This Time It's Different: Speculative Asset Bubbles & Adaptive Expectations

Conor Sheehy

Senior Honors Thesis

Department of Economics
Boston College

Advisor: Harold Petersen

For Caroline and Shawn Sheehy

I would like to thank my advisor, Harold Petersen, for all of the time and wisdom that he has put into this project. I've learned a tremendous amount about the functioning of financial markets from Professor Petersen, and have been honored to pull him out of retirement this year. I could not have asked for a better mentor to work with on the nature of speculative bubbles.

I would like to thank Alan Marcus, Amy Hutton, Arvind Navaratnam, Dmitriy Muravyev, Michael Barry, and Pete Wilson for inspiring me to explore and evaluate the functioning of financial markets. I'd also like to thank Peter Ireland and Rosen Valchev for their comments on this paper. Without your mentorship, I would not have completed this undertaking.

ABSTRACT

Using insights from Hyman Minsky's Financial Instability Hypothesis (FIH), we develop a theoretical framework for how speculative bubbles may materialize in securities markets. Our model and empirical analysis show that agents place undue emphasis on recent experience of risk and returns when developing future expectations.

We use the aggregate investor allocation to equities (aggregate total market capitalization of equities divided by the price of all real liabilities outstanding), Tobin's Q (the aggregate market price of equities divided by the replacement cost of nonfinancial firms' assets), Shiller Total Return Cyclically Adjusted Price to Earnings Ratio (TR CAPE), and Shiller Cyclically Adjusted Price to Earnings Ratio (CAPE) as proxy variables for bubbles. We find statistically significant, negative relationships between all four of these proxy variables and two dependent variables, Subsequent Ten-Year Annualized Cumulative Equity Market Returns (Nominal and Real), and also Subsequent 10-year Average Losses, thereby providing evidence against the Efficient Market Hypothesis and suggesting the possibility of speculative bubbles.

Keywords: Financial Instability Hypothesis (FIH), Efficient Market Hypothesis (EMH), Joint-Hypothesis Problem, Asset-Pricing Model, Mutual Funds, Stock Market, Equities, Debt, Speculative Bubbles, Stability, Business Cycles, Fama-French 3-Factor Model, Black-Scholes Option Pricing, Random Walk, Ten-Year Returns, Real Returns, Systematic Risk, Size Premium, Value Premium, Hyman Minsky, Eugene Fama, Kenneth French, Michael Burry, Capital Asset Pricing Model (CAPM), Risk, Average Loss, Adaptive Expectations, Market Psychology, Aggregate Investor Allocations, Cyclically Adjusted Price to Earnings Ratio, Tobin's Q

I. Introduction

The Efficient Market Hypothesis is by far the most important and commonly accepted theory regarding the functioning of financial markets. Its essential truth that financial markets quickly incorporate new information into prices serves as a bedrock of modern financial theory. However, the assumptions that the hypothesis relies on are questionable. Milton Friedman explains how these two statements can hold simultaneously for such an important theory:

> "In so far as a theory can be said to have 'assumptions' at all, and in so far as their 'realism' can be judged independently of the validity of predictions, the relation between the significance of a theory and the 'realism' of its 'assumptions' is almost the opposite of that suggested by the view under criticism. Truly important and significant hypotheses will be found to have 'assumptions' that are wildly inaccurate descriptive representations of reality, and, in general, the more significant the theory, the more unrealistic the assumptions (in this sense). The reason is simple. **A hypothesis is important if it 'explains' much by little, that is, if it abstracts the common and crucial elements from the mass of complex and detailed circumstances surrounding the phenomena to be explained and permits valid predictions on the basis of them alone. To be important, therefore, a hypothesis must be descriptively false in its assumptions**; The converse of the proposition does not of course hold: assumptions that are unrealistic (in this sense) do not guarantee a significant theory takes account of, and accounts for, none of the many other attendant circumstances, since its very success shows them to be irrelevant for the phenomena to be explained."[1]

One of the core tenets of the Efficient Market Hypothesis is that investors are rational, and maximize expected returns for a given level of expected risk. While this seems to be fairly reasonable, the hypothesis also stipulates that investors have homogenous expectations, and that they incorporate all accessible information into these expectations.

We disagree with this assumption; we believe that people develop irrational expectations. Since the value of financial assets are not directly observable, investors must make guesses as to the actual value of the underlying asset. Investors can vary wildly in their methods of analysis, expected holding periods, risk preferences, as well as any number of other factors. The market is a voting mechanism by which investors cast their ballots and hope that their analysis of a financial asset is closer to true value than everyone else. If they are correct (or lucky) they profit, and if they are wrong (or unlucky) they take losses. As a result, we imagine "market price fluctuations as a consequence of a seesaw or pendulum-like mechanism, by which prices 'orbit' around value, such that $P_t \approx V_t$, most of the time."[2]

We believe that prices can (and do) diverge so far away from this unobservable value that it may constitute a speculative bubble; the presence of which is hypothetically impossible under the

[1] Milton Friedman "The Methodology of Positive Economics" Essays In Positive Economics (Chicago: Univ. of Chicago Press, 1966), pp. 3-16, 30-43.
[2] Sherman, John. "The Efficient Market Hypothesis, the Financial Instability Hypothesis, and Speculative Bubbles". BA, 2014. https://dlib.bc.edu/islandora/object/bc-ir%3A102322

Efficient Market Hypothesis. We hypothesize that this is because investors place undue emphasis on recent experience of returns and risk when developing expectations of the future.

II. Background

Empirical testing of the Efficient Market Hypothesis must be done with regard to an asset pricing model. The difficulty in testing the hypothesis is that pricing models that appear to be empirically invalid do not necessarily invalidate the EMH. Most proponents would argue that the asset pricing specifications are simply incorrect, and that the theoretical basis for these asset pricing tests is still correct. This is the famous joint hypothesis problem.

Since asset pricing tests are incapable of invalidating the hypothesis, we take a different approach. Some of the most commonly used asset pricing models, in particular the Capital Asset Pricing Model[3] and the Black-Scholes Option Pricing Formula[4], both assume that stock prices exhibit Brownian motion (stochastic price movements with drift) which is broadly consistent with the idea that price shocks are the result of new information being incorporated into prices.

The presence of these price shocks from new information could have significant effects when people extrapolate out recent experience of returns and risk when forming expectations of the future. A series of random shocks in one direction could result in a slight divergence in value, however, if "investors start to believe that there is something new about the market"[5], the implication is that future expectations of returns and risk will be biased.[6] Following a series of positive shocks, expected returns will be biased upwards while expectations of risk will be biased downwards, and vice versa. In the paper "Contrarian Investment, Extrapolation, and Risk", the authors evaluated the "Value" factor that is commonly used in many models constructed using Arbitrage Pricing Theory, and hypothesized why high book to market firms may have outperformed low book to market firms over their analysis period:

> "Individual investors might focus on glamour strategies for a variety of reasons. First, they may make judgment errors and extrapolate past growth rates of glamour stocks, such as Walmart or Home Depot, even when such growth rates are highly unlikely to persist in the future. **Putting excessive weight on recent past history, as opposed to a rational prior, is a common judgment error in psychological experiments, and not just in the stock market.** Alternatively, individuals might just equate well-run firms with good investments regardless of price. After all, how can you lose money on Microsoft or

[3]Although every asset pricing model is a capital asset pricing model, the finance profession reserves the acronym CAPM for the specific model of Sharpe (1964), Lintner (1965), and Black (1972) discussed here. Thus, throughout the paper we refer to the Sharpe – Lintner – Black model as the CAPM. (Fama French 2003)

[4] Black, Fischer, and Myron Scholes. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy*, vol. 81, no. 3, 1973, pp. 637–654. *JSTOR*, www.jstor.org/stable/1831029.

[5] Sherman, John. "The Efficient Market Hypothesis, the Financial Instability Hypothesis, and Speculative Bubbles". BA, 2014. https://dlib.bc.edu/islandora/object/bc-ir%3A102322

[6] Sherman, John. "The Efficient Market Hypothesis, the Financial Instability Hypothesis, and Speculative Bubbles". BA, 2014. https://dlib.bc.edu/islandora/object/bc-ir%3A102322

> Walmart? Indeed, brokers always recommend "good" companies, with "steady" earnings and dividend growth."[7]

The authors go on to explain that while institutional investors should be freer from these sorts of judgement biases than individuals, and would ordinarily flock to these types of strategies that appear to be a "free lunch", there are other concerns that they need to deal with:

> "Another important factor is that most investors have shorter time horizons than are required for value strategies to consistently pay off (**De Long *et al.* (1990)** and **Shleifer and Vishny (1990)**). Many individuals look for stocks that will earn them high abnormal returns within a few months, rather than 4 percent per year over the next 5 years. Institutional money managers often have even shorter time horizons. They often cannot afford to underperform the index or their peers for any nontrivial period of time, for if they do, their sponsors will withdraw the funds."[8]

This career risk that institutional investors face is just one of a few reasons why money managers may make decisions that are not in the best interests of their clients when incentives are not perfectly aligned. Note that this is largely because of client demands on managers though. If investment managers were able to act without constraints, they would likely arbitrage a situation where investment expectations are biased, and make money for his/her clients. This major constraint on the "smart money" managers in the marketplace further emphasizes the point that financial asset markets may develop speculative bubbles based on biased expectations because aggregate level asset allocations are not being made by the "smart money"; it is being made by the average person who knows very little about the stock market.

a. Asset Pricing Models

Since the EMH can only be tested against a model of asset pricing, the EMH is essentially impossible to disprove with this method. Any empirical test of the EMH using an asset pricing model that fails to explain returns cannot be used to disprove the EMH; the focus always comes back to the variables utilized in the asset pricing model. According to John Sherman, "All asset-pricing models depend on an accurate assessment of risk to determine value."[9]

Prior to the 1990s, the most commonly accepted asset pricing model was the Capital Asset Pricing Model based on the work of William F. Sharpe (1964)[10], John Lintner (1965)[11], and

---

[7] LAKONISHOK, J. , SHLEIFER, A. and VISHNY, R. W. (1994), Contrarian Investment, Extrapolation, and Risk. The Journal of Finance, 49: 1541-1578. doi:10.1111/j.1540-6261.1994.tb04772.x
[8] LAKONISHOK, J. , SHLEIFER, A. and VISHNY, R. W. (1994), Contrarian Investment, Extrapolation, and Risk. The Journal of Finance, 49: 1541-1578. doi:10.1111/j.1540-6261.1994.tb04772.x
[9] Sherman, John. "The Efficient Market Hypothesis, the Financial Instability Hypothesis, and Speculative Bubbles". BA, 2014. https://dlib.bc.edu/islandora/object/bc-ir%3A102322
[10] Sharpe, William F. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk." *The Journal of Finance*, vol. 19, no. 3, 1964, pp. 425–442. *JSTOR*, www.jstor.org/stable/2977928.
[11] Lintner, John. "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *The Review of Economics and Statistics*, vol. 47, no. 1, 1965, pp. 13–37. *JSTOR*, www.jstor.org/stable/1924119.

Fischer Black (1972)[12]. The model stipulates that the expected return for a given asset i is equal to the risk-free rate (for the holding period) plus beta of the asset times the expected market risk premium over the risk-free rate:

$$E(R_i) = r_f + \beta_i * [E(R_m) - r_f]$$

where $\beta = \frac{covariance\ of\ returns_{i,m}}{variance\ of\ returns_{m}}$ and the United States Treasury Yield (over the holding period) usually acts as a proxy for the risk-free rate. The ex-post version of the test is as follows:

$$R_i = r_f + \beta_i * [R_m - r_f] + \varepsilon$$

Eugene Fama and Kenneth French summarize the attractiveness and issues of the model:

> The attraction of the CAPM is that it offers powerful and intuitively pleasing predictions about how to measure risk and the relation between expected return and risk. **Unfortunately, the empirical record of the model is poor – poor enough to invalidate the way it is used in applications.** The CAPM's empirical problems may reflect theoretical failings, **the result of many simplifying assumptions.** But they may also be caused by difficulties in implementing valid tests of the model. For example, **the CAPM says that the risk of a stock should be measured relative to a comprehensive "market portfolio" that in principle can include not just traded financial assets, but also consumer durables, real estate, and human capital.** Even if we take a narrow view of the model and limit its purview to traded financial assets, is it legitimate to further limit the market portfolio to U.S. common stocks (a typical choice), or should the market be expanded to include bonds, and other financial assets, perhaps around the world? **In the end, we argue that whether the model's problems reflect weaknesses in the theory or in its empirical implementation, the failure of the CAPM in empirical tests implies that most applications of the model are invalid.**[13]

Ordinarily, empirical tests of asset pricing models include the stock market as a proxy for the market because of ease of access to data; this fact alone makes empirical pricing tests extremely questionable because the true theoretical market return is nearly impossible (if not impossible) to observe. Using the expected market premium denoted by $[E(R_m) - r_f]$ is also questionable. The "risk-free" rate of return isn't actually risk-free. There is risk inherent in holding the United States' Federal Treasury Securities. The United States can only sustain a debt burden of so much before it would be unable to pay back its creditors in real terms – a country can only increase tax receipts by so much through explicit taxes and/or seigniorage before economic output falls due to overly burdensome taxes and/or hyperinflation causing instability.

---

[12] Black, Fischer, (1972), Capital Market Equilibrium with Restricted Borrowing, *The Journal of Business*, **45**, issue 3, p. 444-55.

[13] Fama, Eugene F. and French, Kenneth R., The Capital Asset Pricing Model: Theory and Evidence (August 2003). CRSP Working Paper No. 550; Tuck Business School Working Paper No. 03-26. Available at SSRN: https://ssrn.com/abstract=440920 or http://dx.doi.org/10.2139/ssrn.440920

While Fama mentions that the CAPM has been a failure at least empirically, and possibly theoretically, it does not necessarily preclude the presence of another superior asset pricing model that is empirically valid. Arbitrage Pricing Theory is the study of market factors which may close the gap between the predicted returns of a basic model like CAPM and provide portfolio managers with superior returns by closing these gaps. However, the EMH implies that any variable incorporated into an asset pricing model should account for some risk related to the asset being valued. The problem is compounded when variables which appear to carry no risk are incorporated into models show greater explanatory power than those with seemingly solid theoretical risks.

Eugene Fama and Kenneth French created an empirically superior model to the CAPM in 1992.[14] Their 3-factor model of asset pricing in 1992 uses market beta as well as 2 additional factors: "size" (market capitalization) and "value" (book price of equity/market price of equity).

$$E(R_i) = r_f + \beta_{i1} * \left[E(R_m) - r_f\right] + \beta_{i2} * E(SMB) + \beta_{i3} * E(HML)$$

Where SMB is the market capitalization premium (small minus big), and HML is the value premium (high B/M minus low B/M). While empirical tests of this model tend to perform better than the CAPM, the evidence suggesting that these are additional risk factors is weak. Eugene Fama himself has argued that the three-factor model's grounding is very weak:

> "I've spent a good part of the last 40 years testing those models. And a result of a lot of that is the so-called Fama-French three-factor model. **It's widely used both by academics and in industry. [He chuckles.] I'm laughing because the theoretical basis for the model is quite shaky. Basically, we saw these patterns in returns and our motivation was to try to explain them."**[15]

Whether these proxy for risk is also questionable. Fama explains:

> The discussion […] assumes that the asset-pricing effects captured by size and book-to-market equity are rational. For BE/ME (book price of equity to market price of equity), our most powerful expected-return variable, there is an obvious alternative. **The cross-section of book-to-market ratios might result from market overreaction to the relative prospects of firms. If overreaction tends to be corrected, BE/ME will predict the cross-section of stock returns.**[16]

While the theoretical basis for the CAPM seems to make sense (despite its empirical failings) because it captures some element of the systematic risk that portfolios may fall with the market, size and value are both highly questionable in comparison. Do small stocks outperform because they are riskier, or because investment managers have discounted their value irrationally? Do

[14] Fama, Eugene F., and Kenneth R. French. "The Cross-Section of Expected Stock Returns." *The Journal of Finance*, vol. 47, no. 2, 1992, pp. 427–465. *JSTOR*, www.jstor.org/stable/2329112.
[15] Sommer, Jeff. "Eugene Fama: King of Predictable Markets". *The New York Times* https://www.nytimes.com/2013/10/27/business/eugene-fama-king-of-predictable-markets.html
[16] Fama, Eugene F., and Kenneth R. French. "The Cross-Section of Expected Stock Returns." *The Journal of Finance*, vol. 47, no. 2, 1992, pp. 427–465. *JSTOR*, www.jstor.org/stable/2329112.

high B/M firms outperform because there is something inherently riskier about them, or is it because low B/M tend to have overly optimistic prospects that lead to disappointment?[17] Whether or not they proxy for risk, there is cause to believe that there are other additional indicators which may provide insight into returns.

Quantitative analysts spend their time working on developing asset pricing models and arbitrage strategies to find profitable strategies that exploit mispricing in financial markets. They work tirelessly in effort of finding the best combination of indicators that capture variation in returns; they conduct principle component analyses to minimize the number of real-time variables in models to avoid overfitting.

However, the methods by which quantitative analysts manufacture valuation models are often flawed. **Since the volume of data about companies that can be produced in real-time has increased substantially since the late 1980s, there has been a shift to focus on building asset pricing models which incorporate the most cross-sectional information at the expense of the length of the time series.** There is an enormous responsibility of quantitative analysts to balance the amount of real-time information incorporated into a model and the quality/variation in the data going into the models. Since many firms only use data collected since the late-1980s for their models, it is unlikely capturing very many rare event occurrences in the markets. These rare "black-swan" events fall far outside of the ordinarily predicted normal distribution that some asset pricing models use (for the sake of elegance). Rare events are difficult enough to predict with 150 or so years of data; reducing the number of years in the sample can only make the models less effective at making inferences about out of sample data.

While the empirical faults of pricing model building are unfortunate, the real harm caused by the models lies in their widespread applications. Despite the overwhelming evidence showing that CAPM is empirically invalid, investment banks tend to use CAPM with modifications to calculate the discount rate used in valuation models for individual companies. The betas that are inappropriately used in these models should really only be used to evaluate the undiversifiable "risk" of a diversified portfolio of stocks (not individual companies). Regressions that attempt to test covariance of individual stocks with the market often return standard errors that produce extremely large and useless confidence intervals for beta of an individual security. Unscrupulous investment bankers may use these large confidence intervals to their advantage, and deliberately reduce the discount rate in valuation models by using CAPM to calculate cost of equity in order to inflate valuations. Investment practitioners may even use the 3-factor model as a substitute for fundamental stock analysis and good judgement.

