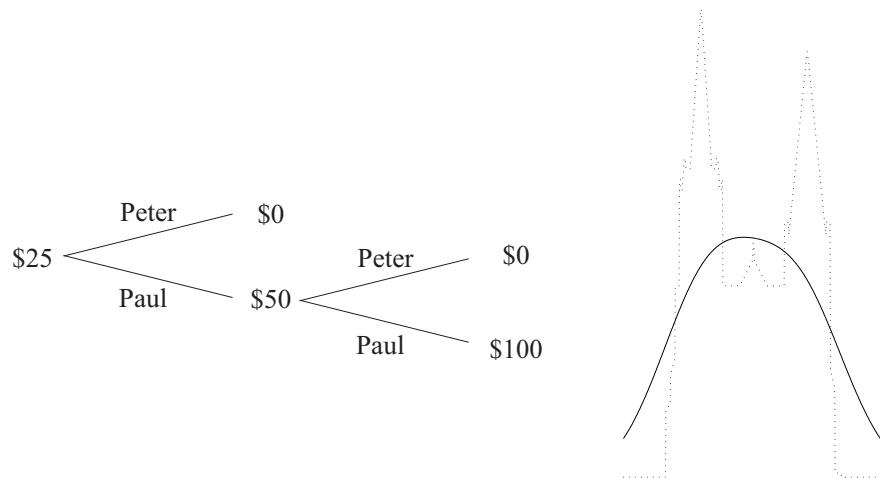


# Testing Lead-Lag Effects under Game-Theoretic Efficient Market Hypotheses

Wei Wu and Glenn Shafer



**The Game-Theoretic Probability and Finance Project**

Working Paper #23

November 17, 2007

Project web site:  
<http://www.probabilityandfinance.com>

## Abstract

A game-theoretic efficient market hypothesis says that a trading strategy will not multiply the capital it risks substantially relative to a specified market index. This implies that the autocorrelation of returns with respect to the index will be small and that a signal  $x$  will have approximately the same lead-lag effect on all traded securities. These predictions do not depend on assumptions about probabilities and preferences. Instead they rely on the game-theoretic framework introduced by Shafer and Vovk in 2001, which unifies statistical testing with the notion of a trading strategy that risks only a fixed capital. In this framework, we reject market efficiency at significance level  $\alpha$  when the capital risked is multiplied by  $1/\alpha$  or more. This approach identifies the same anomalies as the conventional approach: statistical significance for the autocorrelations of small-cap portfolios and equal-weighted indices, as well as for the ability of other portfolios to lead them. Because it bases statistical significance directly on trading strategies, the approach allows us to measure the degree of market friction needed to account for this statistical significance. We find that market frictions provide adequate explanation.

# 1 Introduction

Substantial autocorrelations for returns appear to violate the efficient-markets hypothesis (EMH). Such apparent violations, or *anomalies*, are important to finance theory. By searching for market frictions that can rule out the trading profits that anomalies seem to permit, we arrive at a better understanding of the market.

In this paper, we study autocorrelations and other lead-lag effects using hypotheses that are related to the EMH but are simpler. We call them *game-theoretic efficient-market hypotheses* (GEMHs). A GEMH states that no trading strategy will multiply the capital it risks by a large factor relative to a specific market index. This can be tested directly, with no further assumptions about probabilities or the preferences of investors.

As we show in §2, simple GEMHs predict that the autocorrelations of returns will be small. Though the prediction is not new, it is interesting that we can make it using many fewer assumptions than EMH models require. More importantly, the GEMH approach allows us to measure the *relative efficiency* of a financial market using the level of transaction costs needed to explain away the observed autocorrelation. As we show in §3, the observed anomalies are explained by small levels of transaction costs. This type of analysis is possible because a GEMH allows us to identify a statistical test with a strategy for speculation. Our simple GEMHs also make a less familiar prediction: they tell us that a signal should have approximately the same lead-lag effect on the returns for all securities. The violation of this prediction can be understood as an additional anomaly, which can also be explained by transaction costs.

In §1.1, we explain how our results can help economists overcome their failure to reach consensus concerning the EMH. In §1.2, we explain how the game-theoretic foundation for probability introduced by Shafer and Vovk (2001) motivates GEMHs. In §1.3, we summarize predictions GEMHs make concerning lead-lag effects, and in §1.4, we summarize the comparison of these predictions with actual US stock returns. In §1.5, we relate our results to work by other authors.

## 1.1 Breaking the EMH stalemate

Boudoukh, Richardson, and Whitelaw (1994) have categorized explanations of anomalies into three types: loyalist, revisionist, and heretical. Loyalist explanations point to market frictions that prevent observed prices from matching rational prices. Revisionist explanations propose more complicated models of rational price formation, which may be flexible enough to match observed prices. Heretical explanations abandon the hypothesis that market prices are rational.

The debate between these different viewpoints has been inconclusive. As Andrew Lo (2007) has pointed out, economists have not reached consensus on the validity of the EMH, and tests of the EMH have not even led to agreement on how to measure the relative efficiency of markets. Whenever there seems to be evidence for more friction in some markets than others, a revisionist explains

the evidence away.

One reason the stalemate persists is the complexity of the EMH. From its very beginning in the work of Samuelson (1965) and Fama (1965), the EMH has woven together two threads: the hypothesis that there are no profitable trading strategies and the hypothesis that prices incorporate all relevant information. In theory, we can tie the two threads together by saying that a marginal-utility-weighted price should be a martingale with respect to an objectively valid probability distribution (Lucas Jr. 1978). But in practice, the two still lead in different directions. Loyalists emphasize trading when they argue that the EMH is confirmed whenever apparently profitable trading strategies are ruled out by market frictions (e.g. Malkiel 2003). Revisionists emphasize information when they construct probabilities and investor preferences that match observed prices (e.g., Hameed 1997). Heretics also emphasize information when they argue that these probabilities and investor preferences are implausible (e.g., Shiller 2003).

As Fama (1970) first pointed out, the EMH has a joint hypothesis problem. In order to test the EMH, we must incorporate it into a model that specifies probabilities and utilities for a marginal investor. So it can never be rejected definitively. Even if many models produce the same anomaly, there may be a not-yet-examined model that eliminates the anomaly.

The game-theoretic approach escapes from the joint hypothesis problem by identifying tests with trading strategies. Such tests allow us to identify anomalies and measure the degree of market inefficiency they imply without modeling probabilities and utilities. The degree of market inefficiency implied by a statistically significant anomaly is simply the level of transaction costs required to eliminate the strategy's profit and hence the test's significance. This approach also breaks the EMH stalemate, inasmuch as it makes clear how far the hypothesis of "no profitable trading strategy" takes us. If the level of transaction costs needed to eliminate profits from an anomaly is far below the level that actually exists in the market (as seems to be the case for the anomalies we study in this paper), then we have an explanation for why a reasonably efficient market allows the anomaly to exist. This does not eliminate interest in revisionist explanations for why the anomaly arises, but these explanations, too, must take into account the level of transaction costs the anomaly implies.

## 1.2 The game-theoretic framework

A GEMH is about a particular market (e.g., the market in the stocks monitored by CRSP) and a particular traded instrument  $m$ , such as a currency, a money market account, or a traded market index (e.g., the S&P 500 index). It states that no strategy for trading in the market will multiply the capital it risks by a large factor relative to  $m$ .

The EMH literature has paid little attention to the capital risked by trading strategies. The strategies considered are often self-financing, but because they sell stocks short, the capital they risk may be much larger than meets the eye. So the idea of measuring the profitability of a trading strategy by the factor

by which it multiplies the capital it risks is novel. This idea derives from the game-theoretic analysis of classical probability provided by Shafer and Vovk (2001).

It is conventional to reject a probabilistic hypothesis because of the happening of an event singled out in advance to which the hypothesis gives small probability—say probability less than  $\alpha$ , where  $\alpha$  is a small positive number such as 5%. As Shafer and Vovk show, this is equivalent to rejecting the hypothesis because the expected values it gives, when interpreted as prices, permit a speculator to multiply the capital he risks by  $1/\alpha$ . Multiplying one's capital by 1000, for example, is the same as an event of probability 0.001 happening (Shafer and Vovk 2001, pp. 194–197).

Shafer and Vovk also show that classical theorems such as the law of large numbers and the central limit theorem generalize beyond situations where complete probability distributions for future events (and therefore prices for all possible payoffs) are given, to situations where only some payoffs are priced. The key idea is that we can still test the prices by trying to multiply the capital we risk by a large factor. Predictions such as the law of large numbers are derived from the assumption that strategies for doing so will not succeed. When applied to a financial market rather than to abstract probability theory, these techniques allow us to test the efficiency of the market and to derive predictions from the assumption that it is efficient.

Consider a perfect-information  $N$ -period game between two players, Speculator and Market. Think of the capital risked by a strategy for Speculator as his initial capital in the game, say  $\mathcal{K}_0$ . For simplicity, set  $\mathcal{K}_0$  equal to 1, and assume that Speculator trades in only a single security  $y$ . At the beginning of each period, Speculator chooses a fraction  $Y_n$  of his current capital to invest in  $y$  during the period; he invests the remainder in  $m$ . At the end of each period, Market decides on  $y$ 's simple return  $y_n$ . The following protocol spells out how this determines Speculator's capital process  $\mathcal{K}_0, \mathcal{K}_1, \dots, \mathcal{K}_N$ :

BASIC MARKET PROTOCOL

$$\mathcal{K}_0 := 1.$$

FOR  $n = 1, 2, \dots, N$ :

Speculator announces  $Y_n \in \mathbb{R}$ .

