

Georgia State University

## ScholarWorks @ Georgia State University

---

Mathematics Theses

Department of Mathematics and Statistics

---

Summer 8-18-2010

### Looking in the Crystal Ball: Determinants of Excess Return

Kokou S. Akolly

*Georgia State University*

Follow this and additional works at: [https://scholarworks.gsu.edu/math\\_theses](https://scholarworks.gsu.edu/math_theses)



Part of the [Mathematics Commons](#)

---

#### Recommended Citation

Akolly, Kokou S., "Looking in the Crystal Ball: Determinants of Excess Return." Thesis, Georgia State University, 2010.

[https://scholarworks.gsu.edu/math\\_theses/87](https://scholarworks.gsu.edu/math_theses/87)

This Thesis is brought to you for free and open access by the Department of Mathematics and Statistics at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Mathematics Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact [scholarworks@gsu.edu](mailto:scholarworks@gsu.edu).

# LOOKING IN THE CRYSTAL BALL: DETERMINANTS OF EXCESS RETURN

by

KOKOU S.AKOLLY

Under the direction of Dr Jun Han

## ABSTRACT

This paper investigates the determinants of excess returns using dividend yields as a proxy in a cross-sectional setting. First, we find that types of industry and the current business cycle are determining factors of returns. Second, our results suggest that dividend yield serves a signaling mechanism indicating “healthiness” of a firm among prospective investors. Third we see that there is a positive relationship between dividend yield and risk, especially in the utility and financial sectors. And finally, using actual excess returns, instead of dividend yield in our model shows that all predictors of dividend yield were also significant predictors of excess returns. This connection between dividend yield and excess returns support our use of dividend yield as a proxy for excess returns.

INDEX WORDS: Excess returns, Dividend payout ratio, Dividend yield, Book market value, Industry, fixed effects model, Random effects model.

LOOKING IN THE CRYSTAL BALL: DETERMINANTS OF EXCESS RETURN

by

KOKOU S.AKOLLY

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

In the College of Arts and Sciences

Georgia State University

2010

Copyright by

Kokou Sitsope Mikaeli Akolly

2010

LOOKING IN THE CRYSTAL BALL: DETERMINANTS OF EXCESS RETURN

by

KOKOU S.AKOLLY

Committee Chair: Dr Jun Han

Committee: Dr. Yichuan Zhao

Dr. Gengsheng Qin

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

August 2010

## **DEDICATION**

To my mother, Adjoa Angele Apedjinou.

“Hundreds of dewdrops to greet the dawn,

Hundreds of bees in the purple clover,

Hundreds of butterflies on the lawn,

But only one mother the wide world over.”

George Cooper.

**TABLE OF CONTENTS**

<b>LIST OF TABLES</b>	vi
<b>1. Introduction</b>	1
<b>2. Theory and Model</b>	5
<i>2.1 Theory</i>	5
<i>2.2 Model</i>	6
<b>3. Empirical Analysis</b>	7
<i>3.1 Data Description</i>	7
<i>3.2 General Diagnostics</i>	10
<i>3.3 Empirical Analysis (Using Data Subset)</i>	13
<i>3.3.1 Dividends in Different Industrial Sectors</i>	13
<i>3.3.2 Dividends as a Signaling Mechanism</i>	18
<i>3.3.3 Dividends as Compensation for Risk</i>	23
<i>3.3.4 Excess Return model</i>	26
<b>Conclusion</b>	28

## LIST OF TABLES

<b><u>Table 1: Excess Returns: Dividend Yield, Earning Yield, and Payout (Lamont 1996)</u></b>	3
<b><u>Table 2: Variables description</u></b>	8
<b><u>Table 3: Firms and Industries in AKOLLY30 and DJ30 Indices</u></b>	9
<b><u>Table 4: Dividends in different industries: Regression with Driscoll-Kraay standard errors of excess returns on explanatory variables</u></b>	19
<b><u>Table 5: Mean Dividend Yield by Industry</u></b>	20
<b><u>Table 6: Dividend as signaling mechanism: Regression with Driscoll-Kraay standard errors of log-excess returns on explanatory variables</u></b>	22
<b><u>Table 7: Prediction Errors of Realized Return and Dividend Yield Methods, 1962-1982</u></b>	24
<b><u>Table 8: Dividend as compensation for risk: Regression with Driscoll-Kraay standard errors of excess returns on explanatory variables</u></b>	26
<b><u>Table 9: Excess return model: Regression with Driscoll-Kraay standard errors of excess returns explanatory variables</u></b>	27



## 1. Introduction

Owen Lamont (1998) finds that the  $E/D$  ratio (known as the inverse of the dividend payout ratio<sup>1</sup>) forecasts earnings growth and returns in time series, even controlling for  $P/D$  or book market value. There have been several hedge funds that trade on the idea. They buy stocks with low  $E/D$  ratio assuming that it is good proxy for growth in price and in earnings. Is this still true when we look at cross sectional data? Do differences in  $E/D$  across firms explain differences in earnings growth and future prices? If so, does this hold when we control for book value?

Stocks predictability has been discussed extensively in the finance and statistics literature. Among the most relevant to our research project, is Lamont (1998). In his paper, Lamont finds “that aggregate dividend payout ratio forecasts aggregate excess returns<sup>2</sup> on both stocks and corporate bonds”. Particularly, high corporate profits as well as high stock prices tend to forecast excess returns on equities. The choice of dividend payout ratio as a predictor of earnings and prices is not anodyne. The reasoning behind this choice is the fact that it is connected to business conditions and also because it is a proxy of the “temporary components of earnings” Lamont (1998). Initially, Lamont shows that the dividend payout ratio forecasts excess returns on the bond market. He shows evidence that the dividend payout ratio and the dividend yield measure two different measures of returns: a “price” effect, which compares the current stock prices relative to the long term average and the “profit” effect which compares profits to long term average profits. He also examines the dynamic implications of changes in earnings and dividends on prices.

---

<sup>1</sup> Dividend payout ratio is the fraction of net income a firm pays to its stockholders in dividends:  $\text{Dividend Payout Ratio} = (\text{Dividends}/\text{Net Income for the same period})$ . The part of the earnings not paid to investors is left for investment to provide for future earnings growth. Investors seeking high current income and limited capital growth prefer companies with high Dividend payout ratio. However investors seeking capital growth may prefer lower payout ratio because capital gains are taxed at a lower rate. High growth firms in early life generally have low or zero payout ratios. As they mature, they tend to return more of the earnings back to investors.

<sup>2</sup> **Excess return** = stocks return – Treasury bill rate.

Other relevant literature include Shiller (1984) and Fama and French (1988). Both studies highlight regressions of returns on lagged  $D/P$ <sup>3</sup> and  $E/P$ <sup>4</sup> and find that, even though both have significant explanatory power,  $D/P$  does a better job of explaining returns<sup>5</sup>. In table 1, Lamont displays dividend yield and earning yield regressions, using quarterly data on excess returns, dividends, and earnings from the S&P index, from 1947 to 1994.