The truth is that "the more we learn about financial markets, the more it seems that we are unable to accurately determine the inherent risk of various financial securities.[18]"

b. Hyman Minsky's Financial Instability Hypothesis & Instability Inherent in Capitalist Economies

---

[17] LAKONISHOK, J. , SHLEIFER, A. and VISHNY, R. W. (1994), Contrarian Investment, Extrapolation, and Risk. The Journal of Finance, 49: 1541-1578. doi:10.1111/j.1540-6261.1994.tb04772.x

[18] Sherman, John. "The Efficient Market Hypothesis, the Financial Instability Hypothesis, and Speculative Bubbles". BA, 2014. https://dlib.bc.edu/islandora/object/bc-ir%3A102322

In May of 1992, Hyman Minsky's paper entitled "The Financial Instability Hypothesis" was published; the paper provides a framework to distinguish between stabilizing and destabilizing capitalist debt structures:

> "Three distinct income-debt relations for economic units, which are labeled as hedge, speculative, and Ponzi finance, can be identified. **Hedge financing units are those which can fulfill all of their contractual payment obligations by their cash flows**: the greater the weight of equity financing in the liability structure, the greater the likelihood that the unit is a hedge financing unit. **Speculative finance units are units that can meet their payment commitments on "income account" on their liabilities, even as they cannot repay the principle out of income cash flows.** Such units need to "roll over" their liabilities: (e.g. issue new debt to meet commitments on maturing debt). Governments with floating debts, corporations with floating issues of commercial paper, and banks are typically hedge units. **For Ponzi units, the cash flows from operations are not sufficient to fulfill either the repayment of principle or the interest due on outstanding debts by their cash flows from operations.** Such units can sell assets or borrow. Borrowing to pay interest or selling assets to pay interest (and even dividends) on common stock lowers the equity of a unit, even as it increases liabilities and the prior commitment of future incomes. A unit that Ponzi finances lowers the margin of safety that it offers the holders of its debts. **It can be shown that if hedge financing dominates, then the economy may well be an equilibrium seeking and containing system. In contrast, the greater the weight of speculative and Ponzi finance, the greater the likelihood that the economy is a deviation-amplifying system."[19]**

Minsky goes on further to explain how these various unstable financing regimes come about as the result of prolonged periods of stability:

> "The first theorem of the financial instability hypothesis is that the economy has financing regimes under which it is stable, and financing regimes in which it is unstable. The second theorem of the financial instability hypothesis is that over periods of prolonged prosperity, the economy transits from financial relations that make for a stable system to financial relations that make for an unstable system. **In particular, over a protracted period of good times, capitalist economies tend to move from a financial structure dominated by hedge finance units to a structure in which there is large weight to units engaged in speculative and Ponzi finance.** Furthermore, if an economy with a sizeable body of speculative financial units is in an inflationary state, and the authorities attempt to exorcise inflation by monetary constraint, then speculative units will become Ponzi units and the net worth of previously Ponzi units will quickly evaporate. Consequently, units with cash flow shortfalls will be forced to try to make position by selling out position. This is likely to lead to a collapse of asset values. The financial instability hypothesis is a model of a capitalist economy which does not rely upon exogenous shocks to generate business cycles of varying severity. **The hypothesis**

---

[19] Minsky, Hyman P., The Financial Instability Hypothesis (May 1992). The Jerome Levy Economics Institute Working Paper No. 74. Available at SSRN: https://ssrn.com/abstract=161024 or http://dx.doi.org/10.2139/ssrn.161024

> **holds that business cycles of history are compounded out of (i) the internal dynamics of capitalist economies, and (ii) the system of interventions and regulations that are designed to keep the economy operating within reasonable bounds."[20]**

Minsky's insights about the nature of speculative and Ponzi financing in the late stages of the business cycle align perfectly with those of Michael Burry, who is widely believed to have been the first portfolio manager to bet against the housing market in 2005 by harassing Wall Street banks into selling him credit default swaps (insurance contracts that pay off in the event that the underlying bond defaults):

> In the second quarter of 2005, credit-card delinquencies hit an all-time high—even though house prices had boomed. That is, even with this asset to borrow against, Americans were struggling more than ever to meet their obligations. **The Federal Reserve had raised interest rates, but mortgage rates were still effectively falling—because Wall Street was finding ever more clever ways to enable people to borrow money.** Burry now had more than a billion-dollar bet on the table and couldn't grow it much more unless he attracted a lot more money. So, he just laid it out for his investors: the U.S. mortgage-bond market was huge, bigger than the market for U.S. Treasury notes and bonds. The entire economy was premised on its stability, and its stability in turn depended on house prices continuing to rise. "It is ludicrous to believe that asset bubbles can only be recognized in hindsight," he wrote. "There are specific identifiers that are entirely recognizable during the bubble's inflation. **One hallmark of mania is the rapid rise in the incidence and complexity of fraud.… The FBI reports mortgage-related fraud is up fivefold since 2000." Bad behavior was no longer on the fringes of an otherwise sound economy; it was its central feature.** "The salient point about the modern vintage of housing-related fraud is its integral place within our nation's institutions," he added.[21]

Even in the face of overwhelming evidence that the market was unsustainable, the market kept on chugging until 2007.

> "As early as 2004, if you looked at the numbers, you could clearly see the decline in lending standards. In Burry's view, standards had not just fallen but hit bottom. **The bottom even had a name: the interest-only negative-amortizing adjustable-rate subprime mortgage.** You, the homebuyer, actually were given the option of paying nothing at all, and rolling whatever interest you owed the bank into a higher principal balance. **It wasn't hard to see what sort of person might like to have such a loan: one with no income.**"

Despite all of these very clear issues with the housing market, Burry's investors still doubted him because of the short-term changes in the portfolio's value due to the premiums paid to maintain the credit default swap positions:

---

[20] Minsky, Hyman P., The Financial Instability Hypothesis (May 1992). The Jerome Levy Economics Institute Working Paper No. 74. Available at SSRN: https://ssrn.com/abstract=161024 or http://dx.doi.org/10.2139/ssrn.161024

[21] Lewis, Michael M. The Big Short : [inside the Doomsday Machine]. New York :Simon & Schuster, 2010.

> "Now he had to explain that they had to subtract from that number these & subprime-mortgage-bond insurance premiums. **One of his New York investors called and said ominously, 'You know, a lot of people are talking about withdrawing funds from you.'** As their funds were contractually stuck inside Scion Capital for some time, the investors' only recourse was to send him disturbed-sounding e-mails asking him to justify his new strategy. **'People get hung up on the difference between +5% and -5% for a couple of years,' Burry replied to one investor who had protested the new strategy. 'When the real issue is: over 10 years who does 10% or better annually?** And I firmly believe that to achieve that advantage on an annual basis, I have to be able to look out past the next couple of years.… I have to be steadfast in the face of popular discontent if that's what the fundamentals tell me.' **In the five years since he had started, the S&P 500, against which he was measured, was down 6.84 percent. In the same period, he reminded his investors, Scion Capital was up 242 percent.**"[22]

Despite Burry's significant outperformance of the market based on clear value investing principles, his clients still didn't trust him to make investment decisions because of their unrelenting focus on short term price changes. Burry had the benefit of running a hedge fund that allowed him to lock up investor funds and forbid withdrawals in the face of this change in sentiment among his clients. Even with these provisions, Burry's clients whom he had once respected constantly harassed him and made his job miserable while waiting for the bet to pay off. Investors like Michael Burry are unable to make these kinds of bets within the structure of ordinary mutual funds (excluding separately managed accounts). Mutual fund managers are subject to liquidity requirements and are forbidden from preventing withdrawals of investor capital. This once again underscores how the "smart money" investors are often at the whims of the masses when it comes to asset allocation strategies.

This still doesn't explain why investors were willing to purchase mortgage-backed securities that supported loans made to Speculative Borrowers and Ponzi Borrowers. Creditors lend to Ponzi borrowers because they also believe that prices of the underlying asset will increase and/or the borrower will be able to refinance the loan. Minsky argued that protracted periods of stable good economic conditions will ultimately lead to a greater proportion of debt in the hands of speculators and Ponzi borrowers; we believe that this is because of people irrationally placing emphasis on recent experience when developing their expectations of the future.

c. The Aggregate Investor Allocation to Equities, Tobin's Q, TR CAPE, CAPE, and Long-Term Reversion in Returns

As John Sherman points out, "One of the most fascinating anomalies in securities markets is that of long-term reversals—that a series of high returns tends to be followed by low returns and vice versa."[23] As Eugene Fama explains:

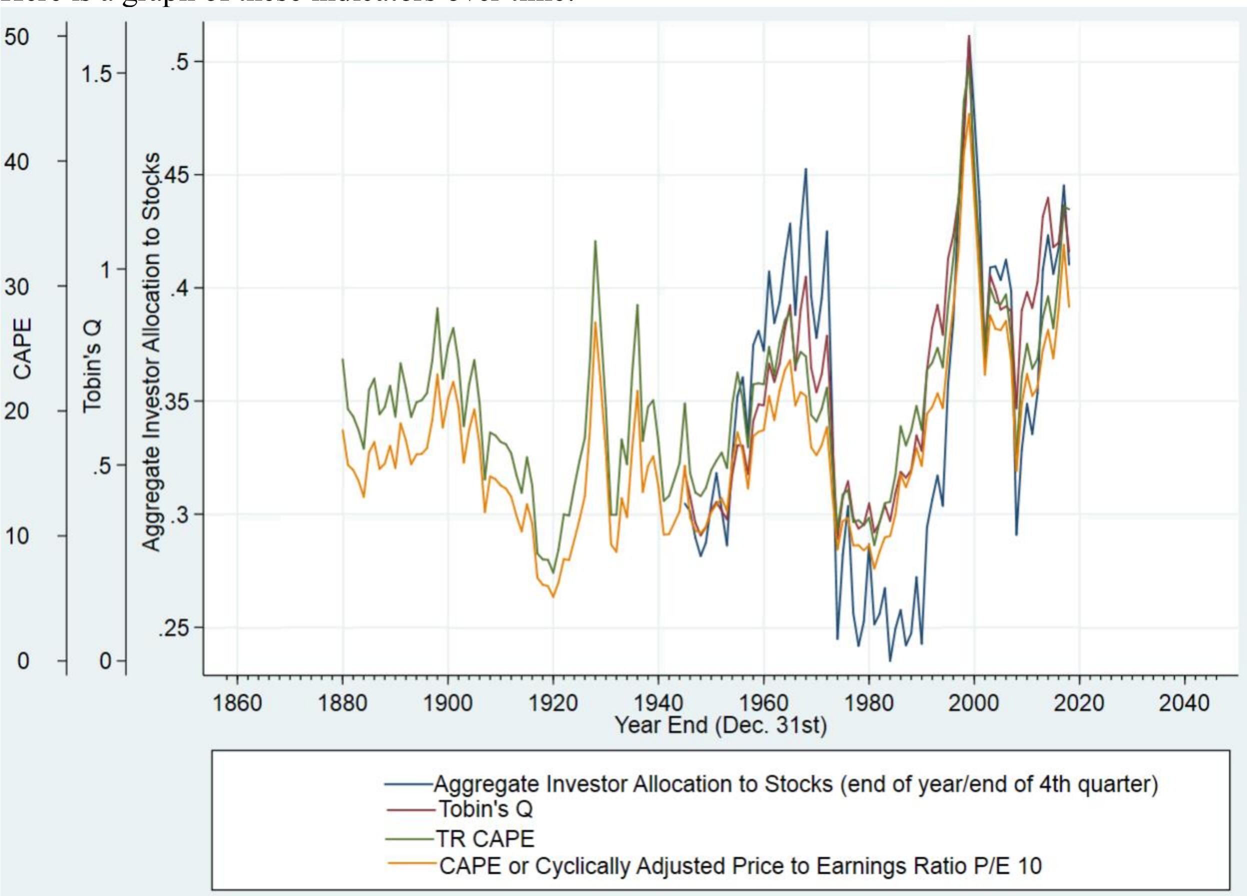
---

[22] Lewis, Michael M. The Big Short : [inside the Doomsday Machine]. New York :Simon & Schuster, 2010.
[23] Sherman, John. "The Efficient Market Hypothesis, the Financial Instability Hypothesis, and Speculative Bubbles". BA, 2014. https://dlib.bc.edu/islandora/object/bc-ir%3A102322

"Ratios involving stock prices have information about expected returns missed by market betas. On reflection, this is not surprising. A stock's price depends not only on the expected cash flows it will provide, but also on the expected returns that discount expected cash flows back to the present. **Thus, in principle the cross-section of prices has information about the cross-section of expected returns.** (A high expected return implies a high discount rate and a low price.) The cross-section of stock prices is, however, arbitrarily affected by differences in scale (or units). But with a judicious choice of scaling variable X, the ratio X/P can reveal differences in the cross-section of expected stock returns. **Such ratios are thus prime candidates to expose shortcomings of asset pricing models – in the case of the CAPM, shortcomings of the prediction that market betas suffice to explain expected returns** (Ball, 1978)[24][25]."

The Aggregate Investor Allocation to Equities, Tobin's Q, TR CAPE, and CAPE are some of these pricing indicators which fit Fama's criteria; coincidentally all 4 show statistically significant negative correlation with subsequent 10-year real returns.

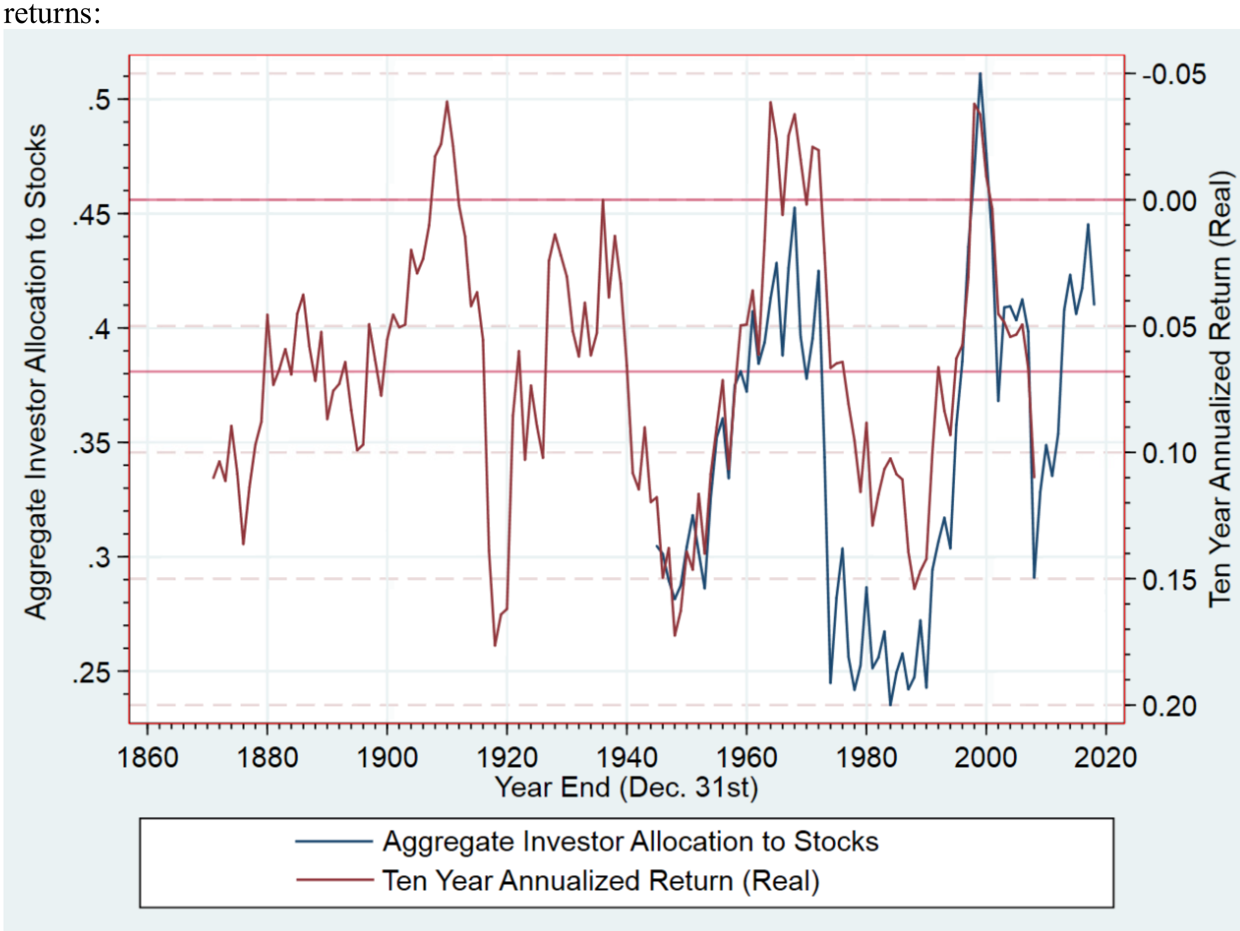
Here is a graph of these indicators over time:



[24] Ball, Ray. 1978. "Anomalies in Relationships Between Securities' Yields and Yield-Surrogates." Journal of Financial Economics. 6:2, pp. 103-126.
[25] Fama, Eugene F. and French, Kenneth R., The Capital Asset Pricing Model: Theory and Evidence (August 2003). CRSP Working Paper No. 550; Tuck Business School Working Paper No. 03-26. Available at SSRN: https://ssrn.com/abstract=440920 or http://dx.doi.org/10.2139/ssrn.440920

The Aggregate Investor Allocation to Equities is of particular interest to us. Financial assets are unique in that someone must be in possession of them at any given time. When investors purchase and sell financial assets, they implicitly make decisions about the relative proportions of various assets that they hold. An investor may decide to allocate 40% of their savings to stocks, and another 60% to corporate bonds. Another may allocate 10% to treasury bills, 70% in mortgage-backed securities, and 20% in cash. Since these assets are always held by someone, we can aggregate the value of all of these securities across asset classes to determine the aggregate allocations to various classes of financial assets. **This is of particular interest because of its relationship to the Efficient Market Hypothesis; rational expectations dictate that an individual investor's allocation to a particular asset class will only increase if this investors' risk-adjusted expected return has increased.** This relationship must also hold in aggregate. The historical average aggregate investor allocation to equities between 1945 and 2018 has been roughly 34.7%. **If historically high levels of the Aggregate Investor Allocation are associated with lower returns and higher levels of risk, this casts doubt upon the rational investor assumption of the EMH.**

Below is a graph of the aggregate investor allocation to stocks and subsequent 10-year real returns:



Notice in the graph the two red reference lines. One corresponds to a real 10-year annualized return of 0%, and the other corresponds to 6.8% real annualized 10-year returns. 6.8% is roughly the geometric average real return from the end of 1871 to the end of 2018 (see appendix).