Market announces  $y_n \in (-1, \infty)$ .

$$\mathcal{K}_n := \mathcal{K}_{n-1}(1 + Y_n y_n).$$

Here we assume that  $\mathcal{K}_n$  and  $y_n$  are measured using the GEMH's  $m$  as the numeraire, or unit of monetary value. If we use some other numeraire instead, then the formula  $\mathcal{K}_{n-1}(1 + Y_n y_n)$  for the capital at the end of period  $n$  must be replaced by  $\mathcal{K}_{n-1}(1 + m_n + Y_n(y_n - m_n))$ , where  $y_n$  and  $m_n$  are returns with respect to the other numeraire; see §2.3.

What does it mean for a strategy for Speculator to risk only the initial capital  $\mathcal{K}_0$  in the game? It means that the move  $Y_1$  specified by the strategy risks no more than  $\mathcal{K}_0$ , the move  $Y_2$  risks no more than  $\mathcal{K}_1$ , and so on. In other words, the strategy never risks bankruptcy. It guarantees that  $\mathcal{K}_n$  will be nonnegative for all  $n$  no matter how Market plays.

In order to test market efficiency or make a prediction based on market efficiency, we select a small positive number  $\alpha$  (our *significance level*) and a strategy for Speculator that risks only the initial unit capital.

- TESTING: We reject efficiency if Market moves in such a way that  $\mathcal{K}_N > 1/\alpha$  when Speculator follows the strategy.
- PREDICTION: If  $A$  is a statement about Market's moves  $y_1, \dots, y_N$  (a bound on their autocorrelation, for example), and if the strategy guarantees that  $\mathcal{K}_N > 1/\alpha$  if  $A$  fails, then we are  $1 - \alpha$  confident that  $A$  will happen.

If we measure  $\mathcal{K}_n$  and  $y_n$  with respect to a numeraire different from our GEMH's  $m$ , we must replace  $\mathcal{K}_N > 1/\alpha$  in both statements with

$$\mathcal{K}_N > (1/\alpha) \prod_{n=1}^N (1 + m_n),$$

because  $\prod_{n=1}^N (1 + m_n)$  is the ratio by which  $m$  grows relative to the numeraire.

Strictly speaking, the requirement that Speculator not risk more than his current capital  $\mathcal{K}_{n-1}$  on the  $n$ th round prevents him from going short in  $y$  (choosing  $Y_n < 0$ ) or in  $m$  (choosing  $Y_n > 1$ ). In the first case, Market can bankrupt him by making  $y_n > -1/Y_n$ ; in the second case he can do so by making  $y_n < -1/Y_n$ . In order to permit Speculator to undertake a prudent amount of short selling, we will impose an upper bound on  $|y_n|$ . But we will make no other assumptions about Market's moves. In particular, we will not assume probabilities for them.

The two players, Speculator and Market, move in the order indicated and see each other's moves. Within minimal constraints (such as a constraint on  $|y_n|$ ), each can move as he pleases. Modeling the market as a single player who can play strategically against Speculator may seem unreasonably pessimistic, but it is appropriate because our mathematical results establish what Speculator can achieve no matter what the market does. The case where Market plays strategically may be the worst case—the most difficult case for Speculator's strategy. But the prediction is also valid if Market does not play strategically.

The intuition underlying the game-theoretic approach is the usual one for market efficiency: opportunities for easy money will already have been eliminated. This intuition is strongest when applied to simple strategies, and we use simple strategies in this paper: momentum strategies, contrarian strategies, and averages of the two. Previous work in the game-theoretic framework by Vovk and Shafer has shown that these simple strategies can enforce  $\sqrt{dt}$  scaling on asset returns (Vovk and Shafer 2003), as well as a CAPM-type relation between observed averages of returns and their observed correlations with market returns (Vovk and Shafer 2007). The present study is the first, however, to apply the game-theoretic approach to autocorrelations and lead-lag effects.

The game-theoretic approach emphasizes the short-run impact of speculation, as opposed to the impact of investors' long-run preferences concerning risk.

The only risk considered is the risk of bankruptcy. The approach does not necessarily contradict equilibrium analyses that make assumptions about investors' preferences and expectations. As Stephen F. LeRoy explained over thirty years ago, "it is likely that changes in the expected rate of return due to changes in estimates of risks are small in comparison with the short-run fluctuations in realized rates of return," so that "we are led to expect on prior grounds that if capital markets are efficient, rates of return will follow a martingale distribution even in the presence of risk-aversion" (LeRoy 1973, p. 445). But empirical validation of long-run equilibrium analyses will require that they explain more than is explained by short-horizon speculative efficiency.

### 1.3 Our predictions

Suppose that in addition to the  $N$  successive returns  $y_1, \dots, y_N$ , Market also announces an earlier return  $y_0$ , and that when he announces  $y_n$ , he also announces a signal  $x_n$ . Then we can compute the *empirical lagged correlation* of  $x$  and  $y$ ,

$$\text{corr}(x, y) := \frac{\sum_{n=1}^N (x_{n-1} - \bar{x}^-)(y_n - \bar{y}^+)}{\sqrt{\sum_{n=1}^N (x_{n-1} - \bar{x}^-)^2} \sqrt{\sum_{n=1}^N (y_n - \bar{y}^+)^2}}, \quad (1)$$

where  $\bar{x}^- = (1/N) \sum_{n=1}^N x_{n-1}$  and  $\bar{y}^+ = (1/N) \sum_{n=1}^N y_n$ . Like other authors, we are interested in  $\text{corr}(x, y)$ , but we will find it convenient to emphasize the simpler quantity

$$\text{eff}(x, y) := \sum_{n=1}^N x_{n-1} y_n, \quad (2)$$

which we call the *lead-lag effect* of the signal  $x$  on the security  $y$ . When the average of the  $y_n$  is close to zero,  $\text{eff}(x, y)$  will be numerically indistinguishable from the numerator of  $\text{corr}(x, y)$ . So under reasonable assumptions on the denominator, a small value for  $\text{eff}(x, y)$  will imply a small value for  $\text{corr}(x, y)$ .

In §2.2, we show that if the returns  $y_n$  are measured with respect to  $m$ , then the GEMH for  $m$  implies that the lead-lag effect  $\text{eff}(x, y)$  will be small. In §2.3, we show that two lead-lag effects involving the same signal  $x$ , say  $\text{eff}(x, y)$  and  $\text{eff}(x, z)$ , will be approximately equal under the GEMH for  $m$  even if the returns  $y_n$  and  $z_n$  are measured with respect to a numeraire different from  $m$ .

Our theory does not require that the  $x_n$  be returns for a security in which Speculator can trade, but we do need the  $x_n$  to be scaled like returns: of order  $\sqrt{dt}$ , where  $dt$  is the length of time in the trading period. We are most interested, of course, in the case where  $x_n = y_n$  so that (1) reduces to  $y$ 's *empirical autocorrelation*,

$$\text{corr}(y, y) := \frac{\sum_{n=1}^N (y_{n-1} - \bar{y}^-)(y_n - \bar{y}^+)}{\sqrt{\sum_{n=1}^N (y_{n-1} - \bar{y}^-)^2} \sqrt{\sum_{n=1}^N (y_n - \bar{y}^+)^2}}, \quad (3)$$

and (2) reduces to  $y$ 's *auto-effect*,  $\text{eff}(y, y) := \sum_{n=1}^N y_{n-1} y_n$ .

## 1.4 Our empirical results

We test the efficiency of the United States stock market by testing our predictions on weekly stock returns for 1962 to 2005. These are our principal findings:

- The game-theoretic approach identifies the same anomalies as the conventional approach. Autocorrelations for portfolios of small-cap stocks are statistically significant, as are the lead-lag effects of portfolios of larger-cap stocks on small-cap stocks.
- Although some of the autocorrelations and lead-lag effects are large, they are accounted for by modest levels of transaction costs. The largest autocorrelation, that for the portfolio consisting of the one-fifth of stocks with the smallest market capitalization, is more than 0.5, but its statistical significance is accounted for by a transaction cost of less than 1.5%. The implied transaction cost for other autocorrelations and lead-lags is much less.
- Game-theoretic statistical tests, like conventional statistical tests, vary in their power. But unlike statistical significance, the implied transaction cost seems to be insensitive to the power of the test.
- As measured by implied transaction costs, the market was noticeably more efficient in 1995–2005 than in 1962–1994.

## 1.5 Connections with other work

Larry Fisher (1966) was the first to point out that thin trading can produce positive autocorrelation for a portfolio’s returns. On a day when the market as a whole moves up or down, the stocks in the portfolio that trade tend to move in the same direction, and so the portfolio does too. The stocks in the portfolio that do not trade that day tend to catch up the next day, and so the portfolio tends to move in that direction again. In recent decades, other authors have pointed out that the same effect can cause lagged returns for one portfolio to have a positive correlation with current returns for another. Such correlations have been called “cross-autocorrelations” (Lo and MacKinlay 1990b,a); we call them simply “lagged correlations.” Stocks with small market capitalization generally trade most thinly, and not surprisingly, portfolios consisting of small-cap stocks show the most substantial autocorrelations and the most substantial correlations with lagged returns for other portfolios. Substantial autocorrelations are also observed for equal-weighted averages of all the stocks in the market, most of which are small-cap stocks (Campbell et al. 1997).

