Recent developments in exchange-traded fund literature

Pricing efficiency, tracking ability, and effects on underlying securities

Received 19 December 2012
Revised 26 December 2012
Accepted 27 December 2012

Narat Charupat and Peter Miu
DeGroote School of Business, McMaster University, Hamilton, Canada

Abstract

Purpose – The purpose of this paper is to provide a brief review of three strands of the literature on exchange-traded funds.

Design/methodology/approach – The paper starts with a review of the history of the growth of exchange-traded funds and their characteristics. The paper then examines the key factors and findings of the existing studies on, respectively, the pricing efficiency, the tracking ability/performance, and the impact on underlying securities of exchange-traded funds.

Findings – Although there has been a substantial amount of research conducted to advance our knowledge on the trading, management, and effect of exchange-traded funds, the findings are still far from conclusive in addressing a number of research questions.

Practical implications – Investors and other market participants will find this review informative in enhancing the understanding of exchange-traded funds.

Originality/value – By highlighting the general theme of the related research findings, the paper provides a systematic review of the existing literature that future researchers can utilize in developing their research agenda.

Keywords Exchange-traded funds, Literature review, Pricing efficiency, Tracking errors, Underlying securities, Pricing, Trade

Paper type Literature review

1. Introduction

Over the past decade, exchange-traded funds (ETFs) have grown from a small, niche index-tracking product to become one of the most successful innovations in the history of investment. As of the end of 2011, the combined assets under management (AUM) of all ETFs traded on exchanges around the world were US$ 1.52 trillion, an increase of approximately 1,400 percent from the assets in 2001[1]. There are now close to 4,000 different ETFs listed on over 50 exchanges.

ETFs have a relatively short history. It is commonly acknowledged that the first ever ETF was the Toronto Stock Exchange Index Participations introduced in 1990 and designed to track the TSE-35 stock index. However, ETFs can trace their conceptual lineage to earlier "basket securities" products whose goal was to enable investors to trade a basket of securities in a cost-effective way. One example of such products is the Index Participation Shares, which came out in 1989 and whose untimely

JEL classification – G10, G12, G14, G23

Managerial Finance
Vol. 39 No. 5, 2013
pp. 427-443

0307-4358
DOI 10.1108/03074351311313816

demise provided an incentive for exchanges to look for a replacement[2]. In 1993, the American Stock Exchange introduced trading on the Standard & Poor's Depositary Receipts (SPDR) whose aim was to track the S&P 500 index. In the years that followed, the SPDR proved to be a huge success and intensified investor demand for low-cost basket-trading vehicles.

In the 1990s, most ETFs were funds that tracked market-wide equity indices, sector indices and fixed-income indices. As the market for these ETFs became saturated, ETF providers (or sponsors) came up with products that were based on other asset classes such as commodities and currencies. Recent innovations (late 2000s) include ETFs that:

- hold physical commodities such as gold and silver;
- are actively managed, rather than passively tracking an index;
- employ leverage to generate returns that are a positive or negative multiple of the index returns; and
- combine a long position (in one index) with a short position (in another index).

The popularity of ETFs is due to their several benefits that distinguish them from other index-tracking products such as conventional mutual funds and closed-end funds. The major benefits are:

- intraday trading;
- tax efficiency;
- low expense ratios; and
- cost transparency.

On the other hand, ETFs have a few disadvantages, one of which is that investors have to pay commissions and bid-ask spreads when they buy/sell them. The bid-ask spreads can be significant for funds that have low liquidity[3]. In addition, some recent ETFs such as actively managed ETFs have higher management fees, while leveraged ETFs are not designed to be held for a long period, which necessitates frequent trading.

The success of ETFs has also attracted interest from researchers. There are three active strands of the literature on ETFs. The first strand examines ETFs' pricing efficiency, i.e. how close market prices are to the funds' net asset values (NAVs) and how quickly price discrepancies disappear. In other words, these studies measure how well the funds' creation/redemption process works in arbitraging away the differences between prices and NAVs. The second strand investigates the performance of ETFs, i.e. how successful they are in achieving their goals. Typically, this is done by measuring ETFs' tracking errors, defined as the differences between the funds' NAV returns and the returns of the underlying indices that the funds track. The tracking errors can be compared among various ETFs and also with competing products such as index mutual funds.

The third strand of literature looks at the effects of ETF trading on related securities such as constituent stocks of the underlying indices and derivatives on the same indices. For the effects on constituent stocks, these studies examine whether there is a change in their trading characteristics (e.g. bid-ask spreads, trading volume, etc.) after the introduction of the ETFs. Such a change may indicate a migration of investors and/or a change in the composition of different kinds of traders who trade the stocks (i.e. the proportions of informed vs uninformed traders). For the effects on derivatives, these studies typically investigate whether the introduction of ETFs facilitates arbitrage activities (i.e. ETFs make it easier for arbitrageurs to short the underlying indices, even on a downtick). If it does, the pricing of the derivatives will be more efficient. Included in this strand of research are studies on the price discovery process. These studies attempt to determine which instrument (i.e. ETFs, derivatives or the underlying stocks) is the first to react to new information, and what role each instrument plays in the process.

In this paper, we summarize the findings of the existing literature. We start with the studies on pricing efficiency in the next section. Then in Section 3, we document the current literature on ETF tracking errors. Section 4 addresses the findings on the effects of ETF trading. Finally, Section 5 concludes.

2. Pricing efficiency

One unique feature of ETFs is the creation/redemption process, under which select traders can purchase (sell) large lots (or creation units) of an ETF directly from (to) the fund issuer at the NAV[4]. Depending on the ETFs, the purchase and sale are done in kind (i.e. using baskets of securities comprising the underlying benchmarks) or in cash. Generally, ETFs that hold the constituent stocks of their underlying indices will use an in-kind process, while ETFs that use derivative securities to replicate the underlying returns will use an in-cash process. The former type of ETFs is common in the US market, while the latter type is more prevalent in the European markets[5].

The creation/redemption process establishes arbitrage bounds for market prices relative to the NAVs. If, for example, the market price of an ETF (which uses an in-kind process) is below its value, traders can buy units of the ETF in the market, redeem them for the underlying basket of securities, and capture the price difference. As a result, in theory at least, the market price of a fund should be close to its NAV.