Tobin's Q was developed by Nobel Laureate James Tobin in 1968. It is the ratio between the market value and replacement value of an asset:

$$Tobin's\ Q = \frac{Market\ Valuation}{Replacement\ Cost}$$

> "One, the numerator, is the market valuation: the going price in the market for exchanging existing assets. The other, the denominator, is the replacement or reproduction cost: the price in the market for the newly produced commodities. We believe that this ratio has considerable macroeconomic significance and usefulness, as the nexus between financial markets and markets for goods and services." [26] [27]

When the market price for a corporation is high relative to its replacement cost of net assets, this may attract new entrants; new entrants could effectively replicate the asset and then sell it in the market for a profit. The ensuing dynamics of competition would lead the market value to come down relative to its replacement cost. While a mean value of 1 for the total market would be expected, the historical average is actually around 0.6. This may be due to measurement error in the replacement cost of nonfinancial firms' assets, which is estimated by the Federal Reserve Bank. It may be that the replacement cost of firms is overstated because GAAP depreciation methods understate depreciation expenses. We will not focus that much on the accounting nuances, but the end result is that replacement cost may biased upwards due to systematic measurement error, which leads Tobin's Q to be biased downwards.

The Cyclically Adjusted Price to Earnings ratio (CAPE) and Total Return CAPE (TR CAPE) are metrics created by Robert Shiller.

$$CAPE = \frac{Market\ Price\ of\ Equities}{Average\ of\ the\ Past\ 10\ years\ of\ Real\ Earnings}$$

Total Return CAPE is nearly identical to CAPE; however, it makes corrections for changes in corporate dividend/buyback policy. TR CAPE and CAPE are interesting because historically high levels of CAPE indicate that the neither the overall level of earnings, nor earnings growth over the period have kept up with price increases over the same period. TR CAPE and CAPE have the most observations in our data since earnings and prices have been accurately recorded for indices since 1871. In contrast, the Aggregate Investor Allocation to Equities and Tobin's Q have only been consistently estimated since 1945.

---

[26]"Asset Markets and the Cost of Capital." James Tobin and W.C. Brainard, 1977, Economic Progress, Private Values and Public Policy

[27]Sherman, John. "The Efficient Market Hypothesis, the Financial Instability Hypothesis, and Speculative Bubbles". BA, 2014. https://dlib.bc.edu/islandora/object/bc-ir%3A102322

d. The Random Walk Assumption

As mentioned earlier, both the Capital Asset Pricing Model[28] and the Black-Scholes Option Pricing Formula assume that stock prices exhibit Brownian motion (stochastic price movements with drift) which is broadly consistent with the idea that price shocks are the result of randomly generated new information being incorporated into prices.

This "random walk" is a fundamental assumption of the EMH, and has been broadly incorporated into financial theory (as exemplified by the CAPM and Black-Scholes models mentioned above). Here is an algebraic representation:

$$V_t = V_{t-1} + \Phi_t + \varepsilon_t$$

Where V_t is the natural log of the value of an asset, Φ_t is the drift trend, and ε_t is a random shock in value which has the following properties:

$$E(\varepsilon_t) = 0 \qquad Var(\varepsilon_t) = \sigma^2 \qquad Cov(\varepsilon_t, \varepsilon_{t-1}) = 0$$

The expected value of the shock ε_t is zero. Price shocks are assumed to have a log-normal distribution. Covariance between price shocks is zero, so price shocks are all independent of one another; historical price changes do not influence prices today.

Our empirical regressions clearly indicate that past price shocks do have some relationship to future shocks, so in reality $Cov(\varepsilon_t, \varepsilon_{t-1}) \neq 0$.

Now in reality, prices do not follow a log-normal distribution. The log distribution of stock returns has fat-tails (kurtosis):

> "From 1916 to 2003, the daily index movements of the Dow Jones Industrial Average do not spread out on graph paper like a simple bell curve. The far edges flare too high: too many big changes. Theory suggests that over that time, there should be fifty-eight days when the Dow moved more than 3.4 percent; in fact, there were 1,001. Theory predicts six days of index swings beyond 4.5 percent; in fact, there were 366. And index swings of more than 7 percent should come once every 300,000 years; in fact, the twentieth century saw forty-eight such days. Truly, a calamitous era that insists on flaunting all prediction. Or, perhaps, our assumptions are wrong."[29]

III. Restructured Model with Endogenously Determined Expected Returns

Now equity returns for individual companies (or the market) can conventionally be thought of this way:

---

[28]Although every asset pricing model is a capital asset pricing model, the finance profession reserves the acronym CAPM for the specific model of Sharpe (1964), Lintner (1965), and Black (1972) discussed here. Thus, throughout the paper we refer to the Sharpe – Lintner – Black model as the CAPM. (Fama French 2003)

[29] Mandelbrot, Benoit; Hudson, Richard L. (2007-03-22). The Misbehavior of Markets: A Fractal View of Financial Turbulence (pp. 11-13).

$$Total\ Return_t = \text{Return from } \Delta \frac{Price_t}{Earnings_t} (holding\ earnings\ constant) +$$
$$Return\ from\ \Delta Earnings_t\ (holding\ the \frac{P}{E} multiple\ constant) + Dividend\ Return$$

Returns over a given period are a function of price changes and dividend returns. Returns from the price change can be further disaggregated into return from a change in the price to earnings ratio (holding earnings constant) and return from a change in earnings (holding the price to earnings multiple constant).

As mentioned earlier, price shocks can have significant effects if we relax the random walk assumption. If people extrapolate out recent experience of returns and risk when forming expectations of the future; a series of random shocks in one direction could result in a slight divergence in value, however, if "investors start to believe that there is something new about the market"[30], future expectations of returns and risk will be biased. This bias in aggregate expected returns and aggregate expected risk leads to price ratios having explanatory power of subsequent returns (because they implicitly contain information about expected returns).

The aggregate investor allocation is shown to explain the most variation in subsequent 10-year annualized real returns in our empirical section; now there is another very intuitive way to think about returns by substituting in the aggregate investor allocation for the price to earnings ratio. Instead of prices being linked to the underlying earnings of an asset, we link equity prices to the outstanding value of cash and bonds outstanding in the market.

Now first, there are particular attributes of financial assets that allow this to be the case. All cash, bonds, and equities must be held by someone at any given point in time. **When investors make explicit decisions about the proportions of each asset they will hold in their individual portfolio, the aggregate market allocations of these assets must implicitly shift through price changes to allow individuals to meet their explicit allocation targets.** When cash and bonds increase in tandem with growth of the real economy, prices of equities must rise commensurately to maintain the same portfolio allocation.

We create a model of equilibrium demand and supply for the aggregate investor allocation:

> In the equities market, demand is a function of expected returns and expected risk. Supply is a function of the # of shares available for purchase in the equities market and new share issuance. Demand shifts are a function of changes in expected returns and changes in expected risk. Supply shifts are a function of new issuance of equity. The vertical axis of the graph is the aggregate allocation to equities. Now demand is downward sloping to the right with a vertical intercept of 100% because the maximum percentage of a portfolio that an individual investor can own is 100%. The demand schedule consists of the target allocations of individual investors to equities. Supply is relatively inelastic and upward sloping since entrepreneurs are slightly responsive to high levels of the aggregate investor allocation; new companies can be sold at high prices

[30] Sherman, John. "The Efficient Market Hypothesis, the Financial Instability Hypothesis, and Speculative Bubbles". BA, 2014. https://dlib.bc.edu/islandora/object/bc-ir%3A102322

relative to the replacement cost of net assets (Tobin's Q is somewhat correlated with the aggregate investor allocation, but the allocation carries more explanatory power). We assume supply is inelastic because ideas that warrant significant new equity issuance are not really responsive to the allocation, as the equity issuance would occur anyway to provide capital to fund the new business idea. Since average net issuance of equity in any given year tends to be very slightly negative because of stock buybacks (at least since the beginning of the 1980s), and expectations are more volatile, aggregate demand shifts tend to have a greater impact on the equilibrium market clearing aggregate investor allocation.

Coincidentally, we can represent returns over any given period with a similar framework to the price to earnings framework described above:

$$\text{Total Return}_t = \text{Return from the change in } \frac{(\text{market value of equities})}{\text{market value of equities} + \text{all other real financial liabilities}^{31}} + \text{Return from change in the supply of bonds and cash (holding the aggregate investor allocation to equities constant)} + \\ + \text{Dividend Return}$$

By thinking of price returns in terms of shifts in the aggregate investor allocation to equities and growth in the supply of bonds and cash (which tend to grow with the economy as a whole), we get a much more realistic picture of how demand and supply functions in financial markets. A significant proportion of subsequent 10-year annualized real returns can be explained in relation to changes in the aggregate investor allocation, much in the way that we think about price to earnings multiples. People intuitively react to high P/E multiples negatively because the earnings yield is implicitly low, and vice versa with low P/E multiples. We think much the same way with regards to the percentage of financial capital invested in the equities market; preferences for individual allocations may mean revert over time, and variation in preferences can be accounted for by a number of factors which we discuss in the second part of the empirical section.

[31] See page 19 in the empirical results data section for a more in-depth explanation of how the aggregate investor allocation is calculated

IV. Empirical Results (2 sections)

a. Subsequent Ten Year Cumulative Annualized Return Regressions

1. Data

Stock return, Shiller CAPE ratios, and the long interest rate come from Robert Shiller's stock market data. This dataset can be found on Yale's website at:
http://www.econ.yale.edu/~shiller/data.htm

Shiller explains how he constructed the dataset we pulled from his website[32]:

> **"Stock market data** used in my book, *Irrational Exuberance* [Princeton University Press 2000, Broadway Books 2001, 2nd ed., 2005] are available for download, **U.S. Stock Markets 1871-Present and CAPE Ratio**. This data set consists of monthly stock price, dividends, and earnings data and the consumer price index (to allow conversion to real values), all starting January 1871. The price, dividend, and earnings series are from the same sources as described in Chapter 26 of my earlier book (*Market Volatility* [Cambridge, MA: MIT Press, 1989]), although now I use monthly data, rather than annual data. Monthly dividend and earnings data are computed from the S&P four-quarter totals for the quarter since 1926, with linear interpolation to monthly figures. Dividend and earnings data before 1926 are from Cowles and associates (*Common Stock Indexes*, 2nd ed. [Bloomington, Ind.: Principia Press, 1939]), interpolated from annual data. Stock price data are monthly averages of daily closing prices through January 2000, the last month available as this book goes to press. The CPI-U (Consumer Price Index-All Urban Consumers) published by the U.S. Bureau of Labor Statistics begins in 1913; for years before 1913 1 spliced to the CPI Warren and Pearson's price index, by multiplying it by the ratio of the indexes in January 1913. December 1999 and January 2000 values for the CPI-U are extrapolated. See George F. Warren and Frank A. Pearson, *Gold and Prices* (New York: John Wiley and Sons, 1935). Data are from their Table 1, pp. 11–14.
>
> As of September 2018, I now also include an **alternative version of CAPE** that is somewhat different. As documented in Bunn & Shiller (2014) and Jivraj and Shiller (2017), changes in corporate payout policy (i.e. share repurchases rather than dividends have now become a dominant approach in the United States for cash distribution to shareholders) may affect the level of the CAPE ratio through changing the growth rate of earnings per share. This subsequently may affect the average of the real earnings per share used in the CAPE ratio. A total return CAPE corrects for this bias through reinvesting dividends into the price index and appropriately scaling the earnings per share. "

Shiller's dataset includes monthly data; however, we use annual end of year data to construct our key dependent variable: Ten Year Annualized Returns (Real & Nominal). We use cumulative annualized returns in order to explore the lagged or "persistence effect" of past price shocks. Ten years is a crude estimation that is roughly close to the average length of an average U.S. business

[32] Robert Shiller's data webpage: http://www.econ.yale.edu/~shiller/data.htm

cycle (5-7 years). Periods significantly longer than 10 years may not be as memorable to investors so there may be less price to price feedback. Shiller's stock data on the S&P 500 goes back until 1871, giving us 137 observations between 1871 and 2008.[33]

Data for the aggregate investor allocation to equities can be calculated using Federal Reserve Data. The formula is calculated by dividing (Nonfinancial corporate business; corporate equities; liability, level + Financial business; corporate equities; liability, level) by (Nonfinancial corporate business; corporate equities; liability, level + Financial business; corporate equities; liability, level + Nonfinancial Corporate Business; Credit Market Instruments; liability + Households and Nonprofit Organizations; Credit Market Instruments; liability, level + Federal Government; Credit Market Instruments; Liability, Level + State and Local Governments, Excluding Employee Retirement Funds; Credit Market Instruments; Liability, Level + Rest of the World; Credit Market Instruments; Liability, Level). The FRED codes are as follows: ((NCBEILQ027S+FBCELLQ027S)/1000)/(((NCBEILQ027S+FBCELLQ027S)/1000)+BCNSDODNS+CMDEBT+FGSDODNS+SLGSDODNS+DODFFSWCMI). The 1000s are there to convert all of the numbers to the same units. A chart of the metric can be found at this link: https://fred.stlouisfed.org/graph/?g=qis. Data on the aggregate investor allocation to equities begins in 1945, leaving us with 73 observations. *Note that this is significantly fewer than with Shiller's TR CAPE or Shiller's CAPE.*

Data for Tobin's Q was calculated using the Federal Reserve Statistical Release Z.1 Financial Accounts of the United States and can be found at: http://www.federalreserve.gov/releases/z1/current/z1.pdf. Tobin's Q ratio can be calculated from the most recent Federal Reserve Flow of Funds release. The ratio is calculated by dividing line 36 of table B.102 by line 33. We calculate Tobin's Q through 2018. The Federal Reserve has data on Tobin's Q beginning in 1945, leaving us with 73 observations. Note that when we complete regressions on 10-year returns, we will drop 10 observations from each one we run since we don't have subsequent ten-year cumulative returns for 2009-2018, despite having other data for these years.

I paraphrase John Sherman's senior thesis, one of professor Petersen's former students, to describe the nature of Tobin's Q:

> Interestingly, Tobin's Q is calculated using data from nonfinancial firms (i.e. it excludes data from banks, insurance companies, etc.). High values of Tobin's Q are presumably a consequence of high growth and/or high profits, which pushes asset prices above their average replacement costs. If markets are efficient, high values of Q should be self-correcting or self-reversing, so we shouldn't get very high values of Q for sustained periods because profits attract new entrants. New firms enter the market and existing firms expand, and as investment in plant and equipment grows, a corresponding increase in the aggregate capital-labor ratio will result. Basic microeconomic theory suggests diminishing returns to capital set in and the profit rate will fall. Therefore, following high values of Q, we should expect lower returns in the following years.[34]

[33] Sherman, John, and Harold Petersen. "The Efficient Market Hypothesis, the Financial Instability Hypothesis, and Speculative Bubbles," pg. 29, 2014. https://dlib.bc.edu/islandora/object/bc-ir%3A102322
[34] Ibid

The theoretical basis for the aggregate investor allocation to equities is also very interesting. Since the allocation should be primarily determined by investors' expected returns and risk, any variation must be accompanied by either change in the market's expectation of return or risk (within the context of all assets available for purchase) according to the Efficient Market Hypothesis. As we will present in our regression results, subsequent return data directly contradicts the idea that higher allocations are met with higher returns and/or less risk, and vice versa (as would be suggested by the EMH).

2. **Dependent Variables for Subsequent Return Regressions**

We constructed cumulative returns in order to test if investors' expectations of risk and return are consistent with the Efficient Market Hypothesis. Specifically, we use:

   a. Ten Year Annualized Returns (Nominal)
   b. Ten Year Annualized Returns (Real – adjusted by CPI in Shiller's data)
   c. Average Annual Losses Over Next 10 Years (only includes years with losses)

3. **Independent Variables for Subsequent Return Regressions**

   a. The Aggregate Investor Allocation to Equities
   b. Tobin's Q
   c. Shiller's TR CAPE (Cyclically Adjusted Price to Earnings ratio adjusted for dividend and buyback payout policy)
   d. Shiller's CAPE
   e. Long Interest Rate (GS10)
   f. Greenspan Put

Our first 4 independent variables all contain aggregate market prices for equities (whether financial, nonfinancial, or all) as well as some other data that informs investors about the asset they are purchasing and/or data about the broader market and past market behavior. The Long Interest Rate comes from Shiller's data, and reflects historical 10-year Treasury yields in the United States. We construct another independent variable: Greenspan Put is a binary variable where every year including and following 1987 is equal to 1, otherwise 0. We construct this to denote Alan Greenspan's suggestion in 1987 that Federal Reserve Monetary Policy will be eased to prop up equity valuations in the event of crises. Subsequent Federal Reserve Chairmen have been believed to follow similar strategies to attempt to prop up equity valuations, however it should be noted that regressions which included the Greenspan Put may just lead to spurious correlation. Since the mid-1990s equity valuations have been historically been on the higher side according to the other independent variables presented; as a result, the binary variable essentially creates a piecewise best fit function that jumps in 1987. Since the greatest stock market collapses have tended to coincide with recessions and reasonable Federal Reserve Policy in recessions would generally include monetary easing, the period following Greenspan's announcement may simply coincide with more effective monetary policy implementations (on average) which have the added effect of leading investors to shift asset allocations towards risky assets; this is exactly what the Federal Reserve Bank intends in recessions as it stimulates greater economic productivity and reduces labor market slack. Despite this, we think that it would be interesting to run regressions with this indicator to see if it adds explanatory power to our other models (after first comparing them without the Greenspan Put).