For the past twenty years, the literature on autocorrelations and lagged correlations of returns has also discussed the opening they provide for “contrarian” trading strategies, which go by this name because they involve buying stocks that have lost value and selling stocks that have gained value. It seems at first paradoxical that a positive autocorrelation for a portfolio’s returns should produce profits for a strategy that bets on failure being followed by success. One of

the contributions of Lo and MacKinlay (1990b) was a resolution of this paradox. Their explanation points to the positive lagged correlation between winners and losers: because the winners went up yesterday, we can expect the losers to go up today.

Our game-theoretic approach casts a different light on this picture. When we consider only strategies that avoid risk of bankruptcy, as the game-theoretic approach requires, the obvious strategy for taking advantage of a positive correlation is most naturally described as a “momentum” strategy. The momentum strategy for taking advantage of a positive autocorrelation for a portfolio’s returns is to make one’s investment in the portfolio proportional to its previous period’s return, thus increasing this investment when the return goes up and decreasing it when the return goes down. More generally, the momentum strategy for taking advantage of a positive correlation between the returns of portfolio A and the lagged returns of portfolio B is to make one’s investment in A proportional to B’s return in the previous period. When B is different from A, the strategy that uses B’s momentum as a signal for investing in A does resemble the “contrarian” strategy of selling winners to buy losers. This is especially true when A and B together comprise all the stocks in the market, and we keep capital not invested in A in a market index. But each game-theoretic momentum strategy relies only on the correlation that inspires it. In §2, we show that this correlation alone is enough to make the strategy work. In §3, moreover, we find empirically that the effectiveness of the strategies tends to be roughly proportional to the magnitude of the correlation. The autocorrelations are larger than the other lagged correlations, and the strategies based on the autocorrelations tend to be most effective—the most profitable and hence the most powerful as statistical tests.

In addition to not trading on particular days, small-cap stocks may also have less timely closing prices, less informed trades, and so on (Lo and MacKinlay 1990b, Chordia and Swaminathan 2000). It may be a matter of taste whether all these consequences of illiquidity should be called inefficiencies. But to the extent that they impede a trading strategy, they can all be thought of as transaction costs, and so our demonstration that observed correlations can be accounted for by modest transaction costs further weakens the case for models that account for them instead by time-varying expected returns (Conrad and Kaul 1988, 1989, Hameed 1997, Bernhardt and Davies 2005).

In addition to positive autocorrelations for portfolios, nontrading can also produce negative autocorrelations for individual stocks (Roll 1984). These autocorrelations are also observed, but they are seldom statistically significant (French and Roll 1986, Campbell et al. 1997). This is also confirmed by our game-theoretic results.

A great deal of work has been done to find the most powerful statistical tests of return correlations under the hypothesis that stock prices follow a random walk. The variance-ratio tests developed by Cochrane (1988), Lo and MacKinlay (1988), and Chow and Denning (1993) are among the most powerful. They test correlations at many lags simultaneously. The game-theoretic tests in this paper can be adapted to deal simultaneously with many lags, and this deserves

exploration, but in this paper we show only that the empirical results obtained by the variance-ratio test can be replicated with game-theoretic tests of single-lag correlations.

There has also been work on the robustness of the statistical significance of observed correlations. Nonparametric tests have been developed by Wright (2000); see also Whang and Kim (2003). Inasmuch as it makes no stochastic assumptions at all, a game-theoretic test is much more robust than a nonparametric test.

## 2 Game-Theoretic Bounds

In this section, we derive bounds for a lead-lag effect  $\text{eff}(x, y)$  from the GEMH for  $m$ . In §2.2, we consider the case where the returns  $y_n$  are measured with respect to  $m$ . In §2.3, we consider the case where they are measured with respect to a numeraire different from  $m$ .

Remember that  $\text{eff}(x, y)$  is an empirical quantity—a function of the observations  $x_0, \dots, x_N$  and  $y_0, \dots, y_N$ . We do not assume a probability distribution for the  $x_n$  and  $y_n$ , and therefore we do not consider a theoretical counterpart for  $\text{eff}(x, y)$ .

### 2.1 Averaging strategies

To obtain a two-sided bound on  $\text{eff}(x, y)$ , we will average two strategies, one giving an upper bound, the other giving a lower bound. In general, as we now explain, an average strategy allows us to assert simultaneously the predictions made by the two strategies individually, but this involves relaxing the significance level.

Suppose  $\alpha > 0$ ,  $A$  is a statement about Reality's moves  $x_0, y_0, \dots, x_N, y_N$ , and Speculator has a strategy guaranteeing both (1)  $\mathcal{K}_n \geq 0$  for  $n = 1, \dots, N$  and (2) either  $A$  happens or  $\mathcal{K}_N \geq 1/\alpha$ . Then we say that the *GEMH predicts  $A$  at level  $\alpha$* , and that the strategy *witnesses* the prediction.

**Lemma 1.** *If the GEMH predicts  $A_1$  at level  $\alpha_1$  and  $A_2$  at level  $\alpha_2$ , then it predicts  $A_1 \cap A_2$  at level  $\alpha_1 + \alpha_2$ .*

*Proof.* Given a strategy witnessing  $A_1$  at level  $\alpha_1$  and another witnessing  $A_2$  at level  $\alpha_2$ , we construct a strategy that witnesses  $A_1 \cap A_2$  at level  $\alpha_1 + \alpha_2$ . This strategy splits Speculator's initial capital into two accounts, one with  $\alpha_1/(\alpha_1 + \alpha_2)$  of it and the other with the remaining  $\alpha_2/(\alpha_1 + \alpha_2)$ ; the first account is played following the strategy that witnesses  $A_1$  at level  $\alpha_1$ , while the second is played following the strategy that witnesses  $A_2$  at level  $\alpha_2$ . The final capital for Speculator is then

$$\frac{\alpha_1}{\alpha_1 + \alpha_2} \mathcal{K}_{1,N} + \frac{\alpha_2}{\alpha_1 + \alpha_2} \mathcal{K}_{2,N}, \quad (4)$$

where  $\mathcal{K}_{1,N}$  is the final capital for the first strategy and  $\mathcal{K}_{2,N}$  is the final capital for the second. If  $A_1 \cap A_2$  fails, then either  $A_1$  fails and so  $\mathcal{K}_{1,N} \geq 1/\alpha_1$

and hence (4) is greater than or equal to  $1/(\alpha_1 + \alpha_2)$ , or else  $A_2$  fails and so  $\mathcal{K}_{2,N} \geq 1/\alpha_2$  and again (4) is greater than or equal to  $1/(\alpha_1 + \alpha_2)$ .  $\square$

This result holds for all protocols considered in this paper and for all protocols considered by Shafer and Vovk (2001).

## 2.2 Lead-lags measured with respect to $m$ should be small.

Let us add to the Basic Market Protocol the requirement that Market obey

$$|x_{n-1}y_n| < \frac{1}{C} \quad (5)$$

for  $n = 1, \dots, N$ . This will allow Speculator to do limited short-selling without risking more than his initial capital.

The inequality that we will now derive involves the function  $\Delta : (-1, \infty) \rightarrow [0, \infty)$  defined by  $\Delta(u) := u - \ln(1 + u)$ . We use two properties of  $\Delta(u)$ :

- When  $u$  is small in absolute value,  $\Delta(u)$  is also small. In fact, it is of order  $u^2$ .
- $\Delta(u) \leq \Delta(-u)$  for  $u \in [0, 1)$ . To verify this, note that  $1/(1 + u) + 1/(1 - u) \geq 2$  for  $u \in [0, 1)$  and then integrate both sides to obtain  $\ln(1 + u) - \ln(1 - u) \geq 2u$ .

**Proposition 1.** Suppose the returns  $y_0, y_1, \dots, y_N$  are measured with respect to  $m$ , and suppose Market is required to obey (5) with  $C > 0$ . Then, for any significance level  $\alpha > 0$ , the GEMH for  $m$  predicts

$$|\text{eff}(x, y)| < \frac{1}{C} \ln \frac{2}{\alpha} + \frac{1}{C} \sum_{n=1}^N \Delta(-C|x_{n-1}y_n|) \quad (6)$$

at level  $\alpha$ . We also have tighter one-sided predictions: the GEMH for  $m$  predicts

$$\text{eff}(x, y) < \frac{1}{C} \ln \frac{1}{\alpha} + \frac{1}{C} \sum_{n=1}^N \Delta(Cx_{n-1}y_n) \quad (7)$$

at level  $\alpha$  and predicts

$$\text{eff}(x, y) > -\frac{1}{C} \ln \frac{1}{\alpha} - \frac{1}{C} \sum_{n=1}^N \Delta(-Cx_{n-1}y_n) \quad (8)$$

at level  $\alpha$ .

