University of the Witwatersrand

PRICE-EARNINGS RATIOS AND STOCK RETURNS: EVIDENCE FROM SELECTED INDICES

Candidate: Justin Colyn—303024

Supervisor: James Britten

Programme: Master of Commerce—Finance

Submission: 31 May 2014

This research report is submitted in *partial* fulfilment of the requirements for the degree of Master of Commerce.

ABSTRACT

This report examines whether price-to-earnings ratio (P/E) and dividend yield influence future price for a select few value-weighted, equity capital market indices, over the period July 1995 through June 2013. The relationships are modelled for the S&P 500, FTSE 100 and FTSE/JSE All Share series indices, using vector autoregression (VAR). Impulse response analysis is then conducted and the forecast error variance is decomposed into its constituent parts. Finally, Granger causality tests are run. The VAR estimates suggest that an increase in P/E leads to an expected increase in price, in two months' time, but only for the S&P 500 and FTSE 100 indices. From the parameter estimates, it is expected that an increase in dividend yield does not significantly impact price except for the FTSE 100 and FTSE/JSE All Share where it has a negative relation. Price is found to be mean-reverting, with shocks to price dying out within three to four months. In general, virtually none of the variance in price can be directly attributed to the valuation ratios. For most of the indices, there is very little evidence of Granger causality, in either direction, between price and P/E and between price and dividend yield. There appears to be Granger causality in both directions between price and P/E in relation to the FTSE100.

DECLARATION

I, Justin Colyn, declare that this research report is my own unaided work. It is submitted in partial fulfilment of the requirements for the degree of Master of Commerce at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other university.

Justin Colyn

31 May 2014

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

Conventional valuation ratios—particularly the price-to-earnings ratio (P/E)—are popular measures of relative value. One often hears analysts throw the term around on television shows discussing which companies offer a reasonable investment opportunity. Market participants frequently use a weighted average P/E for the stocks within a broad market index to evaluate the level of the market as a whole. Usually, the quoted P/E is equal to closing market price per share divided by the company's trailing 12-month diluted earnings per share (EPS).

Campbell and Shiller (1998) note two reasons as to why valuation ratios are so popular. First, very long time series of these ratios exist which are fairly easy to obtain. Second, stock prices are related to *fundamental* measures of firm value. Earnings have been calculated and reported by public companies for over a century so that investors can assess intrinsic value. Is there evidence to support the belief that valuation ratios are good indicators of future investment performance; that they can be used to predict future stock price changes, in contrast to the efficient-markets theory? What does a high price-to-earnings ratio really say about future earnings growth? Does a high dividend-to-price ratio indicate that stock prices will be higher or lower in the future?

From a "micro" viewpoint, the literature does provide evidence of a *contemporaneous* relation between P/E ratios and returns from both cross-sectional and time series studies. The tendency for the stock of firms with low P/E ratios to outperform the stock of firms with high P/E ratios on a risk-adjusted basis—the low P/E effect—has been extensively documented by the likes of Basu (1977). Evidence of predictive power contained in valuation ratios as documented by authors including Campbell and Shiller (1998, 2001) is in line with this forward-looking story. At a "macro" level, the evolution over time of P/E, dividend yield (D/P or DY) as well as stock price can impact the usefulness of these measures of relative value. When active, the process of mean reversion ensures that prices will adjust to bring valuation ratios back to their long-run average. It should be noted that most of the variation in stock returns is due to capital gain because dividends are relatively stable over time (Shen, 2000). This is emphasised by Campbell and Shiller (2001) who note that, in practice, the difference between returns and stock price changes are small. Furthermore, forecasts of returns and price changes are generally very similar. As such, declining P/E ratios should be followed by lower prices—that it, declining returns—and rising P/E ratios should be followed by higher prices—rising returns. High P/E ratios—and thus low earnings yields—result in equity investments becoming less attractive in comparison with other asset classes, contributing to relatively poor short-run performance (Shen, 2000).

This research report investigates the relation between P/E and subsequent price as well between DY and subsequent price. The main aim is to determine whether P/E ratios drive future price and/or whether dividend yields drive future price, and in what direction. Bhargava and Malhotra (2006) investigate whether a (Granger) causal relation exists between P/E and subsequent price or between P/E and subsequent earnings yield for four major Morgan Stanley Capital International (MSCI) indices. Their results indicate that prices rise in response to P/E, but the impact is not very large as only one in four lags across the four indices is statistically significant and positive. Furthermore, there is no statistically significant relation between earnings yield and subsequent price.

In sections two and three, the related academic literature is reviewed. In section two, the dividend discount model (DDM) is explained and the links between price, dividends and earnings are discussed. The issues to bear in mind when using valuation ratios as measures of relative value are mentioned briefly. The next sub-section explores the impact of interest rates, risk premia and volatility on P/E. Section three starts with an overview of the low P/E effect followed by a review of the literature regarding the predicative power of both P/E and of dividend yield. Next, the concept of cointegration is introduced, specifically, of price with dividend yield and of price with P/E. The literature review is concluded with a comparison of two competing theories which describe why it is that valuation ratios fluctuate over time. Section four describes the methodology used in the study. Following Bhargava and Malhotra (2006), vector autoregression (VAR) analysis is performed in order to investigate the dynamic relationships described above. A battery of tests is run to determine whether the assumptions of the regressions hold and whether the systems of equations are stable. Pairs of the endogenous variables are tested for cointegration. Granger causality tests are also performed to determine whether the relations are "causal". Unlike Bhargava and Malhotra (2006) who use standard pair-wise, Wald-type Granger causality tests, this study will follow the surplus-lag approach outlined by Toda and Yamamoto (1995). Furthermore, dividend yield will be considered rather than earnings yield. The empirical component is delimited to the period July 1995 through June 2013—a more recent time period than that considered by Bhargava and Malhotra (2006). It will also be delimited to a select few, pre-determined equity capital market indices which include indices representing the South African market.

2 LINKING DIVIDENDS AND EARNINGS WITH PRICE

2.1 DIVIDEND DISCOUNT MODEL AND P/E

The dividend discount model (DDM) posits that the intrinsic value of a stock is equal to the present value of future expected dividends. Assuming stock price P represents the present value of the stock's future stream of dividends D which grow at a constant rate g, then:

$$P = \frac{D(1+g)}{1+r} + \frac{D(1+g)^2}{(1+r)^2} + \cdots + \frac{D(1+g)^N}{(1+r)^N}. \tag{1}$$

As $N \rightarrow \infty$ and where $g > r$, one can solve for the discount rate—or the cost of equity—represented by r. By solving for r, the constant-growth DDM or Gordon (1959) constant growth model is obtained. Under this model, it is assumed that stock prices and dividends grow at the same, constant rate g. It states that cost of equity is the sum of the dividend yield and capital gain:

$$r = \frac{D(1+g)}{P} + \frac{P_1 - P_0}{P_0} = \frac{D(1+g)}{P} + g, \text{ or} \tag{2}$$

$$P = \frac{D(1+g)}{r-g}. \tag{3}$$

Now, one can solve for the appropriate P/E ratio by dividing both sides by E, such that:

$$\frac{P}{E} = \frac{D(1+g)}{E} \times \frac{1}{r-g}. \tag{4}$$

Another useful element of the relationship between these observable parameters is that it can be utilised to estimate the expected rate of return on the "market" portfolio $E(r_m)$ by simply replacing the variables on the right of eq. (2) with their observable market counterparts. Assuming the market portfolio has a constant pay-out ratio of a, $E(r_m)$ can then be estimated as:

$$E(r_m) = \frac{aE(1+g)}{P} + g = \frac{a(1+g)}{P/E} + g. \tag{5}$$

Graham and Dodd (1934) recommend using a simple moving average of earnings over at least five years in the calculation of the P/E ratio in order to provide a more reliable valuation measure. This is not surprising given that annual earnings are noisy measures of firm value (Campbell and Shiller, 1998) particularly in recessions where earnings per share (EPS) is depressed (Molodovsky, 1953). In their study of all listed firms in the U.K. over the period 1975 through 2003, Anderson and Brooks (2006) find using multiple years of earnings result in better quality predictions of future returns. Bierman (2002) recognises the importance of the P/E ratio to investment practitioners as a measure of the reasonableness of a firm's stock price relative to its earnings, but notes two situations where the P/E ratio may not properly reflect the relative value of the firm: first, where the firm has an extraordinary liability and second, where the firm has an excess asset. Extraordinary liabilities can take the form of damages awarded to another party, unfunded pension liabilities, unfunded medical benefit liabilities and liabilities associated with the acquisition of another company. An excess asset is one where diversion of same will not adversely affect the firm's future cash flow, for example, excess cash. Adjusting for extraordinary liabilities is likely to increase the P/E while adjusting for an excess asset will have the effect of decreasing the P/E.

2.2 INTEREST RATES, RISK PREMIA AND VOLATILITY

In the DDM framework, dividends are discounted at the appropriate cost of equity, which could be estimated using the capital asset pricing model (CAPM), which describes expected returns as the sum of the risk-free rate and the expected market risk premium scaled according to the stock's sensitivity to movements in the broader market. If expected returns are modelled within this framework, then lower returns should be expected when interest rates are low or the market risk premium is low. Rozeff (1984) finds that dividend yield is an approximate measure of the *ex-ante* equity market risk premium on stocks. Recall, the dividend yield term in the DDM is given by the fraction $\frac{D(1+g)}{P}$ in eq. (2). By assuming that the real dividend growth rate is related directly to output growth, Rozeff (1984) uses the *Golden Rule of Accumulation*—a theorem which states that, in equilibrium, the real output growth rate is approximately equal to the real interest rate—to prove that the equity risk premium is approximately equal to the current dividend yield multiplied by one plus the risk-free rate. Due to the fact that the product of dividend yield and the risk-free rate is very small, the equity risk premium is approximately equal to the current dividend yield. Empirical evidence suggests that returns increase monotonically as dividend yield in the prior year increases. Rozeff (1984) argues that high returns occur when the market is deemed very risky and as such equity investors require a high risk premium to continue holding stocks while lower returns are experienced when the market is considered less risky, with equity investors requiring lower returns for holding equities. If the expected return is low then the discounted flow of dividends—or alternatively, free cash flow to equity holders—will be low which, in turn, leads to high P/E ratios. Trevino and Robertson (2002)'s results are consistent with this story. Considering a five-year holding period, when P/E ratios were high, the average returns on Treasury bills (T-bills) and Treasury notes (T-notes) are 4.76% and 5.36%, respectively. The corresponding figures when P/E ratios were low are 6.50% and 6.90%. Using a five-year holding period, the market risk premium over T-bills and T-notes is 4.60% and 4.90%, respectively. When P/E ratios were low, the corresponding figures are 12.07% and 11.67%. Market P/E is highly sensitive to volatility as evidenced by Kane, Marcus and Noh (1996) and by Koutmos (2010). Both studies find that the relation between market P/E and volatility is negative which indicates that market participants demand higher rates of return to compensate them for bearing more risk, which results in lower prices. Kane *et al*. (1996) find that a 1pp increase in volatility from 12% to 13% leads to an expected decrease in the market P/E of 1.8, from 16.8 to 15.0.

3 VALUATION RATIOS AND RETURNS

3.1 CONTEMPORANEOUS RELATION OF VALUATION RATIOS WITH RETURNS

Many studies have examined the relation between accounting variables and stock prices. The possibility that valuation ratios—particularly P/E—could be useful in forecasting future changes in stock prices have been investigated in numerous ways and in numerous countries. Nicholson (1968) studies 189 U.S. stocks over a 25-year period. He segregates the stocks based on earnings yields—the reciprocal of the P/E ratio—and finds that the top 20% earn a mean annual return of 16% while the comparable figure for the bottom 20% is just 3%. These figures are, however, not adjusted for risk. Basu (1977) studies 1,400 industrial, NYSE-listed companies over the period 1956 to 1971 and his results are consistent with those of Nicholson (1968)'s in that low P/E portfolios experience both higher absolute and higher risk-adjusted returns than high P/E portfolios. The mean annual return for the highest (lowest) P/E ratio portfolio was 9.34% (16.30%) while the beta coefficient for the highest (lowest) P/E ratio portfolio was 1.11 (0.98). Basu (1977) notes a near monotonic decline in mean annual return from the highest to the lowest P/E ratio portfolio. He further suggests there are lags and frictions in the adjustment of stock prices to new, publicly available information resulting in P/E ratios possessing "information content". Johnson, Fiore and Zuber (1989) use methodology similar to that of Basu (1977) for a different time period—1979 through 1984—but their results contradict those of Basu (1977). Johnson *et al.* (1989) find that the *low* P/E portfolio earned relatively *low* average returns and was associated with a higher level of total risk. Many investors believe they can earn abnormal returns from stock-picking in some P/E range. The study of Keown, Pinkerton and Chen (1987) investigates the feasibility of a trading strategy based on P/E ratios. Their results suggest that such an investment strategy may not be desirable. In fact, by limiting stock selection to firms with either low or high P/E ratios, investors expose themselves to greater unsystematic risk which cannot be eliminated through diversification, irrespective of portfolio size. In a similar vein, Black and Scholes (1974) show that investors attempting to maximise after-tax returns should be less concerned with stock-picking in some *dividend yield* range and rather focus on diversification because he has a better chance of decreasing portfolio risk by improving diversification than by increasing portfolio expected return by focussing on stocks with certain levels of dividend yield.

A number of studies attempt to explain away the low-P/E effect by challenging the methodology used. As an incomplete explanation of the effect, Ball (1978) argues that Basu (1977)'s assumption of earnings being announced three months after financial year-end may have biased the direction of the anomaly. In fact, Ball and Brown (1968) find 25% of sampled firms did not report within two months of their financial year-end. He argues that the post-announcement excess returns could be contaminated with pre- and at-announcement excess returns. Banz and Breen (1986) examine the relation between the database providing accounting data and returns on portfolios formed using that data. The results suggest *ex-post* selection bias and look-ahead bias “create” the low-P/E effect.

Many studies suggest the low P/E effect stems from the size effect. Banz (1981) finds, on average, small NYSE firms had significantly greater risk-adjusted returns as compared to large NYSE firms over a 40 year sample period. The effect is most evident in the very smallest firms in his sample and is not stable over time. The separate existence of both the size and P/E effects is investigated by Reinganum (1981), using a dataset comprising NYSE-listed common stock over the period 1964 through 1981. He finds both effects present in the returns when considered separately, but not when analysed together. After controlling for an E/P effect, a size effect is present. However, when controlling for a size effect, there is no separate E/P effect. Reiganum (1981) and Banz (1981) both conclude that their findings are consistent with a misspecification of the capital asset pricing model (CAPM) rather than market inefficiency. Roll (1981) offers a possible explanation based on trading infrequency. He argues that the riskiness of small firms, low P/E firms and high dividend yield firms is improperly measured; positive autocorrelation resulting from infrequent trading leads to downward biased estimates of risk and, in turn, causes inflated “risk-adjusted” average returns. Brown, Kleidon and Marsh (1983) draw three important conclusions from their study: (1) there is a linear relationship—in the log of size—between excess returns and market capitalisation, (2) *ex ante* returns attributable to size vary through time and (3) different estimation methodologies can lead to different conclusions regarding the size effect. Blume and Stambaugh (1983) find an upward bias in individual stock returns computed using daily closing prices. They recommend computing returns using a buy-and-hold portfolio approach as it reduces this bias. Their estimates suggest that the full-year size effect—based on such an approach—is about half the size of previous studies and all of the effect is attributable to the month of January. While Reinganum (1981) states that the size effect subsumes the E/P effect, Basu (1983) argues that risk-adjustment used by Reinganum (1981) was defective; it merely conceals the E/P effect and in fact, the E/P ratio effect subsumes the size effect. Cook and Rozeff (1984) re-examined the results of both studies and concluded that both effects are at work and it is quite possible that market value and E/P ratio measure separate aspects of the same *underlying* effect.

3.2 INTERTEMPORAL RELATION OF VALUATION RATIOS WITH RETURNS

Litzenberger and Ramaswamy (1979) develop an after-tax CAPM which incorporates borrowing constraints—related to wealth and to income—and a progressive tax regime. The parameter estimate attaching to D/P suggests a strong, positive, linear relationship between pre-tax expected returns and dividend yield on common equity. Fama and French (1988) note that dividend yield can be approximated by subtracting the constant dividend growth rate from the discount rate in the DDM. This can be shown by re-arranging eq 3 such that:

$$\frac{D(1+g)}{P} = r - g$$

Fama and French (1988) study the power of dividend yields in forecasting stock returns. They regress returns on dividend yield and find that dividend yield explains less than 5% of the variation in returns. Predictive power increases with horizon; dividend yield explains over 25% of the variation in two- to four-year returns. Expected returns are measured by the fitted values obtained from the regressions. Persistence—high, positive autocorrelation—of expected returns leads to disproportionately higher growth in the variance of expected returns relative to growth in the investment horizon. Growth in the residual variance is lessened by the discount rate effect in that a shock to expected return is accompanied by an immediate change in price, in the opposite direction. For example, an increase in expected return is accompanied by an immediate decline in price. On average, the cumulative price impact is almost zero; an increase in expected return implies a higher future price, which is just offset by a decrease in the current price. In effect, Fama and French (1988)'s study indicates that mean reversion is brought about by time variation of expected returns.

Shiller (1984) and Fama and French (1988) perform regression of returns on *either* lagged dividend yield or lagged earnings yield. They note that dividend yield has more explanatory power than earnings yield. They suggest that earnings are more variable as compared to dividends—it well known that dividends are "sticky"—and if this excess variation is unrelated to expected returns, this would imply that earnings yield is a noisier variable. Lamont (1998) sees earnings as informative because earnings have a stronger correlation with business conditions. However, he finds earnings and price have the same, negative relation to future returns whereas dividends and price have opposite relations to future returns. In the long run, Lamont (1998) finds price is the only variable of those he studied that matter when forecasting future returns. Simulations run by Goetzmann and Jorion (1995) using 122 years' worth of data, suggest that tests of long-run predictability may be biased toward finding predictability due to survivorship bias.

Campbell and Shiller (1998) find a negative and statistically significant relationship between P/E and D/P with ten-year returns. They approach the results with caution for two reasons: (1) the prevailing valuation ratios were far removed from their historical averages and (2) it is possible, in hindsight, to choose valuation ratios that "predict" returns even though they may not have been recognised as such in the past. Making use of aggregate annual U.S. data covering 1871 through 2000 as well as aggregate data from 12 other major economies from 1970, Campbell and Shiller (2001) find that D/P and P/E have weak forecasting power with respect to the forecasting of future dividend growth, future earnings growth and future productivity growth. They do find, however, that D/P and P/E are useful in forecasting future stock price changes. Fama and French (1989) find that D/P can be used to forecast both equity and bond returns. Interestingly, D/P and default premium appear to be highly correlated and behave in a similar manner with respect to long-term business conditions. One explanation given is when business conditions are poor, returns must be high to bring about substitution of consumption for investment.

