

A Methodology for Measuring Market Inefficiency, and Its Relationship to the Value Spread *

Aman Rana

24th November, 2024

Abstract

Extending the work of Boguth et al. (2023), we propose a method to measure the level of market inefficiency using unbiasedness regressions. We find that the market is not uniformly efficient and goes through periods of higher inefficiency reverting to efficiency in-line with Lo (2004)'s adaptive market hypothesis. We find that the value spread—the ratio of the average price-to-book of expensive stocks to that of cheap stocks—is cointegrated with market inefficiency. By demonstrating that wider value spreads align with greater market inefficiency, we can empirically support Asness (2024)'s intuition that high spreads during events such as the dot-com bubble and Covid-19 could reflect mispricing rather than fundamental changes in the notion of value. This evidence strengthens the case that the value premium is still alive but has been masked by market inefficiency. We also find that the market's temporary inefficiencies have grown more persistent over time. This has implications for risk-premia based strategies and portfolio management.

*I thank Professor Charles Martineau for his comments. Code and data supporting this analysis is available at: https://github.com/Aman-Rana-02/Irrational_Value_Spreads

Contents

1	Introduction	3
2	Data	5
2.1	Cumulative Log Returns	5
2.2	The Value Spread	6
3	Measurement of Market Efficiency	6
3.1	Unbiasedness Regressions	6
3.2	<i>Beta_SSE</i> Score	8
4	Results	10
5	Discussion	11
A	Appendix	13
A.1	Unbiasedness Regressions	13
A.1.1	Simulated Efficient Market	13
A.1.2	Simulated Inefficient Market	14
A.1.3	R^2 and Intuition Building	14

1 Introduction

Fama (1970) defines markets to be informationally efficient if prices at every moment incorporate all available information about future values rationally and instantaneously. This is an intuitive result if we consider that market participants bid prices up once presented with information that an asset is underpriced, and sell when new information shows an asset as overpriced. When market participants are digesting information, they do so with consideration of a firm's current and future states. We would expect that due to competition, firms get more sophisticated in their forecast and procurement of information, and market gets relatively more efficient over time. Market efficiency is important because in an efficient market investors are compensated for taking systematic risk. The expected returns of risky assets must be higher on average because investors require a risk premium to bear the additional uncertainty associated with risky assets.

'Value' in factor investing is an example of one such systematic risk. Its classical definition is the book-to-market ratio; the book value of equity of a firm divided by its market capitalization. According to Fama and French (1993), firms with high book-to-market ratios (value stocks) tend to outperform those with low book-to-market ratios (growth stocks) over time, as this factor captures a systematic risk associated with the financial distress or other underlying characteristics of value firms. Value is a relative measure, there is no threshold that defines a stock as 'cheap' or 'expensive'.

Asness (2024) defines the value spread as the ratio of the average book-to-market of a portfolio of 'expensive' stocks to the average book-to-market of a portfolio of 'cheap' stocks. The value spread can be interpreted as how much more the market is willing to pay for already 'expensive' stocks than 'cheap' stocks. He uses this measure of value spread in a timeseries, showing that the spread has been increasing steadily for the last decade and proposes that this widening is due to market inefficiency. His claim is that the widening spread has no fundamental backing, the market is irrationally paying more for 'expensive' stocks, and therefore it takes longer for market participants to be appropriately rewarded for holding value risk. His justification lies in his market experience, and that papers that try to justify that the relative definition of 'value' has fundamentally changed, are inconclusive.

In this paper I use an inverted approach to investigate the value spread. I propose a measure for market informational efficiency, and provide a time series of this measure. Asness justifies his

conclusion by contradiction, he presents possible explanations for the value spread, and that they don't work. I am investigating the same phenomena more directly, by constructing a measure of market efficiency and showing its movement with the value spread.

Existing methods for measuring market efficiency typically focus on testing whether stock prices follow a random walk (Fama, 1965) (Lim and Brooks, 2010). However, these methods do not represent the variation of efficiency over time. A more dynamic approach is needed to investigate a timeseries of inefficiency.

Boguth et al. (2023) uses unbiasedness regressions to examine price informativeness around FOMC meetings. This method visually shows us how prices adjust to information flow, in speed, and under or overreaction, using R^2 s and β s over the unbiasedness regression window. Unbiasedness regressions provide us with a tool to measure the market's digestion of information, from which I can construct a measure of market efficiency.

Our estimand is Market Inefficiency, which we define as the relative level of the market's ability to price in future information. Building on this, we introduce the *Beta_SSE* estimate, which measures market efficiency at a point in time. By applying unbiasedness regressions within a rolling window, and constructing the *Beta_SSE* representation, we create a time series that tracks the relative level of market efficiency.

