DEAKIN
UNIVERSITY AUSTRALIA

Faculty of Business and Law
School of Accounting, Economics and Finance

School Working Paper

ECONOMICS SERIES

SWP 2014/9

Market Efficiency in Asian and Australasian Stock Markets: A Fresh Look at the Evidence

Jae Kim, Hristos Doucouliagos and T.D. Stanley

Market Efficiency in Asian and Australasian Stock Markets: A Fresh Look at the Evidence

Jae Kim,* Hristos Doucouliagos and T.D. Stanley

La Trobe University, Deakin University, and Hendrix College

Abstract

Market efficiency is an important feature of successful financial markets. The aim of this paper is to analyze the available evidence on the efficient market hypothesis (EMH). Meta-regression analysis is applied to 1,560 estimates of the Variance Ratio test of the efficiency of Asian and Australasian stock markets. We test if there is evidence of violation of the EMH and we also explain the heterogeneity in the reported test results. Our meta-regression analysis specifically accommodates the possibility of publication selection in favor of accepting the null hypothesis of market efficiency. We find that Asian stock markets are, on average, not informationally efficient. However, market efficiency has improved over time and market capitalization and economic freedom influences stock market efficiency; more developed and less regulated stock markets are more efficient.

Keywords: Random walk, meta-regression, efficient market hypothesis

JEL Codes: G10, G14

Jae H. Kim, Department of Finance, La Trobe University. J.Kim@latrobe.edu.au

Hristos Doucouliagos, Department of Economics, Deakin University, douc@deakin.edu.au

T.D. Stanley, Economics and Business Department, Hendrix College, stanley@hendrix.edu

1. Introduction

The Efficient Market Hypothesis (EMH hereafter) is one of the major theoretical predictions of finance theory. When a stock market is informationally efficient, stock prices fully reflect all available information (Fama, 1970). All freely traded stocks are then correctly priced, given expected risk and returns, as new information is instantly and fully reflected in stock prices. As a result, no arbitrage opportunities exist since excess profits cannot be made from mispriced assets. There has been considerable debate concerning the degree to which stock markets are actually efficient.[1] In light of Lo's (2005) adaptive markets hypothesis, recent studies provide empirical evidence that market efficiency is highly context dependent and dynamic; and that abnormal returns can arise from time to time in response to changing market conditions (see Neely *et al*., 2009; Lim and Brooks, 2011; and Kim *et al*., 2011). Stock market inefficiency can arise from several sources, including structural impediments such as market manipulation (Comerton-Forde and Putniņš, 2014); poor information disclosure and communication (*e.g.,* Shamsuddin and Kim, 2010; Defusco *et al.,* 2010); and market frictions such as transaction and agency costs (Shleifer and Vishny, 1997). There may be behavioral factors that limit efficiency: for example, many investors' decisions may be dominated by fear and greed and not all investors may rationally process available information (*e.g.*, Kahneman and Tversky, 2002). In addition, underdeveloped markets, particularly those in emerging markets and transitional economies, may hinder the efficient

[1] Traders and investors are interested in exploiting any inefficiency in stock market prices by profiting from predictable patterns. Regulators seek to make stock markets efficient because when prices reflect fundamental values they help to allocate new investments to their highest valued use.

communication of new information so that prices do not fully reflect all publicly available information.

Testing for market efficiency has attracted a great deal of attention in the literature. A number of statistical tests have been developed and conducted. The hypothesis has been tested for a large number of stock markets over different time periods. However, as the qualitative reviews by Park and Irwin (2007), Yen and Lee (2008), and Lim and Brooks (2009) document, the evidence is rather mixed and often conflicting. In this paper, we conduct a *quantitative* review of empirical results on stock market efficiency by employing the meta-analysis methodology. Meta-analysis is an effective way of drawing valid inferences from a diverse evidence base that reports conflicting findings (Stanley and Doucouliagos, 2012). It is capable of systematically exploring the differences in empirical findings, by identifying the degree to which they are driven by factors such as cross-country variations or time variations. Meta-analysis can also reveal how the systematic component of market efficiency changes as a function of economic fundamentals, such as the degree of market development and market liberalization. It can also isolate, from the systematic component, data snooping bias, sample selection bias, and measurement errors, which may be associated with individual empirical studies.

Our meta-analysis focuses on the empirical studies which use the variance ratio (VR) test of Lo and MacKinlay (1988), since it is the most popular test for market efficiency or return predictability, with highly desirable statistical properties (see Charles and Darne, 2009). The VR test also provides an appealing and natural measure for the degree of market efficiency or return predictability (Griffin et al., 2010). We focus on Asian and Australasian stock markets, including those of Australia, Bangladesh, China, Hong Kong, India, Indonesia, Japan, Korea, Malaysia, New Zealand, Pakistan, Singapore, Sri Lanka, Taiwan, and

Thailand.[2] This group of stock markets has diverse characteristics, showing varying degrees of development and maturity over time. Our analysis does not cover the U.S. market, since its evolution of market efficiency for the past 100 years is well-documented in the literature (e.g. Gu and Finnerty, 2002; and Kim *et al*., 2011). Our study is the first meta-analytic study of market efficiency based on empirical results from a cross-section of emerging and developed stock markets.

The main findings from our meta-analysis are summarized as follows. We find violation of the EMH for Asian and Australasian stock markets, with stock market inefficiency being higher, in general, for the countries with the least developed and more regulated stock markets. Importantly, we also find that stock market efficiency has improved over time. The paper is set out as follows. Section 2 provides a brief review of market efficiency and presents the details of the VR test. In section 3 we discuss the meta-regression methodology and its attractive features, especially in the context of testing for market efficiency. Section 4 discusses the details of the data used in the meta-analysis, and Section 5 presents the results of the meta-regression analysis. Section 6 concludes the paper.

2. Stock Market Efficiency and the Variance Ratio Test Statistic

In this section, we provide a brief review of the recent literature on stock market efficiency and present the variance ratio test statistic as a measure of the degree of market efficiency. We also discuss its usefulness and limitations as a measure of market efficiency.

2.1. Brief review of stock market efficiency

When a stock market is efficient, prices adjust instantaneously and accurately in response to new information. According to the EMH, all publicly available information is fully reflected

[2] This choice is driven by the Asian countries analyzed in the primary literature.

in stock prices, and no market participants can systematically make abnormal profits (Fama, 1970). When the information set is limited to past prices and returns, the market is said to be weak-form efficient, where the current price reflects all available information from past price history. Whether a stock market is efficient in the weak-form has been a highly contentious issue in finance. While most finance academics believe in weak-form efficiency (Doran *et al*., 2010), predictable patterns of stock returns have been widely observed. For example, Jegadeesh and Titman (1993) document strong momentum effects; and behavioral finance researchers recognize that investor behavior such as overreaction and overconfidence can cause systematic departure from efficiency (*e.g*., De Bondt and Thaler, 1985; and Barber and Ordean, 2001).

The accumulated empirical evidence is rather mixed and conflicting. In their historical survey, Yen and Lee (2008) report empirical findings in support of market efficiency in the 1960s, "mixed evidence" in the 1970s and 1980s, and "refuting evidence" in the 1990s. Park and Irwin (2007) present a survey with similar findings in the context of the profitability of technical trading rules. Harvey (1995) observes that stock returns of emerging markets are generally more predictable than those of developed markets, possibly due to their segmentation from global capital markets. In contrast, Griffin *et al.* (2010) provide evidence that stock returns in emerging markets are as unpredictable as those of advanced markets. With these highly mixed empirical results, stock market efficiency remains one of the most controversial and contested hypothesis in finance.

The EMH also has strong implications to practitioners, investors, and regulators. There is an enormous industry based on the technical analysis of stocks and commodities, with numerous trading strategies employed by traders[3]. Moreover, many mutual fund

[3] According to the bestseller, "Flash Boys," large banks are making billions by placing their program trading computers physically closer to the New York Stock Exchange's computer (Lewis, 2014).

managers try to outperform the market and claim that they have done so. The EMH has also found its way into securities litigation (e.g., Fischel, 1989 and Cornell and Rutten, 2006). These are all in large part inconsistent with the EMH. On the other hand, there are arguments that these observations against the EMH are in fact spurious. For example, Bender et al. (2013) argue that technical analysis rules might reflect imperfectly rational noise trading and stock chart patterns might just be "illusory correlations". Zhang and Jacobsen (2013), find that monthly seasonal effects may not be real, as they could be subject to data snooping bias, noise, and sample selection bias.

As a compromise between the efficient market hypothesis and its behavioral critics, Lo (2004) proposes the *adaptive markets hypothesis*. One of its implications is that market efficiency is highly context dependent and dynamic; and that abnormal returns can predictably arise from time to time due to changing market conditions. Lim and Brooks (2011) provide a comprehensive review of recent empirical studies in the weak-from efficiency of stock market, which strongly support the time-varying nature of return predictability. Kim *et al*. (2011) examine the case of the US stock market over 100 years and provide empirical evidence that return predictability changes over time depending on prevailing market and economic conditions. Based on a cross-sectional study of more than fifty stock markets, Shamsuddin and Kim (2010) find that return predictability depends on a number of measures for equity market development. Hence, the extant literature on the empirical testing of the weak-from market efficiency of stock market holds the view that the degree of market efficiency changes over time depending on economic and market conditions surrounding the market. The latter includes a range of factors such as market fundamentals, regulations, trading technologies, psychology of market participants, political landscape, and the state of the economy.

2.2 **Variance ratio as a measure of stock market efficiency**

As Grossman and Stiglitz (1980) theoretically demonstrate, a perfectly efficient market is not possible. Since Campbell et al. (1997) proposed the notion of relative efficiency, the research focus has moved to measuring the degree of relative efficiency from testing absolute market efficiency. While there are alternative measures of relative efficiency (Griffin et al., 2010), the variance ratio (VR) test of Lo and MacKinlay (1988) is the most popular and appealing, based on autocorrelations of returns[4]. While there have been a number of improvements and extensions (see Charles and Darne, 2009), the test is essentially based on the statistic as a ratio of the variance of k-period stock returns to one-period returns, which can be re-written as a function of return autocorrelation as follows:

$$V(k) = 1 + 2\sum_{j=1}^{k-1}\left(1 - \frac{j}{k}\right)\rho_j, \qquad (1)$$

where ρ_j is the autocorrelation of asset returns of order j. By construction, $V(k)$ is one plus the weighted average of autocorrelations up to order k-1, with positive and declining weights. The main attraction of the VR statistic over its alternatives is that it provides an estimate of the size of return autocorrelations, as well as its overall sign. A value of $V(k)$ greater (less) than 1 indicates the presence of overall positive (negative) autocorrelations up to the order k-1. Due to this property, $V(k)$ is widely used as a measure for the degree of market efficiency or return predictability (*e.g.,* Griffin *et al.,* 2010; Kim et al., 2011).