In practice, the effectiveness of arbitrage and the size of the arbitrage bounds depend on several factors including the amount of transaction costs, the bid-ask spreads and the requirements of the creation/redemption process. Different ETFs impose different rules, especially with respect to the timing of creation/redemption order submission, the length of the settlement period, and the costs. These differences can increase the risk of arbitrage transactions. They also affect the ability of arbitrageurs to hedge such risk[6].

Various studies have investigated the pricing efficiency of ETFs and the ability of arbitrage in minimizing the differences between market prices and NAVs. Typically, these studies examine whether ETFs trade, on average, at a premium/discount to their NAVs, using daily data or (more recently) intraday data. Some studies also look at the speed at which the premium/discount disappears, which indicates how well arbitrage works. The findings can be summarized in a few main points.

First, premiums/discounts are generally small and do not persist, especially for ETFs that track major domestic equity indices. For example, Ackert and Tian (2000) and Elton *et al.* (2002) examine the pricing of the SPDR (which tracks the S&P 500 index) and find no economically significant premium or discount[7]. Elton *et al.* also show that arbitrageurs take advantage of price discrepancies and so the discrepancies disappear within one day. Curcio *et al.* (2004) compare price deviations of the SPDR to those of the Nasdaq 100 Index Tracking Stock (QQQ) and find that the average deviations are small for both ETFs, but the range and the standard deviation of price deviations for QQQ are larger than those for SPDR. Engle and Sarkar (2006) study 21 ETFs on US domestic indices, including major ones such as the S&P 500, the Nasdaq 100, the Dow Jones Industrial Average and the Russell 1000, and also sector indices. They report that price deviations are on average very small and within transaction costs and bid-ask spreads. Their results also show that the volatility of price deviations is related to the volatility of the underlying NAVs, which is consistent with the above-mentioned findings of Curcio *et al.* that price deviations of QQQ are more volatile than those of SPDR (as the Nasdaq 100 index is more volatile than the S&P 500 index).

Second, price deviations are larger and more volatile for ETFs that track international indices, which are normally referred to as international ETFs or country ETFs. This is mainly due to the fact that the NAVs used in the calculations are based on prices from earlier closing times than the US market close. Although the NAVs are adjusted for the prevailing exchange rates, they do not fully incorporate other information arriving at the market during the opening hours of the US market. In addition, the arbitrage mechanism is not effective as trading hours are not the same (between the US market and the countries of the indices). For example, Engle and Sarkar (2006) examine 16 ETFs on international indices (data up to 2000), and report that their average price deviations are much larger than the average for domestic ETFs (0.35 percent vs 0.01 percent). The volatility of their price deviations is also larger. More importantly, Engle and Sarkar show that international ETFs have price deviations that are more persistent than those of domestic ETFs (several days vs several minutes). Using data from a slightly longer period (up to 2002), Delcoure and Zhong (2007) calculate the price deviations of 20 iShares ETFs, each of which tracks an MSCI country-specific index. In their calculations, they attempt to remedy the non-synchronous NAV problem by adjusting the NAVs using the methods proposed by Goetzmann *et al.* (2001) and Engle and Sarkar (2006). Still, they report that iShares generally trade at economically significant premiums between 10 and 50 percent of the time. The premiums, however, are not persistent and disappear within two days. Using data from a later period (i.e. 2002-2005), Ackert and Tian (2008) compare 21 international ETFs to seven domestic ETFs and obtain qualitatively similar results. They also find that, among the 21 international ETFs, those that track emerging-market indices have larger median price deviations with greater volatility than those that track developed-market indices. Moreover, Ackert and Tian report statistically significant first-order autocorrelations of price deviations (an average of 0.20 for developed-market ETFs and 0.41 for emerging-market ETFs). They argue that non-synchronous trading hours can only explain part of the autocorrelations as some ETFs with large autocorrelations are those that track indices in countries in the same time zone as the USA.

The results of the above studies are due mainly to the staleness of the NAVs used in the calculation. They do not directly address the issue of pricing efficiency of international ETFs. This brings us to the third main point of the findings of existing studies, i.e. there is mixed evidence on the pricing efficiency of international ETFs. For example, Madura and Richie (2004) report that, compared to broad-based ETFs, international ETFs are much more prone to extreme price movements, a significant portion of which is reversed in the subsequent trading session (e.g. if an extreme return occurs during normal trading hours, part of it is reversed during after-hour trading). Madura and Richie's results suggest the existence of overreaction, which then leads to return predictability in a subsequent period. On the other hand, Tse and Martinez (2007) compare the volatility of daytime returns to the volatility of overnight returns of 24 international iShares ETFs[8]. They find that the overnight variances are larger than the daytime variances for the iShares that track Asian and European markets, while the opposite is true for iShares that track American markets (i.e. Canada, Mexico and Brazil). They conclude that, because the Asian markets (and to a lesser extent, the European markets) have non-overlapping trading hours with the US markets while the American markets share the majority of trading hours with the US markets, their results suggest that volatility is driven by the release of public information in the foreign markets, rather than by noise trading activity or private information during the US market trading hours. Tse and Martinez also demonstrate that the pricing of international ETFs is efficient, as their prices reflect fundamental information from the underlying indices.

However, a recent study by Levy and Lieberman (2012) again casts doubt on the pricing efficiency of international ETFs. Using intraday data, Levy and Lieberman examine the pricing of 17 international ETFs (nine Asian and eight European) during overlapping trading hours vs non-overlapping trading hours (in a given day in the US market). Their goal is to find out whether there is a difference in the price formation process of these ETFs between the two periods. They report that during the overlapping hours, the NAV returns (i.e. the returns in the foreign markets), rather than the US market returns (as proxied by the S&P 500 index returns), have the biggest influence on the ETF returns. However, when the foreign markets are closed, the US market returns account for a large part of the ETF returns. They then show that this finding is consistent with the hypothesis that during the non-overlapping trading hours, traders overreact to US market sentiment.

