Journal of Empirical Finance 18 (2011) 868–879

Contents lists available at SciVerse ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

Stock return predictability and the adaptive markets hypothesis: Evidence from century-long U.S. data☆

Jae H. Kim[a,*], Abul Shamsuddin[b], Kian-Ping Lim[c]

[a] School of Economics and Finance, La Trobe University, VIC 3086, Australia
[b] Newcastle Business School, The University of Newcastle, Callaghan, NSW 2308, Australia
[c] Labuan School of International Business & Finance, Universiti Malaysia Sabah, Malaysia

ARTICLE INFO

Article history:
Received 5 December 2010
Received in revised form 4 August 2011
Accepted 25 August 2011
Available online 3 September 2011

JEL classification:
G14

Keywords:
Economic bubbles
Economic crises
Adaptive markets hypothesis
Market efficiency
U.S. stock market

ABSTRACT

This paper provides strong evidence of time-varying return predictability of the Dow Jones Industrial Average index from 1900 to 2009. Return predictability is found to be driven by changing market conditions, consistent with the implication of the adaptive markets hypothesis. During market crashes, no statistically significant return predictability is observed, but return predictability is associated with a high degree of uncertainty. In times of economic or political crises, stock returns have been highly predictable with a moderate degree of uncertainty in predictability. We find that return predictability has been smaller during economic bubbles than in normal times. We also find evidence that return predictability is associated with stock market volatility and economic fundamentals.

1. Introduction

The efficient market hypothesis (EMH) grew out of the University of Chicago's business school over 40 years ago. It swayed many academics and policy makers into believing that stock prices fully reflect all available information, and no market participant can systematically make abnormal profit (Fama, 1970). When the information set is limited to past prices, the market is said to be weak-form efficient, and asset return is purely unpredictable from past prices. While most finance academics believe that the market is weak-form efficient (see Doran et al., 2010), there are critics from behavioral finance who document irrational but predictable investor behavior such as overreaction and overconfidence (see, for example, De Bondt and Thaler, 1985; Barber and Odean, 2001) and the momentum effect (Jegadeesh and Titman, 1993). Many commentators even attribute some responsibility for the recent global financial crisis (GFC) to an enduring belief of financial economists and policy makers in the EMH and the self-correcting capacity of markets (see Fox, 2009; Nocera, 2009).

Grossman and Stiglitz (1980) provide a theoretically compelling argument that a perfectly efficient market is impossible because if prices fully reflect all available information, traders would not have any incentive to acquire costly information. Given the impossibility of perfect efficiency, Campbell et al. (1997) propose the notion of relative efficiency, which has led to a shift in research focus from testing the all-or-nothing notion of absolute market efficiency to measuring the degree of market efficiency. There is also a growing empirical literature suggesting that market efficiency varies over time (for a survey, see Lim and Brooks, 2011).

☆ We would like to thank Michael Chng, Steve Easton, Greg Jamieson, Jan Libich, David Prentice, and John Shannon for their insightful comments.
* Corresponding author. Tel.: +61 3 94796616; fax: +61 3 94791654.
E-mail address: J.Kim@latrobe.edu.au (J.H. Kim).

0927-5398/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jempfin.2011.08.002

Lo (2004) proposes a new framework in the form of the adaptive markets hypothesis (AMH), which can help explain the observed time variation in the degree of market efficiency. The AMH is developed by coupling the evolutionary principle with the notion of bounded rationality (Simon, 1955). A bounded rational investor is said to exhibit satisfying rather than optimal behavior. Optimization can be costly, and market participants with limited access to information or abilities to process information are merely engaged in attaining a satisfactory outcome. Lo (2004) argues that a satisfactory outcome is attained not analytically, but through an evolutionary process involving trial-and-error and natural selection. The process of natural selection ensures the *survival of the fittest* and determines the number and composition of market participants and trading strategies. Market participants adapt to the constantly changing environment and rely on heuristics to make investment choices. An important implication of the AMH is that return predictability can arise time to time due to changing market conditions. Therefore, market efficiency may not follow a secular trend toward greater efficiency as anticipated by proponents of the EMH, but instead can vary in a cyclical fashion being "highly context dependent and dynamic" (Lo, 2004). Though a number of recent studies proceed to explain time variation in the degree of return predictability (see Chuluun et al., 2011; Gu and Finnerty, 2002; Lagoarde-Segot, 2009), none of these previous studies explore the role of changing market conditions.

The testable implications of the AMH are twofold. First, the degree of market efficiency fluctuates over time. Second, the degree of market efficiency is governed by market conditions. This paper tests the first implication by tracking the evolution of return predictability of the U.S. stock market over the last century, and the second implication by examining whether the degree of return predictability in the U.S. is dependent upon market conditions as manifested by market crashes, fundamental economic or political crises, economic bubbles and regulatory regimes.[1] We measure the degree of return predictability using three alternative test statistics with superior statistical properties, namely, the automatic variance ratio test of Choi (1999), the automatic portmanteau test of Escanciano and Lobato (2009), and the generalized spectral test of Escanciano and Velasco (2006). In addition, the confidence interval is constructed to gauge the degree of uncertainty associated with return predictability. The above methodological advances provide a more rigorous analysis and results than our predecessors.[2]

We obtain monthly measures of the degree of return predictability from the Dow Jones Industrial Average index over the period from 1900 to 2009, and test whether they are related to different stock market conditions after controlling for macroeconomic fundamentals. Since 1900, the U.S. stock market has experienced a number of exceptional and unexpected events, such as market crashes, economic or political crises, economic bubbles and major regulatory changes. These events have strong implications on the psychology of market participants and the way they incorporate new information to prices, which in turn may generate time variation in the serial correlation of returns as suggested by the AMH.

This paper finds strong evidence in favor of time-varying return predictability of the U.S. stock market and dependence of return predictability on market conditions. Both findings are consistent with the implications of the AMH. In particular, during stock market crashes, no return predictability is observed and an extremely high degree of uncertainty is associated with measures of return predictability. In contrast, during economic or political crises, stock returns are found to be highly predictable with a moderate degree of uncertainty. In times of economic bubbles, the degree of return predictability is found to be lower than in normal times. We also find that return predictability is affected by market volatility and macroeconomic fundamentals such as inflation and interest rates. Contrary to the general findings of past studies, we find a higher degree of return predictability before 1980 and a strong tendency to non-predictability afterwards. The next section presents the details of the data. Section 3 presents the methodology, and Section 4 presents the empirical results and their implications. The conclusion is drawn in Section 5.