To evaluate market efficiency or the presence of return predictability, researchers test for H_0: $V(k) = 1$, implying that all autocorrelations (ρ_js) to order k-1 are zero. When $V(k) = 1$, stock returns are not predictable from their own past. To test H_0: $V(k) = 1$, Lo and MacKinlay (1988) propose the test statistic:

[4] Other autocorrelation-based tests include Box-Ljung type tests (*e.g.*, Escanciano and Lobato, 2009), spectral tests (*e.g.*, Escanciano and Velasco, 2006), and non-parametric tests (*e.g.*, Wright, 2000). The variance ratio test is more widely used in the empirical literature, as it possesses better small sample properties (see Charles et al., 2011).

$$M(k) = \frac{\hat{V}(k) - 1}{se(\hat{V}(k))}, \tag{2}$$

where $\hat{V}(k)$ is the sample estimator for $V(k)$ and $se(\hat{V}(k))$ represents the standard error of $\hat{V}(k)$. Note that Lo and MacKinlay (1988) propose two versions of the above test statistic. The first version (denoted $M_1(k)$) is valid when the asset returns are generated from an identical and independent distribution (i.i.d); and the second version (denoted $M_2(k)$) is valid when the asset returns follow a martingale difference sequence (MDS). This test allows for asset returns with a general form of conditional or unconditional heteroskedasticity, widely observed in asset returns. Note that the two statistics are different only in the form of standard error estimators. Lo and MacKinlay (1988) show that $M_1(k)$ asymptotically follows the standard normal distribution under i.i.d. asset returns, and so does $M_2(k)$ under the MDS.

As mentioned earlier, our meta-analysis exploits sample estimates of $\hat{V}(k)$ reported in the past empirical studies. However, we note that the VR statistic has some limitations as a measure for the degree of market inefficiency. First, it does not capture the effects of the costs associated with transactions and gathering information (Griffin *et al.* 2010; Section 6), which can be high in emerging markets and dissimilar across different international markets. Second, the measure may contain noise from market microstructure and nonsynchronous trading (Boudoukh et al., 1994). Griffin *et al.* (2010) warn that caution should be exercised when VR values are directly compared across different international markets. On these points, we argue that meta-analysis is an effective way of isolating various noises from the fundamental component of the VR estimates and facilitates their comparison in a systematic way. That is, it controls the systematic component of VR estimates over time and across different international markets, isolating the noise and non-fundamental components of VR estimates, including the effects of potential publication bias. For example, the effect of declining transaction and information costs that stock markets have experienced since the

1990s may not be captured in the VR estimates reported in individual studies; but it can be revealed in a meta-regression with the VR estimates showing a downward trend over time. The VR estimates from two dissimilar markets during different time periods may not be directly comparable, but their responses to the market fundamentals such as the market capitalization can be analyzed in the framework of the meta-analysis.

Conrad and Kaul (1988) argue that autocorrelation in stock returns represents time variation of expected returns. However, their model depends heavily on its parametric structure and their result may not be robust to different model specifications. For example, if their model is specified with a constant expected returns allowing for time-varying return predictability, the observed autocorrelation can be regarded as a reflection of market inefficiency. In addition, according to Fama and French (1988a), time-varying equilibrium of expected returns occurs over long horizons, such as 3 to 5 years. We note that the VR test is widely used as a test for short-horizon return predictability, with the value of holding period k typically set to far less than 1 year. It is well-know that the test is not suitable for long-horizon analysis, since the VR statistic is severely biased estimator for $V(k)$ and the test shows undesirable small sample properties when the value of holding period k is high (see, for example, Chen and Deo; 2006).

3. Meta-regression Methodology

The key challenge behind all empirical analyses is making valid inference. Meta-regression analysis (MRA) has been developed to meet this fundamental challenge (Stanley and Doucouliagos, 2012). We employ meta-regression to achieve three tasks: (1) to formally test the EMH, (2) to analyze the distribution of the reported estimates and identify the factors that drive heterogeneity in this literature, and (3) to identify and correct potential publication selection bias. MRA can concurrently inform on each of these dimensions.

3.1 Publication selection

There is much evidence to suggest that researchers often have a preference for reporting empirical results that conform to researcher beliefs and suppress evidence that is at odds with these preferences (Roberts and Stanley, 2005; Stanley and Doucouliagos, 2012). Reported estimates may then be a biased sample of all estimates, potentially resulting in erroneous statistical inferences. Hence, it is important to test for publication selection bias and accommodate potential bias should it be present in the research record. Following Stanley and Doucouliagos (2012), we apply the so-called Funnel-Asymmetry Test Precision-Effect Test (FAT-PET) meta-regression model:

$$VR1_{ij} = \beta_0 + \beta_1 se_{ij} + \varepsilon_{ij}, \tag{3}$$

where $VR1_{ij} = \hat{V}(k) - 1$ ($VR1_{ij}$ is the ith variance ratio minus one reported in the jth study), se_{ij} is its standard error, and ε_{ij} denotes the usual regression error term.[5] See Stanley (2008) and Stanley and Doucouliagos (2012) for further details on this model.

Stanley and Doucouliagos (2012 and 2014) suggest that a more accurate estimate of the underlying effect corrected for publication selection bias can be derived by replacing se_{ij} by se_{ij}^2. Doing so gives the precision-effect estimate with standard error (PEESE) model:

$$VR1_{ij} = \alpha_0 + \alpha_1 se_{ij}^2 + v_{ij}. \tag{4}$$

The logic behind the FAT-PET is as follows. When an empirical literature is free of publication selection bias, then the estimated effect sizes (say VR estimates) will not be correlated with their standard errors (Egger *et al.,* 1997; Stanley and Doucouliagos, 2012). In fact, the validity of regression's conventional t-test requires this independence. In contrast, if

[5] Throughout the paper we measure the effect size as VR – 1, rather than VR, which we denote as VR1. This is for expositional purposes only. The inferences are *identical* if we use VR instead of VR-1, when properly interpreted.

researchers search for estimates that are statistically significant (*e.g*. a rejection of the EMH), they will re-estimate their models until they achieve some 'acceptable' level of statistical significance (*e.g*., statistical significance at the 5% or 10% level).[6] This selection will generate a correlation between an estimated effect and its standard error and will result in a truncated or asymmetric distribution of reported VR estimates (Stanley, 2008). Hence, a test of $\beta_1 = 0$ (known as the FAT or Funnel Asymmetry Test) provides a test of the existence of asymmetry in the estimates and publication selection and the $\beta_1 se_{ij}$ term reflects the impact of publication selection bias. A test of $\beta_0 = 0$, known as the PET or Precision Effect Test, provides a test of the overall existence of market inefficiency in the research record, corrected for publication selection. Also, the estimate of α_0 provides an estimate of the degree to which markets are inefficient.[7]

Although the FAT-PET-MRA model remains the foundation of the tests investigated here, application to variance ratios introduces new challenges. In particular, the variance ratio is known to have small-sample bias (Lo and MacKinlay, 1988), and publication selection may 'go both ways.' Typically, selection bias in empirical economics is in favor of rejecting the null hypothesis of a zero effect (Card and Krueger, 1995; Doucouliagos and Stanley, 2014). In such cases, authors do not report all of the results they uncover. Rather, they select results that are consistent with their prior expectations, conventional theory, or results which they believe have a stronger chance of being published. The effect of this process is that certain findings may be suppressed while others are over-represented. Consequently, publication selection bias may tend to overstate the evidence against the EMH. This could happen if *some* researchers have a prior that the market is inefficient and experiment with

[6] In the case of the EMH, re-estimation can involve taking different samples, time periods, countries, indices, *etc*.

[7] Recall that we use VR -1. Hence, the test of the EMH is the null of $\beta_0 = 0$. If we used VR, then the test of the EMH involves the null of $\beta_0 = 1$.

their models, data, and methods to find a variance ratio that differs from 1 (either greater than 1 or less than 1). However, other researchers may believe that investors are rational and markets are efficient and dismiss some large variance ratios as faulty. In this case, evidence that rejects the null may go unreported. That is, there may be selection in this research literature in both directions: for statistical significance (rejection of EMH) and for statistical insignificance (acceptance of market efficiency). *A priori*, it is not possible to predict the net direction of this bias or whether any net bias is likely to remain. This is an empirical matter about which meta-regression analysis can inform.[8]

An additional complication arises because even in the absence of reporting or publication selection, small-sample bias will cause the effect size, *VR1* (=VR-1), to be correlated with its sample size and thereby inversely with its own standard error. The conventional FAT-PET-MRA (equation (3)) may therefore be affected by this small-sample bias.

Given the above concerns with conventional meta-analysis, we conducted simulations for meta-regression models of the variance ratio tests to accommodate possible publication selection bias. Appendix A provides details of the simulation design and results. The central purpose of these simulations is to investigate the statistical properties of PET for the application to the EMH and thereby insure the validity of our MRA methods in assessing this literature's evidence of market efficiency. Past simulations (*e.g.* Stanley 2008) have only reported the performance of these MRA models of publication bias when there are various incidences of selection for statistical significance—not when selection is for statistical *insignificance*. This too may potentially invalidate the FAT-PET-MRA by adding a yet another correlation of the reported effect to its standard error. Our simulations suggest that using a larger critical value, 3.5, instead of the normal critical value of 1.96 will

[8] Ed Tufte (2006, p.687) famously remarked that t-statistics that fall in the range between 1.6 and 2 lie in the "Zone of Boredom, Ambiguity, and Unpublishability".

accommodate both small-sample bias as well as potential selection for insignificance— see Appendix A for details.

3.2 Heterogeneity

Eqn. (3) can be extended to explain observed variation in the variance ratio test.

$$VR1_{ij} = \beta_0 + \beta_1 se_{ij} + \sum \beta_k Z_{kij} + \varepsilon_{ij}, \tag{5}$$

where $\boldsymbol{Z}_k$ represents variables coded from the studies themselves and exogenous variables on market capitalization and economic freedom coded by us from public information. Other examples of $\boldsymbol{Z}$-variables include: country, level of development, holding period, and data frequency. See Table 3 below for a full list and description.

3.3 Estimation

Eqns. (3) to (5) are estimated by weighted least squares (WLS) in order to accommodate differences in the variances of the VR estimates across studies. Optimal weights for WLS are given by the inverse variance (Hedges and Olkin 1985). For most of the analysis we use w = $1/se_{ij}^2$. However, for robustness we also use *random* effects weights, w = $1/(se_{ij}^2 + \tau^2)$, where τ^2 is the between-study or heterogeneity variance. Stanley and Doucouliagos (2013) demonstrate that WLS MRA is superior to both conventional fixed- and random-effects multiple MRA because WLS MRA has lower bias and mean squared error if there is publication bias and is practically equivalent to random-effects when there is no publication bias.

As can be seen from Table 1 below, we employ multiple estimates from the same study. Dependence within studies is often an issue in meta-regression analysis and can result in downward bias in meta-regression analysis standard errors (Moulton 1990; Cameron, Gelbach and Miller 2008; MacKinnon and Webb 2013). However, it is worth noting that tests of the EMH are different to most empirical research in economics. A common feature of empirical economics is that authors typically estimate numerous versions of a given econometric model; this typically involves alternate specifications, estimators and data samples. This dimension for experimentation and selection is not available to VR tests of the EMH. For example, there can be no specification searching for alternate control variables, which ordinarily generates much excess variation in reported empirical estimates in other areas of economics research. Here, we have multiple estimates but these should be largely statistically independent. For example, authors often report VR tests for several countries, and estimates for different countries can be considered to be statistically independent (Hunter and Schmidt, 2004). Variance ratio testing involves application of a specific formula, and there is little scope for experimentation with alternate specifications, functional forms and estimators. Nevertheless, there are still design choices made by authors, including the choice of countries to analyze, the time period studied, and the holding period for the VRT. Among the VR tests that we find, the intraclass correlation is 0.251, suggesting that there is actually a significant degree of dependence. Hence, we take the potential for data dependence into account by estimating hierarchical models and panel data models that accommodate data dependence.