4. Regression Specifications & Results

We present evidence that is inconsistent with the Efficient Market Hypothesis. First, we assess the significance of these market indicators listed above and their relationship with subsequent 10-year returns (nominal and real). Secondly, we test to see if our indicators may predict years of minimal losses over the following 10-year period. Thirdly, we assess possible determinants of the primary indicator of greatest significance: the aggregate investor allocation to equities.

It is important to note that due to our current specifications, our observations are not independent because ten-year equity returns consist of an overlapping moving average. To correct for autocorrelation, which biases standard errors downwards (and t-statistics upwards), we run a Newey-West (1987)[35] estimator; this is commonly used to correct for specifiable bias between observations in moving average variables. We specify a maximum of 9 lags since 10-year returns reported annually will have at most 9 years of overlapping annual returns. Our Newey-West corrections are stated alongside all of our regressions which require autocorrelation corrections below (those with moving average time series data).

[35] Newey, Whitney K., and Kenneth D. West. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, vol. 55, no. 3, 1987, pp. 703–708. *JSTOR*, www.jstor.org/stable/1913610.

Our specification for subsequent nominal returns regressed on the aggregate investor allocation is as follows:

$$\boldsymbol{TenYearAnnualizedReturnNom} = \boldsymbol{\alpha} + \boldsymbol{\beta} * \boldsymbol{PercentInStocks} + \boldsymbol{\varepsilon}$$

```
. *Regress Ten Year Annualized Nominal Returns on the Aggregate Investor Allocation to Stocks*
. reg TenYearAnnualizedReturnNom PercentInStocks

      Source |       SS           df       MS      Number of obs   =        64
-------------+----------------------------------   F(1, 62)        =    406.35
       Model |   .15270944         1   .15270944   Prob > F        =    0.0000
    Residual |  .023299804        62  .000375803   R-squared       =    0.8676
-------------+----------------------------------   Adj R-squared   =    0.8655
       Total |  .176009245        63  .002793798   Root MSE        =    .01939

------------------------------------------------------------------------------
TenYearAnnual~m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+--------------------------------------------------------------
PercentInStocks |  -.6935454   .0344051   -20.16   0.000    -.7623201   -.6247707
          _cons |   .3442421   .0119586    28.79   0.000     .3203373    .3681469
------------------------------------------------------------------------------

. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnNom PercentInStocks, lag(9)

Regression with Newey-West standard errors      Number of obs     =        64
maximum lag: 9                                  F(  1,        62) =    196.04
                                                Prob > F          =    0.0000

------------------------------------------------------------------------------
                |             Newey-West
TenYearAnnual~m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+--------------------------------------------------------------
PercentInStocks |  -.6935454   .0495337   -14.00   0.000    -.7925619   -.5945289
          _cons |   .3442421   .0200147    17.20   0.000     .3042332     .384251
------------------------------------------------------------------------------
```

The aggregate investor allocation captures 86.6% of variation in ten year annualized nominal stock returns, and is statistically significant at an α=.01 (t-stat=-13.86). Note that based on the model, each additional 1% of financial assets in equities reduces subsequent ten year annualized nominal returns by roughly .7% per year. Let's see how the indicator performs on real returns.

Our specification for subsequent real returns regressed on the aggregate investor allocation is as follows:

$$\boldsymbol{TenYearAnnualizedReturnReal} = \boldsymbol{\alpha} + \boldsymbol{\beta} * \boldsymbol{PercentInStocks} + \boldsymbol{\varepsilon}$$

```
. *Regress Ten Year Annualized Real Returns on the Aggregate Investor Allocation to Stocks*
. reg TenYearAnnualizedReturnReal PercentInStocks

      Source |       SS           df       MS      Number of obs   =        64
-------------+----------------------------------   F(1, 62)        =    140.80
       Model |  .146564614         1  .146564614   Prob > F        =    0.0000
    Residual |   .06453788        62  .001040934   R-squared       =    0.6943
-------------+----------------------------------   Adj R-squared   =    0.6894
       Total |  .211102494        63  .003350833   Root MSE        =    .03226

------------------------------------------------------------------------------
TenYearAnnual~l |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
PercentInStocks |  -.6794485   .0572603   -11.87   0.000    -.7939101   -.5649868
          _cons |   .3011658   .0199026    15.13   0.000     .2613811    .3409505
------------------------------------------------------------------------------

. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnReal PercentInStocks, lag(9)

Regression with Newey-West standard errors      Number of obs     =        64
maximum lag: 9                                  F(  1,        62) =     46.37
                                                Prob > F          =    0.0000

------------------------------------------------------------------------------
                |             Newey-West
TenYearAnnual~l |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
PercentInStocks |  -.6794485   .0997752    -6.81   0.000    -.8788962   -.4800007
          _cons |   .3011658   .0373349     8.07   0.000     .2265345    .3757972
------------------------------------------------------------------------------
```

While the variation in subsequent real returns is lower than in the nominal return regression, it still captures a significant amount of variation in real returns – notice the coefficient is very similar to the coefficient in the nominal return regression. Adjusted R^2 of 68.9% is a fairly good fit, and the corrected t-statistic of -6.81 is still statistically significant on an α=.01 basis.

Since the objective for an investor is to maximize real returns (per unit of risk), we regress on both nominal and real subsequent 10-year returns for all of our indicators. Notice that each additional percent of assets allocated to stocks corresponds with a decline of -0.68% in ten year annualized real returns. We find statistically significant negative correlations for all of our indicators, albeit with less correlation below.

Our specification for subsequent nominal returns regressed on Tobin's Q is as follows:

$$\boldsymbol{TenYearAnnualizedReturnNom} = \boldsymbol{\alpha} + \boldsymbol{\beta} * \boldsymbol{TobinsQ} + \boldsymbol{\varepsilon}$$

```
. *Regress Ten Year Annualized Nominal Returns on Tobin's Q*
. reg TenYearAnnualizedReturnNom TobinsQ

      Source |       SS           df       MS      Number of obs   =        64
-------------+----------------------------------   F(1, 62)        =    182.22
       Model |  .131326294         1  .131326294   Prob > F        =    0.0000
    Residual |  .044682951        62  .000720693   R-squared       =    0.7461
-------------+----------------------------------   Adj R-squared   =    0.7420
       Total |  .176009245        63  .002793798   Root MSE        =    .02685

------------------------------------------------------------------------------
TenYearAnn~m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     TobinsQ |  -.1600386   .0118556   -13.50   0.000    -.1837377   -.1363396
       _cons |    .214524   .0085629    25.05   0.000     .1974071     .231641
------------------------------------------------------------------------------

. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnNom TobinsQ, lag(9)

Regression with Newey-West standard errors      Number of obs     =         64
maximum lag: 9                                  F(  1,        62) =      88.36
                                                Prob > F          =     0.0000

------------------------------------------------------------------------------
             |             Newey-West
TenYearAnn~m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     TobinsQ |  -.1600386   .0170255    -9.40   0.000    -.1940721   -.1260051
       _cons |    .214524   .0103756    20.68   0.000     .1937836    .2352645
------------------------------------------------------------------------------
```

Our specification for subsequent real returns regressed on Tobin's Q is as follows:

$$\boldsymbol{TenYearAnnualizedReturnReal = \alpha + \beta * TobinsQ + \varepsilon}$$

```
. *Regress Ten Year Annualized Real Returns on Tobin's Q*
. reg TenYearAnnualizedReturnReal TobinsQ

      Source |       SS           df       MS      Number of obs   =        64
-------------+----------------------------------   F(1, 62)        =     73.20
       Model |  .114295702         1  .114295702   Prob > F        =    0.0000
    Residual |  .096806792        62    .0015614   R-squared       =    0.5414
-------------+----------------------------------   Adj R-squared   =    0.5340
       Total |  .211102494        63  .003350833   Root MSE        =    .03951

------------------------------------------------------------------------------
TenYearAnn~l |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     TobinsQ |  -.1493014   .0174504    -8.56   0.000    -.1841843   -.1144185
       _cons |   .1691112   .0126038    13.42   0.000     .1439165    .1943058
------------------------------------------------------------------------------

. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnReal TobinsQ, lag(9)

Regression with Newey-West standard errors      Number of obs     =         64
maximum lag: 9                                  F(  1,        62) =      26.12
                                                Prob > F          =     0.0000

------------------------------------------------------------------------------
             |             Newey-West
TenYearAnn~l |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     TobinsQ |  -.1493014   .0292114    -5.11   0.000    -.2076941   -.0909087
       _cons |   .1691112   .0201902     8.38   0.000     .1287515    .2094708
------------------------------------------------------------------------------
```

Notice that once again that our indicator, Tobin's Q, captures less variation in real returns than in nominal returns. An adjusted R^2 of 53.4% is still fairly respectable though. Note while Tobin's Q is statistically significant, it's less significant than the aggregate investor allocation to equities with a t-statistic of -5.11 compared to -6.81 (for the aggregate allocation).

Our specification for subsequent nominal returns regressed on Total Return CAPE is as follows:

$$\boldsymbol{TenYearAnnualizedReturnNom = \alpha + \beta * TRCAPE + \varepsilon}$$

```
. *Regress Ten Year Annualized Nominal Returns on TR CAPE*
. reg TenYearAnnualizedReturnNom TRCAPE

      Source |       SS           df       MS      Number of obs   =       129
-------------+----------------------------------   F(1, 127)       =     69.93
       Model |  .118199008         1  .118199008   Prob > F        =    0.0000
    Residual |  .214676342       127  .001690365   R-squared       =    0.3551
-------------+----------------------------------   Adj R-squared   =    0.3500
       Total |   .33287535       128  .002600589   Root MSE        =    .04111

------------------------------------------------------------------------------
TenYearAnn~m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      TRCAPE |  -.0043597   .0005214    -8.36   0.000    -.0053913    -.003328
       _cons |   .1783462   .0110044    16.21   0.000     .1565705    .2001219
------------------------------------------------------------------------------

. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnNom TRCAPE, lag(10)

Regression with Newey-West standard errors      Number of obs     =        129
maximum lag: 10                                 F(  1,        127) =      30.93
                                                Prob > F          =     0.0000

------------------------------------------------------------------------------
             |             Newey-West
TenYearAnn~m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      TRCAPE |  -.0043597   .0007839    -5.56   0.000    -.0059109   -.0028085
       _cons |   .1783462   .0211014     8.45   0.000     .1365905     .220102
------------------------------------------------------------------------------
```

Our specification for subsequent real returns regressed on Total Return CAPE is as follows:

$$\boldsymbol{TenYearAnnualizedReturnReal = \alpha + \beta * TRCAPE + \varepsilon}$$

```
. *Regress Ten Year Annualized Real Returns on TR CAPE*
. reg TenYearAnnualizedReturnReal TRCAPE

      Source |       SS       df       MS              Number of obs =     129
-------------+------------------------------           F(1, 127)     =   55.16
       Model |  .100909843     1  .100909843           Prob > F      =  0.0000
    Residual |   .23233273   127  .001829392           R-squared     =  0.3028
-------------+------------------------------           Adj R-squared =  0.2973
       Total |  .333242573   128  .002603458           Root MSE      =  .04277

------------------------------------------------------------------------------
TenYearAnn~l |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      TRCAPE |  -.0040282   .0005424    -7.43   0.000    -.0051015    -.002955
       _cons |   .1442785    .011448    12.60   0.000      .121625    .1669321
------------------------------------------------------------------------------

. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnReal TRCAPE, lag(9)

Regression with Newey-West standard errors          Number of obs     =     129
maximum lag: 9                                      F(  1,       127) =   31.73
                                                    Prob > F          =  0.0000

------------------------------------------------------------------------------
             |             Newey-West
TenYearAnn~l |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      TRCAPE |  -.0040282   .0007152    -5.63   0.000    -.0054434    -.002613
       _cons |   .1442785   .0188775     7.64   0.000     .1069234    .1816337
------------------------------------------------------------------------------
```

Notice that once again that Total Return CAPE captures less variation in real returns than in nominal returns. While the adjusted R^2 starts to fall off at 29.7% of variation. TR CAPE is still testing as significant – while the t statistic is too close to that of Tobin's Q to really be able to make a very clear indication of strength, the R^2 of the Tobin's Q regression was fitted much better. It is important to note that there is a significant disparity in observations between these two regressions, which could ultimately be leading to the disparity in fit. For this reason, it is unclear which indicator is stronger because there are no data to test the out of sample observations that are included in the TR CAPE regression on Tobin's Q. The Tobin's Q t-statistic could go either way in the future, especially considering the limitation of the sample. The Tobin's Q and aggregate investor allocation are clearly more comparable because they sample the same years of observations (1945-2008 and subsequent return periods).

It is possible that the 10-year treasury yield (or historical equivalent) could add explanatory power to our regressions, on grounds that investors would accept lower expected returns on stocks if bond yields were lower, so we add the long interest rate to our regressions which include TR CAPE and CAPE.

Our specification for subsequent nominal returns regressed on both Total Return CAPE and the 10-year interest rate is as follows:

$$\boldsymbol{TenYearAnnualizedReturnNom = \alpha + \beta_1 * TRCAPE + \beta_2 * LongInterestRate + \varepsilon}$$

```
. *Regress Ten Year Annualized Nominal Returns on TR CAPE & 10 Year Interest Rate*
. reg TenYearAnnualizedReturnNom TRCAPE LongInterestRate
```

Source	SS	df	MS			
				Number of obs	=	129
				F(2, 126)	=	41.22
Model	.131645923	2	.065822962	Prob > F	=	0.0000
Residual	.201229427	126	.001597059	R-squared	=	0.3955
				Adj R-squared	=	0.3859
Total	.33287535	128	.002600589	Root MSE	=	.03996

TenYearAnnuali~m	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
TRCAPE	-.0041176	.0005136	-8.02	0.000	-.005134	-.0031012
LongInterestRate	.444031	.153025	2.90	0.004	.141199	.746863
_cons	.1528446	.0138438	11.04	0.000	.1254482	.1802411

```
. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnNom TRCAPE LongInterestRate, lag(9)

Regression with Newey-West standard errors      Number of obs     =        129
maximum lag: 9                                  F(  2,       126) =      64.40
                                                Prob > F          =     0.0000
```

TenYearAnnuali~m	Coef.	Newey-West Std. Err.	t	P>|t|	[95% Conf.	Interval]
TRCAPE	-.0041176	.0007953	-5.18	0.000	-.0056915	-.0025437
LongInterestRate	.444031	.3405742	1.30	0.195	-.2299553	1.118017
_cons	.1528446	.0356055	4.29	0.000	.0823824	.2233069

Our specification for subsequent real returns regressed on both Total Return CAPE and the 10-year interest rate is as follows:

$$\boldsymbol{TenYearAnnualizedReturnReal} = \boldsymbol{\alpha} + \boldsymbol{\beta_1} * \boldsymbol{TRCAPE} + \boldsymbol{\beta_2} * \boldsymbol{LongInterestRate} + \boldsymbol{\varepsilon}$$

```
. *Regress Ten Year Annualized Real Returns on TR CAPE & 10 Year Interest Rate*
. reg TenYearAnnualizedReturnReal TRCAPE LongInterestRate
```

Source	SS	df	MS			
				Number of obs	=	129
				F(2, 126)	=	27.44
Model	.101094954	2	.050547477	Prob > F	=	0.0000
Residual	.232147619	126	.001842441	R-squared	=	0.3034
				Adj R-squared	=	0.2923
Total	.333242573	128	.002603458	Root MSE	=	.04292

TenYearAnnuali~l	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
TRCAPE	-.0039998	.0005516	-7.25	0.000	-.0050915	-.0029082
LongInterestRate	.0520977	.164361	0.32	0.752	-.2731679	.3773633
_cons	.1412865	.0148693	9.50	0.000	.1118605	.1707124

```
. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnReal TRCAPE LongInterestRate, lag(9)
```

Regression with Newey-West standard errors	Number of obs	=	129
maximum lag: 9	F(2, 126)	=	21.74
	Prob > F	=	0.0000

TenYearAnnuali~l	Coef.	Newey-West Std. Err.	t	P>|t|	[95% Conf.	Interval]
TRCAPE	-.0039998	.000831	-4.81	0.000	-.0056443	-.0023553
LongInterestRate	.0520977	.3669816	0.14	0.887	-.6741481	.7783434
_cons	.1412865	.0328164	4.31	0.000	.0763438	.2062292

Notice that the adjusted R^2 value drops in both regressions compared to their respective regressions which do not include the long interest rate. After making Newey-West adjustments, the long interest rate does not appear to have any statistical significance. It is for this reason that we do not include the long interest rate in all of our regressions.