*Proof.* To prove (7), consider the momentum strategy for Speculator that sets  $Y_n$  equal to  $Cx_{n-1}$ . The capital process for this strategy,  $\mathcal{K}_n = \prod_{n=1}^N (1 + Cx_{n-1}y_n)$ , is nonnegative by (5). So the strategy witnesses

$$\prod_{n=1}^N (1 + Cx_{n-1}y_n) < \frac{1}{\alpha} \quad (9)$$

at level  $\alpha$ . The constraint (5) also guarantees that  $\Delta(Cy_{n-1}y_n)$  is defined. So we can take the logarithm of both sides of (9) and then substitute  $u - \Delta(u)$  for  $\ln(1 + u)$ , obtaining  $\sum_{n=1}^N Cx_{n-1}y_n < \ln(1/\alpha) + \sum_{n=1}^N \Delta(Cx_{n-1}y_n)$ . Dividing by  $C$ , we obtain (7).

Similarly, the contrarian strategy that sets  $Y_n$  equal to  $-Cx_{n-1}$  witnesses

$$\prod_{n=1}^N (1 - Cx_{n-1}y_n) < \frac{1}{\alpha}$$

at level  $\alpha$ . Again taking the logarithm of both sides, substituting  $u - \Delta(-u)$  for  $\ln(1 - u)$ , and dividing by  $C$ , we obtain (8).

Because  $\Delta(u) \leq \Delta(-u)$ , we can substitute  $\Delta(-C|x_{n-1}y_n|)$  for  $\Delta(Cx_{n-1}y_n)$  in (7) and for  $\Delta(-Cx_{n-1}y_n)$  in (8). Combining these two level  $\alpha$  results using Lemma 1, we obtain

$$|\text{eff}(x, y)| < \frac{1}{C} \ln \frac{1}{\alpha} + \frac{1}{C} \sum_{n=1}^N \Delta(-C|x_{n-1}y_n|)$$

at level  $2\alpha$ . Substituting  $\alpha/2$  for  $\alpha$ , we obtain (6) at level  $\alpha$ .  $\square$

Suppose Speculator trades over a total time  $T$ , divided into trading periods of length  $dt$ , so that  $N = T/(dt)$ . Then under the assumption that the  $x_n$  and the  $y_n$  both scale like  $\sqrt{dt}$ , the qualitative message of the inequality (6) is that

$$|\text{eff}(x, y)| \lesssim \mathbf{a} + \mathbf{b}Tdt, \quad (10)$$

where  $\mathbf{a}$  and  $\mathbf{b}$  are constant with respect to  $T$  and  $dt$ . This is because the term  $(1/C) \ln(2/\alpha)$  does not depend on  $T$  or  $dt$ , while the term

$$(1/C) \sum_{n=1}^N \Delta(-C|x_{n-1}y_n|),$$

because  $\Delta(u)$  is of order  $u^2$ , is of order  $(1/C)N(Cdt)^2 = CTdt$ . On the same principles, we expect the denominator of (1) to grow linearly with  $T$ , and so (10) suggests that

$$|\text{corr}(x, y)| \lesssim \frac{\mathbf{d}}{T} + \mathbf{e}dt,$$

where  $\mathbf{d}$  and  $\mathbf{e}$  are constant with respect to  $T$  and  $dt$ . This confirms our intuition that  $\text{corr}(x, y)$  should be close to zero when  $T$  is large and  $dt$  is small.

### 2.3 Lead-lags not measured with respect to $m$ should be similar.

We now turn to a second implication of the GEMH for  $m$ : when lead-lag effects are measured with respect to a numeraire different from  $m$ , a signal  $x$  will have approximately the same lead-lag effect on a security  $y$  as it has on  $m$ .

To fix ideas, we may suppose that  $m$  is a market index, but that we measure capital and the returns  $x_n$ ,  $y_n$ , and  $m_n$  with respect to cash. In any case, we can formulate the protocol as follows:

PROTOCOL FOR TRADING IN  $y$  AND  $m$  WITH A SIGNAL  $x$

$$\mathcal{K}_0 := 1.$$

Market announces  $x_0$ ,  $y_0$ , and  $m_0$ , all in  $(-1, \infty)$ .

FOR  $n = 1, 2, \dots, N$ :

Speculator announces  $Y_n \in \mathbb{R}$ .

Market announces  $x_n$ ,  $y_n$ , and  $m_n$ , all in  $(-1, \infty)$ .

$$\mathcal{K}_n := \mathcal{K}_{n-1} (1 + m_n + Y_n(y_n - m_n)).$$

If  $\alpha > 0$ ,  $A$  is a statement about Reality's moves  $x_0, y_0, m_0, \dots, x_N, y_N, m_N$ , and Speculator has a strategy that guarantees both (1)  $\mathcal{K}_n \geq 0$  for  $n = 1, \dots, N$  and (2) either  $A$  happens or  $\mathcal{K}_N \geq (1/\alpha) \prod_{n=1}^N (1 + m_n)$ , then we say that the *GEMH relative to  $m$  predicts  $A$  at level  $\alpha$* , and that the strategy *witnesses* the prediction.

We use the function  $\Psi(m, t)$  defined by  $\Psi(m, t) := t - \ln(1 + t/(1 + m))$ . If  $t$  is very small, and  $m$  is of order  $t^{1/2}$ , then  $\Psi(m, t)$  will be of order  $t^{3/2}$ . We set

$$y_n^\dagger = \frac{y_n - m_n}{1 + m_n}. \quad (11)$$

This is  $y$ 's return with respect to  $m$ —i.e.,  $y$ 's return when the price of  $y$  is measured using  $m$  as the numeraire.

**Proposition 2.** *Suppose Market is required to obey*

$$|x_{n-1} y_n^\dagger| < \frac{1}{C} \quad (12)$$

*for some constant  $C > 0$ . Then, for any significance level  $\alpha \in (0, 1]$ , the GEMH for  $m$  predicts*

$$\text{eff}(x, y) - \text{eff}(x, m) < \frac{1}{C} \ln \frac{1}{\alpha} + \frac{1}{C} \sum_{n=1}^N \Psi(m_n, Cx_{n-1}(y_n - m_n)) \quad (13)$$

*at level  $\alpha$  and predicts*

$$\text{eff}(x, m) - \text{eff}(x, y) < \frac{1}{C} \ln \frac{1}{\alpha} + \frac{1}{C} \sum_{n=1}^N \Psi(m_n, -Cx_{n-1}(y_n - m_n)) \quad (14)$$

*at level  $\alpha$ .*

*Proof.* Consider the strategy for Speculator that sets  $Y_n$  equal to  $Cx_{n-1}$ . It witnesses

$$\prod_{n=1}^N (1 + m_n + Cx_{n-1}(y_n - m_n)) < \frac{1}{\alpha} \prod_{n=1}^N (1 + m_n) \quad (15)$$

at level  $\alpha$ . Using (11), we can rewrite condition (12) as

$$|Cx_{n-1}(y_n - m_n)| < 1 + m_n,$$

which implies that each factor on the left-hand side of (15) is positive. So we can take logarithms of both sides, obtaining

$$\sum_{n=1}^N \ln \left( \frac{1 + m_n + Cx_{n-1}(y_n - m_n)}{1 + m_n} \right) < \ln \frac{1}{\alpha}.$$

Because  $\ln((1 + m + t)/(1 + m)) = t - \Psi(m, t)$ , we can rewrite this as

$$\sum_{n=1}^N (Cx_{n-1}(y_n - m_n) - \Psi(m_n, Cx_{n-1}(y_n - m_n))) < \ln \frac{1}{\alpha}.$$

Regrouping and dividing by  $C$ , we obtain (13).

The strategy for Speculator that invests the fraction  $-Cx_{n-1}$  of his capital in  $y$  witnesses (15) with  $C$  replaced by  $-C$ , analogously leading to (14).  $\square$

By averaging the strategies that produce (13) and (14), we can obtain a two-sided inequality on  $\text{eff}(x, y) - \text{eff}(x, m)$ , analogous to the two-sided bound (6) on  $\text{eff}(y, y)$ . We can also obtain bounds on the difference between  $x$ 's lead-lag effect on a security  $y$  and its lead-lag effect a security  $z$ . We omit the details.

When the trading periods are short, so that the returns  $m_n$  are uniformly very small, the quantity  $\text{eff}(x, y) - \text{eff}(x, m)$  bounded by Proposition 2 will approximate  $\text{eff}(x, y^\dagger)$ . This is because  $\text{eff}(x, y) - \text{eff}(x, m) = \sum x_{n-1}(y_n - m_n)$ , while  $\text{eff}(x, y^\dagger) = \sum x_{n-1}(y_n - m_n)/(1 + m_n)$ . But Proposition 1 already told us that  $\text{eff}(x, y^\dagger)$  will be small under the GEMH for  $m$ . So Proposition 2 may be most interesting in cases of monthly or annual returns.

### 3 Empirical Results

In this section, we use the predictions derived in §2 to test two GEMHs for United States stock markets: (1) The *cash GEMH*, which says that a speculator will not multiply the capital he risks by a large factor when capital is measured in cash. (2) The *S&P 500 GEMH*, which says he will not do so when capital is measured relative to the S&P 500 index. We test the cash GEMH because tests of the EMH usually use cash as numeraire. We test the S&P 500 GEMH because efficiency relative to the S&P 500 seems to be the most practicable concrete interpretation of efficiency “relative to the market”; the S&P 500 index is tracked by many mutual and exchange-traded funds, and its 500 stocks account for 70% of the capitalization of publicly traded corporations in the United States. We test the two GEMHs using the predictions about  $\text{eff}(x, y)$  we derived in §2.2 and the predictions about  $\text{eff}(x, y) - \text{eff}(x, z)$  we derived in §2.3.