2. Data

We use the Dow Jones Industrial Average (DJIA) index, daily from January 1900 to June 2009. The index is a price-weighted average of 30 blue-chip stocks, accounting for 25–30% of the total value of U.S. stocks. The component stocks are regarded as the main drivers of the market. The index began with 12 component stocks in 1896, expanded to 20 stocks from 1926, and then to 30 stocks from 1928 to date. The composition changes only occasionally, at the discretion of the managing editor of the Wall Street Journal.[3] For weekly data, we take Wednesday values (if the market is closed on a Wednesday, the Tuesday value is taken). The data are available from the Dow Jones index web site (http://www.djindexes.com/).

Fig. 1 presents the time plots of daily index and log returns for four sub-periods chosen arbitrarily only for clarity of exposition. We observe that the market is heavily affected by a series of major events. The index suffers a sharp decline in 1914 due to World

[1] This study uses return predictability and market inefficiency interchangeably. It is worth noting that the presence of return predictability only alludes to potential market inefficiency because observed return predictability may not be economically exploited in a systematic way due to transaction costs and instability of the return prediction model. Moreover, evidence of return predictability may exist in a world of rational asset pricing with time-varying expected returns.

[2] There are three published papers in the area of time-varying efficiency of the U.S. market. All of these studies focus on first-order return autocorrelation. Gu and Finnerty (2002) employ the Lo and MacKinlay (1988) variance ratio test on the DJIA index with a single arbitrarily selected holding period; Lo (2004) draws evidence from rolling first-order autocorrelation coefficients of monthly returns of the S&P500 index; and Ito and Sugiyama (2009) provide time-varying AR coefficients of monthly returns of the S&P500 index. However, only Gu and Finnerty (2002) proceed to explore the factors associated with autocorrelation of DJIA index returns. Our paper differs from Gu and Finnerty (2002) in two significant ways. First, we use a fully data-dependent method to calculate the optimal holding period for the variance ratio test and estimate its confidence interval with wild bootstrapping. The automatic portmanteau test used in our paper also employs a data-dependent procedure to determine the optimal lag order. In addition to these two autocorrelation-based techniques, we employ the generalized spectral test that is capable of detecting nonlinear return dependence. Second, we examine whether U.S. stock return predictability is related to changing market conditions as hypothesized by the AMH.

[3] The current composition consists of 3M, Alcoa, American Express, AT&T, Bank of America, Boeing, Caterpillar, Chevron, Cisco Systems, Coca-Cola, DuPont, Exxon Mobil, General Electric, Hewlett–Packard, Home Depot, Intel, IBM, Johnson & Johnson, J P Morgan & Chase, Kraft Foods, MacDonald's, Merck, Microsoft, Pfizer, Proctor & Gamble, The Traveller Company, United Technologies, Verizon, Wal-Mart Stores, and Walt Disney.

War I. The market becomes bullish in the 1920s due to economic prosperity as well as excessive optimism of investors before it crashes in October 1929, and then goes through a highly volatile period until the mid-1930s. The market experiences a sharp decline followed by high volatility in the late 1930s and early 1940s with the outbreak of World War II. The DJIA index reaches its pre-1929 level in November 1954, taking about 26 years to recover. From 1950 to 1980, the market is exposed to a number of economic and political crises, including the Soviet Satellite (Sputnik) launch, the Cuban missile crisis, the Vietnam War, and the oil shock in 1974. The market is bullish in the 1980s until it crashes in 1987. The 1990s are characterized by the most bullish period in history, culminating with the dot-com bubble before the 2000s. The DJIA index shows a declining trend and high volatility since the beginning of the subprime mortgage crisis in 2007.

3. Measures of return predictability

This section provides brief descriptions of the statistical tests for return predictability adopted in this paper. Interested readers are pointed to the respective references for full details.

3.1. Automatic variance ratio test

The variance ratio test has been widely used to test for the weak-form efficiency of financial markets, since its development by Lo and MacKinlay (1988). The test statistic can be written as a weighted sum of autocorrelation of stock returns, namely