Our findings show that:

- The market is not uniformly efficient. Instead, it experiences cycles of inefficiency followed by reversion to efficiency.
- Over the last decade, inefficiency has been increasing, suggesting that the market has become less adept at incorporating available information into prices on longer time-scales.
- This rising inefficiency is correlated with the widening value spread, indicating a link between the weakness of the value premium and market inefficiency.

The rest of the paper is structured as follows: Section 2 describes the data used in this analysis. Section 3 outlines the methodology used to measure market inefficiency. Section 4 presents the results of our analysis. Section 5 concludes and discusses the implications of our findings. Appendix A provides additional details on the methodology, simulations to show measure robustness, and illustrative examples of unbiasedness regression results to build intuition.

2 Data

We use Python (Van Rossum and Drake, 2009) to analyze S&P500 price data from Yahoo Finance (Finance, 2024) and the Fama-French 3-factor data library (French, 2024). Yahoo finance provides daily price data of the S&P500 which we resample to monthly returns for the period 1950 - 2024. We get monthly log returns by taking the difference of the log of the adjusted close price of the S&P500 for each month (Figure 1). Our timeseries starts in 1950 based on the premise that the market prior to this was not as sophisticated prior to this date, and is not illustrative for the purposes of this analysis of market efficiency and value spreads (Asness, 2024).

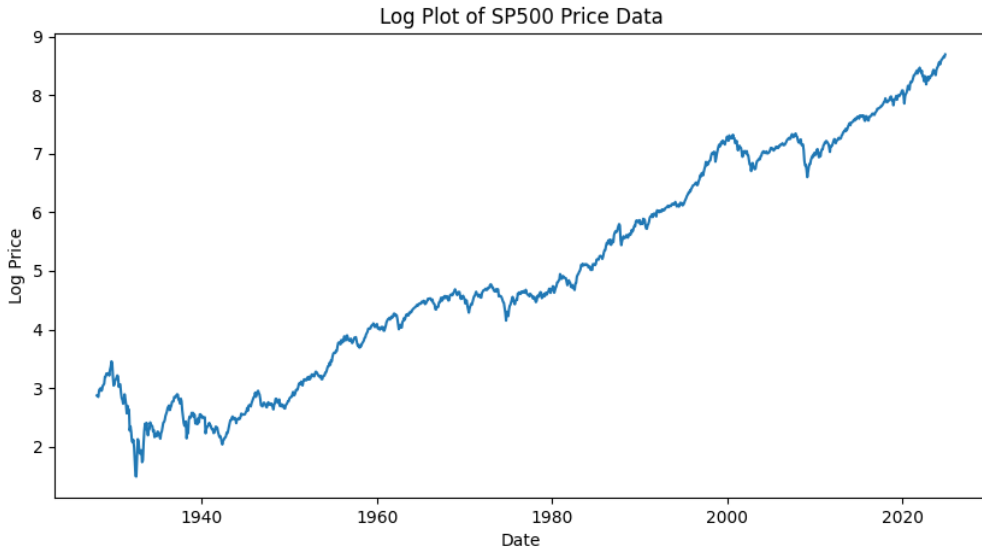


Figure 1: The S&P500 timeseries 1950 - 2024.

2.1 Cumulative Log Returns

Date	Cum Log Return 1	Cum Log Return 2	Cum Log Return 3	...	Cum Log Return 34	Cum Log Return 35	Cum Log Return 36
1950-01-01	0.0154	0.0253	0.0293	...	0.379	0.424	0.459
1950-02-01	0.0100	0.0140	0.0520	...	0.409	0.444	0.436
1950-03-01	0.00412	0.0421	0.0867	...	0.434	0.427	0.408
1950-04-01	0.0380	0.0827	0.0229	...	0.422	0.404	0.380
1950-05-01	0.0446	-0.0151	-0.00672	...	0.366	0.342	0.315

Table 1: Cumulative Log S&P500 Return periods for each period. Each column represents the cumulative log return for a given period starting at the row index, and ending t months ahead.

The unbiasedness regression require windows of cumulative returns. For any month t , we get every forward month's returns up to T months. By taking expanding window sums of the columns of log returns, we get a cumulative log returns matrix (Table 1).

2.2 The Value Spread

The value spread is the ratio of average book-to-market of the most expensive 30% portfolio to the price-to-book of the portfolio of the cheapest 30% of stock portfolio, as per Fama and French (1993). To construct this measure we use Kenneth French’s data library (French, 2024). Using French’s 3x2 sort on book-to-market and market equity, we get the monthly market value weighted average of the book-to-value of the portfolio of the 30% most expensive large-cap stocks and the portfolio of the 30% cheapest large-cap stocks. The value spread is the ratio of these two averages. It represents how much more expensive the expensive stocks are compared to the cheap stocks (Figure 2).