Using 128 years of data, Fisher and Statman (2000) investigate the power of P/E ratios and dividend yield in forecasting future returns over the short- and long-run. They find that valuation ratios produce unreliable forecasts of future returns in the short run; there is no statistically significant relationship between beginning-of-year P/E ratios and returns in the following year or in the following two, non-overlapping years. The same can be said for dividend yield. The performance of the valuation ratios proved to be more reliable over the long-run; regression of 10-year annualised returns—both in nominal and real terms—on trailing 12-month P/E ratios produced a statistically significant negative slope. The relationship is not perfect, however, as the observations are widely scattered about the fitted line. Ang and Bakaert (2007) also investigate the predictive power of dividend yield, with particular focus on long horizons. They employ a long data set for the U.S., the U.K. and Germany as well as a shorter data set for those countries as well as for France. At long horizons dividend yield cannot predict excess returns. It performs best at short horizons but the performance is much improved when the short risk-free rate is added to the regressions. Interestingly, dividend yield is quite good at predicting future cash flow growth rates, which implies that the predictable component of cash flow could be an important source of variation in ratios like P/E and price-to-dividend ratio (P/D). Gupta and Modise (2012) use South African data for the period January 1990 through October 2009 to test for return predictability of P/D and P/E. They fail to reject the null hypothesis of no correlation between the current level of the selected valuation ratios with future changes in price over the short- and the long-term. Like Ang and Bakaert (2007) predictive performance is improved with the inclusion of a variables representing short-term interest rates.

Shen (2000) examines the historical relationship between P/E ratios of U.S. stock indices and the subsequent performance of those indices. Historically, very high P/E ratios are followed by low, short- and long-term returns. In the long-run, high P/E ratios are followed by slow growth in stock prices. In the short-run, performance also deteriorates if high P/E ratio results in earnings yields which are less attractive as compared to the yields on other assets. Trevino and Robertson (2002) examine the relationship of P/E ratios and subsequent returns over different holding periods. Though they find a weak relationship between current P/E ratio and subsequent, short-term average returns, they do find lower, long-term average returns after periods characterised by high P/E ratios. In order to investigate the appropriate investment indicator for less developed markets, Ramcharran (2004) estimates cross-sectional and pooled-data determinants of stock returns for 21 emerging equity markets. He finds the price-to-book ratio (P/B) is more successful in explaining returns than P/E.

If the DDM holds, then a firm's stock price is essentially a measure of the present value of the future, expected cash flows to shareholders. The patterns of these cash flows can vary across firms; for some firms, cash flows may grow in future and for others, cash flows may decline or remain unchanged. Stockholders can only earn returns in excess of the risk-adjusted required rate of return if future cash flows exceed expectations. Danielson and Dowdell (2001) introduce a return-stages model in an effort to quantify investors' expectations from P/B and P/E ratios. The model writes a firm's stock price as a function of three future returns on equity (ROE) components: (1) ROE from the firm's past investment, (2) ROE from firm's new investment made during the growth phase and (3) ROE on all investment made after the growth phase comes to an end. Firms were classified into four groups—growth firms, mature firms, turnaround firms and declining firms—based on their P/B and P/E ratios. Danielson and Dowdell find P/B and P/E ratios are helpful in predicting the patterns firms' future cash flow will follow. Furthermore, the operating performance associated with a given rate of return, differs across the groups. Their results imply that a particular firm's stock return is dependent on how its operating performance compares with expected operating performance as indicated by the firm's P/B and P/E ratios.

4 COINTEGRATION

Variables are said to be cointegrated if they are tied together in a long-run equilibrium relationship (Engle & Granger, 1987). This concept is discussed in greater detail in section 4.2. Barr and Kantor (1999) investigate the possibility of cointegration between stock prices and earnings on the JSE over the period 1980 through 1997. Cointegration was not present between price and earnings where they were the only variables including in the model. After augmenting the models with variables representing the levels of foreign equity markets and exchange rate expectations, it appears that forces drive the levels of the Financial and Industrials Index toward its equilibrium levels with movements away from equilibrium representing opportunities to earn returns in excess of the market. A number of studies have developed hypotheses around investor behavioural finance as an explanation for the cointegration of prices and earnings, particularly interesting are those that model investors' reactions to new information

It has been argued that speculative asset bubbles form as a result of inappropriate responses to new information (Smidt, 1968). Basu (1977) argues that after a string of bad news, investors become overly pessimistic, resulting in companies becoming under-valued, as indicated by very low P/E ratios. Eventually, as it becomes clear that actual earnings will be greater than envisaged in overly pessimistic forecasts, stock prices adjust upwards. Similarly, companies with very high P/E ratios are considered over-valued, resulting in a subsequent downward adjustment in price. The low P/E effect is a special case of this price-ratio hypothesis. Ofer (1975)'s findings are at odds with the conclusions of Smidt (1968) and Basu (1977) in that he finds investors can successfully predict large changes in earnings growth.

De Bondt and Thaler (1985) consider the broader investor over-reaction hypothesis which argues that investors *over*-react to unexpected news. This stems from disproportionately more weight being placed on more recent information as compared to less recent information. The success of contrarian investment strategies—strategies involving buying the stock of companies that have performed poorly in the recent past and selling short those who have performed well (Chan, 1988)—is consistent with this hypothesis. De Bondt and Thaler (1985) find that 36 months after portfolio formation based on P/E, portfolios containing prior "losers" earn, on average, 25% more than portfolios of prior "winners". It is important to note the riskiness of the winners significantly exceeds that of the losers. In line with the results of De Bondt and Thaler (1985) for New York Stock Exchange–listed common stocks, Page and Way (1992) as well as Hsieh and Hodnett (2011) find three years post-formation, winners outperform losers on the JSE for the periods 1974 to 1989 and 1993 to 2009, respectively.

Some studies point to *under*-reaction. The most prominent example of under-reaction is the post-earnings announcement drift; stock prices continue to respond for about a year after the announcement of earnings (Fama, 1998). Ball and Brown (1968) are the first to document the post-earnings announcement drift, finding that the cumulative abnormal return of firms with positive earnings news continue to drift upward while those of firms with negative earnings news continue to drift downward. Rendleman, Jones and Latané (1987) posit that investors do not realise that quarterly earnings changes are serially correlated and therefore, they do not fully appreciate and exploit the information contained in earnings announcements. Due to unexploited information, Bernard and Thomas (1990) show that past earnings have predictive power with respect to abnormal returns at earnings announcements. A number of studies have noted that the original mispricing is followed by correction in price. This story was promoted by Dreman and Berry (1995) and the suggestion is that over-reaction and under-reaction may not be mutually exclusive. Under the mispricing correction hypothesis, initial over-reaction to an event results in mispricing and subsequent under-reaction results in price correction.

Campbell and Shiller (1998)'s results show that history may not necessarily repeat itself. They note that, while valuation ratios have been relatively stable over time, there are two arguments as to why historical patterns may not be useful in guiding the future evolution of these ratios. First, there has been a shift in attitudes toward the stock market. As the baby-boomers account for greater proportions of the economically active population, their attitudes toward risk become more important relative to previous generations. Baby-boomers are said to be more risk seeking, favouring stocks over bonds. Demand for stocks stretches valuations but does not affect future earnings, thereby increasing valuation ratios. High prices result in lower future returns unless stock demand increases towards the end of the holding period. A second reason stems from the work of Modigliani and Cohn (1979); they argue that the market erroneously discounts real dividends at nominal interest rates. This would result in stocks being over-valued in low-inflation periods resulting in low, subsequent realised returns. Shen (2000) notes two additional, popular arguments as to why the market P/E may be permanently higher and why the market may deviate from its historical pattern. The first is that globalisation and technological advancements enable the economy to grow faster which in turn, leads to faster earnings growth. The second argument is that the transaction costs involved with investing in the equity market have fallen. This leads to an increase in P/E in two ways. First, equity investing becomes relatively more attractive thereby increasing demand for stocks and leading to an increase in stock prices and P/E ratios. Second, lower transaction costs result in it becoming easier to diversify among different securities, reducing the riskiness of those investments. This reduction in perceived riskiness leads to investors demanding lower risk premia for holding stocks, leading to increased demand, increased stock prices and higher P/E ratios.

4.1 COINTEGRATION OF PRICE AND DIVIDEND YIELD

If the DDM holds, it is expected that a long run, equilibrium relationship should exist between stock price and dividends. In testing the present-value model of stock prices in the U.S., MacDonald and Power (1995) employ two VAR systems: one system with price and dividends and the second system with price, dividends and earnings retention ratio. For VAR system 1, no cointegration is found. For system 2, one cointegration relation is found among the three variables. Using data from the bourses of France, Germany, Japan, the U.K. and the U.S. covering the period 1973 through 2002, Capelle-Blanchard and Raymond (2004) test for cointegration, making sure to adjust the Engle-Granger two-step procedure for the skewness and excess kurtosis caused by stock market bubbles. Using the results of Campbell and Shiller (1988) that earnings are useful in forecasting future dividends, they include the additional variable of earnings in some of the cointegration tests in an attempt to account for the popularity of share repurchases. Capelle-Blanchard and Raymond (2004) find, *inter alia*, no cointegration between dividend yield and price in any of the markets.

4.2 COINTEGRATION OF PRICE AND EARNINGS

Molodovsky (1953) alludes to mean reversion when he states that stock prices fluctuate around a computable value. Campbell and Shiller (1998, 2001)'s results show that an extremely low market dividend yield reliably forecasts a drop in stock prices. Mean reversion of the market dividend yield results, predominantly, from an adjustment in price—the denominator—rather than in dividends—the numerator. Their results also indicate that an extremely high P/E reliably forecasts lower future returns. Here, again, reversion to the historical mean is predominantly due to changes in price rather than in earnings. This is echoed by Shiller (2002) when he points to subsequent real returns of between 10% and 20% following periods of low U.S. firm valuation ratios—between five and ten—and low or negative real returns following periods of high valuation ratios—between 20 and 25.

In the Campbell and Shiller (1987) model, prices and earnings are identified as being non-stationary time series. The market P/E ratio is, itself, a linear combination of two non-stationary variables and as such, a time series of market P/E ratios may be stationary. The model holds that prices and earnings are cointegrated and as such, time series of market P/E ratios are stationary and mean reverting. Carlson, Pelz and Wohar (2002) show empirically that the "normal" P/E ratio can change over time. The historical mean over the period 1870 through 2000 is estimated at 14.6. Using quarterly P/E data going back to 1945, they identify an upward break in the mean in the fourth quarter of 1992. The first regime has a mean P/E of 12.8 while the second regime has a mean P/E of 22.7. This suggests a higher P/E norm for the U.S. stock market—they estimate the new P/E band to be between 20 and 25. Carlson *et al.* (2002) point to the Nasdaq bubble which burst in 2000 as an example of mean reversion in action; unsustainably high P/E ratios for technology companies with large market capitalisations were followed by a sharp decline in their P/E ratios. They argue that this is consistent with Campbell and Shiller (1998, 2001)'s contention that the market is not perfectly efficient. Becker, Lee and Gup (2012) do not assume that P/E is mean-reverting. They do confirm that it is, however, mean reverting; they use a unit root test which models regime changes using a Fourier function and find that the S&P 500's P/E ratio is stationary around multiple breaks over the period 1871 through 2003.South African studies have generally investigated mean reversion of *returns*. These include, amongst others: Page and Way (1992); Muller (1999); Cubbin, Eidne, Firer and Gilbert (2006); Bailey and Gilbert (2007); Gilbert and Strugnell (2010) as well as Hsieh and Hodnett (2011). The results of these studies confirm the presence of mean reversion of returns on the Johannesburg Stock Exchange (JSE).

Shen (2000) indicates that increased P/E ratios result in lower (earnings) yields on equity investments, resulting in this asset class becoming less attractive relative to other asset classes and therefore causing poorer short-run performance. This story of investment funds flowing to assets with the most attractive yield is consistent with the "Fed model"—the term was coined following a report by the U.S. Federal Reserve (the Fed) noting that stock market earnings yields gravitate toward government bond yields (Weigand & Irons, 2008). In the Fed model, investors benchmark the market earnings yield to the 10-year Treasury note yield. Weigand and Irons (2008) argue that the Campbell-Shiller model and the Fed model cannot simultaneously be valid. According to Tatom (2002) it is well established that the nominal yield on 10-year Treasury notes is non-stationary. If the well-documented high correlation between market earnings yield and T-note yields is due to benchmarking, it is expected that the market E/P—and thus market P/E—would also be non-stationary and not predictably mean-reverting. Weigand and Irons (2008) find that that the market earnings yield and the 10-year Treasury note yield have been cointegrated since 1960, implying that the former is indeed benchmarked to the latter. If the market P/E is non-stationary, this would suggest that it is possible for the market P/E to stay above or below its historical mean for extended periods of time—it no longer exhibits mean-reverting properties. It is not clear whether the deviation from the valuation ratio's historical time series properties is permanent or long-lived but temporary. Koivu, Pennanen and Ziemba (2005) develop a quantitative, dynamic version of the Fed model. They conduct cointegration analysis of U.S., U.K. and German data and examine whether the Fed model has predictive power with respect to forecasting changes in stock prices, earnings and bond yields. For the period 1980 through 2003, it appears the Fed model does have some predictive power, but more so for the U.S. stock market.

Malkiel (2004) notes three criticisms levelled against the Fed model. First, comparing a nominal variable—T-note yield—to one that is essentially real—earnings yield—is inappropriate (Asness, 2003). Earnings and price are both nominal variables making their ratio a real variable. Earnings yield should, therefore, be compared to real interest rates. Second, the Fed model is able to describe how the market P/E is determined, but is not necessarily able to predict future returns. The Campbell-Shiller model has the opposite properties; it is not meant to explain how P/E ratios are determined but is designed to predict future returns. Third, the Campbell-Shiller model's predictions are consistent with the Gordon model; low returns should be expected to follow when P/E ratios are high. It would be quite natural to enquire whether the Fed model is predominantly a U.S. phenomenon. Estrada (2006) collects evidence from 20 developed markets to support his view that the Fed model does not travel well. Of the 20 countries in the sample, earnings yield and bond yields are cointegrated in only two—Ireland and New Zealand. He also finds that in 17 of the 20 countries, P/E ratios by themselves do a better job of forecasting returns than the Fed model.

5 METHODOLOGY

5.1 SAMPLING AND DATA COLLECTION

If P/E ratios are mean-reverting, very high or very low P/E ratios will revert to their historical average. This would suggest that periods of high P/E ratios should be followed by declining returns and periods of low P/E ratios should be followed by rising returns. In order to examine the relationships between P/E and subsequent price and P/E and subsequent dividend yield, equity indices were employed. The FTSE/JSE All Share index (ALSI) represents the South Africa broad-market index, representing 99% of the total market capital value of the JSE's Main Board–listed companies, while the Standard & Poor's 500 (S&P 500) and FTSE 100 represent the U.S. and U.K. equity markets, respectively. The former represents the top 500 listed U.S. companies in terms of market value, while the latter represents the top 100 listed U.K. companies. Use of these two indices enables a comparison with two of the most developed equity markets. Furthermore, the South African equity market has been segmented further by repeating the analysis on indices which are representative of different firm-size bands. The FTSE/JSE Top 40 index, FTSE/JSE Mid-Capitalisation index and FTSE/JSE Small Capitalisation index are comprised of the largest, medium-sized and small-sized JSE-listed companies. The Top 40 is comprised of the top 40 companies in terms of market value, listed on the JSE Main Board. The Mid Cap comprises the next 60 largest companies, while the Small Cap index consists of those firms which are part of the ALSI but are too small to be included in either the Top 40 or Mid Cap. All indices utilised are value-weighted. Monthly index P/E ratios, dividend yields and closing index values for the period July 1995 through June 2013 were retrieved from Thomson Reuters Datastream. Bhargava and Malhotra (2006) argue that monthly data is more appropriate than annual data because many investors take a short-term view of the market. They also suggest that monthly data leads to more reliable results.

5.2 DESCRIPTION OF OVERALL RESEARCH DESIGN

If the time series are non-stationary, it may be that some linear combination of these non-stationary time series is stationary (Engle & Granger, 1987). If such a combination exists, then the non-stationary time series are cointegrated. The stationary linear combination is the cointegrating equation and represents the long-run equilibrium relationship between those variables. The methodology of Johansen (1988, 1991) will be used to test for cointegration. If no cointegration is found, vector autoregression (VAR) will be used to examine the dynamic relationship between the variables. VAR is merely an extension of the standard auto-regressive (AR) model, where more than one endogenous variable is being studied. A *k*-variable VAR system will have *k* equations. Each endogenous variable is used as the dependent variable in its own equation and lagged versions of *all* variables studied are used as independent variables in each equation. For example, the system of equations for testing the relationship between price and P/E would be written as:

$$P_t = c_1 + a_{1,1}P_{t-1} + \cdots + a_{1,p}P_{t-p} + b_{1,1}(P/E)_{t-1} + \cdots + b_{1,p}(P/E)_{t-p} + e_{1,t} \qquad (7a)$$

$$(P/E)_t = c_2 + a_{2,1}P_{t-1} + \cdots + a_{2,p}P_{t-p} + b_{2,1}(P/E)_{t-1} + \cdots + b_{2,p}(P/E)_{t-p} + e_{2,t}. \qquad (7b)$$

In the system of equations, p refers to the lag order. The optimal lag order is chosen such that it minimises the various information criteria including the Akaike information criterion (AIC), Schwarz criterion (SC) and Hannan-Quinn criterion (HQ). The model can be denoted VAR(p). Similarly, the system of equations for testing the relationship between price and dividend yield is written as:

$$P_t = c_1 + a_{1,1}P_{t-1} + \cdots + a_{1,p}P_{t-p} + b_{1,1}(DY)_{t-1} + \cdots + b_{1,p}(DY)_{t-p} + e_{1,t} \qquad (8a)$$

$$(DY)_t = c_2 + a_{2,1}P_{t-1} + \cdots + a_{2,p}P_{t-p} + b_{2,1}(DY)_{t-1} + \cdots + b_{2,p}(DY)_{t-p} + e_{2,t}. \qquad (8b)$$

If a variable has a unit root but is not cointegrated, then the variables with unit roots should be differenced to render them stationary before estimating the model.

.