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

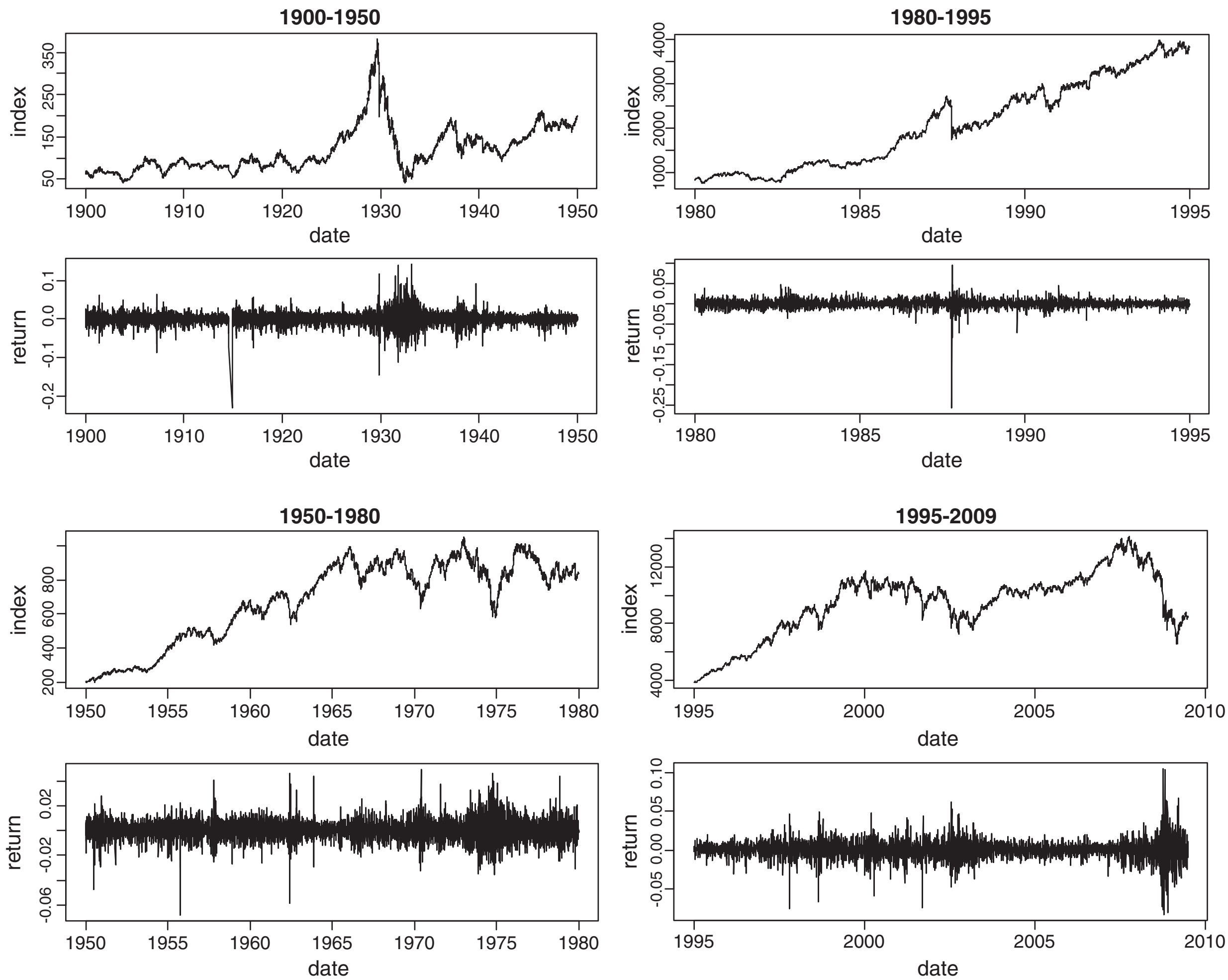

Fig. 1. Time plots of the Dow Jones Industrial Average index and its log return.

where ρ_j is the jth order autocorrelation of the returns and k is the holding period. The $V(k)$ can be estimated as:

$$VR(k) = 1 + 2\sum_{j=1}^{k-1}\left(1-\frac{j}{k}\right)\hat{\rho}_j, \tag{2}$$

where $\hat{\rho}_j$ is the estimator for ρ_j. The null hypothesis of the variance ratio test is $V(k)=1$ (or equivalently, $\rho_j=0$) for all k. The choice of holding period k is often arbitrary and made with little statistical justification. Choi (1999) proposes a fully data-dependent method of estimating the optimal choice $\hat{k}$ for k. Under the null hypothesis of $\rho_j=0$ for all j, Choi (1999) shows that:

$$AVR\left(\hat{k}\right) = \sqrt{T/\hat{k}}\left[VR\left(\hat{k}\right)-1\right]/\sqrt{2}\xrightarrow{d} N(0,1), \tag{3}$$

under the assumption that the stock return is identically and independently distributed. When returns are subject to an unknown form of conditional heteroskedasticity, statistical inference may be invalid, especially in small samples. For example, the confidence interval for $V(k)$ based on $N(0,1)$ may seriously under-estimate the uncertainty associated with estimation. In this paper, we obtain confidence interval for $V(k)$, following the wild bootstrap procedure detailed in Kim (2006, 2009). The author conducts extensive Monte Carlo experiment to show that the wild bootstrap provides accurate statistical inference in small samples under conditional heteroskedasticity.

3.2. Automatic portmanteau test

The portmanteau test is widely used to test for $H_0{:}\rho_j=0$ for all $j=1, \ldots, p$. When the stock return is subject to unknown forms of conditional heteroskedasticity, Lobato et al. (2001) propose a robustified portmanteau test statistic of the following form:

$$Q_p^* = T\sum_{i=1}^{p}\tilde{\rho}_i^2,$$

where $\tilde{\rho}_i^2 = \hat{\gamma}_i^2/\tau_i^2$; $\hat{\gamma}_i^2$ is the estimator for the autocovariance of stock return of order i; and τ_i^2 is the autocovariance of squared stock returns. Escanciano and Lobato (2009) propose an automatic test where the optimal value of p is determined by a fully data-dependent procedure. The test can be written as follows:

$$AQ \equiv Q_{\tilde{p}}^* = T\sum_{i=1}^{\tilde{p}}\tilde{\rho}_i^2, \tag{4}$$

where $\tilde{p}$ is the optimal lag order to be determined by a compromise between Akaike's information criterion and the Bayesian information criterion. The AQ statistic asymptotically follows the chi-squared distribution with one degree of freedom. The null hypothesis of no return autocorrelation is rejected at the 5% level of significance if the value of AQ is greater than 3.84.

3.3. Generalized spectral test

The above two tests are based on autocorrelation, capable of detecting linear dependence only. However, evidence of nonlinearity in stock returns is well-documented (see Lim and Brooks, 2011). When the stock return follows a general martingale difference sequence, its normalized spectral density function is equal to one at all frequencies. Based on this, Escanciano and Velasco (2006) propose a generalized spectral (GS) test, which can capture both linear and nonlinear dependence. Their test involves wild bootstrapping, in a similar manner to the AVR test discussed earlier, where the p-value of the test can be obtained. That is, if the p-value is less than 0.05, the hypothesis of no (linear and nonlinear) return predictability is rejected at the 5% level of significance. For further details of the test, readers are pointed to Escanciano and Velasco (2006).

3.4. Test statistics as measures of the degree of return predictability

The AVR statistic in Eq. (3) is a standardized version of the VR in Eq. (2), which is a weighted sum of autocorrelations with positive and declining weights. As such, it is a natural measure of the degree of return predictability and its direction. A positive (negative) value of the AVR statistic indicates overall positive (negative) autocorrelation in stock return. However, its absolute value is often used since a more efficient price exhibits less autocorrelations in both directions.[4] The statistical significance of return predictability can be evaluated using the 95% confidence interval based on the wild bootstrap for the AVR statistic (see, for details, Kim, 2006, 2009). That is, if the AVR statistic lies outside its 95% confidence interval, it is statistically different from zero at the 5% level of significance, which in turn indicates the presence of statistically significant return predictability. The confidence

[4] A number of recent studies also employ the absolute value of the variance ratio statistic to measure the relative efficiency of international equity markets (see Griffin et al., 2010 and references cited therein).