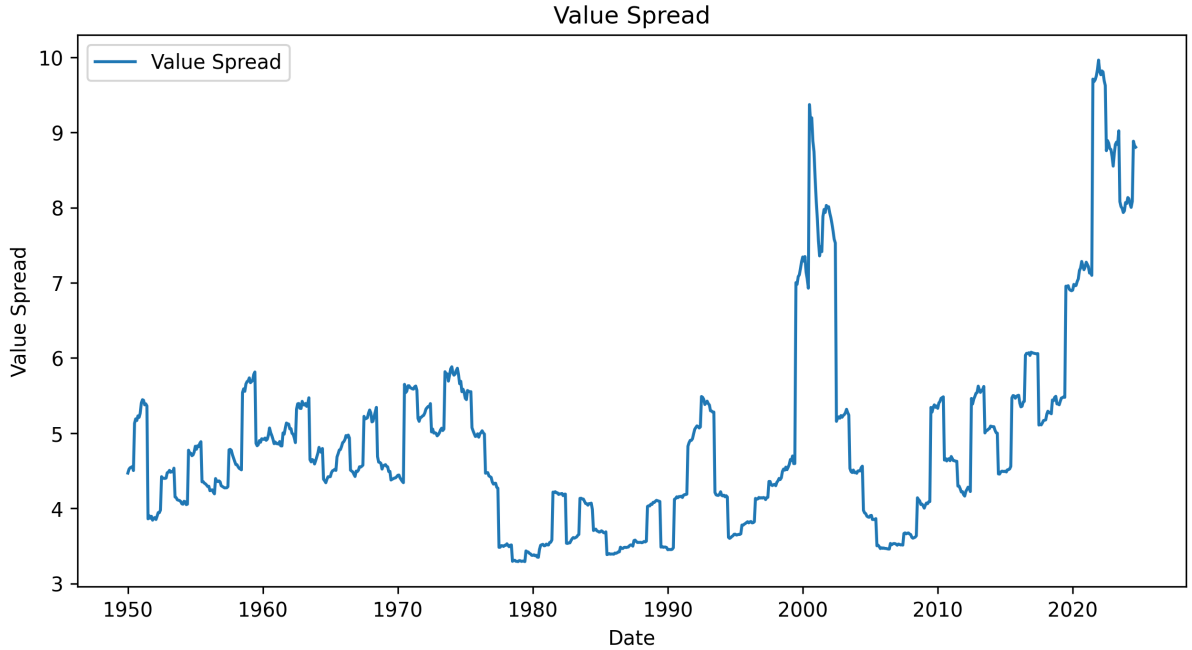


Figure 2: The value spread timeseries 1950 - 2024. Asness (2024) argues the widening spread in the last decade is due to increasing market inefficiency.

3 Measurement of Market Efficiency

In this section, we will introduce the $Beta_SSE$ estimate, which is our proposed measure of market inefficiency based on the unbiasedness regressions framework.

3.1 Unbiasedness Regressions

The unbiasedness regression is a simple OLS regression of the cumulative logarithmic return of the asset over a normalized window of time $SP500_{[0,T]}$, on a subset of the logarithmic cumulative

returns, $SP500_{[0,t]}$, where $t \leq T$, and t is the number of months from the beginning of the window.

1

$$SP500_{[0,T]} = \alpha_t + \beta_t SP500_{[0,t]} + \epsilon_t, \text{ where } 0 \leq t \leq T. \quad (1)$$

For an illustrative example, let $T = 36$, so that the window $[0, T]$ represents a 3-year period, and let us look at the returns of the SP500.

For $t = 1$, the regression is of the form:

$$SP500_{[0,36]} = \alpha + \beta SP500_{[0,1]} + \epsilon \quad (2)$$

From this regression we extract the coefficient β and the R^2 value. The regression is repeated for $t = 2, t = 3, \dots, t = 36$, and the coefficients β are plotted against t , the depth of the cumulative return window. We plot the β coefficients against t to get a sense of the market's ability to process information over time (Figure 3).

