Price discovery on the Johannesburg Stock Exchange: Examining the impact of the SATRIX Top 40 Exchange Traded Fund

Kerry McCullough

School of Accounting, Economics and Finance, New Arts Building,
Pietermaritzburg Campus, University of KwaZulu-Natal, South Africa.
Email:mcculloughk@ukzn.ac.za

Abstract

Price discovery refers to flows of information, describing how and when information is reflected in market prices. It is not unusual for there to be more than one financial instrument in a given market that is linked to, or derived from, a single specific asset. This results in a complex set of relationships where market news is not only reflected by price changes in one asset, but may be reflected in a number of different assets, potentially at different points in time. Given the arbitrage potential of less than instantaneous price changes, studies of price discovery aim to determine which market reflects new information first. This study is the first to examine the price discovery process on the South African Johannesburg Stock Exchange since several significant market changes occurred, including the introduction of Exchange Traded Funds (ETFs). This paper examines the FTSE/JSE Top 40 Index, the Top 40 Index Futures and the SATRIX Top 40 ETF for the period 2003-2015. Findings demonstrate that the introduction of ETFs has not significantly impacted on the direction of the price discovery process. Return to equilibrium price occurs more quickly when the ETF is considered suggesting that the ETF contributes to an improvement in the price discovery process between spot and futures, however, further research is required in order to determine the nature of this relationship. Despite several changes to the market over the course of the most recent decade, the futures market remains the leading point in the price discovery process.

Keywords: Price discovery; SATRIX 40; FTSE/JSE Top 40.

1. Introduction

The significance of the theory of the 'efficient market' is widely acknowledged, and given the decades of work which have gone into developing the theory, methodologies and empirical work related to this topic, in the face of the far-reaching consequences of financial turmoil, it is unsurprising that discussions of market efficiency have grown in importance (Cochrane, 2008). At its most straightforward, market efficiency, as described by the Efficient Market Hypothesis (EMH), refers to asset prices reflecting market information; that is, a market is efficient if prices reflect all information. In a strongly efficient market the incorporation of information into prices should be instantaneous. However, this is not always the case, and so understanding exactly how market information is incorporated into prices is the foundation of studies in price discovery (Booth, So and Tse, 1999; Fleming, Ostdiek and Whaley, 1996), where price discovery refers to the flow of information, specifically in terms of how and when information is reflected in market prices. Chen and Chung (2012) speak to the significance of price discovery by highlighting that the process of price discovery is considered to be one of the significant contributions that the addition of derivatives trade has had on capital markets.

In efficient markets the price of an asset should reflect all publically available information. The prices of securities and their derivatives should therefore reflect this information simultaneously. If this were not the case, opportunities to make arbitrage profits would arise. Empirical research on the EMH, however, has shown that even ostensibly efficient markets are not efficient all the time (Cochrane, 2008). There are several reasons that markets may not be fully efficient, summarised by Stout (2003: p. 667) as, "including (but not limited to) … pricing anomalies, demand inelasticity, excessive volatility, delayed information response, and consistently superior traders." The consequence of this is that it has become necessary to give cognisance to the fact that the instantaneous and simultaneous pricing mechanism required under the EMH is not an entirely accurate or realistic depiction of how market prices actually respond to news events. Markets have been shown empirically to reach various levels of this fully efficient ideal, suggesting that efficiency achieved is neither static nor constant. Given that in the absence of market efficiency arbitrage becomes possible, the more modern approach to 'efficiency' analysis is to examine the flows of information within or between markets rather than trying to comment on some snapshot of efficiency (or lack thereof) (Cochrane, 2008).

This potential for price changes to occur at different times means that price discovery may be more generally described as the transmission of market information into the prices of assets linked by some collective commonality. This common element is usually examined empirically under two distinct conditions: the first being where a single common asset is traded on multiple markets, for example cross-listed shares; and secondly, where a single market offers multiple asset types all derived from one specific common asset (adapted from So and Tse, 2004). The commonality of linked assets, and not linked markets, is the focus of the research presented here. The link between the ideas of 'information in prices' and 'price discovery' comes down to understanding that new market information becomes incorporated into asset prices through the trading behaviours of market participants. When traders choose one asset over another to engage in trading, that asset should reflect price changes ahead of others through the forces of supply and demand balancing. Various reasons for these leads and lags are presented as possible explanations for price discovery, and are contained in the following four hypotheses (adapted from Chu, Hsieh and Tse, 1999).

Firstly, more leveraged securities tend to offer a better risk-return trade-off, which in turn attracts investors seeking assets with these characteristics. This is known as the *leverage hypothesis*. This hypothesis lends itself to a futures-led price discovery process, and as such a futures-led finding would indicate that informed traders are trading predominately on the futures market. The second hypothesis speaks to the costs involved in trading, and suggests that the market with lower costs will attract greater trade activity and thus lead the price discovery process. This is known as the *trading cost hypothesis*. Thirdly, the *uptick hypothesis* speaks to the fact that in many markets short sales may only take place if the last price change was non-negative. If, for example, the spot market is subject to this constraint (set by the market and/or other regulatory body) while the futures market is not, the futures market would be more likely to take the lead in the price discovery relationship. Lastly, the *marketwide information hypothesis* suggests that baskets of securities such as indices are usually designed along principles of diversification. As such, the assets linked to these basket index securities, such as futures and ETFs, allow traders to focus on market-related information as the security-level information has been accounted for in the basket formation. The implication of this is that securities linked to basket indices would lead the price discovery process.

Of particular interest in this paper, is the growing body of price discovery literature which examines the impact the introduction of a new instrument has on the existing price discovery process of a given market (Schlusche, 2009). While a large body of literature exists which examines the relationship between spot and futures markets (including Bose, 2007; Floros, 2009; and, Kumar and Chaturvedula, 2013) this only provides part of the picture, as the price discovery relationship may be affected by changes to the market environment such as by the inclusion of additional securities linked to the same underlying asset.

