Multi-Variable Regression Analysis
For the Prediction of Equity Returns Over 10 Year Periods

Arjun Singh Jaswal

Spring 2007

Professor Edward Tower, Faculty Advisor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University.

Duke University
Durham, North Carolina
2007

Acknowledgements

I would like to give special thanks to my faculty advisor, Professor Edward Tower for his guidance of this work, for his continuing collaboration and deep interest in my research, and most of all for his immense patience in reviewing my work and research. Without his tutelage this work would have been impossible.

Abstract

The use of 5 variables is examined in order to forecast *ex ante* the total return from holding equities over 10 year periods. The 5 variables are a moving average of Campbell and Shiller's P/E ratio, Robert B. Barsky and J. Bradford De Long's log price predictor, a function of James Tobin's q, the rate of change of GDP over 30 years and the rate of change of cash flow over 10 years. The significance of these variables is explained by considering them individually, simultaneously and finally under the architecture suggested by David Hirshleifer.

I. Introduction

The business of predicting equity returns has been one of considerable activity for many years now. Various pundits have claimed to possess an ability to predict consistently the movement of prices and returns from equities, and have offered a wide spectrum of formulae and methods to do so. A short time ago, with the recovery of the stock market from its lows of 2001, there was considerable optimism over the future of equity returns. However, ever since equity returns began to plateau last year, the pessimists have begun to return to the mainstream. As a result, once again, there is much debate about the future of equity returns from the S&P500.

In this paper I have reviewed briefly several pieces of academic research over the past years that propose methods of formulae to predict equity returns. This includes John Y. Campbell and Robert Shiller's proposal to use a moving P/E ratio, Robert B. Barsky and J. Bradford De Long's use of a log price predictor and James Tobin's q. I have also reviewed the work of David Hirshleifer, in which he outlines the general framework of investor psychology and their participation in the markets.

While I have extensively used the formulae proposed by the above authors in several regressions that appear in this paper, they have been modified extensively to fit the general framework that is proposed by Hirshleifer. In addition, I have proposed the use of variables such as the rate of change of GDP and of cash flows as predictors. I believe that the simultaneous use of these variables has resulted in a more coherent and reliable prediction of equity returns than is obtained by the use of any of these variables (especially in their original form) individually.

As a result, we come up with a general framework that can be used to predict equity returns for moving 10 year periods. It should be noted that we try to predict the *total* return from holding equities, i.e. the return to be obtained from both capital appreciation as well as dividend accumulation. This is because this gives a truer picture of the wealth earned by investors, and indicates to a better extent investor decision making. This method is also consistent with that followed by prominent works that attempt to predict equity returns, such as Modigliani and Cohn (1979) and Campbell and Shiller (1988).

The question of course is, why attempt to predict returns over 10 year periods specifically? The answer is a little hard to give, but mostly has to do with the predictability of returns. Equity returns are a notoriously hard thing to predict, and many authors and analysts have been made or broken in this field. Indeed, academicians such as Barberis, Shleifer and Vishny (1998) believe that returns follow a random walk and cannot be predicted. Most others though have come to settle on a middle path. The prevailing consensus, if there is such a thing in the debate on predicting equity returns, is that returns might be very hard to predict in the short run, but in the long run can be approximated by complex models and a judicious use of market experience.

While any period might seem arbitrary by itself, 10 years seems to be a long enough period to predict equity returns over. Most investors realize that it is hard to consistently make good returns in the market, and that in order to be reasonably confident of making returns one needs to be invested in the market for more than just a short period of time. As will be explained later on in this paper, 10 years is a time span that most investors seem to consider the medium-term for investing. Hence, in this paper, we attempt to predict the rate of return over a 10 year period. This decision is in part validated by the decision by other eminent authors, such as Campbell and Shiller (1998) and Harney and Tower (2003) to use the same time period in their estimation of future equity returns. A statistical study of this issue is presented later in this paper.

II. Literature review

In this section I have reviewed the methods and work of those authors whose theories have been taken into consideration in writing this paper.

1) David Hirshleifer

As compared to many other authors, David Hirshleifer prefers to treat the subject of predicting equity returns at a much more subjective level. In his work "Investor Psychology and Asset Pricing", Hirshleifer prefers to sketch a theoretical framework for the investor psychology that is at play in the markets. In addition, the paper serves as a survey of a wide range of literature by other authors in the field of equity valuation. This is very useful from a point of view of understanding the various biases that effect not only investor valuation, but also misevaluation in the markets.

Traditional models of asset valuation stress a rational approach where no biases exist. In such a system expected returns are a function of the risk borne by an investor. If actual returns do happen to deviate from expected returns, this is due to a tangible event that passes, not due to misevaluation on the investor's part. However, the author stresses that "expected returns are related to *risk* and to *investor misevaluation*" (p. 1534). In the author's opinion, it is this error by investors that causes the market to fluctuate in an unsteady manner, as well as causes the formation of bubbles.

One natural way to identifying mispricing according to the author is to include a benchmark value for any asset class being considered (p. 1556). For example, the author suggests that benchmarking for the stock market would mean measuring the return on a "cheap" security *vis-à-vis* that on an "expensive" security. It is natural that different classes of the same asset usually perform in a given way in the long-run. A departure from these patterns would indicate potential mispricing by investors.

Most importantly, the author also suggests predictability based on past returns (p. 1558). He points out that in many asset classes there are positive short-lag autocorrelations and negative long-lag autocorrelations. He suggests that the short-run phenomena can be explained by the presence of momentum in the markets and that the long-run phenomena can be explained by overreaction followed by a correction in the markets. The author points out a pattern of "post-event return continuation".

One possible explanation is that investor actions vary according to the availability of news for public versus private events (p. 1559). The author suggests that “a firm’s decision whether and when to engage in the event depends on whether there is any misevaluation”, and that the firm manipulates “information reported to investors in order to induce misevaluation”. Hence, the author tries to explain the aforementioned post-event return continuation by suggesting that there may be insider trading or privileged information in action. If this is the case, the argument can be made that there must be some evidence of trading on privileged information in the period leading up to the event, in the behavior of relevant market variables. In that case, the use of momentum variables could be beneficial to approximate the post-event effect.

When prices do overshoot their ‘justified’ price, there must be a period of correction. It is this correction that induces the long-run negative autocorrelation. The author also attempts to explain this by citing research indicating that the impulse response function to a favorable initial shock, the ‘private information signal’, is hump shaped (p. 1568). As more information arrives in the market, confidence in the private signal increases, resulting in overconfidence in the signal and an overreaction. Eventually, perfect information is available in the market and investors take a more rational position, resulting in the hump shape of the response. This hypothesis of investor behavior can help us evaluate our results later in this paper.

While all of the above imply momentum trading implicitly, the evidence for momentum trades is even stronger in some research that tackles the subject directly[1] (p. 1562). It seems that investors change their behavior in parallel with each other and in some cases indulge in positive feedback trading in response to news events. Finally, the author points out that there is some evidence for habit formation in the markets (p. 1571), which would also encourage momentum trading. It seems that investors who have been lucky in their recent activity become more risk-tolerant, while those who have suffered bad luck become more risk-averse.

---

[1] See also, Grinblatt, Titman, and Wermers (1995).

Finally, one should note what the bell curve implies for momentum variables. Since imperfect news is widespread in the market, and since this is corrected only in the long term, one should average momentum variables over very long periods of time in order to approximate the underlying trend in the market.

I have made extensive use of Hirshleifer's theories to both modify the variables proposed by other eminent academicians as well as to explain several of my results. This will be covered in later sections.

2) Robert J. Shiller and John Y. Campbell

In his paper titled "Stock Prices, Earnings and Expected Dividends", Robert Shiller focuses on developing a framework to predict future dividend payouts using a weighted average of moving average earnings as well as dividends, and current stock price. He thus introduces two variables – the P/E ratio and the P/D ratio – that he uses later on in his career to predict equity returns[2].