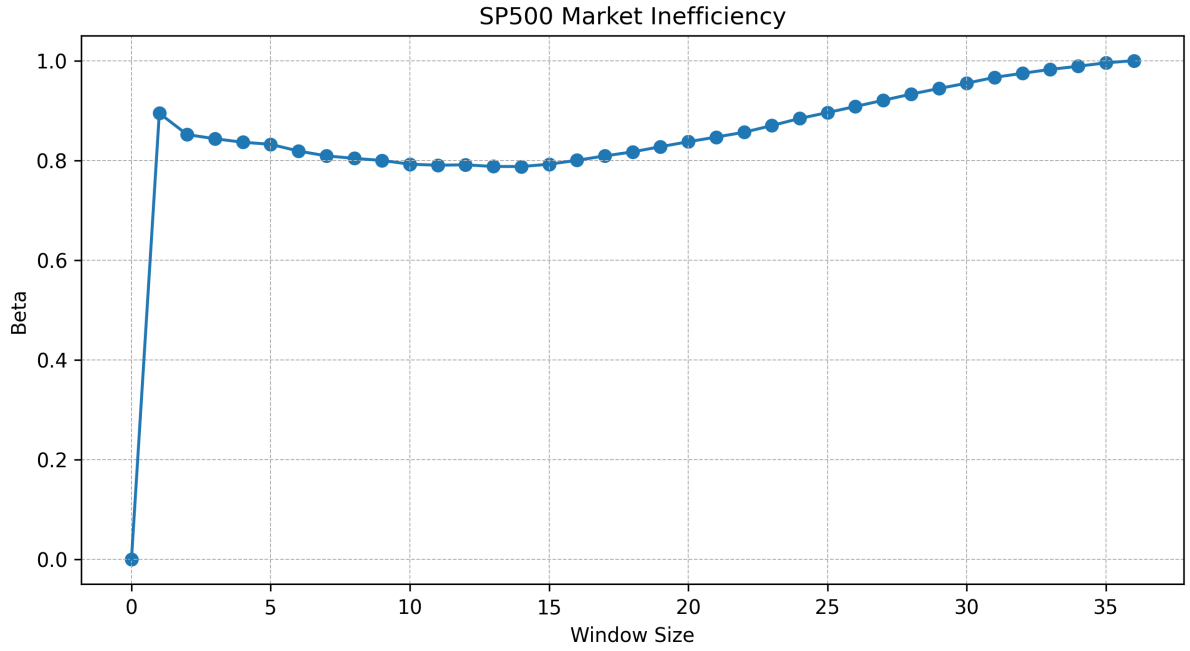


Figure 3: An unbiasedness regression on the SP500 (1950 - 2024) showing the market's tendency to overreact.

$\beta_t = 1$ for all t when log prices move efficiently, following a random walk. When $\beta_t = 1$ the

¹SP500_[0,1] is the matrix of returns of the SP500 from January 1950 to February 1950, February 1950 to March 1950, ..., August 2024 to September 2024, September 2024 to October 2024.

SP500_[0,2] would be the matrix of cumulative returns from January 1950 to March 1950, February 1950 to April 1950, ..., August 2024 to October 2024, and so on for the entire dataset.

partial returns from 0 to t provide a forecast of the return from t to T (Mincer and Zarnowitz, 1969) that is as efficient as possible and does not need to be amplified or dampened. If $\beta_t < 1$, the partial return is dampened in the total returns, therefore there was a temporary component in prices which decays representing an “overreaction” (Barclay and Hendershott, 2003). Conversely, if $\beta_t > 1$, the partial return is amplified in the total return, suggesting the market had improperly processed the information relevant in pricing the T sized window, and has underreacted.

In this paper we will be focusing on the β_t coefficient, but an explanation of the R^2 is provided in the Appendix for completeness and intuition building. We now have the foundation to construct a score for market efficiency.