In the South African context, the Johannesburg Stock Exchange (JSE) has experienced several significant changes over the course of the last decade. These include several technological advances such as the introduction of the Millennium Exchange trading platform for equities in 2013; and structural changes including incorporating as a public listed company in 2005 (JSE, 2013). The introduction of ETFs and their resultant spread across the globe (JSE, 2010) may be argued to be a one of the more notable developments in financial markets in (relatively) recent history. The first ETF listed in South Africa was the SATRIX 40 in 2000. Since then the SATRIX 40 has grown to a market capitalisation of over R6.3 billion in January 2016 (SATRIX, 2016). Given the size and popularity of ETFs as an investment vehicle, examining the price discovery process between spot, futures and ETFs in the South African equity market poses an interesting question: has the introduction of ETFs impacted on the price discovery process between the spot and futures markets? By examining the influence of ETFs on this markets price discovery process, this paper contributes to deepening our understanding of information flows on the JSE.

2. The importance of Exchange Traded Funds

In any given market, it is not unusual for there to be more than one financial instrument based on a single asset. Equities may trade as single stand-alone shares; they may be used as the underlying asset tied to a financial derivative such as a futures or option contract; they may be grouped together and traded as an indexed asset; or they may be used as a tracking benchmark for other assets such as ETFs. What this results in is a complex set of relationships where market news is not necessarily reflected simultaneously in the prices of all linked assets, but where news reflects in the prices of different assets, at different points in time. When different classes of assets are all tied to the same underlying, it may be found that one of those linked assets incorporates new information better,

and more quickly, than its linked counterparts. When informed traders choose to trade in one asset over another, the price of that asset is likely to lead the price discovery process (Chu, Hsieh and Tse, 1999). Due to the arbitrage potential which these dispersed price changes represent, studies of price discovery aim to determine which asset reflects information first as a means of determining where the leading indicator of price movements for a particular group of linked assets is located.

While a price discovery analysis of linked assets may be performed on single shares, using an index-based approach allows the overall equity market to be examined. This approach is more comprehensive due to the fact that an index is a statistical measure of the changes in a portfolio of shares that represent some portion or sector of an overall market. Futures contracts are available on both individual shares and on various share-based indices. However, Individual Stock Futures trade volumes tend to be smaller than the trades in Index Futures (Alexander, 2008b). This provides additional support for using a basket of securities as the basis of the group of linked assets examined.

ETFs provide an alternative to index investment, given that direct investment in an index is often not possible. "ETFs are securities backed by a pool of assets, the return on which is expected to track a specific benchmark as closely as possible" (Kosev and Williams, 2011, p. 51). Investors in ETF shares are able to trade on the stock exchange throughout the day (Tse, Bandyopadhyay and Shen, 2006). Their notable growth both globally and in South Africa speaking to their popularity (JSE, 2010). Because ETFs are a fairly recent financial market innovation (Kosev and Williams, 2011), empirical research examining their impact on the price discovery process is limited. Their popularity, however, indicates that their presence should not be ignored. The increased popularity of ETFs suggests that it is a possibility that ETFs are being chosen as a preferred investment vehicle, which may theoretically result in ETFs incorporating information more efficiently, potentially resulting in ETFs making a significant contribution to the price discovery process.

Specifically, ETFs are expected to improve the pricing efficiency of the futures markets. This is made increasingly likely due to the fact that ETFs, unlike mutual funds, trade throughout the day on the stock exchange in a manner similar to normal stocks in that their prices are free to reflect relevant information during trading hours. Due to this, it may be argued that access to another linked asset, in this instance ETFs, may result in increased attention being given to arbitrage trading thereby reducing any mispricing and improving price discovery. Further,

any liquidity gains in the spot market arising from ETF trade may also reduce mispricing and improve price discovery between the spot and futures markets (Deville, Gresse and de Séverac, 2014).

3. Literature review

Price discovery is acknowledged to be influenced by a number of factors, including market depth, liquidity, volume and volatility. Markets which have greater liquidity, lower transaction costs and fewer restrictions are likely to contribute more to the price discovery process (Chen and Chung, 2012). In studies examining the relationship between spot and futures, it is often found that futures markets contribute more to price discovery (Schlusche, 2009). However, results are often mixed, with some studies noting that the spot market may lead at times, and that bidirectional causality may also be found.

For example, Bose (2007) examined the Indian stock market's price discovery process using daily closing prices of the futures contract on S&P CNX Nifty Index and the underlying index values from the NSE, from March 2002 to September 2006. It was found that there were clear flows of information between these two markets, with futures appearing to lead the spot market. In examining the long-run relationship between the two, some evidence of bidirectional informational flows between the markets was noted. Kumar and Chaturvedula (2013) also examined the Indian market, in this instance using individual stocks, and found evidence of the spot market leading the price discovery process. In explaining their findings they examined the clientele of the respective markets and concluded that due to the low percentage of institutional investors active in the futures market, informed traders were showing preference for the spot market, hence its role in leading prices (Kumar and Chaturvedula, 2013).

Of interest here, however, are studies which speak to the impact the inclusion of ETFs has on the informational flows of various markets. Internationally several papers have examined the effect ETFs have had on the relationship between spot and futures markets. Ivanov, Jones and Zaima (2013) examined the DJIA, S&P 500, S&P 400, NASDAQ 100 and Russell 2000 markets and found that the impact of the introduction of a linked ETF had involved a two-fold effect where the price discovery process of the futures market was negatively affected while that of the spot market was positively affected.