4. Quantifying the Research Record on Market Efficiency

The search for studies and data coding followed the MAER-NET guidelines for meta-regression analysis (Stanley *et al.* 2013). Specifically, we conducted a comprehensive search for studies that tested the EMH for Asian and Australasian stock markets. Numerous search engines were used: Econlit, Google Scholar, Academic Search Complete, Business Source Complete, Science Direct, Scopus, Web of Science, and Wiley Online Library. Keywords used included ‘variance ratio’, ‘efficient market hypothesis’, ‘random walk’, ‘predictability’, ‘stock markets’, ‘market efficiency’, and ‘market inefficiency’. In addition, we pursued cited references from studies, and we also physically checked numerous economics, finance and accounting journals. The database searches were terminated June 2014.

The empirical literature on the EMH is enormous. In order to make sense from a diverse literature with diverse findings, it is essential to construct a set of comparable estimates. Our criteria for inclusion were as follows. First, the study had to report an estimate of the variance ratio. There are other tests that explore the EMH. However, in order to ensure comparability, only those that report estimates of VR are included in our dataset. For example, we do not consider studies on the predictability of stock returns using macro variables. Second, the study had to report an estimate using Asian or Australasian data. Some studies report estimates also for other countries. These estimates are excluded from the meta-analysis. Third, the study had to be published. Some researchers prefer to include the so-called ‘grey literature’, such as working papers and doctoral thesis. Others prefer to stick to the published literature.[9] Fourth, for practical reasons, we exclude any study that was not written in English. Fifth, we focus on overall market efficiency and thereby exclude any estimates for individual stocks to avoid potentially overwhelming volatility and unreliability, particularly if individual stocks in emerging countries experience thin trading; aggregate indices are less vulnerable to thin trading. Also, testing efficiency of individual stocks is likely to be statistically less powerful, since no account is taken of cross stock correlations. Finally, we exclude estimates from multiple variance ratio tests (MV) as we cannot normally recover VR values from the MV statistic and the MV test statistic does not have a standard error estimator. In addition, it is difficult to justify an MV statistic as an appealing measure of predictability.

[9] It is often claimed that including unpublished studies reduces or even eliminates publication selection bias. Stanley and Doucouliagos (2012) argue that this belief is mistaken.

This search process identified 38 studies that report a variance ratio for Asian or Australasian stock markets. Although other tests statistics are reported in the research literature, the variance ratio test is by far the most prevalent. However, only 29 report sufficient information from which we could calculate the standard error of the VR.[10] The standard error is necessary in order to explore whether the EMH literature is afflicted by publication selection bias and it is also needed to properly weight the reported findings. Many studies report only the statistical significance of the VR without reporting the associate standard errors. The 29 studies report a total of 1,560 VR estimates.[11] We make only one further adjustment to the quantifiable relevant empirical record—outliers. 14 estimates have an absolute value of the standardized residual greater than 3.5 from the FAT-PET, Eqn (3), and these outliers are removed from further statistical investigation.[12] The 1,560 estimates used in our study are the population of comparable VR estimates for Asia and Australasia. While these come from 29 studies, the sample size is large. Moreover, our analysis makes use of 92 statistically independent samples. Hunter and Schmidt (2004) show that samples from different countries can be treated as being statistically independent, even if analysis of such samples is reported in the same study. In other words, the 29 studies can effectively be treated as 92 distinct cases that report 1,560 estimates of the EMH. Consequently, we have a high degree of confidence in the quality of the sample for inference purposes.

In addition to collecting data on the VR and its standard error, we also coded several other study characteristics (see Table 3 below). Two of the authors coded the studies and then checked each other's coding. All included studies are referenced in Appendix B.

---

[10] Some studies focus only on the statistical significance of the VR but do not report standard errors. The results from these studies are broadly in line with our own findings.

[11] The full dataset can be download for replication and extension from www.deakin.edu.au/meta-analysis.

[12] We also removed a single extreme value with a very large VR1 (in excess of 8) that was estimated with poor precision.

Table 1 lists the studies that form our meta-analysis dataset, including the countries studied and the number of estimates from each study. The first study was published in 1992 (Lee) and the two most recent studies in 2013 (Guidi and Gupta and Youssef and Galloppo). Table 2 presents the country distribution of the estimates and the average value of the VR for each country.

Tables 1 and 2 about here

Figure 1 presents the *VR1* estimates in the form of a funnel plot, illustrating the association between *VR1* and its precision, where precision is measured as the inverse of *VR1's* standard error. The funnel plot is a convenient way of illustrating the distribution of the reported findings. It can highlight outliers and influential observations and it can also, potentially, highlight publication selection bias (see Stanley and Doucouliagos, 2010, 2012). Estimates that stand out are potentially leverage points or outliers. Severe selection bias would cause a noticeably asymmetric distribution of results.

Figure 1 about here

The funnel graph reveals a long tail of relatively high *VR1* values, even after the very large positives VR values are removed. Asymmetry in effect sizes is typical of meta-data sets, and it is often an outcome of publication selection bias and/or heterogeneity inherent in different samples (Roberts and Stanley, 2005; Stanley and Doucouliagos, 2012). Figure 1 suggests that publication bias might be an issue in this literature, with the long tail suggesting preference for reporting violation of the EMH.

About 21 percent of the estimates report a VR1 less than 0 and the remainder (about 79 percent) report a VR1 that is greater than 0. The weighted-average VR1 is 0.13; this value is shown as the vertical line in the figure.

Table 3 lists the variables used in the multiple MRA. Some are dummy (0/1) variables (*Weekly*, *Monthly* and the various country dummy variables), while others are continuous

(*Standard error*, *Average year*[13], *Holdingperiod*, *MarketCap* and *EcoFreedom*). The binary variables *Weekly* and *Monthly*[14] reflect the data frequency used (with daily as the base). We also include the length of the holding period *k* used to construct the VR; recall Eqn. (1). The time horizon can potentially inform on the degree of predictability.[15] These three variables, *Weekly, Monthly* and *Holdingperiod*, are included in the MRA to capture potential differences in predictability of investment horizons. Stock price predictability may increase with the investment horizon as has been reported in some prior studies (*e.g.*, Fama and French, 1988b). In particular, by applying the *VR* test to U.S. stock return, Poterba and Summers (1988) report negative serial correlation in the long horizon and positive serial correlation in the short horizon.

Table 3 about here

As noted earlier, the VR can be estimated assuming either homoscedasticity or heteroscedasticity. The difference affects only the calculation of the standard error and not the value of the VR or *VR1*. Hence, we do not include this research design choice in the MRA as a moderator variable.[16] *Average year* is included to investigate whether the reported VR estimates have been changing over time. Generally, if stock markets are becoming more

---

[13] In unreported regressions we also included a dummy for the Asian financial crisis, assigning a value of 1 for estimates that relate purely to the post-financial crisis period. This variable was never statistically significant in any of the MRA models.

[14] Of the 1,560 VR1 estimates included in our sample, 434 relate to monthly data and 526 relate to weekly data, with the remainder relating to daily data. We include in *Monthly* 13 observations for quarterly returns and 11 observations for annual returns.

[15] We also calculated this length as fractions of monthly data, *i.e.*, the length of the holding period for weekly and daily frequencies was converted into monthly. For example, k=20 for daily data is converted into a one month holding period.

[16] If a dummy variable for heteroscedastic VR estimates is included in the MRA, it is never statistically significant.

(less) efficient over time, then the expected or fundamental component of VR should be falling (rising).[17]

Finally, we explore the effects of the degree of market capitalization and the degree of economic freedom prevalent in the countries and time periods sampled.[18] Market capitalization data was collected from the World Bank Development Indicators and from individual stock exchanges. Data on economic freedom was collected from the Fraser Institute (Economic Freedom of the World).[19] Economic freedom is a measure of the degree to which market forces are allowed to allocate resources and the degree to which regulations hinder this operation. The series ranges from 1 to 10, with higher values denoting higher levels of economic freedom. This is an aggregate measure of several factors, including legal structure and protection of property rights, freedom to trade and regulation of markets.

Both the market capitalization and the economic freedom series were matched with the countries and time periods used by the primary studies. Table 2 reports the average market capitalization and economic freedom for the samples used by authors for each of the countries included in our data.

Market capitalization is expected to be inversely related to market inefficiency; the more capitalized is a market the more efficient it should be, *ceteris paribus*. Market capitalization is a measure of equity market development. Underdeveloped stock markets are more likely to contain opportunities for stock market predictability. Underdeveloped

---

[17] An exception to this interpretation occurs if the variance ratio tends to be less than one, as we find among monthly frequencies.

[18] Market capitalization and economic freedom are highly correlated with a correlation coefficient of 0.85. Hence, we treat these variables as alternate measures rather than including both in the MRA; neither is statistically significant when both are included in the MRA.

[19] Kaminsky and Schmukler (2008) construct a new database of market liberalization. However, their index does not cover many of the countries in our database. The Economic Freedom of the World index is more comprehensive and available for all the countries in our database.

markets will contain a larger proportion of small and illiquid stocks, they are more likely to be characterized by thin trading and there is a greater likelihood of market manipulation.

Similarly, the greater the degree of economic freedom (market liberalization), the more efficient should be stock markets. That is, if regulations restrict the operation of markets in terms of their role in price revelation and market clearing function, then restrictions on economic freedom should, *ceteris paribus*, reduce market efficiency. Regulations impose frictions on markets and hence restrict traders from profiting from mispriced assets. Hence, we expect a negative coefficient on the MRA variables, market capitalization and economic freedom.

5. Results

We commence the analysis by calculating basic averages of all *VR1* estimates; these are presented in Table 4. Columns 1 and 2 report the simple and weighted averages, respectively. Column 3 reports the FAT-PET (Eq. 3) and column 4 presents the PEESE model results (Eq. 4). As is normally the case when there is publication selection, the weighted average and the selection bias corrected meta-averages are significantly lower than the simple average. The t-statistics exceed the higher 3.5 critical value for all averages, be they simple, weighted, or publication bias corrected. All averages suggest a rejection of the EMH for Asian and Australasian stock markets and the FAT (coefficient on standard error) is statistically significant (column 3), suggesting some net positive publication bias in this literature. All of these findings are confirmed in a multiple MRA context after other research and market dimensions are considered. See Table 5 below.

Table 4 about here

WLS estimates of the multiple MRA model, Eqn. (5), are presented in Table 5. We construct these models so that the baseline represents daily returns in Australia and New Zealand.[20] That is, when all explanatory (or moderator) variables are zero, the constant estimates *VR1* for daily returns in Australasia in 1992 (the sample mean). Column 1 controls for potential publication bias, average year, whether the data is monthly or weekly (with daily as the base), the holding period length used to construct the VR (*HoldingPeriodH= HoldingPeriod* - the Australasian mean) and market capitalization (*MarketCapH=MarketCap* - the Australasian mean). Column 2 adds the dummy variable *LesserIncome*. Column 3 explores further the difference between countries by including fourteen country dummies; it serves as our core MRA model.[21]

Columns 4 to 8 explore the robustness of these results to various alternate models.[22] Column 4 uses Robust Regression. Columns 5 and 6 report results from unbalanced panel-data MRA models that include study level effects (see Nelson and Kennedy 2009; Stanley and Doucouliagos 2012). Column 5 reports the results from the fixed-effects panel data WLS model, while column 6 present results from the random-effects panel data WLS model. Both the random and fixed effects panel data models are weighted using inverse variance weights. Column 7 reports results using 'random effects' MRA, using modified inverse variance weights, $1/(se_{ij}^2+\tau^2)$; where τ^2 is the between-study or heterogeneity variance.[23] While conventional practice among meta-analysts argues in favor of the random effects model, there

[20] In unreported regressions we used an alternate specification where the baseline of the MRA is high income countries, using the World Bank's classification system; Australia, New Zealand, Singapore, Taiwan, Hong Kong, Japan and South Korea. The results are essentially the same as those reported in the text.