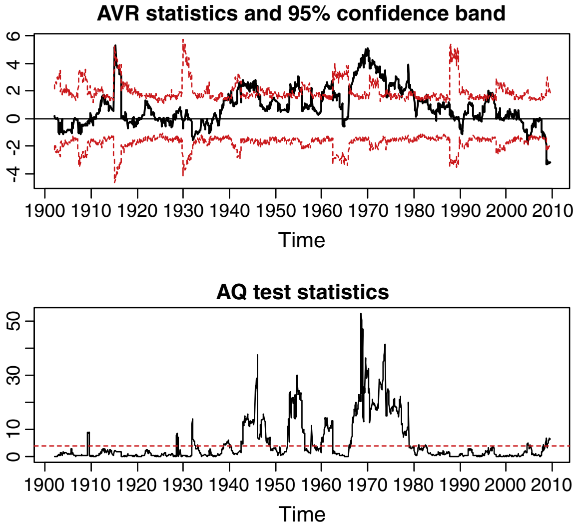

Fig. 2. The linear autocorrelation tests (daily, two-year window). Note: The thick line in the first graph plots the *AVR* (automatic variance ratio) statistics, and the broken red lines represent the 95% confidence intervals. The second graph plots the *AQ* (automatic portmanteau) test statistic, with the horizontal line corresponding to the 5% critical value of 3.84.

interval can also be used to measure the degree of uncertainty associated with return predictability. That is, a wider (narrower) interval is an indication of a higher (lower) degree of the associated uncertainty.

The *AVR* statistic, however, can be problematic when the autocorrelations of different signs are compensated. This can be overcome by using the *AQ* statistic in Eq. (4), which is based on the sum of the squared return autocorrelations and is also a natural measure of the degree of return predictability. As stated before, the *GS* test is used to detect the presence of nonlinear return predictability. Note that the *AQ* test is an asymptotic test, while the *AVR* and *GS* tests are small-sample tests based on wild bootstrapping. The three tests adopted in this paper represent a balanced mixture of tests with different statistical properties.

To obtain monthly measures of the degree of stock return predictability, we apply these tests to the whole sample using the moving-subsample window of a fixed length over the grid of months. Taking a two-year window as an example, we calculate the test statistics using the data from the first trading day of January 1900 to the last trading day of December 1901, and then move the window that covers the period from the first trading day of February 1900 to the last trading day of January 1902. We continue this process to the end of the data set and obtain monthly measures of return predictability to June 2009.[5] For daily data, we have considered window lengths of one to four years; for weekly data, we use window lengths of five and ten years. The results however are found not to be sensitive to the different choices of window length. For simplicity, we only report the results using a two-year window length for daily data and a five-year window length for weekly data.

4. Empirical results

In this section, we discuss the results of the tests given in the previous section, and conduct a regression analysis to determine whether market conditions drive stock return predictability, while controlling for macroeconomic and other potential factors.

4.1. Time-varying return predictability

Fig. 2 provides evidence from daily DJIA returns using a two-year window length. The first panel of this figure presents the *AVR* statistics along with the 95% confidence band. It is clear that the statistics fluctuate over time. There is a clear tendency that the *AVR* statistic is higher over 1935–1980, especially in the 1960s and 1970s, with a strong downward trend since 1970. The *AVR* statistic is outside the confidence band around 1914 (World War I) and also on many occasions from the late 1930s to the late 1970s, an eventful period with many economic and political crises. After 1980, the *AVR* statistic goes outside the confidence band on only three occasions: the late 1990s (dot-com bubble), around 2005 (housing bubble), and the GFC. The last two cases are the only two in the history where the *AVR* statistic has been negative with statistical significance, indicating the presence of overall negative return autocorrelation. This may represent a higher degree of panic or overreaction by market participants than in past crises.

Paying attention to the width of the confidence band, the band ranges from approximately −2 to 2 during normal times but is wider during crash times. The confidence band is exceptionally wide during the 1929 crash and the 1987 crash. These stock

[5] Note that this process does not constitute sequential testing for market efficiency. We merely test for the statistical significance of stock return predictability at the monthly frequency, allowing return predictability to change from month to month.

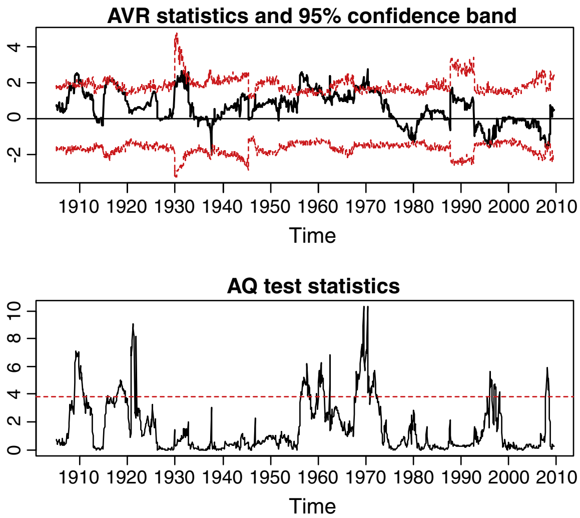

Fig. 3. The linear autocorrelation tests (weekly, five-year window). Note: The thick line in the first graph plots the *AVR* (automatic variance ratio) statistics, and the broken red lines represent the 95% confidence intervals. The second graph plots the *AQ* (automatic portmanteau) test statistic, with the horizontal line corresponding to the 5% critical value of 3.84.

market crashes are associated with statistically insignificant *AVR* statistics with an exceptionally high degree of uncertainty. In contrast, during economic and political crises, the width of the confidence band is moderate (with the exception of World War I[6]), but the *AVR* statistics are statistically significant on many occasions.

The second panel plots the *AQ* statistics along with the horizontal line corresponding to its 5% critical value. The overall result is similar to the case of *AVR*, where the time-varying nature of return predictability is evident. As with the *AVR* test, the *AQ* statistic shows a high degree of predictability between 1940 and 1980 with statistical significance, but becomes substantially smaller and overall statistically insignificant afterwards. It shows no predictability during the 1929 and 1987 stock market crashes, but does show statistically significant predictability around 2005 and 2009, which may be related to the U.S. housing bubble and the subsequent GFC.

Fig. 3 presents the case of weekly returns using a five-year window length. Overall, the results are similar to the case of daily data: a higher degree of return predictability is evident before 1980 than after 1980. Both tests show a moderate degree of return predictability in the mid to late 1990s, the time of the dot-com bubble. The *AVR* and *AQ* tests show higher predictability just before the sub-prime mortgage crisis. Return predictability is low and statistically insignificant during the two market crashes in 1929 and 1987.

