New Testing Procedures to Assess Market Efficiency with Trading Rules

Bell, Peter N

University of Victoria

15 March 2013

Computing in Economics and Finance
19th International Conference

New Testing Procedures to Assess Market Efficiency with Trading Rules

Peter Bell

Department of Economics, University of Victoria
PO Box 1700 STN CSC, Victoria, BC, V8W 2Y2, Canada
pnbell@uvic.ca 250 588 6939

New Testing Procedures to Assess Market Efficiency with Trading Rules

Abstract: This paper presents two computational techniques and shows that these techniques can improve tests for market efficiency based on profit of trading rules. The two techniques focus on interval estimates for expected profit per trade, in contrast to the standard approach that emphasizes point estimates for profit per trade (Daskalakis, 2013; Marshall, Cahan, & Cahan, 2008). The first technique uses confidence intervals to determine if the expected profit is significantly different from zero. The second technique uses moving-window resampling, a procedure of drawing sub-samples that overlap and move incrementally along a time series, to determine if the expected profit is sensitive to sample selection. The paper develops formal testing criteria based on each technique and uses simulation to establish existence results about the tests for efficiency: the standard approach can give false negative results and the new tests can give correct negative or correct positive results. Using a random walk, I show situations where the standard approach incorrectly determines that a market is inefficient whereas the new techniques do not make this error; the standard approach can be fooled by randomness of profit. Using a mean reverting process and a trading rule designed to exploit mean reversion, based on Bollinger bands, I show that the new techniques can correctly recognize an inefficient market. Since the new testing procedures can correctly identify an efficient or inefficient market, with an error rate discussed in the paper. These results support Fama's (1970) position that trading rules *can* form the basis for the theory of efficient markets. This definition of an efficient market in terms of trading profit is timely given the current dominance of algorithmic trading in secondary markets.

JEL Classification C63, D84, G, G14

Keywords: Efficient market, trading rule, expected profit, testing procedure, confidence interval, moving window, resampling, random walk, mean reversion

1 Introduction

> "It's not the 'Information' that counts; It's a superior *interpretation* of it that counts" (Strategic Economic Decisions, 2013).

A "trading rule" is a process that specifies how traders interpret information and establish speculative positions. A formal definition is provided by Skouras (2001). Trading rules attempt to exploit predictability, yet an efficient market is unpredictable according to the Efficient Market Hypothesis (Malkiel, 2003). Therefore, it is possible to use trading rules to assess whether a market is efficient or not.

The standard approach to testing efficiency with trading rules is exemplified by Marshal, Cahan, & Cahan (2008), who correct for data snooping, compare profits in real markets to simulated ones, and consider different time periods. Their approach is state of the art. However, they do not consider dispersion of profit per trade when assessing significance of profit; they focus on average return per trade, which is a point estimate. In contrast, Daskalakis (2013) does include information about dispersion when testing for efficiency by using the Sharpe ratio; however, he uses daily returns rather than profit per trade. In Section 3 I explain how the standard approach has developed from theory of finance, and how my approach connects back to the fundamental and historically important concept of a "Fair Game".

The objective of this paper is to establish that the standard approach can be improved by adapting two computational techniques to test for significance of profits. First, I propose a test based on Confidence Intervals (CI) for profit per trade. Second, I propose a test based on the variability of average-profit as the sample changes in a "moving window" resampling procedure. By presenting formal decision criterion for each test, I am able to explore the statistical properties of the tests in simulation. I establish that the new techniques can correctly identify an efficient or inefficient market with low error rates. These results serve my broader objective of arguing that the Fair Game concept is more useful than other concepts of efficiency for trading.

Although trading is fundamentally related to theories of finance, there is insufficient discussion of the statistical properties of tests for efficiency based on trading profits. Levich & Thomas (1991) present a bootstrapping procedure to assess if trading profits are statistically significant but this destroys non-linear dependence structures in the data by resampling at random. It is undesirable to resample at random because research on technical analysis has shown that non-linear patterns have significant predictive power (Lo, Mamaysky, & Wang, 2000; Gençay, 1998). Twenty years ago, Kendrick (1993) stated that we need new research techniques to deal with increasing automation of financial markets and this call is gaining greater urgency today as information becomes available about the profitability of high frequency trading (Clark-Joseph, 2013). My paper contributes to this debate on market efficiency by clarifying the importance of trading in historical theories of finance and developing new testing procedures that can be used for any trading rule that has a clear definition of profit or loss per trade.