This raises the question: Is it better to use the P/E ratio or the P/D ratio? While in the aforementioned paper the author prefers to rely on the P/D ratio, it occurs to me that price-earnings ratios should be a better dependent variable in models used to predict stock price fluctuations. In fact, various authors have demonstrated that stock price is independent of dividend policy. For example, according to the Modigliani-Miller theorem, the value of the firm is unaffected by how the firm is financed, including the decisions management makes to take on equity or debt, or to issue dividends (1958). They argue that if a company decides on a lower payout ratio for the current period, then it does so to either reinvest in itself, to pay back debt, or to buyback shares from the market. All of these actions result in an implicit cash flow to investors. Reinvesting in oneself must result in a higher payout in subsequent periods. Similarly, paying back debt or buying back stock from the markets results in a higher value for stock that an investor

[2] See also, Shiller (1996) and Campbell and Shiller (1998).

holds. In the reverse case, when dividends are raised in the current period, this must result in a lower payout in future periods. Hence, in the long-term, shareholders must be compensated or charged for all changes in dividend payout policy.

We would naturally expect that adding dividend payout ratios to the equation would increase variability in the predictor, due to the uncertainty associated with the payout ratio. Hence dividends should show more variability than earnings do.

However, it seems the situation is the other way around. In Exhibit 1 it is apparent that while dividends seem to hug a definite trend line with an R-squared of 0.83, the earnings fluctuate about a trend line with a lower R-squared of 0.73. Thus it seems that dividends are indeed a more stable variable and measure of fundamental value.

In my opinion there are two reasons for this – one psychological and the other economic. While earnings must be a more fundamental measure of value, shareholders care only about what they receive in the form of a cash flow to them, i.e. dividends. While they may understand that foregoing dividends in the present would theoretically entitle them to higher payouts at a future date, they are still uncertain about whether these future payouts would actually materialize. After all, management is prone to make human mistakes; their plans for using their financial resources may not pan out. As such, they perceive a lower risk in current dividends than future dividends. As pointed out by James E. Walter, the level of future cash flows from operations may not be independent of the dividend payout policy (1963). Since management must for the most part award shareholders with the cash flow of their choosing, they choose to pay a stable dividend year after year. In years of lower earnings, the owners feel compensated by healthy dividend, while in years of higher earnings management can lower the dividend a little. Inherently, this makes dividends a more stable variable.

The other possible reason for this may be investors' holding periods. While in the long-term shareholders should be compensated for forgoing a higher dividend payout, in practice this may not occur during the investment horizon of the average shareholder, often a period as short as 10-15 years. Theoretically, he should still be compensated for this by capital appreciation in a sale of the stock to another investor. However, the buyer may not have the same confidence (from the logic above) as the seller in a higher future dividend payout.

However, I still propose to use the P/E ratio as a predictor of future returns, since it leads to a theoretically more rigorous proposal. This is validated by Shiller's own decision later in his career to use the P/E ratio to predict equity returns. In a paper he wrote just before his testimony before the Board of Governors of the Federal Reserve, he demonstrates that the moving P/E ratio is a great predictor of a later change in the real price (Shiller 1996). He believed that a moving average P/E ratio could explain as much as 51% of the variation in the log ratio of prices (10 years from present to present). An even more compelling case is made in Campbell and Shiller (1998), where they conclusively demonstrate the robustness of this measure. In addition, the P/E formula is a perfect example of the framework offered by Hirshleifer; it assesses the market by considering a measure of current price benchmarked by a long term indicator of value (earnings).

3) Robert B. Barsky and J. Bradford De Long

In their work titled "Why Does the Stock Market Fluctuate", the authors argue that to use the traditional Gordon formula to value equities investors must estimate the growth rate of dividends. While this growth rate may approach a constant value in the very long-term, as is considered in much of academia, investors have a much shorter time frame. They consider this growth rate of dividends to be time varying, and must estimate it based on their past experience and knowledge of the current market conditions. In their opinion, the long run is in fact simply an aggregation of short run events, and hence investor decision making in the short run should be given special importance.

Hence, the authors argue that the traditional "static" Gordon formula that is usually applied must be replaced with a more dynamic formula, in which the future growth rate of dividends is in fact driven by current conditions of the market. This growth rate will resemble closely a "distributed lag on past growth rates, with slowly declining weights". They suggest:

$$\frac{\partial p_t}{\partial d_t} = 1 + \left[\frac{1}{r-g_t}\right] \frac{\partial g_t}{\partial d_t}$$

Where lowercase "p" and "d" reflect logs of prices and dividends, g is the future growth rate of dividends, r is the discount rate, and t represents the time period in reference. Upon integrating this we are able to arrive at:

$$p_t = d_t - \ln(r - g_t)$$

The authors use this formula to predict movements in the log price. However, as shown later, this can also be used to estimate the rate of return from the market.

It should be noted that the inclusion of this measure makes our study far more complete. Most other measures that have been suggested by eminent authors and academicians, such as q, the P/E and P/D ratios, and so on, are based on either momentum trades or use a variation of the Gordon formula over very long periods. They hence try to eliminate the noise that is inherent in stock price movement, in order to estimate the fundamental value of equities. The above study, however, uses exactly the apposite approach. In claiming that the long run is simply an aggregation of short run events, and in proposing formulae that estimates a dynamic rate of growth, the authors attempt to estimate exactly the same noise that has been ignored by other academicians. This is what makes their method unique, and prompted an inclusion in this study.

4) James Tobin

In 1969 James Tobin proposed a framework to analyze decisions relating to monetary theory in the economy. In doing so he proposed the use of q, a variable to measure corporations' inclination to invest. I will go into this in more detail later.

Tobin defines his q as the ratio of marginal efficiency of capital relative to reproduction cost, to the return from holding capital. The claim is that for a company to consider investing in new capital, the marginal efficiency of capital relative to reproduction cost must be at least slightly greater than the return from simply holding capital. If it weren't so, an investor could simply purchase capital and hold it in order to post a gain. In the very long term Tobin's q would approach an equilibrium value of 1.

One of the key assumptions made by Tobin (p. 16) is that "spending decisions and portfolio decisions are independent". He argues that "as savers, people decide how much to add to their wealth", and as "portfolio managers, they decide how to distribute the net worth they already have". Unfortunately, this may be more than a little idealistic. As outlined by classic microeconomics, investors who save today do so at the cost of consumption today. However, their decision to consume or save today must be affected by the return they expect to make on the assets that are available to them for investment, i.e. upon interest rates. Further, the decisions that portfolio managers make have implications towards the returns that can be expected from these portfolios. Although there is no *direct* linkage between wealth allocation and spending, this causality through interest rates has implications on the independence of variables that Tobin works with. This would probably merit further investigation beyond the scope of this paper.

Tobin has also ignored the return from holding money. Even in the simple two asset "money-capital economy" suggested by Tobin, an investor has the option of holding money as well as capital. The 'marginal efficiency of capital' that an investor may purchase is done so not just at the cost of purchasing capital, but perhaps at the cost of simply holding money as well, i.e. the option of doing nothing with one's wealth.

There is another frequently cited argument against using Tobin's q to value the markets. What is it that differentiates a functional, profitable company from simply an aggregation of capital? As argued by Epstein and others, it's the intangibles such as human capital and training. One often cited argument against using Tobin's q is that it does not account for such intangibles in measuring value in the economy. In classical macroeconomics, there are two inputs to production in the economy: capital and labor. While even intangible capital such as patents and goodwill can be valued and bought, the same is not necessarily true for human capital, training or management quality. Yet, Tobin's definition of q attempts to measure value by ignoring labor in both the numerator and denominator of the formula. Even modern interpretations of Tobin's q suffer from this fault. The numerator, i.e. the total market value of the company, must obviously be affected by the labor quality of the company. Yet, the denominator is defined simply as the total asset value (or book value) of the company. The book value of a company is defined as the assets that shareholders would receive upon liquidation of the company. Clearly, if a company were to be liquidated, it would not receive in payments any of the labor of the former company! Even if the labor were bonded by contracts that transferred from the former company on to the shareholders (which in itself is highly unlikely), the very act of splitting up the human capital between the shareholders would decrease their value due to the lost synergies of working together.

A better definition of Tobin's q would include the marginal efficiency of labor relative to the cost of labor in the numerator, and the return from labor in the denominator (and probably some other adjustments that are surely beyond my faculties!).

Another point of contention in academic circles is the theoretical value for Tobin's q. It is widely accepted that the capital stock of individual companies as well as the wider economy must increase at the natural rate of growth of the wider economy. In order for this to happen, q must be at least marginally greater than 1, so that companies would see more profitability in investing in new capital rather than simply holding on to their existing assets. If this weren't so, there would be no addition of capital to the aggregate economy. However, in Tobin's view the value of q must be less than 1, for investment-saving equality to occur as described in the LM framework (p. 22-23). This agrees with the historical long-term average of 0.65. Yet another point of view is that in the long run, the equilibrium value of q must equal 1. This would however imply a no growth economy, or at the very least an economy with no growth in capital. This situation is very unlikely to occur, at least until a very distant future. Hence it seems that the theoretical value of q is very much under debate.