In contrast, Deville *et al.* (2014) found that the inclusion of ETFs had not improved the efficiency of the French CAC 40 market, nor had ETFs contributed to a decrease in the mispricing between spot and futures markets. So and Tse

(2004) found that the ETF on the Hang Seng market did not contribute to price discovery, and that it was the futures market which maintained the leading role in this process. Hasbrouck (2003) concurred, showing that the futures market led the S&P500 Index even when the ETF was incorporated into the analysis. However, it was noted that the ETF was contributing to the price discovery process, albeit in a small, yet statistically significant, fashion.

The first South African study on this topic by Ferret and Page (1998) examined four different South African indices from 1990 to 1996. They showed that information appeared to flow from the futures market into the spot market more strongly than the reverse. Fedderke and Joao (2001a) examined the ALSI 40 Index from 1996 to 1998, and noted that the long-run equilibrium was not affected by the emerging market crises of 1997 and 1998. In a second paper Fedderke and Joao (2001b) specifically examined the JSE All Share Index and its related futures contracts to determine where price discovery took place. Using intraday data and end-of-day data from 1996 to 1998, they too concluded that the price discovery process was led by the futures market.

Leng (2002) examined the All Share 40 Index using daily index and futures prices from 1996 to 2001. The results suggested that the South Africa market was inefficient before the Asian crisis, but that efficiency had increased during the crisis. The futures market responded first to market news. However, during the crisis there was a period during which the spot market took on the price discovery role. This, it was argued, was due to market uncertainty during the crisis period (Leng, 2002). More recently, Floros (2009) used daily data for the FTSE/JSE Top 40 Index and associated futures contract from 2002 to 2006 to examine this relationship using multiple methods. The study confirmed that a long-run relationship exists between the two, however, in this instance a bidirectional information flow was observed.

The existing literature in South Africa therefore suggests that the futures market has historically led the spot market, although the changing nature of price discovery direction was noted by periods of spot-led and bidirectional flow. However, the most recent period examined was until 2006. This paper aims to contribute to the existing literature by examining the price discovery process in South Africa with a larger and more recent data set, in order to address two key objectives. Firstly, the relationship between the FTSE/JSE Top 40 Index and futures in the most recent decade is determined. Secondly, the impact that the SATRIX Top 40 ETF has on the price discovery process between the FTSE/JSE Top 40 Index and associated futures contract is examined.

4. Data

Daily closing prices on the FTSE/JSE Top 40 Index, the Futures contracts on the FTSE/JSE Top 40 Index, and the SATRIX 40 ETF which tracks the FTSE/JSE Top 40 were obtained from Bloomberg's Professional Trading Platform. Observations for the index, futures and ETF series were taken from January 2003 through to March 2015. Collectively these three instruments represent a substantial portion of the potential market for the Top 40 Index and related products in South Africa. For example, using the JSE first quarter report of 2014, as at 10 January 2014 the Top 40 Index had a total index market capitalisation amounting to 84.39% of the ALSI market capitalisation. Shares contained in the top 40 index are screened to ensure liquidity is sufficient for trade of each component. Further, the SATRIX 40 is acknowledged to be the largest market capitalisation weighted ETF in South Africa (Profile Media, 2014). It is logical to examine whether its popularity has impacted the price discovery process of the market in South Africa.

The FTSE/JSE Top 40 index and ETF final daily trade occurs at 17:00 while the futures final trade occurs at 17:30. While one would ideally use a common trading time, the unavailability of historical tick data made the creation of a 'matched-closing-time' sample impossible over the long-term time frame used in this analysis. The benefits of examining a sample spanning more than a decade was considered preferable to examining only a few months of daily observations, which would have been the case if a 'matched-closing-time' was enforced. In South Africa, the study by Fedderke and Joao in 2001(b) accounted for the time difference as they had access to intraday data and were consequently able to look at both unmatched and matched closing-time data sets. However, their overall conclusions were the same under both the daily and intraday data sets.
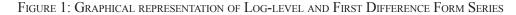
Futures contracts on the FTSE/JSE Top 40 expire in March, June, September and December on the 3rd Thursday of those months. The futures series in this study were created using a 1-day prior to expiry rollover to the next contract. McCullough, Murray and Strydom (2014) demonstrate in their paper examining rollover effects that the choice of rollover date may influence the results of causality type analysis. They conclude that using a rollover date closer to the expiry of the near contract is preferable due to the fact that futures contracts tend to be more highly traded in their expiry month, and so those prices are considered to contain relevant informational content which should be incorporated. The use of a near-expiry rollover is further confirmed by Kumar and Chaturvedula (2013); So and Tse (2004) and, Srinivasan and Bhat (2009).
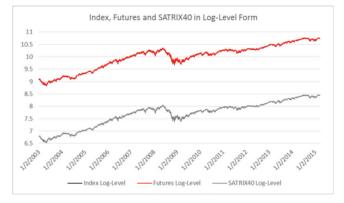
Each series of closing prices was converted into log form. The use of logarithmic transformation is commonly seen in studies of price discovery, and is popular in econometrics in general due to the fact that it removes trends which may obscure the relationship between two variables (Asteriou and Hall, 2011).

5. Methodology

5.1. Unit root tests and order of integration

Analysis begins by examining stochastic time series St (the index or spot series), Ft (the futures series), and Et (the ETF series) for the presence of a unit root using the Augmented Dicky-Fuller (ADF) Test. This is done firstly in log level form; and secondly, on the first differenced form (FDF) of each log price series. The *a priori* expectation is that each series is non-stationary in levels and integrated of order one, I(1). Figure 1 below shows a graphical representation of the series in log-levels, and then in FDF in order to demonstrate why this *a priori* expectation holds.