The  $R^2$  associated with dividend yield and future returns are 0.05 and 0.02 respectively. As such Lamont (1998) claims that dividend yield is a better forecasting variable than future returns. Fama and French (1988), explain this finding by the fact that “Earnings are more viable than dividends, if this variability is unrelated to the variation in expected returns;  $E/P$  is a noisier measure of expected returns than  $D/P$ ”.

Lamont (1998) compares the forecasting power of  $E/P$  and  $D/P$  by putting both explanatory variables in the same regression. The result, as can be seen from Table 1( in the third row), is that dividend yield is still positive and significant, while earning yield is significant but bears the unexpected negative sign. Lamont (1998) explains these results: “The explanation must be that higher variability of earnings mentioned by Fama and French (1988) is actually related to returns.” In their study, Kandel and Stambaugh (1996) use a novel approach. They consider the sample evidence about monthly stock returns<sup>7</sup>

---

<sup>3</sup>The dividend yield or the dividend-price ratio on a company stock is the company's annual dividend payments divided by its market cap, or the dividend per share, divided by the price per share. It is often expressed as a percentage. Its reciprocal is the Price/Dividend ratio. Instead, dividends paid to holders of common stock are set by management, usually in relation to the company's earnings. There is no guarantee that future dividends will match past dividends or even be paid at all. Due to the difficulty in accurately forecasting future dividends, the most commonly-cited figure for dividend yield is the current yield which is calculated using the following formula:  
**Current Dividend Yield = (Most Recent Full-Year Dividend/Current Share Price).**

<sup>4</sup> The  $E/P$  ratio is the inverse of  $P/E$  ratio (price-to-earnings ratio) of a stock (also called its "P/E", or simply "multiple") is a measure of the price paid for a share relative to the annual net income or profit earned by the firm per share. It is a financial ratio used for valuation: a higher  $P/E$  ratio means that investors are paying more for each unit of net income, so the stock is more expensive compared to one with lower  $P/E$  ratio. The  $P/E$  ratio has units of years, which can be interpreted as "number of years of earnings to pay back purchase price", ignoring the time value of money. In other words,  $P/E$  ratio shows current investor demand for a company share. The reciprocal of the  $P/E$  ratio is known as the earnings yield. The earnings yield is an estimate of expected return to be earned from holding the stock.

<sup>5</sup> remember that  $D/P * P/E = D/E$

predictability from the perspective of a risk-averse Bayesian investor. They consider an investor on the last day of 1993, who must put his funds between the value weighted portfolio of the NYSE and one

**Table 1: Excess Returns, Dividend Yield, Earning Yield, and Payout (Lamont 1996)<sup>6</sup>**

*Dependent variable:  $R_{m,t+1} - R_{f,t+1}$*

		Coefficient			
		(standard error)			
		$D_t/P_t$	$E_t/P_t$	$D_t/E_t$	$R^2$
<b>Quarterly data</b> <b>1947Q1-1994Q4</b> <b>N=192</b>		0.057 (0.018)			0.05
			0.013 (0.007)		0.02
		0.158 (0.044)	-0.044 (0.017)		0.08
		0.074 (0.019)		0.008 (0.003)	0.08

month Treasury bills. The investor is provided with the coefficients of the following regression:

$$r_t = x'_{t-1} b + \epsilon_t \quad (1)$$

<sup>6</sup> **Notes:**  $R_{m,t+1} - R_{f,t+1} = 4 * (\ln(\text{CSTIND}_{t+1}) - \ln(\text{CSTIND}_t)) - 4 * (\ln(\text{USTIND}_{t+1}) - \ln(\text{USTIND}_t))$  in quarterly data, where CSTIND is an index of total return (including reinvested dividends) on the S&P Composite Index and USTIND is an index of total return on one-month T-bills, as of the last day of quarter t. D, E, and P are the quarterly dividends, earnings, and end-of-period's stock price levels reported by Standard and Poor's Statistical Service. Earnings and dividend share 4-quarter totals, paid out in the four quarters including quarter t. All regression in this and other tables include a constant term, not shown.

where  $\mathbf{r}_t$  is the continuously compounded NYSE return<sup>7</sup> in month  $t$ , in excess of the continuously compounded T-bill<sup>8</sup> rate for that month,  $\mathbf{x}_{t-1}$  is a vector of “predictive” variables that are observed at the end of month  $t-1$ ,  $\mathbf{b}$  is vector of coefficients, and  $\epsilon_t$  is the regression disturbance in month  $t$ . They ask the question: “If the investor is also given the most recent vector of the predictive variables  $\mathbf{x}_t$ , to what extent does the investor’s asset allocation decision depend on it? The empirical set up allows Kendal and Stambaugh (1996) to observe how the Bayesian investor who uses the “sample evidence to update prior beliefs” utilizes past information about market influences not only to allocate his assets but also to predict future returns.

The objective of this paper is to investigate if cross sectional in earning yields and dividend yield forecast cross sectional in average returns, even controlling for  $B/P$ <sup>9</sup> (book-market value). Utilizing the theory provided by Gordon (1962), I set up an empirical model where dividend yield across industrials sectors and determines that industries and current economy might be factors in determining excess returns. I consider dividend yield as a variable to signal company’s success. I then concluded that, if that was the case then firms that are successful will tend to not pay as much dividend as firms that are not. Since “unhealthy” firms need to attract investors, thus they pay more dividends to compensate for the lackluster

---

<sup>7</sup> Return is the ratio of money gained or lost (whether realized or unrealized) on an investment relative to the amount of money invested.  $r = \ln(P_t/P_{t-1})$

<sup>8</sup> **Treasury bills** (or **T-Bills**) are Government’s bond issued the US. Treasury. Many regard Treasury bills as the least risky investment available to U.S. investors.

<sup>9</sup> The B/P is the inverse of market-to-book ratio, a financial ratio used to compare a company's book value to its current market price. Book value is an accounting term denoting the portion of the company held by the shareholders; in other words, the company's total tangible assets less its total liabilities. The calculation can be performed in two ways, but the result should be the same each way. In the first way, the company's market capitalization can be divided by the company's total book value from its balance sheet. The second way, using per-share values, is to divide the company's current share price by the book value per share (i.e. its book value divided by the number of outstanding shares). As with most ratios, it varies a fair amount by industry. Industries that require more infrastructure capital (for each dollar of profit) will usually trade at P/B ratios much lower than, for example, consulting firms. P/B ratios are commonly used to compare banks, because most assets and liabilities of banks are constantly valued at market values. A higher P/B ratio implies that investors *expect* management to create more value from a given set of assets, all else equal (and/or that the market value of the firm's assets is significantly higher than their accounting value). P/B ratios do not, however, directly provide any information on the ability of the firm to generate profits or cash for shareholders’ his ratio also gives some idea of whether an investor is paying too much for what would be left if the company went bankrupt immediately. For companies in distress, the book value is usually calculated without the intangible assets that would have no resale value. In such cases, P/B should also be calculated on a "diluted" basis, because stock options may ill vest on sale of the company or change of control or firing of management.

performance but also to hide the “weaknesses”. Second, I notice that industries that are risky tend to pay more dividends than those that are not as risky and this observation is complemented by the results based on data of utilities and the financial sector. Finally, I run a regression of excess returns on all the variables I have studied. The finding was that industry type, age of the company, earning/price ratio, and dividend yield matters in predicting returns. On the other hand, economic cycle doesn’t seem to influence the prediction of returns.