Tobin does offer (perhaps unknowingly) a possible explanation to the fact that the historical average of 0.65 is far less than one. In his analysis of an economy including money, capital and government securities, he asserts that if the return on money is fixed by an external authority such as perhaps the central bank that can force the market return on physical capital to diverge from its inherent value. As seen in Table 2 in his paper, one possible explanation is that a historic increase in the return on money has hence caused a decrease in the value of q to below the equilibrium value of 1.

While it may seem that I am not very optimistic about the use of Tobin's q in predicting equity returns, this is not entirely true. It has indeed been demonstrated to be a very good predictor of returns in research done by several prominent authors. Matthew Harney and Edward Tower demonstrate that "q beats all variants of the PE ratio for predicting real rates of return" (2003). Indeed, in predicting the return over a 10 year period they achieve an R-squared of 0.536, which is a bit higher than what Shiller could achieve with the moving average P/E ratio. This case is further built by Smithers and Wright in their work.[3] Finally, one should also note that Tobin's q fits the framework suggested by Hirshleifer by benchmarking price against a fundamental measure of value, the net worth of all the firm's assets.

While some deficiencies may exist is the definition of q, several of them are of a nature that cannot be resolved by a mathematical adjustment to q. For example, it is hard to put a value of the synergies or managerial expertise of a firm. Due to its demonstrated success in predicting equity returns, I have used Tobin's q, albeit with some modifications, as a predictor of equity returns.

[3] See for example, "Valuing Wall Street: Protecting Wealth in Turbulent Markets", Smithers and Wright (2000).

III. Methods and formulae

In this section I have outlined all the basic formulae that are used in this regression analysis. The basic regression analysis is given by:

10 year equity return = f (Shiller's P/E ratio, the Dynamic log Price predictor, the Q ratio, % Δ of GDP, % Δ of Cash Flows)

1) Measuring equity return

In this study our central purpose is to predict the total return from holding equities for periods of 10 years. The time series calculated contains real data, and was adjusted to account for dividend payouts during the period of holding by the investor. That is, the 10 year return is given by:

$$R = \text{10 year equity return} = \frac{P_t - P_{t-10} + (D_t + D_{t-1} + D_{t-2} + \ldots + D_{t-10})}{P_{t-10}}$$

2) P/E ratio

Campbell and Shiller's (1998) P/E ratio has demonstrated much success as a predictor of 10 year equity returns. In its classic form, as it was proposed by Shiller early in his career, the ratio obeys one of Hirshleifer's requirements for a good predictor of equity returns: it is a ratio of some measure of market value (in this case, the price) against a benchmark value for the asset (in this case, the earnings). However, I wished to modify the ratio in order to make it fit another requirement of Hirshleifer's overall framework of investor psychology. Hence, I used a moving 30 year average of earnings instead of just current earnings, to calculate the P/E ratio, i.e.

$$P/E\,(30)_t = P_t \,/\, \frac{(E_{t-1} + E_{t-2} + E_{t-3} + \ldots + E_{t-30})}{30}$$

This has 2 effects: First, it causes the P/E ratio to act as a momentum variable as well. Second, it leads to a more reliable value for the benchmark. By averaging the value used for earnings, the P/E ratio in any one given year does not over-respond to a sudden fluctuation of company performance in that year.

The question of course arises: What is the appropriate time period to average earnings over? In order to test the performance of the variables in predicting equity returns, I have used three time periods, besides of course, the present. These are: the short, medium and long term. Of course, to use these I must assign to these periods exact numbers. I made the decision to define a long-term investor as someone who stays invested in the markets for a period of 30 years. The reasoning was that most persons cannot start investing until they begin saving, which usually is at the age of 25 or so. A long-term investor stays invested in the markets for the entire duration of his ability to save. However, in 30 years or so, when this person is in their mid to late 50s, he or she must start cashing in the investments in order to pay for expenses such as retirement, health care or children's college. Hence, 30 years seems like a reasonable time-frame for the long-term investor.

Similarly, most investors realize that it is almost impossible to time the markets. As such, even a short-term investor must stay invested for at least a couple of years to be confident of making any returns on his investment. I thus decided that 5 years was an appropriate time frame for the short-term investor. Finally, 10 years seemed like an appropriate time frame for the medium-term investor, being that it is between 5 years and 30 years.

In the case of the Shiller P/E ratio, it is most appropriate to consider the longest time frame possible. In order to use a benchmark of value we must ensure that the value we use approaches as best it can some fundamental measure of performance. Thus, we can clearly not use a single year to measure earnings, since the random nature of earnings would result in too much variability. It is advisable to use as long a period as possible, so that the average of earnings approaches what is truly its fundamental value. As we will see soon, this is true for most variables that attempt to approximate the underlying trend in the market.

In a similar vein, it may be interesting to revisit Hirshleifer's hypothesis of earnings manipulation and the reliability of reported data. We can almost be sure that there is some manipulation of earnings reports in the market. This would mean that there are some individuals with privileged information, who in all probability would act upon this information. This would cause Hirshleifer's "bell curve" of returns to become more skewed. As a result, returns in the market due to news events related to earnings would normalize only in the long-term, once all participants have similar information. In order to ensure the most reliable data, we should use the longest period possible for the P/E ratio. Hence, I have decided to use a 30 year time span for the moving average of earnings that is used in the P/E ratio.

The time span we have decided to use has indeed been validated by the work of several prominent authors. Harney and Tower have demonstrated that of the three time periods of 10, 20 and 30 years, it is the 30 year time period that performs the best in predicting future returns. Indeed, Shiller himself uses a 30 year time span to average earnings in his later works, such as in the deposition before the Federal Reserve Board of Governors.

3) Dynamic log Price predictor

In their work, Barsky and De Long (1993) suggest the formula:

$$p_t = d_t - \ln(r - g_t)$$

The left hand side of the above equation is supposed to be a predictor of the log of the real price of the S&P500. In order to calculate this I took the discount rate to be the long term return on equity, calculated for the period 1871-2006 to be 4.28%.

It should be noted that the authors consider the expected future growth rate of dividends to be a "distributed lag on past growth rates, with slowly declining weights" (p. 293). They define the growth rate by:

$g_t = g_{t-1} + (1-\Theta)\varepsilon_t$

Here g_{t-1} represents the growth rate in the previous period (or the perceived "permanent" growth rate), ε_t represents the departure from momentum and Θ represents the level of the departure. They introduce the 2nd term in the function for growth to account for the 'noise' that exists in the time series for the fundamental growth rate of dividends in the economy. The authors find experimentally that the formula given above fits actual values to a surprisingly good extent for values of Θ above 0.95, with an optimum value of 0.989 (p. 301 and 306). For the purposes of this study, the best values were obtained by setting Θ to 0.96. Finally, the authors define ε_t as $\Theta^t g_0$, i.e. a function of past dividend growth.

I modified this variable to define 'the Dynamic log Price estimator', in order to predict the return from holding the index. In order to estimate the underlying fundamental trend in the market, the log price predictor defined above is averaged over 30 years. Then, the variable under consideration is given by:

$$DP = \text{Dynamic log Price estimator} = \frac{(p_t + p_{t-1} + p_{t-2} + \ldots + p_{t-30})}{30} - p$$

Here p is defined as the actual current price level. Hence, the level of departure of the log natural of the price prediction from the log natural of actual price indicates the expected future return from equities.

Why use a moving average of the price estimator at all? As outlined in the sections above, we must estimate the fundamental value of the price estimator in order to eliminate the noise that is persistent in such models. It may seem that this is contradictory to our earlier review of Barsky and De Long's (1993) work. However, this is not true. In their work, the authors define the error rate to be a function of past dividend growth. Hence, in the model used by the authors above, the expected future rate of dividend growth is purely a function of past dividend growth. One would expect that even such a model, professed to account for short-term fluctuations in the time series, would contain error terms. Hence, we are justified in averaging this variable over a long period of time. In this case, once again we use the longest possible period in order to approximate the most fundamental value of the log price estimator used by the authors.