FIGURE 1: GRAPHICAL REPRESENTATION OF LOG-LEVEL AND FIRST DIFFERENCE FORM SERIES

The close relationship between the index and futures series is evident in their log-level form in the first panel of Figure 1 where the series plot so closely together that they are difficult to distinguish from one another. The ETF series, although different in 'size', follows a closely matched path to the index and index-futures series. A clear upward trend is noted, with the financial crises of 2008-2009 evidenced by the 'dip' during that time. In FDF the trend disappears, and a more stationary-looking process in all three series is observed.

5.2. Testing for cointegration

The tests for cointegration and subsequent formulation of the Vector Error Correction Model (VECM) are presented in two parts: the first examining the Spot-Futures two-variable case; and the second examining the three-variable Spot-Futures-ETF case. Johansen's cointegration is used due to its ability to identify multiple cointegrating relationships, and both Trace and Maximum Eigenvalue (ME) tests are estimated.

5.3. Testing for price discovery

5.3.1. Lag selection for the VECM

An appropriate lag length for the VECM must be determined (Brooks, 2008). The relationship between spot and futures is modelled firstly with an unrestricted Vector Autoregression VAR, and this is used to run a likelihood ratio test to determine an optimal lag length. Observations are daily trade days, with 5 trade days per week and so 5 lags would represent 1 week of trade, 20 lags a month, *etc*. Given that the lag-selection is being made for an error-correction model, it is economically sensible to assume that any 'correction' would occur within a month of trading, suggesting that the number of lags employed should not exceed 20. In line with Chu *et al.* (1999) the resulting VECM is run under several economically sensible lag length choices, and the effect, if any, on the interpretation of results is noted.

5.3.2. The VECM

The VECMs used to describe price discovery in the South African equities market are given below. In both cases the models are adapted from those discussed by Asteriou and Hall (2011). The relationship between Spot and Futures prices is described by the following two-variable VECM:

$$\Delta S_t = \alpha_1 + \alpha_S \varepsilon_{t-1} + \sum_{i=1}^{p} \beta_{11} \Delta S_{t-i} + \sum_{i=1}^{p} \beta_{12} \Delta F_{t-i} + \varepsilon_{St}$$

$$\Delta F_t = \alpha_2 + \alpha_F \varepsilon_{t-1} + \sum_{i=1}^{p} \beta_{21} \Delta S_{t-i} + \sum_{i=1}^{p} \beta_{22} \Delta F_{t-i} + \varepsilon_{Ft}$$

In the three-variable case, the VECM examines the Spot, Futures and ETF series as follows:

$$\Delta S_t = \alpha_1 + \alpha_S \hat{\varepsilon}_{t-1} + \sum_{i=1}^{p} \beta_{11} \Delta S_{t-i} + \sum_{i=1}^{p} \beta_{12} \Delta F_{t-i} + \sum_{i=1}^{p} \beta_{13} \Delta E_{t-i} + \varepsilon_{St}$$

$$\Delta F = \alpha_2 + \alpha_F \hat{\varepsilon}_{t-1} + \sum_{i=1}^{p} \beta_{21} \Delta S_{t-i} + \sum_{i=1}^{p} \beta_{22} \Delta F_{t-i} + \sum_{i=1}^{p} \beta_{23} \Delta E_{t-i} + \varepsilon_{Ft}$$

$$\Delta E_t = \alpha_3 + \alpha_E \hat{\varepsilon}_{t-1} + \sum_{i=1}^{p} \beta_{31} \Delta S_{t-i} + \sum_{i=1}^{p} \beta_{32} \Delta F_{t-i} + \sum_{i=1}^{p} \beta_{33} \Delta E_{t-i} + \varepsilon_{Et}$$

The two-variable VECM results in a 2×2 matrix, hence the notation β_{11}, β_{12}, β_{21} and β_{22} of the coefficients associated with the lagged terms within the first system. The three-variable VECM which includes ETF prices forms a 3×3 matric, and so the notation on the lagged terms in that system ranges from β_{11} to β_{33}. 'ρ' refers to the number of lags. The lagged terms within each equation allow comment to be made on past prices impact on current prices. For example, lagged futures prices significant in the spot equation would indicate that previous futures prices influenced the spot price formation, that is, futures prices have lead the spot price. Alternatively, spot prices will be said to 'lead' futures prices if lagged values of the spot price can be used to predict current and future values of the futures series better than using lagged values of the futures prices alone (Alexander, 2008a). Being able to determine in which market the point of price discovery lies may be considered the pivotal finding in this model and determining this relies on being able to establish where new information is first reflected – the changed futures price or the changed spot price (Mahalik, Acharya and Babu, 2009).

The speed of adjustment coefficients, represented as α_S, α_F and α_E in the VECMs above, represent the rate of adjustment towards equilibrium. At least one of the speed of adjustment coefficients should be statistically significantly different from zero in the presence of cointegration as this indicates that deviations from the long-run equilibrium position are corrected for in the short run (Bose, 2007). The absolute values of these coefficients should not be too large, as a finding greater than 1 would indicate a more than 100% adjustment to the long-run equilibrium, which in turn suggests the variables move apart and would not be cointegrated (Enders, 2010).

6. Results

6.1. Unit root tests and order of integration

Results in Table 1 are from the ADF test with trend and intercept. 'Level' here refers to the series in log-level form, while 'FDF' refers to the same series in first difference form. In all instances the form of the test (with or without intercept and/or trend) did not affect the conclusion drawn: that all price series are non-stationary and contain a unit root; and that each price series is I(1). Both AIC (Akaike Information Criterion) and SIC (Schwarz Information Criterion) criteria were examined in order to estimate an appropriate lag length, these being the most commonly used (Enders, 2010). Conclusions drawn, however, were not sensitive to the number of lags indicated by these criteria, and so Table I shows the ADF test on each series using 5 lags, this representing an economically sensible period of 1 week of trade. Results were further confirmed by both Phillips-Perron (PP) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests. However, these approaches reached the same conclusion and so their output is not shown here.