2 Methods

In this section I provide details to allow the work to be replicated. Further information on why I make each modeling choice is provided in Section 3.

2.1 Trading Rules

A trading rule is a function that assigns a position to a trader at every point in time. As in other research, the position is either long or short one unit of the risky asset. This is a very limited view of trading because it does not consider position sizing, leverage, or risk management. I use two trading rules: Variable Moving Average and Bollinger Bands.

The Variable Moving Average Rule ("VMA Rule") is widely used by researchers and defined by Brock, Lakonishok, & LeBaron (1992). The rule uses two moving averages, a fast one with short lag length and a slow one with large lag length. The positions are: long when the fast average is above the slow average, short when the fast average is below the slow average. For this research I use lag lengths *10* and *50*.

There has been criticism of the VMA Rule because the arbitrary choice of lag lengths raises questions of data snooping. Sullivan, Timmerman, Timmerman, & White (1999) provide a method that uses resampling to check if results are robust to snooping. In an earlier paper, I present a visualization technique to address the same concern (Bell, 2010). Since the purpose of this paper is to extend the standard approach, I use the VMA for an arbitrary pair of lag lengths.

The technical indicators known as Bollinger Bands (BB) are widely used by traders. They are a type of moving standard deviation calculated as:

$$\overline{\sigma(t)} = \frac{1}{N}\sum_{i=1}^{N}\left(P(t-i) - \overline{P(t)}\right)^2 \text{ where } \overline{P(t)} = \frac{1}{N}\sum_{i=1}^{N} P(t-i). \quad (1)$$

I use lag length $N=20$ for all calculations with BB. Further information about these indicators is available from Bollinger Capital Management (2013).

Although the BB indicators are widely known, they have not been used before in research on trading efficiency. This is surprising because the BB provide a reliable way to exploit mean reversion. I introduce a new trading rule called the Bollinger Band Rule ("BB Rule") and define the conditions under which a trader would start or end each trade as follows:

- Start *Short* when: $P(t) \geq \overline{P(t)} + 2\overline{\sigma(t)}$; End *Short* when: $P(t) \leq \overline{P(t)} + \frac{1}{2}\overline{\sigma(t)}$.
- Start *Long* when: $P(t) \leq \overline{P(t)} - 2\overline{\sigma(t)}$; End *Long* when: $P(t) \geq \overline{P(t)} - \frac{1}{2}\overline{\sigma(t)}$.

The BB Rule is designed to exploit mean reversion: it takes long positions when the price is very low, and short positions when price is very high. The concept of *very high* or *low* is made precise using the BB indicator. As with other trading rules, the BB Rule can be refined; however, it serves the purposes of this paper by successfully exploiting mean reversion in prices.

2.2 Price Simulations

I use stochastic processes to represent asset prices in an efficient and inefficient market. Unless otherwise stated, each simulation has length *1000* with initial price *P(1)=100*. All calculations in this paper are conducted in discrete time.

To simulate an efficient market, I use a Random Walk (RW) where price changes are drawn from a standard normal distribution. This represents an efficient market because it is a martingale. However, it has several problems: the RW can take negative values and does not match fat tails observed in real data. It remains as an open challenge to extend the methods in this paper to different price processes.

For an inefficient market, I use an Ornstein-Uhlenbeck (OU) process. The price changes are defined as:

$$\Delta P(t) = \theta(P(t-1) - 100) + \varepsilon \tag{2}$$

Where $\theta = -0.1$ and $\varepsilon \sim N(0,1)$ is standard normal. The purpose of simulating an inefficient market with mean reversion is that it can be exploited by the BB Rule. However, mean reversion is only one type of predictability. Although all markets with (sufficient) predictability are inefficient, not all types of predictability are equally easy to exploit; what is the best trading rule for a market where prices are generated by an autoregressive model with several lags?

2.3 Confidence Intervals

When testing for efficiency, I use two well-known techniques to identify plausible values for the center of a probability distribution. The "Parametric method" is the classic result based on *t*-statistics: if $X(t) \sim N(\mu, \sigma)\ \forall t$ then $[\bar{X} - 1.96\frac{\bar{\sigma}}{\sqrt{n}}, \bar{X} + 1.96\frac{\bar{\sigma}}{\sqrt{n}}]$ contains population mean with 95% significance level. A reference for this formula and extensive discussion of confidence intervals is available in Neyman (1937, P. 376). The "Non Parametric method" is based on quantiles: if $X(t) \sim i.i.d. \forall t$ then $[Q^-(n^-), Q^+(n^+)]$ contains the population median with 95% significance level. Hettmansperger & Sheather derive this method and provide details on how to calculate n^+ and n^- (1986).