4) The Q predictor

In order to estimate the fundamental q in the underlying market trend, one should consider a moving average value for Tobin's q. For purposes of this study we define:

Q predictor = $Q_{30} - Q$

Here Q is the present level of Tobin's q and Q_{30} is the value of Tobin's averaged over the past 30 years. Hence, the level of departure of the moving average q from the current level of q is an indicator of the expected future level of returns from equities.

Once again, we must justify the time period chosen to average Tobin's q over. As usual, we can find our justification in Hirshleifer's theory for the market's reaction to news events. We should note that the denominator for Tobin's q contains the market value of all of the company's physical assets, or in our case, all of the economy's physical assets. However, as noted before, there is much motivation for individual companies to misreport the value of their physical assets in order to meet market expectations. However, one would expect that after a significant period of time, the value of reported earnings would be corrected to their true value due to SEC action, self-correction or shareholder litigation.

In addition, like any other indicator of market value, the value of q tends to vary over time. In order to estimate the fundamental value of Tobin's q for the market, while keeping in view the investor experience that has built up in recently passed periods, one must average q over the longest possible period of time. Hence, we average q over 30 years, which we consider for purposes of this study to be the long term.

5) Rate of change of GDP

In my search for variables that ought to be good predictors of equity return, I refocused on what essentially I was trying to measure. By measuring the rate of return on equities, we approximate the rate at which the physical companies that exist in the economy are growing. Clearly, the aggregate of all firms in the economy is growing at the same rate as the economy, which should indicate that the rate of growth of economy should be a good measure of the return from equities.

Some might make the argument that the rate of change of GDP is not a valid indicator of equity returns in a study that focuses on a narrow index such as the S&P500. They might say that the S&P500 ignores most of the small firms that exist in the economy, and that may be included by a wider index such as the Wilshire 5000. As a result of the now famous 'size effect' that postulates a faster appreciation in small firms, while the Wilshire 5000 may grow at a rate approaching that of the GDP, the S&P500 must of course grow at a slower rate.

However, we find that the Wilshire 5000 has a long term growth rate of about 3.93%, whereas the S&P500 has a *higher* long term growth rate of about 4.5%. It looks like proponents of the size effect in the Wilshire 5000 do not have much credence in their argument. More importantly, the economy has grown at a compounded annual rate of about 6.5% since 1929. Thus, it seems that the growth rate of the S&P500 is indeed a closer approximation of the growth rate of the economy than the Wilshire 5000.

Even if the Wilshire 5000 were a better representation of the economy, it is safe to suggest that the limiting rate in the long run of the return on equities should still be the rate at which the GDP grows, as is pointed out by Shiller (1998). Hence, it should be a safe assumption to make that the rate of change of GDP is a good indicator of the expected return on equities. A fuller discussion of index selection will follow in the section on 'Data and Methods'.

Once again the now familiar question challenges us: What time frame should we use in order to calculate the rate of change of GDP? It would be natural to claim that the market would have a more skewed "bell-curve" for news that is more ambiguous in nature. This is almost always the case with events that affect GDP as well as reported GDP figures. These figures are prone to being revised for several years after they are released, since they are in essence simply guesstimated by the Bureau of Economic Affairs. One would expect that in the long term these variations should average themselves out. In addition, in the long term these errors in estimation would (hopefully) have been corrected for by the BEA. Therefore, one would expect that knowledgeable investors would have a reaction to GDP movements would be most defined in the long term. Hence, I have chosen to use a 30 year time span (the long term) to calculate the rate of change of GDP. The exact formula of the variable used is given below:

$$\text{GDP predictor} = \frac{(\text{GDP})_t - (\text{GDP})_{t-30}}{(\text{GDP})_{t-30}}$$

6) **Rate of Change of Cash Flows**

According to the classic discounted cash flow models, the value of a security is given by:

$$\text{DCF value} = \frac{(\text{Cash flows})_1}{(1+r)^1} + \frac{(\text{Cash flows})_2}{(1+r)^2} + \frac{(\text{Cash flows})_3}{(1+r)^3} + \ldots + \frac{(\text{Cash flows})_t}{(1+r)^t}$$

Where r is the discount rate and t is the time period referenced.

Hence, the level of cash flows from a security is an indicator of the price of the security. This must mean that the rate of change of cash flows should be an indicator of the rate of change of price, i.e. the return from the equity. Thus, I have used the rate of change of cash flows as a predictor of the 10 year equity return.

The exact definition of the variable used is given below:

$$\text{Cash flow predictor} = \frac{(\text{Cash flows})_t - (\text{Cash flows})_{t-10}}{(\text{Cash flows})_{t-10}}$$

This gives the growth rate of the cash flows over the 10 year period.

Obviously, the question arises: Why 10 years? As explained earlier in this paper, an implication of Hirshleifer's hypothesis of privileged information and data manipulation suggests that for such information one must use the longest possible time frame. One would expect that cash flows should be a much more concrete indicator of value than GDP is, since they are reported by individual companies rather than guesstimated by a body such as the Bureau of Economic Analysis. However, it is possible that this very feature of cash flows reduces their reliability. It is a well known fact that companies have an incentive to manipulate cash flow figures in order to project good health. In fact, these days it has become common for companies to review income and cash flow reports several years after they were released due to investigation by the SEC and other governing bodies. As a result, investor confidence in self-reported figures has always been a bit lower than in other measures of value. However, one would expect that over a long period of time, errors in these reported figures would correct themselves either due to SEC action or due to self-corrections by the reporting companies. In addition, over long periods of time these errors in reporting should average themselves out. Hence, it is best to use the longest possible time frame to calculate the rate of change of cash flows.

It should be noted that I have used a 10 year span instead of a 30 year span due to the lack of sufficient data. As will be noted later, data for cash flows was available only from 1946 to 2002, which gave us 57 years of data. Using a 10 year span would result in 47 periods for which the above variable could be calculated. If we were to use a 30 year span, this would result in only 27 periods of data, which would not give us statistically significant results.

IV. Data sources and Methods

This section details the sources for all the data used in this study as well as discusses index selection for the purposes of regression analysis. All the results are summarized in Exhibit 2 at the end of this report.

1) Index selection

In order to test the 10 year return from holding equities one must first decide on which index to use. The United States has several major indices that are actively traded by millions of investors everyday. Some of these are the Dow Jones 30, S&P500, NASDAQ, the Chicago Mercantile Board and the Wilshire 5000. However, most of these were simply too narrow to be considered for the purposes of this study. For example, the Dow Jones 30 contains only 30 stocks. A study of this index would suffer from major selection and survival biases. Similarly, the NASDAQ contains only stocks related to the field of technology. A study focusing on this index would suffer from biases due to the selection of only one industry. Hence, the only real choices we have are the S&P500 and the Wilshire 5000.

A short discussion of the comparison between these two indices was made in the section above that detailed the use of GDP as a predictor of the 10 year return from equities. In the end I made the decision to use the S&P500 as the index of reference for this study. There were several reasons for this.

First, as shown above, it appears as though the rate of growth of the S&P500 is a more realistic representation of the rate at which the GDP grows. This is important in light of the use of macroeconomic variables in this study. Second, as pointed out by Siegel, the S&P500 represents the vast majority of the capitalization of all firms in the economy. This supports our claim that the S&P500 is representative of the wider economy.

Third, and most importantly, using the S&P500 as our index of choice made available to us a wide range of ready to use datasets that focused on this index. For the purposes of index data, we had available the updated version of the dataset used by Robert Shiller (2002). This provided us with the price of the S&P500 from 1871 to 2006, as well as earnings and dividend figures for the period, on a monthly basis. The dataset also provides other supporting data, such as the consumer price index and the real value of the price of the index, as well as of earnings and dividends. We hence have available monthly data for a total of 136 years, i.e. a total of 1632 periods.

As noted before, the prices were adjusted to account for dividend issues and hence the index used represented the *total* return from holding equities.

For purposes of this study only annual data was used. The value used for a particular year was its beginning of January value. Hence, a total of 136 periods of data was used.

2) Shiller's P/E(30) and the Dynamic log Price estimator

The time series for these two variables was calculated using the series provided by Robert Shiller's aforementioned dataset. This is in general agreement with the work of Shiller as well as Barsky and De Long. Albeit Barsky and De Long use the dataset published by Shiller in 1989, it seems fair that their methods should continue to be consistent with the updated dataset published available at Robert Shiller's website. Only real values were used in calculating the P/E ratio as well as the dynamic log Price estimator.