Our specification for subsequent nominal returns regressed on CAPE is as follows:

$$\boldsymbol{TenYearAnnualizedReturnNom = \alpha + \beta * CAPE + \varepsilon}$$

```
. *Regress Ten Year Annualized Nominal Returns on CAPE*
. reg TenYearAnnualizedReturnNom CAPE

      Source |       SS       df       MS              Number of obs =     129
-------------+------------------------------           F(1, 127)     =   58.50
       Model |  .104977131     1  .104977131           Prob > F      =  0.0000
    Residual |  .227898219   127  .001794474           R-squared     =  0.3154
-------------+------------------------------           Adj R-squared =  0.3100
       Total |   .33287535   128  .002600589           Root MSE      =  .04236

------------------------------------------------------------------------------
TenYearAnn~m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        CAPE |  -.0043116   .0005637    -7.65   0.000    -.0054271   -.0031961
       _cons |   .1621629   .0099696    16.27   0.000     .1424349    .1818909
------------------------------------------------------------------------------

. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnNom CAPE, lag(9)

Regression with Newey-West standard errors          Number of obs     =    129
maximum lag: 9                                      F(  1,        127) =  24.86
                                                    Prob > F          = 0.0000

------------------------------------------------------------------------------
             |             Newey-West
TenYearAnn~m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        CAPE |  -.0043116   .0008648    -4.99   0.000    -.0060229   -.0026003
       _cons |   .1621629   .0199546     8.13   0.000     .1226763    .2016495
------------------------------------------------------------------------------
```

Our specification for subsequent real returns regressed on CAPE is as follows:

$$\boldsymbol{TenYearAnnualizedReturnReal} = \boldsymbol{\alpha} + \boldsymbol{\beta} * \boldsymbol{CAPE} + \boldsymbol{\varepsilon}$$

```
. *Regress Ten Year Annualized Real Returns on CAPE*
. reg TenYearAnnualizedReturnReal CAPE

      Source |       SS           df       MS      Number of obs   =       129
-------------+----------------------------------   F(1, 127)       =     53.59
       Model |  .098887675         1  .098887675   Prob > F        =    0.0000
    Residual |  .234354898       127  .001845314   R-squared       =    0.2967
-------------+----------------------------------   Adj R-squared   =    0.2912
       Total |  .333242573       128  .002603458   Root MSE        =    .04296

------------------------------------------------------------------------------
TenYearAnn~l |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        CAPE |  -.0041847   .0005716    -7.32   0.000    -.0053158   -.0030535
       _cons |   .1326201   .0101098    13.12   0.000     .1126146    .1526256
------------------------------------------------------------------------------

. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnReal CAPE, lag(9)

Regression with Newey-West standard errors      Number of obs     =        129
maximum lag: 9                                  F(  1,         127) =    24.55
                                                Prob > F          =     0.0000

------------------------------------------------------------------------------
             |             Newey-West
TenYearAnn~l |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        CAPE |  -.0041847   .0008446    -4.95   0.000     -.005856   -.0025133
       _cons |   .1326201   .0176741     7.50   0.000     .0976462     .167594
------------------------------------------------------------------------------
```

Notice that neither the adjusted R^2, nor the t statistic for CAPE drop off between the real and nominal regressions. Like TR CAPE, CAPE comes up as statistically significant when testing on many more observations than Tobin's Q and the Aggregate Investor Allocation to Equities (129 obs. Vs. 64). Just like TR CAPE, this makes them impossible to directly compare, however when TR CAPE and CAPE are regressed on the same 64 observations (1945-2008) as Tobin's Q and the Aggregate Investor Allocation to Equities, they capture less variation over the time period, with CAPE coming in behind TR CAPE.

As mentioned in the TR CAPE & long interest rate regressions, the long interest rate does not add any considerable explanatory power; this also holds true in the CAPE regressions below.

Our specification for subsequent nominal returns regressed on both CAPE and the 10-year interest rate is as follows:

$$TenYearAnnualizedReturnNom = \alpha + \beta_1 * CAPE + \beta_2 * LongInterestRate + \varepsilon$$

```
. *Regress Ten Year Annualized Nominal Returns on CAPE & 10 Year Interest Rate*
. reg TenYearAnnualizedReturnNom CAPE LongInterestRate

      Source |       SS           df       MS      Number of obs   =       129
-------------+----------------------------------   F(2, 126)       =     37.51
       Model |  .124221061         2  .06211053    Prob > F        =    0.0000
    Residual |   .20865429       126  .001655986   R-squared       =    0.3732
-------------+----------------------------------   Adj R-squared   =    0.3632
       Total |   .33287535       128  .002600589   Root MSE        =    .04069

------------------------------------------------------------------------------
TenYearAnnuali~m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------------+------------------------------------------------------------
            CAPE |   -.004127   .0005442    -7.58   0.000     -.005204    -.00305
LongInterestRate |   .5267472   .1545197     3.41   0.001     .2209573   .8325371
           _cons |   .1346077   .0125324    10.74   0.000     .1098064   .1594089
------------------------------------------------------------------------------

. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnNom CAPE LongInterestRate, lag(9)

Regression with Newey-West standard errors      Number of obs     =       129
maximum lag: 9                                  F(  2,        126) =     56.57
                                                Prob > F          =    0.0000

------------------------------------------------------------------------------
                 |             Newey-West
TenYearAnnuali~m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-----------------+------------------------------------------------------------
            CAPE |   -.004127   .0008357    -4.94   0.000    -.0057808   -.0024733
LongInterestRate |   .5267472    .327136     1.61   0.110    -.1206453     1.17414
           _cons |   .1346077   .0331969     4.05   0.000     .0689119    .2003034
------------------------------------------------------------------------------
```

Our specification for subsequent real returns regressed on both CAPE and the 10-year interest rate is as follows:

$$\boldsymbol{TenYearAnnualizedReturnReal = \alpha + \beta_1 * CAPE + \beta_2 * LongInterestRate + \varepsilon}$$

```
. *Regress Ten Year Annualized Real Returns on CAPE & 10 Year Interest Rate*
. reg TenYearAnnualizedReturnReal CAPE LongInterestRate
```

Source	SS	df	MS			
				Number of obs	=	129
				F(2, 126)	=	27.03
Model	.100037553	2	.050018777	Prob > F	=	0.0000
Residual	.23320502	126	.001850833	R-squared	=	0.3002
				Adj R-squared	=	0.2891
Total	.333242573	128	.002603458	Root MSE	=	.04302

TenYearAnnuali~l	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
CAPE	-.0041396	.0005754	-7.19	0.000	-.0052782	-.003001
LongInterestRate	.1287601	.1633575	0.79	0.432	-.1945196	.4520397
_cons	.1258844	.0132492	9.50	0.000	.0996647	.1521042

```
. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnReal CAPE LongInterestRate, lag(9)
```

```
Regression with Newey-West standard errors      Number of obs     =        129
maximum lag: 9                                  F(  2,        126) =      19.12
                                                Prob > F          =     0.0000
```

TenYearAnnuali~l	Coef.	Newey-West Std. Err.	t	P>|t|	[95% Conf.	Interval]
CAPE	-.0041396	.0009019	-4.59	0.000	-.0059245	-.0023547
LongInterestRate	.1287601	.3485908	0.37	0.712	-.5610909	.8186111
_cons	.1258844	.0299891	4.20	0.000	.0665368	.185232

Now we turn our attention to the models that include the Greenspan Put. As mentioned earlier, the correlation may be spurious, and regression fit may be high; this does not necessarily suggest a causal relationship, as all of the valuation indicators have been at historically high valuation levels on average since 1987. However, we may find some interesting results from the regressions.

Our specification for subsequent nominal returns regressed on both the aggregate investor allocation and the Greenspan Put is as follows:

$$\boldsymbol{TenYearAnnualizedReturnNom = \alpha + \beta_1 * PercentInStocks + \beta_2 * GreenspanPut + \varepsilon}$$

```
. *Regressing Ten Year Annualized Nominal Returns on %Invested in Stocks and the Greenspan Put*
. reg TenYearAnnualizedReturnNom PercentInStocks GreenspanPut
```

Source	SS	df	MS			
				Number of obs	=	64
				F(2, 61)	=	200.91
Model	.152811282	2	.076405641	Prob > F	=	0.0000
Residual	.023197963	61	.000380294	R-squared	=	0.8682
				Adj R-squared	=	0.8639
Total	.176009245	63	.002793798	Root MSE	=	.0195

TenYearAnnual~m	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
PercentInStocks	-.6892188	.0356056	-19.36	0.000	-.7604165	-.6180211
GreenspanPut	-.0027323	.00528	-0.52	0.607	-.0132902	.0078256
_cons	.3437087	.0120739	28.47	0.000	.3195655	.3678519

```
. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnNom PercentInStocks GreenspanPut, lag(9)

Regression with Newey-West standard errors          Number of obs     =        64
maximum lag: 9                                      F(  2,        61) =    122.21
                                                    Prob > F          =    0.0000
```

TenYearAnnual~m	Coef.	Newey-West Std. Err.	t	P>|t|	[95% Conf.	Interval]
PercentInStocks	-.6892188	.046034	-14.97	0.000	-.7812694	-.5971682
GreenspanPut	-.0027323	.0068981	-0.40	0.693	-.016526	.0110613
_cons	.3437087	.0193739	17.74	0.000	.3049683	.3824491

The inclusion of the Greenspan Put adds no additional explanatory power. It is highly likely that the Aggregate Investor Allocation to Equities already incorporates information from the Greenspan Put when determining the allocation. As a result, the two variables are likely to be collinear, with the aggregate allocation including much more other information that is statistically significant. This may be why the Greenspan Put is not statistically significant in this regression.

Our specification for subsequent real returns regressed on both the aggregate investor allocation and the Greenspan Put is as follows:

$$\boldsymbol{TenYearAnnualizedReturnReal} = \boldsymbol{\alpha} + \boldsymbol{\beta_1} * \boldsymbol{PercentInStocks} + \boldsymbol{\beta_2} * \boldsymbol{GreenspanPut} + \boldsymbol{\varepsilon}$$

```
. *Regressing Ten Year Annualized Real Returns on %Invested in Stocks and the Greenspan Put*
. reg TenYearAnnualizedReturnReal PercentInStocks GreenspanPut

      Source |       SS           df       MS      Number of obs   =        64
-------------+----------------------------------   F(2, 61)        =     75.63
       Model |  .150437155         2  .075218578   Prob > F        =    0.0000
    Residual |  .060665339        61  .000994514   R-squared       =    0.7126
-------------+----------------------------------   Adj R-squared   =    0.7032
       Total |  .211102494        63  .003350833   Root MSE        =    .03154

------------------------------------------------------------------------------
TenYearAnnual~l |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
PercentInStocks |  -.7061283   .0575789   -12.26   0.000    -.8212644   -.5909922
   GreenspanPut |   .0168488   .0085384     1.97   0.053    -.0002248    .0339223
          _cons |   .3044551    .019525    15.59   0.000     .2654123    .3434978
------------------------------------------------------------------------------

. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnReal PercentInStocks GreenspanPut, lag(9)

Regression with Newey-West standard errors      Number of obs     =        64
maximum lag: 9                                  F(  2,        61) =     35.44
                                                Prob > F          =    0.0000

------------------------------------------------------------------------------
                |             Newey-West
TenYearAnnual~l |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
PercentInStocks |  -.7061283   .0847997    -8.33   0.000    -.8756959   -.5365607
   GreenspanPut |   .0168488   .0154232     1.09   0.279    -.0139919    .0476894
          _cons |   .3044551    .034531     8.82   0.000      .235406    .3735042
------------------------------------------------------------------------------
```

Once again, the Greenspan Put variable carries no additional statistically significant explanatory power. Any power that it adds to the following regressions could ultimately be from chance correlation, which divides the time series into a piecewise function where both sides have the same coefficient/slope for the allocation to equities.

Our specification for subsequent nominal returns regressed on both Tobin's Q and the Greenspan Put is as follows:

$$\boldsymbol{TenYearAnnualizedReturnNom} = \boldsymbol{\alpha} + \boldsymbol{\beta_1} * \boldsymbol{TobinsQ} + \boldsymbol{\beta_2} * \boldsymbol{GreenspanPut} + \boldsymbol{\varepsilon}$$

```
. *Regressing Ten Year Annualized Nominal Returns on Tobin's Q and the Greenspan Put*
. reg TenYearAnnualizedReturnNom TobinsQ GreenspanPut
```

Source	SS	df	MS			
				Number of obs	=	64
				F(2, 61)	=	235.77
Model	.15584843	2	.077924215	Prob > F	=	0.0000
Residual	.020160815	61	.000330505	R-squared	=	0.8855
				Adj R-squared	=	0.8817
Total	.176009245	63	.002793798	Root MSE	=	.01818

TenYearAnn~m	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
TobinsQ	-.2146139	.0102275	-20.98	0.000	-.2350649	-.1941628
GreenspanPut	.0525005	.006095	8.61	0.000	.0403128	.0646881
_cons	.2327418	.0061724	37.71	0.000	.2203993	.2450843

```
. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnNom TobinsQ GreenspanPut, lag(9)

Regression with Newey-West standard errors      Number of obs     =        64
maximum lag: 9                                  F(  2,        61) =    110.75
                                                Prob > F          =    0.0000
```

TenYearAnn~m	Coef.	Newey-West Std. Err.	t	P>|t|	[95% Conf.	Interval]
TobinsQ	-.2146139	.0144897	-14.81	0.000	-.2435877	-.18564
GreenspanPut	.0525005	.0088702	5.92	0.000	.0347634	.0702375
_cons	.2327418	.0091239	25.51	0.000	.2144975	.2509861

Our specification for subsequent real returns regressed on both Tobin's Q and the Greenspan Put is as follows:

$$\boldsymbol{TenYearAnnualizedReturnReal = \alpha + \beta_1 * TobinsQ + \beta_2 * GreenspanPut + \varepsilon}$$

```
. *Regressing Ten Year Annualized Real Returns on Tobin's Q and the Greenspan Put*
. reg TenYearAnnualizedReturnReal TobinsQ GreenspanPut
```

Source	SS	df	MS		
				Number of obs =	64
				F(2, 61) =	115.13
Model	.166891356	2	.083445678	Prob > F =	0.0000
Residual	.044211137	61	.000724773	R-squared =	0.7906
				Adj R-squared =	0.7837
Total	.211102494	63	.003350833	Root MSE =	.02692

TenYearAnn~l	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
TobinsQ	-.229228	.0151454	-15.14	0.000	-.2595131	-.198943
GreenspanPut	.076888	.0090258	8.52	0.000	.0588399	.0949362
_cons	.1957915	.0091404	21.42	0.000	.1775141	.2140689

```
. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnReal TobinsQ GreenspanPut, lag(9)

Regression with Newey-West standard errors      Number of obs     =         64
maximum lag: 9                                  F(  2,        61) =      40.50
                                                Prob > F          =     0.0000
```

TenYearAnn~l	Coef.	Newey-West Std. Err.	t	P>|t|	[95% Conf.	Interval]
TobinsQ	-.229228	.0261429	-8.77	0.000	-.281504	-.1769521
GreenspanPut	.076888	.0135442	5.68	0.000	.0498047	.1039714
_cons	.1957915	.0194533	10.06	0.000	.1568921	.2346908

This regression is particularly interesting because it is actually fit slightly better than the regression using the aggregate equity allocation. Both Tobin's Q and the Greenspan Put are statistically significant after NW corrections. However, as noted earlier we cannot draw conclusions about causality about the Greenspan Put, as **there are numerous omitted variables that may have changed over the period**. For example, if we had created a binary variable that included and followed the flash crash of 1987, we would have had the same exact regression results. If we go down the rabbit hole, we could have also constructed a binary variable that denoted years including and following the release of the Stanley Kubrick film, "Full Metal Jacket". Once again, the results would be the exactly the same. The era of "high valuations" that began in the 1990s coincides with this **artificial indicator**, and a much longer time series would be required to draw significant conclusions about the nature and cause of these elevated valuations.

These remarks also apply to the subsequent return regressions on variations of TR CAPE and CAPE that include the Greenspan Put below.

Our specification for subsequent nominal returns regressed on both Total Return CAPE and the Greenspan Put is as follows:

$$\boldsymbol{TenYearAnnualizedReturnNom = \alpha + \beta_1 * TRCAPE + \beta_2 * GreenspanPut + \varepsilon}$$

```
. *Regressing Ten Year Annualized Nominal Returns on TR CAPE and the Greenspan Put*
. reg TenYearAnnualizedReturnNom TRCAPE GreenspanPut

      Source |       SS           df       MS      Number of obs   =       129
-------------+----------------------------------   F(2, 126)       =     59.31
       Model |  .161419856         2  .080709928   Prob > F        =    0.0000
    Residual |  .171455494       126  .001360758   R-squared       =    0.4849
-------------+----------------------------------   Adj R-squared   =    0.4768
       Total |   .33287535       128  .002600589   Root MSE        =    .03689

------------------------------------------------------------------------------
TenYearAnn~m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      TRCAPE |  -.0059915   .0005501   -10.89   0.000    -.0070802   -.0049028
GreenspanPut |    .057236   .0101558     5.64   0.000      .037138    .0773339
       _cons |   .2011106   .0106677    18.85   0.000     .1799996    .2222217
------------------------------------------------------------------------------

. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnNom TRCAPE GreenspanPut, lag(9)

Regression with Newey-West standard errors      Number of obs     =        129
maximum lag: 9                                  F(  2,        126) =     28.92
                                                Prob > F          =     0.0000

------------------------------------------------------------------------------
             |             Newey-West
TenYearAnn~m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      TRCAPE |  -.0059915   .0007901    -7.58   0.000    -.0075551   -.0044279
GreenspanPut |    .057236   .0138615     4.13   0.000     .0298044    .0846675
       _cons |   .2011106   .0203463     9.88   0.000     .1608459    .2413754
------------------------------------------------------------------------------
```

Our specification for subsequent real returns regressed on both Total Return CAPE and the Greenspan Put is as follows:

$$\boldsymbol{TenYearAnnualizedReturnReal} = \boldsymbol{\alpha} + \boldsymbol{\beta_1} * \boldsymbol{TRCAPE} + \boldsymbol{\beta_2} * \boldsymbol{GreenspanPut} + \boldsymbol{\varepsilon}$$

```
. *Regressing Ten Year Annualized Real Returns on TR CAPE and the Greenspan Put*
. reg TenYearAnnualizedReturnReal TRCAPE GreenspanPut
```

Source	SS	df	MS		
				Number of obs =	129
				F(2, 126) =	46.54
Model	.141589299	2	.07079465	Prob > F =	0.0000
Residual	.191653274	126	.001521058	R-squared =	0.4249
				Adj R-squared =	0.4158
Total	.333242573	128	.002603458	Root MSE =	.039

TenYearAnn~l	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
TRCAPE	-.0056113	.0005816	-9.65	0.000	-.0067624	-.0044603
GreenspanPut	.0555277	.0107373	5.17	0.000	.0342789	.0767766
_cons	.1663635	.0112785	14.75	0.000	.1440437	.1886834

```
. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnReal TRCAPE GreenspanPut, lag(9)
```

Regression with Newey-West standard errors — Number of obs = 129
maximum lag: 9 — F(2, 126) = 28.08
Prob > F = 0.0000

TenYearAnn~l	Coef.	Newey-West Std. Err.	t	P>|t|	[95% Conf.	Interval]
TRCAPE	-.0056113	.0007509	-7.47	0.000	-.0070974	-.0041253
GreenspanPut	.0555277	.0145257	3.82	0.000	.0267818	.0842737
_cons	.1663635	.0176037	9.45	0.000	.1315262	.2012008