TABLE 1: ADF TESTS ON LOG-LEVEL AND FIRST DIFFERENCE SPOT, FUTURES AND ETF SERIES

	ADF Test Statistic	**Prob.**	**Result**
Spot Level	-1.812	0.6987	Spot prices contain a unit root.
Spot FDF	-25.136*	0.0000	Spot series Integrated of Order 1 I(1).
Futures Level	-1.832	0.6893	Future prices contain a unit root.
Futures FDF	-25.123*	0.0000	Futures series Integrated of Order 1 I(1).
ETF Level	-1.831	0.6894	ETF prices contain a unit root.
ETF FDF	-25.082*	0.0000	ETF series Integrated of Order 1 I(1).

*Significant at 1%; **Significant at 2.5%; *** Significant at 5%

6.2. Cointegration

The number of cointegrating relationships is determined by both the Trace and Maximum Eigen-value (ME) Tests. All versions of these tests are considered (with/without trend and/or intercept). Results, as shown in Table 2, are as expected. These are consistent in all economically sensible forms of the model, and further, are not sensitive to lag length selection. The less economically plausible results (no-trend-no-intercept and quadratic-trend-with-intercept) are ignored as per Asteriou and Hall (2011). There is clear evidence of 1 cointegrating long–run relationship existing between spot and futures prices; while 2 long-run cointegrating relationships are present in the spot, futures and ETF combination.

TABLE 2:COINTEGRATION TESTS

	II Panel A		II Panel B	
	Spot-Futures Cointegration		**Spot-Futures-ETF Cointegration**	
	Trace Test	ME Test	Trace Test	ME Test
Number of Cointegrating Relationships	1	1	2	2

6.3. Lag length criterion tests

Table 3 shows the results of the Lag Length Selection Tests run under AIC and SIC. Each test is at the 5% level and the * represents the optimal number of lags indicated by each criterion. In both the two-variable and three-variable case there is some variation in lag length under the different criteria used, although this is not unusual (Asteriou and Hall, 2011). Results for lags which were not significant under these measures were removed. One may argue for preferring the SIC due to its superior large sample properties (Enders, 2010). However, one may also argue for using the AIC due to it being more commonly used in time series analysis (Asteriou and Hall, 2011). In all instances the lag lengths indicated fall within the economically sensible bounds discussed in section 5.3.1.

TABLE 3: LAG LENGTH CRITERIA TESTS

Lag Length	III Panel A		III Panel B	
	Spot-Futures VECM		**Spot-Futures-ETF VECM**	
	AIC	SIC	AIC	SIC
3	-14.9087	**-14.8809***	-22.9422	-22.8827
4	-14.9129	-14.8772	-22.9685	**-22.8912***
6	**-14.9152***	-14.8637	-22.9900	-22.8769
14	-14.9142	-14.7992	**-22.9966***	-22.7407

* Indicates the Lag Length Selected by each Information Criteria

Initially each VECM was formed with the smaller lag length suggested by SIC, as the VECM output becomes extensive as more lags are included. However, the VECM's formed with these shorter SIC lag lengths contained some autocorrelation in the resulting residual series. Consequently, the longer AIC lag lengths were used in final estimation, resulting in the issue of autocorrelation in the residuals being resolved. As a result, the spot-futures VECM is estimated with 6 lags, and the spot-futures-ETF VECM with 14 lags. Although these models

residuals both continued to have non-normal errors, Srinivasan (2011) noted that "Johansen (1988) tests are shown to be fairly robust to [the] presence of non-normality (Cheung and Lai, 1993) and heteroscedastic disturbances (Lee and Tse, 1996)." Non-normality with fat tails and the presence of heteroscedasticity in the residuals due to ARCH effects are fairly typical for return series, and so while an estimated model should have Gaussian error terms (Asteriou and Hall, 2011), normality of the errors is not a necessary condition for results to be valid when one is examining any of the VAR models (Leutkepohl, 2011).

6.4. VECM price discovery

The VECM estimates are presented in Tables 4, 5 and 6 below. Lags which are insignificant across all assets examined have been removed from these tables to improve their readability. Standard Errors are in () and t-statistics are in [].

The sample size used in the Spot-Futures and Spot-Future-ETF VECM contain just over 3000 observations per asset. As a result of this large sample size, the degrees of freedom on the t-statistics are taken at the ∞ level, which represents a critical value of 2.3263 at the 1% level (*), with less consideration given to the 2.5% level of 1.9600 (**) and the 5% level of 1.6449 (***). The emphasis on the 1% level is due to the fact that a large sample is associated with decreasing standard errors and increasing t-statistics, potentially inflating the acceptance region (Brooks, 2008). As result, the 10% level of significance is ignored (associated with a critical value of 1.2816), and greater weight given to results significant at the 1% level.

TABLE 4: ERROR CORRECTION TERM ESTIMATES AND COINTEGRATING EQUATIONS

	IV Panel A		IV Panel B		
	Futures-Spot VECM		Futures-Spot-ETF VECM		
	Futures	Spot	Futures	Spot	ETF
Error Correction Term1 (probability)	**-0.1415188*** (0.005)	-0.0660431 (0.178)	0.1569345 (0.178)	0.0967178 (0.398)	**0.3735443*** (0.001)
Error Correction Term2 (probability)	- -	- -	0.0047207 (0.961)	-0.0208325 (0.825)	**-0.2529063*** (0.007)
	Cointegrating Equation		**Coint. Eq 1**	**Coint. Eq 2**	
Log-Level Spot	1.0000		1.0000	0.0000	
Log-Level Futures	-1.001841* (0.00119)		-1.001765* (0.0000)	-1.001717* (0.0000)	
Log-Level ETF	-		0.0000	1.0000	
Constant Term	0.023174		0.022575	2.322053	
Half Life	4.53 days		1.48 days	2.38 days	

*Significant at 1%; **Significant at 2.5%; *** Significant at 5%

Table 4 Panel A shows evidence of a price discovery process existing between spot and futures markets, indicated by the statistically significant ECT in the Futures cointegrating equation. Only if both error correction terms in this bivariate VECM were significant would bidirectional causality be concluded. The value of -0.1415 indicates that 14.15% of the adjustment to return to equilibrium takes place each period. In this two-variable case normalisation takes place against the futures series.