Fig. 4 presents the p-values for the *GS* test for both daily and weekly data. As with the linear *AVR* and *AQ* tests discussed above, there is a strong tendency that the null of no return predictability is rejected at the 5% level of significance before 1980, especially in daily data between the 1960s and the 1970s. The inferential outcomes of the *GS* test are largely consistent with the *AVR* and *AQ* tests, which is an indication that nonlinear dependence has not been a strong feature of the U.S. stock return.[7]

If the market was efficient for the entire period, the *AVR* statistics should be outside the 95% confidence band only 5% times purely by random chance. Hence, if the AMH is true and the market becomes inefficient from time to time depending on market conditions, then this proportion should be substantially higher than 5%. This is based on the assumption that each successive case is independent, which may be arguable. We have calculated the proportion, and it is 23.4% for the daily data (Fig. 2) and 11.0% for weekly data (Fig. 3). Thus, the market has not been efficient for the whole period and market efficiency has changed over time. This evidence supports the AMH.

4.2. Return predictability and market conditions

Lo (2004) argues that the degree of return predictability varies with changes in market conditions. However, he neither suggests any specific indicators of market conditions nor any refutable predictions about the direction of the relation between return predictability and the indicators of market conditions. Given the lack of a well-structured model, we regress monthly measures of return predictability against a range of commonly used economic fundamentals and dummy variables for the states of the market. As discussed in Section 3.4, we use the absolute value of the *AVR* statistic (|*AVR*|) as a measure of the degree of return predictability and the width of the 95% confidence interval for the *AVR* statistic (*CI*) as a measure of uncertainty associated with return

[6] The market was closed for four months during World War I, and this may have spuriously resulted in exceptionally wide confidence intervals for this time.
[7] The incidence of return predictability occurs more frequently in daily rather than weekly data, which is consistent with the findings of Kim and Shamsuddin (2008) for Asian markets.

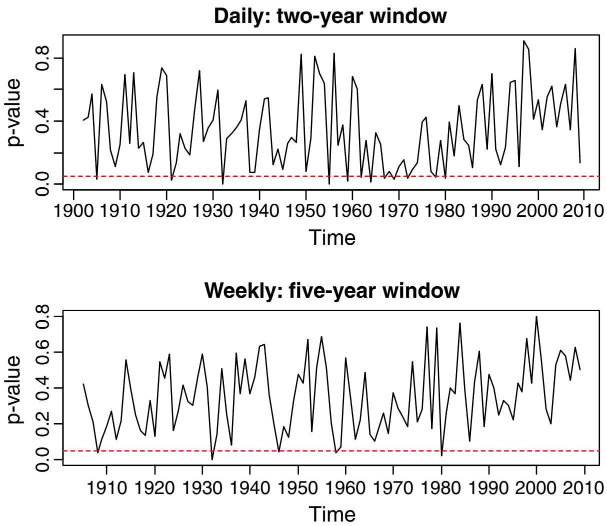

Fig. 4. p-values of the generalized spectral test. Note: the horizontal lines in each graph correspond to 0.05.

predictability. Again, similar results are evident when the measures of predictability from different window lengths are used. The regression analysis is also conducted with the *AQ* statistic as a measure of return predictability, but the results are not fully reported since qualitatively similar results are obtained.

We use dummy variables for economic bubbles, stock market crashes, and economic or political crises. We consider two sets of dummy variables as listed in Tables 1 and 2. In Table 1, the crash dummy is based on a narrower definition of market crashes, where stock markets crash purely by investors' panic with no association to any particular economic or political event. The crisis dummy is used to indicate a major economic or political crisis. In Table 2, the crash dummy variable follows the definition of Bordo (2003): a stock market crash is defined as a 20% or higher decline in prices from peak to trough, while the crisis dummy refers to an economic or political crisis that was not accompanied by such price decline. The bubble dummy variables are identically defined for both tables. The reason for using two sets of dummy variables for crash and crisis is to examine whether the results are robust to different definitions of a stock market crash. We also include a dummy variable for the expansionary phase of the U.S. business cycle, following the business cycle dates of the National Bureau of Economic Research (NBER, 2010). Santa-Clara and Valkanov (2003) observe that the U.S. stock market return is higher under Democratic presidencies than Republican presidencies, and this difference is primarily attributed to *surprises* in economic policies during Democratic presidencies. We use a control dummy for Democratic presidencies to examine whether the degree of return predictability is also sensitive to such regime switches in economic policy.

The economic variables include the interest rate, proxied by 10-year Treasury bond yield; the inflation rate, measured by the logarithmic first difference of the monthly consumer price index; and the aggregate market price-earnings ratio. The first two represent U.S. macroeconomic fundamentals and the last is an average valuation measure of the S&P500 listed firms, which rises in bull markets periods and decreases in bear markets periods.[8] As a measure of total market risk, we calculate the realized monthly volatility of the DJIA index returns, which is a purely non-parametric measure of market risk (see, for example, Andersen et al., 2003), calculated as the square root of the sum of daily return squares over a month.

The regression results are tabulated in Table 3. The results in the column labeled *REG*1 are based on the dummy variables in Table 1; while those labeled *REG*2 are based on the dummy variables in Table 2. For all cases, the regression requires lagged dependent terms to ensure the residuals mimic a white noise, based on the Ljung–Box *Q* statistics. For both regressions, |*AVR*| reacts to crises and bubbles with statistical significance. The estimated coefficients indicate a higher degree of return predictability during the crises and lower degree of return predictability during bubble times. The crash dummy is statistically insignificant, indicating no return predictability during crash times. Again for both regressions, |*AVR*| is negatively related to inflation and its lagged value, indicating that stock return predictability is lower during high inflation periods. High inflation raises inflation uncertainty (Ball, 1992), which in turn clouds the decision-making of market participants and makes return forecasting difficult. Return predictability is also related to market volatility and its lagged value with statistical significance. The |*AVR*| statistic is positively related to the interest rate (statistically significant in *REG*1 only), indicating that market return is more predictable when the interest is higher.

As for the confidence interval (CI) equation, only the *bubble*, *crash*, and *crisis* dummy variables and the lagged dependent terms are found to provide statistically adequate regression with a white noise error term. In *REG1*, *CI* responds positively to both the *crash* and the *crisis* dummies with statistical significance, but the coefficient of the *crash* dummy is substantially larger

[8] The data for these variables are obtained from http://www.econ.yale.edu/~shiller/. The data descriptions can be found in Shiller (2005, p. 5). The price-earnings ratio for the S&P500 composite index is used here as a proxy for the DJIA price-earnings ratio due to data limitations.