A total of 136 periods of data was available for means of calculating these two variables. Hence, 106 (136 minus 30) periods of data was available for the Dynamic log Price estimator and for the Shiller moving average P/E ratio.

3) Tobin's q

The data series for Tobin's q is available for access at Stephen Wright's website, and is an updated version of the dataset used in Smithers and Wright in "Valuing Wall Street" (2000). The values used from this dataset are taken from the column titled 'Non Financial Equity q'. A total of 103 periods of data was available to us, from 1900 to 2003. Hence, for the average of q over 30 years, a total of 73 periods of data was available.

4) GDP and Real GDP

The values for GDP and Real GDP were obtained from the Bureau of Economic Analysis (BEA, 2006). For the purposes of regression analysis in this study only real values were used. A total of 78 periods of annual data is available to us, from 1929 to 2006. Hence, for the 30 year rate of change of Real GDP, 48 periods of data was available to us.

5) Cash flows

Figures for the cash flow levels were obtained from the aforementioned Smithers and Wright dataset (2000). These were converted to real values using the price index in the Shiller (2002) dataset. A total of 57 periods of data were available to us, from 1946 to 2002. Hence, for the 10 year rate of change of real Cash flows, a total of 47 periods of data were available to us.

V. Results and Analysis

In this section I have first tested the robustness of each of the five variables under consideration individually. They were tested for 2 measures: (1) is the time period we have selected for measuring the particular variable appropriate? (2) Does the variable act as a robust predictor of the 10 year return on equities?

I have then tested the variables simultaneously as predictors of the 10 year real return on equities. Finally, we have tried to answer the question whether 10 years was an appropriate time frame to predict equity returns over.

1) Shiller's P/E(30)

In order to test whether we chose the appropriate time period to average earnings over, I tested Shiller's P/E ratio calculated for the earnings averaged over time frames of 5, 10 and 30 years, in addition to a simple P/E ratio, as predictors of the 10 year return on equities. The regression, in general, is given by:

10 year return = C + a (P/E ratio)

Here C and a are unknown constants. The results are given in Exhibit 3. The plots for the significant regressors are attached at the end of this report, as Exhibits 4 through 9.

We can see from the results that while even the simple P/E ratio is a significant predictor of the 10 year return, the predicting power of the ratio increases with the period over which earnings are averaged. Clearly, we were correct in our decision to consider the longest possible time period to average earnings over. As the time span increases, the earnings average gets closer to the fundamental trend, leading to a better estimate of the fundamental P/E ratio of the index.

We can also note that the results are more or less consistent with Hirshleifer's hypothesis of how event news is disseminated into the market. Consequently, the market should have a bell shaped response to news events. From the coefficients of the variables above, we can see that the magnitude of the market response first increases and then decreases with time. Albeit the response to earnings news appears to be negative, the results are still in agreement with Hershleifer's hypothesis.

It should be noted that all the coefficients in the above regression analysis have negative coefficients. This must mean that investors and analysts consider an increasing P/E ratio as a departure from market fundamentals, and consider a return to lower P/E levels necessary. This is manifested in the form of decreasing prices, which results in lower returns. Thus, increasing P/E ratios indicate lower returns in the time ahead, and decreasing P/E ratios indicate higher returns in the time ahead. Hence, P/E ratios are a contrarian indicator of future return from equities.

2) Dynamic log Price estimator

In order to test whether we chose the appropriate time period to average earnings over, I tested the estimator averaged over time frames of 5, 10 and 30 years, in addition to a simple estimator using only present period data, as predictors of the 10 year return on equities. The regression, in general, is given by:

10 year return = C + a (Dynamic log Price estimator)

Here C and a are unknown constants. The results are given in Exhibit 10. The plot for the significant regressor is attached at the end of this report, as Exhibits 11.

We can see from the results that only the Dynamic log predictor averaged over the longest time period is a significant predictor of the future return from equities. This is in agreement with our earlier analysis. As the time span increases, the estimation of the fundamental trend in the market gets better.

It should be noted that the coefficient of the only significant variable in this analysis is positive. Hence, as the value of the predictor increases, the expected rate of return increases. How can we justify this? We have defined the Dynamic log Price predictor as:

$$\begin{aligned} DP &= p_t - p \\ &= d_t - \ln(r - g_t) - p \\ &= \ln\left(\frac{D_t}{r - g_t}\right) - \ln P \\ &= \ln\left[\frac{D_t}{P(r - g_t)}\right] \\ &= \ln\left(\frac{1}{1 + g_t}\right) \end{aligned}$$

Because, by the Gordon formula: $P_o = \frac{D_1(1+g)}{r - g}$

Hence, $DP = -\ln(1 + g_t)$

From the coefficient we have obtained, we can say that when this above term increases, then the expected return from equities increases. Hence:

$-\ln(1 + g_t)\uparrow$ implies that $R\uparrow$

Then $\ln(1 + g_t)\downarrow$ implies that $R\uparrow$

Then $1 + g_t\downarrow$ implies that $R\uparrow$

Then $g_t\downarrow$ implies that $R\uparrow$

In essence, when the predicted growth rate for the current period decreases, then the expected return from equities increases. This means that growth in the current period is a contrarian indicator for future returns. This can possibly be because investors may believe that the growth rate for dividends follows a mean-reverting path, and that a low current growth rate for dividends must mean that the future growth rate of dividends must be higher, resulting in a higher future return from equities.

3) Q predictor

In order to test whether we chose the appropriate time period to average Tobin's q over, I tested the estimator averaged over time frames of 5, 10 and 30 years, as predictors of the 10 year return on equities. The regression, in general, is given by:

10 year return = C + a (Q predictor)

Here C and a are unknown constants. The results are attached as Exhibit 12 at the end of this report. The plot for the significant regressor is displayed in Exhibits 13 attached at the end of this report.

From the results we can see that the only significant variable is the one that was calculated by using the longest span of time. This is in agreement with our earlier analysis; we were justified in our decision to average Tobin's q over 30 years.

We can also see that the coefficient of the significant variable is positive. Hence, an increase in the Q predictor results in an increase in the expected return from equities. How can we justify this? We define the Q predictor as:

Q predictor = $Q_{30} - Q$

Hence, when the Q predictor increases, then the difference between Q_{30} and current Q increases. This must mean that investors interpret an increasing Q predictor to mean that the current level for Tobin's q is below historical lows. If one is to believe that Tobin's q follows a mean reverting path, then this must indicate to investors that the current level of Tobin's q must increase in the coming years.

Now, Tobin's q is defined as the ratio between the market value of the company and the cost of purchasing the physical assets of the company in the open market. We know that the value of the physical assets of a company changes very slowly over short periods of time, and even over a period of 10 years shows nowhere near the variability as shown by the price of the company's shares. If Tobin's q shows a consistent increase over a period of several years, then we can safely assume that most of this increase must be due to an increase in the price of the stock. Hence, an increase in the predictor used above must indicate an increase in the future return from holding equities.

4) Rate of Change of GDP

In order to test whether we selected the appropriate time span to calculate the rate of change of GDP over, I tested the rate of change of GDP over time spans of 1, 5, 10 and 30 years as predictors of the 10 year return from equity. In general, the regression is given by:

Return from equity = C + a (Rate of change of GDP)

Here C and a are unknown constants. The results are given in Exhibit 14 and the plot for the significant regressor is shown in Exhibit 15.

We can see that in general the predictive power of the rate of change of GDP increases with the time span over which the rate of change is calculated. This can be seen by the increasing R-squared of the variable with increasing time spans. It also seems like the response of the market to news events concerning GDP obey Hirshleifer's bell curve, with the response first increasing and then decreasing with time.

However, for purposes of coefficient analysis we should consider only the rate of change of GDP over 30 years, since this is the only statistically significant predictor. We can see that this coefficient is negative, indicating that as the rate of change of GDP increases, the rate of return from equities over the next 10 years decreases. This could be because an increasing GDP, especially at a rate that is above historical averages, might indicate to investors that the future growth of GDP must decrease in order to revert back to historical averages. Since, as discussed before, the rate of change of GDP is an implicit limit on the rate of return from equities, this must indicate that the future rate of return from equities is going to decline. Hence, the rate of change of GDP is a contrarian indicator of the future return from equities.