2.4 Resampling Procedure

Moving Averages and Bollinger Bands are calculated by forming a series of samples that overlap and move incrementally along a time series. At each time step new data is added and old data is dropped from the sample. This is the fundamental idea of the moving window resampling procedure that I use to test market efficiency. Shasha & Zhu (2004) show that the moving window technique is valuable in many settings, especially finance.

The procedure is well known in information sciences and was given some theoretical justification by Chu (1995), who showed that a "sliding window" approach is an effective way to detect changes in the distribution of data when it is piecewise stationary. Although Chu's approach is different from mine because he compares a sample of data to a reference sample, the sliding and moving window are fundamentally similar approaches because they both draw overlapping samples that move incrementally along a time series.

The moving window takes data $\{P(t)\}$ and creates a sequence of samples $\{S(i)\}$. If $X(t)$ is univariate then the i^{th} sample, $S(i)$, will be a vector of price observations. The procedure requires two variables: the "window size" and the "step size". The window size is the length of each sample, which I denote w. The window size is Chu's h variable (1995, p. 148). The step size is the distance between each sample of data, which I denote k. The step size is Chu's k variable (1995, p. 153). Therefore, the i^{th} sample can be described as the interval:

$$S(i) = [P(1 + k * i),\ P(W + k * i)] \tag{3}$$

Where i takes integer values, including zero. If the total length of price is N, then the maximum value for i is $(N\text{-}w)/k$. This is also the total number of samples that the resampling procedure creates. For simulations in Section 4.3 I use $N = 10{,}000$, $w = 500$, $k = 100$, which gives *95* different samples.

3 Calculation

In this section I develop my testing procedures from theories of market efficiency. The standard approach to testing efficiency compares trading returns against market returns. This comparison is based on an influential theory of asset prices: the submartingale model. Fama wrote that profits from trading could be positive, but no larger than the market returns in expectation under the submartingale model (1970). Despite this strong theoretical basis, the standard approach remains informal and poorly justified in terms of the statistical properties of the tests. To improve on this situation, I argue that a more fundamental view of market efficiency can be used to motivate tests for efficiency.

The bedrock concept of efficiency is arguably Bachelier's first principle which states that speculation should be a Fair Game: if a market is efficient then it is not possible to make positive expected profit by trading (Fama, 1970). Bachelier proposed this idea in 1900 while working on

French bond markets. It had a large influence on subsequent research. When Samuelson proved that prices follow a martingale, he called a crucial theorem the "Theorem of Fair-Game Futures Pricing" (1965, p. 44). When Fama described the broad implications of efficient markets, he put great emphasis on the concept of a Fair Game; he was careful to argue that Fair Game is an implication of an efficient market, rather than a cause (1970, p. 385). However, I argue that we should return to the Fair Game concept of efficiency when testing trading rules because of the direct prediction on trading profits.

The theoretical basis for my tests is the Fair Game principle. To repeat, if a market is efficient then it is not possible to make positive expected profit by trading. The converse of this statement is empirically testable: if there is a rule with positive expected profit then market is not efficient. This leads to the following null and alternative hypotheses:

$$H_0\text{: "Market is efficient"}; \qquad H_1\text{: "Market is inefficient".} \tag{4}$$

Technically speaking, we can prove that a market is inefficient or fail to prove it is inefficient. Researchers should be cautious when saying the market is efficient because there could always be another trading rule that gives positive profit. Therefore, I suggest the following terminology: if a researcher rejects the null, they can say that the market is inefficient; if a researcher fails to reject the null, they can say the market is efficient but this should be understood to mean "no evidence of inefficiency" or "efficient with respect to given trading rule".

The variable of interest in the standard approach is daily returns (Daskalakis, 2013). This variable allows researchers to test the submartingale model by comparing daily returns for trading and the market. It can even be extended to other settings where trades do not occur on a daily basis. Yet this approach is limited because it requires calculation of returns per unit of time. I suggest this is not a meaningful concept in some contexts, such as high frequency trading, where traders can earn profits or loss in several milliseconds. What is the market return at a millisecond time scale? This is one reason that I use the Fair Game principle for efficiency.