## 2. Theory and Model

### 2.1 Theory

Gordon (1962) asserts that “the value of an investment opportunity is the expected future receipts it ownership provides discounted at the rate of profit required on the investment”. When applied to stock price, the expected future receipts are the dividends:

$$p = \int_0^{\infty} d_t e^{-kt} dt \quad (2)$$

where  $k$  is the rate at which the corporation’s future dividends are discounted at the end of  $t = 0$  to arrive at their present value?  $p$  is the price of a stock at the end of time  $t = 0$ .  $d_t$  is the expected dividend a share of stock is supposed to pay in period  $t$ . Now let consider  $Y_t$  being the income<sup>10</sup> expected to be earned by a share of stock in period  $t$ . Let  $b$  be the fraction of income the corporation is expected to retain, the rest will be paid as dividends to shareholders. If a fraction  $b$  of the income is retained then the dividend will grow at a rate  $br$ . We can then say:

$$d_t = Y_t(1-b)e^{tbr} \quad (3)$$

Substituting this equation into the former, we get:

---

<sup>10</sup> Residual income of a firm after adding total revenue and gains and subtracting all expenses and losses for the reporting period. Net income can be distributed among holders of common stock as a dividend or held by the firm as an addition to retained earnings

$$p = \int_0^{\infty} Y_0 (1 - b) e^{brt} e^{-kt} dt \quad (4)$$

For  $k > br$ , we can integrate and get:

$$p = \frac{Y_0 (1 - b)}{k - br} \quad (5)$$

## 2.2 Model

The model proposed by Gordon (1962) will be the theoretical basis for the studies conducted in this paper.

The Gordon (1962) model of stock price states:

$$P_t = \frac{d_t}{r_t - g_t} \quad (6)$$

where  $p_t$  is stock price,  $d_t$  is the dividend,  $g_t$  is the growth rate of dividends and  $r$  is the discount rate, and both are fixed constants. It is easy to see that  $d/p_t = r - g_t$ . I can thus say that  $d/p_t$  is a measure of both the discount rate and of future dividend growth. Campbell and Shiller (1988) have a dynamic, stochastic Gordon model:

$$d_t p_t = E_t[\sum_{j=0}^{\infty} \rho^j r_{t+1+j}] - E_t[\sum_{j=0}^{\infty} \rho^j g_{t+1+j}] + k \quad (7)$$

where  $k$  is not an important constant and  $\rho$  is a fixed parameter related to the mean ratio of price to dividend. In the Gordon model I know that

$$r_t = d_t/p_t + g \quad (8)$$

hence in the dynamic model I can say:

$$E_t[r_{t+1}] = (d_t p_t) - E_t[\sum_{j=0}^{\infty} \rho^j r_{t+1+j}] + E_t[\sum_{j=0}^{\infty} \rho^j g_{t+1+j}] - k \quad (9)$$

The stock price  $p$  will adjust endogenously, in order for the expected returns equal the rate of return. I can see that dividend yield is a good measure of expected returns, however for only one period, since it maybe correlated to future returns and future dividend growth.

In my empirical approach, I choose to insert both  $E/D$  and  $P/D$  in the model, since they measure two different aspects of returns.

For a company, we have the following general model:

$$r_{i,t+1} = \alpha + \beta(B/P_{it}) + \gamma (E/D)_{it} + \delta (P/D)_{it} + \dots + \epsilon_{it+1} \quad (10)$$

The general model is the foundation, on which different models will be constructed throughout the paper. It is used as an illustration of a cross-sectional regression. Throughout the paper, the book market value ( $b/p$ ), ( $e/d$ ) and ( $p/d$ ), will be replaced by their log counterparts, other independent variables such as age, industry, risk, volatility, and interactions variables made of combinations of the previously cited variables. Their coefficients will be studied in terms of “sign” and “significance”.

### **3. Empirical Analysis**

#### **3.1 Data Description**

The source of the data used in this study is the Wharton Research Data Services. It includes annual data for the years 1995 until 2005 for 30 companies. There are nine variables included in the dataset and they are described in Table 2. Note, AKOLLY30 is an index that was created to mimic the DOW JONES 30. All the DJ30 industries are represented in the AKOLLY30 index with the same proportions. The main difference between the two indices is the companies included.

The AKOLLY30 index uses companies that have more information available as opposed to the companies present in the DJ30, while maintain the same proportion of various industries.

**Table 2: Variables description**

<b>Variables</b>	<b>Description</b>
<b><i>Dividend yield (dp)</i></b>	Dividend yield: I divided the annual dividends by the year-end price of one share. Let $d$ be the monthly dividend, I define $d_t$ as the annual dividends in year $t$ , i.e. the sum of all dividends issued throughout the year, hence the dividend yield is $d_t/p_t$
<b><i>Log dividend yield (ldp)</i></b>	Logarithmic version of the dividend yield: $\ln(1 + d_t/p_t)$ .
<b><i>E/P ratio(ep)</i></b>	It was calculated using the year-end earnings divided the year-end price per share
<b><i>Log e/p ratio(lep)</i></b>	Logarithmic version of the e/p ratio: $\ln(1 + e_t/p_t)$ .
<b><i>Book market value(bm)</i></b>	I use the year-end data. The book market-value is $b_t/p_t$ , where $p_t$ is the price per share at the end of year $t$ and $b_t$ is the book value of year $t$ .
<b><i>Log book market value (lbm)</i></b>	Logarithmic version of the book market value: $\ln(1 + b_t/p_t)$
<b><i>Excess returns (eret)</i></b>	in this paper returns mean excess returns above the federal fund rates. The reason being, the risk free rate has its own behavior, and it is not our intent to model that in this paper. Returns due to risk is what crucial to study in



	this paper. The formula for excess simple returns is $(p_t+d_t)/p_{t-1}$
<b>Log excess returns ( leret)</b>	Logarithmic version of the excess returns: $\ln((p_t+d_t)/p_{t-1})$ .
<b>year</b>	Year ranging from: 1995-2005
<b>age</b>	Number of years after IPO.

Table 3 lists the different companies in the various industries in the respective indices.