5) Rate of Change of Cash Flows

In order to test whether we selected the appropriate time span to calculate the rate of change of cash flows over, I tested the rate of change of cash flows over time spans of 1, 5, 10 and 30 years as predictors of the 10 year return from equity. In general, the regression is given by:

Return from equity = C + a (Rate of change of Cash Flows)

Here C and a are unknown constants. The results are given in Exhibit 16 and the plot for the significant regressor is shown in Exhibit 17.

It seems that we were right in estimating that the longer the time span over which the rate of change of cash flows is calculated, the more the predictive power of the variable. This can be observed by the increasing R-squared of the variable with increasing time spans. Once again, it seems that the variable obeys Hirshleifer's bell curve. As the time span over which the rate of change is calculated increases, the market response to growth in cash flows first increases and then decreases. Once again, note that we do not use the rate of change of cash flows over a 30 year period due to lack of sufficient periods of data.

We can also note that all the variables above have positive coefficients, save the rate of change of cash flows over a 1 year period. However, as a variable that is strongly insignificant as a predictor of the 10 year return on equities, it can be easily ignored. Clearly then, an increase in the rate of growth of cash flows indicates an increase in the future rate of return from equities.

Why might this be? One possible explanation is that investors have faith in the management abilities of the companies' leadership. When investors observe that a company has increased its cash flows to its shareholders, they might take this to be an indication that the company is operating at the peak of its financial abilities. Normally, a company issues dividends, or increases its dividends, when it cannot find any more productive stream to invest in (including reinvesting in it self). Thus, an increase in cash flows from dividends indicates to investors that companies are doing well. Obviously, this will cause them to buy up more of the stock, thus driving up prices and hence returns.

6) Multi-variable regression analysis

I now tested all of the aforementioned variables as predictors of the 10 year real return from holding equities. The following function was tested:

10 year equity return = f (Shiller's P/E ratio, the Dynamic log Price predictor, the Q ratio, % Δ of GDP, % Δ of Cash Flows)

The results were as shown in Exhibit 18, and the plot for this response is attached as Exhibit 19 at the end of this report.

We can see from the results that the combination of all the variables discussed is a significant and good predictor of the real return from holding equities for 10 years. It seems that we have been validated in our theoretical analysis of the above variables. However, a closer analysis of the results obtained is necessary.

We can see from the P-values obtained that several of the variables seem to be insignificant predictors of the future return from equities. This seems to be contradictory to our earlier analysis, where each variable individually tested to be a significant predictor in the regression. The seeming incoherence is due to the dynamics of how these variables behave in a regression simultaneously. In this case, the values obtained above do not necessarily indicate an insignificant predictor, but in fact indicate than a particular variable may be deemed unnecessary when analyzed in tandem with the other variables. In order to truly gauge the importance of any one variable, one must still analyze that particular variable in isolation from other predictors.

For example, in the above analysis, it seems that both the rate of change of GDP and the Dynamic log Price predictor are insignificant predictors of the return from equities. If we remove these from the multi-variable analysis above we get the results shown in Exhibit 20. The plot for this response is shown in Exhibit 21.

As can be seen from the results, the predictive power of the combination falls significantly, even though only seemingly insignificant variables were removed from the multi-variable analysis. This justifies the claim that in order to establish the true significance of a variable as a predictor of the return from holding equities, it should be analyzed in isolation from other predictors.

The question of course arises as to why the rate of change of GDP and the Dynamic log Price predictor seem to be insignificant when used in combination with the other variables. Considering the strong predictive power displayed by these variables when we analyzed them individually, especially for the rate of change of GDP, this is an important question. The answer probably lies in the fact that there must be certain correlations between some of the variables that have been used in this analysis. This could either be due to causality from one of the variables to another in the analysis above, or due to causality from some unknown variable to two or more of the variables used in the multi-variable regression. In either case, this makes our analysis considerably more complex; one of the basic principles of multi-variable regression analysis is that the variables being used must be independent of each other.

However, this does not necessarily pose a problem for us. This is because the above study does not necessarily claim direct causality from any of the variables to the 10 year real return from equities. Instead, the implicit claim is that certain movements in this real return are accompanied by certain movements in the independent variables. The causality could certainly be from some other unknown source.

As a side note, while this may begin to sound like a justification for data mining, it is certainly not so. Even though we might not be certain of the underlying cause of market movements in the economy, especially in a context where modeling can be very complex and include hundreds of cause-and-effects, we can analyze how and why certain movements in some variables are interpreted by investors, as well as how this understanding in investors causes them to participate in a certain way in the market. If we can explain conclusively this chain of investor psychology, then we can be justified in using the above variables in our analysis.

For example, we are certainly not sure of how and why real GDP falls and rises, since there are literally innumerable causes of such changes. However, we do know that when GDP rises up in a period, investors believe that it raises the prospects of it falling in subsequent periods, since GDP must follow a mean reverting path. Hence, investors believe that the limiting rate for growth in equity returns must fall in subsequent periods, resulting in a fall in expected returns. Hence, without knowing what the underlying cause of the movement in GDP is, we can use this movement to justify a movement in the expected return from equities. This is true even when what caused the GDP to change may have an effect on another variable in our analysis, such as say Tobin's q.

What does this mean for the particular analysis under consideration? It means that while the rate of change of GDP and the Dynamic log Price predictor may appear to be insignificant in the multi-variable regression, this is only because of the dynamics of their interaction with other variables being used as predictors. We must continue to consider all of the five variables in tandem with each other, and test their significance as predictors of equity return only in isolation from one another.

Finally, it is also important to conduct a coefficient analysis of the results obtained. The signs of the coefficients obtained are in general the same as those obtained when the predictors were analyzed individually, except for the sign of the coefficient for the Q predictor (i.e. $Q_{30} - Q$). However, as argued before, this is probably due to the interaction of this variable with the others contained in the regression analysis. In order to conduct a coefficient sign test, one must consider the variables in isolation.

7) Period of equity returns

Now that we have decided on our list of variables, we come back a full circle to answer one of the questions we posed at the start of this study: What is the appropriate time period to measure returns over? In order to test our assumption that 10 years is an appropriate time period, I conducted a regression analysis of the return over various time periods against the variables outlines above. The results are shown in Exhibit 22.

The initial expectation was that the explanative power observed for the return over 1, 3 or 5 years would be significantly less than the explanative power for returns over a 10 year period. While correct in this assumption, we are pleasantly surprised by the high values of R squared obtained for predicting returns over periods of 1, 3 and 5 years as well. It seems like not does our multi-variable analyses predict returns over long periods of time, but it can also do so over the short and medium terms.

VI. Conclusions

Before, we close the door on this study we must consider what the purpose of this study was. While there has been much research by prominent academicians on predicting future equity returns, most of these have focused on using indicators that did not truly consider the psychology of market participants. In the case that they did, the authors did not seem to emphasize on how and why they came to settle upon the methods they now promote. For example, while Shiller (2002) does argue that it is necessary to average earnings over a very long period of time, he considers nearly any selection of time period to be arbitrary and does not put much effort into justifying why he chose to average them over 30 years specifically.

In this study, however, emphasis was laid on fitting each of the variables into an overall framework of investor psychology, as is extensively described by Hirshleifer (2001). For example, the exact reasoning for why 30 years seems to be an appropriate time span to average earnings for the purposes of a P/E ratio was explained in the frame of reference of investor psychology. Many other variables that have in the past been used to justify price levels were modified to fit this psychology as well, such as those offered by Barsky and De Long (1993). In addition, the results obtained from such an analysis were interpreted not only using traditional tools such as an explanation of market forces, but also using tools such as Hirshleifer's bell curve and his hypothesis about event news reactions by market participants.

We found that all the variables discussed in this paper are theoretically appropriate as predictors of 10 year total return from holding equity, and that our results were mostly in line with our expectations.

No study is truly worth its salt unless it can claim that the methods proposed therein are an improvement over the understanding that existed to that point. Hence, the claim here is that a true understanding and application of investor psychology is essential to a better forecasting ability for the markets.