Finally, we discuss the pricing efficiency of leveraged ETFs, which are a new but very popular type of ETFs. The goal of these leveraged ETFs is to generate daily returns that are in a positive or a negative multiple of the daily returns on an underlying index. Currently, the available multiples are $+2\times$, $+3\times$, $-1\times$, $-2\times$ and $-3\times$. In order to generate the promised returns, the funds use leverage, which is typically obtained through derivatives such as futures contracts, forward contracts and total-return swaps. As a result, these funds generally do not hold the constituent stocks of the underlying indices. This has an implication on their creation/redemption process (and thus the arbitrage activity) because they have to use an in-cash process rather than an in-kind process. Charupat and Miu (2012) measure premiums/discounts of 15 leveraged ETFs that are based on the S&P 500, Nasdaq 100 and Russell 2000 indices. They find that the price deviations, on average, are small and within the range of transaction costs and bid-ask spreads. In addition, they report that the price deviations follow particular patterns. For example, bull funds (i.e. those with positive multiples) trade at a discount more often than bear funds (i.e. those with negative multiples). Also, price deviations of bull (bear) funds are negatively (positively) correlated with the returns on their own underlying index. Charupat and Miu attribute this behavior partly to the funds' daily exposure adjustments, which have to be done at the end of each trading day in order to maintain their leverage ratios.

3. Tracking ability and performance

In the literature on ETFs, tracking error is typically defined as the deviation of the return on the NAV of an ETF from the corresponding return on its underlying benchmark index. Unlike price deviations, which are typically expected to be within the arbitrage bounds given the creation/redemption process of ETFs (as examined above), any deviations of the returns on NAV from those of their underlying benchmarks could accumulate over time and thus significantly affect the long-term performance of the ETFs. In Section 3.1, we first examine the factors (either within or outside the control of the ETFs' issuers) that dictate the tracking errors and performance of ETFs. We then introduce the commonly adopted measures of tracking errors in the literature in Section 3.2. Finally, in Section 3.3, we highlight the findings of a number of empirical studies on the tracking errors and performance of ETFs.

3.1 Factors dictating the tracking errors and performance of ETFs

Management fees. Management fees represent the explicit costs of managing the ETFs. They are typically charged on a daily basis by the fund issuers at annualized rates, usually referred to as the "expense ratios", which are based on the average daily NAVs. One of the selling points of ETFs is their low management fees as compared with many open-end index mutual funds. Unlike mutual funds, ETFs are traded on exchanges and thus the issuers do not need to provide transfer agency services and other services to unit holders. The expense ratios of those ETFs tracking broadly diversified domestic equity indices are typically very low (e.g. the net expense ratio of SPDR is 0.09 percent per annum in 2011), while those tracking domestic sector and international indices tend to have higher expense ratios (BlackRock, 2012)[9]. Ceteris paribus, the higher the expense ratio, the larger is the tracking error and the more an ETF is expected to underperform the underlying index. Nevertheless, management fees are not the only sources of tracking errors.

Transaction costs. The transaction costs incurred in buying and selling of the AUM are not reflected in the fund's expense ratio. Note that these are not the transaction costs (e.g. commissions and brokerage fees) paid by investors of ETFs in buying and selling units of ETFs on the exchanges. ETFs using the in-kind creation/redemption process tend to incur lower transaction costs than open-end index mutual funds, which need to purchase (liquidate) the securities underlying the benchmark indices when there is a net positive (negative) fund flow. However, similar to index mutual funds, in order to ensure their portfolios of the constituent securities can mimic those of the benchmarks, ETFs are subject to transaction costs associated with changes in the indices' compositions. Transaction costs are generally higher for ETFs tracking indices that are more volatile and with underlying assets that are less liquid. Moreover, it is also specific to the design and management of the benchmark index. According to Gastineau (2002), the equity index funds tracking the Russell 2000 index (and, to a lesser extent, those tracking the S&P 500 index) are significantly affected by the changes in the index composition. Gastineau estimates that the annual changes in the Russell 2000 index composition may lead to a 25-30 percent annual turnover of stocks in rebalancing the portfolios, and thus resulting in an annual transaction cost of 2-3 percent. Besides the revision of an equity index composition, a number of corporate actions of the constituent companies may also lead to the requirement of portfolio rebalancing of an index tracking fund and thus resulting in transaction costs for the fund. For example, Frino *et al.* (2004) find that tracking errors of index funds on the S&P 500 index can be influenced by the amount of share issuances/repurchases and spin-offs of the constituent companies through their effects on the Index Divisor in the computation of the S&P 500 index value.

Dividends and cash holding. For the ETFs (e.g. S&P SPDR, NASDAQ QQQ, and Dow Jones DIA) that are organized as unit investment trusts (as opposed to the open-end mutual fund structure), the fact that dividends from the constituent stocks are accrued in a cash account before they are distributed by the fund issuers to the unit holders (typically on a quarterly basis) can also result in tracking errors. The delay in the reinvestment of the dividends can result in a "drag" on the performance of the ETF. The higher the dividend yield, the longer the time delay, the higher the underlying index return, the more negative is the impact on the ETF returns. The impact of the delay in the reinvestment of dividends has been documented in the literature on ETFs and index tracking funds. For example, Elton *et al.* (2002) show that the tracking errors of SPDR can be explained by the accrual of its dividends in a non-interest-bearing account. Frino *et al.* (2004) verify this effect with a sample of S&P 500 index funds.

Choice of replication strategy. The specific replication technique adopted by an ETF can also affect its tracking errors. There is typically a trade-off between tracking errors and performance in the choice of a suitable replication strategy. A *full-replication* (or *exact-replication*) strategy, which involves the holding of all the constituent securities in the specific weights, can ensure the minimum tracking errors. It, however, may result in an unacceptable level of transaction costs since any changes in the index composition will involve the trading of all the constituent securities. On the other hand, by holding only a *representative* subset of the constituent securities, an ETF may reduce the amount of transaction costs. It, however, will be subject to larger tracking errors. In managing the funds, some issuers therefore choose to adopt various forms of optimization (non-replication) strategies (Canakgoz and Beasley, 2008; Corielli and Marcellino, 2006). The use of optimization strategies typically results in higher tracking errors than full-replication strategies. While the replications are generally quite accurate, the chance of any mismatch is expected to be higher around the time of a change in the index composition. Besides, the return volatility of the underlying index can amplify any replication errors and thus the tracking errors of the fund. Rompotis (2006) documents a positive relation between tracking errors and the return volatility in an empirical study on iShares ETFs.