The variable of interest under the Fair Game view of efficiency is profit per trade. Therefore, I introduce a simple measure of profit per trade defined as:

$$\Pi = P^S - P^B. \tag{4}$$

Here Π is profit per trade, P^S is the selling price, and P^B is the buying price. Buy low – sell high. Consider a profitable short: start trade by selling at a high price ($P^S > 0$), end trade by buying at a low price ($P^S > P^B > 0$). Although this simple variable serves my purposes here, it can be made more sophisticated by including information about transaction costs, position sizing, or leverage – three concepts which the literature on trading and efficiency has given insufficient consideration to thus far.

The Fair Game concept implies profits for trading are zero in expectation. This is a statement about the center of the distribution of profits, which can be assessed using the mean or median. However, it is insufficient to use a point estimate for either because it does not include information about the dispersion of profits. Therefore, I argue that a CI for mean or median calculated using the Parametric of Non-Parametric methods from Section 2.3 can be used to assess whether expected profits are positive. This is a basic test for statistical significance. This simple testing approach has not been used in prior research, possibly because authors have not focused on profit per trade and the Fair Game concept of efficiency.

The testing procedure that I propose is the same for either type of CI:

CI Test: To test if a market is efficient, use the following decision criteria

- If CI contains zero then fail to reject null hypothesis;
- If CI does not contain zero then reject null hypothesis.

In order for the CI methods to be valid, profit must be identically and independently distributed. That is, Π must have a stationary distribution. Results reported by Marshall et al. (2008) and Daskalakis (2013) state that trading profits vary greatly between sample periods. This suggests the distribution of profit is changing over time for a given trading rule and market. Therefore, I develop another testing procedure designed to address problems with the CI methods caused by non-stationary data.

The Fair Game concept does not require that the distribution of profits must be stationary over time. The principle simply requires that the profits have positive expected value at all times. This is my interpretation from Fama (1970). Therefore, I use the resampling approach from Section 2.4 to calculate a sequence of estimates of average profit per trade $\{\Pi(i)\}$.

Resampling Test: To test if a market is efficient, use the following decision criteria

- If $\Pi(i)$ takes positive and negative values then fail to reject null hypothesis;
- If $\Pi(i)$ takes positive values for all samples then reject null hypothesis.

This Resamping Test is novel because it focuses on performance across samples. The literature has not developed theory to explain how profits over time relate to efficiency. Often researchers will claim the market is getting *more* or *less* efficient, whereas my Resampling Test would suggest the market is efficient and trading profits are simply non stationary.

There are several ways to refine the Resampling Test, such as using the CI for each sample or calculating a test of significance over the sequence of sample-averages. However, this is a situation where the literature on trading and efficiency is lagging practice. Given the widespread use of moving indicators by practitioners, we need new tools to understand the theoretical properties of statistical estimation in a moving-window context.

4 Results

In this section I use simulation to demonstrate that the new Tests can correctly identify efficient or inefficient markets. I also report the accuracy of the CI test in simulation and show that the standard approach can reach an incorrect conclusion about market efficiency.

4.1 Efficient Market

The results in this section are based on the RW and VMA Rule described in Section 2. To begin, I report results for a particular price path chosen for illustrative purposes: the average return and Sharpe ratio for the trading rule exceed that of the buy-hold strategy. The standard approach would infer that this is an inefficient market. However, the RW is a martingale which is *the* classic model for an efficient market. Therefore, the results in Table I establish that the standard approach can reach an incorrect conclusion; it can be fooled by the randomness of trading profits.

Table I

Results for Standard Approach to Efficiency Test

This table provides estimates of returns to a trading rule and a buy-hold strategy. The price is simulated as a Random Walk. The "Trading Rule" is the VMA rule with lag lengths 10 and 50. The "Average Return" is arithmetic average of profit. The "Sharpe Ratio" is the ratio of average profit to standard deviation of profit. The standard approach is to claim the market is inefficient if the trading rule has larger Average Returns and Sharpe Ratio.

Strategy	Average Return	Sharpe Ratio
Buy-Hold	0.01	0.01
Trading Rule	1.52	0.24

Although the trading rule achieves large profits with lower variance (relatively) than the market, the CI for profit per trade reveals that these profits are not significantly different from zero. Both Parametric and Non-Parametric CI contain zero. Therefore, Table II shows that the CI Test can correctly identify an efficient market using either the Parametric or Non-Parametric CI.

Table II

Confidence Intervals for Profit per Trade in Efficient Market

This table provides estimates of the CI for one simulation: RW with length 1000, VMA trading rule with lag lengths 10 and 50. The "Method" is the type of CI calculation. "Parametric" is a CI for population average based on the *t*-statistic. "Non-Parametric" is a CI for population median based on quantiles. The "Upper CI" and "Lower CI" are the values of the bounds of the CI for each method. The "Decision" is a test of the null hypothesis that profits are equal to zero and the market is efficient.