If cointegration is present then the vector error correction model (VECM) form should be used. The structure of a VECM differs from that of the unrestricted VAR in two respects: (1) the variables are differenced and (2) an error correction mechanism (ECM) u_{t-1} is added into both equations. To estimate the ECM, the co-integrating equation is estimated by ordinary least squares (OLS). Consider the system of equations represented by equations 7a and 7b; P is regressed on P/E as in:

$P_t = a + b(P/E)_t + u_t$, which can be re-arranged as (9a)

$$u_t = P_t - a - b(P/E)_t. \qquad (9b)$$

The ECM is then just a lagged version of u_t. The same is done for the price–dividend yield relation. The empirical specifications must be checked for stability.

The methodology up to this point enables an exploration of the relation between variables, but does not establish whether a relation is "causal"—as opposed to purely contemporaneous. A Granger causality test will be implemented to test for pair-wise causality between price and P/E as well as between price and dividend yield. P/E is said to *Granger*-cause P if P can be better predicted using the histories of both P/E and P than it can by using P alone. The Toda and Yamamoto (1995) surplus-lag approach will be used. The procedure utilises a VAR system estimated in *levels* with additional lags of the dependent variables treated as exogenous to the system. After estimating the VAR, a Wald test of parameter restrictions is run. This methodology is versatile as it can be used regardless of whether the process is (trend-) stationary, integrated of an arbitrary order or even cointegrated of an arbitrary order. The surplus-lag approach avoids so-called “pre-test testing”.

6 RESULTS

6.1 TESTING FOR STRUCTURAL CHANGE

Table 1 contains abbreviations and descriptions of the variables used in the analysis to follow. To identify multiple, unknown break points in a time series, the sequential Bai and Perron (1998) test is used. It tests for $l+1$ versus l sequentially determined breaks. The test allows for a maximum of five breaks, by employing 15% trimming and using the 10% significance level. The error distributions are allowed to differ across breaks. Before running the test, the variable of interest is regressed on a constant by ordinary least squares (OLS) with a heteroscedasticity- and autocorrelation-consistent (HAC) covariance matrix. The break dates determined for each index’s variables are contained in table 2. The Bai-Perron tests detect no structural breaks for any of the ALSI’s variables or the Small Cap’s variables. Where present, the break points for the price series are close together in time across indices; around mid-2005 and in the third quarter of calendar year 2010. The P/E series for the FTSE 100 has one break—July 2004—while the Mid Cap’s P/E series also has one break—August 2008. The only dividend yield series which experiences structural change is that of the Mid Cap index. When the breaks are re-estimated by repartitioning, the dates are identical.

Table 1: Variables and their descriptions

VARIABLE	DESCRIPTION
LPI	Natural logarithm of index closing value
LPE	Natural logarithm of index P/E
LDY	Natural logarithm of index dividend yield
DLPI	First difference of LPI
DLPE	First difference of LPE
DLDY	First difference of LDY

Table 2: Sequentially determined break-points

Indices	LPI		LPE		LDY	
	1	2	1	2	1	2
FTSE/JSE All Share						
FTSE 100			2004M07			
S&P 500	2010M05					
FTSE/JSE Top 40	2005M07	2010M09				
FTSE/JSE Mid Cap	2005M07		1998M08		2000M10	
FTSE/JSE Small Cap						

6.2 TESTING FOR UNIT ROOTS

Drawing inferences from the parameter estimates obtained from any regression procedure estimated using least squares methodology is done on the basis that the time series variables are stationary. A time series whose mean and autocovariances are not dependent on time are said to be weakly—or covariance—stationary. Differencing operations may render a series stationary. Where this is the case, the series is said to be integrated with order *d*—I(*d*). The order of integration refers to the number of unit roots contained in the series, or alternatively, the number of differencing operations that must be performed on the series to render it covariance stationary. A series which is stationary in levels is referred to as I(0), whilst a series which is non-stationary in levels, but stationary in first-differenced form is referred to as I(1). The stationarity of a series can be determined by a running unit root test such as the popular Augmented Dickey-Fuller (ADF) test. The null hypothesis of the ADF test is that the series has a unit root, i.e. the series is non-stationary.

Table 3: Unit roots test results—levels, broad-market indices

Variable	Test equation add-ins	Lambda	Test statistic	Critical values		
				1%	5%	10%
FTSE/JSE All Share						
LPI	Intercept, trend	N/A	-2,2000	-4,0013	-3,4309	-3,1391
LPE	Intercept	N/A	-2,9464	-3,4607	-2,8748	-2,5739
LDY	Intercept	N/A	-2,6390	-3,4607	-2,8748	-2,5739
FTSE 100						
LPI	Intercept	N/A	-2,0955	-3,4607	-2,8748	-2,5739
LPE	None	50%	-2,9464	-4,4900	-3,9300	-3,6500
LDY	Intercept, trend	N/A	-2,2463	-4,0013	-3,4309	-3,1391
S&P 500						
LPI	Intercept	N/A	-2,1773	-3,4607	-2,8748	-2,5739
LPE	None	83%	-2,5964	-4,3500	-3,7750	-3,4850
LDY	Intercept, trend	N/A	-2,7509	-4,0013	-3,4309	-3,1391

Where there are structural breaks—in the level and/or trend—of a series, the ADF test may erroneously indicate the presence of a unit root. Perron (1989) developed modified unit root test procedures which allow for a one-time change in the intercept and/or slope of the series. Where necessary in this study, Perron's "case B" methodology is followed to estimate the test statistics. To account for the structural change, each time series is regressed on an intercept, a linear time trend variable and a trend variable which takes the value of zero up to an including the break-date followed by 1, 2, 3… from the break-date to the end of the sample period. The residuals from the first regression are extracted and the ADF test is run on them.

Table 4: Unit root tests results—levels, ALSI sub-indices

Variable	Test equation add-ins	Lambda	Test statistic	Critical values 1%	5%	10%
FTSE/JSE Top 40						
LPI	None	55%	-2,3960	-4,4950	-3,9350	-3,6500
LPE	Intercept	N/A	-3,0556	-3,4607	-2,8748	-2,5739
LDY	Intercept	N/A	-2,7035	-3,4607	-2,8748	-2,5739
FTSE/ JSE Mid Cap						
LPI	None	56%	-2,1608	-4,4950	-3,9350	-3,6500
LPE	None	18%	-3,5281	-4,2450	-3,6200	-3,3200
LDY	None	30%	-2,7304	-3,4607	-2,8748	-2,5739
FTSE/JSE Small Cap						
LPI	None	55%	-1,5785	-4,4950	-3,9350	-3,6500
LPE	Intercept	N/A	-3,6082	-3,4607	-2,8748	-2,5739
LDY	Intercept	N/A	-2,7559	-3,4607	-2,8748	-2,5739

Table 5: Unit root test results—first difference, broad-market indices

Variable	Test equation add-ins	Lambda	Test statistic	Critical values 1%	5%	10%
FTSE/JSE All Share						
DLPI	Intercept, trend	N/A	-14,3180	-4,0015	-3.4310	-3.1391
DLPE	Intercept	N/A	-14,6956	-3.4609	-2.8749	-2,5740
DLDY	Intercept	N/A	-14,0546	-3,4609	-2.8749	-2,5740
FTSE 100						
DLPI	Intercept	N/A	-13,5211	-3,4609	-2,8749	-2,5740
DLPE	None	50%	-15,4575	-4,4900	-3,9300	-3,6500
DLDY	Intercept, trend	N/A	-14,3778	-4,0015	-3,4310	-3,1391
S&P 500						
DLPI	Intercept	N/A	-13,1094	-3,4609	-2,8749	-2,5740
DLPE	None	83%	-15,2382	-4,3500	-3,7750	-3,4850
DLDY	Intercept, trend	N/A	-16,1115	-4,0015	-3,4310	-3,1391

The Dickey-Fuller test equation contains neither an intercept nor a linear trend because the first equation acts as a filter. The test statistics estimated are then compared to the asymptotic critical values derived by Perron and as corrected by Perron and Vogelsang (1993). The new critical values are, in part, dependent upon how far into the sample period the break occurs—a parameter known as lambda, λ. The results of the ADF and modified ADF tests are contained in tables 3 to 6. Price is found to be I(1) across all of the indices. P/E is found to be stationary in levels for all FTSE/JSE indices whilst being I(1) for the FTSE 100 and S&P 500 indices. Dividend yield is I(0) for the ALSI, Top 40 and Small Cap indices and I(1) for the FTSE 100, S&P 500 and Mid Cap indices. It it worthwhile noting that in instances the null hypothesis is rejected at only a 10% level of significance and, as such, the results of the ADF test should be approached with caution. This is true for the Mid Cap's P/E series as well as for the dividend yield series of FTSE/JSE indices.

Table 6: Unit root test results—first difference, ALSI sub-indices

Variable	Test equation add-ins	Lambda	Test statistic	Critical values		
				1%	5%	10%
FTSE/JSE Top 40						
DLPI	None	55%	-14,6998	-4,4950	-3,9350	-3,6500
DLPE	Intercept	N/A	-14,5954	-3,4609	-2,8749	-2,5740
DLDY	Intercept	N/A	-13,9034	-3,4609	-2,8749	-2,5740
FTSE/ JSE Mid Cap						
DLPI	None	56%	-12,6475	-4,4950	-3,9350	-3,6500
DLPE	None	18%	-16,7617	-4,2450	-3,6200	-3,3200
DLDY	None	30%	-13,9034	-3,4609	-2,8749	-2,5740
FTSE/JSE Small Cap						
DLPI	None	55%	-1,5785	-4,4950	-3,9350	-3,6500
DLPE	Intercept	N/A	-3,6082	-3,4607	-2,8748	-2,5739
DLDY	Intercept	N/A	-2,7559	-3,4607	-2,8748	-2,5739

6.3 VAR PARAMETER ESTIMATES

6.3.1 <u>Price and P/E</u>

Tables 7 and 8 contain the vector autoregression parameter estimates where index closing prices and price-to-earnings ratios are the endogenous variables. Tables 7 and 8 also contain the *t*-statistics associated with the coefficients—in the grey rows—as well as the (adjusted) coefficients of multiple determination for the broad-market indices and ALSI sub-indices, respectively. The VARs were estimated with an exogenous constant term. In an attempt to capture the adverse effect of the financial crisis of 2008 on stock prices and valuation ratios, an exogenous indicator variable (FINCRISIS) is included, which takes the value of 1 for each month between 2007M08 and 2009M03, inclusive, and 0 otherwise. In addition, transitory blip dummies—of the form (…0, 0, +1, -1, 0, 0…)—were added as exogenous variables until the residuals were "well-behaved". The purpose of the blip dummies is to reduce the impact of *random* shocks which disappear in the next period or in the next few periods and which may result in the residual distribution deviating substantially from the theoretical distribution.

ALSI. Looking first at the price equation, the coefficient of the first lag of price is relatively small and positive, but statistically insignificantly different from zero. The second and third lags are negative but not statistically significant. The lagged versions of P/E all have small, positive coefficients, suggesting that an increase in P/E leads to an expected increase in price. The increases are not statistically significant, however. With respect to the P/E equation, the first and third lags of price are negative but statically insignificant. The first and third lags of P/E are positive and relatively large in economic terms and both are significant with at least 90% confidence. The second lag has a negative coefficient but the *t*-statistic is very small, indicating insignificance.

Table 7: VAR parameter estimates—price and P/E, broad market indices

	FTSE/JSE ALSI		FTSE100		S&P500	
	DLPI	LPE	DLPI	DLPE	DLPI	DLPE
DLPI(-1)	0,09952	-0,03283	0,22402	-0,02477	0,24146	-0,30667
	0,89576	-0,25368	2,74341	-0,22230	2,94912	-2,59866
DLPI(-2)	-0,00946	0,00022	-0,17323	-0,22380	0,00805	0,03524
	-0,09343	0,00187	-2,21502	-2,09742	0,11141	0,33824
DLPI(-3)	-0,04962	-0,13883	-0,01600	-0,23434	-0,11133	-0,48797
	-0,49351	-1,18552	-0,20580	-2,20930	-1,47514	-4,48573
DLPI(-4)			0,19466	0,33092	0,00512	0,17657
			2,58282	3,21825	0,06942	1,66117
(D)LPE(-1)	0,03161	0,20080	0,06418	0,33141	-0,00916	0,58679
	0,34445	1,87900	1,18113	4,47004	-0,19534	8,68662
(D)LPE(-2)	0,02999	-0,00108	0,12201	0,16366	-0,01828	0,06454
	0,36301	-0,01119	2,50607	2,46397	-0,52074	1,27523
(D)LPE(-3)	0,06328	0,19836	-0,00633	0,21491	0,07595	0,28626
	0,76441	2,05738	-0,12575	3,12903	2,19617	5,74300
(D)LPE(-4)			-0,07128	-0,21327	0,00158	-0,12179
			-1,35432	-2,97012	0,04631	-2,47120
C	0,01157	0,00393	0,00396	0,00063	0,00620	0,00354
	2,89585	0,84432	1,67135	0,19388	2,09201	0,82898
FINCRISIS	-0,02762	-0,02050	-0,02130	-0,00962	-0,03009	-0,01402
	-2,27162	-1,44769	-2,64033	-0,87392	-2,84727	-0,92063
R^2	0,19204	0,26709	0,46240	0,56793	0,37745	0,67894
Adj. R^2	0,13060	0,21100	0,37970	0,50145	0,28167	0,62954

FTSE 100. In the price equation, the coefficients attaching to the first and fourth lags of price are similar in size and are both significant at 1%. The second lag has the opposite sign and is significant at 5%. The second lag of P/E has positive and statistically significant coefficient, suggesting an increase in P/E leads to a meaningful increase in price, two months into the future. None of the other coefficients are significant. In the P/E equation the price variable coefficients are all negative except for the fourth lag. The second and fourth lags have statistically significant coefficients. The coefficients of the lagged P/E variables are large and mostly positive. These significant coefficients indicate that P/E is—to some extent—dependent on its recent history.

S&P 500. With respect to the price equation, the coefficient attaching to the first lag is highly significant and positive. The third lag of P/E is the only lag of that variable which is statistically significant and it has a small, positive coefficient estimate. An increase in P/E also leads to an expected, albeit smaller, increase in price, in three months' time. In the P/E equation, the first and third lags of price are highly significant and negative, suggesting an increase in price leads to an expected decrease in subsequent P/E. P/E appears to be highly dependent on its past values as well with the first, third and fourth lags of P/E being highly significant.

Top 40. In the price equation none of the lagged price variables are significant nor are any of the P/E variables. In the P/E equation, none of the price-variable coefficients are significant, whilst the first and third lags of P/E are significant at the 1% level.

Mid Cap. There are no statistically significant price variable coefficients in the price equation, however, the first lag of P/E is statistically significant and the coefficient is positive. Once again, it appears that an increase in price is expected to follow on from an increase in P/E, here one period in the future. In the P/E equation, the lagged price variables have negative coefficients, which are statistically indistinguishable from zero. The first lag of P/E has a positive, statistically and economically significant coefficient.

Table 8: VAR parameter estimates—price and P/E, ALSI sub-indices

	TOP 40		**MID CAP**		**SMALL CAP**	
	DLPI	LPE	DLPI	LPE	DLPI	LPE
DLPI(-1)	0,12729	-0,10221	0,07031	-0,00487	0,33005	0,38432
	1,17889	-0,82451	0,80016	-0,04519	4,95621	2,37678
DLPI(-2)	0,02131	0,04703	-0,11706	-0,00700	-0,10487	-0,10898
	0,21982	0,42259	-1,43924	-0,04519	-1,54236	-0,66004
DLPI(-3)	-0,02940	-0,15706				
	-0,30804	-1,43329				
LPE(-1)	0,04282	0,32605	0,21075	0,19670	0,00991	0,29629
	0,47802	3,17054	3,58159	2,72811	0,33934	4,17928
LPE(-2)	-0,03816	-0,15533	0,03762	0,03585	0,02262	-0,04475
	-0,48753	-1,72852	0,70149	0,54546	0,90740	-0,73907
LPE(-3)	0,06977	0,28437				
	0,89520	3,17794				
C	0,01006	0,00367	0,01259	0,00073	0,01085	-0,00064
	2,50970	0,79720	3,63761	0,17083	3,40943	-0,08278
FINCRISIS	-0,02652	-0,01929	-0,02262	-0,02039	-0,02728	-0,01536
	-2,13094	-1,35049	-2,14169	-1,57616	-2,67826	-0,62089
R^2	0,23156	0,32678	0,28107	0,55730	0,38286	0,38235
Adjusted R^2	0,16851	0,27154	0,23023	0,52600	0,32560	0,32504

Small Cap. In the price equation, the first lag of price has a relatively large, positive coefficient which is also highly significant. Both lags of P/E produce positive coefficients, though they are statically indistinguishable from zero. In the P/E equation, the coefficient attaching to the first lag of price is positive, relatively large in size and highly significant. Likewise, the coefficient of the first lag of P/E is also positive and statistically significant.

The constant terms in each price equation across all of the indices are of similar size and are all statistically meaningful at conventional levels of significance. The average, adjusted R^2 across the price equations is 25.27%. For the broad-market indices, the figure is 26.40% and lower at 24.14% for the ALSI sub-indices. The average, adjusted R^2 for the P/E equations is 41.08%. The average, adjusted R^2 for the broad-market indices is higher than for the ALSI sub-indices; 44.75% and 37.42% respectively. The exogenous financial crisis dummy was only significant in the price equations and the coefficients were of a similar—small—size with a negative coefficient. This is as expected.

6.3.2 Price and Dividend Yield

Tables 9 and 10 contain the VAR parameter estimates where index closing prices and dividend yield are the endogenous variables in the systems of equations for the broad-market indices and the ALSI sub-indices, respectively. Here, again, an exogenous constant term is included in each of the specifications.

ALSI. In the price equation, all of the lagged price variables have positive coefficients, but are all statistically indistinguishable from zero. The coefficient of the first lag of dividend yield is negative and statistically significant at a 5% level, implying that an increase in dividend yield leads to an expected decrease in price, one month in the future. Turning attention to the dividend yield equation, the first lag of price has a statistically significant and positive coefficient, implying an increase in price leads to an expected increase in dividend yield. The coefficient on the first lag of dividend yield is highly significant and positive, suggesting movements in dividend yield are related to their immediate past values.

FTSE 100. With respect to the price equation for the FTSE 100, the third lag of price has a negative coefficient which is significant at the 5% level while the fourth lag of price has a positive coefficient which is significant at the 1% level. The first and third lags of dividend yield have negative coefficients which are significant at less than the 5% level. The second lag has a positive coefficient which is marginally significant. In the second equation where dividend yield is the dependent variable, the fourth lag of price is negative and significant at 10%. Both the first and third lags of dividend yield have positive coefficients which are highly significant. The second lag of dividend yield has a negative coefficient, which is less significant; significant at only the 10% level.