The key research question of this paper is whether this analysis is impacted at all by including the SATRIX 40 into the price discovery relationship. To this end, Table 4 Panel B shows the VECM constructed with all three variables: spot, futures and the ETF. Constructing a three-variable VECM requires that some consideration be given to which variables fall into the two cointegrating equations, as the choice of which variable to normalise needs to be justified. In the two-variable spot-futures case it is commonly accepted that normalisation be applied to the futures price, and so in the three-variable VECM it makes sense that one of the two cointegrating relationships describes the spot-futures case. Given that our *a priori* expectation is of a futures-led market, normalisation was structured such that the first cointegration equation still examines the spot-futures scenario as in the two-variable case, while the second cointegrating equation speaks to the relationship between the Futures and ETF. In all cointegrating equations (displayed in the lower section of Table 4) the parameter values are as expected, being highly statistically significant and close to 1.

Statistical significance of at least one ECT in each of the two cointegrating equations is a requirement of the error correction terms from the three-variable VECM presented in Table 4 Panel B. This is seen to be present by the significance of both of the ECT terms in the ETF equation column. A finding of two significant error correction terms, one positive and one negative, is what is expected in a three-variable VECM. It can be seen that both are significant with values that are less than 1, demonstrating a gradual adjustment towards equilibrium.

In this instance the rates of correction to the equilibrium error are much larger (faster) than those indicated under the Spot-Futures case (37.35% and 25.29% compared to the earlier 14.15%). That is, when more variables involved in the price discovery process are incorporated into the analysis, the speed of adjustment back to equilibrium is shown to occur more quickly than in the two-variable VECM. A faster adjustment back to equilibrium, where mispricing exists for a shorter period of time, may be interpreted as an improvement in the price discovery process between spot and futures markets. This is also seen in

the half-life[1] estimations, contained in the final row of Table 4, where the number of days it takes to half the gap between actual and equilibrium price is seen to decrease from 4.53 days in the 2-variable estimation, down to 1.48 days when the ETF is included. Finding that mispricing decreases is in line with Deville *et al*. (2014) who found that there had been a reduction in spot-futures mispricing following the introduction of ETFs, however, they discussed that this may be attributed to structural changes to the market (indirect effects) following the introduction of ETFs, rather than as a direct effect of the availability of ETF trade.

The next two tables present the VECM findings of short-run causality between the spot and futures markets (Table 5); followed by those of the spot, futures and ETF markets (Table 6). Given the fairly lengthy lag length employed (6 lags in the two-variable case and 14 in the three-variable case), only statistically significant lags are displayed.

TABLE 5: DETERMINING THE IMPACT OF PRICE CHANGES: THE SPOT-FUTURES VECM
(*5 FDF LAGS*)

	SPOT	FUTURES
Panel A: Lags Significant at 1%, 2.5% and 5%.		
SPOT Lag 1	**-0.3637***	-
SPOT Lag 2	-0.2407**	-
FUTURES Lag 1	**0.3671***	-
FUTURES Lag 2	0.2179***	-
Panel B: Lags Significant at 1% only.		
SPOT Lag 1	**-0.3637***	-
FUTURES Lag 1	**0.3671***	-

*Significant at 1%; **Significant at 2.5%; *** Significant at 5%

Table 5 shows clear evidence of the spot-futures relationship being led by the futures market, where evidence of the futures price impacting on the spot price is seen by the significant first lag of the futures price impacting on the spot series, as well as to a lesser extent as the second lag impacting on the spot series. That is, previous futures prices are seen to influence current spot prices. The reverse, however, is not true. Previous spot prices are seen to influence the current spot price at lag 1 and to a lesser extent at lag 2. That is, the previous day's spot price is, not surprisingly, seen to be an influence on the current spot

---

[1] Half-life is estimated as $\log(0.5)/\log(1- |ECT|)$ in line with Hammoudeh, Nguyen and Sousa (2014).

price. However, there is no feedback noted from the spot price into the futures price. This result shows that the market in South Africa continues to be led by the futures market when considering the relationship between the FTSE/ JSE Top 40 Index and Futures. The finding that the market remains futures led augments previous studies done on the South African market.

TABLE 6: DETERMINING THE IMPACT OF PRICE CHANGES: THE SPOT-FUTURES-ETF VECM (*13 FDF LAGS*)

	SPOT	FUTURES	ETF
Panel A: Lags Significant at 1%, 2.5% and 5%.			
SPOT Lag 1	**-0.4091***	-	-
SPOT Lag 2	-0.3716**	-	-
SPOT Lag 3	-0.3351***	-	-
SPOT Lag 4	-0.3436**	-0.3523***	-
SPOT Lag 11	-	-0.2882***	-
FUTURES Lag 1	**0.3772***	-	**0.3771***
FUTURES Lag 2	-	-	0.2352***
FUTURES Lag 3	-	-	-
FUTURES Lag 11	0.2219***	0.2668**	0.2299***
ETF Lag 1	-	-	**-0.6039***
ETF Lag 2	-	-	**-0.3765***
ETF Lag 3	-	-	-0.2702**
ETF Lag 4	0.2235***	-	-
ETF Lag 13		0.1310**	
Panel B: Lags Significant at 1% only.			
	SPOT	FUTURES	ETF
SPOT Lag 1	**-0.4091***	-	-
FUTURES Lag 1	**0.3772***	-	**0.3771***
ETF Lag 1	-	-	**-0.6039***
ETF Lag 2	-	-	**-0.3765***

*Significant at 1%; **Significant at 2.5%; *** Significant at 5%

Looking to the lagged effects described in Table 6 Panel A, it is evident from the significant lags in the spot equation that the spot series is still influenced by its own past values. Further, the statistically significant lagged futures prices in the spot equation demonstrates that there is a still significant influence on the spot series by the futures series. There is some indication of a small contribution made by ETFs to the price discovery process as the ETF on lag 4 is seen to have an influence on spot prices, albeit only a weak one at the 5% level.