Method	Upper CI	Lower CI	Decision
Parametric	4.20	-1.16	Efficient Market
Non-Parametric	2.85	-2.37	Efficient Market

To explain the results in another way, I report a histogram of profit per trade in Figure 1. This histogram shows the probability of various levels of profit for the VMA rule. The distribution reveals several interesting features that are not apparent from the earlier results. In particular, the profit has large positive skewness: most trades incur a small loss but a few trades make large profits. It is important to consider this information because traders may have different preferences. Skouras (2001) made an important contribution in this direction.

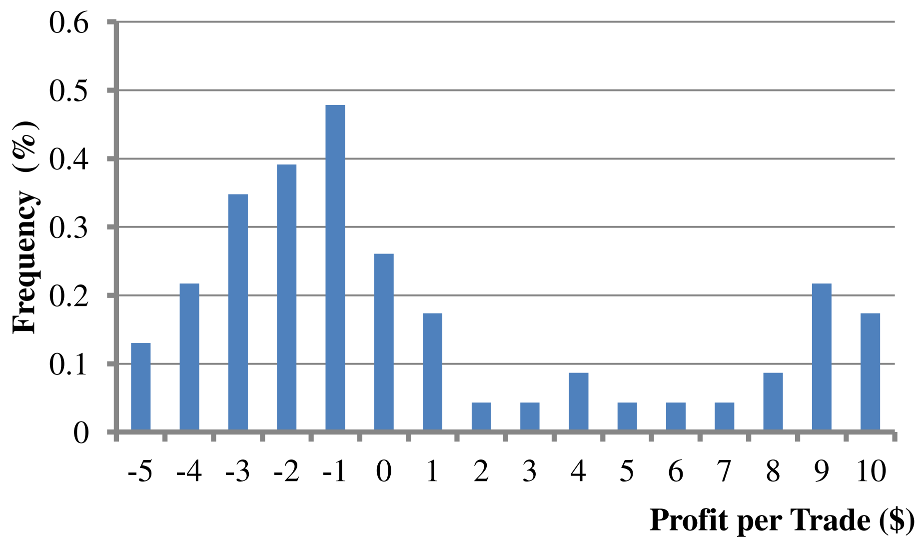

Figure 1. Histogram of Profit in Efficient Market. This figure shows the distribution of profit per trade for the VMA trading rule in a Random Walk that has length 1000. The "Profit per Trade" is the value of profit for a single trade in this simulation. The "Frequency" is the fraction of times that each level of profit is observed. The total number of trades is 24 and the histogram is smoothed by a nearest neighbor procedure.

The results presented so far pertain to a single path of RW simulation. To provide further evidence that the CI Test is superior to the standard approach, I explore the "significance level" for the CI Test in Table III. The significance level is also known as the "size of the test", one minus the error rate. I only report the significance level for the Parametric CI. I calculate significance for several values of "Number of Time Steps" to establish how the accuracy of the test relates to the number of price observations.

Table III reveals that the confidence level is close to the theoretical value, *95%*, for large samples. This suggests that a central limit theorem applies to the distribution of profits, which is a novel result that should encourage further inquiry into the theory of the distribution of profits. It is also important to note that the test has low significance for samples with 100 observations. This may be due to the low number of trades incurred by the trading rule, but it is a cautionary tale for researchers who attempt to make inference based on small data sets.

Table III

Significance Level for Parametric CI Test in Efficient Market

This table provides estimates of the proportion of time that the test correctly identifies an efficient market. These results are specific to the Parametric CI for the VMA trading rule and RW price model. The "Significance Level" or "Size of Test" is one minus the probability of Type 1 error, which are a type of error that occur when the test incorrectly rejects the null hypothesis. The null hypothesis is that profits are equal to zero and the market is efficient. The "Number of Time Steps" is the length of the time series used in the simulation; for each value, we simulate 1000 different paths and calculate the number of times the test is correct. The "Number of Trades" is the average number of trades that occurs for the VMA rule with time series of particular length.

Number of Time Steps	100	500	1000	5000
Significance Level	0.543	0.867	0.91	0.947
Number of Trades	5	10	21	125

The results in this section show that the CI Test can correctly identify an efficient market with a low error rate. They also show that the standard approach can make an important error. These are both novel results for the literature. Although the results presented here are limited to the RW model of prices and VMA Rule, the methods can be extended to other models of prices and trading rules.