Compounding effect of leveraged and inverse ETFs. In addition to the four factors mentioned above, the compounding of daily returns can also result in tracking errors for leveraged and inverse ETFs. Given the objective to maintain the target leverage ratios in generating the daily returns, these ETFs need to adjust their exposures to the benchmark indices at the end of each trading day. This daily rebalancing of exposure results in tracking errors in the compounded returns over holding periods that are longer than one day. Because of this compounding effect, the holding-period returns on leveraged or inverse ETFs will not be equal to the returns on the index multiplied by the target leverage ratio even if there is no tracking error on the daily return (see the illustrations by Cheng and Madhavan (2009) for details). The longer the holding period and the more volatile the return, the stronger the compounding effect on the tracking errors of leveraged and inverse ETFs. Moreover, the effect is stronger for bear funds than bull funds.

3.2 Measures of tracking errors

Tracking errors of ETFs can be measured in a similar fashion as that of mutual funds (Roll, 1992; Pope and Yadav, 1994). The commonly adopted tracking error measures are:

- the average absolute difference between the return on the fund and that of the underlying benchmark index (TE_1);
- the root-mean-square deviation of the return on the fund from that of the benchmark (TE_2);

- the standard deviation of the difference between the return on the fund and that of the benchmark (TE_3); and
- the standard error of the regression of the returns on the funds on that of the benchmark (TE_4).

Specifically:

$$TE_1 = \frac{1}{T}\sum_{t=1}^{T}\left|r_t^{ETF} - r_t^{I}\right| \tag{1}$$

where r_t^{ETF} and r_t^{I} are, respectively, the return of the ETF and the underlying benchmark index on day t, and T is the length of the time period under consideration. The measures based on the root-mean-square deviation (TE_2) and the standard deviation (TE_3) can be expressed as:

$$TE_2 = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(r_t^{ETF} - r_t^{I}\right)^2} \tag{2}$$

$$TE_3 = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left[\left(r_t^{ETF} - r_t^{I}\right) - \left(\bar{r}^{ETF} - \bar{r}^{I}\right)\right]^2} \tag{3}$$

where $\bar{r}^{ETF}$ and $\bar{r}^{I}$are, respectively, the sample mean returns on the ETF and the underlying benchmark index over the time period under consideration. Note that, unlike TE_1 and TE_2, TE_3 could be zero even if the ETF consistently over- or underperforms the index. Finally, TE_4 can be estimated as the standard deviation of the residuals (ε_t) of the following regression:

$$r_t^{ETF} = \alpha + \beta \cdot r_t^{I} + \varepsilon_t \tag{4}$$

Moreover, a positive (negative) estimated value of the intercept (α) will suggest the ETF outperforms (underperforms) the underlying index. Note that, TE_3 will be identical to TE_4 if the estimated value of the slope coefficient (β) equals to unity. On the other hand, TE_2 will be identical to TE_4 if the estimated values of the intercept (α) and the slope coefficient (β) equal to zero and unity, respectively.

3.3 Empirical findings on tracking errors and performance

There are a few general themes from the findings of the existing empirical studies. Despite their generally low expense ratios and *tax-friendly* design (as discussed below), ETFs do not necessarily outperform their index mutual fund competitors even on an after-tax basis. There is empirical evidence suggesting that ETFs and index mutual funds are substitutes of each other (Agapova, 2010). Nevertheless, they are far from perfect substitutes and they appeal to different types of investors according to their investment strategies/philosophies, the length of their investment horizon, their tax rates, and the transaction costs they are subject to.

Because of its popularity among investors, SPDR has attracted a lot of attention from researchers starting from the early 2000s. Elton *et al.* (2002) find that SPDR underperforms the S&P 500 index by an average of 28 basis points (bps) per annum from 1993 to 1998. Otherwise, tracking errors are found to be very small. They attribute the 28 bps underperformance to the expense ratio of SPDR and the delay in the reinvestment of dividends (from the constituent stocks) that are held in a non-interest-bearing account before their distributions. Over the same time period, SPDR also underperforms low-cost index mutual funds by an average of 18 bps, which Elton *et al.* interpret as the *value of immediacy* (since only the former, but not the latter, can be purchased and sold at any time during the trading day). In their study on SPDR, Poterba and Shoven (2002) highlight the tax-friendliness of ETFs. The in-kind creation/redemption design allows ETFs to substantially reduce (or even completely avoid) the distribution of realized capital gains, to which buy-and-hold investors of an open-end index mutual are subject as a result of the fund manager selling appreciated shares. Poterba and Shoven find that, from 1993 to 2000, the average capital-gain distribution on SPDR is 3 bps as opposed to the 48 bps distribution from the Vanguard Index 500 fund. Nevertheless, based on plausible ordinary-income tax rates, the tax advantage is not sufficient to compensate for the lower pre-tax performance of SPDR. In an after-tax basis, SPDR is still outperformed by the Vanguard Index 500 fund.

The lower pre-tax performance of SPDR compared to their low-cost index mutual fund competitors may be attributed to its inflexible replication strategy in the pursuit of the smallest tracking errors (Gastineau, 2004; Blume and Edelen, 2002). Given the imbalance of the demand/supply subsequent to the announcement of the change in the index composition, those stocks of low floating supply that are added to (removed from) the index experience abnormally high negative (positive) return after they are added (removed). In mimicking the portfolio of the underlying index up to the date when the index composition change becomes effective, SPDR is subject to additional transaction costs due to this price pressure, which translates into a lower performance compared with some of their mutual fund competitors that adopt a more aggressive approach by modifying their portfolios upon (or even prior to) the official announcement of the change in index composition. The analysis performed by Blume and Edelen (2002) on the S&P 500 index from 1995 to 2000 suggests that this more aggressive approach can result in an enhancement of the annual return by about 26 bps. The resulting tracking errors are, however, expected to be much larger than those of the more conservative replication strategy adopted by SPDR.

The above performance comparison has ignored the transaction costs encountered by ETF investors when they trade on the exchanges. Similar to the trading of any stocks listed on the exchanges, ETF investors are subject to bid-ask spreads, brokerage fees and commissions. The amount and significance of these transaction costs are specific to each type of investors (e.g. retail vs institutional) and their trading strategy and horizon (e.g. short- vs long-term investors). Investors of index mutual funds are not subject to these transaction costs. Together with the difference in tax efficiency as mentioned above, which is also investor-specific, the choice between ETFs and conventional index mutual funds is far from a straight-forward task. Actually, the coexistence of both instruments may be explained by a clientele effect (Agapova, 2010; Dellva, 2001; Guedj and Huang, 2009). Specifically, ETFs are generally considered to be a more efficient investment product for:

- long-term or institutional investors who are subject to lower trading costs; and
- tax-sensitive investors.