3.2 *Beta_SSE* Score

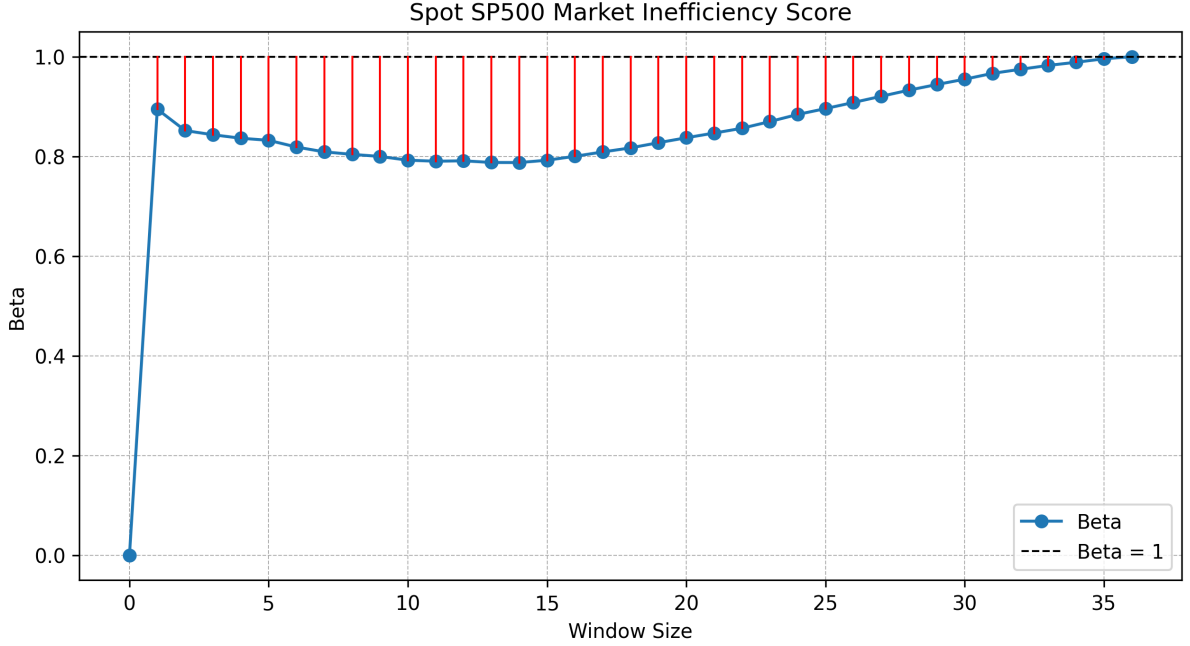


Figure 4: The *Beta_SSE* score visualization for the SP500 (1950 - 2024). The *Beta_SSE* score is a measure of market efficiency at a point in time, and is the sum of the squares of the red distances. To construct a timeseries of the *Beta_SSE* score, we run this regression for shingled five-year windows and plot their sums.

We have established the unbiasedness regression infrastructure, and that a $\beta_t < 1$ indicates overreaction, a $\beta_t > 1$ indicates underreaction, and $\beta_t = 1$ indicates efficiency. To construct a score that measures the level of market inefficiency at a point in time, we have to create a spot representation of the β_t graph for a time period.

We define the *Beta_SSE* score as the sum of the squared differences between β_t and the

horizontal line at $\beta = 1$, for all t in the window. Since a $\beta_t = 1$ indicates efficiency, the $Beta_SSE$ score will be lower when the market is more efficient and larger when the market is less efficient, whether by under- or over-reaction. We provide a visual representation of the $Beta_SSE$ score for the S&P500 over our entire data window in Figure 4.

To construct a time series of these scores, we split our entire period of S&P500 returns in shingled 5-year windows. Each window is 5 years long, and the window is moved forward by 1 month at a time. We run the same sequence of unbiasedness regressions as we did for the entire period in Figure 3, however on each window independently, and calculate the $Beta_SSE$ score for each window. So $Beta_SSE_t$ is constructed from the unbiasedness regressions on windows of the SP500 returns from $t - 60$ to t . Note that our unbiasedness regressions extend out to $t + 36$, so $Beta_SSE_t$ score isn't tradable at time t but at time $t + 36$.² This is by design, a measure of market inefficiency at a time t is to be representative of the market's ability to price the T months ahead.

$Beta_SSE_t$ is calculated as follows:

$$Beta_SSE_t = \sum_{i=0}^{36} (\beta_{t,i} - 1)^2 \quad (3)$$

Where $\beta_{t,i}$ is the β coefficient from the unbiasedness regression from a window starting on the date t , at the i th depth of the unbiasedness window.

²Consider the rolling window [January 2000, January 2005], the observation starting at January 2005 ($t = 0$ on January 2005) requires the window of returns from January 2005 to January 2008 to make $SP500_{[0,36]}$.

4 Results

Using the $Beta_SSE$ score, we construct a time series of the level of market efficiency as seen in Figure 5.