Table 9: VAR parameter estimates—price and dividend yield, broad market indices

	FTSE/JSE ALSI		FTSE 100		S&P 500	
	DLPI	DLDY	DLPI	DLDY	DLPI	DLDY
DLPI(-1)	0,03188	0,21335	0,08574	0,02195	0,34095	-0,16736
	0,30136	2,01866	0,88059	0,20120	3,64152	-1,83075
DLPI(-2)	0,05060	-0,00870	0,07702	-0,07532	-0,00896	0,00470
	0,47835	-0,08230	0,78235	-0,68275	-0,01018	0,05476
DLPI(-3)	0,00925	0,00615	-0,19130	0,14327	0,03788	0,01460
	0,09221	0,06144	-1,97879	1,32255	0,42262	0,16685
DLPI(-4)			0,25474	-0,17530	0,00045	-0,06911
			2,68022	-1,64595	0,00530	-0,82688
DLDY(-1)	-0,21840	0,53810	-0,20467	0,28825	-0,10362	0,78099
	-2,18155	5,38082	-2,19858	2,76333	-1,07469	0,82961
DLDY(-2)	-0,00530	0,05819	0,15387	-0,17190	0,01581	0,02657
	-0,05540	0,60849	1,74650	-1,74120	0,16662	0,28674
DLDY(-3)	0,01355	0,05174	-0,18298	0,22020	0,09884	-0,07153
	0,15737	0,60149	-2,04667	2,19808	1,04271	-0,77284
DLDY(-4)			0,14237	-0,06090	0,09262	-0,14575
			1,62038	-0,61852	0,99058	-1,59657
C	0,00993	-0,00275	0,00424	-0,00140	0,00468	-0,00014
	2,64533	-0,73427	1,82455	-0,53719	1,78921	-1,59657
FINCRISIS	-0,01924	0,01632	-0,02116	0,01174	-0,02305	-0,00135
	-1,72045	1,46092	-2,68416	1,32870	2,57314	-0,05303
R^2	0,38903	0,52409	0,47358	0,42864	0,54266	0,47186
Adjusted R^2	0,25336	0,45426	0,39591	0,34435	0,45740	0,37340

S&P 500. The first lag of price in the price equation has a positive coefficient which is statistically significant at 1%. The lags of dividend yield are mostly positive, but are—statistically—no different to zero. In the dividend yield equation, the first lag of price has a negative coefficient which is statistically significantly different from zero with 90% confidence. The lagged versions of dividend yield included in the equation are not significant suggesting dividend yield is not dependent on immediate past values of the series.

Top 40. In the price equation, all lagged price variables have positive coefficients; however, none are statistically significant. The coefficient estimates for the first and second lags of dividend yield are negative, while the coefficient attaching to the third lag is positive. None of the lagged versions of dividend yield have statistically significant coefficients, as evidenced by the small *t*-statistics.

Mid Cap. With respect to the price equation, the first lag of price is highly significant with a positive coefficient. In the price equation, none of the dividend yield variables have statistically meaningful coefficients. Where dividend yield is the dependent variable, the second lag of price has a negative coefficient which is significant at a 5% level of significance. The second lag of dividend yield has a negative coefficient which is significant at 10%.

Table 10: VAR parameter estimates—price and dividend yield, ALSI sub-indices

	TOP 40		MID CAP		SMALL CAP	
	DLPI	DLDY	DLPI	DLDY	DLPI	DLDY
DLPI(-1)	0,10736	0,20754	0,25741	-0,09735	0,26828	-0,20423
	0,99374	1,82284	3,26359	-0,90702	3,66416	-1,79636
DLPI(-2)	0,05827	0,02200	-0,01093	-0,23372	-0,09424	-0,11857
	0,58518	0,20961	-0,18462	-2,18440	-1,13109	-1,06211
DLPI(-3)	0,04117	-0,09749			-0,02661	0,08990
	0,40578	-0,91170			-0,37557	0,81693
DLPI(-4)					-0,12033	0,16509
					-1,77554	1,56878
DLDY(-1)	-0,12990	0,55525	-0,01093	0,03756	-0,06767	0,05660
	-1,34256	5,44544	-0,18620	0,46629	-1,53037	0,82420
DLDY(-2)	-0,02332	0,08243	0,00315	-0,13550	0,05425	-0,14165
	-0,27545	0,92391	0,05783	-1,83063	1,27612	-2,14582
DLDY(-3)	0,02198	-0,04520			-0,03243	-0,20119
	0,24774	-0,48332			-0,77153	-3,08243
DLDY(-4)					-0,01314	0,08903
					-0,30437	1,32809
C	0,00858	-0,00103	0,01038	0,00345	0,01279	-0,00462
	2,18671	-0,24925	2,99426	0,73131	3,78408	-0,88062
FINCRISIS	-0,01765	0,01598	-0,02318	0,01047	-0,03213	0,03888
	-1,49712	1,28614	-2,13226	0,70776	-3,03796	2,36779
R^2	0,35161	0,53640	0,22571	0,28659	0,34829	0,42206
Adjusted R^2	0,24414	0,45956	0,18334	0,24755	0,30528	0,38392

Small Cap. Where price is the dependent variable, the first lag has a positive coefficient which is highly significant whereas the fourth lag of price has a negative coefficient which is significant, but only with 90% confidence. None of the dividend yield variables are significant in the statistical sense. In the dividend yield equation, the first lag of price has a negative, statistically significant coefficient. The second and third lags of the dividend yield are both negative whilst the former is significant at 5% and the latter significant at 1%.

The constant terms across the indices are similar in size. Furthermore, they are statistically significant in each of the price equations. The financial crisis dummy variable has a negative coefficient in each of the price equations. The size of the coefficients is similar across indices. The financial crisis dummy variable is significant at 1% for each of FTSE 100, S&P 500, Mid Cap and Small Cap indices. The variable is significant at a 5% level of significance for the All Share index but is not statistically significant for the Top 40 index. The average, adjusted R^2 across the price equations is 30.66%. For the broad-market indices, the figure is 36.89% and lower at 24.43% for the JSE sub-indices. The average, adjusted R^2 for the dividend yield equations is 37.72%. The average, adjusted R^2 for the broad-market indices is higher than for the JSE sub-indices; 39.07% and 36.37% respectively.

6.4 STABILITY DIAGNOSTICS

6.4.1 Price and P/E

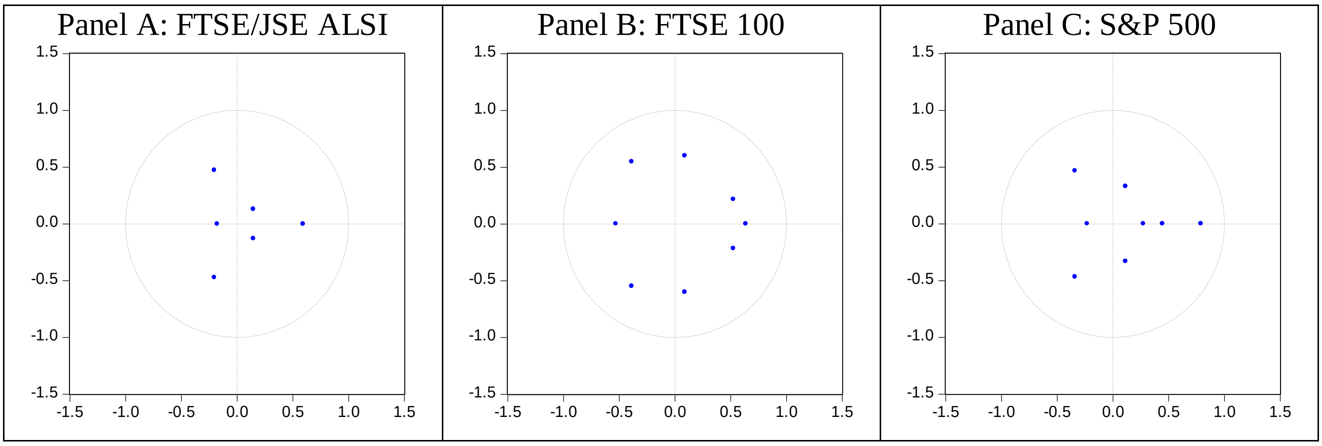

Figure 1: Inverse roots of AR characteristic polynomial—price and P/E, broad-market indices

Figures 1 and 2 contain graphs of the inverse roots of the autoregressive (AR) characteristic polynomial of the estimated VARs for the broad-market indices and ALSI sub-indices, respectively. For a VAR system to be stable—that is, stationary—all the roots must lie within the unit circle. A VAR system that is not stable would result in inferences being invalid. All of the roots in each of the panels of figures 1 and 2 lie within the unit circle and thus the systems of equations each satisfy the stability condition.

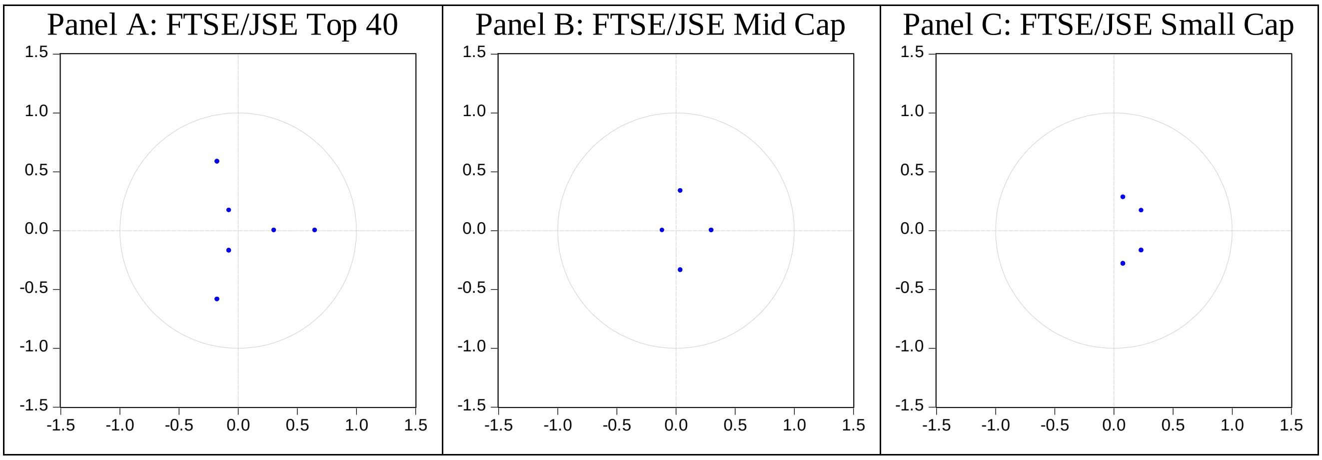

Figure 2: Inverse roots of AR characteristic polynomial—price and P/E, ALSI sub-indices

6.4.2 Price and Dividend Yield

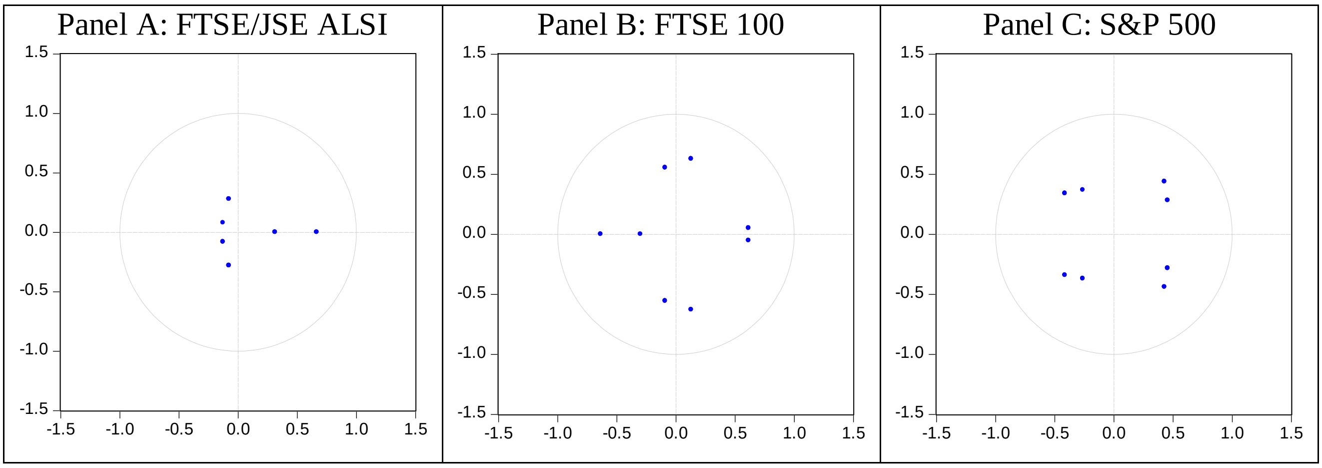

Figure 3: Inverse roots of AR characteristic polynomial—price and DY, broad-market indices

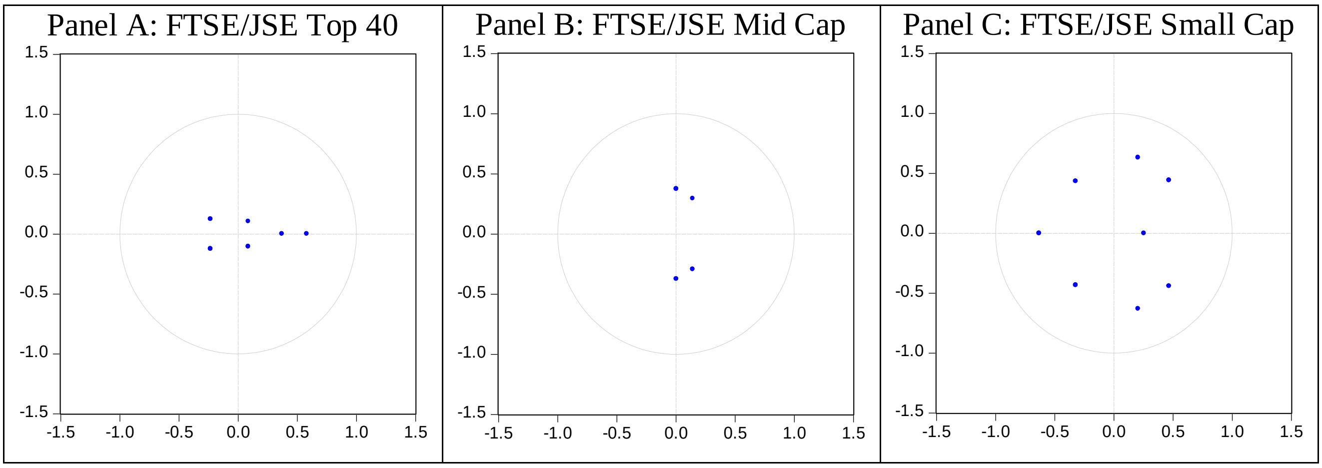

Figure 4: Inverse roots of AR characteristic polynomial—price and DY, ALSI sub-indices

Figures 3 and 4 contain the inverse AR roots of the VAR systems for the broad-market indices and the ALSI sub-indices, respectively, where price and dividend yield are the endogenous variables.

The inverse roots in each panel of both figures show that the roots lie within the unit circle and thus these VAR systems also satisfy the stability condition.

6.5 RESIDUAL DIAGNOSTICS

VAR systems are estimated by OLS and as such, the classical assumption of independent and identically distributed residuals must be checked. Two major problems with regression assumptions are heteroscedasticity—i.e. non-constant variance—and autocorrelation of the error terms. Should either of these problems be present, the OLS estimators are still unbiased and consistent, but are no longer efficient. The normality test chosen is the multivariate extension of the Jarque-Bera test, which is a joint test of skewness and kurtosis. The test's null hypothesis is that the residuals are normally distributed—the distribution is both symmetrical and mesokurtic.

The Doornik-Hansen orthogonalisation method is employed, whereby the factorisation matrix is given by the inverse square root of the residual correlation matrix. This method of orthogonalisation is widely used; the Doornik-Hansen approach is popular because the statistics are not dependent upon the order or scale of the variables, not that this is a problem here. The multivariate version of the White test is used to test for heteroscedasticity because, again, it is widely used in the literature. In the particular version of the test employed, the test statistic is estimated by regression of each cross product of the residuals on the cross products of the regressors. Next, the joint significance of the regression is tested using both the levels as well as the squares of the regressors with the null hypothesis being that the residuals are homoscedastic. A multivariate Lagrange multiplier (LM) test is used to identify the presence of any autocorrelation in the residuals. The null hypothesis for the test is that there is no serial correlation up to a certain lag.

6.5.1 Price and P/E

For the broad-market indices, the Jarque-Bera test shows that the residuals are normally distributed for the FTSE 100 and the S&P500. Whilst the ALSI's residuals are not mesokurtic, they are symmetrically distributed about the mean. The residuals for each of the indices are homoscedastic, as evidenced by the White test. The null hypothesis of the autocorrelation LM test is rejected in two—of ten—instances for the ALSI, one instance for the FTSE 100 and three instances for the S&P 500. It should be noted that normality of the residual distribution is not a necessary condition for most of the tests related to VAR systems (Lütkepohl, 2011). The residuals of the Top 40 appear to be both symmetrical as well as mesokurtic. The distribution of the residuals for the Mid Cap index and

Small Cap index are symmetrical. For each of the indices, the residuals have constant variance. There is no autocorrelation present in the residuals for the Mid Cap index, however, there are a few instances where the null hypothesis of the autocorrelation LM test are rejected.