We report *game-theoretic p-values*. Recall that our predictions all involve a significance level  $\alpha$ . The prediction (6), for example, is of the form

$$\text{lower bound depending on } \alpha \leq \text{eff}(x, y) \leq \text{upper bound depending on } \alpha. \quad (16)$$

The lower bound is negative, the upper bound is positive, and both are small for reasonable values of  $\alpha$ . So the inequality says that  $\text{eff}(x, y)$  will be small. The bounds tighten as we make  $\alpha$  larger, and unless  $\text{eff}(x, y)$  is exactly zero, the bounds will be violated for a sufficiently large value of  $\alpha$ . As usual in statistical testing, we write  $\hat{\alpha}$  for this value. It is the *game-theoretic p-value*.

When (6) or one of the other inequalities we derived in §2 is violated, the corresponding strategy for Speculator has produced  $\mathcal{K}_N > 1/\alpha$ , and so our test has rejected the hypothesis of efficiency. Because of the approximations used to derive the inequality there will be some values of  $\alpha$  for which the inequality is satisfied but  $\mathcal{K}_N > 1/\alpha$  still holds. So the observed value of  $1/\mathcal{K}_N$ , which we could alternatively consider the p-value, will be slightly smaller than the p-value we report. We use the value of  $\alpha$  for which the inequality is just violated instead of  $1/\mathcal{K}_N$  as our p-value because it can be considered the “observed significance” of the empirical quantity that violates the inequality, whether this quantity be an auto-effect or a difference between lead-lag effects.

Our game-theoretic p-values differ in one striking way from classical p-values. In classical testing, a p-value is a probability and is therefore always less than 1. In game-theoretic testing, the strategy for multiplying the capital risked may actually lose money, yielding a value of  $\mathcal{K}_N$  less than 1 and hence a p-value greater than 1. Aside from its indication that the auto-effect or difference between lead-lag effects is not large enough to reject the hypothesis of market efficiency, no importance should be attached to the p-value being greater than 1.

In §3.1, we describe the data we use for our tests. In §3.2, we test our two GEMHs by applying the inequalities in Proposition 1 to auto-effects; here we use individual stocks as well as indices and size portfolios. In §3.3, we use the transaction costs needed to eliminate our profits to measure the efficiency of the market in stocks of different levels of capitalization. In §3.4, we test the S&P 500 GEMH by looking at lead-lags between size portfolios.

### 3.1 Data

We use daily stock returns based on closing prices for all publicly traded firms on the NYSE/AMEX/NASDAQ daily tapes maintained by the Center for Research in Security Prices (CRSP) for the period beginning July 18, 1962, and ending December 28, 2005. Our data for market indices and exchange-traded funds (ETFs) are also obtained from CRSP. So that we can compare our results with those of authors who studied CRSP data for 1962–1994, we split the 1962–205 period into two periods, 1962–1994 and 1995–2005.

Like Lo and MacKinlay (1990b) and other authors, we use weekly returns so as to minimize microstructure effects. At the daily and intraday levels, there

is a substantial amount of nontrading and nonsynchronous trading. This effect is weaker at the weekly level; Forester and Keim (2000) estimate the likelihood for a typical stock going untraded for five consecutive days to be 0.42 percent. We measure returns from Wednesday to Wednesday. This is consistent with the literature, although some authors use Tuesdays rather than Wednesdays. Delistings, including those caused by liquidations and mergers, are common in the database, and a delisting price and return is often provided. When it is provided, we use it to adjust the last historical return.

We create five size portfolios. We change their composition each year, based on the market capitalization of stocks the preceding December. All securities in the CRSP universe are used to find breakpoints. We consider both equal-weighted returns for the portfolio, calculated by taking the simple average of the stocks in the portfolio, and value-weighted returns, calculated weighting individual returns with market capitalization at the end of the previous week.

### 3.2 Statistical significance of autocorrelations

Most individual stock returns are weakly negatively autocorrelated, while returns on some indices and on portfolios of small-cap stocks are strongly positively autocorrelated (Campbell et al. 1997). Our game-theoretic tests give the same results as conventional tests: autocorrelations for individual stocks are usually not statistically significant, while autocorrelations for portfolios and indices often are.

We report uncentered autocorrelations, obtained by dividing the auto-effect  $\text{eff}(y, y)$  by the denominator of (3):

$$\text{uncentered autocorrelation} := \frac{\sum_{n=1}^N y_{n-1} y_n}{\sqrt{\sum_{n=1}^N (y_{n-1} - \bar{y}^-)^2} \sqrt{\sum_{n=1}^N (y_n - \bar{y}^+)^2}}. \quad (17)$$

We also report bounds on these autocorrelations obtained by dividing by the same denominator the bounds on the auto-effects given by (7) and (8) for  $\alpha = 5\%$ . We also give three game-theoretic p-values:

- The two-sided value  $\hat{\alpha}$  obtained by setting the two sides of (6) equal.
- The one-sided value  $\hat{\alpha}_m$  obtained by setting the two sides of (7) equal.
- The one-sided value  $\hat{\alpha}_c$  obtained by setting the two sides of (8) equal.

In our tables, two asterisks on an autocorrelation indicate rejection at the 5% level by the two-sided test; one asterisk indicates rejection at the 5% level by a one-sided test but not by the two-sided test.

In the case where  $x_n = y_n$ , Proposition 1 assumes that  $|y_{n-1} y_n| \leq 1/C$  for  $n = 1, \dots, N$ . Subject to this condition, we want to choose  $C$  to make our inequalities as tight as possible, thus making the game-theoretic p-values as

Table 1: Autocorrelations of weekly returns for 423 individual stocks, 1962–1994, tested under the cash GEMH

	Game-theoretic p-values				
	Mean	Median	St. Dev.	Min	Max
$\hat{\alpha}$	1.5	1.5	0.77	0.065	12
$\hat{\alpha}_m$	3.5	1.4	12	0.24	159
$\hat{\alpha}_c$	1.0	0.85	0.87	0.031	8.7
<i>Uncentered autocorrelations, with predicted bounds, <math>\alpha = 5\%</math>, <math>C = 5</math></i>					
	Mean	Median	St. Dev.	Min	Max
autocorrelation	-0.02	-0.02	0.04	-0.14	0.13
lower bound	-0.27	-0.23	0.15	-0.83	-0.080
upper bound	0.27	0.23	0.15	0.080	0.83

small as possible. To gain insight about the choice of  $C$ , consider the equation

$$\ln \hat{\alpha} = \ln 2 - \underbrace{\left| C \sum_{n=1}^N y_{n-1} y_n \right|}_{a} + \underbrace{\sum_{n=1}^N \Delta(-C |y_{n-1} y_n|)}_{b}$$

obtained by setting the two sides of (6) equal. The term  $a$  is linear in  $C$ , while the term  $b$  is approximately quadratic in  $C$ . When  $C$  is relatively small,  $a > b$ . But as  $C$  grows,  $b$  catches up and exceeds  $a$ . So as  $C$  increases,  $\hat{\alpha}$  first declines and then rises.

**Individual stocks.** For the 10% most volatile stocks,  $C$  cannot be chosen large enough to give interesting bounds. The remaining 90% satisfy  $|y_{n-1} y_n| \leq 1/C$  for  $C = 5$ . But few of them have significant auto-effects.

Altogether, 10,869 different stocks appear in the CRSP database during the period July 18, 1962, to December 28, 1994. But only 470 stocks were listed for the entire period, with no consecutive missing weekly returns. Dropping the 10% most volatile, we are left with 423 stocks. Table 1 summarizes the autocorrelations for these 423 stocks when returns are measured relative to cash, as well as the 5% bounds from (7) and (8), together with corresponding game-theoretic p-values for the cash GEMH. The autocorrelations are generally negative, close to zero, and not statistically significant. None of the two-sided p-values are less than 5%, and only four of the one-sided p-values (both corresponding to contrarian strategies) are less than 5%.

Results are similar for (1) the smaller set of stocks that had strictly no missing weekly returns for 1962–1994, (2) the S&P 500 GEMH for 1962–1994, and (3) both GEMHs for 1995–2005. In all three cases, very few autocorrelations for individual stocks are significant.

**Indices and size portfolios.** We also consider autocorrelations for three indices (CRSP equal-weighted, CRSP value-weighted, and the S&P 500) and five size portfolios. Our results under the cash GEMH (Tables 2, 3, and 4) are again consistent with results from conventional tests: the autocorrelations are

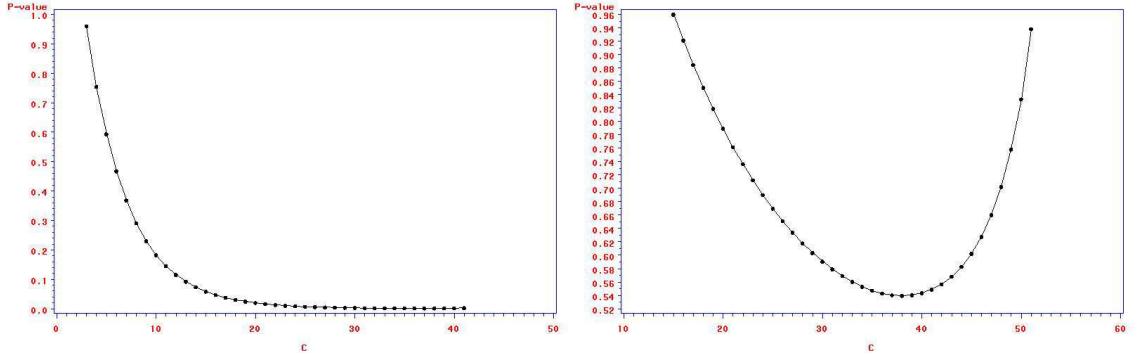


Figure 1: Plot of  $\hat{\alpha}$  and  $C$  for the cash GEHM: equal-weighted CRSP on left; value-weighted on right

statistically significant in the case of the CRSP equal-weighted index and the smaller size portfolios, with the greatest significance observed for the smallest size portfolios. The pattern of statistical significance for size portfolios is similar under the S&P 500 GEMH (Table 5).