[21] Note that we pool all estimates for China. In unreported regressions, we considered separate dummies for Shanghai and Shenzen indices but these results are not qualitative different to those presented in the text.

[22] We do not report OLS results here as these are not recommended for meta-regression models; OLS treats every observation equally and ignores the fact that precision varies across estimates and studies.

[23] The term 'random effects' in column 7 refers to the weights used, whereas in column 6 it refers to the normal econometrics usage, as a panel data estimator.

is growing concern about this practice, particularly in the presence of publication selection (Stanley and Doucouliagos, 2012 and 2013). Finally, there is no consensus on how clustering should be treated in the very unbalanced data used in MRA. Column 8 reports the results from a multilevel, linear hierarchical model that is one way of handling any data dependence within studies. In Table 5 we use market capitalization as the key external variable. We also re-estimate these models replacing market capitalization with economic freedom. For the sake of space, only the coefficient on economic freedom is reported in panel B of Table 5 (the full results are available from the authors).

Tables 5 and 6 about here

Most of the results from the MRA models are essentially the same (see Table 5). In particular, the results for *Standard error*, *Average Year*, and *Monthly* are very robust. However, *HoldingPeriodA* is unstable, being statistically significant in some models but not in others and also changing sign in some cases. Although *Weekly* is not robustly greater than the base, daily, all columns of Table 5 estimate the average VR1 to be positive at the weekly frequency—see the bottom of Table 5.

The MRA coefficients can be used to estimate the average VR for different data frequencies. The bottom panel of Table 5 reports these VR1 estimates and tests of the null of the EMH for daily data ($VR1_d$), weekly data ($VR1_w$), and monthly data ($VR1_m$) for the base (Australasia). These are evaluated at sample means. Table 6 reports these tests of the EMH for each country; for each country there is some evidence of violation of the EMH. In general, for daily and weekly data there is evidence of positive autocorrelation (momentum) in Asian and Australasian stock markets. For monthly data the results suggest negative autocorrelation or mean reversion.

Recall that the VR1 is essentially an estimate of the weighted sum of all the autocorrelations up to the *k* holding period, with the greatest weight placed on lags 1 and 2.

Hence, we can interpret it as a correlation coefficient. Our results show that the degree of stock market inefficiency is largest among several of the least developed countries, namely Indonesia, the Philippines and Sri Lanka; with Sri Lanka being the most inefficient. However, market inefficiency is also relatively large in Taiwan. Moreover, some of the estimates suggest that the *absolute* value of the VR1 is greater with monthly data. Thus, we can conclude that market inefficiency is larger for the longer data frequency than it is for the shorter data frequency. While these estimates do not directly quantify profits net of transaction costs, they do suggest that the degree of predictability is potentially of economic significance, on average.

6. Discussion

Several important findings emerge from the MRA models.

Standard Error is always statistically significant with a positive coefficient, suggesting publication selection bias. This finding is the opposite of what one would expect from small-sample bias alone. In Appendix A, we show that the variance ratio increases with larger samples size, but our MRA results show that VR1 increases with the standard error. *Ceteris paribus*, the standard error will be inversely related to the sample size. Thus, rather than a preference to support the conventional view, EMH, there appears to be some selective reporting of variance ratios that are greater than one (VR1 > 0), relative to those that are less than one. This does not mean that all or even most researchers engage in this practice. It takes only a small minority of negative *VR1* estimates to be suppressed to be detected statistically. Selection bias can inflate the evidence against or in favor of the EMH. Hence, it is important to accommodate or correct this bias. This is what the MRA models in Table 5 achieve; they provide tests for the EMH after allowing for selection bias, enabling more valid inferences to be made.

The coefficient on *Average Year* is negative and is robustly statistically significant, suggesting that the VR ratio falls by about 0.07 per decade. We offer three explanations for this finding. First, it may simply reflect the so-called 'Prometheus effect', or the 'decline effect,' whereby it has been observed that empirical effects are declining over time (see, for example, Ioannidis 2008). Declining effects are often attributed to initial studies reporting much larger effects as a result of selection of results that favor novel findings; for example, the earlier studies might have been eager to report a rejection of the EMH. The subsequent literature then finds smaller effects as a broader range of estimates are published. Schwert (2001) speculates that this has indeed occurred in the EMH literature.

A second and arguably more plausible explanation is that this 'decline effect' may reflect structural changes that have resulted in less momentum in stock returns and thereby improved stock market efficiency over time. This is consistent with the results from the US stock market (Kim *et al*., 2011; Ito and Sugiyama, 2009; Gu and Finnerty, 2002). For example, technical advancements in trading systems and platforms, improved transaction flows in stock exchanges, high frequency trading, reduced bid-ask spreads, greater media coverage, increased liquidity, financial products such as exchange-traded funds and greater general awareness of stock markets, can all contribute to increased stock market efficiency. A third explanation for declining effects was advanced by Schwert (2001) who claims that: "… even if the anomalies existed in the sample period in which they were first identified, the activities of practitioners who implement strategies to take advantage of anomalous behavior can cause the anomalies to disappear (as research findings cause the market to become more efficient)." In essence, Schwert (2001) is arguing that when the market is inefficient in a particular way, the inefficiency is to some degree self-correcting. Once a particular pattern of returns is highlighted and published by researchers, arbitragers can exploit them for profit, which tends to dissolve the pattern in question. Nonetheless, investor irrationality and market inefficiency may persist indefinitely (Shliefer and Summers 1990). A final point worth noting is that the negative sign on *Average Year* indicates that stock market efficiency improves over time. This finding is consistent with Lo's (2004) adaptive markets hypothesis.

Monthly always has a negative and statistically significant coefficient in the MRA, suggesting that, on average, the use of monthly data results in VR estimates less than one—Tables 5 and 6. Monthly data reveals negative autocorrelation in stock returns for the majority of countries, *i.e.* mean reversion. This means that the variance in stock returns is less than proportional to the investment horizon, whereas market efficiency requires this variance to be proportional. In contrast, on average, the use of daily and weekly data results in positive autocorrelation in stock returns and VR estimates greater than one. The use of monthly observations means that the analysis focuses on patterns that repeat, at most, every two months. We would expect that lagged prices will be less important for explaining future prices the longer is interval at which prices are observed, because serial correlation should decay as the lag length increases.

The two exogenous variables, *MarketCap* and *Ecofreedom* both have the expected negative sign in all specifications except column 2 which adds a dummy variable for lesser income countries. *MarketCap* and *Ecofreedom* are not significant in this specification because these factors tend to have the same effect on market efficiency as economic development. In column 2 of Table 6, it seems that *LesserIncome* captures the effect of all of these forces. Otherwise, the MRA confirms that more developed stock markets, as measured by market capitalization, are also more efficient. Similarly, the MRA confirms that countries with less regulation (more market liberalization) experience greater market efficiency. Our results broadly confirm those of Kaminsky and Schmukler (2008), who find that financial cycles are dampened in the long run by "improvements in property rights, transparency, and overall contractual environment", all important components of economic freedom.[24]

The development level variable, *Lesserinc*, has a positive and statistically significant coefficient, confirming that stock markets are less efficient in the less developed nations. The coefficients on the individual country dummies provide more detailed evidence on the degree to which stock market efficiency varies across countries. In general, stock markets in the less developed countries (*e.g.*, Sri Lanka, Indonesia and the Philippines) are less efficient than those from higher income countries. Kim and Shamsuddin (2008) report similar findings in their investigation of the weak-form efficiency of Asian stock markets using the VR test.

7. Summary and conclusions

The efficiency of markets is one of the cornerstones of finance theory, with profound implications for the functioning of markets and the role of regulators. When markets are efficient, prices reflect fundamental values and hence they allocate scarce funds to their highest valued use. Its prominence in economics and finance notwithstanding, there continues to be considerable debate in the theoretical and empirical literature regarding the EMH. For example, Engel and Morris (1991: 21) conclude that " … the evidence on mean reversion is mixed. Thus, more evidence is needed before declaring the stock market inefficient." This paper explores the EMH in Asian and Australasian markets. We apply meta-regression analysis to 1,560 estimates of the EMH that use the Variance Ratio test across 16 nations. We intentionally adopt a cross-country comparison, enabling us to analyze stock market efficiency from a relative (comparative) perspective.

[24] Kaminsky and Schmukler (2008) find that there is an increase in the short turn but in the long run, cycles are less pronounced.

Our results indicate that the weight of the evidence from Variance Ratio tests is a rejection of the EMH for Asian nations. Stock market efficiency is particularly weaker in the less developed and more regulated economies. An interesting pattern of results emerges among the different data frequencies. We find small to medium sized inefficiencies among daily and weekly returns and negative autocorrelation among monthly returns. Our results suggest that there is a degree of inefficiency in Asian stock markets and hence potentially some room for technical analysis and mutual fund managers to outperform the market.[25]

Market efficiency is often viewed as a final steady state. However, perhaps it is more appropriate to view market efficiency as a process rather than a state. Viewed this way, it is clear from the meta-analysis that stock markets in Asian are becoming more efficient. We find that efficiency has been improving over time and that market capitalization and economic freedom (market linearization) both increase stock market efficiency. These factors mean that it is dubious whether there remain opportunities to outperform the stock market, except in the least developed stock markets and the less liberal nations.

Our focus in this paper has been on Asian stock markets and the Variance Ratio test of the EMH. Meta-analysis could be profitably employed to other regions, especially other emerging economies such as those in the Middle East and Latin America. A particularly important extension would be to apply meta-regression analysis to other tests of the EMH, such as the stock market predictability literature that uses regression analysis.

General References

Angrist, J. D. and Pischke, J. S. (2008) *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, Princeton.

Bachelier, L. 1900. *Théorie de la Speculation*, Paris.

Barber, B. M. and Odean, T. (2001) Boys will be boys: Gender, overconfidence, and common stock investment, *Quarterly Journal of Economics* **116**(1), 261-292.

[25] Recall however that transaction costs and taxation considerations are *not* considered in VR tests.