4.2 Inefficient Market

The results in this section are based on the OU and BB methods described in Section 2. To begin, I report the histogram of profits for a single price path in Figure 2. The distribution is very different from Figure 1 because the profits are centered on positive values, which suggest the BB rule is successfully exploiting the inefficiency in the market.

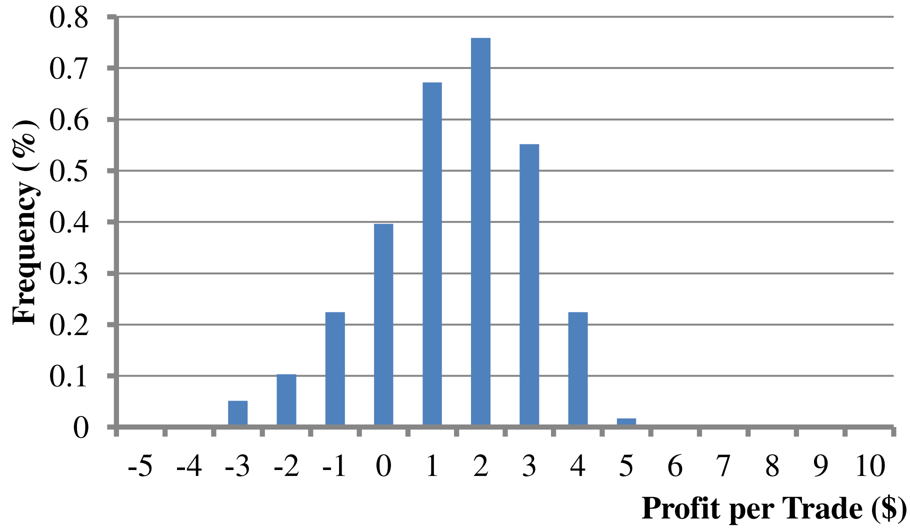

Figure 2. Histogram of Profit in Inefficient Market. This figure shows the distribution of profit per trade for the BB trading rule in an OU process that has length 1000. The "Profit per Trade" is the value of profit for a single trade in this simulation. The "Frequency" is the fraction of times that each level of profit is observed. The total number of trades is 58 and the histogram is smoothed by a nearest neighbor procedure.

Results of the CI Test for the BB Rule are reported in Table IV. The confidence intervals for both mean and median profit do not contain zero. This demonstrates that CI Test is able to correctly identify an inefficient market.

Table IV

Confidence Intervals for Profit per Trade in Inefficient Market

This table provides estimates of the CI for one simulation: OU with length 1000, BB trading rule with lag length 20. The "Method" is the type of CI calculation. "Parametric" is a CI for population average based on the *t*-statistic. "Non-Parametric" is a CI for population median based on quantiles. The "Upper CI" and "Lower CI" are the values of the bounds of the CI for each method. The "Decision" is a test of the null hypothesis that profits are equal to zero and the market is efficient.

Method	Upper CI	Lower CI	Decision
Parametric	1.73	1.10	Inefficient Market
Non-Parametric	2.21	1.14	Inefficient Market

As in Section 4.1, I explore the error rate for the Parametric CI Test in the inefficient market. However, I report "power" for the test in this case because the null hypothesis is false. Power represents the ability of the test to correctly identify an inefficient market, one minus the error rate. The results in Table V show that the CI Test has a very low error rate for relatively small samples (500 price observations). The power may be unrealistically high due to the fact that I used a trading rule and stochastic process that were designed for each other.

Table V

Power for Parametric CI Test in Inefficient Market

This table provides estimates of the proportion of time that the test correctly identifies an inefficient market. These results are specific to the Parametric CI for the BB trading rule and OU price model. The "Statistical Power" is one minus the probability of Type 2 error, which are a type of error that occur when the test incorrectly fails-to-reject the null hypothesis. The null hypothesis is that profits are equal to zero and the market is efficient. The "Number of Time Steps" is the length of the time series used in the simulation; for each value, we simulate 1000 different paths and calculate the number of times the test is correct. The "Number of Trades" is the average number of trades that occurs for the VMA rule with time series of particular length.

Number of Time Steps	100	500	1000	5000
Statistical Power	0.76	0.99	1.0	1.0
Number of Trades	6	29	58	265

The results in this section provide strong support for the Fair Game theory of market efficiency because they show that trading rules can correctly identify an inefficient market.