**Table 3: Firms and Industries in AKOLLY30 and DJ30 Indices**

<b>AKOLLY30 companies</b>	<b>DOWJONES 30</b>	<b>Industries</b>
<b>Avery Dennison corp.</b>	3M	Industrials
<b>Ashland</b>	Alcoa	Materials
<b>Allstate Corp</b>	American Express	Financials
<b>Sprint Nextel corp.</b>	AT&T	Telecommunication
<b>Berkshire Hathaway</b>	Bank of America	Financials
<b>Eaton Corp.</b>	Boeing	Industrials
<b>HoneyIII</b>	Caterpillar	Industrials
<b>Valero Energy</b>	Chevron	Energy
<b>Apple</b>	Cisco Systems	Information Technology
<b>Pepsi co.</b>	Coca-cola	Consumer Staples
<b>Dow Chemical</b>	Du Pont	Materials

<b>Halliburton co.</b>	Exxon Mobil	Energy
<b>General Dynamics</b>	General Electric	Industrials
<b>Dell</b>	Hewelet Packard	Information Technology
<b>McAfee</b>	IBM	Information Technology
<b>Intuit</b>	Intel	Information Technology
<b>Stryker corp.</b>	Johnson & Johnson	Healthcare
<b>Morgan Stanley</b>	JP Morgan Chase	Financials
<b>Tyson Foods</b>	Kraft Foods	Consumer Staples
<b>Target</b>	McDonald's	Consumer Discretionary
<b>Boston Scientific</b>	Merck	HealthCare
<b>Citrix Systems</b>	Microsoft	Information Technology
<b>Bristol-Myers</b>	Pfizer	Healthcare
<b>Safeway</b>	Procter &Gamble	Consumer Staples
<b>Office Depot</b>	The Home Depot	Consumer Discretionary
<b>Capital one</b>	Travelers	Financials
<b>Raytheon</b>	United Technologies co.	Industrials
<b>Qwest</b>	Verizon Communications	Telecommunication Serv.
<b>Walgreen</b>	Wal-Mart	Consumer Staples
<b>Viacom</b>	Walt Disney	Consumer Discretionary

### 3.2 *General Diagnostics*

Due to the cross-sectional nature of our data we have a choice between two different methodologies: fixed effects and random effects analysis.

Fixed effects will help us explore the relationship between the predictor and outcome variables within a company. Each company has its own characteristics, which can have an impact on returns. The most important assumption we make when using fixed effects within a company is that they may possibly impact the predictor or outcome variables, hence the need to control for it. This explains the assumption of correlation between error term and variables that may impact returns. In other words, the fixed effect model takes care of the time-invariant characteristics, in order for us to get the explanatory variables' net effect. The other assumption of the fixed effect model is "that those time-invariant characteristics are unique to the individual and should not be correlated with other individual's characteristics" (Oscar Reyna).

Unlike the fixed effect model, random effect model will assume that variations across companies are random and uncorrelated with the independent variables that are part of the model. The crucial distinction between fixed and random effects is whether the unobserved individual effect embodies elements that are correlated with the regressors in the model, not whether these effects are stochastic or not" (Green 2008). Do I have reason to believe that differences across companies have some influence on our dependent variables? If yes, then I should use the random effect model.

We use the Hausman's test to decide which model to use. The null hypothesis is that the adequate model is the random effect model, and the alternative is the fixed effects model. The **Hausman test** or **Hausman specification test** is a statistical test in econometrics named after Jerry A. Hausman. The Hausman's test compares a more efficient model against a less efficient model but consistent model, to check if the more efficient model also gives consistent results. We initially estimate a random effects model, save the coefficients then compare them to the fixed model coefficients. The test tests the null hypothesis that the coefficients estimated by the efficient random effect estimator and the one by the consistent fixed effects estimator are the same. Using the Hausman test we find that the test statistic is 0.001, which suggests that we use fixed effects. We then, test for random effects by using the Breusch

Pagan Lagrangian Multiplier test; whose null hypothesis is that variance across entities is zero (i.e. random effect is appropriate). The **Breusch–Pagan LM test** (named after Trevor Breusch and Adrian Pagan) is used to test for heteroscedasticity in a linear regression model. It tests whether the estimated variance of the residuals from a regression are dependent on the values of the independent variables.

Suppose that we estimate the equation

$$y = \beta_0 + \beta_1 x + u. \quad (11)$$

We can then estimate  $\hat{u}$ , the residual. Ordinary least squares constrain these so that their mean is 0, so we can calculate the variance as the average squared values. A simpler method is to regress the squared residuals on the independent variables, which is the Breusch–Pagan test:

$$\hat{u}^2 = \beta_0 + \beta_1 x + v. \quad (12)$$

If an F-test confirms that the independent variables are jointly significant then I can reject the hypothesis of homoscedasticity, implying that there is heteroscedasticity.

The Breusch–Pagan test tests for conditional heteroscedasticity. It is a chi-squared test: the test statistic is  $n\chi^2$  with  $k$  degrees of freedom. If the Breusch–Pagan test shows that there is conditional heteroscedasticity, it can be corrected by using the Hansen method, using robust standard errors, or re-thinking the regression equation. In the context of seemingly unrelated regression, Breusch-Pagan (1979) proposed a Lagrange Multiplier (LM) statistic, which is valid for fixed  $N$  as  $T \rightarrow \infty$  and is given by

$$LM = T \sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij}^2 \quad (13)$$

where  $\hat{\rho}_{ij}$  is the sample estimate of the pair wise correlation of the residuals

$$\hat{\rho}_{ij} = \hat{\rho}_{ji} = \frac{\sum_{t=1}^T \hat{u}_{it} \hat{u}_{jt}}{\sqrt{(\sum_{t=1}^T \hat{u}_{it}^2)(\sum_{t=1}^T \hat{u}_{jt}^2)}} \quad (14)$$

And  $\hat{u}_{it}$  is the estimate of  $u_{it}$ . LM is asymptotically distributed as chi-squared with  $N(N-1)/2$  degrees of freedom under the null hypothesis of interest. However, this test is likely to exhibit substantial size distortions in cases where  $N$  is large and  $T$  is finite- a situation that is often encountered in empirical applications, primarily due to the fact that the Lm statistic is not correctly centered for finite  $T$  and the bias is likely to be worse with large  $N$ . With a test statistic of 0.06, I can reject the null and confirm our previous conclusion.

Next, we test for time fixed effect. We run a test to see if all time dummies are jointly equal to 0 (which is the null hypothesis) of an F-test, if they are then I do not need a time fixed effect model. With a test statistic of 0.0000, I reject the null, which indicates that a time fixed effect model is needed.

If there is a time fixed effect, it means that the variable “year” affect returns. It is safe to conclude that “year” might also affect dividends since dividends are also part of the return.

### 3.3 *Empirical Analysis (Using Data Subset)*

#### 3.3.1 *Dividends in Different Industrial Sectors*

As we have seen in the preceding part, the importance of the variable “companies” could be significant. We can hypothesize that the industry that the company belongs to affects its dividend (thus its returns). After all, intuitively it makes sense: we can hypothesize that, the fact that some industries have specific characteristics when it comes to their customers, resources, competition and market share structure, might affect their dividends (thus returns). Powell (1999) in his survey of corporate managers on how they view dividend policy, explores the relevance of industry in dividend policy among others explanations such as agency theory and tax preference explanation. He studies the view of managers about dividends in

different industries. He concluded that, industry does not have any influence on managers' point of view. This contradicts with the findings of Baker, Farrelly and Edelman (1985) that find out that there are concrete difference between utilities and manufacturing. Those results could be explained by greater competitive environment in utilities and regulations which has made the utilities industry riskier than before. Michel (1979) showed that there is a strong link between industry classification and dividend payout. Dhrymes and Kurz (1967), McCabe (1979) and Michel (1979) have all shown evidence that a firm's industry influence its dividend payout. The view of Higgins (1972) contradicts those of Michel. He finds no evidence that industry classification affects dividends. Rozeff (1982) solves the questions several ways, initially he examine the residual across all firms. The input of the data is done in order of industry. If there industry effects, they will manifest themselves by nonrandom variables, i.e. the residuals of the same sign and same industry group will be clustered together. He uses the DW<sup>11</sup> (Durbin-Watson) test; the test statistic of 1.88 shows that there some randomness in the residuals which shows that there is no industry effects. Secondly, he selects 8 different industries with a minimum of thirty companies each. He ran regression on each of them, then he use Chow tests<sup>12</sup> to test if all the coefficients are equal. The F-statistics ends up being too small and he ends up failing to reject the hypotheses that each industry has the