Table 1
Dates for crashes, crises, and bubbles.*

Events	Dates
Market crashes	
1929	1929:10–1933:03
1987	1987:10–1989:12
Economic or political crises	
World War I	1914:06–1918:09
World War II	1939:09–1945:08
Soviet satellite launch	1957:10–1958:05
Cuban missile crisis	1962:10–1963:10
Vietnam war	1959:03–1974:04
Oil shock	1973:10–1974:03
Iraq–Kuwait war	1990:08–1991:02
Sub-prime mortgage	2007:12–2009:06
Bubbles	
Roaring 1920s	1921:08–1929:08
Booming 1980s	1982:08–1987:08
Dot-com	1999:01–2000:12
US housing	2005:01–2006:12

*We are guided by the information in the timeline of the Dow Jones index web site (http://www.djindexes.com/) and the economic and political event dates provided by Wikipedia (http://en.wikipedia.org/wiki/Main_Page).

than that of the *crisis* dummy. This means that the degree of uncertainty is higher than normal times during the periods of crash and crisis, but exceptionally so during crash times. In *REG*2, a higher degree of uncertainty associated with return predictability is also evident during stock market crashes.

In summary, the regression results provide evidence of higher return predictability during economic or political crises. In contrast, during stock market crashes, no return predictability is observed and a higher degree of uncertainty is associated with return predictability. It is also found that return predictability is smaller during bubble times than in normal times. Return predictability is also related to other economic and market factors such as the inflation, interest rates, and market volatility.

4.3. Further analysis and robustness checks

In this paper, time-varying return predictability is estimated using a variance ratio test statistic over moving-subsample windows. As an alternative, we consider the test for evolving efficiency proposed by Zalewska-Mitura and Hall (1999). This test is based on Kalman filtering of AR(1) model of stock return with a GARCH(1,1)-in-mean structure in variance. This AR(1) coefficient, assumed to follow a random walk, measures the degree of time-varying return predictability. We have fitted the model to daily DJIA log return, and reported daily AR(1) coefficients with the 95% confidence band in Fig. 5.[9] The figure shows statistically significant return predictability with positive AR(1) coefficients from 1940s to 1970s, and declining predictability from 1980s, followed by negative and statistically significant AR(1) coefficients around the GFC. These features are largely consistent with those observed from Fig. 2.

In Section 4.2, we examine how the return predictability is related to changing market conditions and economic variables, using the regression analysis. To better understand their dynamic interactions, we have fitted a VAR model with five endogenous variables (|*AVR*|, *P*/*E Ratio*, *Interest Rate*, *Inflation*, and *Market Volatility*) and the dummy variables as given in Table 3. We conduct the generalized impulse response analysis proposed by Pesaran and Shin (1998).[10] Dynamic responses of |*AVR*| to the shocks of the economic variables are found to be negligible, while its responses to market conditions (measured by the coefficients of the dummy variables) are qualitatively similar to those reported in Table 3. This reinforces our regression results that return predictability is driven largely by changing market conditions. The details of these VAR model results can be obtained from the authors on request.

As for the stock market characteristics, it would be desirable to include market development measures such as liquidity and market size in the regression analysis (in Table 3). Using daily volume data for the DJIA index available from October 1928 to June 2009, we have calculated monthly values of Amihud's (2002) illiquidity ratio (ILR).[11] The ILR is included in the regression models as an explanatory variable. However, it is found to be statistically insignificant for both *REG*1 and *REG*2. These results

[9] We would like to thank Zalewska-Mitura for sharing her GAUSS code.

[10] The generalized impulse response analysis does not require orthogonalization of shocks and is invariant to the ordering of the variables in the VAR model.

[11] Market capitalization data are not available for this sub-period. Hence, the turnover ratio, the ratio of trading volume to market capitalization, cannot be used as an alternative proxy of liquidity. The ILR shows a downward trend over the sample span. The ILR is high and very volatile prior to 1950, particularly during the Great Depression and the WW II. After 1950 the ILR declines steadily and remains at a low level with little variation since the 1970s. The evidence of low illiquidity since the 1970s can be attributed to innovation in information technology and financial deregulations. The ILR values can be obtained from the authors on request.

Table 2
Alternative dates for crashes, crises, and bubbles.*

Events	Dates	
Market crashes		
Rich man's panic	1903:01	1903:11
World financial crisis	1906:01	1907:10
War scare	1912:09	1914:07
World War I	1916:11	1918:11
Disinflation, disarmament	1919:10	1921:08
Roaring 20s and policies	1929:09	1933:02
Tight monetary policy	1937:07	1938:03
World War II	1939:09	1942:04
Post war slump	1946:05	1949:06
Bretton Woods	1968:11	1970:06
Oil shock	1973:10	1974:12
October 1987 crash	1987:09	1987:11
Dot Com bust	2001:01	2002:09
Sub-prime mortgage	2007:12	2009:06
Political crises (with no market crashes)		
Soviet satellite launch	1957:10	1958:05
Cuban missile crisis	1962:10	1963:10
Vietnam war	1959:03	1974:04
Iraq–Kuwait war	1990:08	1991:02
Bubbles		
Roaring 1920s	1921:08	1929:08
Booming 1980s	1982:08	1987:08
Dot-com bubble	1997:03	2000:12
US housing bubble	2005:01	2007:05

*The market crash and bubble years are first identified by Bordo (2003) and the beginning or ending month of a crash/bubble is determined by examining the historical DJIA series.

are not tabulated in this paper since the findings for the other variables are insensitive to the inclusion of illiquidity measure in the regression models.

4.4. Breaking point of 1980

It is clear from Fig. 2 that, in general, the degree of return predictability is higher before 1980, implying that market efficiency has improved since then. This is consistent with the findings of Gu and Finnerty (2002) but is in contrast with those of Lo (2004) and Ito and Sugiyama (2009). There are two points that may support our findings of the breaking point of 1980. First, the U.S. stock market implemented a series of innovations, especially in the 1960s and the 1970s.[12] The innovations include automation and modernization of trading technology; permitting foreign brokers/dealers to operate on the floor; abolishment of fixed commissions; new regulations on corporate disclosure and auditing; extended trading hours; and the establishment of the New York futures market. These innovations must have contributed to the improved market efficiency after 1980. Following Gu and Finnerty (2002), we have included dummy variables for the market innovations in our regression, but none of them are found to be statistically significant. It seems that the innovations exert rather gradual effects and their partial effects are difficult to identify using dummy variables. Second, the breaking point of 1980 is also consistent with "Great Moderation," a decline in the volatility of macroeconomic fundamentals in the U.S. from the early 1980s (see, for example McConnell and Perez-Quiros, 2000). With a more stable economy and predictable regulatory environment, it is likely that the stock market becomes more efficient.