4.3 Resampling Procedure

In this section I report results for both the efficient and inefficient markets on the same figure. Figure 3 shows the average profit for the VMA Rule in an efficient market and BB Rule in an inefficient market as the sample moves across the time series. These results are comparable to Sections 4.1 and 4.2.

The results of the Resampling Tests are both correct, as follows. The sequence of average profits, $\Pi(i)$, for the VMA Rule is not always positive so the Resampling Test correctly infers that the RW is efficient. The sequence of average profits for the BB Rule is always positive so the Test correctly infers that the OU market is inefficient. Also, it is interesting to note that $\Pi(i)$ has a tendency to take large positive values for the VMA Rule, which is similar to the positive skew observed in Figure 1.

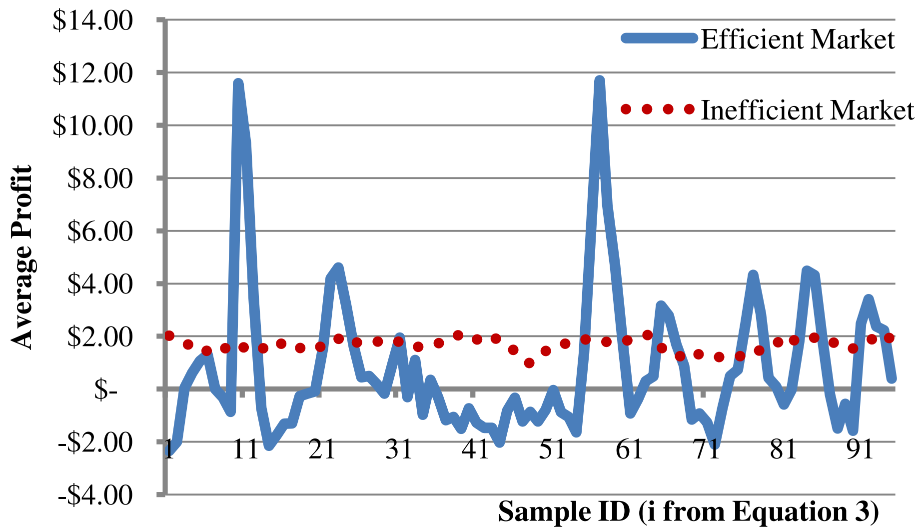

Figure 3. Average Profit per Trade under Moving Window Resampling Procedure. This figure shows how average profit changes as the sample changes. The "Efficient Market" is the result for the VMA trading rule with the RW process. The "Inefficient Market" is the result for the BB trading rule with the OU process. The "Average Profit" is the average value of profit for a single trade. The "Sample ID" is index (i) from the Equation 3 that describes the moving window resampling procedure. Each sample has window length equal to $w=500$, step size $k=100$, and total length of simulation $N=10,000$; this gives 95 different samples.

The results of this section establish that the moving window procedure can be used to correctly identify an efficient or inefficient market. This implies that there is an economically significant relationship between the distribution of profits over time and market efficiency.

4 Discussion

In this section I relate the results to theories of market efficiency, discuss external validity, and point out new topics for research. The results of this paper show that trading rules can correctly identify efficient or inefficient markets. This means the original idea of a Fair Game has empirical validity, which is important because the Fair Game concept has been neglected in favour of the submartingale model.

The Fair Game concept of efficiency is that profits for trading are zero. This is similar to an absolute return measure of investment performance. In contrast, the submartingale concept of efficiency implies that profits for trading are no larger than the market returns. This is similar to a relative return measure of investment performance. There is debate on which measure of investment performance is appropriate for a given context. The results of my paper show that it is appropriate to use absolute returns for trading rules, rather than relative returns. Further, the analysis of absolute returns allows inference about market efficiency.

I have attempted to establish external validity for the results of this paper by reporting the statistical properties of the tests. However, this validity is limited by a ubiquitous problem in finance: how well does my model of prices reflect reality? I used RW and OU processes, but these are not particularly realistic models. Further, Brock et al. (1992) showed that profits for trading rules in simulated markets are very different from profits based on real data, which further reduces the external validity of my results. The good news is that the methods developed here (CI Test, Resampling Test) can be applied to more realistic models for asset prices for a broad set of trading rules.

Another limitation of the validity of these results is the way I have modeled traders' decision making. I propose a single rule and allow agents no discretion, which is a simplified version of reality. Skouras (2001) and others have showed that it is possible to use artificial intelligence to find trading rules endogenously in data, which raises questions about learning that are being addressed in other areas of economics, such as: the El Farol Bar problem (Arthur, 1994) or endogenous beliefs (Kurz & Motolese, 2011). Again, the good news is that the methods developed in this paper can be used with endogenous trading rules to determine if profits are significantly positive.