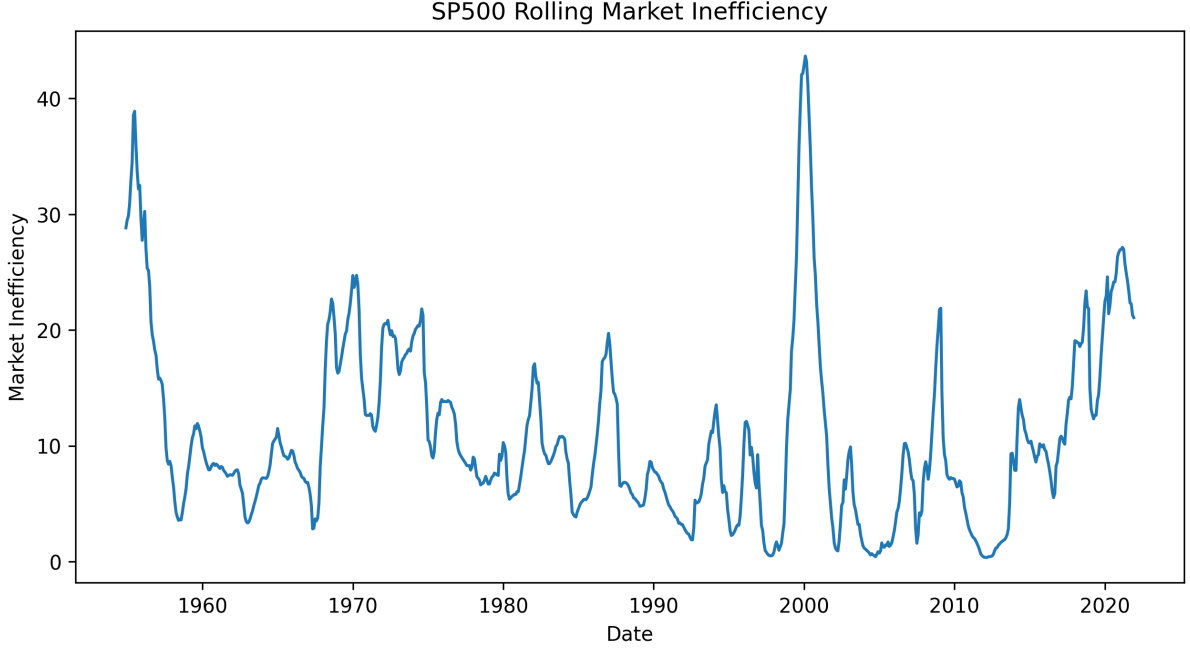


Figure 5: The $Beta_SSE$ score time series of the SP500 (1950 - 2024) showing mean reversion and periods of inefficiency.

We see the market is inefficient around the dot-com bubble, the 2008 financial crisis, and the COVID-19 pandemic, which is in line with our expectations. We can visually assess that the market is not uniformly efficient. Instead, it goes through periods of higher inefficiency followed by reversion to efficiency. We also see that the market has been increasingly inefficient in the last decade, which is in line with Asness (2024).

Empirically we can test the mean reversion through an Augmented Dickey-Fuller test (Cheung and Lai, 1995) on the $Beta_SSE$ time series, which is significant at the 0.1% level.

Figure 6 shows the value spread and the $Beta_SSE$ score time series. We see visually that the two are cointegrated, which we test with the Engle-Granger two-step cointegration test, which is significant at the 0.1% level.

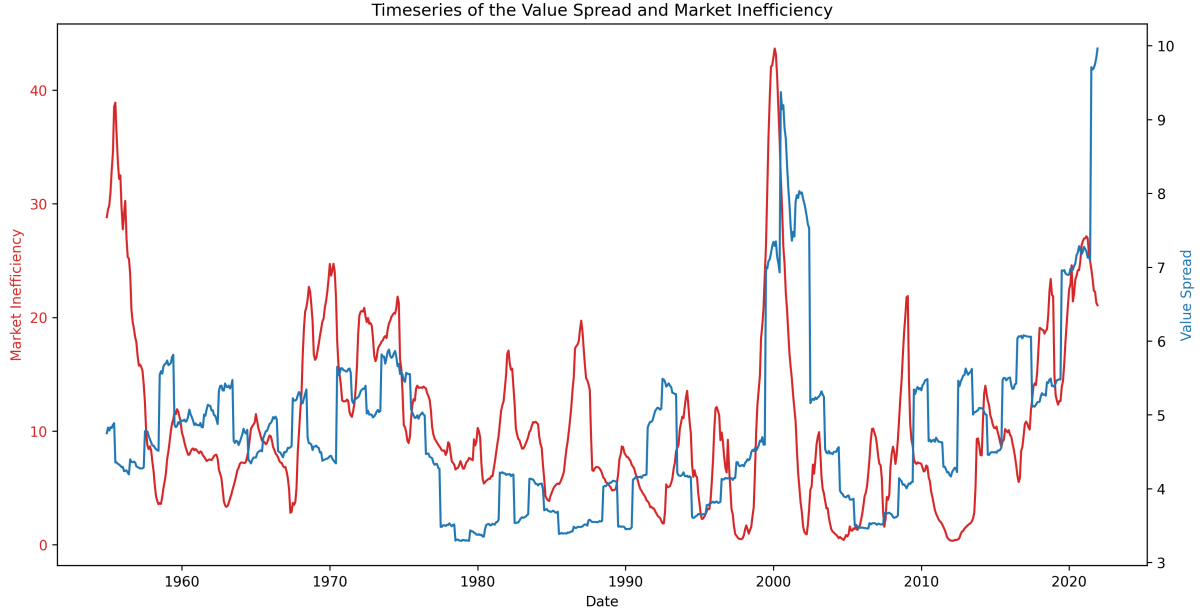


Figure 6: The value spread, and market inefficiency move together, and have been increasing in tandem for the last decade.