On the other hand, index mutual funds are more suitable for:

- short-term or retail investors who are subject to higher trading costs; and
- investors with tax-exempt or tax-deferred retirement accounts.

Based on the same line of argument, Kostovetsky (2003) develops a couple of models to quantify the impact of investor trading preferences, tax implications, and other characteristics in dictating the choice between these two types of instruments.

Most of the empirical studies on ETF's tracking errors and performance focus on ETFs tracking the domestic and international equity market indices. In addition, most of them conduct comparisons with conventional index mutual funds. For example, Rompotis (2005) investigates 16 matching pairs of ETFs and index funds tracking a number of domestic and international equity market indices from 2001 to 2002 and concludes that the two types of instruments have essentially the same degree of tracking errors. Aber *et al.* (2009) consider four equity index tracking ETFs with data covering up to the end of 2006. The results present a mixed picture regarding the relative performance between ETFs and index mutual funds, although the mutual funds tend to beat their ETF counterparts in terms of tracking errors. Among other things, Rompotis (2006) examines the tracking errors of 73 domestic, international, and sector iShares in 2005-2006 and documents significant tracking errors, especially for those tracking international indices. In examining 42 equity ETFs tracking different industry sectors from 2006 to 2008, Qadan and Yagil (2012) find that tracking error is positively correlated with the daily volatility of the ETF. One of the most comprehensive empirical studies in terms of its coverage of ETFs was conducted by Svetina (2010). He examines a total of 584 ETFs tracking domestic/international equity, domestic/international fixed income, real estates and other indices up to the end of 2007. The performances of the ETFs are found to be statistically indistinguishable from their corresponding competing index mutual funds.

There has been an increased interest in empirical studies on US-listed (and, to a lesser extent, European-listed) ETFs tracking foreign equity market indices. For example, both Johnson (2009) and Shin and Soydemir (2010) recently conduct tracking errors analysis for a number of iShares MSCI country ETFs. Based on a sample 20 iShares MSCI country ETFs over the period of 2004-2007, Shin and Soydemir (2010) find that tracking errors display persistence (especially for those ETFs tracking Asian markets) and that the change in the exchange rate is a significant source of tracking errors. Although these two studies cover several emerging equity markets, most of the foreign countries being considered are developed markets. Blitz and Huij (2012) focus their attention on seven US- and European-listed ETFs that track the conventional broad emerging markets indices, namely MSCI EM and S&P EM BMI. Using data from the inceptions of these global emerging market ETFs up to end of 2010, they find that these ETFs exhibit much higher levels of tracking errors than those tracking developed foreign market indices. Moreover, the ETFs that do not adopt a full-replication strategy are more prone to high tracking errors during periods of high cross-sectional dispersion in stock returns.

There has also been a growing literature on the performance and tracking errors of ETFs listed on exchanges outside the US and Europe (e.g. Gallagher and Segara (2006), study ETFs in Australia; Lin and Chou (2006), examine the first ETF in Taiwan; Prasanna (2012), studies the ETFs in India). These ETFs typically track the respective local equity market indices. In some of these countries, the development of ETFs is still in its early stage and thus they have not yet been embraced to the same extent as in the US and Europe.

There is also a small but growing literature on the performance and tracking errors of leveraged and inverse ETFs. Most of the empirical studies focus on the impact of the compounding effect which is specific to these types of ETFs and cover funds tracking the equity market indices. The general conclusion is that the compounding effect plays an important (if not the most important) role in explaining the deviation from the target leverage returns over holding periods that are longer than a few months. Lu *et al.* (2009) examine the performance of four pairs of bull and bear funds tracking, respectively, the Dow Jones Industrial Average, the S&P 500 Index, the Nasdaq 100 Index, and the Russell 2000 Index, using return data up to the end of 2008. They find that tracking errors increase with the length of the holding period. Bull (bear) funds have difficulty tracking the underlying benchmark indices at the respective stated leverage multiples for holding periods of one-year (one-quarter) or longer. A more recent study is conducted by Rompotis (2012) on 28 leveraged and 40 inverse Proshares funds tracking domestic/international equity markets, sectors, and commodities. Based on the results, Rompotis concludes that the funds under consideration may not meet their target leveraged returns even on a daily basis. Interestingly, he also documents a day-of-the-week effect where the deviations between the returns on the funds and the daily target returns tend to be lower on Wednesday than other days of the week[10]. While most of the studies on the performance of these funds focus on the impact of the compounding effect, Shum and Kang (2012) propose a way to disentangle the effects of compounding, management factors, and price deviation on the performance of leveraged and inverse ETFs. They find that, under certain circumstances, the impact of management factors on performance can in fact outweigh that of the compounding effect[11].

4. Effects on underlying securities

Many studies have examined the effect of the introduction of ETFs on their underlying indices. A major impetus for the research in this area is papers by Subrahmanyam (1991) and Gorton and Pennacchi (1993). Both studies create models to explain the attractiveness of "basket securities" (which can be index-linked securities or index futures contracts). The idea is that because composite securities are diversified (and thus have no firm-specific risk), liquidity traders face lower expected losses to insiders by trading in composite securities, rather than in individual securities that make up the composites. As a result, the models predict a migration of liquidity traders to composite securities, thus causing individual securities to be traded mainly by informed traders and thus to be less liquid and to have higher adverse-selection costs (in the form of wider bid-ask spreads). In addition, the models also imply a positive relationship between the degree of diversification of a composite security and its attractiveness to liquidity traders.

In contrast to the predictions of the above models, it is argued that the introduction of a basket security such as an ETF opens up another avenue for arbitrage. This argument is based on the work of Fremault (1991), who shows that stock index futures contracts allow arbitrageurs to establish positions that benefit from price discrepancies between the index futures and the constituent stocks. With the introduction of ETFs, arbitrageurs can now also conduct arbitrage using the ETFs. The arbitrage activity will increase the liquidity and the pricing efficiency of the individual stocks.