Table 11: Normality and heteroscedasticity test p-values—price and P/E

Indices	Normality			Hetero-scedasticity
	Skewness	Kurtosis	Jarque-Bera	White
FTSE/JSE All Share	0,7482	0,0143	0,0593	0,4956
FTSE 100	0,1738	0,3457	0,2291	0,2555
S&P 500	0,1081	0,1924	0,1013	0,1198
FTSE/JSE Top 40	0,9482	0,6685	0,9229	0,1088
FTSE/JSE Mid Cap	0,4084	0,0006	0,0023	0,1094
FTSE/JSE Small Cap	0,5384	0,0014	0,0063	0,8347

Table 12: Autocorrelation LM test p-values—price and P/E

Lag	FTSE/JSE All Share	FTSE 100	S&P 500	FTSE/JSE Top 40	FTSE/JSE Mid Cap	FTSE/JSE Small Cap
1	0,0070	0,6924	0,0000	0,0230	0,1839	0,0000
2	0,5711	0,0681	0,0088	0,2875	0,4880	0,5671
3	0,8879	0,6936	0,3578	0,7632	0,3864	0,0231
4	0,4395	0,0033	0,5465	0,2874	0,1614	0,4355
5	0,2521	0,4198	0,2176	0,0900	0,5341	0,4809
6	0,0403	0,2644	0,6182	0,0156	0,0738	0,9400
7	0,2806	0,7173	0,0716	0,5141	0,7042	0,2179
8	0,0972	0,5815	0,7404	0,5016	0,9230	0,0148
9	0,8793	0,8848	0,9853	0,8072	0,5080	0,0085
10	0,7820	0,6421	0,5845	0,5439	0,0682	0,4627

6.5.2 Price and Dividend Yield

Table 13: Normality and heteroscedasticity test p-values—price and DY

Indices	Normality			Hetero-scedasticity
	Skewness	Kurtosis	Jarque-Bera	White
FTSE/JSE All Share	0,5037	0,0070	0,0236	0,5737
FTSE 100	0,2062	0,0784	0,0829	0,1134
S&P 500	0,9214	0,0231	0,1031	0,2551
FTSE/JSE Top 40	0,5712	0,8331	0,8293	0,4611
FTSE/JSE Mid Cap	0,1495	0,0000	0,0000	0,1595
FTSE/JSE Small Cap	0,2106	0,0000	0,0000	0,1544

Where price and dividend yield are the endogenous variables, the distribution of the S&P 500's residuals appear to be jointly symmetrical and mesokurtic with constant variance. The residuals for the ALSI and FTSE 100 are only symmetrical and homoscedastic. The null hypothesis of the autocorrelation LM test is rejected in a few instances but fails to be rejected in the majority of cases.
The Jarque-Bera test indicates that the residuals of the Top 40 are normally distributed. The residuals for the Mid Cap and Small Cap exhibit no skewnesss. All three indices have homoscedastic residuals. There is no evidence of serial correlation in the Top 40's residuals; the probability values exceed 10% in the majority of cases.

Table 14: Autocorrelation LM test p-values—price and DY

Lag	FTSE/JSE All Share	FTSE 100	S&P 500	FTSE/JSE Top 40	FTSE/JSE Mid Cap	FTSE/JSE Small Cap
1	0,4248	0,0479	0,0272	0,2891	0,2639	0,7410
2	0,2358	0,0668	0,0144	0,2950	0,7595	0,9673
3	0,0211	0,3383	0,5292	0,0117	0,4688	0,0362
4	0,1632	0,4041	0,2710	0,1190	0,1055	0,8310
5	0,2907	0,0469	0,2261	0,5161	0,5695	0,9178
6	0,4875	0,0041	0,7651	0,0807	0,5115	0,3716
7	0,3817	0,3020	0,3710	0,3261	0,5214	0,9413
8	0,2043	0,2568	0,5654	0,6203	0,4814	0,0770
9	0,0477	0,6287	0,1821	0,1981	0,4466	0,0151
10	0,8527	0,5288	0,3987	0,8367	0,4626	0,6181

6.6 TESTING FOR COINTEGRATION

The Johansen cointegration test is performed using the estimated VARs allowing for a linear deterministic trend in the levels of the data. The trace test statistics tests the null hypothesis of r cointegrating equations against the alternative hypothesis of k cointegrating equations, where k refers to the number of endogenous variables for $r = 0, 1, \dots, k-1$. It is important to account for any structural change when running the Johansen test. To this end, the $H_l(r)$ test of Johansen, Mosconi and Nielsen (2000) is used where appropriate. The asymptotic distribution of this test differs from that of the usual trace test. Also, the asymptotic critical values are dependent on how far into the sample period the break-point occurs as well as on $p - r$, where p is the number of endogenous variables in the system and r the cointegration rank—the number of cointegrating relations. To construct the modified trace test statistics, a number of exogenous variables are added. First, a linear time trend variable is included. Second, a dummy variable is added which takes the value of zero up to and including the break-point and one otherwise. This variable has the same lag order as the relevant

VAR. Third, the interaction of this variable and the linear time trend variable is also included. Fourth, an indicator variable which takes the value of one at the beginning of the next regime and zero otherwise. Fifth, lagged versions of up to the lag order of the Johansen test, i.e. one lag less than that of the VAR itself. The trace test statistics and modified trace test statistics along with their corresponding 5% critical values are contained in tables 15 and 16. Strictly speaking, it is necessary for two variables to be integrated of the same order for cointegration to be present. As such, the results are only shown for those pairs of variables where cointegration is possible. In relation to price and P/E, neither of the pairs of variables appear to be cointegrated. Similarly, no cointegrating relations were found in relation to any of the indices in table 16. These results show that it is unnecessary to re-write the VARs in VECM form.

Table 15: Cointegration test results—price and P/E

Indices	None			At most one		
	Test Statistic	Critical Value	Prob.	Test Statistic	Critical Value	Prob.
FTSE 100	15,1088	37,4180	0,9677	7,0846	18,9020	0,8262
S&P 500	19,2863	33,6940	0,5857	5,7510	16,6750	0,5905

Table 16: Cointegration test results—price and DY

Indices	None			At most one		
	Test Statistic	Critical Value	Prob.	Test Statistic	Critical Value	Prob.
FTSE 100	5,6328	15,4947	0,7384	0,1109	3,8415	0,7391
S&P 500	16,7183	25,8721	0,4363	4,3668	12,5180	0,6886
FTSE/JSE Mid Cap	35,9812	47,9530	0,3975	15,6542	24,8190	0,4388

6.7 IMPULSE RESPONSE FUNCTIONS AND VARIANCE DECOMPOSITION

The impulse response functions (IRF) are able to show how a series evolves over time after a shock in each of the innovation terms. They plot the response over time of each of the endogenous variables to a shock in each of the equations. The impulses are transformed using the Cholesky factor obtained from the residual covariance matrix, adjusted for degrees of freedom. The accumulated IRF is the cumulative sum of the IRF. Not only do IRFs illustrate the effect that a one-time shock in, say, P/E has on future price, but they can also confirm whether a VAR system is stable. For stationary VARs, the impulse response should die out whilst the accumulated response should tend towards some non-zero constant. The reported IRFs are quantitatively similar to those where the Cholesky ordering is swapped. Variance decomposition analysis is also performed. Forecast error variance decomposition separates the variance of an endogenous variable into its component parts by determining how much of the variance of each endogenous variable can be explained by shocks to its own equation as well as shocks to the other endogenous variable's equation. In doing so, it is possible to see the extent of the variation in say, price, that is caused by P/E when these are the only two variables considered.

6.7.1 Price and P/E

In relation to price, the accumulated impulse response to a shock in price is positive. The accumulated IRF flattens outs at or about four to six months. There is a delayed response in price to shock in P/E with a positive cumulative impact. It increases gradually, peaking around four months. The largest cumulative impacts are for the Small Cap and Mid Cap indices. The impact on the FTSE 100 is relatively large while the impact is relatively small for the Top 40. The cumulative sum is initially negative for the S&P 500 but starts to increase at a decreasing rate at around three months.

In relation to P/E, a shock to price results in a positive, cumulative impact across all of the indices. In most cases, the initial increase is quite sharp, followed by a more gradual increase which flattens off at about four months for the ALSI and three months for the Small Cap and Mid Cap indices. For both the FTSE 100 and the Top 40, there is an initial increase, a pull-back and then another increase finally tending toward their long-run averages at or about five months. A shock to price with respect to the S&P 500 leads to an initial increase until about three months at which point it decreases, eventually flattening out around the ten-month mark. A shock with respect to P/E in the S&P 500 system results in a cumulative increase in P/E over time. A similar shape is found for the ALSI, but the trend breaks down between two and four months. The Top 40 follows the same shape as the ALSI, but there is a second trend break at about five months. The accumulated IRF flattens out at about three months for the Small Cap and at four months for the Mid Cap and FTSE 100.

Across all indices, virtually none of the variance in price is due to P/E. Between 56% and 55% of the variation in the ALSI's P/E is due to price. The figures are somewhat similar for the Top 40. Approximately 47% of the variation in Mid Cap's P/E can be explained by price. Between 33% and 37% and between 16% and 27% of the variance in P/E is due to price for the S&P 500 and FTSE 100 indices, respectively.

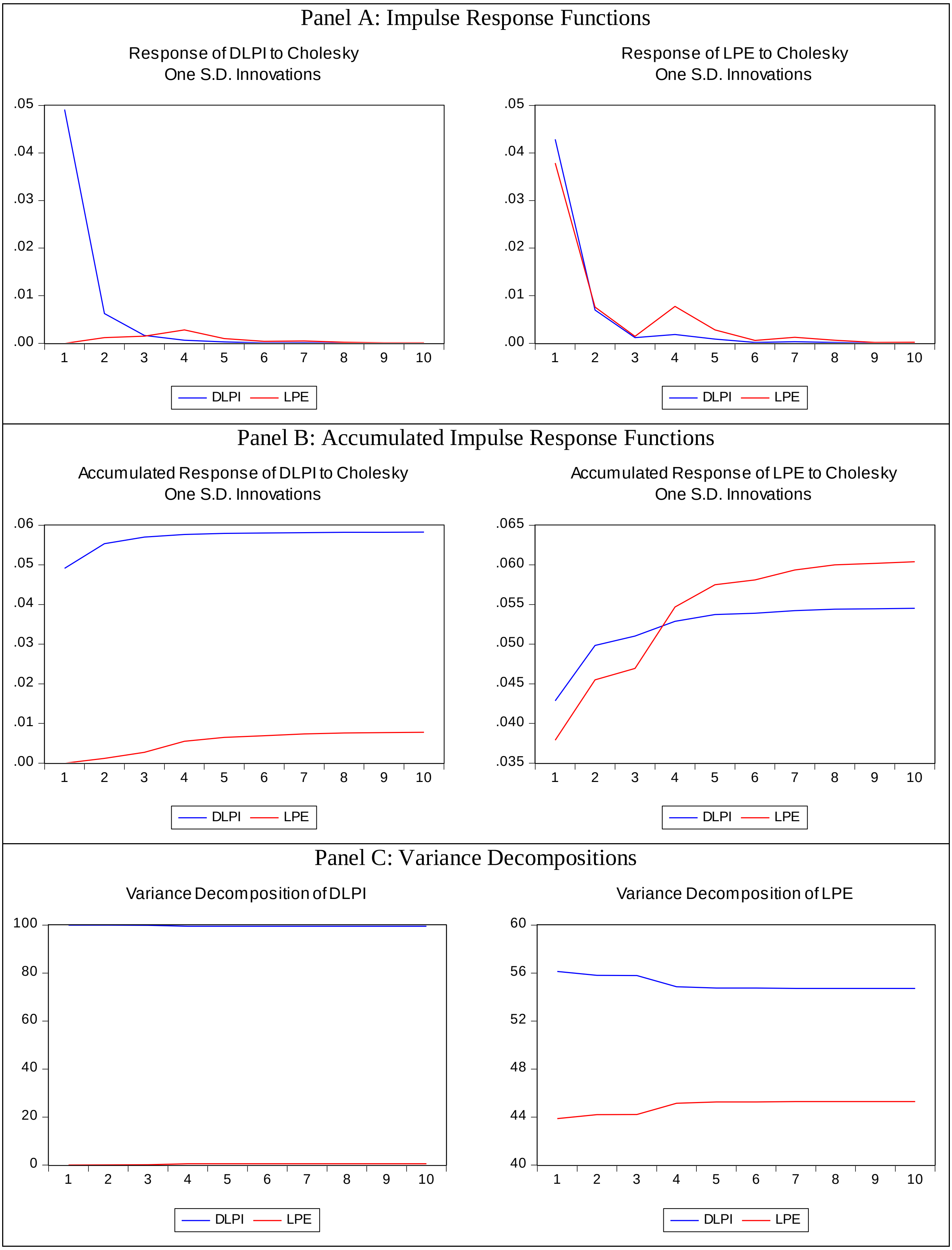

Figure 5: IRFs, accumulated IRFs and variance decompositions—price and P/E, ALSI

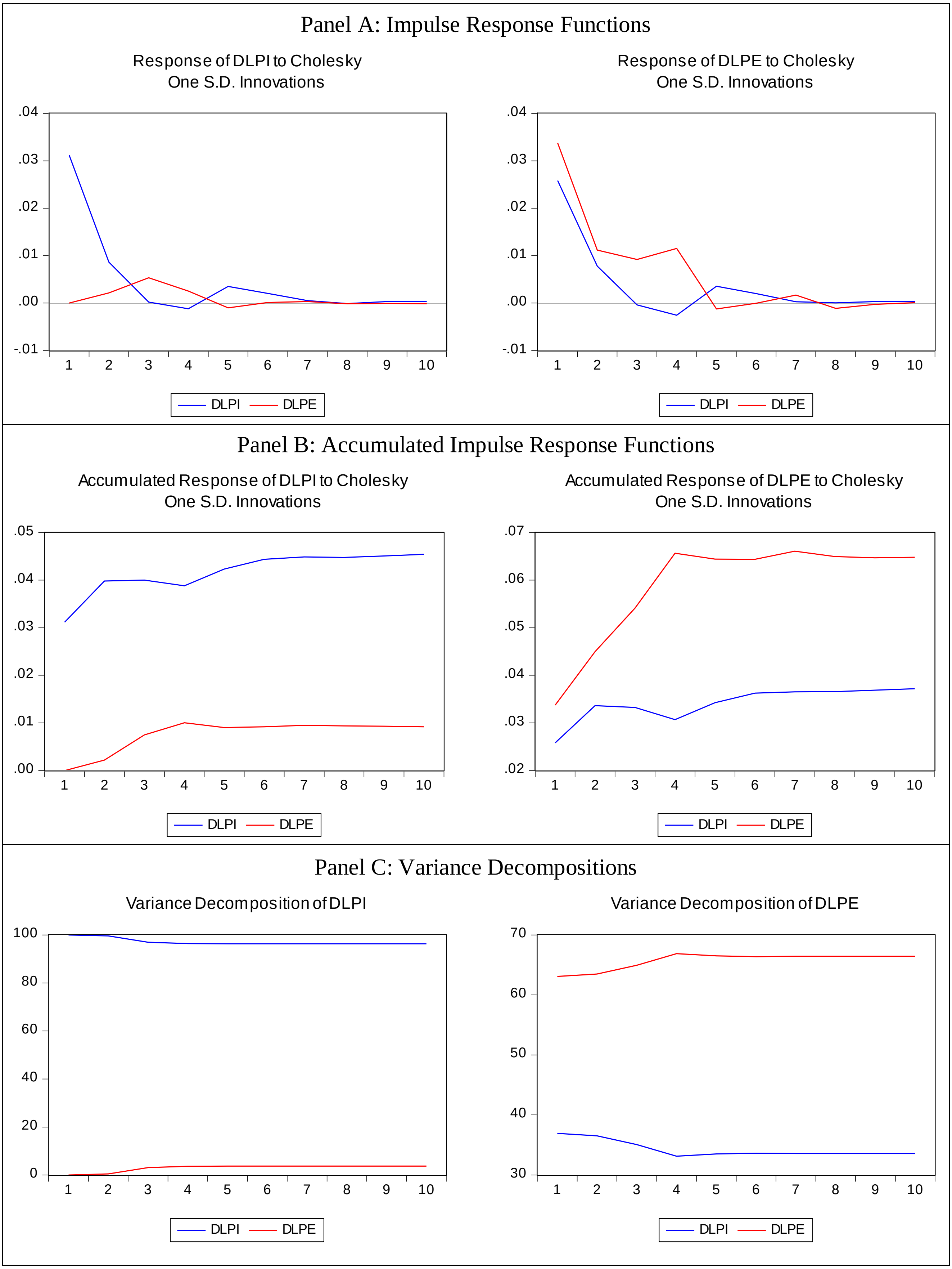

Figure 6: IRFs, accumulated IRFs and variance decompositions—price and P/E, FTSE 100

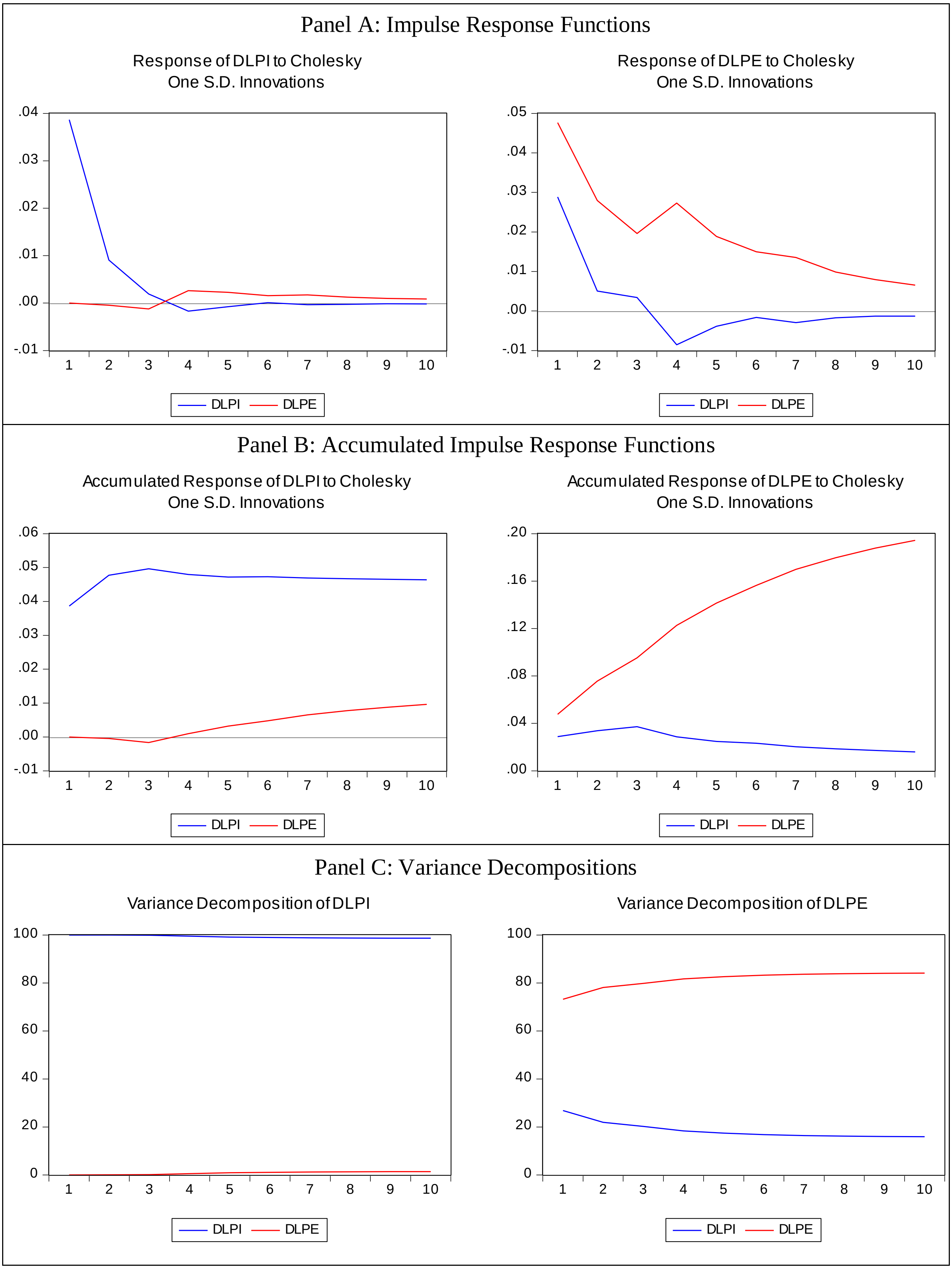

Figure 7: IRFs, accumulated IRFs and variance decompositions—price and P/E, S&P 500