Bender, J. C., Osler, C. L. and Simon, D. (2013) Noise trading and illusory correlations in US equity markets, *Review of Finance* **17**, 625-652.
Bom, P. R. D. and Ligthart, J. E. (2008) How productive is public capital? A meta-analysis. CESifo Working Paper No. 2206.
Boudoukh J.,M. P. Richardson, and R. F. Whitelaw (1994) A tale of three schools: insights on autocorrelations of short-horizon stock returns. *Review of Financial Studies* **7**, 539–73.
Campbell, J.Y., Lo, A.W. and MacKinlay, A.C. (1997) *The Econometrics of Financial Markets*. Princeton University Press, Princeton.
Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008) Bootstrap-based improvements for inference with clustered errors, *Review of Economics and Statistics* **90**(3), 414-427.
Charles, A. and Darné, O. (2009) Variance ratio tests of random walk: An overview, *Journal of Economic Surveys* **23**, 503-527.
Charles, A. Darne, O. and Kim, J. H. (2011) Small sample properties of alternative tests for martingale difference hypothesis, *Economics Letters* **110**, 151-154.
Chen, W. W. and Deo, R. S. (2006) The variance ratio statistic at large horizons, *Econometric Theory* 22, 206–234.
Choi, I. (1999) Testing the random walk hypothesis for real exchange rates, *Journal of Applied Econometrics* 14, 293–308.
Chow, K. V. and Denning, K. C. (1993) A simple multiple variance ratio test, *Journal of Econometrics* 58, 385–401.
Comerton-Forde, C and Putniņš, T. J. (2014) Stock price manipulation: prevalence and determinants, *Review of Finance* **18**, 23–66.
Conrad, J. S., and Kaul, G. (1988) Time varying expected returns. *Journal of Business* **61**:409–25.
Cornell, B. and Rutten, J. C. (2006) Market efficiency, crashes and securities litigation, *Tulane Law Review* **81**, 442-471.
De Bondt, W. F. M. and Thaler, R. H. (1985) Does the stock market overreact? *Journal of Finance* **40**(3), 793-805.
DeFusco, R., Mishra, S. and Raghunandan, K. (2010) Changes in the information efficiency of stock prices: Additional evidence, *The Financial Review* **45**(1), 153–165.
Dominguez, M. A. and Lobato, I. N. (2003) Testing the martingale difference hypothesis, *Econometrics Review* **22**, 351–377.
Doran, J. S., Peterson, D. R. and Wright, C. (2010) Confidence, opinions of market efficiency, and investment behavior of finance professors, *Journal of Financial Markets* **13**(1), 174-195.
Doucouliagos, H. and Stanley, T. D. (2013) Theory competition and selectivity: Are all economic facts greatly exaggerated? *Journal of Economic Surveys* **27**, 316-39.
Escanciano, J. C. and Lobato, I. N. (2009) An automatic portmanteau test for serial correlation, *Journal of Econometrics* **151**, 140–149.
Escanciano, J.C. and Velasco, C. (2006) Generalized spectral tests for the martingale difference hypothesis, *Journal of Econometrics* **134**, 151–185.
Fama, E. F. (1970) Efficient capital markets: A review of theory and empirical work, *Journal of Finance* **25**(2), 383-417.
Fama, E. and French, K. (1988a) Permanent and temporary components of stock prices, *Journal of Political Economy* **96**, 246-273.
Fama, E. and French, K. (1988b) Dividend yield and expected stock returns, *Journal of Financial Economics* **22**, 3-25.
Fischel, D. R. (1989) Efficient capital markets the crash and the fraud on the market theory, *Cornell Law Review* **74**, 907-922.

Garcia, P., Hudson, M. A. and Waller, M .L. (1988) The pricing efficiency of agricultural futures markets: an analysis of research results, *Southern Journal of Agricultural Economics* 119-130.

Griffin, J. M., Kelly, P. J. and Nardari, F. (2010) Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets, *Review of Financial Studies* **23**(8), 3225-3277.

Grossman, S.J. and Stiglitz, J.E. (1980) On the impossibility of informationally efficient markets. *The American Economic Review* **70** (3), 393–408.

Gu, A. Y. and Finnerty, J. (2002) The evolution of market efficiency: 103 years daily data of Dow, *Review of Quantitative Finance and Accounting* **18**(3), 219–237.

Gunther, M. and Herath, S. (2010) Efficiency of the real estate market: a meta-analysis, *17th Annual European Real Estate Society Conference*. ERES: Conference. Milan, Italy, 2010.

Harvey, C. (1995) Predictable risk and returns in emerging markets, *Review of Financial Studies* **8**, 773–816.

Hunter, J. E. and Schmidt, F. L. (2004) *Methods of Meta-Analysis: correcting error and bias in research findings*, New York: Sage.

Ioannidis, J. P. A. (2008) Why most true discovered associations are inflated, *Epidemiology* **19**, 640–648.

Ito, M. and Sugiyama, S. (2009) Measuring the degree of time varying market inefficiency, *Economics Letters* **103**(1), 62–64.

Jegadeesh, N. and Titman, S. (1993) Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* **48**(1), 65-91.

Kaminsky, G. L. and Schmukler, S. L. (2008) Short-run pain, long-run gain: financial liberalization and stock market cycles, *Review of Finance* **12**, 253–292.

Keef, S. P. and Roush, M. L. (2007) A meta-analysis of the international evidence of cloud cover on stock returns, *Review of Accounting and Finance* 6(3), 324-338.

Kim, J. H. (2006) Wild bootstrapping variance ratio tests, *Economics Letters*, **92**, 38-43.

Kim, J. H. (2009) Automatic variance ratio test under conditional heteroskedascity, *Finance Research Letters* **6**(3), 179-185.

Kim, J. H. and Shamsuddin, A. (2008) Are Asian stock markets efficient? evidence from new multiple Variance Ratio tests, *Journal of Empirical Finance* **15**(3), 518-532.

Kim, J. H., Shamsuddin, A. and Lim, K.-P. (2011) Stock return predictability and the adaptive markets hypothesis: Evidence from century-long U.S. data, *Journal of Empirical Finance* **18**, 868-879.

Lewis, M. (2014) *Flash Boys*. W.W. Norton, New York.

Lim K-P. and Brooks, R. D. (2011) The evolution of stock market efficiency over time: A survey of the empirical literature, *Journal of Economic Surveys* **25**, 69-108.

Lo, A. W. (2004) The adaptive markets hypothesis: market efficiency from an evolutionary perspective, *Journal of Portfolio Management* 15-29.

Lo, A. W. and MacKinlay, A. C. (1988) Stock market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies* **1**(1), 41-66.

Lo, A. W. and MacKinlay, A. C. (1989) The size and power of the variance ratio test in finite samples: a Monte Carlo investigation, *Journal of Econometrics* 40, 203–238.

Lo, A. and MacKinlay, A. C. (1990). An econometric analysis of nonsynchronous trading, *Journal of Econometrics* **45**, 181–212.

Mandelbrot, B. (1963) The variation of certain speculative prices, *Journal of Business* **36**, 394-419.

Moulton, B. R. (1990) An illustration of a pitfall in estimating the effects of aggregate variables on micro unit, *The Review of Economics and Statistics* **72**(2), 334–338.

Neely, C.J.,Weller, P.A., Ulrich, J., 2009. The adaptive markets hypothesis: evidence from the foreign exchange market. Journal of Financial and Quantitative Analysis 44 (2), 467–488.

Park, C. H. and Irwin, S. H. (2007) What do we know about the profitability of technical analysis? *Journal of Economic Surveys* **21**, 786-826.

Poterba J. A., and Summers, L. H. (1988) Mean reversion in stock prices, *Journal of Financial Economics* **22**, 27-59.

Reilly, D., Lucey, B. and Gurdgiev, C. (2011) Real estate and the stock market: a meta-regression analysis. Paper presented at the 2011 Cambridge MAER-NET Colloquium, Cambridge University, September 16-18th, England.

Richardson, M. and Smith, T. (1991) Tests of financial models in the presence of overlapping observations, *Review of Financial Studies* **4**, 227–254.

Samuelson, P. (1965) Proof that properly an anticipated prices fluctuate randomly, *Industrial Management Review* **6**, 41–9.

Schwert, G. (2003) Anomalies and market efficiency. In G. Contantinides, et. al., *Handbook of the Economics of Finance* (pp. 937-972). Amsterdam: North Holland.

Shamsuddin, A. and Kim, J. H. (2010) Short-horizon return predictability in international equity markets, *The Financial Review* **45**(2), 469-484.

Shiller, R. J. (1984) Stock prices and social dynamics, *Brookings Papers in Economic Activity* **2**, 457-498.

Shleifer, A. and Vishny, R. W. (1997) The limits of arbitrage, *Journal of Finance* **52**, 35−55.

Shliefer, A, and L. H. Summers (1990) The noise trader approach to finance, *Journal of Economic Perspectives* **4**, 19-33.

Stanley, T. D. (2008) Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection, *Oxford Bulletin of Economics and Statistics* **70**(1), 103-127.

Stanley, T. D. and Doucouliagos, H. (2010) Picture this: A simple graph that reveals much ado about research, *Journal of Economic Surveys* **24**(1), 170-191.

Stanley, T. D. and Doucouliagos, H. (2012) *Meta-Regression Analysis in Economics and Business*. Oxford: Routledge.

Stanley, T. D., Doucouliagos, H., Giles, M., Heckemeyer, J., Johnston, R. J., Laroche, P., Nelson, J. P., Paldam, M., Poot, J., Pugh, G., Rosenberger, R., Rost, K. (2013) Meta-analysis of economics research reporting guidelines, *Journal of Economic Surveys* **27**(2), 390-394.

Stanley T. D. and Doucouliagos, H. (2013) Better than random: Weighted least squares meta-regression analysis. SWP, Economics Series 2013-2, Deakin University.

Stanley, T. D. and Doucouliagos, H. (2014) Meta-regression approximations to reduce publication selection bias, *Research Synthesis Methods* **5**, 60-78.

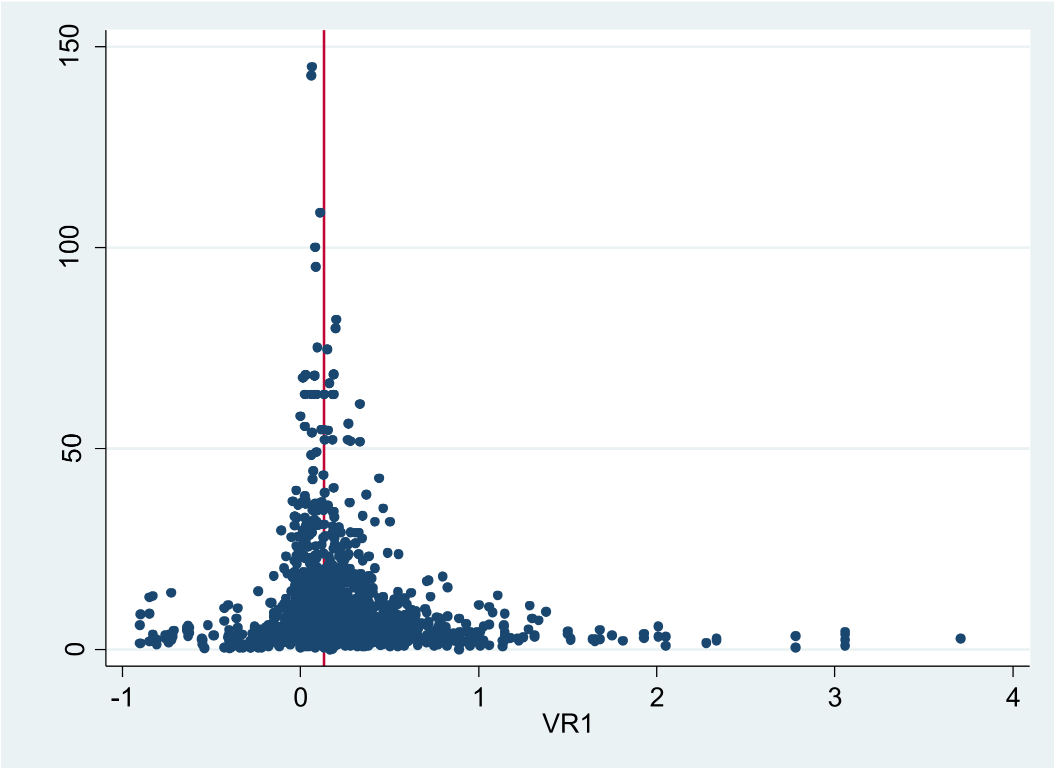
Tufte, E. (2006) *American Political Science Review* Top Twenty Citations, *American Political Science Review* **100**(4), 667-687.