References

1. Barberis, N., A. Shleifer, and R. Vishny. "A model of Investor Sentiment." *Journal of Financial Economics*, Vol. 49, Iss. 3, p. 307-343. September 1998.
2. Barsky, Robert B., and J. Bradford De Long. "Why Does the Stock Market Fluctuate" *Quarterly Journal of Economics*, Vol. 108, Iss. 2, 291-311. May 1993.
3. Campbell, John Y. and Robert J. Shiller. "Valuation Ratios and the Long Run Stock Market Outlook". *Journal of Portfolio Management*, Vol. 24, Iss. 2. December 1988.
4. Campbell, John Y., and Robert J. Shiller. "Stock Prices, Earnings and Expected Dividends". *The Journal of Finance*, Vol. XLIII, No. 3. July 1998.
5. Epstein, Gene. "Smither's Contention About Q Is All Fouled Up". *Barron's*. Vol. 80, No. 20, p. 45. May 15, 2000.
6. Epstein, Gene. "The Q Controversy: Gene Epstein Replies". *Barron's*. Vol. 80, No. 23, p. 68. June 5, 2000.
7. Grinblatt, M., S. Titman, and R. Wermers. "Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior". *American Economic Review* Vol. 85, Iss. 5, 1088-1105. December 1995.
8. Harney, Matthew, and Edward Tower. "Rational Pessimism: Predicting equity returns using Tobin's q and price-earning ratios". *Journal of Investing*, Vol. 12, Iss. 3, p. 58. September 2003.
9. Hirshleifer, David. "Investor Psychology and Asset Pricing". *The Journal of Finance*, Vol. LVI, No. 4. August 2001.
10. Modigliani, Franco and Merton H. Miller. "The Cost of Capital, Corporation Finance and the Theory of Investment". *American Economic Review*, Vol. 8, No. 3, p. 261-297. June 1958.
11. Modigliani, Franco, and Richard A. Cohn. "Inflation, Rational Valuation and the Market." *Financial Analysts Journal*, Vol. 35, Iss. 2, p. 24. March/April 1979.
12. *National Economic Accounts*, Bureau of Economic Analysis. Dataset available at http://www.bea.gov/national/index.htm#gdp. Retrieved on February 25th, 2007.

13. Shiller, Robert J. "Price Earning Ratios as Forecasters of Returns". Available online at http://www.econ.yale.edu/~shiller/online.htm. July 1996. Retrieved on November 8, 2006.
14. Shiller, Robert J. *Irrational Exuberance* data, 2002. Updated version available at http://www.econ.yale.edu/~shiller/data/ie_data.htm. Retrieved on September 28th, 2006.
15. Siegel, Jeremy J. *Stocks for the Long Run*, 3rd edition. New York: McGraw-Hill, 2002.
16. Smithers, Andrew, and Stephen Wright, *Valuing Wall Street: Protecting Wealth in Turbulent Markets*. New York: McGraw-Hill, 2000. Updated dataset available at http://www.econ.bbk.ac.uk/faculty/wright/.
17. Tobin, James. "A general equilibrium approach to monetary theory". *Journal of Money, Credit and Banking*, Vol. 1, No. 1, p. 15-29. February 1969.
18. Walter, James E. "Dividend Policy: Its Influence on the Value of the Enterprise". *The Journal of Finance*, Vol. 18, No. 2, p. 280-291. May 1963.

Exhibit 1: Stability of real earnings and dividends

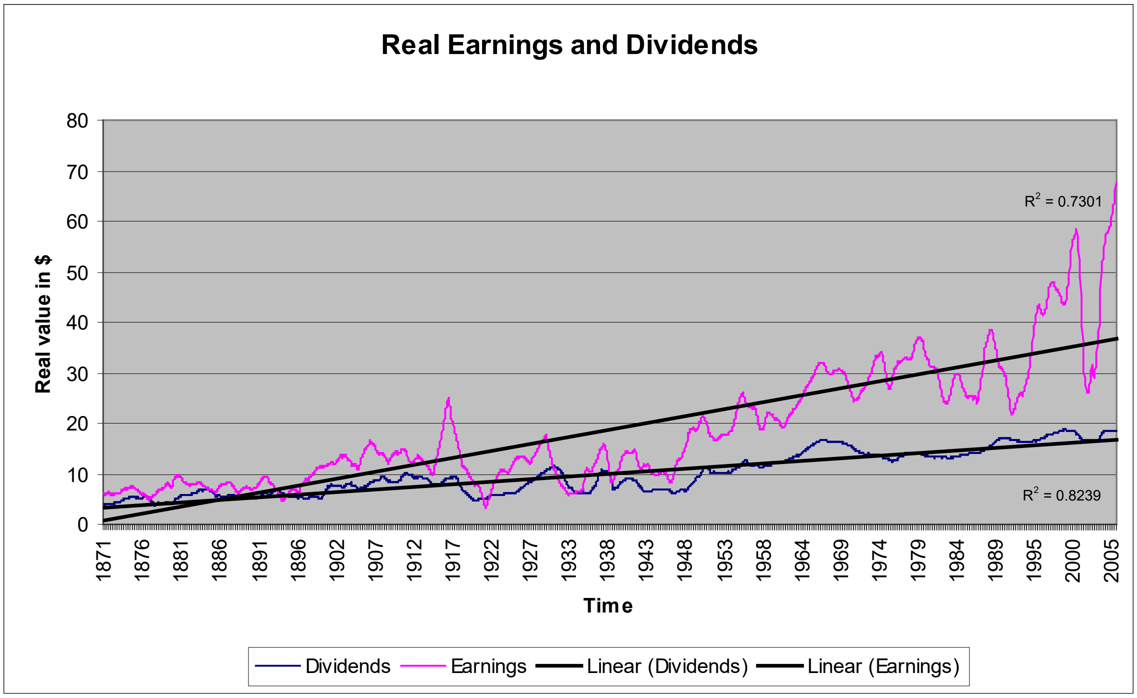

Exhibit 2: Data sources

Variable	Start year	End year	# of periods	Source
S&P500	1871	2006	136	Shiller (2002)
Earnings	1871	2006	136	Shiller (2002)
Dividends	1871	2006	136	Shiller (2002)
Price index	1871	2006	136	Shiller (2002)
Tobin's q	1900	2003	103	Smithers and Wright (2000)
Cash flows	1946	2002	57	Smithers and Wright (2000)
GDP	1929	2006	78	BEA (2006)

Exhibit 3: P/E ratios as predictors of the 10 year return

Variable	Coefficient	R-squared	t-ratio	P-value
Simple P/E	-0.0767	0.1669	-4.98	<0.0001
P/E (5)	-0.0801	0.2307	-5.97	<0.0001
P/E (10)	-0.0839	0.2615	-6.35	<0.0001
P/E (30)	-0.0737	0.3146	-6.57	<0.0001

Exhibit 4: Simple P/E as a predictor for equity returns

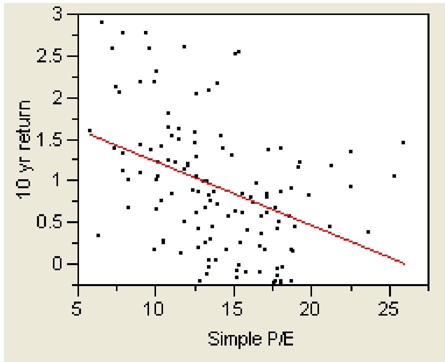

Exhibit 5: P/E (5) as a predictor for equity returns

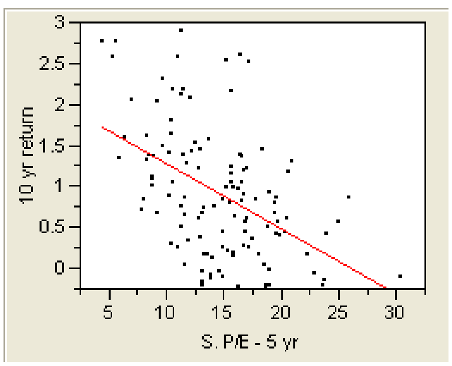

Exhibit 6: P/E (10) as a predictor for equity returns

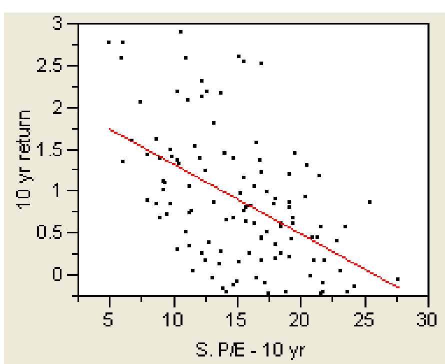

Exhibit 7: P/E (30) as a predictor for equity returns

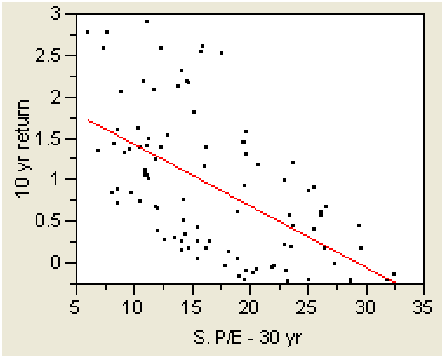

Exhibit 10: The Dynamic log Price predictor as a predictor of the 10 year return

Variable	Coefficient	R-squared	t Ratio	P-value
$p_t - p$	-0.0229	0.0058	-0.85	0.3951
$p_t (5) - p$	-0.0162	0.0025	-0.55	0.5823
$p_t (10) - p$	0.0002	2.209e-7	0.01	0.9960
$p_t (30) - p$	0.1364	0.0854	2.98	0.0037