With conflicting predictions between the migration argument and the arbitrage activity argument, the effects of the introduction of ETFs, on the underlying stocks are indeed an empirical question. The empirical evidence so far appears to be mixed. There are studies that show that the liquidity of the underlying stocks improves, which is consistent with the effect of arbitrage activity. For example, Hegde and McDermott (2004) report that the introduction of the ETFs on the Dow Jones Industrial Average index and the Nasdaq 100 index result in an improvement in liquidity of the constituent stocks of both indices during the first 50 days of trading in both ETFs. Hegde and McDermott show that the improvement in liquidity comes from the decline in adverse selection costs. Similarly, Richie and Madura (2007) measure the impact of the introduction of the Nasdaq 100 ETF on the constituent stocks. They find that the liquidity of the constituent stocks increases, as evidenced by narrower spreads, higher trading volume and a higher number of trades. The liquidity improvement is more pronounced for the stocks that have lower weights in the index. The systematic risk of the stock also declines. Other studies that report liquidity improvement for the constituent stocks include Yu (2005) and De Winne *et al.* (2011).

On the other hand, there are studies that support the adverse-selection predictions of Subrahmanyam (1991) and Gorton and Pennacchi (1993). For example, Van Ness *et al.* (2004) find that after the introduction of the ETF on the Dow Jones Industrial Average, the bid-ask spreads of the index's 30 constituent stocks increase "relative to those of matching stocks"[12]. The results can be interpreted as being consistent with uninformed traders migrating to the ETF. However, Van Ness *et al.* do not find consistent changes in the adverse-selection components of the constituent stocks relative to the matching stocks. More recently, Hamm (2010) uses data on 32 sector ETFs and 31 broader-based ETFs to observe the migration activity of uninformed traders when sector ETFs (which are less diversified) are introduced and compares it to when broader-based ETFs are introduced. The results show that uninformed traders are more likely to switch to an ETF if it is more diversified, as they expect to encounter fewer informed traders in a more diversified ETF than in a less diversified ETF. Hamm's results also provide weak support for the hypothesis that stocks underlying a broader-based ETF end up having higher adverse-selection cost than stocks underlying a sector ETF do.

Finally, there is a small but growing literature on the effects of the end-of-day rebalancing of exposures of leveraged and inverse ETFs on the volatilities and prices of the underlying securities of the benchmark indices (Trainor, 2010; Charupat and Miu, 2012; Haryanto *et al.*, 2012). In order to achieve their objectives of delivering daily returns at constant (positive or negative) multiples of the underlying benchmarks, leveraged and inverse ETFs need to rebalance their exposures on a daily basis typically close to the end of the trading day. When there is positive (negative) return on an index in a particular day, the issuers or their swap counterparties of both the bull and bear funds tracking that index will have to increase (reduce) their exposures on the index by buying (selling) the underlying securities. The larger the magnitude of the daily return of the index and the larger the amount of AUM, the more the rebalancing demand (see the derivations by Cheng and Madhavan (2009) for details). Given that the trades related to this rebalance demand concentrate near the end of the trading day, they may affect the price volatility and even the prices of the underlying securities. The effect is expected to be more significant for those indices where the aggregated AUM of their tracking leveraged ETFs are substantial relative to the depth of the markets of the underlying securities[13].

The findings from the limited empirical research are however mixed. Trainor (2010) does not find any changes in the intraday patterns and levels of S&P 500 index volatility that are related to the growing leverage ETF market over time. Besides, he cannot detect any changes in the propensity of price momentum over the last 30 min of trading before and after the introduction of leveraged ETFs to the market. He concludes that leveraged ETFs do not have any substantial effect on the S&P 500 index. Using intraday data from 2006 to 2011, Haryanto *et al.* (2012), however, find that end-of-day rebalancing of leveraged and inverse equity tracking ETFs has a statistically significant impact on the end-of-day volatility of their sample of 346 US blue-chip stocks. Nevertheless, the effect is only considered to be economically significant for those trading days with large price swings (3 percent or higher). After controlling for the expected level of volatility, the rebalancing demand can still explain on average 31-60 percent of the end-of-day volatility on these days. In another recent study, Charupat and Miu (2012) show that the pricing deviation of leveraged and inverse ETFs can be partly explained by the end-of-day rebalancing of exposure. The effect is especially strong for those funds tracking the Russell 2000 index, which, among the indices under consideration, has the highest amount of aggregate exposure adjustments relative to the amount of daily trading in the index.

5. Conclusions

In this paper, we start with a brief review of the history of the growth of ETFs since their introduction to the market in early 1990s. We also review some of the key characteristics of ETFs, namely intraday trading, in-kind creation/redemption, and tax efficiency, that differentiate them from other similar index tracking instruments. We also briefly describe different types of ETFs, including the leveraged ETFs which have become one of the fastest growing types of ETFs.

We then focus our attention on three strands of research on ETFs, namely, their pricing efficiency, their tracking ability and performance, and their effects on underlying securities. We describe the key factors and the underlying economics of these three strands of research. We also summarize the key findings of the existing literature on each of these strands with an attempt to highlight the general theme of the key conclusions.

We expect to witness the continuous growth of the literature not only in these three strands of research but also in other related research areas as ETFs are being embraced by more and more investors (not only in North America and Europe, but also in other countries) as a viable low-cost investment vehicle. The diversity of the literature is also expected to grow with the innovations of the industry. To the extent that new products like exchange-traded notes, actively managed funds, leveraged and inverse ETFs become more popular among the investment community, they will naturally attract more attention from future researchers.

Notes

1. *Source*: BlackRock (2012). The numbers reported by BlackRock include other exchange-traded products which are different in structure and operation from ETFs. An example of these exchange-traded products is exchange-traded notes (ETNs), which are unsecured debt securities issued by banks to track certain market indices. Accordingly, ETNs are backed only by the credit of the issuing banks, and not by portfolios of securities. Currently, however, the size of the market for ETNs and other exchange-traded products is only a small fraction of the size of the ETF market.