The statistically significant ETF lags and futures lags provide evidence that past futures prices, as well as past ETF prices, influence the prevailing ETF price. In the futures equation, however, there are no statistically significant lags at the 1% level feeding back to the futures price. In the futures column there is now some weak evidence (at the 2.5% and 5% level) to suggest that there may be some influence of the 4th and 11th lag spot prices and 13th lag ETF price on futures prices, which may imply some weak bicausality between these markets. Given the significance level and the size of the sample, however, this interpretation is not given the same value as the findings at the 1% level.

Given the sample size and associated preference for the 1% significance level, both Table 5 and Table 6 include a Panel B component which shows only those lags significant at the 1% level. These panels show that it is the futures market which leads the price discovery process, even in the three-variable case; and that this effect is primarily seen occurring in a single day. In addition, the inclusion of the ETF indicates that ETF prices are being led by their own previous prices as well as by previous futures prices. This impact on the ETF by the futures market interestingly shows that the futures market in South Africa leads the price discovery process even when the ETF is included in the analysis. These results mirror those of So and Tse (2004) who also found that ETFs did not contribute to the price discovery process in the Hang Seng market.

6. Conclusions

The objective of this research was to examine the influence ETF trade, in the form of the SATRIX 40, has had on the price discovery relationship between the spot and futures market for the FTSE/JSE Top 40 Index. Previous studies in South Africa only examined the spot-futures relationship. A long-run cointegrating relationship between the FTSE/JSE Top 40 spot, futures and ETF markets, which is expected of financial instruments related to a common base, was shown to exist.

Guided by the importance of ETF trade, given this instrument's ability to allow investors to trade easily in a basket security as compared to the index proper, this study sought to add to the knowledge of price discovery in South Africa by examining the price discovery process between spot, futures and the ETF. Internationally there had been mixed reports on the influence exchange traded products had had on information transmission in different markets. Results in this South African context showed that the market in South Africa continues to be led by the futures market even when the impact of the ETF is

considered. Past futures prices were shown to influence both current spot and current ETF prices.

This finding speaks to the continued importance of futures derivative trade in the South African context, specifically in the context of this market's contribution to price discovery in the market. While trade in ETF instruments has grown considerably over the last few years, the impact of this asset's popularity on price discovery in South Africa is not evident over the studied time period.

The continued dominance of the futures market may be influenced by the fact that the market for futures closes later than the spot and ETF markets. However, as discussed in the introduction, this extra half hour of price information is a reality of the information environment of the South African market. The later closing time on the futures may mean that the futures market contains more information than the spot and ETF markets due to its longer trading times. One may speak to the timing difference as a 'limitation' of the data set. However, the trading times of these South African markets are a reflection of the reality of these markets' information environment. As such, that timing differential as a factor influencing any lead-lag relationship is not necessarily a result caused simply by the time mismatch, but rather may be argued to be due to what that the timing difference implies: more trading time results in more information (in this case, in the futures market). The question of whether this futures-led finding remains true within a set of common trading times is an avenue for future research. Given the complexities of working with tick data, however, such research falls outside the confines of this paper.

While the day-to-day impact shows clearly that the price discovery process between the index, index-futures and ETF is futures led, there is weak evidence to suggest that the ETF market influences informational movements in the equities market, suggesting that this matter be revisited in the future should the market for ETFs continue to grow in popularity. The finding that the return to equilibrium price occurs more quickly when the ETF is considered in the price discovery process suggests that the ETF is contributing to an improvement in the price discovery process and a reduction in spot-futures mispricing between the FTSE/JSE Top 40 spot and futures markets. Further research is required in order to determine the nature of this relationship. That is, the exact nature of the direct and/or indirect effects of the SATRIX 40 on spot-futures mispricing in the South African market requires further investigation. Another avenue of future research is to augment these price discovery findings with a detailed analysis of the volatility characteristics and volatility spillover effects of these markets.

Biographical Notes

Kerry McCullough is a Lecturer in Finance at the University of KwaZulu-Natal, South Africa. She holds a Master's degree in Finance and a Diploma in Insolvency Law and Practice. She is currently enrolled for a PhD in Finance examining information transmission on the JSE. Her research interests include price discovery, volatility spillover, commodity markets, and higher education focused action research.

Acknowledgments

References

Alexander, C. (2008a). *Practical financial econometrics*. Chichester, John Wiley & Sons.

Alexander, C. (2008b). *Pricing, Hedging and Trading Financial Instruments*. Chichester, John Wiley & Sons.

Asteriou, D. and Hall, S.G. (2011). *Applied econometrics*. 2nd Edition. Basingstoke, Palgrave Macmillan.

Booth, G.G., So, R.W. and Tse, Y. (1999). Price discovery in the German equity index derivatives markets. *Journal of Futures Markets*, vol 19, pp. 619-643.

Bose, S. (2007). Contribution of Indian index futures to price formation in the stock market. *Money & Finance*, vol 3, no. 1, February 2007, pp. 39-56.

Brooks, C. (2008). *Introductory econometrics for finance*. 2nd Edition, Cambridge, Cambridge University Press.