4.5. Adaptive markets hypothesis, behavioral finance, and data snooping

The AMH of Lo (2004) claims that return predictability is "highly context dependent and dynamic," which turns out to be one of the main findings of this paper. Since market crashes and bubbles represent extreme events that change market ecologies, our results are consistent with the AMH in that changing market conditions do matter for return predictability. While the stock market may be efficient during normal times, it can become temporarily inefficient depending on the underlying market conditions. A related study by Cooper et al. (2004) finds that short-run momentum profits depend on the state of the market, though these authors only consider up- and down-market states measured by lagged market returns.

From the standpoint of the behavioral finance, return predictability is a reflection of investors' underreaction or overreaction to news (see, for example, Daniel et al., 1998; Hong and Stein, 1999). Our finding of higher return predictability during economic and political crises indicates that such misreactions are particularly strong during times of economic and political crises. This

[12] For details, see http://www.nyse.com/about/history/timeline_chronology_index.html.

Table 3
Regression of return predictability and the state of market and economic fundamentals.

	*REG*1		*REG*2	
	|*AVR*|	*CI*	|*AVR*|	*CI*
Bubble	−0.057*	−0.023	−0.050*	−0.025
Crash	−0.056	0.322#	0.003	0.067#
Crisis	0.096*	0.083*	0.062*	0.027
Economic expansion	0.025		0.020	
Democratic president	−0.012		0.004	
Inflation	−0.023*		−0.018#	
Inflation(−1)	−0.024#		−0.019#	
Interest rate	0.006#		0.002	
P/E ratio	−0.001		−0.001	
Market volatility	1.047*		0.938*	
Market volatility(−1)	−0.971#		−1.065*	
AR	0.94	0.89	0.95	0.93
R^2	0.93	0.85	0.93	0.85
Q	0.38	0.14	0.33	0.42

Notes:
*REG*1 refers to the regression with the *bubble*, *crash*, and *crisis* dummy variables as defined in Table 1; *REG*2 uses those defined in Table 2.
|*AVR*|: absolute value of the automatic variance ratio statistic; *CI*: width of the 95% confidence interval for the *AVR* statistic; *Crisis*: dummy variable for political or economic events; *Crash*: dummy variable for stock market crashes; *Bubble*: dummy variable for bubbles (see Tables 1 and 2 for details); *Interest rate*: yield of the 10-year U.S. government bond (Shiller, 2005, p. 5); *P/E Ratio*: price-earnings ratio for the S&P 500 composite index (Shiller, 2005, p. 5); *Inflation*: logarithmic first difference of monthly consumer price index; *Market volatility*: realized volatility of the DJIA index returns; AR: sum of the coefficients for the lagged dependent variables; *Democratic president*: dummy variable for Democratic presidencies; *Economic expansion*: dummy variable for the expansionary phase of the U.S. business cycle (obtained from NBER, 2010).
Q: p-value of the Ljung–Box test statistic of order 10.
Standard errors are calculated using White's heteroskedasticity-consistent estimator.
* Indicates statistical significance at the 5% level.
\# Indicates statistical significance at the 10% level.

complements a previous finding that stock market responses to macroeconomic news do vary over the U.S. business cycle (Andersen et al., 2007) and, in particular, investors underreact to news in a down-market (Hameed and Kusnadi, 2006).

Our findings imply that the outcomes of the past empirical studies on weak-form market efficiency are dependent on the market conditions at the timing of the respective studies. That is, many past empirical outcomes may be subject to data-snooping bias. For example, the same set of stock prices evaluated at different points in time can result in completely different inferential outcomes since the respective market conditions that determine return predictability are different. This point may explain why past empirical studies on the efficiency of international stock markets have provided rather mixed and controversial results (see, for example, Yen and Lee, 2008).

Recently, Malkiel (2003; p.80) argues that “As long as stock markets exist, the collective judgment of investors will make mistakes. … as a result, … predictable patterns in stock returns can appear over time and even persist for short periods.” Timmermann (2008) reports empirical evidence that stock returns show modest levels of local predictability only during certain pockets of time, while they are unpredictable most of the time. Our empirical results confirm that this elusive stock return predictability is driven by changing market conditions. One natural question that arises is whether one can develop a profitable trading rule based on the outcomes of our measures of return predictability. Since predictability is neither a necessary nor a sufficient condition for profitability as Brown (2008) notes, this point can be further explored in future research. We note that the recent study by Neely et al. (2009) provides evidence for the AMH in the context of technical trading rules for the foreign exchange market.

5. Concluding remarks

We examine the degree of return predictability of the U.S. stock market using the century-long Dow Jones Industrial Average index. As measures of the degree of return predictability, we use the statistics from the automatic variance ratio and automatic portmanteau tests. To detect possible nonlinear dependence in stock returns, the generalized spectral test has been implemented. We obtain monthly time-varying measures of return predictability by applying these tests to moving subsample windows over monthly grids. A regression analysis is conducted to determine how these measures of return predictability are related to changing market conditions and economic fundamentals.

We find evidence that return predictability fluctuates over time and is governed largely by changing market conditions. It is found that during market crashes, no return predictability is evident, possibly due to extreme degree of associated uncertainty. However, during economic and political crises, a high degree of return predictability with a moderate degree of uncertainty is observed. During bubble times, return predictability and its uncertainty are found to be lower than normal times. However, the relation between return predictability and the state of the market may not be exploited economically because it is difficult to predict crashes, the start and end of bubbles, or the timing and duration of a crisis. We find evidence that inflation, risk-free rates, and stock market volatility are important factors that influence stock return predictability over time.

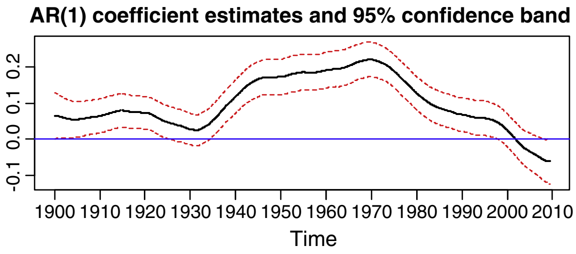

Fig. 5. An alternative test for the evolving efficiency.

Contrary to the general findings of past empirical and survey studies, we find evidence that the U.S. market has become more efficient after 1980. This is plausible, given that the U.S. market has implemented various measures of market innovation in the 1960s and 1970s, and that U.S. macroeconomic fundamentals have become much more stable since 1980. In addition, apart from the sub-prime lending crisis, there have been fewer occurrences of major economic and political crises after 1980 than before. Overall, our finding is in line with the adaptive markets hypothesis, which argues that dynamic market conditions govern the degree of stock market efficiency.