---

<sup>11</sup> The Durbin–Watson statistic is a test statistic used to detect the presence of autocorrelation in the residuals from a regression analysis. It is named after James Durbin and Geoffrey Watson. If  $e_t$  is the residual associated with the observation at time  $t$ , then the test statistic is

$$d = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2}.$$

Since  $d$  is approximately equal to  $2(1-r)$ , where  $r$  is the sample autocorrelation of the residuals,<sup>[1]</sup>  $d = 2$  indicates no autocorrelation. The value of  $d$  always lies between 0 and 4. If the Durbin–Watson statistic is substantially less than 2, there is evidence of positive serial correlation. As a rough rule of thumb, if Durbin–Watson is less than 1.0, there may be cause for alarm. Small values of  $d$  indicate successive error terms are, on average, close in value to one another, or positively correlated. If  $d > 2$  successive error terms are, on average, much different in value to one another, i.e., negatively correlated. In regressions, this can imply an underestimation of the level of statistical significance.

<sup>12</sup> The Chow test is a statistical and econometric test of whether the coefficients in two linear regressions on different data sets are equal. The Chow test was invented by economist Gregory Chow. In econometrics, the Chow test is most commonly used in time series analysis to test for the presence of a structural break. In program evaluation, the Chow test is often used to determine whether the independent variables have different impacts on different subgroups of the population.

same regression coefficient as the one in the pool sample model. Thirdly, he takes residuals from the regression of each individual industry and sees if they fall randomly around the regression line; if they do then there are no industry effects. Using a 5% confidence level, eight industries out of sixty four have residuals that are nonrandom. Those industries include Home Appliance, Fast Food Service, Securities Brokerage and Retail. Finally, he uses dummy variables for the eight industries mentioned previously and runs the regression. Six out the eight prove to be significant.

There are several issues with Rozeff's approach. First, when he runs regression on industries dummies without other variables that might affect dividends among the explanatory variables, there is a model misspecification problem: There is some variability in dividends that cannot be captured by the dummy variable "industry" alone. Second, the industry classification by Rozeff is flawed in the sense that some of them can be grouped together, for instance the Home Appliance industry and the Toiletries/ cosmetics industries can be grouped together under the "consumer products" industries.

In this paper we solve those two shortcomings, by initially choosing four distinct industries. We make sure they do not have anything in common and are distinct in each of their main characteristics. The four industries selected are: Financials, Healthcare, Information Technology and Utilities.

We picked 19 companies from each of the four industries. Dummy variables are created from each industry: 1 for financial industry, 2 for healthcare, 3 for information technology and 4 for utilities. We include dummy variables for all the years based on the following grouping: 1 from 1995 to 1997, 2 from 1998 to 1999, 3 for 2000 to 2001, 4 for 2002 to 2003, and 5 for 2004 to 2005.

We regress excess return on: dividend yield,  $E/P$  ratio, book-market value. We add the variables "*dyear*" and "*ind*", but also some interaction variables such as "*ind x ldp*" and "*dyear x ldp*", where *ind* is the dummy variable for industry and *dyear* is the dummy variable for year. The rationale being that the coefficient of *dyear.ldp*, not only tells us about returns, but also informs on the interaction between

“*dyear*” and “*ldp*”, in other words between year and dividend yield. In other words, does economic cycle (variable “*dyear*”) affect dividend yield (thus returns) on average. The same reasoning motivated the inclusion of the interaction variable “*ind.ldp*”. Does industry influence dividend yield? If yes, to what extent?

We start by running the Hausman’s test, in order to choose between fixed or random model. The test statistic of 0.000, suggests that we use a fixed effect model. We then checked for cross-sectional dependence and contemporaneous correlation: based on the **Pesaran CD test**. Pesaran (2004) has proposed the following alternative to the Breusch-Pagan LM test:

$$CD = \sqrt{\frac{2}{N(N-1)}} \left( \sum_{i=1}^{N-1} \sum_{j=i+1}^N \sqrt{T_{ij}} \hat{\rho}_{ij} \right) \quad (15)$$

Where  $T_{ij} = \#(T_i \cap T_j)$  (i.e. the number of common time series observations between units  $i$  and  $j$ ),

$$\hat{\rho}_{ij} = \hat{\rho}_{ji} = \frac{\sum_{t \in T_i \cap T_j} (\hat{u}_{it} - \bar{\hat{u}}_i)(\hat{u}_{jt} - \bar{\hat{u}}_j)}{\sqrt{\left[ \sum_{t \in T_i \cap T_j} (\hat{u}_{it} - \bar{\hat{u}}_i)^2 \right] \left[ \sum_{t \in T_i \cap T_j} (\hat{u}_{jt} - \bar{\hat{u}}_j)^2 \right]}} \quad (16)$$

and

$$\bar{\hat{u}}_i = \frac{\sum_{t \in T_i \cap T_j} \hat{u}_{it}}{\#(T_i \cap T_j)} \quad (17)$$

The test revealed a cross sectional dependence. In order to solve our cross sectional dependence problem, we will use the **Driscoll and Kray** standard errors’ method. Let consider the following linear regression model:

$$y_{it} = x'_{it} \theta + \varepsilon_{it}, i = 1, \dots, N \text{ and } t = 1, \dots, T$$



The dependent variable  $y_{it}$  is a scalar,  $x_{it}$  is a  $(K+1) \times 1$  vector of independent variables whose first element is 1, and  $\theta$  is  $(K+1) \times 1$  vector of unknown coefficients.  $i$  is the cross-sectional unit (“individuals”) and  $t$  denotes time. We assume that the regressors

and the scalar disturbances are uncorrelated. We can thus consistently estimate  $\theta$  by least squares (OLS) regression:

$$\hat{\theta} = (X'X)^{-1}X'y \quad (18)$$

Driscoll and Kraay standard errors for the coefficient estimates are the square roots of the diagonal entries of the asymptotic (robust) covariance matrix.

$$V(\hat{\theta}) = (X'X)^{-1}S_T(X'X)^{-1} \quad (19)$$

where

$$\hat{S}_T = \hat{\Omega}_0 + \sum_{j=1}^{m(T)} w(j, m) [\hat{\Omega}_j + \hat{\Omega}_j'] \quad (20)$$

where  $m(T)$  is the lag length up to which residuals may be autocorrelated. The  $(K+1) \times (K+1)$  matrix  $\hat{\Omega}_j$

is defined as :

$$\hat{\Omega}_j = \sum_{t=j+1}^T h_c(\hat{\theta}) h_{t-j}(\hat{\theta})' \text{ with } h_c(\hat{\theta}) = \sum_{i=1}^{N(t)} h_{it}(\hat{\theta}) \quad (21)$$

By relying on cross-sectional averages, standard errors estimated by this approach are consistent independently of the panel’s cross-sectional dimension  $N$ . Estimating the covariance matrix using this approach yields standard errors that are robust to very general forms of cross-sectional and temporal dependence.