Cochrane, J.H. (2008). Introduction for Gene Fama. Presented at Fama's Talk on 'The History of the Theory and Evidence on the Efficient Markets Hypothesis' at the University of Chicago's Graduate School of Business. 10th October 2008.

Chen, W. and Chung, H. (2012). Has the introduction of S&P500 ETF options led to improvements in price discovery of SPDRs? *The Journal of Futures Markets*, vol 32, no 7, pp. 683-711.

Chu, Q.C., Hsieh, W.G. and Tse, Y. (1999). Price discovery on the S&P500 index markets: an analysis of spot index, index futures and SPDRs. *International Review of Financial Analysis*, vol 8, no 1, pp. 21-34.

Deville, L., Gresse, C. and de Séverac, B. (2014). Direct and indirect effects of index ETFs on spot-futures pricing and liquidity: evidence from the CAC 40 Index. *European Financial Management*, vol 20, no 2, pp. 352-373.

Enders, W. (2010). *Applied econometric time series,* 3rd Edition, Hoboken, NJ, John Wiley & Sons.

Fedderke, J. and Joao, M. (2001a). Arbitrage, cointegration and efficiency in financial markets in the presence of financial crisis. *South African Journal of Economics*, vol 69, pp. 366-384.

Fedderke, J. and Joao, M. (2001b). Price discovery in South African financial markets: investigating the relationship between South Africa's stock index futures market and the underlying market. *Journal for Studies in Economics and Econometrics*, vol 25, no 2, pp. 1-23.

Ferret, A. and Page, M.J. (1998). Cointegration of South African index futures contracts and the underlying spot market indices. *Journal for Studies in Economics and Econometrics*, vol 22, no 1, pp. 69-90.

Fleming, J., Ostdiek, B. and Whaley, R. (1996). Trading costs and the relative rates of price discovery in stock, futures and option markets. *The Journal of Futures Markets*, vol 16, pp. 353-387.

Floros, C. (2009). Price discovery in the South African stock index futures market. *International Research Journal of Finance and Economics*, vol 34, pp. 148-159.

Hasbrouck, J. (2003). Intraday price formation in US equity index markets. *The Journal of Finance*, vol 58, pp. 2375-2400.

Hammoudeh, S., Nguyen, D.K. and Sousa, R.M. (2014). What explains the short-term dynamics of the prices of CO2 emissions? Working Paper Series University of Minho NIPE WP/04/2014. Available online at: http://www.nipe. eeg.uminho.pt/Uploads/NIPE_WP_04_2014.pdf. Accessed 18th July 2016.

Ivanov, S.I., Jones, F.J. and Zaima, J.K. (2013). Analysis of DJIA, S&P500, S&P400, NASDAQ 100 and RUSSELL 2000 ETFs and their influence on price discovery. *Global Finance Journal*, 2013, pp 171–187.

JSE (2010). The South African and Global Landscape of Exchange Traded Funds (ETFs). *Johannesburg Stock Exchange*, Available online at: https://www.jse.co.za/content/JSEEducationItems/The%20South%20African%20 and%20global%20landscape%20of%20ETFs%20(1).pptx. Accessed 28th April 2016.

JSE (2013). JSE Overview. *Johannesburg Stock Exchange*, Available online at: https://www.jse.co.za/about/history-company-overview. Accessed 27th January 2016.

Kosev, M. and Williams, T. (2011). Exchange-traded funds. *RBA Bulletin*, March, pp. 51-59.

Kumar, K. and Chaturvedula, C. (2013). Price leadership between spot and futures markets. *Journal of Applied Finance and Banking*, vol 3, no 1, pp 93–107.

Leng, H.M.J. (2002). The South African share index futures and share markets: efficiency and causality revisited. *Journal for Studies in Economics and Econometrics*, vol 26, no 3, pp 1–18.

Luetkepohl, H. (2011). Vector autoregressive models. European University Institute Department of Economics Working Papers, ECO2011/30. Available online at: http://cadmus.eui.eu/bitstream/handle/1814/19354/ECO_2011_30.pdf?sequence=1. Accessed 27th July 2016.

Mahalik, M.K., Acharya, D. and Babu, M.S. (2009). Price discovery and volatility spillovers in futures and spot commodity markets: some empirical evidence from India. Paper Presented to Fourth Annual International Conference on Public Policy and Management. August 9-12. Available online at: http://www.igidr.ac.in/pdf/publication/PP-062-10.pdf. Accessed 30th July 2010.

McCullough, K., Murray, M. and Strydom, B. (2014). Market efficiency and price discovery: a comment on futures rollover practices. *Journal of Contemporary Management*, vol 11, pp. 367-384.

Profile Media (2014). *Profile's stock exchange handbook: 2014 – Issue 1*, Pietermaritzburg, Profile Media.

SATRIX (2016). SATRIX Top 40. *SATRIX,* Available online at: http://www.satrix.co.za/fund-research-product.php?type=ETF&fund=satrixALSI40Fund. Accessed 27 January 2016.

Schlusche, B. (2009). Price formation in spot and futures markets: exchange traded funds vs. index futures. *The Journal of Derivatives*, vol 17, pp. 26-40.

So, R.W. and Tse, Y. (2004). Price discovery in the Hang Seng index markets: index, futures, and the tracker fund. *Journal of Futures Markets*, vol 24, pp. 887-907.

Srinivasan, P. and Bhat, K.S. (2009). Spot and futures markets of selected commercial banks in India: what causes what? *International Research Journal of Finance and Economics*, vol 31, pp. 28-40.

Stout, L.A. (2003). The mechanisms of market inefficiency. *Journal of Corporation Law*, vol 28, pp. 635-639.

Tse, Y., Bandyopadhyay, P. and Shen, Y. P. (2006). Intraday price discovery in the DJIA index markets. *Journal of Business Finance & Accounting*, vol 33, pp. 1572-1585.