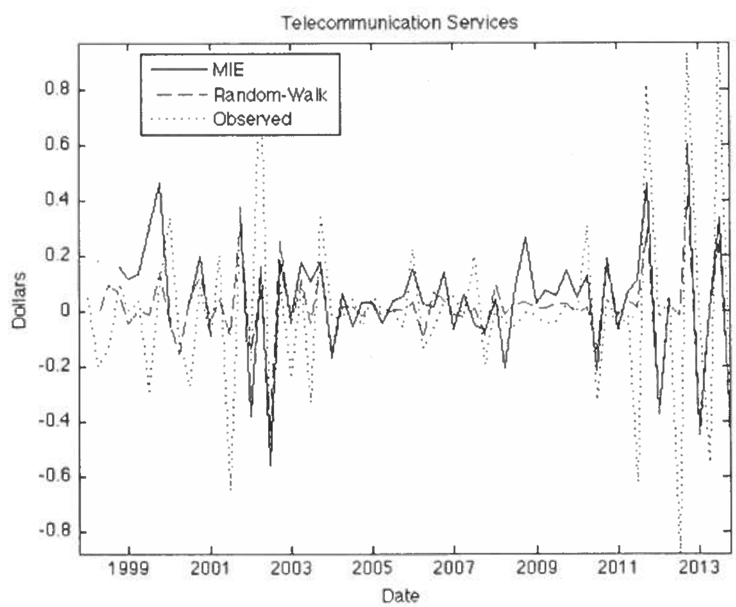

Financials
MIE
Random-Walk
Observed
5
4
3
2
1
0
-1
-2
-3
-4
Dollars
1999
2001
2003
2005
2007
2009
2011
2013
Date

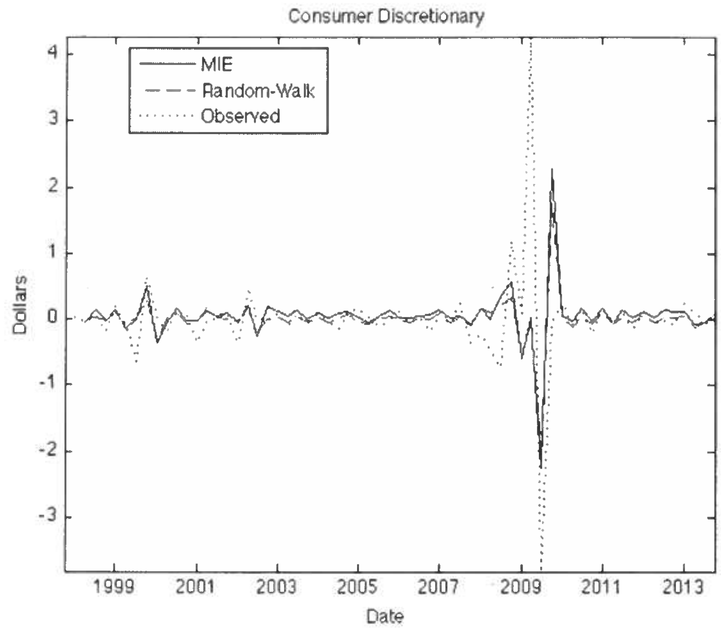

Health Care
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
Dollars
MIE
Random-Walk
Observed
1999
2001
2003
2005
2007
2009
2011
2013
Date

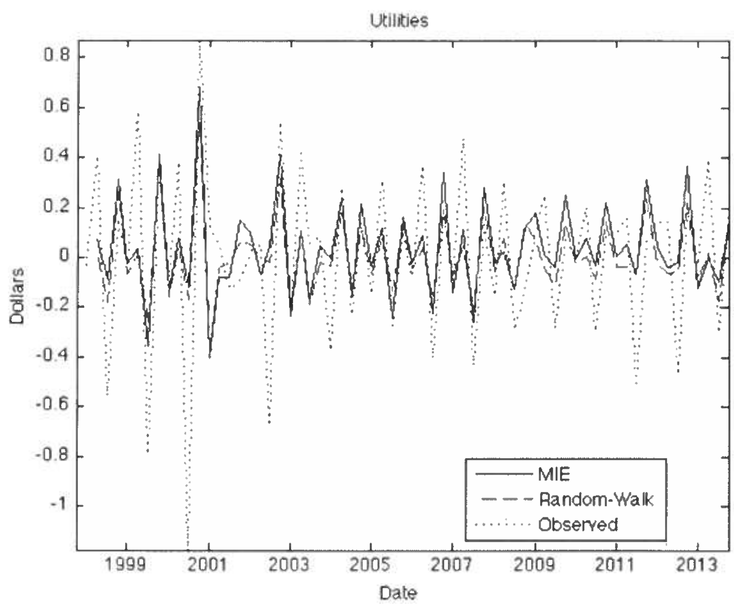

Industrials
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
Dollars
MIE
Random-Walk
Observed
1999
2001
2003
2005
2007
2009
2011
2013
Date


IT
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
Dollars
MIE
Random-Walk
Observed
1999
2001
2003
2005
2007
2009
2011
2013
Date


Materials
MIE
Random-Walk
Observed
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
Dollars
1999
2001
2003
2005
2007
2009
2011
2013
Date


Telecommunication Services
MIE
Random-Walk
Observed
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
Dollars
1999
2001
2003
2005
2007
2009
2011
2013
Date





B) Ferreira and Santa Clara's Decomposition

1+R(t+1)= 1+ CG(t+1)+DY(t+1), where R is the return on an equity, CG is the capital gain, and DY is the dividend yield, with (t+1) indicating the time period from point t to point (t+1).

1+R(t+1) = 1+ (P(t+1)-P(t))/P(t)+DY(t+1), where P(t) is the price of the stock at time t

$1+R(t+1)=1+(P(t+1)/P(t))-(P(t)/P(t))+DY(t+1)$

$1+R(t+1)=1+P(t+1)/P(t)-1+DY(t+1)$

$1+R(t+1)=P(t+1)/P(t) * E(t+1)E(t)/E(t+1)e(t)+DY(t+1)$, where $E(t)$ is EPS at time t

Ferreiria and Santa-Clara go on to create an earnings term and a P/E term. However, at this point it becomes apparent that the model is unsuited to extracting MIE. To extract MIE would involve isolating $E(t+1)$ and solving for it in the period between when the earnings are accrued and when the earnings report is released. The means by which $E(t+1)$ is introduced to the function, however, prohibits such isolation. A different method for estimating MIE was thus required.