2. The Index Participation Shares used futures contracts to synthetically create returns that mimic those of the S&P 500 index. Shortly after they started trading on the American Stock Exchange and the Philadelphia Stock Exchange, a federal court ruled that they were similar to futures contracts and, as a result, would have to be traded on futures exchanges. Their trading on the two stock exchanges was shut down, and they ceased to exist. For more details on the history of ETFs, see Gastineau (2010, Chapter 2).
3. As has been well documented, the liquidity of an ETF does not depend solely on the number of shares outstanding and the trading activity of that ETF. This is because market makers can create new shares by delivering the underlying portfolio to the issuer. However, because a creation of shares has to be done in a multiple of "creation unit", which typically is between 50,000 and 100,000 shares, this process may be of little benefit to retail investors who trade in small lots. For a discussion of ETF liquidity in practice, see Abner (2010, Chapters 9 and 10).
4. Select traders are commonly designated brokers and market makers. In the US market, these select traders are also known as authorized participants.
5. As will be discussed below, leveraged ETFs (in both the US and the European markets) typically use derivatives to generate their returns. This necessitates an in-cash creation/redemption process.
6. For example, some ETFs have a cut-off time for creation/redemption order submission at noon on a trading day. Since creation/redemption is done at the NAV, which is determined at the end of the day, there is a time difference during which price risk exists. The earlier the cut-off time, the larger the risk.
7. Ackert and Tian actually compare SPDR's prices to the S&P 500 index itself. In order to make a proper comparison, they adjust the prices of the SPDR by the amounts of dividends accumulated in the fund.
8. Tse and Martinez define daytime returns as: log(closing price on day t) $-$ log(opening price on day t), and overnight returns as: log(opening price on day t) $-$ log(closing price on day $t-1$).
9. In addition to the fees charged by the fund issuer, the expense ratio also covers other costs payable to different parties for providing services in the management of the ETF, e.g. administration fees, custodian fees, licensing fees, and trustees fees.
10. Wang (2009) finds that the daily tracking errors of leveraged and inverse ETFs are far from random. They are serially correlated and can be predicted by an autoregressive model. Wang also explores how this predictability can be exploited to enhance the trading profitability involving leveraged and inverse ETFs.
11. Most of the studies on the performance and tracking errors of leveraged and inverse ETFs examine US-listed funds. Charupat and Miu (2011) investigate the performance of a number of these funds listed on the Canadian stock exchange. They find that their sample of Canada-listed funds in general track their underlying benchmark indices better than some of the US-listed funds as documented in the literature.
12. Van Ness *et al.*'s results show that the bid-ask spreads of the 30 constituent stocks indeed decline after the introduction of the ETF. However, the bid-ask spreads of the matching stocks decline by a larger magnitude. As a result, the spreads of the constituent stocks increase relative to those of the matching stocks.
13. This effect has also attracted a lot of attention from the media (Keefe, 2009; Lauricella *et al.*, 2008; Zweig, 2009). It seems that market participants have mixed impressions regarding the significance of the impact.

References

Aber, J.W., Li, D. and Can, L. (2009), "Price volatility and tracking ability of ETFs", *Journal of Asset Management*, Vol. 10 No. 4, pp. 210-221.

Abner, D.J. (2010), *The ETF Handbook: How to Value and Trade Exchange Traded Funds*, Wiley, Hoboken, NJ.

Ackert, L.F. and Tian, Y.S. (2000), "Arbitrage and valuation in the market for Standard and Poor's depositary receipts", *Financial Management*, Vol. 29, pp. 71-87.

Ackert, L.F. and Tian, Y.S. (2008), "Arbitrage, liquidity, and the valuation of exchange traded funds", *Financial Markets, Institutions & Instruments*, Vol. 17, pp. 331-362.

Agapova, A. (2010), "Conventional mutual index funds versus exchange-traded funds", *Journal of Financial Markets*, Vol. 14, pp. 323-343.

BlackRock (2012), *The ETP Landscape Global Handbook Q4 2011*, BlackRock Advisors (UK) Limited, London.

Blitz, D. and Huij, J. (2012), "Evaluating the performance of global emerging markets equity exchange-traded funds", *Emerging Markets Review*, Vol. 13, pp. 149-158.

Blume, M.E. and Edelen, R.M. (2002), "On replicating the S&P 500 index", The Rodney L. White Center for Financial Research, Paper No. 08-02, The Wharton School, University of Pennsylvania, Philadelphia, PA.

Canakgoz, N.A. and Beasley, J.E. (2008), "Mixed-integer programming approaches for index tracking and enhanced indexation", *European Journal of Operation Research*, Vol. 196, pp. 384-399.

Charupat, N. and Miu, P. (2011), "The pricing and performance of leveraged exchange-traded funds", *Journal of Banking & Finance*, Vol. 35 No. 4, pp. 966-977.

Charupat, N. and Miu, P. (2012), "The pricing efficiency of leveraged exchange-traded funds: evidence from the US markets", *Journal of Financial Research* (in press).

Cheng, M. and Madhavan, A. (2009), "The dynamics of leveraged and inverse exchange-traded funds", *Journal of Investment Management*, Vol. 7, pp. 43-62.

Corielli, F. and Marcellino, M. (2006), "Factor based index tracking", *Journal of Banking & Finance*, Vol. 30, pp. 2215-2233.

Curcio, R.J., Lipka, J.M. and Thornton, J.H. Jr (2004), "Cubes and the individual investor", *Financial Services Review*, Vol. 13, pp. 123-138.

Delcoure, N. and Zhong, M. (2007), "On the premiums of iShares", *Journal of Empirical Finance*, Vol. 14, pp. 168-195.

Dellva, W.L. (2001), "Exchange-traded funds not for everyone", *Journal of Financial Planning*, April, pp. 110-124.

De Winne, R., Gresse, C. and Platten, I. (2011), "Liquidity and risk sharing benefits from the introduction of an ETF", working paper.

Elton, E.J., Gruber, M.J., Comer, G. and Li, K. (2002), "Spiders: where are the bugs?", *Journal of Business*, Vol. 75 No. 3, pp. 453-472.

Engle, R.F. and Sarkar, D. (2006), "Premiums-discounts and exchange traded funds", *Journal of Derivatives*, Vol. 13, pp. 27-45.

Fremault, A. (1991), "Stock index futures and index arbitrage in a rational expectations model", *Journal of Business*, Vol. 64, pp. 523-547.

Frino, A., Gallagher, D.R., Neubert, A.S. and Oetomo, T.N. (2004), "Index design and implications for index tracking", *Journal of Portfolio Management*, Vol. 30 No. 2, pp. 89-95.