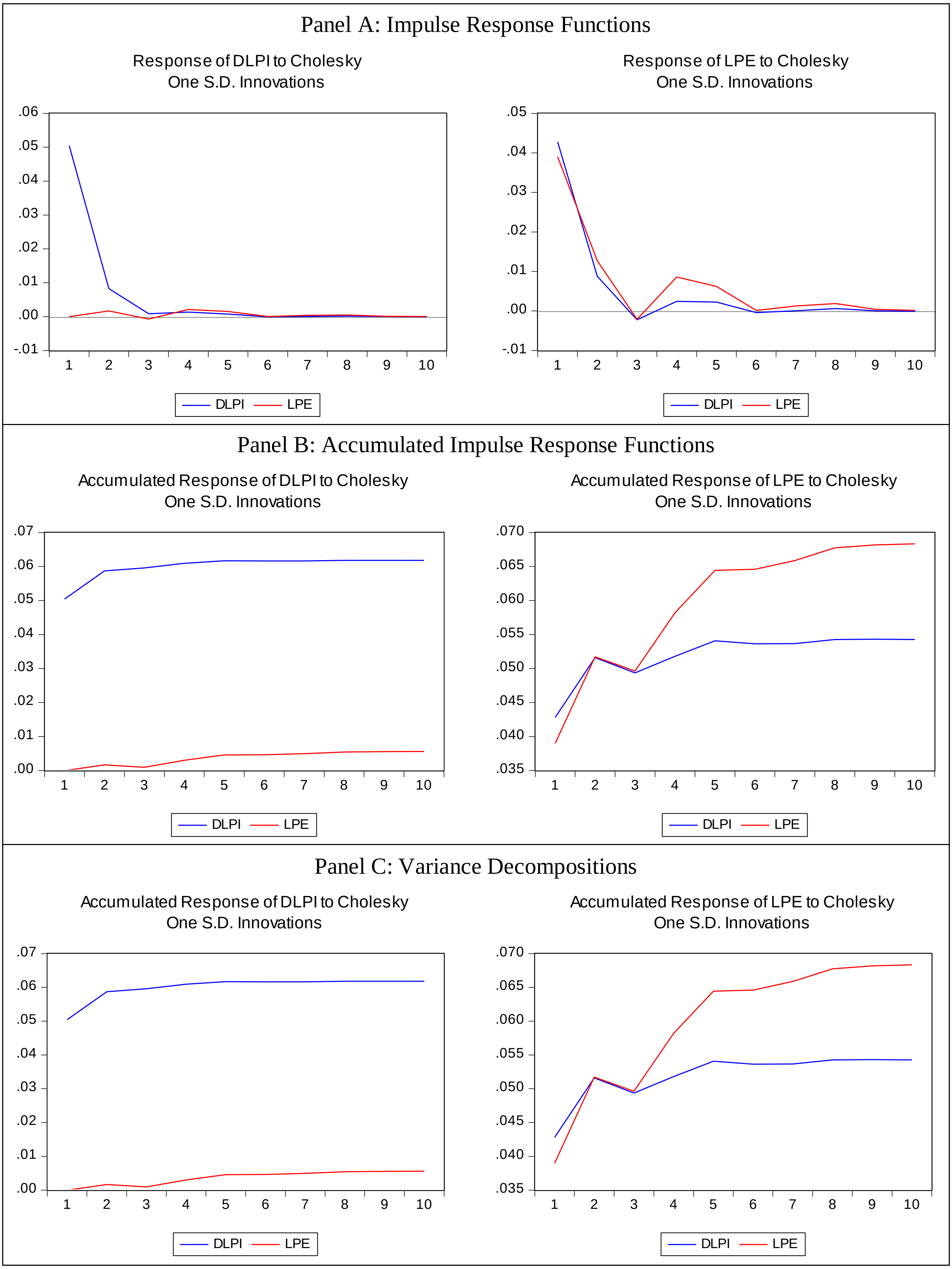

Figure 8: IRFs, accumulated IRFs and variance decompositions—price and P/E, Top 40

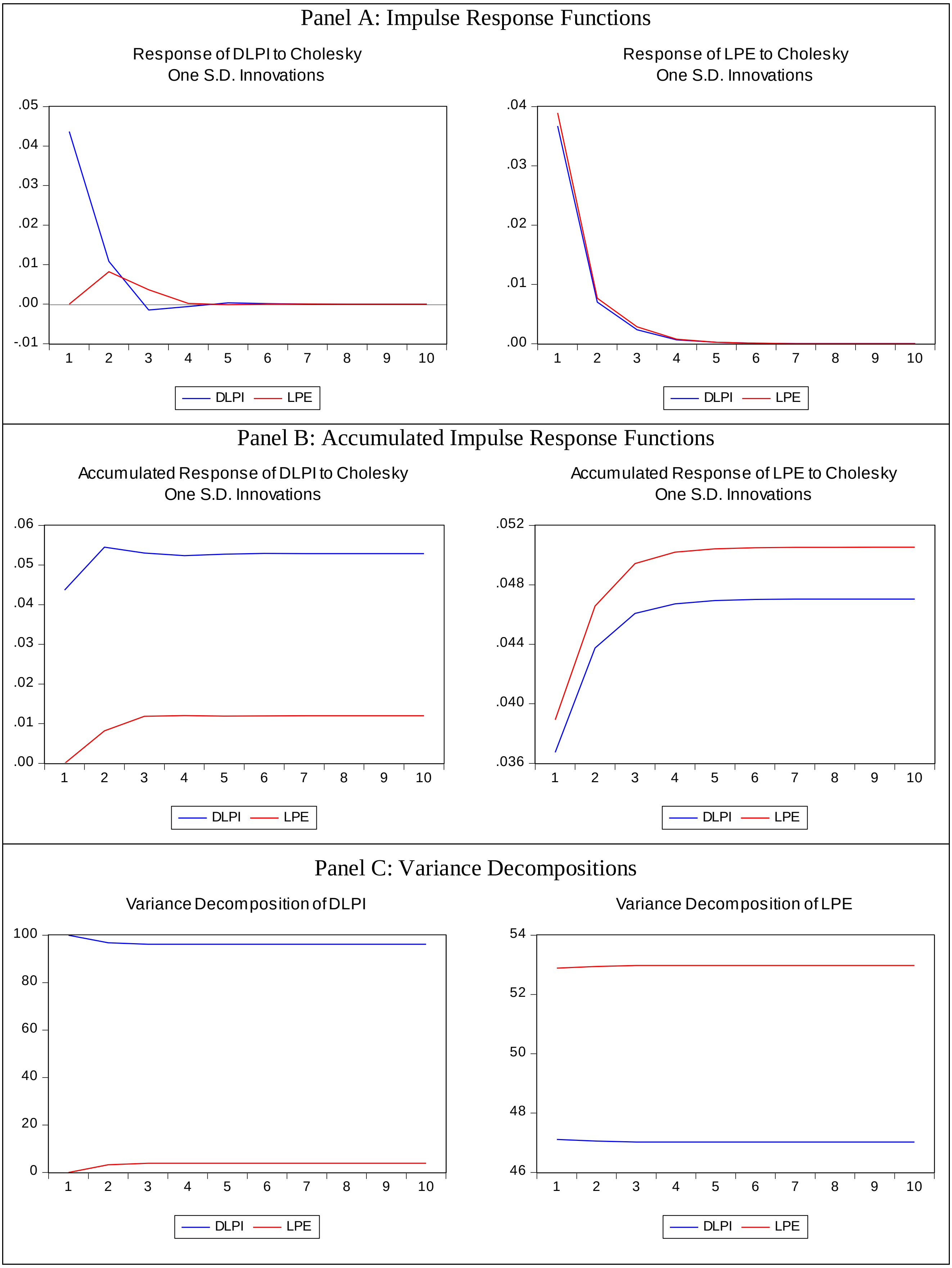

Figure 9: IRFs, accumulated IRFs and variance decompositions—price and P/E, Mid Cap

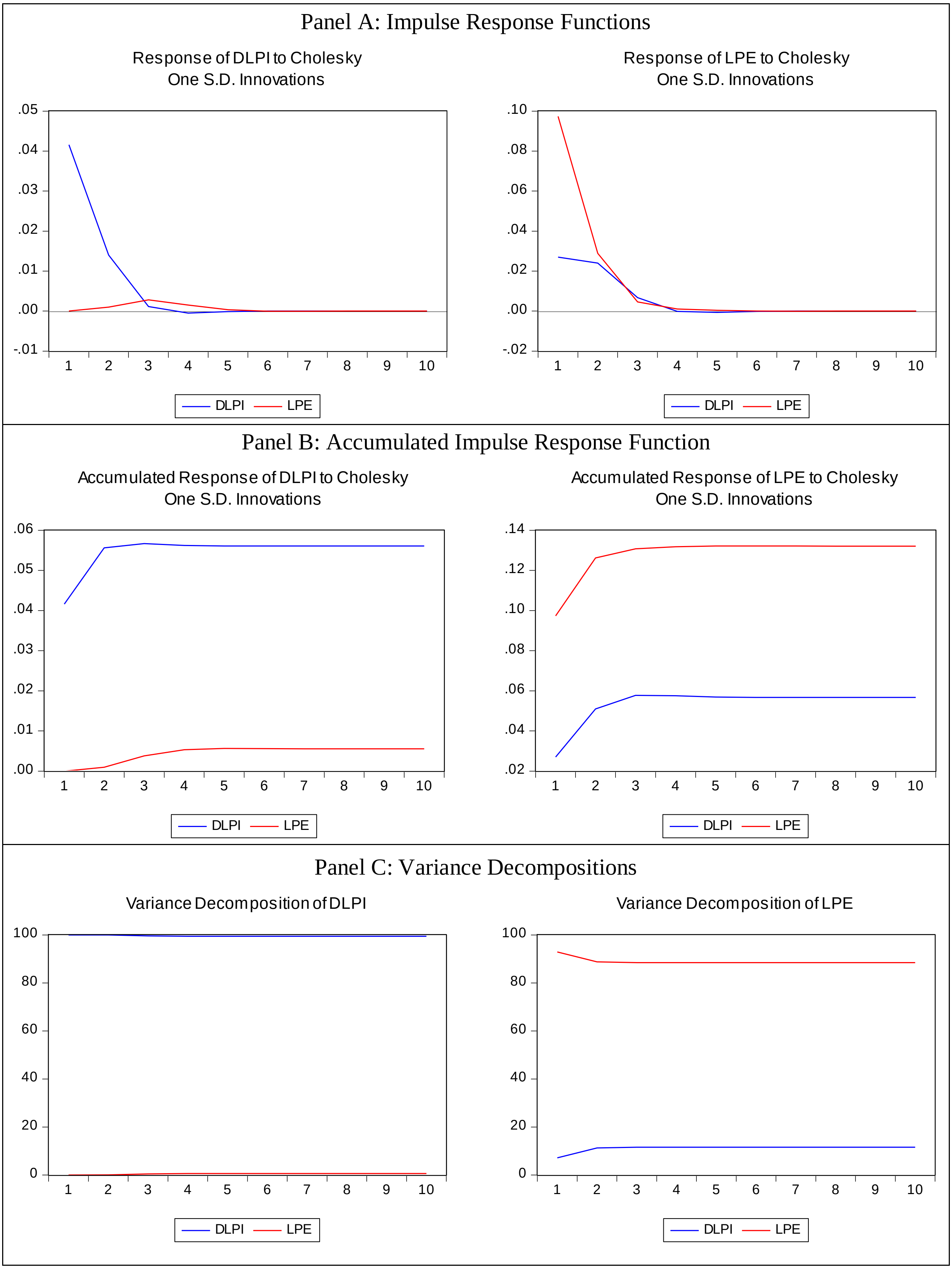

Figure 10: IRFs, accumulated IRFs and variance decompositions—price and P/E, Small Cap

6.7.2 Price and Dividend Yield

It is interesting to note that shocks traced by the IRFs in this section generally take more time to die out as compared to the previous section. When there is a shock to *price*, there is, generally, a positive cumulative impact on price. For the ALSI and Top 40, there is a gradual increase with a flattening-out at or about four months. The FTSE 100 and S&P 500 experience an initial increase, a pull-back, followed by an increase before flattening out at around six months. The accumulated IRF for the Mid Cap peaks at two months and flattens out at four months. The Small Cap experiences a peak at two months, a decline and then a sideways movement from six months. A shock applied to the *dividend yield* has a negative cumulative impact on price. The cumulative impact flattens out at around five months. The impact for the FTSE is relatively small. S&P 500 closing index value experiences an initial decline which turns and eventually becomes positive at five months. A shock in dividend yield has essentially no cumulative impact on price for the Mid Cap and a small, negative impact on price for the Small Cap.

A shock to *price* results in a negative cumulative impact in dividend yield across all indices. The impact for the FTSE 100 is almost constant over time. For the S&P 500, the cumulative impact is initially positive, declines somewhat before becoming flat from 5 months. For the Mid Cap, the cumulative impact flat-lines at 4 months. The accumulated IRF of the Small Cap has two humps. In most cases, a shock to *dividend yield* is a mirror image of a shock to price. The cumulative impact in relation to the ALSI is steep and positive. For the FTSE 100, the accumulated IRF is essentially flat. The function for the S&P 500 has a distinct hump and then becomes flat. Most of the cumulative impact for the Mid Cap index is felt in the first two to three months, after which the function settles at a lower level than at the first month. The accumulated IRF for the Small Cap is essentially flat.

As far as variance decomposition goes, the variance in price due to dividend yield is essentially zero in each of the indices. The variance in dividend yield due to price ranges from 57% to 63% for the ALSI, 53% to 57% for the FTSE 100 and 53% to 55% for the S&P 500. Unsurprisingly, the figures for the Top 40 are similar to those of the ALSI. In comparison to the Top 40 index, the variation in dividend yield due to price for each of the Mid Cap and Small Cap indices is much lower; 26% to 27% and 30 to 34%, respectively.