Our specification for subsequent nominal returns regressed on both CAPE and the Greenspan Put is as follows:

$$TenYearAnnualizedReturnNom = \alpha + \beta_1 * CAPE + \beta_2 * GreenspanPut + \varepsilon$$

```
. *Regressing Ten Year Annualized Nominal Returns on CAPE and the Greenspan Put*
. reg TenYearAnnualizedReturnNom CAPE GreenspanPut

      Source |       SS           df       MS      Number of obs   =       129
-------------+----------------------------------   F(2, 126)       =     57.79
       Model |  .159253558         2  .079626779   Prob > F        =    0.0000
    Residual |  .173621792       126  .001377951   R-squared       =    0.4784
-------------+----------------------------------   Adj R-squared   =    0.4701
       Total |   .33287535       128  .002600589   Root MSE        =    .03712

------------------------------------------------------------------------------
TenYearAnn~m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        CAPE |  -.0065904   .0006131   -10.75   0.000    -.0078036   -.0053771
GreenspanPut |   .0676855   .0107847     6.28   0.000     .0463429     .089028
       _cons |   .1879944   .0096572    19.47   0.000      .168883    .2071058
------------------------------------------------------------------------------

. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnNom CAPE GreenspanPut, lag(9)

Regression with Newey-West standard errors      Number of obs     =        129
maximum lag: 9                                  F(  2,        126) =     30.11
                                                Prob > F          =     0.0000

------------------------------------------------------------------------------
             |             Newey-West
TenYearAnn~m |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        CAPE |  -.0065904   .0008835    -7.46   0.000    -.0083388    -.004842
GreenspanPut |   .0676855   .0119985     5.64   0.000     .0439408    .0914302
       _cons |   .1879944   .0196414     9.57   0.000     .1491246    .2268642
------------------------------------------------------------------------------
```

Our specification for subsequent real returns regressed on both CAPE and the Greenspan Put is as follows:

$$\boldsymbol{TenYearAnnualizedReturnReal} = \boldsymbol{\alpha} + \boldsymbol{\beta_1} * \boldsymbol{CAPE} + \boldsymbol{\beta_2} * \boldsymbol{GreenspanPut} + \boldsymbol{\varepsilon}$$

```
. *Regressing Ten Year Annualized Real Returns on CAPE and the Greenspan Put*
. reg TenYearAnnualizedReturnReal CAPE GreenspanPut
```

Source	SS	df	MS			
				Number of obs	=	129
				F(2, 126)	=	54.66
Model	.154814092	2	.077407046	Prob > F	=	0.0000
Residual	.178428482	126	.001416099	R-squared	=	0.4646
				Adj R-squared	=	0.4561
Total	.333242573	128	.002603458	Root MSE	=	.03763

TenYearAnn~l	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
CAPE	-.0064978	.0006215	-10.46	0.000	-.0077277	-.0052679
GreenspanPut	.0687066	.0109329	6.28	0.000	.0470706	.0903425
_cons	.1588413	.00979	16.22	0.000	.1394672	.1782155

```
. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey TenYearAnnualizedReturnReal CAPE GreenspanPut, lag(9)
```

Regression with Newey-West standard errors — maximum lag: 9

Number of obs	=	129
F(2, 126)	=	32.52
Prob > F	=	0.0000

TenYearAnn~l	Coef.	Newey-West Std. Err.	t	P>|t|	[95% Conf.	Interval]
CAPE	-.0064978	.0008057	-8.06	0.000	-.0080923	-.0049033
GreenspanPut	.0687066	.0131195	5.24	0.000	.0427434	.0946697
_cons	.1588413	.0157004	10.12	0.000	.1277706	.1899121

The aggregate allocation to equities is of particular interest to us, not simply because it appears that it may have the best correlation to subsequent real returns over the following 10 years, but because the efficient market hypothesis' implications directly affect the metric. The efficient market hypothesis' claim that all investors pursue the highest return for an equivalent measure of risk has a corollary: **in equilibrium the aggregate allocation to equities should only move higher if its expected returns per unit of expected risk increase in aggregate.** Either expected return should be greater (holding expected risk constant) or expected risk should be lower (holding expected returns constant) when the allocation is higher. The regression of annualized real returns over the subsequent 10 years above clearly shows that when the investor allocation is high, investors' subsequent returns are lower; however, if risk declined over the following period by more than returns declined, the shift upwards in the equity allocation could still be considered rational under efficient market theory. We construct a measure of risk below to test if historically high values of the aggregate investor allocation have been correlated with lower amounts of risk.

We construct the average annual loss as a measure of risk. First, we identify years with gains and years with losses. Then we create a binary variable equal to 1 for years with losses and multiply it by the LN(return) vector to create a vector which only includes the size of losses in years of losses. For each year, we then average the size of those losses over the subsequent 10 years.

Our Specification for average annual nominal losses (in years where losses occur) over the subsequent 10-year period regressed on the aggregate investor allocation is as follows:

$$\boldsymbol{AverageAnnualLossOverNext10 = \alpha + \beta * PercentInStocks + \varepsilon}$$

```
. *Regress Average Annual Loss (for years with losses) over the Next 10 Years on % Invested In Stocks*
. *This is to determine if % Invested In Stocks successfully predicts periods of lower than average risk*
. reg AverageAnnualLossOverNext10 PercentInStocks

      Source |       SS           df       MS      Number of obs   =        64
-------------+----------------------------------   F(1, 62)        =    245.56
       Model |  .026805655         1  .026805655   Prob > F        =    0.0000
    Residual |  .006768004        62  .000109161   R-squared       =    0.7984
-------------+----------------------------------   Adj R-squared   =    0.7952
       Total |  .033573659        63  .000532915   Root MSE        =    .01045

------------------------------------------------------------------------------
AverageAnnua~10 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
PercentInStocks |  -.2905729   .0185428   -15.67   0.000    -.3276396   -.2535063
          _cons |   .0683643   .0064451    10.61   0.000     .0554806    .0812479
------------------------------------------------------------------------------

. *Run Newey-West Estimator to correct for overlapping Moving Average*
. newey AverageAnnualLossOverNext10 PercentInStocks, lag(9)

Regression with Newey-West standard errors      Number of obs     =        64
maximum lag: 9                                  F(  1,        62) =    229.06
                                                Prob > F          =    0.0000

------------------------------------------------------------------------------
                |             Newey-West
AverageAnnua~10 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
----------------+-------------------------------------------------------------
PercentInStocks |  -.2905729   .0191989   -15.13   0.000     -.328951   -.2521948
          _cons |   .0683643   .0066221    10.32   0.000     .0551268    .0816018
------------------------------------------------------------------------------
```

The regression above demonstrates that the aggregate equity allocation has a statistically significant negative relationship with the average annual nominal loss over subsequent 10-year periods (for years with losses). Since losses are measured as zero or negative, the negative sign on the regression coefficient indicates that losses are significantly greater following years of a high allocation to equities.

While volatility is a common proxy for risk we disagree with its use as an indicator of risk. In an efficient market where actual returns are log-normally distributed, volatility as a primary indicator of risk would make sense since there is an equal chance of excess returns and excess declines over any given period. Volatility would be important because it would be an indicator of how far away portfolio value could diverge from the long-run return over a given period of time. While portfolio returns should revert back to this hypothetical mean return for portfolios of equivalent risk over a long enough time period, people are concerned about short term deviations because our lives are short. Another issue with volatility is that it increases when the size of positive excess returns increases (relative to historical excess returns); rational investors would generally accept this as a good thing. We believe that investors are ultimately concerned about the risk of losses (divergence below long-run expected returns) because investors may be forced to sell at unfavorable valuations due to unforeseen life circumstances.

If we take this as a proxy for risk, it is implied that high equity allocations are in fact shown to have greater risk. A high equity allocation should only be consistent with higher expectations of return and/or lower expectations of risk. Our analysis shows that these expectations are likely to be irrational, as subsequent returns are lower and risk is greater. Below we conduct an empirical analysis of the factors which explain the level of the aggregate investor allocation:

b. Aggregate Investor Allocation to Equities Regressions

1. **Data is the same as above.**

2. **Dependent Variables for Aggregate Investor Allocation Regressions**
    a. Aggregate Investor Allocation to Equities

3. **Independent Variables for Aggregate Investor Allocation Regressions**
    a. Inflation Rate (current year)
    b. Inflation Rate over the Past 10 Years (annualized)
    c. Civilian Unemployment (U-3)
    d. The Ten-Year U.S. Treasury Yield (GS10) or Historical Equivalent
    e. Greenspan Put
    f. Expected Return (Koyck Lag λ=1%)
    g. Expected Return (Koyck Lag λ=3%)
    h. Nominal Expected Earnings Growth (Koyck Lag λ=1%)
    i. Nominal Expected Earnings Growth (Koyck Lag λ=3%)
    j. Real Expected Earnings Growth (Koyck Lag λ=1%)
    k. Real Expected Earnings Growth (Koyck Lag λ=3%)

Since we have shown that a higher aggregate investor allocation is associated with lower returns and greater risk, we now turn our attention to the determinants of this metric. We use the current year inflation rate, civilian unemployment, and the inflation rate over the prior 10 years (annualized) in addition to the variables mentioned earlier.

We construct Koyck distributed lag functions for a number of variables which denote expectations for the associated metric. $\boldsymbol{\lambda}$ denotes the proportional weight of the current year. $(\mathbf{1-\lambda})$ denotes the proportional weight of the previous expectation. For each variable being tested we test two versions of the lag where λ=1% and 3%. When lambda is included in the Koyck distributed lag variables below it does not include the percentage symbol. All of the expectations are annualized. The functional form of the Koyck lag for expected nominal returns is as follows:

$$\boldsymbol{ERKoyck0\lambda_t} = \lambda * \boldsymbol{ln}\left(\frac{\boldsymbol{S_t}}{\boldsymbol{S_{t-1}}}\right) + (\mathbf{1}-\lambda) * \boldsymbol{ERKoyck0\lambda_{t-1}}$$

$\boldsymbol{ER_t}$ is the expected future return at time ***t***, and $\boldsymbol{ER_{t-1}}$ is the expected future return at time ***t – 1.*** $\boldsymbol{S_t}$ is the nominal portfolio value (with dividends reinvested) at time ***t***, and $\boldsymbol{S_{t-1}}$ is the nominal portfolio value (with dividends reinvested) at time ***t – 1.*** We use the natural log to calculate returns because it is additively symmetric while arithmetic returns are not.

The functional form of the Koyck lag for expected real earnings growth is as follows:

$$\boldsymbol{EEG0\lambda Real_t} = \lambda * \boldsymbol{ln}\left(\frac{\boldsymbol{Real\ Earnings_t}}{\boldsymbol{Real\ Earnings_{t-1}}}\right) + (\mathbf{1}-\lambda) * \boldsymbol{EEG0\lambda Real_{t-1}}$$

$\boldsymbol{EEG0\lambda Real_t}$ is the expectation of future real earnings growth at time ***t***, and $\boldsymbol{EEG0\lambda Real_{t-1}}$ is the expectation of future real earnings growth at time ***t – 1***. $\boldsymbol{Real\ Earnings_t}$ is the current year's real earnings growth, and $\boldsymbol{Real\ Earnings_{t-1}}$ is the prior year's real earnings.

The functional form of the Koyck lag for expected nominal earnings growth is as follows:

$$\boldsymbol{EEG0\lambda Nominal_t} = \lambda * \boldsymbol{ln}\left(\frac{\boldsymbol{Nominal\ Earnings_t}}{\boldsymbol{Nominal\ Earnings_{t-1}}}\right) + (\mathbf{1}-\lambda) * \boldsymbol{EEG0\lambda Nominal_{t-1}}$$

$\boldsymbol{EEG0\lambda Nominal_t}$ is the expectation of future nominal earnings growth at time ***t***, and $\boldsymbol{EEG0\lambda Nominal_{t-1}}$ is the expectation of future nominal earnings growth at time ***t – 1***.

Now since the Koyck weights for the present year are very low, this means that it can take a significant period of time for past years to drop out. When evaluating an approach to the model, we considered using higher weights on the present year, and dropping the earliest data in the sample when running regressions. We decided against this because the Koyck expected return was far too volatile, and frequently fell below zero in the event of a market crash or a prolonged decline. Expected long-run returns should always be positive, otherwise no one would hold the asset, and the aggregate investor allocation to equities would be zero.

What we decided to do instead may be problematic: we maintained low weights on the current year, and entered in the geometric average return from the end of 1871 to the end of 2018 as the first expected Koyck return in the data.

Theoretically this is not reasonable – someone in 1871 simply could not have perfect foresight of 150 years of data and then decided what their expected return is was going to be. **However, we insert the geometric average return in the first year because our analysis is concerned with time variation in expected returns due to recent experience being built into expectations.** Had we used these small weights on the current year and utilized the first year's return, this also would have theoretical problems; clearly one year of returns is not enough to develop expectations for the long-run. We figured that by using the geometric average, any deviation from a "reasonable" long-run return would be captured by our Koyck distributed lags.

4. Regression Specifications and Results

Our specification for the aggregate investor allocation regressed on Nominal Expected Earnings Growth (Koyck Lag λ=1%) is as follows:

$$\boldsymbol{PercentInStocks = \alpha + \beta * EEG01Nominal + \varepsilon}$$

```
. *Regressing % in Stocks on Expected Nominal Earnings Growth (Koyck .01)
. reg PercentInStocks EEG01Nominal
```

Source	SS	df	MS			
				Number of obs	=	73
				F(1, 71)	=	4.55
Model	.020973192	1	.020973192	Prob > F	=	0.0364
Residual	.327389014	71	.004611113	R-squared	=	0.0602
				Adj R-squared	=	0.0470
Total	.348362205	72	.004838364	Root MSE	=	.06791

PercentInS~s	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
EEG01Nominal	3.286876	1.541182	2.13	0.036	.2138463	6.359905
_cons	.2212987	.0589567	3.75	0.000	.1037424	.3388551

Our specification for the aggregate investor allocation regressed on Nominal Expected Earnings Growth (Koyck Lag λ=3%) is as follows:

$$\boldsymbol{PercentInStocks = \alpha + \beta * EEG03Nominal + \varepsilon}$$

```
. *Regressing % in Stocks on Expected Nominal Earnings Growth (Koyck .03)
. reg PercentInStocks EEG03Nominal
```

Source	SS	df	MS			
				Number of obs	=	73
				F(1, 71)	=	2.68
Model	.012669186	1	.012669186	Prob > F	=	0.1061
Residual	.335693019	71	.004728071	R-squared	=	0.0364
				Adj R-squared	=	0.0228
Total	.348362205	72	.004838364	Root MSE	=	.06876

PercentInS~s	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
EEG03Nominal	1.186208	.7246502	1.64	0.106	-.2587036	2.631119
_cons	.2906226	.0347073	8.37	0.000	.2214182	.359827

Our specification for the aggregate investor allocation regressed on Real Expected Earnings Growth (Koyck Lag λ=1%) is as follows:

$$\boldsymbol{PercentInStocks = \alpha + \beta * EEG01Real + \varepsilon}$$

```
. *Regressing % in Stocks on Expected Real Earnings Growth (Koyck .01)
. reg PercentInStocks EEG01Real
```

Source	SS	df	MS			
				Number of obs	=	73
				F(1, 71)	=	16.04
Model	.064186842	1	.064186842	Prob > F	=	0.0002
Residual	.284175363	71	.00400247	R-squared	=	0.1843
				Adj R-squared	=	0.1728
Total	.348362205	72	.004838364	Root MSE	=	.06327

PercentInS~s	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
EEG01Real	9.968397	2.489238	4.00	0.000	5.004998	14.9318
_cons	.179944	.0420947	4.27	0.000	.0960096	.2638785

Our specification for the aggregate investor allocation regressed on Real Expected Earnings Growth (Koyck Lag λ=3%) is as follows:

$$\boldsymbol{PercentInStocks = \alpha + \beta * EEG03Real + \varepsilon}$$

```
. *Regressing % in Stocks on Expected Real Earnings Growth (Koyck .03)
. reg PercentInStocks EEG03Real
```

Source	SS	df	MS
Model	.064592177	1	.064592177
Residual	.283770029	71	.003996761
Total	.348362205	72	.004838364

Number of obs	=	73
F(1, 71)	=	16.16
Prob > F	=	0.0001
R-squared	=	0.1854
Adj R-squared	=	0.1739
Root MSE	=	.06322

PercentInS~s	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
EEG03Real	3.354449	.8344212	4.02	0.000	1.69066	5.018237
_cons	.2857901	.0166803	17.13	0.000	.2525305	.3190498

The expected real earnings growth metrics seem to have some explanatory power on the Aggregate Investor Allocation to Equities. We move on to more extensive models below.