Gallagher, D.R. and Segara, R. (2006), "The performance and trading characteristics of exchange-traded funds", *Journal of Investment Strategy*, Vol. 1 No. 2, pp. 49-60.

Gastineau, G.L. (2002), "Equity index funds have lost their way", *Journal of Portfolio Management*, Vol. 28 No. 2, pp. 55-64.

Gastineau, G.L. (2004), "The benchmark index ETF performance problem", *Journal of Portfolio Management*, Vol. 30 No. 2, pp. 96-103.

Gastineau, G.L. (2010), *The Exchange-Traded Funds Manual*, 2nd ed., Wiley, Hoboken, NJ.

Goetzmann, W.N., Ivkovic, Z. and Rowenhorst, K.G. (2001), "Day trading international mutual funds: evidence from closed-end funds", *Journal of Financial and Quantitative Analysis*, Vol. 36, pp. 287-239.

Gorton, G.B. and Pennacchi, G.G. (1993), "Security baskets and index-linked securities", *Journal of Business*, Vol. 66 No. 1, pp. 1-27.

Guedj, I. and Huang, J. (2009), "Are ETFs replacing index mutual funds?", working paper.

Hamm, S.J.W. (2010), "The effect of ETFs on stock liquidity", doctoral dissertation, The Wharton School, University of Pennsylvania, Philadelphia, PA.

Haryanto, E., Rodier, A., Shum, P. and Hejazi, W. (2012), "Intraday share price volatility and leveraged ETF rebalancing", working paper.

Hegde, S.P. and McDermott, J.B. (2004), "The market liquidity of DIAMONDS, Q's and their underlying stocks", *Journal of Banking & Finance*, Vol. 28, pp. 1043-1067.

Johnson, W.F. (2009), "Tracking errors of exchange traded funds", *Journal of Asset Management*, Vol. 10 No. 4, pp. 253-262.

Keefe, J. (2009), "Strange end-of-day phenomenon sets pundits thinking", *Financial Times*, 10 May.

Kostovetsky, L. (2003), "Index mutual funds and exchange-traded funds", *Journal of Portfolio Management*, Vol. 29 No. 4, pp. 80-92.

Lauricella, T., Pulliam, S. and Gullapalli, D. (2008), "Are ETFs driving late-day turns?", *Wall Street Journal*, 15 December.

Levy, A. and Lieberman, O. (2012), "Overreaction of country ETFs to US market returns: intraday vs daily horizons and the role of synchronized trading", *Journal of Banking & Finance* (in press).

Lin, A. and Chou, A. (2006), "The tracking error and premium/discount of Taiwan's first exchange traded fund", *Web Journal of Chinese Management Review*, Vol. 9 No. 3, pp. 1-21.

Lu, L., Wang, J. and Zhang, G. (2009), "Long term performance of leveraged ETFs", working paper.

Madura, J. and Richie, N. (2004), "Overreaction of exchange-traded funds during the bubble of 1998-2002", *Journal of Behavioral Finance*, Vol. 5 No. 2, pp. 91-104.

Pope, P.F. and Yadav, P.K. (1994), "Discovering errors in tracking error", *Journal of Portfolio Management*, Vol. 20 No. 2, pp. 27-32.

Poterba, J.M. and Shoven, J.B. (2002), "Exchange-traded funds: a new investment option for taxable investors", *American Economic Review*, Vol. 92 No. 2, pp. 422-427.

Prasanna, P.K. (2012), "Performance of exchange-traded funds in India", *International Journal of Business and Management*, Vol. 7 No. 23, pp. 122-143.

Qadan, M. and Yagil, J. (2012), "On the dynamics of tracking indices by exchange traded funds in the presence of high volatility", *Managerial Finance*, Vol. 38 No. 9, pp. 804-832.

Richie, N. and Madura, J. (2007), "Impact of the QQQ on liquidity and risk of the underlying stocks", *Quarterly Review of Economics and Finance*, Vol. 47, pp. 411-421.

Roll, R. (1992), "A mean/variance analysis of tracking error", *Journal of Portfolio Management*, Vol. 18 No. 4, pp. 15-24.

Rompotis, G.G. (2005), "An empirical comparing investigation on exchange traded funds and index funds performance", working paper.

Rompotis, G.G. (2006), "Evaluating the performance and the trading characteristics of iShares", working paper.

Rompotis, G.G. (2012), "A survey on leveraged and inverse exchange-traded funds", *Journal of Index Investing*, Vol. 2 No. 4, pp. 84-95.

Shin, S. and Soydemir, G. (2010), "Exchange-traded funds, persistence in tracking errors and information dissemination", *Journal of Multinational Financial Management*, Vol. 20, pp. 214-234.

Shum, P.M. and Kang, J. (2012), "Leveraged and inverse ETF performance during the financial crisis", working paper.

Subrahmanyam, A. (1991), "A theory of trading in stock index futures", *Review of Financial Studies*, Vol. 4 No. 1, pp. 17-51.

Svetina, M. (2010), "Exchange traded funds: performance and competition", *Journal of Applied Finance*, Vol. 20 No. 2, pp. 130-145.

Trainor, W.J. Jr (2010), "Do leveraged ETFs increase volatility", *Technology and Investment*, Vol. 1, pp. 215-220.

Tse, Y. and Martinez, V. (2007), "Price discovery and informational efficiency of international iShares funds", *Global Finance Journal*, Vol. 18, pp. 1-15.

Van Ness, B.F., Van Ness, R.A. and Warr, R.S. (2004), "The impact of the introduction of index securities on the underlying stocks: the case of the Diamonds and the Dow 30", in Lee, C.F. (Ed.), *Advances in Quantitative Analysis of Finance and Accounting*, World Scientific Publishing, Singapore.

Wang, Z. (2009), "Market efficiency of leveraged ETFs", working paper, Federal Reserve Bank of New York, New York, NY.

Yu, L. (2005), "Basket securities, price formation and informational efficiency", working paper, Notre Dame University, South Bend, IN.

Zweig, J. (2009), "Will leveraged ETFs put cracks in market close?", *The Wall Street Journal*, 18 April.

Corresponding author
Peter Miu can be contacted at: miupete@mcmaster.ca