5 Discussion

In this paper, we introduced the $Beta_SSE$ score, a measure of market inefficiency based on the unbiasedness regressions framework. The score follows patterns of well established market inefficiencies, such as the dot-com bubble, the 2008 financial crisis, and the COVID-19 pandemic. We find that the market is not uniformly efficient, but instead goes through cycles of inefficiency followed by reversion to efficiency. Over the last decade, inefficiency has been increasing, suggesting that the market has become less adept at using available information for present value discounts for future forecasts. This rising inefficiency is correlated with the widening value spread, indicating a strong link between the weakness of the value premium and market inefficiency. Causality between the two is not established, and can be investigated in a future work. However, this paper compliments the intuition presented in Asness (2024) that the value premium has been weakening over the last decade because of market inefficiency. A time series of market inefficiency is useful, since if market inefficiency can be forecasted it has implications for risk management. Portfolios looking to profit from the value premium may want to consider the prevailing level of market inefficiency when constructing their portfolios, adjusting their exposure during periods of inefficiency.

References

- Asness, C. S. (2024, August). The less-efficient market hypothesis. *50th Anniversary Issue of The Journal of Portfolio Management, Forthcoming*. Available at SSRN: <https://ssrn.com/abstract=4942046> or <http://dx.doi.org/10.2139/ssrn.4942046>.
- Barclay, M. J. and T. Hendershott (2003, Oct). Price discovery and trading after hours. *Review of Financial Studies* 16(4), 1041–1073.
- Boguth, O., A. J. Fisher, V. Gregoire, and C. Martineau (2023, October). Noisy fomic returns? information, price pressure, and post-announcement reversals. Available at SSRN: <https://ssrn.com/abstract=4131740> or <http://dx.doi.org/10.2139/ssrn.4131740>.
- Cheung, Y.-W. and K. S. Lai (1995). Lag order and critical values of the augmented dickey–fuller test. *Journal of Business & Economic Statistics* 13(3), 277–280.
- Fama, E. F. (1965). Random walks in stock market prices. *Financial Analysts Journal* 21(5), 55–59.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance* 25(2), 383–417.
- Fama, E. F. and K. R. French (1993, Feb). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33(1), 3–56.
- Finance, Y. (2024). Historical data for s&p 500 index (^gspc). <https://finance.yahoo.com/quote/%5EGSPC/history>. Accessed: 2024-11-23.
- French, K. R. (2024). Data library. Accessed: 2024-10-16.
- Lim, K.-P. and R. Brooks (2010, Mar). The evolution of stock market efficiency over time: A survey of the empirical literature. *Journal of Economic Surveys* 25(1), 69–108.
- Lo, A. (2004, 10). The adaptive markets hypothesis: Market efficiency from an evolutionary perspective.
- Mincer, J. and V. Zarnowitz (1969). ISBN, 0–870.
- Van Rossum, G. and F. L. Drake (2009). *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.

Appendix

A Appendix

A.1 Unbiasedness Regressions

In this Appendix we build on the intuition behind the unbiasedness regressions and the measurement of market inefficiency. We provide plots from simulated efficient and inefficient markets, and show how the unbiasedness regressions can be used to measure the degree of inefficiency in the market.

A.1.1 Simulated Efficient Market

In this section we simulate an efficient market, where returns are normally distributed with a mean of 0 and a standard deviation of 0.1. They are independent random events.