Yen, G. and Lee, C. F. (2008) Efficient market hypothesis (EMH): past, present and future, *Review of Pacific Basin Financial Markets and Policies* **11**(2), 305-329.

Wright, J. H. (2000) Alternative variance–ratio tests using ranks and signs, *Journal of Business* and *Economic Statistics* **18**, 1–9.

Zhang, C. Y. and Jacobsen, B. (2013) Are monthly seasonals real? A three century perspective, *Review of Finance* **17**, 1743–1785.

Figure 1: Funnel Graph, Variance Ratio Estimates



Note: The vertical line illustrates the value of the weighted-average variance-ratio minus one (*VR1*) (0.13), using the inverse variance as weights. Outliers removed.

Table 1: Studies included in the meta-regression analysis

Authors	Countries	Number of estimates	Average VR (range)
Alam, Hasan & Kadapakkam (1999)	Bangladesh, Hong Kong, Malaysia, Sri Lanka, Taiwan	40	1.23 (0.45 - 2.32)
Ayadi & Pyun (1994)	Korea	40	0.96 (0.10 - 4.71)
Chakraborty (2006)	Pakistan	42	1.39 (1.07 – 1.62)
Chang & Ting (2000)	Taiwan	87	1.55 (0.57 – 3.40)
Chen & Jarrett (2011)	China	32	1.25 (1.07 – 1.60)
Cheung & Coutts (2001)	Hong Kong	14	1.07 (0.84 – 1.22)
Claessens, Dasgupta & Glen (1995)	India, Indonesia, Korea, Malaysia, Pakistan, Philippines, Taiwan, Thailand	16	1.24 (0.99 – 1.74)
Cohen (1999)	Japan	12	1.04 (0.90 – 1.23)
Darrat & Zhong (2000)	China	20	1.26 (0.87 – 1.52)
Fuss (2005)	Indonesia, Korea, Malaysia, Philippines, Taiwan, Thailand	112	1.61 (0.92 – 4.06)

Groenewold & Ariff (1998)	Australia, Hong Kong, Indonesia, Japan, Korea, Malaysia, New Zealand, Singapore, Taiwan	36	1.79 (0.88 – 3.34)
Guidi & Gupta (2013)	Indonesia, Malaysia, Philippines, Singapore, Thailand	40	1.15 (0.98 -1.38)
Hassan & Chowdhury (2008)	Bangladesh	4	1.34 (1.26 – 1.40)
Hiremath & Kamaiah (2010)	India	8	1.06 (1.04 – 1.09)
Huang (1995)	Hong Kong, Indonesia, Japan, Korea, Malaysia, Philippines, Singapore, Thailand, Taiwan	105	1.27 (0.6 – 2.81)
Hung (2009)	China	48	1.15 (0.97 – 1.68)
Islam & Khaled (2005)	Bangladesh	18	1.30 (0.99 – 1.92)
Karemera, Ojah & Cole (1999)	Hong Kong, Indonesia, Korea, Malaysia, Philippines, Singapore, Taiwan, Thailand	128	0.98 (0.17 – 2.13)
Kawakatsu & Morey (1999)	India, Korea, Thailand, Malaysia, Philippines	59	1.44 (0.74 – 9.26)
Lai, Balachandher & Nor (2003)	Malaysia	56	1.36 (1.09 – 1.89)
Lee (1992)	Australia, Japan	32	1.23 (0.98 – 1.77)
Lee, Chen & Rui (2001)	China	16	1.21 (1.06 – 1.47)
Lock (2007)	Taiwan	16	1.25 (0.97 – 1.78)
Long, Payne & Feng (1999)	China	4	1.21 (1.10 – 1.30)
Lu & Wang (2007)	China	47	1.06 (0.46 -2.01)
Rashid (2006)	Pakistan	24	0.85 (0.60 – 1.17)
Patro & Wu (2004)	Australia, Hong Kong, Japan, Singapore,	360	1.12 (0.75 – 1.42)
Youssef & Galloppo (2013)	China, India, Indonesia	46	1.17 (0.91 – 1.39)
Worthington & Higgs (2006)	Australia, China, Hong Kong, Indonesia, Japan, Korea, Malaysia, New Zealand, Pakistan, Philippines, Singapore, Sri Lanka, Taiwan, Thailand	120	1.18 (0.17 – 2.06)

Table 2: Individual Country Estimates, Variance Ratio Tests

Country	Number of studies (estimates)	Average VR1	Average market capitalization	Average economic freedom
Australia	4 (118)	0.096	50.26	7.19
Bangladesh	3 (30)	0.309	2.54	5.21
China	6 (168)	0.176	23.84	5.56
Hong Kong	8 (191)	0.029	116.74	8.52
India	6 (60)	0.203	34.87	5.81

Indonesia	8 (84)	0.585	16.35	6.27
Japan	6 (136)	0.085	82.98	7.11
Korea	8 (105)	0.131	44.67	6.13
Malaysia	10 (138)	0.272	160.65	6.99
New Zealand	2 (12)	0.732	38.64	7.63
Pakistan	4 (76)	0.203	15.73	5.58
Philippines	7 (74)	0.419	36.62	6.04
Singapore	6 (139)	0.170	133.53	8.01
Sri Lanka	2 (16)	0.794	14.10	5.61
Taiwan	9 (166)	0.351	65.62	6.95
Thailand	7 (76)	0.378	44.34	6.68

Notes: Average VR1 calculated by the authors from the included studies. VR1 = VR -1.

Table 3: Multiple MRA variables, descriptions and mean and standard deviation

Variable	Description	Mean (standard deviation)
VR1	Variance-ratio, the dependent variable	0.224 (0.42)
Standard error	Standard error of the VR1	0.243 (0.51)
Average year	Mean year of the data used normalized to 1992	0.000 (6.28)
Weekly	Data frequency is weekly (daily is the base). Binary variable.	0.331 (0.47)
Monthly	Data frequency is monthly (daily is the base). Binary variable.	0.280 (0.45)
Holdingperiod	The length of the holding period.	18.098 (59.11)
HoldingperiodA	The length of the holding period minus the mean for Australasia, 29.850	2.030 (59.11)
Marketcap	Average degree of market capitalization	74.864 (59.49)
MarketcapA	Average degree of market capitalization minus the mean for Australasia, 48.12	25.692 (59.49)
EcoFreedom	Average economic freedom	6.783 (0.96)
EcoFreedomA	Average economic freedom minus the mean for Australasia, 7.236	-0.444 (0.96)
LesserIncome	Binary variable for middle and low income countries	0.477 (0.50)
China	Dummy variable with a value of 1 if the data relate to China.	0.122 (0.33)
Korea	Dummy variable with a value of 1 if the data relate to Korea.	0.064 (0.19)
Hong Kong	Dummy variable with a value of 1 if the data relate to Hong Kong.	0.098 (0.30)
India	Dummy variable with a value of 1 if the data relate to India.	0.040 (0.19)
Indonesia	Dummy variable with a value of 1 if the data relate to Indonesia.	0.054 (0.23)
Japan	Dummy variable with a value of 1 if the data relate to Japan.	0.085 (0.28)
Malaysia	Dummy variable with a value of 1 if the data relate to Malaysia.	0.088 (0.28)
Philippines	Dummy variable with a value of 1 if the data relate to Philippines.	0.048 (0.21)
Singapore	Dummy variable with a value of 1 if the data relate to Singapore.	0.092 (0.29)
Taiwan	Dummy variable with a value of 1 if the data relate to Taiwan.	0.102 (0.30)
Thailand	Dummy variable with a value of 1 if the data relate to Thailand.	0.048 (0.21)
Australasia	Dummy variable with a value of 1 if the data relate to Australia or New Zealand.	0.081 (0.27)

Table 4. Average and Meta-Average Variance Ratio and Publication Selection Bias Tests
(Dependent variable is the *VR1*)

	Simple average (1)	Weighted average (2)	Publication bias corrected: FAT-PET (3)	Publication bias corrected: PEESE (4)
Average VR1	0.224*	0.130*	0.102*	0.129*
	(21.22)	(33.98)	(20.84)	(33.60)
Standard error	-	-	0.761*	-
			(8.99)	
Standard error squared	-	-	-	0.300*
				(2.54)
Adjusted R^2	-	-	0.049	0.01

Notes: The dependent variable in all columns is the variance-ratio minus 1, *VR1*. The number of observations is 1,560. Column 1 reports the simple (unweighted) average variance-ratio. Column 2 reports the weighted average variance-ratio, with the inverse variance as weights (fixed effects weights). Column 3 reports the results from the FAT-PET model (Eq. 3), while column 4 reports the results from the PEESE model (Eq. 4). Columns 3 and 4 are estimated using weighted least squares (WLS), with the inverse variance as weights. Parentheses report t-statistics. * denotes statistical significance at least at the 5 percent level. We raise the critical value for the PET coefficient to 3.5 to allow for small-sample bias (see Appendix A).

Table 5. Multiple Meta-Regression Results
(Dependent variable is the *VR1*)