The graphs in Figure 1 show  $\hat{\alpha}$  for the cash GEMH as a function of  $C$  for the CRSP equal-weighted and value-weighted indices for the 1962–1994 period. Each graph extends from the point where  $\hat{\alpha}$  dips below 1 to the largest permitted value of  $C$ ,  $1/\max|y_{n-1}y_n|$ . The graphs suggest that  $C = 40$  is reasonable for both indices. For the S&P 500 index,  $C = 40$  is again permitted, but the graph (not shown) is less interesting, because no value of  $C$  approaches significance. So we use  $C = 40$  for all three indices for 1962–1994. For 1995–2005, we choose  $C = 148$  based on a similar analysis. The indices have smaller extreme returns in this period.

Table 2 reports autocorrelations and game-theoretic p-values under the cash GEMH for the three indices. Only the autocorrelation of the CRSP equal-weighted index is significant. These results are consistent with results from conventional tests. Recall that the game-theoretic p-values can be interpreted in terms of the capital for the corresponding strategy in the absence of transaction costs. For example, the p-value  $\hat{\alpha}_m = 0.00013$  for the CRSP equal-weighted index indicates that our momentum strategy for this index would have turned \$1 into at least  $\$(1/0.00013) = \$7692$  between 1962 and 1994. Because of the approximation involved in deriving the inequality (7), this is an understatement; as the panel of the table labelled “Final capital” states, this momentum strategy actually produces a final capital of \$7830.92. Holding \$1 in the index, in contrast, would have produced only \$445.

Although the autocorrelation of returns for the CRSP equal-weighted index is smaller in 1995–2005 than in 1962–1994 (0.29 as opposed to 0.35), Table 2 shows it to be slightly more significant in the second period. This is likely due to

Table 2: Autocorrelations of weekly returns for indices, tested under the cash GEMH

1962–1994	<b>CRSP Equal-Weighted</b>	<b>CRSP Value-Weighted</b>	<b>S&amp;P 500</b>
<i>Game-theoretic p-values, C = 40</i>			
$\hat{\alpha}$	1.7E-3	0.54	1.7
$\hat{\alpha}_m$	1.3E-4	0.19	0.70
$\hat{\alpha}_c$	2.7E5	22	4.5
<i>Uncentered autocorrelations, with predicted bounds, <math>\alpha = 5\%</math>, C = 40</i>			
autocorrelation	0.35**	0.08	0.03
lower bound	-0.20	-0.15	-0.13
upper bound	0.14	0.14	0.13
<i>Final capital for different strategies starting with \$1</i>			
momentum	\$7830.92	\$5.33	\$1.43
contrarian	\$0.00	\$0.05	\$0.22
average	\$3915.46	\$2.69	\$0.82
hold index	\$445.22	\$26.50	\$7.98
1995–2005	<b>CRSP Equal-Weighted</b>	<b>CRSP Value-Weighted</b>	<b>S&amp;P 500</b>
<i>Game-theoretic p-values, C = 148</i>			
$\hat{\alpha}$	4.9E-4	26	6.8
$\hat{\alpha}_m$	1.1E-4	1.3E3	3.2E3
$\hat{\alpha}_c$	1.5E6	1.2	0.33
<i>Uncentered autocorrelations, with predicted bounds, <math>\alpha = 5\%</math>, C = 148</i>			
autocorrelation	0.29**	-0.06	-0.09
lower bound	-0.14	-0.13	-0.13
upper bound	0.13	0.17	0.17
<i>Final capital for different strategies starting with \$1</i>			
momentum	\$12918.09	\$0.00	\$0.00
contrarian	\$0.00	\$0.69	\$2.69
average	\$6459.05	\$0.35	\$1.35
hold index	\$15.56	\$3.36	\$2.73

Table 3: Autocorrelations of weekly equal-weighted returns for size portfolios, tested under the cash GEMH

1962–1994	<b>Small 1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>Big 5</b>
<i>Game-theoretic p-values, C = 32</i>					
$\hat{\alpha}$	7.0E-6	1.7E-3	0.018	0.041	0.15
$\hat{\alpha}_m$	3.0E-6	1.5E-4	6.7E-4	5.1E-3	0.055
$\hat{\alpha}_c$	2.1E6	1.8E5	9.7E4	3.3E3	61
<i>Uncentered autocorrelations, with predicted bounds, <math>\alpha = 5\%</math>, C = 32</i>					
autocorrelation	0.51**	0.37**	0.32**	0.25**	0.15
lower bound	-0.15	-0.21	-0.25	-0.21	-0.17
upper bound	0.14	0.15	0.15	0.16	0.16
1995–2005	<b>Small 1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>Big 5</b>
<i>Game-theoretic p-values, C = 79</i>					
$\hat{\alpha}$	0.018	7.3E-3	0.066	1.3	4.9
$\hat{\alpha}_m$	2.0E-5	3.4E-3	0.030	0.60	7.8
$\hat{\alpha}_c$	1.8E8	6.3E2	91	8.3	1.6
<i>Uncentered autocorrelations, with predicted bounds, <math>\alpha = 5\%</math>, C = 79</i>					
autocorrelation	0.56**	0.39**	0.21*	0.06	-0.03
lower bound	-0.48	-0.22	-0.18	-0.16	-0.17
upper bound	0.19	0.22	0.18	0.16	0.18

Table 4: Autocorrelations of weekly value-weighted returns for five size portfolios, tested under the cash GEMH

1962–1994	<b>Small 1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>Big 5</b>
<i>Game-theoretic p-values, C = 30</i>					
$\hat{\alpha}$	9.0E-5	0.016	0.82	0.11	0.82
$\hat{\alpha}_m$	2.0E-5	4.7E-4	1.7E-3	0.011	0.38
$\hat{\alpha}_c$	3.7E5	1.5E5	5.9E5	1.7E3	4.8
<i>Uncentered autocorrelations, with predicted bounds, <math>\alpha = 5\%</math>, C = 30</i>					
autocorrelation	0.43**	0.34**	0.30*	0.23*	0.07
lower bound	-0.17	-0.26	-0.38	-0.23	-0.18
upper bound	0.14	0.15	0.16	0.16	0.18
1995–2005	<b>Small 1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>Big 5</b>
<i>Game-theoretic p-values, C = 108</i>					
$\hat{\alpha}$	6.2E-4	8.5E-4	0.031	4.0	0.97
$\hat{\alpha}_m$	4.0E-6	3.2E-4	0.012	1.6	1.3E2
$\hat{\alpha}_c$	4.3E8	2.7E4	633	8.6	0.25
<i>Uncentered autocorrelations, with predicted bounds, <math>\alpha = 5\%</math>, C = 108</i>					
autocorrelation	0.46**	0.35**	0.18**	0.03	-0.09
lower bound	-0.29	-0.16	-0.14	-0.14	-0.14
upper bound	0.15	0.15	0.14	0.14	0.16

the power of the test; the smaller number of weeks in the second period makes the test less powerful, but larger value of  $C$  works in the opposite direction, perhaps more strongly.

Table 3 reports results for equal-weighted returns from our five size portfolios. The smaller portfolios have significant positive correlations, with the p-value  $\hat{\alpha}$  increasing monotonically with the size of the portfolio. Table 4, for value-weighted returns, shows a similar pattern. This is again consistent with conventional tests.

The most important point is that the S&P 500 index, which is now traded and could have been replicated with low transaction costs even in the earlier period, does not show statistically significant autocorrelation under the cash GEMH. This is also true for other traded indices, including the Rydex S&P Equal-Weighted ETF (RSP), which consists of large S&P 500 stocks, and iShares S&P SmallCap 600 Index Fund (IJR) (details not reported here).