The advantage of running a regression with Driscoll and Kray standard errors is that in addition to fixing cross sectional dependence, we do not have to worry about heteroskedacity and autocorrelation. The output of the regression is displayed in Table 4.

As expected, all the coefficients are significant, recall that  $ydp$  is the product of the year dummy and dividend yield, while  $idp$  is the product of industry dummy variable and dividend yield. The significance of the interaction variables' coefficients implies that there is some interaction effect between the dividend yield and the two dummy variables. This suggests that the industry to which a firm belongs to affects its dividend thus its returns, which was our hypothesis earlier. The interaction effect between years and dividend yield could be hypothesized as older firms pay less dividend than younger firms. Why? This would make sense if paying dividend (thus high dividend yield) was in fact a signal mechanism; a way of attracting investors and getting the trust of the market. To answer those questions, we look at the mean dividend yield by industry. Table 5 displays the mean dividend yield by industry. There seems to be a vast difference between dividend yields paid by industries. How can that be explained? Why do financials and utilities have the highest dividends yields? We address those questions in two folds. Initially we hypothesize that dividends might be in fact a signaling mechanism; second we check if dividends can be considered a compensation for risk.

### ***3.3.2 Dividends as a Signaling mechanism***

In this part of the paper, the following hypothesis will be tested using data: “Can dividend yield be a signaling mechanism, to attract investors?” It is fair to assume that most investors, use past performance as a predictor of future performance. Companies that have been successful and profitable in the past do not have to pay dividends, since they have “nothing to prove”. The central idea is that firms that are not as profitable as the average firms have to compensate for their mediocre performance. To verify our Hypothesis, we need a proxy for survival.

**Table 4: Dividends in different industries: Regression with Driscoll-Kraay standard errors of excess returns on explanatory variables**

Variables	Coefficients (DK standard errors)
<i>dividend yield ( log)</i>	-22.27** (9.12)
<i>book market value(log)</i>	-1.83*** (.36)
<i>industry x dividend yield</i>	4.42** (2.17)
<i>year x dividend yield</i>	1.33*** (.48)

**Notes:** \*\*\*: significant at  $p < 0.01$  level.

\*\* : significant at  $p < 0.05$  level.

The rationale being that a firm that has survived past recessions, competition, and change in its industry is more equipped to survive future challenges. Everything else equal, investors are willing to invest in that firm. It is fair to assume that everything else equal, investors will more likely invest in firms that have been around longer. For instance, most casual investors are more likely to invest in IBM than in a new start up in Silicon Valley. If we follow that rationale, we could use age of the company as a proxy to ability to survive in competitive market place. Let age be the number of years since the IPO<sup>13</sup>.

<sup>13</sup> According to Gregoriou (2006) an **initial public offering (IPO)** referred to simply as an "offering" or "flotation," is when a company (called the *issuer*) issues common stock or shares to the public for the first time. They are often issued by smaller, younger companies seeking capital to expand, but can also be done by large privately-owned companies looking to become publicly traded. In an IPO the issuer may obtain the assistance of an underwriting firm, which helps it determine what type of security to issue (common or preferred), best offering price and time to bring it to market. An IPO can be a risky investment. For the individual investor, it is tough to predict

**Table 5: Mean Dividend Yield by Industry**

<b>Industries</b>	<b>Mean dividend yield</b>
<i>Financials</i>	0.023
<i>Healthcare</i>	0.001
<i>Information Technology</i>	0.004
<i>Utilities</i>	0.043

Our assumption is that, in a capitalist market, firms that have been profitable survive through the years, while firms that have not been profitable either are bankrupt and disappeared or are taken over by more profitable companies. In summary I assume a sort of natural selection, where stronger (more profitable) firms survive, while the weaker (less profitable) firms are bankrupt or taken over. Miller and Modigliani (1961) showed that price per share is often an after effect of dividend rate. Powell (1999) affirms that due to information asymmetry between the corporate managers who has an information advantage and the investors, dividend might be a way to communicate the firm's future profitability, or in a lesser extent its potential future profitability. On the other hand, investors will use the dividend announcement as a measure of the stock price. Are dividends perfect signals? Easterbrook (1994), affirms that increases in

---

what the stock or shares will do on its initial day of trading and in the near future since there is often little historical data with which to analyze the company. Also, most IPOs are of companies going through a transitory growth period, and they are therefore subject to additional uncertainty regarding their future value.

dividends are “ambiguous”, unless the market can distinguish between growing firms and firms with lack of investment chances. To illustrate this point, Soter, Brigham, and Evanson (1996) give the example of the parent company of Florida Power and Light Company. In 1994, the parent company announced a 32% reduction in their quarterly dividend, the stock dropped by 20%. It is only after it was determined that the reduction in dividend was made not because of a problem in cash flow, but because of strategic reasons to invest in the company's growth. It is only after investors figured out this was the case that the stock recovered. Savov (2006) discusses the dividend-signaling hypothesis by analyzing the post-announcement performance of German companies. She specifically looks at characteristics such as earnings, level of earnings, assets growth, and capital expenditures. She also looks at the annual stock return and sees if signaling theory can be the cause of the negative correlation between the dividend decision and the stock price performance. Savov gets results that go against the theory. She finds out that there is not “any evidence that dividend increases or convey information about the future operating performance”. Companies do not perform significantly better after the announcement of dividends' paying. “Not only does earnings not increase after a dividend increase, they do in fact decrease” Savov (2006). The limitations of the models used by Savov and Modigliani, is the fact that their proxy for successful companies is flawed: For instance using percentage change in dividend payments as a measure of success or profit has its limit as I showed it earlier. We consider 30 companies from the utilities industry. We create a variable “*age*” which is the age of the company at the time the dividend, earnings and other characteristics are released. Compared to the regression in the previous part, we use the same variables and interaction variables, but this time *industry* and *year* variables are not considered, the variable “*age*” is introduced. We regress *excess returns* on *dividend yield*, *book market value*, *E/P ratio*, *age* and *age x dividend yield*. We control for industry by choosing firms from the same industry. The interaction variable *age x dividend yield* is expected to be significant; since it will confirm our hypothesis that dividends are in fact paid to signal “corporate health”. The Hausman test statistic of 0.000, suggests

that we should use a fixed effect model. The Pesaran cross-sectional dependence test suggests that there is a cross-sectional dependency. We will use Driscoll and Kraay standard errors to fix it. The output is in Table6.