Variables	WLS (1)	WLS (2)	WLS (3)	Robust (4)	FE Panel (5)	RE Panel (6)	Random effects (7)	Multi-level (8)
Constant	0.108*	0.054*	0.046*	0.093*	0.017	0.018	0.098*	0.016
	(8.70)	(4.27)	(4.16)	(11.38)	(1.42)	(1.54)	(4.36)	(1.41)
Standard error	1.024*	1.079*	1.106*	0.599*	1.416*	1.402*	0.342*	1.563*
	(9.21)	(9.13)	(11.65)	(9.87)	(17.12)	(7.10)	(4.54)	(4.17)
Average year	-0.003*	-0.006*	-0.007*	-0.007*	-0.007*	-0.007*	-0.012*	-0.007*
	(-3.03)	(-5.47)	(-6.39)	(-12.71)	(-6.76)	(-7.34)	(-7.66)	(-7.07)
Monthly	-0.168*	-0.172*	-0.191*	-0.163*	-0.143*	-0.152*	-0.204*	-0.146*
	(-10.86)	(-10.67)	(-12.03)	(-14.16)	(-8.46)	(-8.95)	(-10.18)	(-8.71)
Weekly	0.028	0.033*	-0.006	-0.009	0.022*	0.010	0.028	0.017
	(1.45)	(1.70)	(-0.39)	(-1.15)	(1.92)	(0.89)	(1.61)	(1.58)
HoldingPeriodA	-0.000	-0.000	-0.000	0.002*	-0.001*	-0.001*	0.001*	-0.001*
	(-0.31)	(-0.97)	(-1.30)	(10.43)	(-4.18)	(-4.33)	(5.31)	(-4.38)
MarketCapA	-0.0004*	-0.0001	-0.002*	-0.001*	-0.001*	-0.002*	-0.002*	-0.001*
	(-4.18)	(-1.11)	(-4.60)	(-6.16)	(-5.69)	(-6.28)	(-4.08)	(-5.97)
LesserIncome	−	0.085*	−	−	−	−	−	−
		(5.85)						
Bangladesh	−	−	-0.062*	-0.024	-0.022	-0.037	0.087	-0.027
			(-2.61)	(-0.72)	(-0.45)	(-0.76)	(1.23)	(-0.56)
China	−	−	0.025	0.069*	-0.051*	0.004	0.119*	-0.030
			(1.00)	(5.94)	(-2.62)	(0.20)	(3.35)	(-1.58)
India	−	−	0.086*	0.088*	0.082*	0.088*	0.125*	0.085*
			(5.05)	(5.63)	(3.47)	(3.84)	(3.03)	(3.65)
Indonesia	−	−	0.110*	0.125*	0.113*	0.116*	0.274*	0.114*
			(4.01)	(7.66)	(5.21)	(5.23)	(6.63)	(5.29)
Malaysia	−	−	0.288*	0.243*	0.248*	0.268*	0.328*	0.255*
			(6.84)	(11.01)	(8.20)	(8.78)	(6.42)	(8.51)
Philippines	−	−	0.155*	0.160*	0.157*	0.161*	0.242*	0.158*
			(8.55)	(10.54)	(7.90)	(7.93)	(6.08)	(8.03)
Sri Lanka	−	−	0.275*	0.540*	0.273*	0.273*	0.576*	0.273*
			(3.87)	(22.10)	(8.47)	(8.33)	(8.58)	(8.53)
Thailand	−	−	0.116*	0.121*	0.112*	0.116*	0.183*	0.114*
			(5.09)	(7.51)	(5.30)	(5.39)	(4.76)	(5.41)
Pakistan	−	−	0.027	0.073***	-0.023	-0.016	0.151***	-0.021
			(0.99)	(4.30)	(-0.84)	(-0.61)	(3.55)	(-0.79)
Korea	−	−	-0.046	-0.003	-0.007	-0.010	-0.066*	-0.008
			(-1.31)	(-0.21)	(-0.34)	(-0.46)	(-1.86)	(-0.36)
Hongkong	−	−	0.200*	0.121*	0.176*	0.195*	0.173*	0.183*
			(3.89)	(4.49)	(4.65)	(5.16)	(2.94)	(4.90)
Japan	−	−	-0.043*	-0.060*	-0.047*	-0.043*	-0.029	-0.046*
			(-2.48)	(-4.91)	(-2.95)	(-2.67)	(-0.94)	(-2.87)
Singapore	−	−	0.242*	0.181*	0.220*	0.235*	0.198*	0.226*
			(6.88)	(9.63)	(8.51)	(9.04)	(4.59)	(8.81)
Taiwan	−	−	0.200*	0.158*	0.162*	0.170*	0.249*	0.164*
			(5.86)	(12.00)	(8.38)	(8.76)	(7.79)	(8.61)
$VR1_d$	0.108*	0.054*	0.046*	0.094*	0.017	0.018	0.098*	0.016
	(8.70)	(4.27)	(4.16)	(11.38)	(1.42)	(1.54)	(4.36)	(1.41)
$VR1_w$	0.136*	0.087*	0.040*	0.084*	0.038*	0.028*	0.125*	0.034*
	(6.51)	(3.95)	(2.30)	(7.47)	(2.37)	(1.74)	(4.92)	(2.13)
$VR1_m$	-0.060*	-0.117*	-0.145*	-0.070*	-0.126*	-0.134*	-0.106*	-0.129*
	(-3.18)	(-5.75)	(-7.22)	(-4.86)	(-6.10)	(-6.41)	(-3.81)	(-6.30)

Adjusted R^2	0.139	0.178	0.359	0.728	0.371	0.354	0.510	
B: EcofreedomA	-0.025*	0.004	-0.069*	-0.092*	-0.039*	-0.066*	-0.081*	-0.053*
	(-3.90)	(0.45)	(-3.02)	(-8.25)	(-2.06)	(-3.61)	(-2.51)	(-2.86)

Notes: The dependent variable in all columns is the variance-ratio minus 1. The number of observations is 1,554. Cell entries in parentheses report t-statistics. All estimates use weighted least squares (WLS), with the inverse variance as weights and using robust standard errors for columns (1)-(3). Columns (4)-(6) test the robustness of the basic WLS findings using robust regression (4), fixed-effects panel methods (5), random-effects panel (6), random effects weights (7), and multi-level (8) methods. $VR1_d$, $VR1_w$, and $VR1_m$ denote the estimated Variance Ratio using daily, weekly and monthly data for the base (Australasia), respectively, evaluated at sample means; figures in brackets are t-statistics testing VR1 = 0. *denotes statistically significant at least at the 5% level, one-tail test. R^2 in columns (4) and (5) are for variations among the reported t-values in this research literature. Panel B reports the coefficient on economic freedom, replacing this with market capitalization. This model uses the same specification but for the sake of brevity the coefficients on the other variables are not reported.

Table 6. Country Specific MRA tests of the EMH, Daily, Weekly and Monthly Data

Country	Daily data $\mathbf{VR1_d}$ (1)	Weekly data $\mathbf{VR1_w}$ (2)	Monthly data $\mathbf{VR1_m}$ (3)
Bangladesh	0.060*	0.054*	-0.131*
	(4.49)	(2.74)	(-6.21)
China	0.104*	0.099*	-0.086*
	(5.10)	(4.28)	(-3.36)
Korea	0.002	-0.003	-0.188*
	(0.07)	(-0.09)	(-5.19)
Hong Kong	0.029*	0.023	-0.162*
	(2.30)	(1.22)	(-7.42)
India	0.152*	0.146*	-0.039*
	(10.87)	(7.88)	(-1.89)
Indonesia	0.211*	0.205*	0.020
	(9.66)	(7.36)	(0.75)
Japan	-0.053*	-0.058*	-0.243*
	(-4.86)	(-3.38)	(-11.49)
Malaysia	0.156*	0.151*	-0.035
	(7.94)	(6.60)	(-1.48)
Philippines	0.220*	0.215*	0.030
	(15.00)	(10.80)	(1.39)
Sri Lanka	0.377*	0.371*	0.186*
	(5.44)	(5.24)	(2.64)
Singapore	0.152*	0.147*	-0.038*
	(14.38)	(8.33)	(-1.97)
Taiwan	0.224*	0.219*	0.034
	(6.47)	(6.64)	(0.90)
Thailand	0.173*	0.167*	-0.018
	(7.99)	(6.47)	(-0.66)
Australasia	0.046*	0.040*	-0.145*
	(4.16)	(2.30)	(-7.22)

Notes: The cells report estimates using the MRA coefficients from Table 5, column 3, evaluated at the mean of the samples for average year and market capitalization. Figures in brackets are t-statistics. $VR1_d$, $VR1_w$, and $VR1_m$ denote the estimated Variance Ratio using daily, weekly and monthly data, respectively. * denotes statistically significant at least at the 5% level, one-tail test.

APPENDIX A: Meta-Regression Analysis for Variance Ratio Tests

The purpose of this Appendix is to report simulation results from meta-regression models of variance ratio tests that accommodate possible publication selection bias. Past studies have investigated similar MRA methods when regression coefficients or transformations of regression coefficients are being summarized and modeled (*e.g.*, Stanley, 2008; Stanley, Jarrell and Doucouliagos, 2010). Although the FAT-PET-MRA model that is developed and employed in these papers remains the foundation of the tests investigated here (Stanley and Doucouliagos, 2012), application to variance ratio testing introduces new challenges. In particular, the variance ratio is known to have small-sample bias (Lo and MacKinlay, 1988) and publication selection may 'go both ways.' Typically, publication selection is for statistical significant results (Card and Krueger, 1995; Doucouliagos and Stanley, 2013), which in this application implies a rejection of market efficiency. Some researchers may have priors that the market is inefficient and experiment with their models, data ranges, and methods to find a variance ratio that differs from 1 (either greater than 1 or less than 1). Other researchers may believe that the investors are rational and markets are efficient and dismiss some large variance ratios as faulty. That is, there may be selection across different studies in both directions, for statistical significance and for statistical insignificance.

Secondly, even in the absence of reporting or publication selection, small-sample bias will cause the effect size, VR1= VRT-1, to be correlated with its sample size and thereby inversely with its own standard error. The conventional FAT-PET-MRA (equation (3)) will therefore be affected by this small-sample bias, which may or may not distort the precision-effect test (PET) of the overall market efficiency (MEPET). The central purpose of these simulations is to investigate the statistical properties of MEPET for this application and thereby insure the validity of our MRA methods in assessing the research literature's evidence of market efficiency. Appendix Table 1 reports this small-sample bias.

Past simulations have only reported the performance of these MRA models of publication bias when there are various incidences of selection for statistical significance but not when selection may be for statistical *insignificance*. This too may potentially invalidate the FAT-PET-MRA by adding a yet another correlation of the reported effect to its standard error. We do not wish to suggest that these challenges invalidate the precision-effect test. In fact, the below simulations findings support the use of PET for application to variance-ratio testing. In this Appendix, we merely wish to address these challenges directly and thereby ensure the validity of a modified precision-effect test for market-efficiency research.

These simulations are designed to be realistic and to follow the statistical properties that we observe in this particular research literature on market efficiency. First, returns are generated randomly assuming that prices follow a random walk,

$$r_t = 0.035 + \varepsilon_t, \tag{A1}$$

where ε_t is $N(0, 0.035^2)$. This allows us to gage the size and type I errors of MEPET.

Appendix Table 1 reports the average small-sample bias of the variance ratio over 10,000 replications. Appendix Table 1 merely corroborates the known small-sample bias of variance ratio testing. Although this bias is thought to be rather small (Lo and MacKinlay, 1988), it is of sufficient magnitude for our FAT-PET-MRAs to detect and be influenced by it.

Appendix Table 1: Small-Sample Bias of the Variance Ratio
(10,000 replications)

VR1 Sample size	Bias of the Variance Ratio
25	-0.077
50	-0.038
100	-0.019
200	-0.011
500	-0.0043
1000	-0.0020

As a result of this small-sample bias, our precision-effect test (H_0: $\beta_0 = 0$) for equation (3) has inflated type I errors when conventional critical values are used—see Appendix Table 2. To investigate the power of PET, returns are generated randomly assuming that they are correlated to their past values, AR(1). Next, either 100, 200, 500, 800 or 1000 returns are used to calculate the Lo-MacKinlay variance ratio test statistics (q=2). This process is repeated any number of times to represent the number of tests (MRA sample size) reported by a research literature.

Appendix Table 2 reports the observed size of MEPET (*i.e.*, the frequency that market efficiency is rejected even though it is true) for various MRA samples sizes. The MRA sample size represents the number of tests reported in a given empirical literature. As Appendix Table 2 shows, the inflation of the type I error increases as PET has access to a larger number of reported tests. Therefore, if we do not modify PET to accommodate these inflated type I errors, it would be unreliable for this application. Rejecting H_0: $\beta_0 = 0$ might then imply either market inefficiency or a mistaken rejection of market efficiency. Thus, it is crucial that this type I error inflation be controlled if PET is to be used.