Exhibit 11: p_t (30) – p as a predictor of equity returns

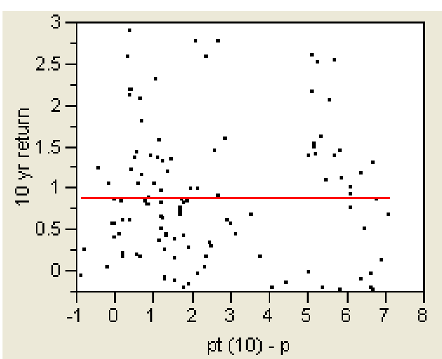

Exhibit 12: $Q_{avg} - Q$ as a predictor of the 10 year return

Variable	Coefficient	R-squared	t Ratio	P-value
$Q_5 - Q$	-0.3776	0.0077	-0.84	0.4044
$Q_{10} - Q$	0.1643	0.0076	0.81	0.4213
$Q_{30} - Q$	0.7209	0.1194	2.97	0.0042

Exhibit 13: Q_{30} – Q

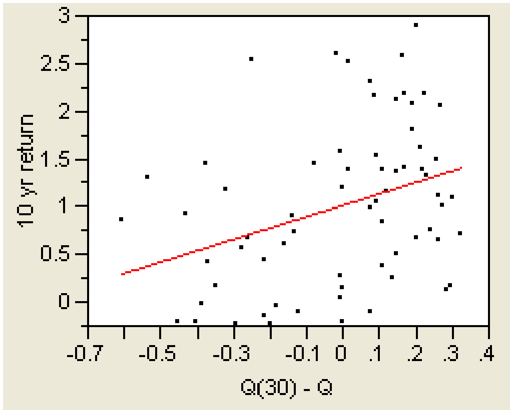

Exhibit 14: % Δ GDP as a predictor of return

Variable	Coefficient	R-squared	t-ratio	P-value
% Δ in GDP over 1 year	-0.6401	0.0017	-0.34	0.7375
% Δ in GDP over 5 years	-0.7006	0.0232	-1.20	0.2333
% Δ in GDP over 10 years	0.1280	0.0020	0.34	0.7364
% Δ in GDP over 30 years	-0.7266	0.2712	-3.66	0.0008

Exhibit 15: % Δ in GDP over 30 years

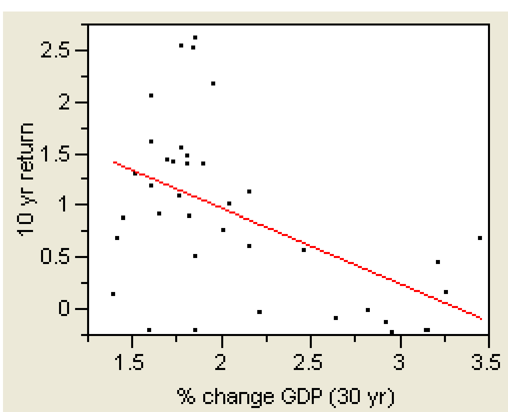

Exhibit 16: % Δ in Cash Flows

Variable	Coefficient	R-squared	t-ratio	P-value
% Δ in CF over 1 year	-0.0876	0.0014	-0.26	0.7940
% Δ in CF over 5 years	0.2775	0.0442	1.43	0.1608
% Δ in CF over 10 years	0.3508	0.2338	3.45	0.0014
% Δ in CF over 30 years	0.2552	0.5279	4.61	0.0002

Exhibit 17: % Δ in CF over 10 years

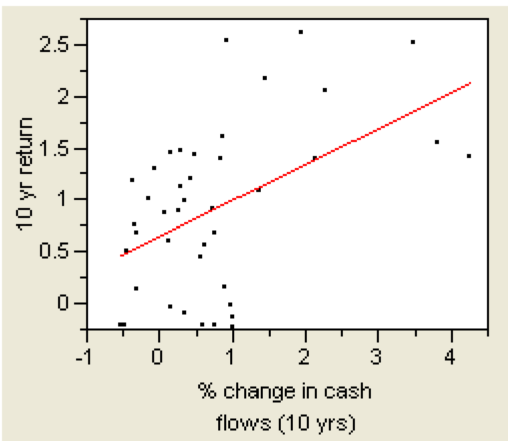

Exhibit 18: Multi-variable analysis

R squared: 0.6667
Adjusted R squared: 0.6146
F Ratio: 12.7999
Prob > F: <0.0001

Variable	Coefficient	t Ratio	P-value
P/E (30)	-0.0794	-1.97	0.1161
p_t (30) – p	0.0678	0.57	0.5732
Q_{30} – Q	-1.0534	-1.56	0.1279
% Δ in GDP over 30 years	-0.0908	-0.38	0.7081
% Δ in CF over 10 years	0.2614	3.17	0.0034

Exhibit 19: Multi-variable analysis response

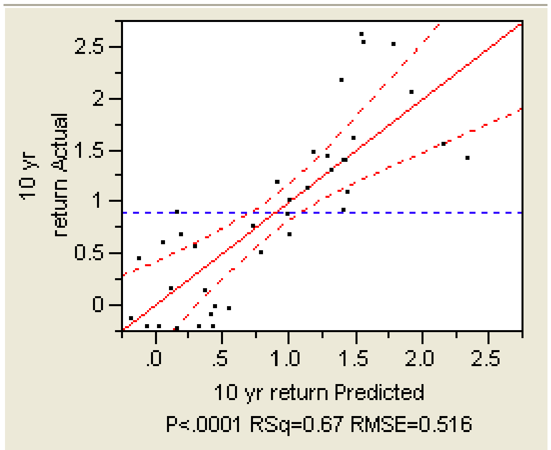

Exhibit 20: Multi variable analysis including only significant predictors

R squared: 0.5938
Adjusted R squared: 0.5609
F Ratio: 18.0293
Prob > F: <0.0001

Variable	Coefficient	t Ratio	P-value
P/E (30)	-0.1163	-3.09	0.0039
$Q_{30} - Q$	-1.3571	-2.05	0.0481
% Δ in CF over 10 years	0.2509	3.10	0.0037

Exhibit 21: Response

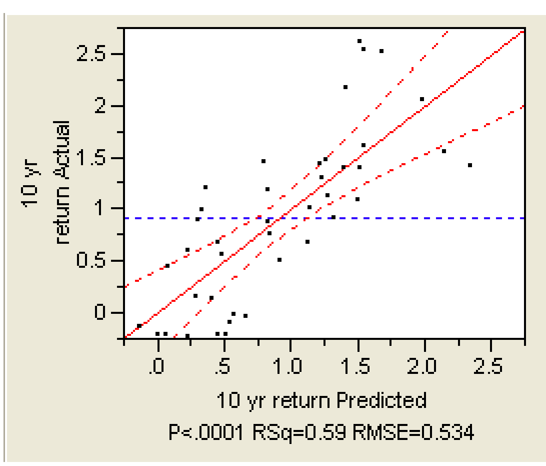

Exhibit 22: Period of equity returns

Dependent variable	R squared	F Ratio	Prob > F
1 year return	0.6101	11.8937	<0.0001
3 year return	0.4428	6.0404	<0.0003
5 year return	0.4929	7.1943	<0.0001
10 year return	0.6667	12.7999	<0.0001