One topic that I feel deserves further study is the distribution of profit per trade. I provided techniques to determine if the center is significantly different from zero or if the distribution is stationary, but these basic questions can be elaborated on. Is there a relationship between duration of trade and profitability? How does the distribution of profit vary across trading rules and data generating processes? Do traders have preferences over these distributions? This topic has theoretical motivation from the Fair Game concept and will allow researchers to tackle new problems in industry, such as the profitability of high frequency trading

A final topic that deserves further research is the moving window. In fact, I would like to suggest a different way to think about asymptotic statistics: not an ever-increasing sample size, rather, an ever-increasing number of samples of a given size. This focus on the number of samples can be achieved by using a moving window and reflects a concern with *out of sample* performance. This resampling approach is widely used in practice but not widely used in theory, which may trace back to Fisher's emphasis on independent samples (Zilliak and McCloskey, 2007). The moving window does not give independent samples, but it does provide an effective way to assess how sample selection effects estimation results. This represents an opportunity for science to *pull* ideas from business, which can be a fruitful approach to research.

References

Arthur, B. (1994). Inductive Reasoning and Bounded Rationality, *American Economic Review, 84*(2), 406–411.

Bell, P. (2010). *Introduction to the Profit Surface* (MPRA Working Paper). Retrieved from http://mpra.ub.uni-muenchen.de/26812/

Bollinger Capital Management (2013). Bollinger Bands | Tutorial. Retrieved from http://www.bollingerbands.com/services/bb/

Brock, W., Lakonishok, J., & LeBaron, B. (1992). Simple Technical Trading Rules and the stochastic properties of Stock Returns, *Journal of Finance*, *47,* 1731–1764.

Clark-Joseph, A. (2013). *Exploratory Trading*. (Unpublished job market paper). Harvard University, Cambridge MA.

Chu, C. S. (1995). Time series segmentation: A sliding window approach, *Information Sciences, 85*(1), 147–173.

Daskalakis, G. (2013). On the efficiency of the European carbon market: New evidence from Phase II*, Energy Policy, 54,* 369–375.

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work, *The Journal of Finance, 25*(2), 383–417.

Gençay, R. (1998). Optimization of technical trading strategies and the profitability in security markets, *Economics Letters, 59*, 249–254.

Hettmansperger, T.P., & Sheather, S.J. (1986). Confidence intervals based on interpolated order statistics, *Statistics & Probability Letters, 4*(2), 75–79.

Kendrick, D. (1993). Research Opportunities in Computational Economics, *Computational Economics*, *6*, 257–314.

Kurz, M. & Motolese, M. (2011). Diverse beliefs and time variability of risk premia, *Economic Theory, 47*(2), 293–335.

Levich, R., & Thomas, L. (1993) The significance of technical–trading rules profits in the foreign exchange market: A bootstrap approach, *Journal of International Money and Finance, 12*(5), 451–474.

Lo, A., Mamaysky, H., & Wang, J. (2000). Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation, *The Journal of Finance, 55*(4), 1705–1765.

Marshall, B. R., Cahan, R. H., & Cahan, J. M. (2008). Can commodity futures be profitably traded with quantitative market timing strategies? *Journal of Banking & Finance, 32*, 1810–1819.

Malkiel, B. (2003). The Efficient Market Hypothesis and Its Critics, *The Journal of Economic Perspectives, 17*(1), 59–82.

Neyman, J. (1937). Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability, *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences, 236*(767), 330–380.

Samuelson, P. (1965). Proof That Properly Anticipated Prices Fluctuate Randomly, *Industrial Management Review, 6*(2), 41–49.

Shasha, D., & Zhu, Y. (2004). *High Performance Discovery in Time Series: Techniques and Case Studies*. New York: Springer.

Skouras, S. (2001). Financial returns and efficiency as seen by an artificial technical analyst, *Journal of Economic Dynamics & Control, 25*, 213–244.

Strategic Economic Decisions (2010). Home - Strategic Economic Decisions. Retrieved from http://www.sedinc.com/

Sullivan, R., Timmerman, R., Timmerman, A., & White, H. (1999). Data–Snooping, Technical Trading Rule Performance, and the Bootstrap, *The Journal of Finance, 54*(5), 1647–1691.

Ziliak, S.T. & McCloskey, D.N. (2007). *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives.* Ann Arbor: University of Michigan Press.