**Table 6: Dividend as signaling mechanism: Regression with Driscoll-Kraay standard errors of log-excess returns on explanatory variables**

Variables	Coefficients (DK standard errors)
<i>Earning price ratio(log)</i>	-0.04 (0.12)
<i>Dividend yield(log)</i>	-2.37 (2.99)
<i>Book market value(log)</i>	-1.31*** (.32)
<i>Age</i>	0.024 (0.016)
<i>Age x log dividend yield</i>	-0.015 (0.02)

**Notes:** \*\*\*: significant at p<0.01 level.

After the DK Standard error regression, surprisingly only the book market value is significant. The interaction variable *age x log-dividend yield* is not significant. This could be explained by the fact that, investors prefer book-market value which is based on the balance sheet and the stock market, rather than the number of years the firm has existed. The book market value has two components, the value of the company in the books, which included asset, liabilities; and stock price which reflect the sentiment of the

market at the time. It is then justified that book-market value might be a better proxy of a firm's health, since it encompasses verifiable assets and market's sentiment at the exact time. Perhaps the use of book market value might be a good proxy for profit over.

In Table 6, after the utilities industry, financials was the second industry with the highest average dividend yield. It appears that risk might be a factor. Paying dividend might be a way of compensating for risk.

### ***3.3.3 Dividends as Compensation for Risk***

Rozeff (1984) studies dividend yield and equity risk premiums. He also looks at random walk and bond-stock spread. He uses the interaction between the three topics to study the dividend yield. If his hypothesis is right then, we can use dividend yield to measure equity risk premiums. The equity risk premium tells us how much additional return and investor want in exchange of owning the stock. When thinking about acquiring a stock, the anticipated returns that are being priced into the stock are unknown. Finding an estimate of the risk premium will help us choose between buying the stock or consider a less risky instrument such as Treasury bills. Rozeff (1984) uses two methods to estimates equity risk premium. First, the realized return method which was first introduced by Ibbotson and Sinquefeld (1982): they replace the unobservable risk premium by an average of their past actual market return less the actual Treasury bill return. Due to yearly variations, the realized risk premium varies from period to period. Ibbotson and Sinquefeld fix that deficiency by averaging the realized risk premiums over 56 years and getting a mean of 8.3%. One of the shortcomings of the method is the fact that, we have to assume a stationary risk premium distribution through time. The sample used has data that might be irrelevant to today's market. The authors defend themselves by mentioning the fact that the time series used comes from a stationary distribution, which justifies the averaging over extended time periods. The other method mentioned is the dividend yield method which is of interest to us in this paper. The author assumes that

growth rate of dividends is related to the economy's rate of output. This implies that dividend growth is equal to rate of interest. The dividend yield method of estimating risk states that: equity risk premium<sup>14</sup> is proportional to dividend and inversely proportional to Treasury bill interest rate.

This means that fluctuations in the equity risk premium can be measured by fluctuation in the dividend yield, since there is not much fluctuation in the Treasury bill's interest rate. Comparisons between the two different methods are displayed in Table 7.

**Table 7: Prediction Errors of Realized Return and Dividend Yield Methods, 1962-1982**

Predictor	Mean Error <sup>a</sup>	Mean absolute error <sup>b</sup>	Mean Square Error <sup>c</sup>
<i>Realized return method</i>	-0.076	0.1378	0.0317
<i>Dividend yield method</i>	-0.0181	0.1334	0.249

<sup>a</sup> Defined as arithmetic mean over 21 annual observations of  $(A-F_t)$  where A= actual risk premium and  $F_t$ =forecasted risk premium.

<sup>b</sup> Defined as arithmetic mean of  $|A-F_t|$ .

<sup>c</sup> Defined as arithmetic mean of  $(A-F_t)^2$ .

The forecasts of the realized return method have mean of 9.94%, while the mean realization was 2.34%. On the other hand the dividend yield method had a mean 4.15% which is very close to 2.34%. The average error for the realized return method is 13.8% versus 13.3% for the dividend yield methods. Our hypothesis, that dividend might be a compensation for risk is thus justified. Using the mean square error criterion, the dividend yield method's MSE is .0249 compared to the .0317 of the realized risk return

<sup>14</sup>  $ERP_t \approx \frac{d_t}{1+r_{bill}}$  (where  $ERP_t$  is the equity risk premium and  $r_{bill}$  is the Treasury bill interest rate).



methods. This has shown us the positive relationship between the dividend yield and the equity risk premium. This might explain why the utilities and the financial industries have the highest dividends. We hypothesize that dividend yield is in fact a compensation to the investor for taking risk. We will study and verify this hypothesis by observing the volatility of different companies in various sectors.

We observe 19 firms in 5 different industries. Here we will define volatility or risk by variance in price per share; we then run a regression of excess returns on *E/P ratio*, *dividend yield*, *book market value*, *volatility* and *volatility x dividend yield*. The Hausman's test suggest that I use a fixed effect model( test statistic of 0.000).The Pesaran CD test reveal a cross sectional dependence( test statistic of 0.000).To fix it we will use the Driscoll and Kay variances. The output is in Table 8, *volatility*'s coefficient became insignificant and was dropped, however the interaction variable's coefficient *riskyield* which is the product between dividend yield and volatility is significant which tell us that there is some interaction effect between the dividend yield and the volatility and it affects excess returns.

The significance of the coefficients of *the log-dividend yield*, *log-book market value* were expected, since it has been established earlier that they both affect excess returns. The significance of the interaction's variable (*volatility x dividend yield*) coefficient, tells us that dividend yield is indeed a compensation for risk, and it affects excess returns. This means that holding everything else constant, a company with high volatility in its stock price, will pay more dividends, in order to attract investors. This could also be explained by a self-fulfilling prophecy: A company noticing its lagging stock price, will pay more than average dividends in order to attract investors, his in turn will shoot up the stock price, since more investors are willing to buy it, thus inflating the stock. However later on, some investors will catch up to it and rapidly sell their share, which in turn will drive down the price and the cycle, goes on.

**Table 8: Dividend as compensation for risk: Regression with Driscoll-Kraay standard errors of excess returns on explanatory variables**

Variables	Coefficients (DK standard errors)
<i>dividend yield (log)</i>	-5.90*** (2.72)
<i>book market value (log)</i>	-1.81*** (0.37)
<i>dividend yield x volatility</i>	0.043*** (0.020)

**Notes:** \*\*\*: significant at  $p < 0.01$  level.

### 3.3.4 Excess Return Model

Earlier we have investigated variables that might influence dividend yield. We found out that volatility, years in industry, year's economy and type of industry all influence dividend yield. Can we extend that conclusion to excess returns? It would be a legitimate extension since we know that dividend is included into returns. We include all the variables I have studied earlier that affect dividend yield into this last regression. We are expecting better predictability. The Hausman's test suggest a fixed effect model and the Pesaran CD test reveals a cross sectional dependence. We use the DK variances methods; our output is in Table 9.