Table 5: Autocorrelations of weekly returns for size portfolios, tested under the S&P 500 GEMH

1962–1994	<b>Small 1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>Big 5</b>
<i>Equal-weighted, C = 150</i>					
autocorrelation	0.30	0.22	0.19	0.13	0.10
$\hat{\alpha}$	2.2E-8	1.9E-5	1.2E-3	0.076	0.89
$\hat{\alpha}_m$	5.8E-11	5.8E-6	4.9E-4	0.037	0.44
$\hat{\alpha}_c$	6.8E16	1.2E7	1.6E4	70	2.5
<i>Value-weighted, C = 200</i>					
autocorrelation	0.28	0.21	0.18	0.12	0.03
$\hat{\alpha}$	1.4E-8	5.4E-6	6.4E-5	0.077	1.9
$\hat{\alpha}_m$	2.7E-12	9.6E-7	1.7E-5	0.032	0.94
$\hat{\alpha}_c$	6.6E20	8.2E8	1.9E5	168	1.1
1995–2005	<b>Small 1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>Big 5</b>
<i>Equal-weighted, C = 64</i>					
autocorrelation	0.22	0.10	0.11	0.04	0.07
$\hat{\alpha}$	2.3	0.76	0.86	1.6	1.7
$\hat{\alpha}_m$	0.046	0.37	0.43	0.79	0.85
$\hat{\alpha}_c$	3.3E3	4.2	2.9	1.5	1.2
<i>Value-weighted, C = 87</i>					
autocorrelation	0.22	0.19	0.16	0.06	0.02
$\hat{\alpha}$	2.8	0.16	0.33	1.3	2.0
$\hat{\alpha}_m$	0.018	0.066	0.16	0.66	0.99
$\hat{\alpha}_c$	4.2E4	62	11	1.9	1.0

Table 5 reports on autocorrelations for the five size portfolios when returns are measured with respect to the S&P 500 index and tested under the S&P 500 GEMH. The correlations are smaller than for the cash GEMH, but they still tend to be statistically significant for the smaller size portfolios. Again,

the degree of the statistical significance depends not so much on the size of the autocorrelation as on the power of the test. Our tests are more powerful when there are more weeks in the period (1962–1994 is about three times as long as 1995–2005) and when extreme values of  $|y_{n-1}y_n|$  are less extreme, so that we can choose  $C$  larger. The extremes are greater in 1962–1994 for returns with respect to cash, but greater in 1995–2005 for returns with respect to the S&P 500.

### 3.3 Relative efficiency and economic significance

It is widely understood that the statistical significance of an autocorrelation does not measure its economic significance. We cannot say that one market is less efficient than another because the autocorrelation of its returns are more statistically significant. Nor is it clear that the magnitude of an autocorrelation of returns measures inefficiency. As we noted in the introduction, the absence of a convincing measure of relative efficiency has frustrated economists.

In our game-theoretic approach, however, there is a natural way to measure the inefficiency of a market. An autocorrelation is statistically significant in this approach when its size permits a speculator, in the absence of transaction costs, to multiply the capital he risks by 20 or more. It is natural to measure inefficiency by the level of transaction costs required to eliminate this money-making opportunity.

Consider, for example, the effect of transaction costs on Speculator's capital in the Basic Market Protocol. At the beginning of period  $n$ , Speculator plans to invest  $\mathcal{K}_{n-1}Y_n$  in the security  $y$ . But he already has  $\mathcal{K}_{n-2}Y_{n-1}(1 + y_{n-1})$  invested in  $y$ . So the amount of capital he shifts between  $y$  and  $m$  at the beginning of period  $n$  is  $|\mathcal{K}_{n-1}Y_n - \mathcal{K}_{n-2}Y_{n-1}(1 + y_{n-1})|$ . If we assume that the cost of the shift is a fraction  $\tau$  of the amount shifted, then his capital at the end of the period is given by

$$\mathcal{K}_n := \mathcal{K}_{n-1}(1 + Y_n y_n) - |\mathcal{K}_{n-1}Y_n - \mathcal{K}_{n-2}Y_{n-1}(1 + y_{n-1})|\tau \quad (18)$$

instead of  $\mathcal{K}_n := \mathcal{K}_{n-1}(1 + Y_n y_n)$ . The economic significance of an autocorrelation that is statistically significant at the 5% level can be measured by the value of  $\tau$  that reduces  $\mathcal{K}_N$ , determined recursively by  $\mathcal{K}_0 = 1$  and (18), to 20 or even to 1. Let us write  $\tau_{20}$  for the value of  $\tau$  that reduces  $\mathcal{K}_N$  to 20, and  $\tau_1$  for the value that reduces it to 1. We may call  $\tau_1$  the *economic significance* of the result of our game-theoretic statistical test.

Table 6 reports in basis points implied transaction costs for our momentum strategy for equal-weighted and value-weighted size portfolios during the 1962–1994 period. Table 7 reports implied transaction costs for CRSP equal-weighted index in the same fashion. All the transaction costs in these tables are lower than rates documented in the finance literature. For example, Lesmond et al. (1999) report that average round-trip transaction costs from 1963 to 1990 were 103 and 12 basis points for small and large decile firms, respectively. Roll's measure of bid-ask spread is 434 and 31 basis points for small and large decile

Table 6: Implied transaction costs for weekly returns of size portfolios, 1962–1994

Cash EMH	<b>Small 1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>Big 5</b>
<i>Equal-weighted, C = 32</i>					
autocorrelation	0.51	0.37	0.32	0.25	0.15
$\tau_{20}$	119	62	44	22	0
$\tau_1$	146	87	67	45	21
<i>Value-weighted, C = 30</i>					
autocorrelation	0.43	0.34	0.30	0.23	0.07
$\tau_{20}$	90	52	36	16	0
$\tau_1$	117	77	60	40	3
S&P 500 EMH	<b>Small 1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>Big 5</b>
<i>Equal-weighted, C = 150</i>					
autocorrelation	0.30	0.22	0.19	0.13	0.10
$\tau_{20}$	45	24	15	1	0
$\tau_1$	50	31	22	10	1
<i>Value-weighted, C = 200</i>					
autocorrelation	0.28	0.21	0.18	0.12	0.03
$\tau_{20}$	40	22	13	1	0
$\tau_1$	45	27	20	8	0

Table 7: Implied transaction costs for weekly returns of CRSP equal-weighted index, cash EMH

	<b>1962–1994</b>	<b>1995–2005</b>
	$C = 40$	$C = 148$
autocorrelation	0.35	0.29
$\tau_{20}$	52	36
$\tau_1$	73	49

firms, respectively. We may conclude that the statistical significance of the autocorrelations of returns for size portfolios is fully accounted for by transaction costs.

The implied transaction costs in the table are approximately linearly related to the autocorrelations, suggesting that the autocorrelations can also serve as a crude but useful measure of relative efficiency. P-values, in contrast, are not a meaningful measure of relative efficiency, because they depend so much on the power of the test.

Tables 8 and 9 give values of  $\tau_1$  for tests of lead-lag effects  $\text{eff}(x : y)$  and  $\text{eff}(x^\dagger : y^\dagger)$  for our five size portfolios. When  $x = y$ , these tests reduce to the tests of  $\text{eff}(y : y)$  and  $\text{eff}(y^\dagger : y^\dagger)$  that we have already considered, and for this reason, the main diagonals of these tables match the rows of Table 6 that give  $\tau_1$  for value-weighted returns under the cash EMH and the S&P 500 EMH,

respectively.

Table 8: Values of  $\tau_1$  for tests of the lead-lag effect  $\text{eff}(x : y)$  with  $C = 30$ , as  $x$  and  $y$  range over our five size portfolios, with value-weighted returns, 1962–1994

	$y = 1$	$y = 2$	$y = 3$	$y = 4$	$y = 5$
$x = 1$	117	84	61	37	1
$x = 2$	103	77	58	37	1
$x = 3$	97	77	60	41	4
$x = 4$	85	71	57	40	4
$x = 5$	58	51	44	34	3

Table 9: Values of  $\tau_1$  for tests of the lead-lag effect  $\text{eff}(x^\dagger : y^\dagger)$  with  $C = 200$ , as  $x$  and  $y$  range over our five size portfolios, with value-weighted returns, 1962–1994

	$y = 1$	$y = 2$	$y = 3$	$y = 4$	$y = 5$
$x = 1$	45	24	12	2	0
$x = 2$	47	27	15	4	0
$x = 3$	50	33	20	8	0
$x = 4$	43	30	19	8	0
$x = 5$	0	0	0	0	0

### 3.4 Statistical significance of lead-lag differences

The S&P 500 GEMH does not directly rule out a signal observed at time  $n - 1$  helping predict the change in the S&P 500 index from  $n - 1$  to  $n$ . It says only that such a signal should not help us make more money than we could make by holding the S&P 500 index. As we learned from Proposition 2 in §2.3, this implies that the signal should have approximately the same effect on every stock or portfolio as it has on the S&P 500 index itself.

The most widely discussed lead-lag effects are those for size portfolios: large stocks lead small stocks. So we use our five size portfolios to test Proposition 2. We use value-weighted returns, measured with respect to cash, for the period from 1962 to 1994. Table 10 reports  $\text{eff}(x, y)$  for 1962–1994 as  $x$  and  $y$  both range over our five size portfolios and three indices. Here we designate the size portfolios, from smallest to largest, as 1, 2, 3, 4, and 5, the CRSP equal-weighted and value-weighted indices as  $e$  and  $v$ , and the S&P 500 index as  $m$ . For example, the last entry in the top row, 0.013, for example, is  $\text{eff}(1, m)$ , where 1 is the smallest size portfolio, and  $m$  is the S&P 500 index.