In this paper, we conduct empirical evaluation of the AMH in the context of the U.S. stock market. While a century-long U.S. daily stock price series presents a unique opportunity to examine the AMH and our results are strongly suggestive to other financial markets, it is of interest to examine if the same results can be applicable to other financial prices. Another possible future extension is to adopt alternative ways of assessing predictability, such as evaluating out-of-sample predictability of stock returns (see, for example, Timmermann, 2008). We leave these lines of research for future studies.

References

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5 (1), 31–56.
Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71 (2), 579–625.
Andersen, T.G., Bollerslev, T., Diebold, F.X., Vega, C., 2007. Real-time price discovery in global stock, bond and foreign exchange markets. Journal of International Economics 73 (2), 251–277.
Ball, L., 1992. Why does high inflation raise inflation uncertainty? Journal of Monetary Economics 29 (3), 371–388.
Barber, B.M., Odean, T., 2001. Boys will be boys: gender, overconfidence, and common stock investment. Quarterly Journal of Economics 116 (1), 261–292.
Bordo, M., 2003. Stock market crashes, productivity boom busts and recessions: some historical evidence. Rutgers University, Working Paper.
Brown, S.J., 2008. Elusive return predictability: discussion. International Journal of Forecasting 24 (1), 19–21.
Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton.
Choi, I., 1999. Testing the random walk hypothesis for real exchange rates. Journal of Applied Econometrics 14 (3), 293–308.
Chuluun, T., Eun, C.S., Kiliç, R., 2011. Investment intensity of currencies and the random walk hypothesis: cross-currency evidence. Journal of Banking & Finance 35 (2), 372–387.
Cooper, M.J., Gutierrez, R.C., Hameed, A., 2004. Market states and momentum. Journal of Finance 59 (3), 1345–1365.
Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. Journal of Finance 53 (6), 1839–1885.
De Bondt, W.F.M., Thaler, R.H., 1985. Does the stock market overreact? Journal of Finance 40 (3), 793–805.
Doran, J.S., Peterson, D.R., Wright, C., 2010. Confidence, opinions of market efficiency, and investment behavior of finance professors. Journal of Financial Markets 13 (1), 174–195.
Escanciano, J.C., Lobato, I.N., 2009. An automatic Portmanteau test for serial correlation. Journal of Econometrics 151 (2), 140–149.
Escanciano, J.C., Velasco, C., 2006. Generalized spectral tests for the martingale difference hypothesis. Journal of Econometrics 134 (1), 151–185.
Fama, E.F., 1970. Efficient capital markets: a review of theory and empirical work. Journal of Finance 25 (2), 383–417.
Fox, J., 2009. The Myth of the Rational Market: A History of Risk, Reward, and Delusion on Wall Street. HarperCollins Publishers, New York.
Griffin, J.M., Kelly, P.J., 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., Stiglitz, J.E., 1980. On the impossibility of informationally efficient markets. The American Economic Review 70 (3), 393–408.
Gu, A.Y., Finnerty, J., 2002. The evolution of market efficiency: 103 years daily data of Dow. Review of Quantitative Finance and Accounting 18 (3), 219–237.
Hameed, A., Kusnadi, Y., 2006. Stock return cross-autocorrelations and market conditions in Japan. Journal of Business 79 (6), 3029–3056.
Hong, H., Stein, J.C., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. Journal of Finance 54 (6), 2143–2184.
Ito, M., Sugiyama, S., 2009. Measuring the degree of time varying market inefficiency. Economics Letters 103 (1), 62–64.
Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48 (1), 65–91.
Kim, J.H., 2006. Wild bootstrapping variance ratio tests. Economics Letters 92 (1), 38–43.
Kim, J.H., 2009. Automatic variance ratio test under conditional heteroskedasticity. Finance Research Letters 6 (3), 179–185.
Kim, J.H., Shamsuddin, A., 2008. Are Asian stock markets efficient? Evidence from new multiple variance ratio tests. Journal of Empirical Finance 15 (3), 518–532.
Lagoarde-Segot, T., 2009. Financial reforms and time-varying microstructures in emerging equity markets. Journal of Banking & Finance 33 (10), 1755–1769.
Lim, K.P., Brooks, R.D., 2011. The evolution of stock market efficiency over time: a survey of the empirical literature. Journal of Economic Surveys 25 (1), 69–108.
Lo, A.W., 2004. The adaptive markets hypothesis: market efficiency from an evolutionary perspective. Journal of Portfolio Management 30 (5), 15–29.
Lo, A.W., 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.
Lobato, I.N., Nankervis, J.C., Savin, N.E., 2001. Testing for autocorrelation using a modified Box–Pierce Q test. International Economic Review 42 (1), 187–205.
Malkiel, B.G., 2003. The efficient market hypothesis and its critics. Journal of Economic Perspectives 17 (1), 59–82.
McConnell, M., Perez-Quiros, G., 2000. Output fluctuations in the United States: what has changed since the early 1980's? The American Economic Review 90 (5), 1464–1476.
NBER, 2010. US Business Cycle Expansions and Contractions, the Business Cycle Dating Committee of the National Bureau of Economic Research, Retrieved on 15 February 2010from http://www.nber.org/cycles.html2010.

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.

Nocera, J., 2009. Poking Holes in a Theory on Markets, New York Times (June 5, 2009). http://www.nytimes.com/2009/06/06/business/06nocera.html?scp=1&sq=efficient%20market&st=cse2009.

Pesaran, H.H., Shin, Y., 1998. Generalized impulse response analysis in linear multivariate models. Economics Letters 58 (1), 17–29.

Santa-Clara, P., Valkanov, R., 2003. The Presidential puzzle: political cycles and the stock market. Journal of Finance 58 (5), 1841–1872.

Shiller, R.J., 2005. Irrational Exuberance, 2nd edition. Princeton University Press, Princeton.

Simon, H.A., 1955. A behavioral model of rational choice. Quarterly Journal of Economics 69 (1), 99–118.

Timmermann, A., 2008. Elusive return predictability. International Journal of Forecasting 24 (1), 1–18.

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

Zalewska-Mitura, A., Hall, S.G., 1999. Examining the first stages of market performance: a test for evolving market efficiency. Economics Letters 64 (1), 1–12.