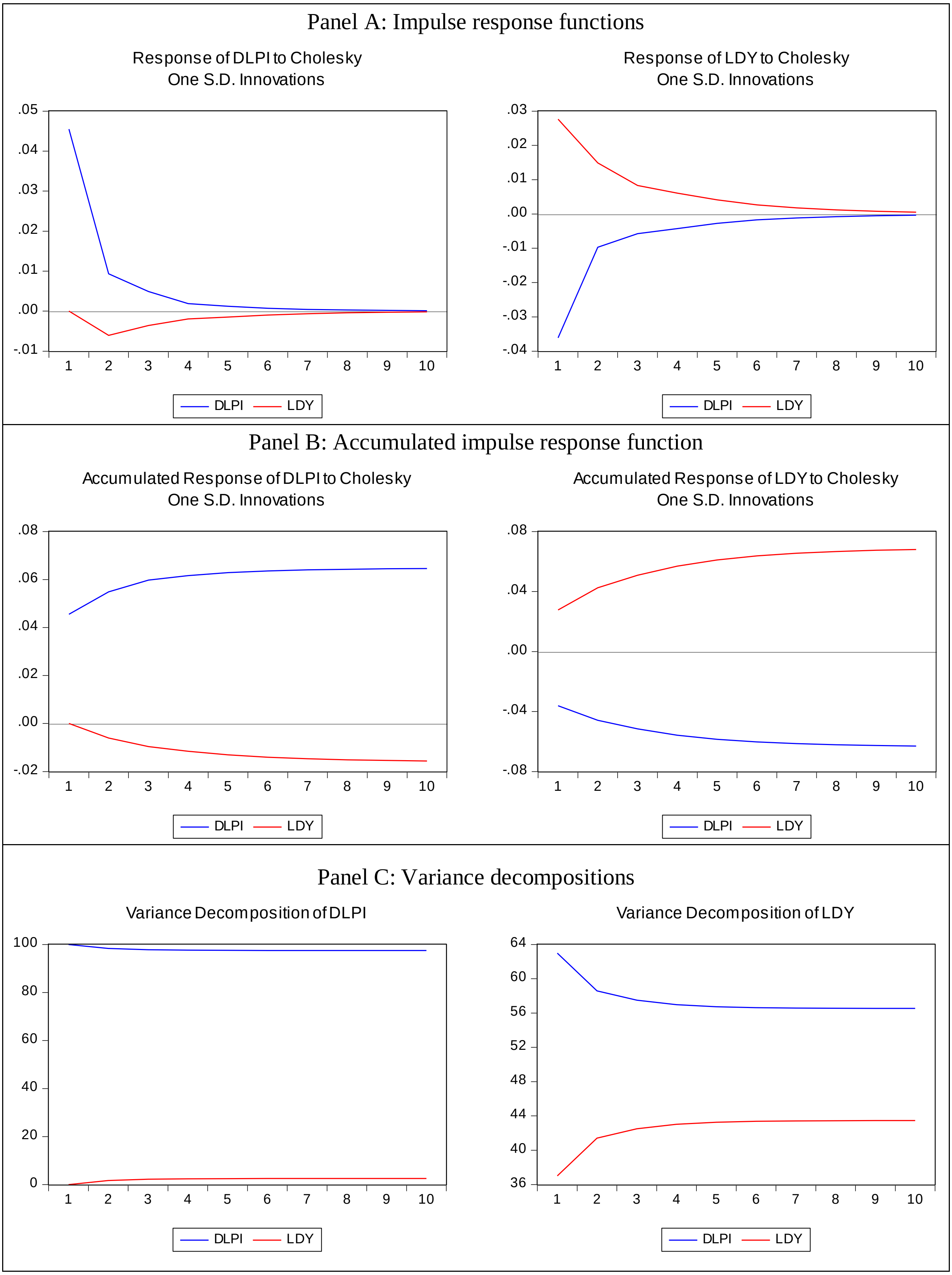

Figure 11: IRFs, accumulated IRFs and variance decompositions—price and DY, ALSI

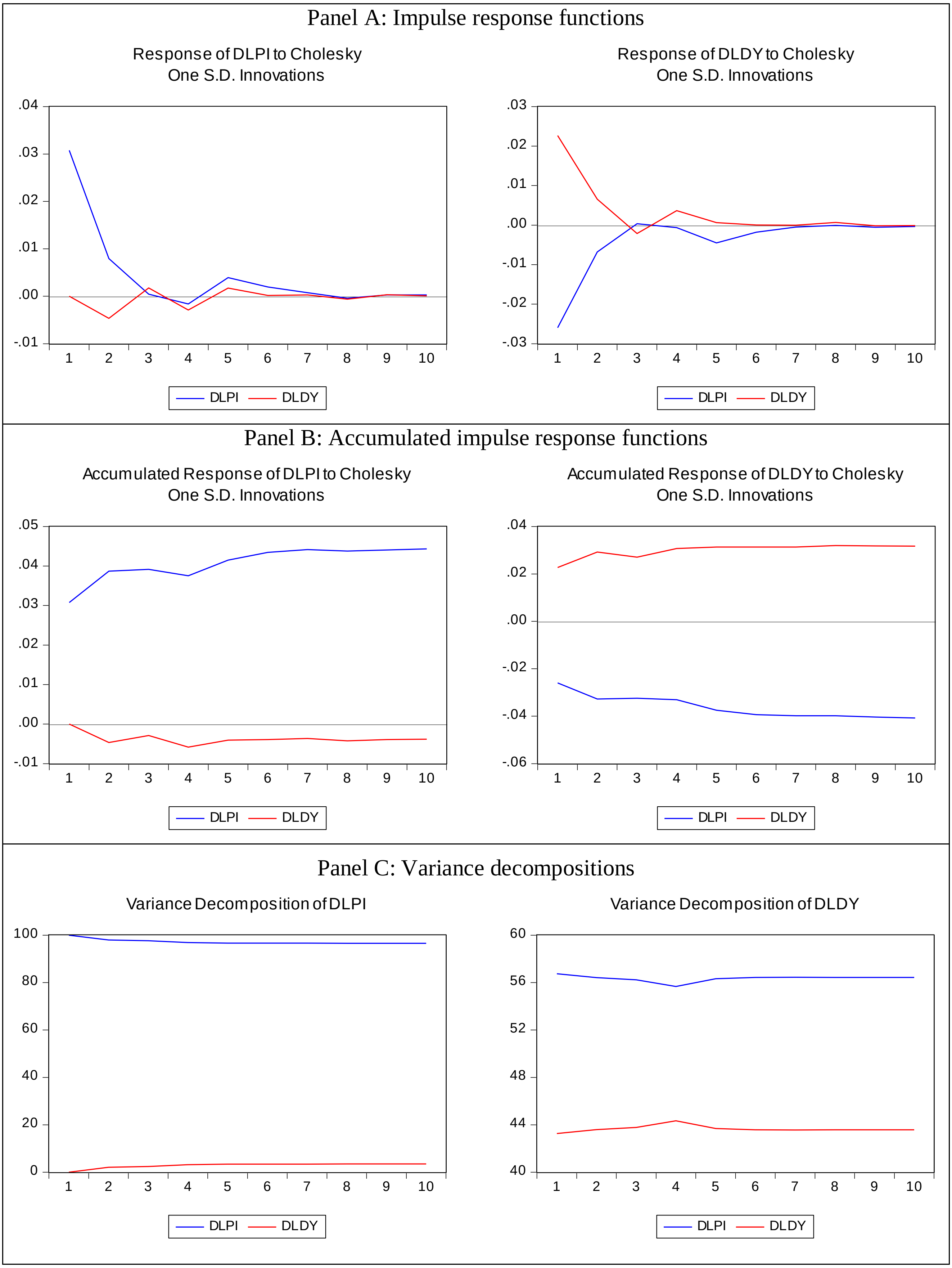

Figure 12: IRFs, accumulated IRFs and variance decompositions—price and DY, FTSE 100

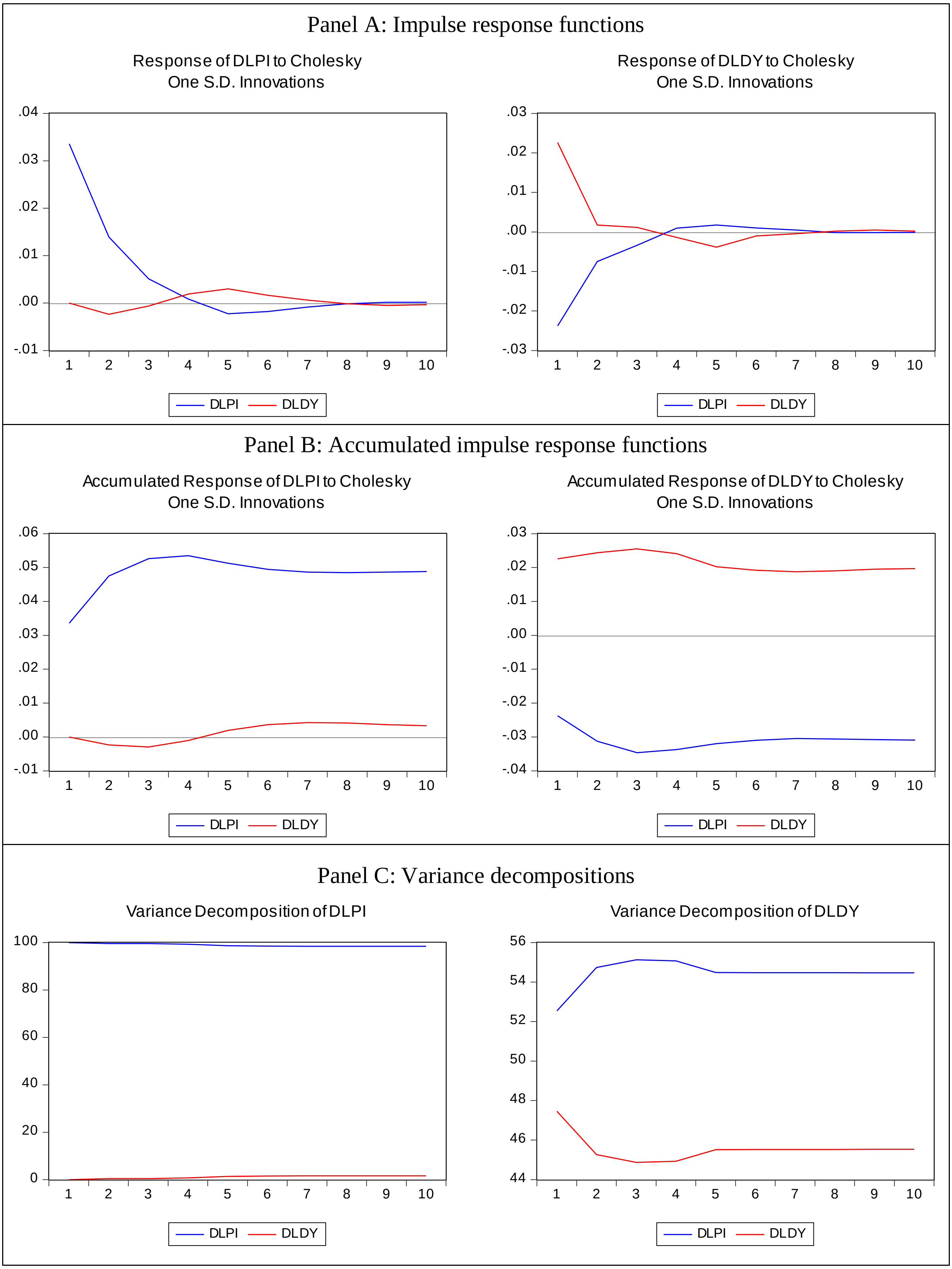

Figure 13: IRFs, accumulated IRFs and variance decompositions—price and DY, S&P 500

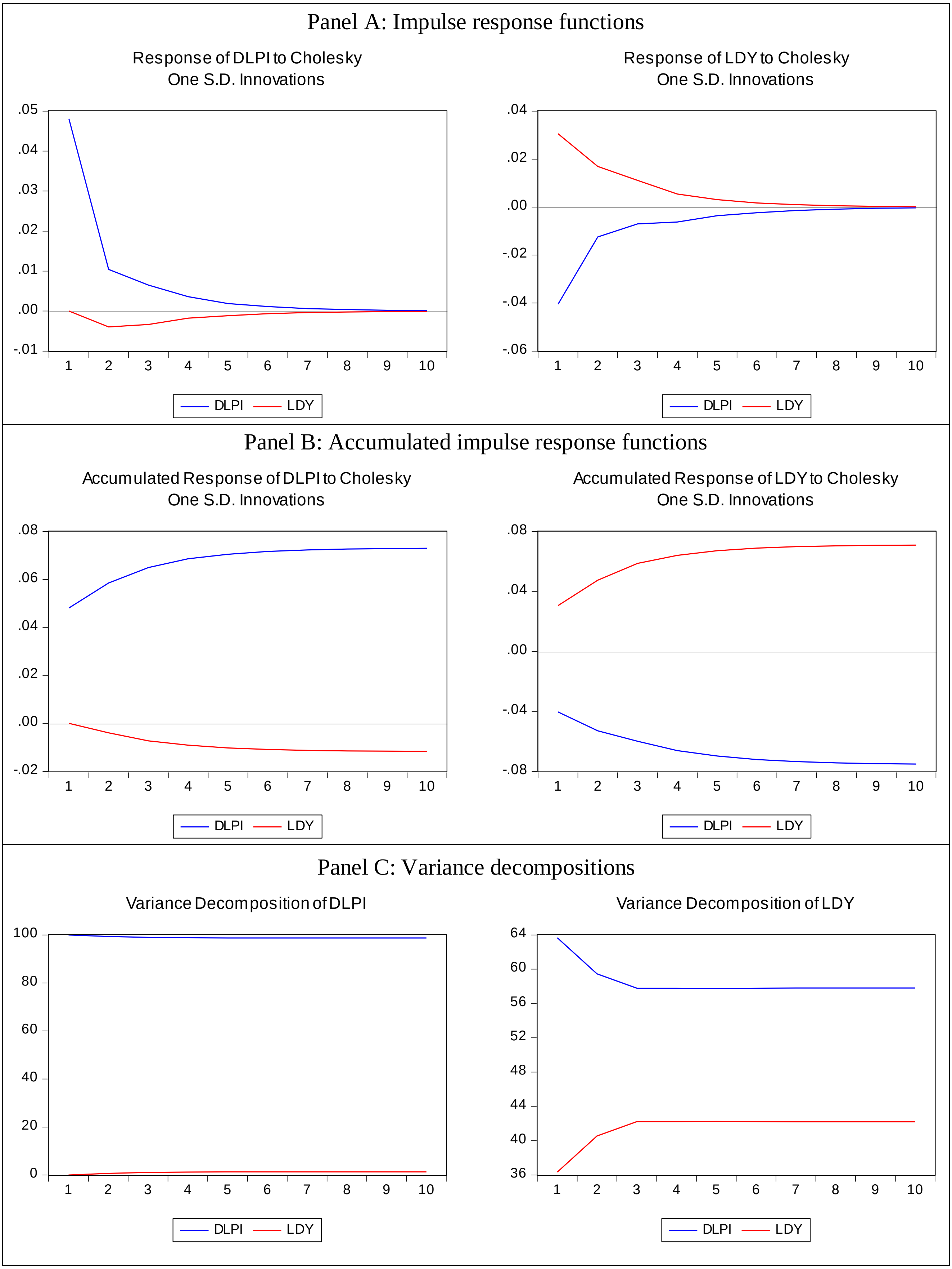

Figure 14: IRFs, accumulated IRFs and variance decompositions—price and DY, Top 40

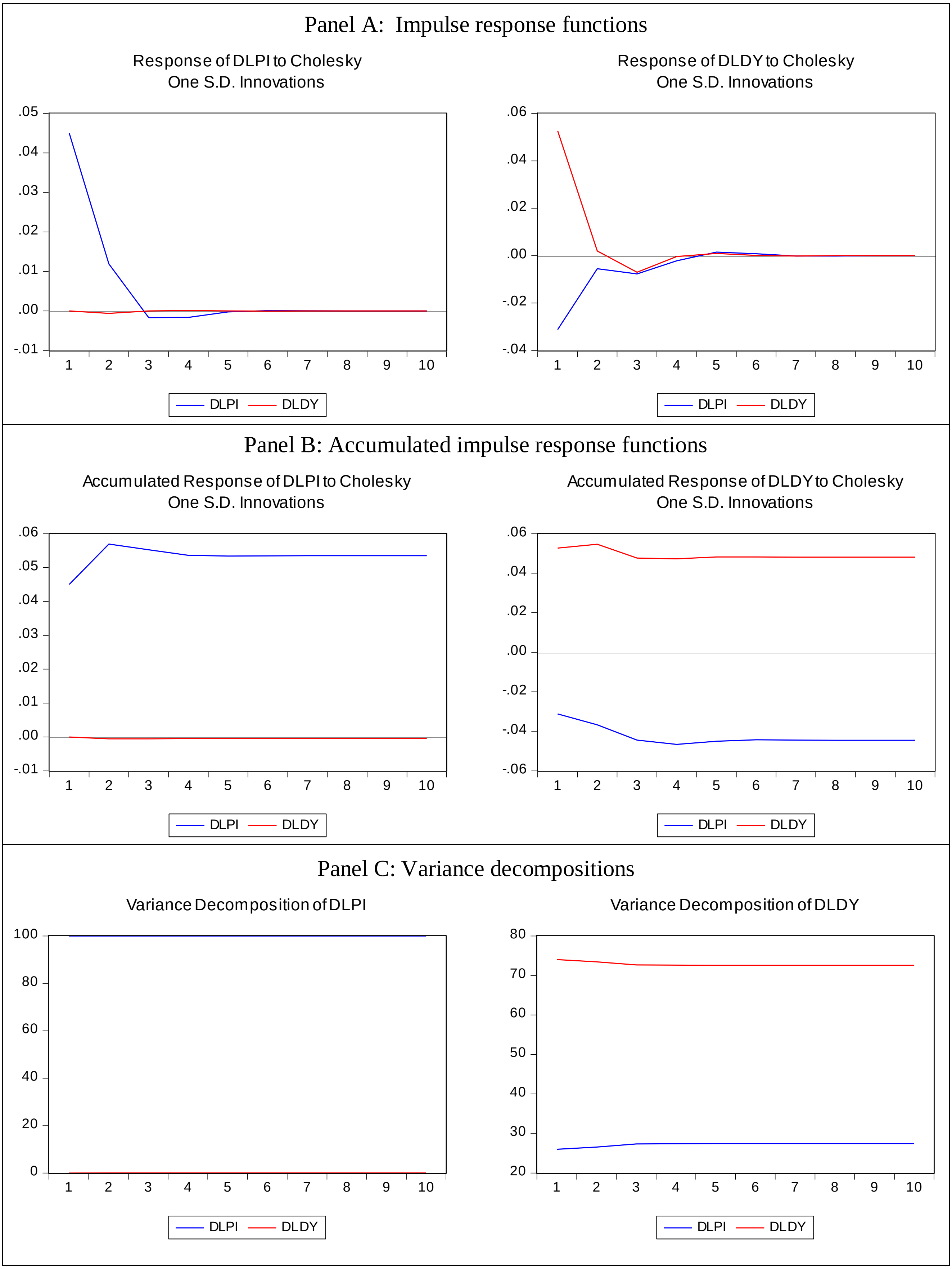

Figure 15: IRFs, accumulated IRFs and variance decompositions—price and DY, Mid Cap

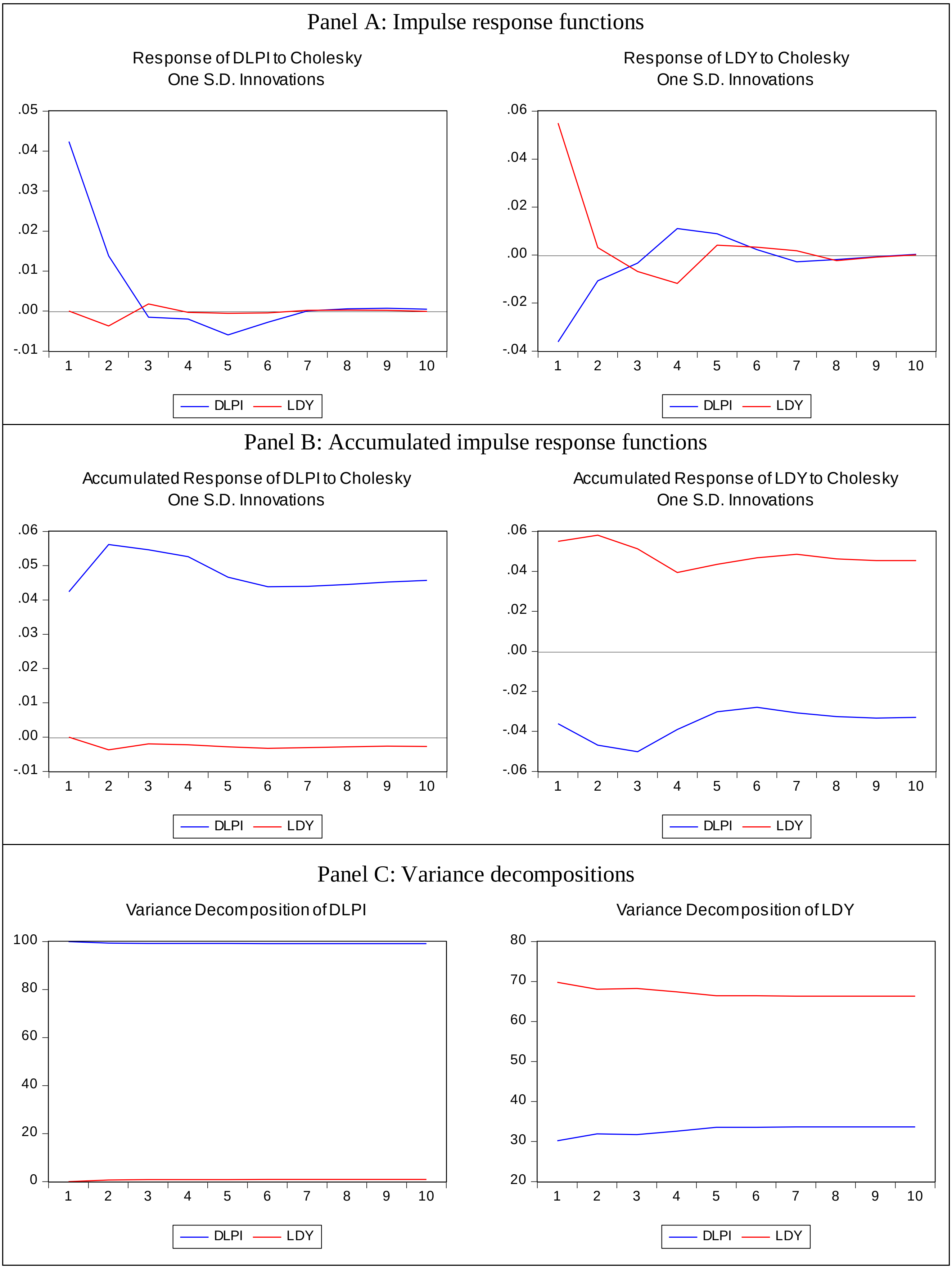

Figure 16: IRFs, accumulated IRFs and variance decompositions—price and DY, Small Cap

6.8 TESTING FOR GRANGER CAUSALITY

The Toda and Yamamoto (1995) surplus-lag approach is used in testing for Granger causality. Conditioning the choice of model for the Granger causality test on a test for cointegration is an example of pre-test testing; the causality test essentially is actually a random combination of the two tests. The causality test statistic is a random, weighted sum of the test statistic had VAR been chosen and the test statistic had VECM been chosen. The size of the weights are dependent upon the properties of the initial test and, as such, the power of the ultimate test is compromised (Giles, 2011a). Clarke and Mirza (2006) find that over-fitting—as in the Toda-Yamamoto methodology—results in better control over type I error probability.