Appendix Table 2: Type I Errors of MEPET
(10,000 replications)

MRA sample size	H_0: $\beta_0 = 0$ Critical Value=1.96	H_0: $\beta_0 = 0$ Critical Value=3.50
50	0.056	0.002
100	0.075	0.002
250	0.128	0.004
500	0.206	0.010
1000	0.351	0.024

The last column of Appendix Table 2 reports the observed size of MEPET when the critical value is 3.5. That is, we consider H_0: $\beta_0 = 0$ to be rejected only when its t-value is 3.5 or larger. Note how all observed type I errors are less than the nominal value (0.025) for the conventional critical value = 1.96. For this reason our modified market efficiency precision-effect test (MEPET) uses a critical value of 3.5.

In practice, empirical research is likely to be more complex than what we have simulated thus far. As discussed above, some researchers might not report all of the test results that they obtain; thereby selecting what is reported. Appendix Table 3 displays the stimulations results when proportions of statistical significant and statistically insignificant VRTs are selected. In our meta-analysis of market efficiency, 38% of the reported findings are statistically significant. Thus, these simulations assume that 40% could be selected to be statistically significant, as a worse-case scenario. Likewise, we also allow 40% to be selected to be statistically insignificant. To encompass the range of possible selection we assume that either: 40/40/20%, 60/0/40%, 0/60/40% or 0/0/100% are selected to be statistically significant, insignificant or not selected, respectively. Although all returns are generated randomly, we control for selection by forcing the above percent of the VRs in our MRA sample to be statistically significant (or insignificant). In the ‘no selection’ condition, the first VR1 (significant or not) obtained is added to the MRA sample. In a second set of simulations, we allow returns to be correlated (ρ=.1 or ρ=.2).

Our simulations are made further realistic by using an MRA sample size of 1,000, which is approximately the number of VRTs found in our meta-analysis (1,339). Appendix

Table 3 shows that there are acceptably small type I errors ($\leq$ 0.024) when there is no heterogeneity or moderate heterogeneity (I^2=56%)—see column 1. Furthermore, in all these cases, MEPET is able to detect even small departures from a random walk (ρ=.1)—see column 2 of Appendix Table 3. When there is only moderate excess heterogeneity and 1,000 reported variance-ratio tests, MEPET is a nearly perfect test. Nonetheless, there remains one crucial research dimension that might reverse or qualify these results—large excess heterogeneity

Appendix Table 3: Publication Selection, Type I Error and Power of MEPET
(MRA sample size = 1,000 with 1,000 replications)

Hetero-geneity*	Incidence of Publication Selection†	1:Random Walk Level of MEPET Critical value= 3.5	2: Correlated (ρ=.1) Power of MEPET Critical value= 3.5	3: Correlated (ρ=.2) Power of MEPET Critical value= 3.5
None	No selection	.024	1.000	1.000
	0/40/60%	.022	1.000	1.000
	40/0/60%	.000	1.000	1.000
	40/40/20%	.000	1.000	1.000
I^2=56%	No selection	.007	1.000	1.000
	0/40/60%	.000	1.000	1.000
	40/0/60%	.004	1.000	1.000
	40/40/20%	.000	1.000	1.000
I^2=95%	No selection	.000	.131	1.000
	0/40/60%	.000	.000	.000
	40/0/60%	.085	.883	1.000
	40/40/20%	.000	.000	.000

* Heterogeneity is measured by $I^2=\sigma_h^2/(\sigma_h^2+\sigma_\varepsilon^2)$ (Higgins and Thompson, 2002)

† The incidence of publication selection is reported as the percent that are selected to be statistically significant/insignificant/not selected.

Excess heterogeneity is found in all meta-analyses of economics research (Stanley and Doucouliagos, 2012). That is, the observed variation among reported estimates (or tests) is larger than what one could expect from the reported standard errors alone. Past simulations have found that such excess heterogeneity is the most important characteristic in a research literature (Stanley, 2008; Stanley et al., 2010). The simulations reported in Appendix Table 3 generate excess heterogeneity by adding a second random term to the equation that generates the returns—recall Appendix equation (1). To quantify this heterogeneity, we use Higgins and Thompson's (2002) $I^2=\sigma_h^2/(\sigma_h^2+\sigma_\varepsilon^2)$; where σ_h^2 is the between-study heterogeneity variance and σ_ε^2 is the within-study sampling variance. I^2 measures the proportion of the total variation due to unexplained heterogeneity. Although the exact value of I^2 varies for each random sample, Appendix Table 3 reports it average value when there is either no selection or balanced selection (40/40/20%).

Moderate excess heterogeneity (I^2=56%) causes no problem for MEPET. In fact, the Type I errors are smaller, and the power remains perfect—100%. However, extreme heterogeneity (I^2=95%) offers another challenge for MEPET. When there such large heterogeneity, power can fall precipitously and the Type I errors rises (8.5%) if there is only selection for significant VRT findings—see column 1 of Appendix Table 3. Unfortunately, our meta-analysis finds an observed I^2 of 94%. Because the Type I error can be larger than the conventional α=.05 with such extreme heterogeneity, we believe that something further must be done to be absolutely sure that our findings imply that a rejection of market rationality.

Several remedies come to mind. First, we could increase the critical value and thereby drive all the observed type I errors to zero. Simulations show that a critical value of 10 is more than sufficient. Yet, our MEPET has a t-value much larger than 10, thereby ensuring that our evidence against market efficiency is not a type I error (recall Table 4).

The only remaining concern is the low power of MEPET for some incidences of publication selection when there is extreme heterogeneity. Given that large power is found for both versions of MEPET only when there is selection only for statistically significant VRTs (column 2 Appendix Table 3), our meta-analysis of this market efficiency seems to indicate that there is publication selection for the rejection of market efficiency. However, MEPET power would be expected to increase as the correlation among returns increase. Column 3 of Appendix Table 3 reports the power of MEPET when the first-order correlation among returns is .2. In this case, power for the no selection case increases to 100%. Thus, the exact incidence of selection is not clear from our meta-analysis. Although not reflected in these simulations, our findings are also consistent with smaller incidences of selection in both directions as long as the larger proportion is selected for statistical significance.

Appendix B – Studies included in the Meta-Analysis
(n = number of estimates used from each study)

1. Alam, M. I., Hasan, T. and Kadapakkam, P. J. (1999) An application of variance-ratio test to five Asian stock markets, *Review of Pacific Basin Financial Markets and Policies* **2**(3), 301-315. (n=40)
2. Ayadi, O. F. and Pyun, C. S. (1994) An application of variance ratio test to the Korean securities market, *Journal of Banking and Finance* **18**, 643-658. (n=40)
3. Chakraborty, M. (2006) Market efficiency for the Pakistan stock market: evidence from the Karachi stock exchange, *South Asia Economic Journal* **7**, 67-81. (n=42)
4. Chang, K. P. and Ting, K-S. (2000) A Variance Ratio test of the random walk hypothesis for Taiwan's stock market, *Applied Financial Economics* **10**, 525-532. (n=87)
5. Chen, F. and Jarrett, J. E. (2011) Financial crisis and the market efficiency in the Chinese equity markets, *Journal of the Asia Pacific Economy* **16**(3), 456-463. (n=32)
6. Cheung, K. G. and Coutts, J. A. (2001) A note on weak form market efficiency in security prices: evidence from the Hong Kong Stock Exchange, *Applied Economics Letters* **8**, 407-410. (n= 14)
7. Claessens, Dasgupta and Glen (1995) Stock return behavior in emerging stock markets*, The World Bank Economic Review* **9**, 131-151. (n=16)
8. Cohen, B. H. (1999) Derivatives, volatility and price discovery, *International Finance* 2(2), 167-202. (n=12)
9. Darrat, A. F. and Zhong, M. (2000) On testing the random-walk hypothesis: a model-comparison approach, *The Financial Review* **35**, 105-124. (n=20)
10. Fuss, R. (2005) Financial liberalization and stock price behaviour in Asian emerging markets, *Economic Change and Restructuring* **38**, 37–62. (n = 112)
11. Groenewold, N. and Ariff, M. (1998) The effects of de-regulation on share-market efficiency in the Asia-Pacific, *International Economic Journal* **12**(4), 23-47. (n=36)
12. Guidi, F. and Gupta, (2013) Market efficiency in the ASEAN region: evidence from multivariate and cointegration tests, *Applied Financial Economics* **23**(4), 265-274. (n=40)
13. Hassan, M. K. and Chowdhury, S. S. H. (2008) Efficiency of Bangladesh stock market: evidence from monthly index and individual firm data, *Applied Financial Economics* **18**, 749–758. (n= 4)
14. Hiremath, G. S. and Kamaiah, B. (2010) Some further evidence on the behaviour of stock returns in India, *International Journal of Economics and Finance* **2**(2), 157-167. (n=8)
15. Huang, B-N. (1995) Do Asian stock market prices follow random walks? evidence from the variance ratio test, *Applied Financial Economics* **5**, 251-256. (n=105)
16. Hung, J-C. (2009) Deregulation and liberalization of the Chinese stock market and the improvement of market efficiency, *The Quarterly Review of Economics and Finance* **49**, 843–857. (n=48)
17. Islam, A. and Khaled, M. (2005) Tests of weak-form efficiency of the Dhaka stock exchange, *Journal of Business, Finance and Accounting* **32**(7), 1613-1624. (n=18)
18. Karemera, D., Ojah, K. and Cole, J. A. (1999) Random walks and market efficiency tests: evidence from emerging equity markets. *Review of Quantitative Finance and Accounting* **13**, 171-188. (n=128)

19. Kawakatsu, H. and Morey, M. R. (1999) Financial liberalization and stock market efficiency: an empirical examination of nine emerging market countries, *Journal of Multinational Financial Management* **9**, 353-371. (n=59)
20. Lai, M-M., Balachandher, K.G. and Nor, F. M. (2003) An examination of the random walk model and technical trading rules in the Malaysian stock market, *Quarterly Journal of Business and Economics* **41**(1), 81-104. (n=56)
21. Lee, U. (1992) Do stock prices follow random walk? *International Review of Economics and Finance* **1**(4), 315-327. (n =32)
22. Lee, C. F., Chen, G.-M. and Rui, O.M. (2001) Stock returns and volatility on China's stock market, *The Journal of Financial Research* **24**, 523-543. (n=16)
23. Lock, D.B. (2007) The Taiwan stock market does follow a random walk, *Economics Bulletin* **7**(3), 1-8. (n=16)
24. Long, M. D., Payne, J. D. and Feng, C. (1999) Information transmission in the Shanghai equity market, *The Journal of Financial Research* **22**(1), 29-45. (n=4)
25. Lu, C. and Wang, K. (2007) Integrating A- and B-share markets in China: the effects of regulatory policy changes on market efficiency, *Review of Pacific Basin Financial Markets and Policies* **10**(3), 309–328. (n=47)
26. Patro, D. K. and Wu, Y. (2004) Predictability of short-horizon returns in international equity markets, *Journal of Empirical Finance* **11,** 553-584. (n=360)
27. Rashid, A. (2006) Random walk tests for KSE-100 index: evidence and implications, *Business Review* **1**(1), 80-95. (n=24)
28. Youssef, A. and Galloppo, G. (2013) The efficiency of emerging stock markets: evidence from Asia and Africa, *Global Journal of Business Research* **7(**4), 1-7. (n=46)
29. Worthington, A. and Higgs, H. (2006). Weak-form market efficiency in Asian emerging and developed equity markets: Comparative tests of random walk behaviour, *Accounting Research Journal* **19**(1), 54-63. (n=120)