Of all the variables studied, the significant ones are *log dividend yield*, *log book market value*, *industry* and *industry x year x volatility x dividend yield*. The dividend yield (log), which was expected since dividends, is part of returns. The significance of the book market value (log) can be explained by the fact

**Table 9: Excess return model: Regression with Driscoll-Kraay standard errors of excess returns explanatory variables**

Variables	Coefficients (DK standard errors)
<i>dividend yield (log)</i>	-6.23*** (2.37)
<i>book market value (log)</i>	-1.83*** (0.40)
<i>industry</i>	0.72*** (0.055)
<i>industry x year x volatility x dividend yield</i>	0.0033*** (0.0012)

**Notes:** \*\*\*: significant at p<0.01 level.

that book value and current stock price are good predictors of company's health. The book value is the summary of all the known assets of the firms; hence it carries weight in investors' decision in buying or selling a stock. The current stock market price can be thought of as an opinion poll of the market about a particular firm. Most investors are risk averse; they will not buy a stock that other investors think of as less than average. Table 9 confirms that the industry a firm belongs to influence its excess returns. This was true for dividend yield, so we expected it for excess returns. Depending on the year or the month, specific industries are doing better than others. It is only fair to think that at a particular time, investors might think some industry is worth more of their capital than another one. Stock price being driven by the mechanism of supply and demand, the stock price of the industry in favor is more likely to go up. For instance, during recessions, discounters such as Wal-Mart's stock price goes up, because investors assumes that being in recession will force people from Whole food Market to Wal-Mart which is cheaper

competitor. Finally, the interaction variable between industry, year, volatility and dividend yield is significant. First it tells us that those variables influence each other. Second, the variable that deserves attention is *year*, since it is the only that has not been studied so far. Depending on the year, stocks prices might be rallying or retracting. *year* encompasses more parameters than can affect the markets and consumers, thus returns. For instance depending on the year, we might have the Federal Reserve lower interest rate. We might be in election's year or in a recession or in a war, all of which can affect a firm's performance in particular and the economy as a whole. All the variables that we have hypothesized as being determinants of returns have been proven to be significant.

## **Conclusion**

Our goal in this paper was to investigate the cause of changes in dividend, and the predictability of excess returns. We have shown that industries affect dividend yield. Especially industries such as utilities and financials; who tend to pay more dividends but also are the riskiest. In that case, it seems that dividend is in fact a compensation for risks. Supporting Rozeff's conclusion, in a cross-sectional setting, that there is a strong relationship between dividend and risk. Finally, we discuss the hypothesis that dividend might be a signaling mechanism; firms use it to signal to investors their economic "health". Companies' financial health is based on income statement and balance sheet. Companies that have good balance sheet will likely pay less dividends to their investors. On the other hand, firms with poor balance sheet will most likely increase dividend payment to investors in order to compensate for their poor performance, but also to attract more investors. All these hypotheses are consistent with each other. In the last part of the paper, we use all these variables to see if they can help predict excess returns. Even though all the variables we hypothesized about were significant, we had an R-squared of 21%. Our study has several implications. First, as far as trading strategies goes, it can help traders hedge their position in the equity market. Our study would have prevented people from investing Enron. Enron being one of the oldest companies on the NYSE, prompted investors to buy the stock. Our research shows that the age of the company is

insignificant in predicting returns; book market value is a much better predictor. If investors have paid attention to the book market value as our study suggested, they would have noticed accounting irregularities and avoid the stock. In our study, we showed that dividends were compensations for risk. A good illustration of that fact is the collapse of BP's stock price. BP being in a risky industry-the energy industry- has a history of paying dividend. As a result, the stock price was high. However, the oil spill in the Gulf of Mexico on April 22<sup>nd</sup> 2010 sent a shockwave through the market, lowering BP's stock price. It is naïve to think that the oil spill directly affected BP's stock price. After all, BP had a profit of 6 billion in the first quarter of 2010 costs associated with the oil spill would be merely a fraction of that amount. The stock price decreased rapidly (more than 40% as of June 21, 2010), only after there were rumors that BP might not pay dividends at the end of the year. As long BP was paying dividends, investors were willing to take the risk. If it stops, there is no incentive for them to take on that risk, so they will sell the stock, which will drive down the price. For future research, we suggest that historical profit margin be used as a proxy for the ability for a firm to survive in a competitive environment. Historical profit margin data would have been good predictor since it is related to profit but also it is relative to the revenues of the firm: it is a variable that will also capture the ability of the corporate managers to be efficient in their resources' allocation. Other predictor, worth researching is the fragmentation of the industry. Are there high barriers to entry? How concentrated is the industry? Do Industries that are made of few firms pay less or more dividends? Finally it would have been better if we could find a better measure of volatility other than variance which is not standardized; this would have made the model even better in terms of predictability.

## References

Baker, Kent, Gail Farelly, and Richard Edelman, “A Survey of Management Views on Dividend Policy”, *Financial Management*, 1985, Vol.14, No. 3, pp 78-84.

Breusch, T.S.; Pagan, A.R. , "Simple test for heteroscedasticity and random coefficient variation". *Econometrica* (The Econometric Society), 1979, 47 (5): 1287–1294.

Campbell, John Y. and Robert Shiller, “The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors”, *Review of Financial Studies*, 1988a, Vol1, No 3 pp 195-228.

Dhrymes, PJ and, Mordecai Kurz, “Investment, Dividend, and external finance behavior of firms:, *NBER chapter*, 1967, pp 427- 486.

Driscoll, J.C. and A.C. Kraay, “Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data”. *Review of Economics and Statistics*, 1998, Vol. 80, pp549-560.

Fama, Eugene and Kenneth French, “Dividends Yields and Expected Stock Returns”, *Journal of Financial Economics*, 1988, 22, pp.3-25.

Gordon, M.J., the Investment, Financing, and Valuation of the Corporation, *Irwin*, Homewood, Ill. 1962.

Hausman, J.A. “Specification Tests in Econometrics”, *Econometrica*, 1978, 46 (6), 1251-1271,

Higgins, Robert, “Dividend Policy and Increasing Discount Rates: A Clarification”, *Journal of Financial and Quantitative Analysis*, 1972, Vol.7, No.3 pp 1757-1762.

Ibbotson, R.G., and R.A. Sinquefield, “Stocks, Bonds, Bills and Inflation: The Past and The Future”, *Financial Analysts Research Foundation*, 1982.

Kandel, Shmuel and Robert Stambaugh, “On the Predictability of Stock Returns: An Asset-Allocation Perspective”, *Journal of Finance*, 1996, pp385-424.

Lamont, Owen, "Earnings and Expected Returns", *Journal of Finance*, Vol. 53, No 5 (Oct.1998), pp.1563-1587.

Michel, Allen, "Industry Influence on Dividend Policy", *Financial Management*, 1979, Vol.8, No. 3, pp 22-26.

Miller, M., and F. Modigliani, "Dividend Policy Under Asymmetric Information," *Journal of Finance*,40, 1031-1051.

Pesaran, "General diagnostic tests for cross section dependence in panels", *Cambridge Working Papers in Economics*, (2004), No. 0435.

Peterson, Pamela, "*Analysis of Financial Statements*", Wiley, (1999), p. 92.

Powell Gary, "How Corporate Managers View Dividend Policy", *Quarterly Journal of Business and Economics*, 1999, Vol.38.

Savov, Sava, "Dividend Changes, Signaling, and Stock Price Performance", *Mannheim Finance Working Paper*, 2006, No 3.

Shiller, Robert, "Stock Prices and Social Dynamics", *Brookings Papers on Economic Activity*, 1984:2, 457-498.

Silvey, S. D., "The Lagrangian Multiplier Test", *Annals of Mathematical Statistics* 30, 1959, 389-407.

Soter, Brigham, Evanson," The Dividend cut heard around the World", *Journal of Applied Corporate Finance*, 1996.