Giles (2011b) outlines the procedure succinctly. First, the maximum order of integration m for the pair of time series must be determined. If one of the time series is integrated of order zero—it is stationary—and the other time series is integrated of order one—it must be differenced to be $I(0)$—then $m = 1$. Next, a VAR must be run "in levels"—that is, the time series must not be differenced—with lag order p. Should the residuals not be independently distributed, p should be increased until autocorrelation is no longer a problem. Then, m lags of each variable are added to each equation in the system. Finally, a standard Wald test is used to test the null hypothesis that the first p lagged values—not $p + m$ lags—of the second equation's dependent variable are zero in the first equation and, similarly, whether the first p lagged values of the first equation's dependent variables are zero in the second equation. Using equation 7a as an example, the null hypothesis would be given by: $b_{1,1} = \cdots = b_{1,p} = 0$. Similarly, for equation 7b, the null hypothesis would be given by: $a_{2,1} = \cdots = a_{2,p} = 0$. The Wald test statistics are asymptotically χ^2–distributed with p degrees of freedom.

Where two time series are cointegrated, Granger causality must be present in one or both directions. Rejection of the null hypothesis would indicate the presence of Granger causality. In contrast, if Granger causality is present, the time series are not necessarily cointegrated. The probability values of the Wald tests are presented in table 17 for the systems of equations with price and P/E and table 18 for the systems with price and dividend yield.

6.8.1 Price and P/E

There is no evidence of Granger-causality in either direction between price and P/E for the ALSI and for the Top 40. In contrast, Granger-causality runs both ways for the FTSE 100. The test results suggest price Granger-causes P/E in relation to the S&P 500 and the Small Cap index, whilst P/E Granger-causes price in relation to the Mid Cap index.

Table 17: Granger causality test *p*-values—price and P/E

Indices	Null hypotheses	
	LPE does not Granger-cause LPI	LPI does not Granger-cause LPE
FTSE/JSE All Share	0,9882	0,8579
FTSE 100	0,0864	0,0003
S&P 500	0,3310	0,0000
FTSE/JSE Top 40	0,8188	0,5917
FTSE/JSE Mid Cap	0,0026	0,9654
FTSE/JSE Small Cap	0,7441	0,0188

6.8.2 Price and Dividend Yield

Table 18: Granger causality test *p*-values—price and DY

Indices	Null hypotheses	
	LDY does not Granger-cause LPI	LPI does not Granger-cause LDY
FTSE/JSE All Share	0,2651	0,4002
FTSE 100	0,0140	0,3356
S&P 500	0,3721	0,2215
FTSE/JSE Top 40	0,7442	0,3364
FTSE/JSE Mid Cap	0,8390	0,1035
FTSE/JSE Small Cap	0,2259	0,1241

The results contained in table 18 show no evidence of Granger-causality in either direction in relation to the ALSI, S&P 500, Top 40, Mid Cap and Small Cap. With respect to the Mid Cap, the failure to reject the null hypothesis of no Granger-causality running from price to dividend yield is borderline in that the probability value just exceeds 10%. It does seem as though dividend yield Granger-causes price in relation to the FTSE 100.

7 CONCLUSION

The aim of this research report is to determine whether two popular measures of relative value—P/E and dividend yield—impact subsequent price and to what extent. VAR analysis enables an exploration of the dynamics relations between the variables within a non-structural framework. In addition, the VARs are used to determine whether the relations are causal—in the Grangerian sense. A number of equity indices are used in the analysis; three broad-market indices and three JSE indices which are subsets of the local broad-market index.

From the price–P/E VAR estimates, an increase in price is followed by an initial increase in price and, subsequently, a decrease in price. The pattern for the ALSI and the ALSI sub-indices is similar. The subsequent decline in price is consistent with mean reversion of price. The parameter estimates show that P/E does not have a significant impact on price, except in relation to the FTSE 100 and S&P 500. For these indices, an increase in P/E leads to an expected increase in price in two months' time. This is consistent with Bhargava and Malhotra (2006)'s findings on two fronts: (1) price increases in response to an increase in P/E and (2) the impact is not very pronounced. When decomposing the variation in price into variance due to P/E and variance due to price, very little of the variance in price is due to P/E. A substantial portion of the variation in P/E is due to price, which is consistent with the findings of Campbell and Shiller (1988). The fraction of total variance in P/E due to price declines from the Top 40 index to the Mid Cap index and from the Mid Cap index to the Small Cap index. In most cases, there is no evidence of Granger causality—in either direction—between price and P/E. The only instance where Granger causality is present is from P/E to price for the Mid Cap Index.

With respect to the price–dividend yield VARs, the price equations paint a mixed picture. For the S&P 500, the Mid Cap index and the Small Cap index, the first lag of price has a statistically significant, positive coefficient. The ALSI and Top 40 have no significant price variables. For the FTSE 100, the third lag of price has a negative coefficient which is significant at the 5% level and the fourth lag of price has a positive coefficient and is statistically different from zero at the 1% level. With respect to the broad-market indices, dividend yield has a statistically significant and negative impact on price for the ALSI and FTSE 100. None of the lagged dividend yield–variables in the S&P 500 system are statistically significant. None of the DY-variables in the price equations of the ALSI sub-indices have statistically significant coefficients. Variance decomposition of price shows that the fraction of the variation in price due to dividend yield is virtually zero. The variation of dividend yield due to price is approximately 50% to 60% for most of the indices. For the Mid Cap and

Small Cap indices, the figure is closer to 30%. It is worth noting that price plays a greater role in the variance of dividend yield than in the variance of P/E. In most instances, there is no evidence of Granger causality in either direction. Dividend yield does appear to Granger-cause price in relation to the Mid Cap index, though.

It is interesting to note that there does not appear to be a meaningful difference between the Small Cap index—representing the smallest JSE-listed firms and the Top 40 index—representing the largest JSE-listed firms—in terms of the nature of the relationship between the valuation ratios and price as well as with the how they evolve over time. It is also interesting that qualities usually associated with the stocks of growth companies—higher P/E ratio and lower dividend yield—seem to reflect higher prices in future.

Overall, the impact of the two, popular valuation ratios on future prices is not as great as expected, especially in a South African context. This suggests that investment practitioners can afford to focus less on these ratios as measures of relative value and that ratio analysis should be, at best, one step in a more rigorous investment analysis process. This research report adds value in its use of a relatively long, recent time period using local data. It is also adds value in that it allows for a comparison between South Africa's equity capital market with that of two developed countries' markets. Future research may wish to extend this comparison to include other developing markets; for example, a comparison of the markets of BRICS, CIVETS or "fragile five" nations. It may be instructive to model central bank "intervention" in the capital markets by means of intervention dummy variables; for example, it may be that the Fed's current bond-buying programme has driven down the yields on most investable assets and, in turn, driven up the valuations of companies' common shares. The fall in stock prices around the world seen in June 2013 following the Fed's announcement of "tapering" may support this view. It may also be useful to carry out a liquidity adjustment because some traded indices are far more liquid than others. Small-sample corrections of the test statistics—by means of bootstrapping of their distribution—may improve the quality of results obtained.

8 REFERENCES

Anderson, K., & Brooks, C. (2006). The long-term price-earnings ratio. *Journal of Business Finance & Accounting, 33*(7), 1063-1086.

Ang, A., & Bakaert, G. (2007). Stock return predictability: Is it there? *The Review of Financial Studies, 20*(3), 651-707.

Asness, C. (2003). Fight the Fed model. *The Journal of Portfolio Management, 30*(1), 11-24.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica, 66*(1), 47-78.

Bailey, G., & Gilbert, E. (2007). The impact of liquidity on mean reversion of share returns of the JSE. *Investment Analysts Journal, 66*, 19-30.

Ball, R. (1978). Anomalies in relationships between securities' yields and yield-surrogates. *Journal of Financial Economics, 6*(2), 103-126.

Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research, 6*(2), 159-178.

Banz, R. W., & Breen, W. J. (1986). Sample-dependent results using accounting and market data: Some evidence. *The Journal of Finance, 41*(4), 779-793.

Barr, G., & Kantor, B. (1999). Price-earnings ratios on the Johannesburg Stock Exchange—Are they a guide to value? *South African Journal of Accounting Research, 13*(1), 1-23.

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *The Journal of Finance, 32*(3), 663-682.

Basu, S. (1983). The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence. *Journal of Financial Economics, 12*(1), 129-156.

Becker, R., Lee, J., & Gup, B. E. (2012). An empirical analysis of mean reversion of the S&P 500's P/E ratios. *Journal of Economics and Finance, 36*(3), 675-690.

Bernard, V. L., & Thomas, J. K. (1990). Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics, 13*(4), 305-340.

Bhargava, V., & Malhotra, D. K. (2006). Do price-earnings ratios drive stock values? *The Journal of Porfolio Management, 33*(1), 86-92.

Bierman, H. (2002). The pice-earnings ratio. *The Journal of Portfolio Management, 28*(4), 57–60.

Black, F., & Scholes, M. (1974). The effects of dividend yield and dividend policy on common stock prices and returns. *Journal of Financial Economics, 1*(1), 1-22.

Blume, M. E., & Stambaugh, R. E. (1983). Biases in computed returns: An application to the size effect. *Journal of Finanical Economics, 12*(3), 387-404.

Brown, P., Kleidon, A. W., & Marsh, T. A. (1983). New evidence on the nature of size-related anomalies in stock prices. *Journal of Financial Economics, 12*(1), 33-56.

Campbell, J. Y., & Shiller, R. J. (1987). Cointegration and tests of present value models. *Journal of Political Economy, 95*(5), 1062-1088.

Campbell, J. Y., & Shiller, R. J. (1998). Valuation ratios and the long-run stock market outlook. *The Journal of Portfolio Management, 24*(2), 11-26.

Campbell, J. Y., & Shiller, R. J. (2005). Valuation ratios and the long-run stock market outlook: An update. In R. H. Thaler (Ed.), *Advances in Behavioral Finance* (Vol. II, pp. 173-201). Princeton: Princeton University Press.

Capelle-Blanchard, G., & Raymond, H. (2004). Empirical evidence on periodically collapsing stock price bubbles. *Applied Economics Letters, 11*(1), 61-69.

Carlson, J. B., Pelz, E. A., & Wohar, M. E. (2002). Will valuation ratios revert to historical means? *The Journal of Porfolio Management, 28*(4), 23-35.

Chan, K. C. (1988). On the contrarian investment strategy. *The Journal of Business, 61*(2), 147-163.

Clarke, J. A., & Mirza, S. (2006). A comparison of some common methods for detecting Granger noncausality. *Journal of Statistical Computation and Simulation, 76*(3), 207-231.

Cook, T. J., & Rozeff, M. S. (1984). Size and earnings/price ratio anomalies: One effect or two? *The Journal of Financial and Quantitative Analysis, 19*(4), 449-466.

Cubbin, E., Eidne, M., Firer, C., & Gilbert, E. (2006). Mean reversion on the JSE. *Investment Analysts Journal, 63*, 39-47.

Danielson, M. G., & Dowdell, T. D. (2001). The return-stages valuation model and the expectations within a firm's P/B and P/E ratios. *Financial Management, 30*(2), 93-124.

De Bondt, W. F., & Thaler, R. H. (1985). Does the stock market overreact? *The Journal of Finance, 40*(3), 793-805.

Dreman, D. N., & Berry, M. A. (1995). Overreaction, underreaction and the low-P/E effect. *Financial Analysts Journal, 51*(4), 21-30.

Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: Representation, estimation and testing. *Econometrica, 55*(2), 251-276.

Fama, E. F. (1998). Market efficiency, long-term returns and behavioral finance. *Journal of Financial Economics, 49*(3), 283-306.

Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics, 22*(1), 3-25.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics, 25*(1), 23-49.

Fisher, K. L., & Statman, M. (2000). Cognitive biases in market forecasts. *The Journal of Portfolio Management, 27*(1), 72-81.

Gilbert, E., & Strugnell, D. (2010). Does survivorship bias really matter? An empirical investigation into its effects on the mean reversion of share returns on the JSE (1984-2007). *Investment Analysts Journal, 72*, 31-42.

Giles, D. (2011, April 29). *Testing for Granger Causality*. Retrieved December 11, 2013, from Econometrics Beat: Dave Giles' Blog: http://davegiles.blogspot.com/2011/04/testing-for-granger-causality.html

Giles, D. (2011, October 25). *VAR or VECM When Testing for Granger Causality?* Retrieved December 11, 2013, from Econometrics Beat: http://davegiles.blogspot.ca/2011/10/var-or-vecm-when-testing-for-granger.html

Goetzmann, W., & Jorion, P. (1995). A longer look at dividend yields. *The Journal of Business, 68*(4), 483-508.

Gordon, M. J. (1959). Dividends, earnings, and stock prices. *The Review of Economics and Statistics, 41*(2), 99-105.

Graham, B., & Dodd, D. (1934). *Security Analysis* (1st ed.). New York City: McGraw-Hill.

Gupta, R., & Modise, M. (2012). Valuation ratios and stock return predictability in South Africa: Is it there? *Emerging Markets Finance and Trade, 48*(1), 70-82.

Hsieh, H.-H., & Hodnett, K. (2011). Tests of the overreaction hypothesis and the timing of mean reversals on the JSE Securities Exchange (JSE): The case of South Africa. *Journal of Applied Finance and Banking, 1*(1), 107-130.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control, 12*(2), 231-254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica, 59*(6), 1551-1580.

Johansen, S., Mosconi, R., & Nielsen, B. (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend. *Econometrics Journal, 3*(2), 216-249.

Kane, A., Marcus, A. J., & Noh, J. (1996). The P/E multiple and market volatility. *Financial Analysts Journal, 52*(4), 16-24.

Keown, A. J., Pinkerton, J. M., & Chen, S. N. (1987). Portfolio selection based upon P/E ratios: Diversification, risk decomposition and implications. *Journal of Business Finance & Accounting, 14*(2), 187-198.

Koivu, M., Pennanen, T., & Ziemba, W. T. (2005). Cointegration analysis and the Fed model. *Finance Research Letters, 2*(4), 248-259.

Koutmos, D. (2010). The P/E multiple and market volatility revisited. *International Research Journal of Finance and Economics, 43*, 7-16.

Litzenberger, R. H., & Ramaswamy, K. (1979). The effect of personal taxes and dividends on capital asset pricing: Theory and empirical evidence. *Journal of Financial Economics, 7*(2), 163-195.

Lütkepohl, H. (2011). *Vector autoregressive models.* Working Paper EUI ECO 2011/30, European University Institute.

MacDonald, R., & Power, D. (1995). Stock prices, dividends and retention: Long-run relationships and short-run dynamics. *Journal of Empirical Finance, 2*(2), 135-151.

Malkiel, B. G. (2004). Models of stock market predictability. *The Journal of Financial Research, 27*(4), 449-459.

Modigliani, F., & Cohn, R. A. (1979). Inflation, rational valuation and the market. *Financial Analysts Journal, 35*(2), 24-44.

Molodovsky, N. (1953). A theory of price-earnings ratios. *The Analysts Journal, 9*(5), 65-80.

Muller, C. (1999). Investor overreaction on the Johannesburg Stock Exchange. *Investment Analysts Journal, 49*, 5-17.

Nicholson, S. F. (1968). Price ratios in relation to investment results. *Financial Analysts Journal, 24*(1), 105-109.

Ofer, A. R. (1975). Investors' expectations of earnings growth, their accuracy and effects on the structure of realized rates of return. *The Journal of Finance, 30*(2), 509-523.

Page, M. J., & Way, C. V. (1992). Stock market over-reaction: The South African evidence. *Investment Analysts Journal, 36*, 35-49.

Perron, P. (1989). The great crash, the oil price shock and the unit root hypothesis. *Econometrica, 57*(6), 1361-1401.

Perron, P., & Vogelsang, T. J. (1993). Erratum: The great crash, the oil price shock and the unit root hypothesis. *Econometrica, 61*(1), 248-249.

Ramcharran, H. (2004). Returns and pricing in emerging markets. *The Journal of Investing, 30*(3), 45–56.

Reinganum, M. R. (1981). Misspecification of capital asset pricing: Empirical anomalies based on earnings' yields and market values. *Journal of Financial Economics, 9*(1), 19-46.

Rendleman, R. J., Jones, C. P., & Latané, H. A. (1987). Further insight into the standardized unexpected earnings anomaly: Size and serial correlation effects. *The Financial Review, 22*(1), 131-144.

Rozeff, M. S. (1984). Dividend yields are equity risk premiums. *The Journal of Portfolio Management, 11*(1), 68-75.

Shen, P. (2000). The P/E ratio and stock market performance. *Economic Review, 85*(4), 23-36.

Shiller, R. J. (1984). Stock prices and social dynamics. *Brookings Papers on Economic Activity*(2), 457-510.

Shiller, R. J. (2002). The irrationality of markets. *The Journal of Psychology and Financial Markets, 3*(2), 87-93.

Smidt, S. (1968). A new look at the random walk hypothesis. *The Journal of Financial and Quantitative Analysis, 3*(4), 235-261.

Tatom, J. A. (2002). Stock prices, inflation and monetary policy. *Business Economics, 37*(4), 7-19.

Toda, H. Y., & Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics, 66*(1), 225-250.

Trevino, R., & Robertson, F. (2002). P/E ratios and stock market returns. *Journal of Financial Planning, 15*(2), 76-84.

Weigand, R. A., & Irons, R. (2008). Compression and expansion of the market P/E ratio: The Fed model explained. *The Journal of Investing, 17*(1), 55-64.