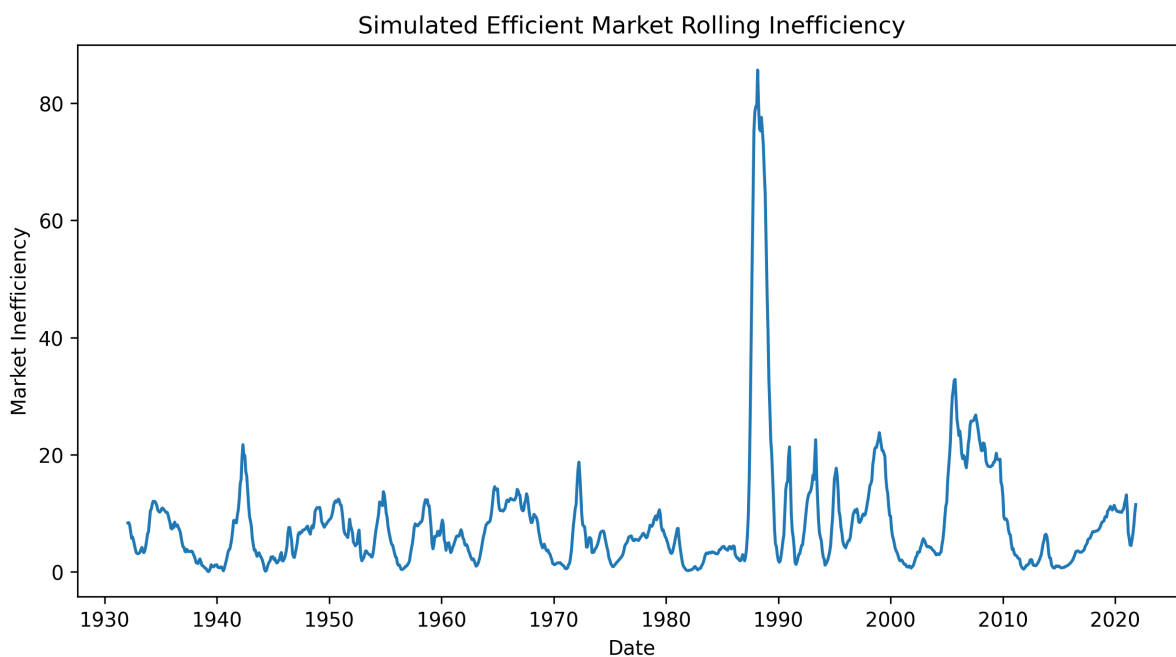


Figure 7: Simulated Efficient Market

In Figure 7 we see the rolling inefficiency of the simulated efficient market. Notice that the level remains low, with only small spikes in inefficiency.

A.1.2 Simulated Inefficient Market

In this section we simulate an inefficient market, where returns are an AR(1) process:

$$r_t = \phi r_{t-1} + \epsilon_t \quad (4)$$

where $\phi = 0.5$ and $\epsilon_t \sim N(0, 0.1)$. The resulting series of market inefficiency is shown in Figure 8. We see that there are much more frequent spikes of inefficiency, and the level of inefficiency is much higher than in the efficient market.

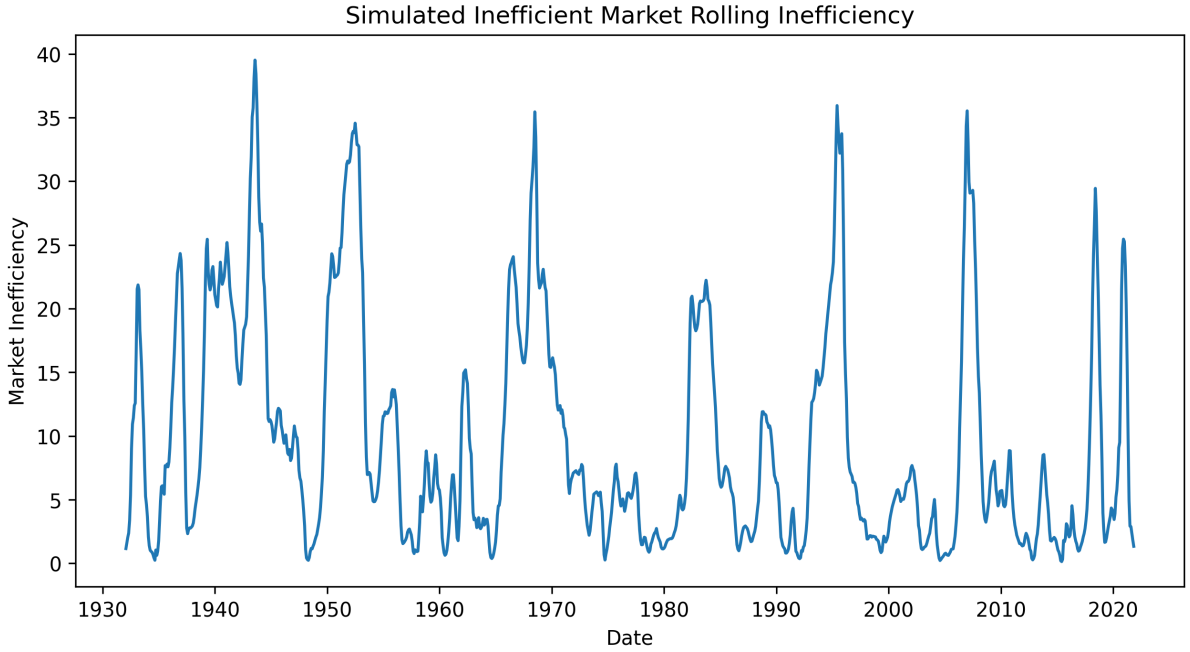


Figure 8: Simulated Inefficient Market

A.1.3 R^2 and Intuition Building

The R^2 value of the unbiasedness regression is a measure of how well the partial returns from 0 to t predict the total return from t to T . While β_t measures the efficiency with which the market processes information, R^2 measures the information content of the partial returns.

I provide an example of the R^2 and β_t values around 'Earnings Months' and 'Non-Earnings Months' in the SP500. I define 'Earnings Months' as the months where the SP500 companies typically report their earnings (January, April, July, October), and 'Non-Earnings Months' as the remaining months.