The lead-lag effects are roughly proportional, by (1), to the corresponding correlations, and as it happens they even have the same order of magnitude. So their most salient properties are those that have been noted in the literature for correlations of returns for size portfolios: (1) lead-lag effects involving the same portfolio being led (in the same column) are about the same, and (2)

Table 10: Values of  $\text{eff}(x, y)$  as  $x$  and  $y$  range over our five size portfolios, with value-weighted returns, 1962–1994

	$1_n$	$2_n$	$3_n$	$4_n$	$5_n$	$e_n$	$v_n$	$m_n$
$1_{n-1}$	0.38	0.29	0.22	0.15	0.028	0.25	0.047	0.013
$2_{n-1}$	0.36	0.28	0.22	0.16	0.032	0.24	0.051	0.017
$3_{n-1}$	0.35	0.29	0.24	0.18	0.046	0.25	0.064	0.030
$4_{n-1}$	0.32	0.27	0.23	0.18	0.048	0.23	0.065	0.031
$5_{n-1}$	0.23	0.20	0.18	0.15	0.041	0.18	0.054	0.024
$e_{n-1}$	0.34	0.28	0.23	0.17	0.044	0.25	0.062	0.028
$v_{n-1}$	0.24	0.22	0.19	0.15	0.041	0.19	0.055	0.024
$m_{n-1}$	0.22	0.20	0.18	0.15	0.038	0.18	0.051	0.021

lead-lag effects are largest when the portfolio being led is smallest (is in the first column). This contrasts sharply with Proposition 2, which predicts that the lead-lags involving the same lead (in the same row) should be about the same.

Proposition 2 states two inequalities, and the one that risks being violated is (13), which says that  $\text{eff}(x, y)$  should not be too much larger than  $\text{eff}(x, m)$ . The game-theoretic p-value for testing whether  $x$  leads  $y$  too much more than it leads  $m$  is the value of  $\alpha$  that satisfies

$$\text{eff}(x, y) - \text{eff}(x, m) = \frac{1}{C} \ln \frac{1}{\alpha} + \frac{1}{C} \sum_{n=1}^N \Psi(m_n, Cx_{n-1}(y_n - m_n)).$$

These values are given in Table 11 for the five size portfolios. Except when  $y$  is the largest size portfolio (the last column), the differences  $\text{eff}(x, y) - \text{eff}(x, m)$  are highly statistically significant. The corresponding values of  $\tau_1$  are given in Table 12.

Table 11: Game-theoretic p-values for  $\text{eff}(x, y) - \text{eff}(x, m)$  with  $C = 55$ , for  $x = 1, \dots, 5$  and  $y = 1, \dots, 5$ , value-weighted returns, 1962–1994

	$y = 1$	$y = 2$	$y = 3$	$y = 4$	$y = 5$
$x = 1$	4.4E-9	5.4E-7	1.8E-5	6.8E-4	0.44
$x = 2$	2.0E-8	1.0E-6	2.0E-5	5.8E-4	0.43
$x = 3$	5.6E-8	1.2E-6	1.8E-5	4.3E-4	0.42
$x = 4$	3.5E-7	3.5E-6	2.9E-5	5.0E-4	0.40
$x = 5$	2.7E-5	7.8E-5	2.4E-4	1.5E-3	0.39

As we mentioned at the end of §2.3, there is little difference, for weekly returns, between testing  $\text{eff}(x, y) - \text{eff}(x, m)$  using Proposition 2 and testing  $\text{eff}(x, y^\dagger)$  using Proposition 1. Moreover, the momentum strategies for testing  $\text{eff}(x, y^\dagger)$  are the same as the momentum strategies for testing  $\text{eff}(x, y)$ . So it is not surprising that Table 12 is not too different from Table 8.

Table 12: Values of  $\tau_1$  for Game-theoretic p-values for  $\text{eff}(x, y) - \text{eff}(x, m)$  with  $C = 55$ , for  $x = 1, \dots, 5$  and  $y = 1, \dots, 5$ , value-weighted returns, 1962–1994

	$y = 1$	$y = 2$	$y = 3$	$y = 4$	$y = 5$
$x = 1$	133	102	79	56	15
$x = 2$	118	93	75	54	14
$x = 3$	107	88	73	54	14
$x = 4$	94	81	68	52	14
$x = 5$	67	61	55	45	13

## 4 Conclusion

Returns for portfolios that are dominated by small-cap stocks have positive autocorrelations. Conventional statistical tests, based on models that give probabilities for changes in stock prices, show that these autocorrelations are statistically significant. Such portfolios are also led by other portfolios, and this too is statistically significant under the same models.

We have shown that these results can be replicated within a framework where tests are based not on probability models but on the principle that a strategy for speculation will not multiply the capital it risks by a large factor. By tying statistical testing directly to trading, this justifies measuring the relative efficiency of a market and the economic significance of a statistically significant result by the level of transaction costs needed to eliminate the statistical significance.

The transaction costs needed to eliminate statistical significance for the autocorrelation of small-cap portfolios is modest, and that this same modest level of transaction costs also explains the leading of small-cap portfolios by other portfolios. So we may say that the statistically significant correlations are not economically significant.

We have considered only the efficiency of the market in publicly listed stocks in the United States. But our concept of relative efficiency provides a needed tool for comparing different markets, across financial instruments and across nations. It may also be useful for measuring the substantive significance of statistically significant failures of models for prediction in other domains.

## Acknowledgments

Vladimir Vovk suggested the strategies used in §2. Ren-Raw Chen and Feng Zhao gave important advice concerning related literature, data analysis, and presentation.

## References

- Bernhardt, D., R. J. Davies. 2005. Portfolio cross-autocorrelation puzzles. UIUC and Babson Working Paper.

- Boudoukh, J., M. P. Richardson, R. F. Whitelaw. 1994. A tale of three schools: Insights on autocorrelations of short-horizon stock returns. *Rev. Finan. Stud.* **7** 539–573.
- Campbell, J. Y., A. W. Lo, A. C. MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton University Press, New Jersey.
- Chordia, T., B. Swaminathan. 2000. Trading volume and cross-autocorrelations in stock returns. *J. Finance* **55** 913–935.
- Chow, K. V., K. C. Denning. 1993. A simple multiple variance ratio test. *J. Econometrics* **58** 385–401.
- Cochrane, J. H. 1988. How big is the random walk in GNP? *J. Polit. Economy* **96** 893–920.
- Conrad, J., G. Kaul. 1988. Time varying expected returns. *J. Bus.* **61** 409–425.
- Conrad, J., G. Kaul. 1989. Mean-reversion in short-horizon individual security return. *J. Finan. Econ.* **29** 365–384.
- Fama, E. F. 1965. Random walks in stock market prices. *Financial Analysts Journal* **21** 55–59.
- Fama, E. F. 1970. Efficient capital markets: A review of theory and empirical work. *J. Finance* **25** 383–417.
- Fisher, L. 1966. Some new stock market indexes. *J. Bus.* **39** 191–225.
- Forester, S. R., D. B. Keim. 2000. Direct evidence of non-trading of NYSE and AMEX stocks. D. B. Keim, W. T. Ziemba, eds., *Security Market Imperfections in Worldwide Equity Markets*. Cambridge University Press, New York.
- French, K. R., R. Roll. 1986. Stock return variances: The arrival of information and the reaction of traders. *J. Finan. Econ.* **17** 5–26.
- Hameed, A. 1997. Time-varying factors and cross-autocorrelations in short-horizon stock returns. *J. Finan. Res.* **20** 435–458.
- LeRoy, S. 1973. Risk aversion and the martingale property of stock prices. *Int. Econ. Rev.* **14** 436–446.
- Lesmond, D. A., J. P. Ogden, C. A. Trzcinka. 1999. A new estimate of transaction costs. *Rev. Finan. Stud.* **12** 1113–1141.
- Lo, A. W. 2007. Efficient markets hypothesis. L. Blume, S. Durlauf, eds., *The New Palgrave: A Dictionary of Economics*, 2nd ed. Palgrave Macmillan, London.
- Lo, A. W., A. C. MacKinlay. 1988. Stock prices do not follow random walks: Evidence from a simple specification test. *Rev. Finan. Stud.* **1** 41–66.
- Lo, A. W., A. C. MacKinlay. 1990a. An econometric analysis of nonsynchronous trading. *J. Econometrics* **45** 181–211.
- Lo, A. W., A. C. MacKinlay. 1990b. When are contrarian profits due to stock market overreaction? *Rev. Finan. Stud.* **3** 175–205.
- Lucas Jr., R. E. 1978. Asset prices in an exchange economy. *Econometrica* **46** 1429–1446.
- Malkiel, B. G. 2003. The efficient market hypothesis and its critics. *J. Econ. Perspect.* **17** 59–82.
- Roll, R. 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. *J. Finance* **39** 1127–1140.

- Samuelson, P. A. 1965. Proof that properly anticipated prices fluctuate randomly. *Ind. Manage. Rev.* **6** 41–50.
- Shafer, G., V. Vovk. 2001. *Probability and Finance: It's Only a Game*. Wiley, New York.
- Shiller, R. J. 2003. From efficient markets theory to behavioral finance. *J. Econ. Perspect.* **17** 83–104.
- Vovk, V., G. Shafer. 2003. A game-theoretic explanation of the  $\sqrt{dt}$  effect. Working Paper # 5, www.probabilityandfinance.com.
- Vovk, V., G. Shafer. 2007. The game-theoretic capital asset pricing model. To appear in the *Int. J. Approx. Reason.*. See also Working Paper # 1, www.probabilityandfinance.com.
- Whang, Y.-J., J. Kim. 2003. A multiple variance ratio test using subsampling. *Econ. Letters* **79** 225–230.
- Wright, J. H. 2000. Alternative variance ratio tests using ranks and signs. *J. Bus. Econ. Statist.* **18** 1–9.