Our specification for the aggregate investor allocation regressed on the inflation rate, U-3, the 10-year treasury (GS10), Greenspan Put, Expected Return (Koyck Lag λ=1%), and Real Expected Earnings Growth (Koyck Lag λ=1%) is as follows:

$$\boldsymbol{PercentInStocks = \alpha + \beta_1 * InflationRate + \beta_2 * CivUnemployment + \beta_3 * LongInterestRate + \beta_4 * GreenspanPut + \beta_5 * ERKoyck01 + \beta_6 * EEG01Real + \varepsilon}$$

```
. reg PercentInStocks InflationRate CivUnemployment LongInterestRate GreenspanPut ERKoyck01 EEG01Real
```

Source	SS	df	MS
Model	.281010945	6	.046835157
Residual	.063563616	64	.000993181
Total	.344574561	70	.004922494

Number of obs	=	71
F(6, 64)	=	47.16
Prob > F	=	0.0000
R-squared	=	0.8155
Adj R-squared	=	0.7982
Root MSE	=	.03151

PercentInStocks	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
InflationRate	.406676	.1826087	2.23	0.029	.0418731	.7714789
CivUnemployment	-.0612897	.3037391	-0.20	0.841	-.6680783	.5454989
LongInterestRate	-1.612037	.2127576	-7.58	0.000	-2.037069	-1.187004
GreenspanPut	-.0933793	.0137667	-6.78	0.000	-.1208814	-.0658771
ERKoyck01	17.92264	1.645895	10.89	0.000	14.63459	21.21069
EEG01Real	3.940566	1.500939	2.63	0.011	.9420952	6.939037
_cons	-1.145791	.1380525	-8.30	0.000	-1.421583	-.8699995

Our specification for the aggregate investor allocation regressed on the inflation rate, U-3, the 10-year treasury (GS10), Greenspan Put, Expected Return (Koyck Lag λ=3%), and Real Expected Earnings Growth (Koyck Lag λ=3%) is as follows:

$$\boldsymbol{PercentInStocks = \alpha + \beta_1 * InflationRate + \beta_2 * CivUnemployment + \beta_3 * LongInterestRate + \beta_4 * GreenspanPut + \beta_5 * ERKoyck03 + \beta_6 * EEG03Real + \varepsilon}$$

```
. reg PercentInStocks InflationRate CivUnemployment LongInterestRate GreenspanPut ERKoyck03 EEG03Real
```

Source	SS	df	MS			
				Number of obs	=	71
				F(6, 64)	=	26.46
Model	.245565949	6	.040927658	Prob > F	=	0.0000
Residual	.099008611	64	.00154701	R-squared	=	0.7127
				Adj R-squared	=	0.6857
Total	.344574561	70	.004922494	Root MSE	=	.03933

PercentInStocks	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
InflationRate	.3449966	.2309307	1.49	0.140	-.1163405	.8063337
CivUnemployment	.3061761	.4113871	0.74	0.459	-.5156642	1.128016
LongInterestRate	-1.486662	.2726273	-5.45	0.000	-2.031298	-.9420267
GreenspanPut	-.0299034	.0128497	-2.33	0.023	-.0555738	-.0042331
ERKoyck03	5.413998	.7285784	7.43	0.000	3.958496	6.869501
EEG03Real	1.9733	.6208256	3.18	0.002	.7330581	3.213541
_cons	-.1343653	.0755511	-1.78	0.080	-.285296	.0165654

Our specification for the aggregate investor allocation regressed on the inflation rate, U-3, the 10-year treasury (GS10), Greenspan Put, and Expected Return (Koyck Lag λ=1%) is as follows:

$$\boldsymbol{PercentInStocks = \alpha + \beta_1 * InflationRate + \beta_2 * CivUnemployment + \beta_3 * LongInterestRate + \beta_4 * GreenspanPut + \beta_5 * ERKoyck01 + \varepsilon}$$

```
. reg PercentInStocks InflationRate CivUnemployment LongInterestRate GreenspanPut ERKoyck01
```

Source	SS	df	MS			
				Number of obs	=	72
				F(5, 66)	=	51.41
Model	.277309585	5	.055461917	Prob > F	=	0.0000
Residual	.071197358	66	.001078748	R-squared	=	0.7957
				Adj R-squared	=	0.7802
Total	.348506943	71	.004908548	Root MSE	=	.03284

PercentInStocks	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
InflationRate	.5545544	.1806177	3.07	0.003	.1939395	.9151694
CivUnemployment	-.3173457	.3013379	-1.05	0.296	-.9189863	.2842949
LongInterestRate	-1.718076	.217397	-7.90	0.000	-2.152123	-1.284029
GreenspanPut	-.0953025	.0141137	-6.75	0.000	-.1234814	-.0671235
ERKoyck01	18.71101	1.669313	11.21	0.000	15.37812	22.0439
_cons	-1.130941	.1430987	-7.90	0.000	-1.416646	-.8452347

Our specification for the aggregate investor allocation regressed on the inflation rate, U-3, the 10-year treasury (GS10), Greenspan Put, and Expected Return (Koyck Lag λ=3%) is as follows:

$$\boldsymbol{PercentInStocks = \alpha + \beta_1 * InflationRate + \beta_2 * CivUnemployment + \beta_3 * LongInterestRate + \beta_4 * GreenspanPut + \beta_5 * ERKoyck03 + \varepsilon}$$

```
. reg PercentInStocks InflationRate CivUnemployment LongInterestRate GreenspanPut ERKoyck03
```

Source	SS	df	MS
Model	.231268377	5	.046253675
Residual	.117238566	66	.001776342
Total	.348506943	71	.004908548

Number of obs	=	72
F(5, 66)	=	26.04
Prob > F	=	0.0000
R-squared	=	0.6636
Adj R-squared	=	0.6381
Root MSE	=	.04215

PercentInStocks	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
InflationRate	.5105188	.238812	2.14	0.036	.0337153	.9873224
CivUnemployment	-.1892226	.4121393	-0.46	0.648	-1.012085	.63364
LongInterestRate	-1.591612	.287841	-5.53	0.000	-2.166305	-1.016918
GreenspanPut	-.0258501	.0134458	-1.92	0.059	-.0526955	.0009952
ERKoyck03	5.427385	.7646578	7.10	0.000	3.900697	6.954074
_cons	-.0719803	.0779089	-0.92	0.359	-.2275305	.0835698

Our specification for the aggregate investor allocation regressed on the annualized inflation rate over the past 10 years, U-3, the 10-year treasury (GS10), Greenspan Put, and Expected Return (Koyck Lag λ=1%) is as follows:

$$\boldsymbol{PercentInStocks = \alpha + \beta_1 * Past10YearsInflationAnnualized + \beta_2 * CivUnemployment + \beta_3 * LongInterestRate + \beta_4 * GreenspanPut + \beta_5 * ERKoyck01 + \varepsilon}$$

```
. reg PercentInStocks Past10YearsInflationAnnualized CivUnemployment LongInterestRate GreenspanPut ERKoyck01
```

Source	SS	df	MS
Model	.284984159	5	.056996832
Residual	.063522784	66	.000962466
Total	.348506943	71	.004908548

Number of obs	=	72
F(5, 66)	=	59.22
Prob > F	=	0.0000
R-squared	=	0.8177
Adj R-squared	=	0.8039
Root MSE	=	.03102

PercentInStocks	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Past10YearsInflationAnnualized	-1.589304	.3691103	-4.31	0.000	-2.326256	-.8523514
CivUnemployment	-.6858577	.2898149	-2.37	0.021	-1.264492	-.1072236
LongInterestRate	-.3680635	.2744638	-1.34	0.185	-.9160482	.1799212
GreenspanPut	-.0580999	.0144423	-4.02	0.000	-.0869349	-.0292649
ERKoyck01	11.00428	1.934407	5.69	0.000	7.142108	14.86645
_cons	-.4555884	.1681488	-2.71	0.009	-.7913084	-.1198685

Our specification for the aggregate investor allocation regressed on the annualized inflation rate over the past 10 years, U-3, the 10-year treasury (GS10), Greenspan Put, and Expected Return (Koyck Lag λ=3%) is as follows:

$$\boldsymbol{PercentInStocks = \alpha + \beta_1 * Past10YearsInflationAnnualized + \beta_2 * CivUnemployment + \beta_3 * LongInterestRate + \beta_4 * GreenspanPut + \beta_5 * ERKoyck03 + \varepsilon}$$

```
. reg PercentInStocks Past10YearsInflationAnnualized CivUnemployment LongInterestRate GreenspanPut ERKoyck03
```

Source	SS	df	MS			
				Number of obs	=	72
				F(5, 66)	=	41.96
Model	.26510745	5	.05302149	Prob > F	=	0.0000
Residual	.083399493	66	.001263629	R-squared	=	0.7607
				Adj R-squared	=	0.7426
Total	.348506943	71	.004908548	Root MSE	=	.03555

PercentInStocks	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Past10YearsInflationAnnualized	-2.278536	.395425	-5.76	0.000	-3.068028	-1.489045
CivUnemployment	-.8491326	.3522688	-2.41	0.019	-1.55246	-.1458052
LongInterestRate	.1354658	.2998884	0.45	0.653	-.4632807	.7342124
GreenspanPut	-.0099332	.0113933	-0.87	0.386	-.0326805	.0128142
ERKoyck03	2.144087	.7179421	2.99	0.004	.7106697	3.577505
_cons	.2781538	.0752446	3.70	0.000	.127923	.4283845

Now we drop the Greenspan Put variable. Our specification for the aggregate investor allocation regressed on the annualized inflation rate over the past 10 years, U-3, the 10-year treasury (GS10), and Expected Return (Koyck Lag λ=1%) is as follows:

$$\boldsymbol{PercentInStocks = \alpha + \beta_1 * Past10YearsInflationAnnualized + \beta_2 * CivUnemployment + \beta_3 * LongInterestRate + \beta_4 * ERKoyck01 + \varepsilon}$$

```
. reg PercentInStocks Past10YearsInflationAnnualized CivUnemployment LongInterestRate  ERKoyck01
```

Source	SS	df	MS			
				Number of obs	=	72
				F(4, 67)	=	57.05
Model	.269407861	4	.067351965	Prob > F	=	0.0000
Residual	.079099082	67	.001180583	R-squared	=	0.7730
				Adj R-squared	=	0.7595
Total	.348506943	71	.004908548	Root MSE	=	.03436

PercentInStocks	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Past10YearsInflationAnnualized	-2.240873	.3673425	-6.10	0.000	-2.974091	-1.507654
CivUnemployment	-1.215093	.286004	-4.25	0.000	-1.78596	-.6442267
LongInterestRate	.2864329	.2448155	1.17	0.246	-.2022208	.7750866
ERKoyck01	4.513156	1.181695	3.82	0.000	2.154483	6.87183
_cons	.0977933	.1071037	0.91	0.364	-.1159867	.3115732

Our specification for the aggregate investor allocation regressed on the annualized inflation rate over the past 10 years, U-3, the 10-year treasury (GS10), and Expected Return (Koyck Lag λ=3%) is as follows:

$$\boldsymbol{PercentInStocks = \alpha + \beta_1 * Past10YearsInflationAnnualized + \beta_2 * CivUnemployment + \beta_3 * LongInterestRate + \beta_4 * ERKoyck03 + \varepsilon}$$

```
. reg PercentInStocks Past10YearsInflationAnnualized CivUnemployment LongInterestRate ERKoyck03

      Source |       SS           df       MS      Number of obs   =        72
-------------+----------------------------------   F(4, 67)        =     52.45
       Model |  .264146951         4  .066036738   Prob > F        =    0.0000
    Residual |  .084359993        67  .001259104   R-squared       =    0.7579
-------------+----------------------------------   Adj R-squared   =    0.7435
       Total |  .348506943        71  .004908548   Root MSE        =    .03548

------------------------------------------------------------------------------------------------
               PercentInStocks |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------------------------+----------------------------------------------------------------
Past10YearsInflationAnnualized |  -2.339574   .3884808    -6.02   0.000    -3.114984   -1.564163
               CivUnemployment |  -.9840466   .3158933    -3.12   0.003    -1.614572   -.3535209
              LongInterestRate |   .2415173   .2736196     0.88   0.381    -.3046297    .7876643
                     ERKoyck03 |   1.769478   .5741394     3.08   0.003     .6234909    2.915465
                         _cons |   .3132953   .0634243     4.94   0.000        .1867    .4398907
------------------------------------------------------------------------------------------------
```

Notice that the explanatory power only decreased slightly when eliminating the Greenspan Put. This seems to suggest that both the Greenspan Put and past annualized 10-year inflation rate may act as proxies for periods where monetary policy is effective. Since the Greenspan Put is marked at a relatively arbitrary date (1987), we drop the Greenspan Put since the past 10 years of inflation annualized is likely a better proxy.

Our specification for the aggregate investor allocation regressed on the inflation rate, U-3, the 10-year treasury (GS10), and Expected Return (Koyck Lag λ=1%) is as follows:

$$\boldsymbol{PercentInStocks = \alpha + \beta_1 * InflationRate + \beta_2 * CivUnemployment + \beta_3 * LongInterestRate + \beta_4 * ERKoyck01 + \varepsilon}$$

```
. reg PercentInStocks InflationRate CivUnemployment LongInterestRate ERKoyck01
```

Source	SS	df	MS			
				Number of obs	=	72
				F(4, 67)	=	31.74
Model	.228123093	4	.057030773	Prob > F	=	0.0000
Residual	.12038385	67	.001796774	R-squared	=	0.6546
				Adj R-squared	=	0.6340
Total	.348506943	71	.004908548	Root MSE	=	.04239

PercentInStocks	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
InflationRate	.2754753	.2269178	1.21	0.229	-.1774545	.7284051
CivUnemployment	-1.173669	.3527861	-3.33	0.001	-1.877833	-.4695052
LongInterestRate	-1.003877	.2451246	-4.10	0.000	-1.493148	-.5146064
ERKoyck01	9.554196	1.256374	7.60	0.000	7.046462	12.06193
_cons	-.3635637	.1122314	-3.24	0.002	-.5875785	-.1395488

Our specification for the aggregate investor allocation regressed on the inflation rate, U-3, the 10-year treasury (GS10), and Expected Return (Koyck Lag λ=3%) is as follows:

$$\boldsymbol{PercentInStocks = \alpha + \beta_1 * InflationRate + \beta_2 * CivUnemployment + \beta_3 * LongInterestRate + \beta_4 * ERKoyck03 + \varepsilon}$$

```
. reg PercentInStocks InflationRate CivUnemployment LongInterestRate ERKoyck03
```

Source	SS	df	MS			
				Number of obs	=	72
				F(4, 67)	=	30.40
Model	.224702692	4	.056175673	Prob > F	=	0.0000
Residual	.123804251	67	.001847825	R-squared	=	0.6448
				Adj R-squared	=	0.6235
Total	.348506943	71	.004908548	Root MSE	=	.04299

PercentInStocks	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
InflationRate	.441937	.240837	1.84	0.071	-.0387757	.9226497
CivUnemployment	-.5306304	.3793285	-1.40	0.166	-1.287773	.2265124
LongInterestRate	-1.35558	.2655323	-5.11	0.000	-1.885584	-.8255748
ERKoyck03	4.517626	.6126144	7.37	0.000	3.294843	5.74041
_cons	.011349	.066029	0.17	0.864	-.1204453	.1431434

Interpretations of Results

a. Investor/Expert Conflict of Interest – Division of funds by asset class, Long term nature of the indicators, and Career Risk/Lack of a "Permanent Capital Base"

Since most funds tie investment manager salary and bonuses to assets under management (AUM), and they are generally restrained to a particular asset class, managers may advocate to keep assets within the asset classes that they manage in order to maximize short term compensation, despite lower expected returns due to overbidding of financial assets. While the

Even large asset managers which have AUM across the whole spectrum of asset types may face restrictions when optimally allocating capital. The client-facing side of the business, wealth management, is often disjoint from fund managers. Since all clients have different needs based primarily on household income/wealth, and age, wealth managers will often advocate for a certain type of portfolio mix based off of long-run historical return, risk/volatility metrics, and individualized needs, despite evidence suggesting that asset returns may be relatively predictable in the medium term (10-15 years). **Since equities tend to outperform all other asset classes over long periods of time, there is a tendency for a large percentage of financial capital to cluster in equities because of a systemic failure of wealth managers to rebalance client portfolios.**

b. Portfolio Rebalancing

If the evidence seems to indicate that prices are getting to be too high, why don't people decide to sell? Why don't people periodically rebalance toward fixed portfolio weights? The evidence seems to suggest that wealth managers and their clients just "let it ride", and since equity returns tend to be greater than most other assets the equity weight steadily climbs up. There may be a belief among many people that equities always outperform in the long-run, but then you would generally expect the equity allocation to be much higher if this were true.

We think that the people fail to periodically rebalance portfolios because it is difficult to sell when prices are increasing or to buy when prices are falling. In *Manias, Panics, and Crashes*, Charles Kindleberger claims that "There is nothing as disturbing to one's well-being and judgment as to see a friend get rich."[36] When everyone around you is getting rich while you wait on the sidelines, it can be very isolating and difficult to stick with your strategy. Being comfortable with being a contrarian is one of the most useful attributes that a fund manager can have (of course when it is backed up by well thought out research).

The world can be very cruel to people who say that the party's over – which is effectively what you are doing when you sell out positions; company executives absolutely hate short sellers because there is a perception that they want the company to fail. More often times, short sellers believe that investor expectations are too high, even in scenarios where managers meet all of their own goals, and so they take short positions.

---

[36] Kindleberger, Charles P., and Aliber, Robert Z., *Manias, Panics, and Crashes: A History of Financial Crises*, 6th edition, 2011, p. 30.

Imagine then how people react when an investor says that they aren't holding any equities whatsoever. If there is a perception that short sellers want individual companies to fail, then people may extrapolate this and suggest that getting out of the market or shorting it means that you want all businesses to fail. While contrarian investors are just trying to correct mispricing, everyone around them treats them with suspicion because of a lack of understanding about the reasons for investment/divestment.

In order to achieve materially different returns from the market you must differentiate your positions from that of the overall market. In the pursuit of superior risk-adjusted returns, an investor has to think that they are smarter than the overall market at pricing securities. When you take positions that significantly differentiate your portfolio from the rest of the market people may look at you odd. No one likes someone who is smarter than them, and you have to think that you are smarter than most to succeed and try differentiated strategies. People will interpret your decision to differentiate as an insult because they feel you saying that everyone else is wrong. Perhaps this is why many people have a specious attitude that connects a person's wealth with their supposed level of intelligence. When you have money at risk in the market, and there are rich people making the same decision as you, believing that they are smart is comforting because it allows you to rationalize making the same decision.

c. Chance Correlation

Do stock returns follow a random walk, or are the findings on the aggregate investor allocation to equities, Tobin's Q, TR CAPE, and CAPE compelling? A skeptical manager might reasonably think as follows: might this just be another case of finding relationships which have happened to hold in the past, such as that of the "Super Bowl Indicator," (which held up through 2008) but may not hold in the future? **Before urging clients to act on these indicators, wealth managers need compelling arguments as to why these relationships hold, need evidence that they will continue to hold after first being brought to attention**.

Wealth managers need accurate estimates of expected returns over the next ten years using these indicators. I can do my best to give my client full information, but what would I give as my best estimate if forced to give a single number and a confidence interval? Is our level of confidence high enough to rely on these regressions rather than the long-run average return?

Our regression on subsequent ten-year annualized returns utilizing the aggregate investor allocation to equities had a 95% confidence interval for our coefficient; This shows that each additional percent of total capital (shown in the denominator in the model section) in equities leads to reduction of subsequent annualized real return over ten years between -0.5% and -0.9%.

d. Limitations

Our paper is ultimately limited by practical difficulties in data analysis related to observing the theoretical market portfolio. While basic models such as CAPM utilize the equity market as a proxy for the diversified market portfolio, our study is also limited in that we do not observe all constituent elements of the market portfolio. While the CAPM and Black-Scholes model both utilize the United States Treasury Yield as a proxy for the risk-free rate which is ultimately tied to U.S. Treasury bond issuance, our theoretical model differs from others in that we consider non-equity, non-treasury bond/bill financial assets' value when evaluating subsequent return data. Evaluating equity valuations within a fully contextualized investment environment is conducive to finding more theoretically plausible relationships because asset allocation choices are generally made with regard to the entire spectrum of asset class opportunities. Excluding anyone of these from an analysis can lead to significant distortions, since financial assets are slightly imperfect substitutes – while various financial assets carry individual characteristics which make them particularly attractive to various investor classes, the ultimate pursuit of any coldly rational investment activity is to maximize expected returns with minimal expected risk.

V. Discussion & Conclusions

The Efficient Market Hypothesis presents a compelling way to understand securities markets, and the tremendous difficulty in outperforming the market on a risk-adjusted basis. However, while the EMH presents a straightforward way to understand how market prices incorporate new information, anomalies in models of asset pricing which do not carry risks present a hiccup in the rationalist theory that expected returns are developed to reward investors for taking on commensurate levels of risk. It is foolish to deny the recurrences of speculative mania driven by greed (or fear) and short-sightedness. While speculative bubbles do not necessarily occur often, they appear infrequently and may expand and deflate over long periods of time or in an explosive fashion. It is because of this tendency of capital markets, that it is of absolute necessity to research how financial assets prices are reached.

We use the aggregate investor allocation to equities as a proxy measure for overvaluation in securities markets (as well as Tobin's Q, Shiller TR CAPE, and Shiller CAPE). Between 1945 and 2008 (year-end), historically high values for the aggregate investor allocation to equities have been highly correlated with much lower real returns over subsequent 10-year time periods (1945-1955 through 2008-2018). High values for equity allocation to equities are also correlated with higher average losses over the subsequent 10-year periods. Following high allocation to equities, we find lower average returns and greater losses. All of the other indicators listed also show a statistically significant inverse relationship with subsequent real returns as well, though with varying levels of correlation and different sample sizes. The Efficient Market Hypothesis does not account for this relationship, and we feel that ultimately Hyman Minsky's Financial Instability Hypothesis presents a more realistic view of booms and busts in securities markets. While Eugene Fama argues that there is no effective way to test for speculative bubbles, we disagree; the aggregate investor allocation to equities acts as a highly significant indicator for subsequent returns, and in particular this indicator is significantly related to return and risk expectations that place an emphasis on recent experience. Our analysis of expectation development shows that investors mistakenly perceive a combination of higher returns and lower risk for future returns following periods of relative prosperity, and vice versa.

While we do not exhaustively go into Minsky's thoughts about the various types of debt financing that drive instability, this empirical evidence substantiates his contention that "stability is destabilizing". Periods of high returns are predictably followed by lower returns; this stands in contrast to the idea that stock prices exhibit Brownian motion as forwarded by efficient market proponents. While there is no rule which stipulates that stock prices must exhibit some form of mean reversion in pricing, we can make an argument for it using our analysis. Our analysis shows that a significant portion of variation in stock returns over subsequent 10-year periods can be explained by variation in the aggregate allocation to stocks.

It is important to note that while this metric provides a great deal of insight into future real returns, another significant component of real returns is the growth in real liabilities in the economy, which is of course not a real-time variable. It can only be observed ex-post, and so it is of no use to investors today. However, the theoretical implications of this other component are very useful for understanding the various factors at play that determine stock returns. Investors can guess how the aggregate allocation to equities may change over the next 10 years, and forecast real liability growth to develop expectations of return given various assumptions to better evaluate their current asset allocation strategies.

Various economists and financial economists have brought forward the proposition that investors may have improved return prospects when investing during recessionary periods because they are willing to accept cyclical risks. This is an interesting proposition that casts doubt on the idea that investors are acting irrationally, as we propose. When you're in a crumbling castle and you're unable to see the light out, should those who have faith be duly rewarded? This is truly the epitome of Warren Buffet's famous quote, "be fearful when others are greedy and greedy when others are fearful."

But what about economies that did crumble? The stock market during the Weimar Hyperinflation is just one example. In an attempt to avoid hyperinflation, investors sunk their money into the stock market and drove respectable dividend yields down to below 1%, and investors never recouped their money.

Investors who stayed invested certainly took on significant risks, and they were harmed when the system fell apart. When setting expectations of the future, nothing is truly certain; this conundrum is difficult to assess. **It is important to note however, that equity prices did not suddenly jump back up to historical values following events like the end of WWII.**

When looking back in economic history, there have clearly been instances of speculative bubbles fueled by irrational behavior (whether driving asset prices up or down significantly); the Dutch tulip bubble, the 1920s Florida Housing Boom, the 1929 stock market bubble, the U.S. Market Crash of 1987, the rise and fall of Japanese stocks and real estate in the 1990s, the dot-com bubble, & the ownership society bubble are some examples of positive bubbles. Negative bubbles include the slump in stock prices during WWII, and the market bottom in 1982 during the Volcker Disinflation. While our proxy is a valuable indicator for equity overvaluation and undervaluation, it is because of these types of out of sample observations that we cannot conclude that it is the only metric that ought to be considered when evaluating whether or not financial assets deviate significantly from intrinsic value.

Bibliography

Ball, Ray. 1978. “Anomalies in Relationships Between Securities’ Yields and Yield-Surrogates.” Journal of Financial Economics. 6:2, pp. 103-126.

Black, Fischer, and Myron Scholes. “The Pricing of Options and Corporate Liabilities.” *Journal of Political Economy*, vol. 81, no. 3, 1973, pp. 637–654. *JSTOR*, www.jstor.org/stable/1831029.

Black, Fischer, Capital Market Equilibrium with Restricted Borrowing, The Journal of Business, 45, issue 3, 1972, p. 444-55.

Fama, Eugene F. “Efficient Capital Markets: A Review of Theory and Empirical Work.” *The Journal of Finance*, vol. 25, no. 2, 1970, pp. 383–417. *JSTOR*, www.jstor.org/stable/2325486.

Fama, Eugene F. and French, Kenneth R., The Capital Asset Pricing Model: Theory and Evidence (August 2003). CRSP Working Paper No. 550; Tuck Business School Working Paper No. 03-26. Available at SSRN: https://ssrn.com/abstract=440920 or http://dx.doi.org/10.2139/ssrn.440920

Fama, Eugene F., and Kenneth R. French. “The Cross-Section of Expected Stock Returns.” *The Journal of Finance*, vol. 47, no. 2, 1992, pp. 427–465. *JSTOR*, www.jstor.org/stable/2329112.

James Tobin and W.C. Brainard, Asset Markets and the Cost of Capital." 1977, Economic Progress, Private Values and Public Policy

Kindleberger, Charles P., and Aliber, Robert Z., *Manias, Panics, and Crashes: A History of Financial Crises*, 6th edition, 2011, p. 30.

Lakonishok, J. , Shleifer, A. and Vishny, R. W. (1994), Contrarian Investment, Extrapolation, and Risk. The Journal of Finance, 49: 1541-1578. doi:10.1111/j.1540-6261.1994.tb04772.x

Lewis, Michael M. The Big Short : [inside the Doomsday Machine]. New York :Simon & Schuster, 2010.

Lintner, John. “The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets.” *The Review of Economics and Statistics*, vol. 47, no. 1, 1965, pp. 13–37. *JSTOR*, www.jstor.org/stable/1924119.

Mandelbrot, Benoit; Hudson, Richard L. (2007-03-22). The Misbehavior of Markets: A Fractal View of Financial Turbulence (pp. 11-13).

Milton Friedman "The Methodology of Positive Economics" Essays In Positive Economics (Chicago: Univ. of Chicago Press, 1966), pp. 3-16, 30-43.

Minsky, Hyman P., The Financial Instability Hypothesis (May 1992). The Jerome Levy Economics Institute Working Paper No. 74. Available at SSRN: https://ssrn.com/abstract=161024 or http://dx.doi.org/10.2139/ssrn.161024

Robert Shiller’s data webpage: http://www.econ.yale.edu/~shiller/data.htm

Sharpe, William F. “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk.” *The Journal of Finance*, vol. 19, no. 3, 1964, pp. 425–442. *JSTOR*, www.jstor.org/stable/2977928.

Sherman, John, and Harold Petersen. “The Efficient Market Hypothesis, the Financial Instability Hypothesis, and Speculative Bubbles,” pg.29, 2014. https://dlib.bc.edu/islandora/object/bc-ir%3A102322

Sommer, Jeff. "Eugene Fama: King of Predictable Markets". *The New York Times* https://www.nytimes.com/2013/10/27/business/eugene-fama-king-of-predictable-markets.html

Appendix

Summary Statistics

Aggregate Investor Allocation to Stocks (%) (end of year/end of 4th quarter) ((N

	Percentiles	Smallest		
1%	.2352095	.2352095		
5%	.242804	.2418698		
10%	.2513441	.2420979	Obs	74
25%	.2874888	.242804	Sum of Wgt.	74
50%	.3504783		Mean	.3467576
		Largest	Std. Dev.	.0694842
75%	.4072572	.452534		
90%	.4284154	.468883	Variance	.0048281
95%	.452534	.4762677	Skewness	.0985146
99%	.5112216	.5112216	Kurtosis	1.967084

Tobin's Q

	Percentiles	Smallest		
1%	.3114934	.3114934		
5%	.3374382	.3200855		
10%	.3535823	.3280141	Obs	74
25%	.4276936	.3374382	Sum of Wgt.	74
50%	.721739		Mean	.7145651
		Largest	Std. Dev.	.2964451
75%	.9086709	1.182222		
90%	1.085685	1.196638	Variance	.0878797
95%	1.182222	1.332919	Skewness	.4527678
99%	1.594446	1.594446	Kurtosis	2.506084

Cyclically Adjusted Price to Earnings Ratio P/E10
or TR CAPE (Corrected for Chan

	Percentiles	Smallest		
1%	8.093561	7.060923		
5%	10.54989	8.093561		
10%	11.69074	8.124581	Obs	139
25%	15.42706	8.595793	Sum of Wgt.	139
50%	20.3122		Mean	20.52893
		Largest	Std. Dev.	7.159536
75%	24.19377	36.82863		
90%	28.70736	39.71281	Variance	51.25895
95%	33.59094	44.81105	Skewness	.8331475
99%	44.81105	47.58769	Kurtosis	4.442043

Cyclically Adjusted Price to Earnings Ratio P/E10
or CAPE

	Percentiles	Smallest		
1%	5.989668	5.122184		
5%	8.072249	5.989668		
10%	9.257637	6.098468	Obs	139
25%	11.89576	6.287087	Sum of Wgt.	139
50%	16.37848		Mean	17.03341
		Largest	Std. Dev.	6.864903
75%	20.97858	33.30734		
90%	26.4923	36.97887	Variance	47.1269
95%	28.33287	40.57696	Skewness	1.032358
99%	40.57696	43.77258	Kurtosis	4.715962

Long Interest Rate (10 year from Shiller)

	Percentiles	Smallest		
1%	.0191	.0188		
5%	.0225	.0191		
10%	.0246	.0195	Obs	148
25%	.03235	.0197	Sum of Wgt.	148
50%	.03835		Mean	.045223
		Largest	Std. Dev.	.022569
75%	.0508	.1138		
90%	.0778	.1167	Variance	.0005094
95%	.091	.1257	Skewness	1.830866
99%	.1257	.1459	Kurtosis	6.762829

Civilian Unemployment Rate

	Percentiles	Smallest		
1%	.027	.027		
5%	.035	.031		
10%	.039	.034	Obs	72
25%	.0445	.035	Sum of Wgt.	72
50%	.055		Mean	.0573889
		Largest	Std. Dev.	.0163065
75%	.066	.085		
90%	.079	.093	Variance	.0002659
95%	.085	.099	Skewness	.7234764
99%	.108	.108	Kurtosis	3.450264

Expected Return from Koyck Lag= .01 (8.5% is
Arithmetic Average from 1871-2018)

	Percentiles	Smallest		
1%	.0719871	.0718248		
5%	.0745813	.0719871		
10%	.0757587	.073025	Obs	148
25%	.0782796	.0735019	Sum of Wgt.	148
50%	.0822452		Mean	.0824918
		Largest	Std. Dev.	.0053609
75%	.0862773	.0943176		
90%	.0897534	.0948932	Variance	.0000287
95%	.0910984	.0958694	Skewness	.2900821
99%	.0958694	.0967806	Kurtosis	2.510309

Expected Return from Koyck Lag= .03 (8.5% is
Arithmetic Average from 1871-2018)

	Percentiles	Smallest		
1%	.0581301	.055517		
5%	.0635853	.0581301		
10%	.0656804	.0601475	Obs	148
25%	.0722549	.0616596	Sum of Wgt.	148
50%	.0819737		Mean	.0833519
		Largest	Std. Dev.	.0138946
75%	.0948249	.1146367		
90%	.1005204	.1149498	Variance	.0001931
95%	.104204	.1189862	Skewness	.3073426
99%	.1189862	.1210264	Kurtosis	2.410419

Real Expected Earnings Growth (Koyck=.01) (Real EEG)

	Percentiles	Smallest		
1%	.0040205	.0015603		
5%	.0108492	.0040205		
10%	.0118304	.0076187	Obs	147
25%	.015152	.008053	Sum of Wgt.	147
50%	.0172292		Mean	.016565
		Largest	Std. Dev.	.0033235
75%	.0187205	.0213606		
90%	.0200368	.0214346	Variance	.000011
95%	.0207238	.0216858	Skewness	-1.317436
99%	.0216858	.023951	Kurtosis	6.145224

Real Expected Earnings Growth (Koyck=.03) (Real EEG)

	Percentiles	Smallest		
1%	-.0241643	-.0282741		
5%	-.0004808	-.0241643		
10%	.0022751	-.0111058	Obs	147
25%	.0107836	-.0092411	Sum of Wgt.	147
50%	.0181595		Mean	.0159888
		Largest	Std. Dev.	.009739
75%	.0224722	.0291684		
90%	.0260095	.0297751	Variance	.0000948
95%	.0280285	.0317751	Skewness	-1.483782
99%	.0317751	.0351789	Kurtosis	6.773432

Nominal Expected Earnings Growth (Koyck=.01) (Nominal EEG)

	Percentiles	Smallest		
1%	.0201464	.019602		
5%	.0242845	.0201464		
10%	.0262004	.0204148	Obs	147
25%	.030286	.0221144	Sum of Wgt.	147
50%	.0333604		Mean	.0340697
		Largest	Std. Dev.	.0061456
75%	.0384232	.0455166		
90%	.043483	.0457904	Variance	.0000378
95%	.0447992	.0458888	Skewness	.0585381
99%	.0458888	.0468113	Kurtosis	2.541132

Real Expected Earnings Growth (Koyck=.01) (Real EEG)

	Percentiles	Smallest		
1%	.0040205	.0015603		
5%	.0108492	.0040205		
10%	.0118304	.0076187	Obs	147
25%	.015152	.008053	Sum of Wgt.	147
50%	.0172292		Mean	.016565
		Largest	Std. Dev.	.0033235
75%	.0187205	.0213606		
90%	.0200368	.0214346	Variance	.000011
95%	.0207238	.0216858	Skewness	-1.317436
99%	.0216858	.023951	Kurtosis	6.145224

Real Expected Earnings Growth (Koyck=.03) (Real EEG)

	Percentiles	Smallest		
1%	-.0241643	-.0282741		
5%	-.0004808	-.0241643		
10%	.0022751	-.0111058	Obs	147
25%	.0107836	-.0092411	Sum of Wgt.	147
50%	.0181595		Mean	.0159888
		Largest	Std. Dev.	.009739
75%	.0224722	.0291684		
90%	.0260095	.0297751	Variance	.0000948
95%	.0280285	.0317751	Skewness	-1.483782
99%	.0317751	.0351789	Kurtosis	6.773432

Past 10 Years Inflation Annualized

	Percentiles	Smallest		
1%	-.0323669	-.0324502		
5%	-.0268273	-.0323669		
10%	-.0229136	-.0309233	Obs	138
25%	.0092444	-.0300764	Sum of Wgt.	138
50%	.0244494		Mean	.0230301
		Largest	Std. Dev.	.0299409
75%	.0419401	.0813434		
90%	.0631555	.0813812	Variance	.0008965
95%	.0731032	.0865932	Skewness	-.1080774
99%	.0865932	.0866581	Kurtosis	2.384363

Average Annual Loss over next 10 Years (Including Only Years with Losses)

	Percentiles	Smallest		
1%	-.1297393	-.1317376		
5%	-.0913641	-.1297393		
10%	-.08144	-.1290485	Obs	138
25%	-.0529176	-.1290485	Sum of Wgt.	138
50%	-.0351887		Mean	-.0390613
		Largest	Std. Dev.	.0286508
75%	-.017971	-.0028959		
90%	-.0074124	-.0028959	Variance	.0008209
95%	-.0028959	-.0028959	Skewness	-1.145708
99%	-.0028959	-.0028959	Kurtosis	4.407303

Ten Year Annualized Return (Real)

	Percentiles	Smallest		
1%	-.0385598	-.0388744		
5%	-.0237871	-.0385598		
10%	-.0092299	-.0380158	Obs	138
25%	.0366089	-.0338463	Sum of Wgt.	138
50%	.0651877		Mean	.0667538
		Largest	Std. Dev.	.0505527
75%	.1035609	.1627147		
90%	.1389841	.1641179	Variance	.0025556
95%	.1495471	.1724507	Skewness	-.1078399
99%	.1724507	.1764039	Kurtosis	2.52658

Ten Year Annualized Return (Nominal)

	Percentiles	Smallest		
1%	-.009305	-.0135801		
5%	.0137493	-.009305		
10%	.031201	-.0063588	Obs	138
25%	.0577243	.0011194	Sum of Wgt.	138
50%	.0833229		Mean	.0907444
		Largest	Std. Dev.	.0495387
75%	.1314292	.1806133		
90%	.1633475	.1886476	Variance	.0024541
95%	.1780842	.1897052	Skewness	.2191058
99%	.1897052	.1948497	Kurtosis	2.257802