Equity market momentum: A synthesis of the literature and suggestions for future work

Avanidhar Subrahmanyam

Anderson School, UCLA, Los Angeles, CA 90095-1481, United States

ABSTRACT

I review the literature on equity market momentum, a seminal and intriguing finding in finance. This phenomenon is the ability of returns over the past one to four quarters to predict future returns over the same period in the cross-section of equities. I am able to document about ten different theories for momentum, and a large volume of empirical work on the topic. I find, however, that after a quarter century following the discovery of momentum by Jegadeesh and Titman (1993), we are still no closer to finding a discernible cause for this phenomenon, in spite of the extensive work on the topic. More needs to be done to develop tests that are focused not so much on testing one specific theory, but on ruling out alternative explanations.

Introduction

One of the key lines of research in finance is to understand the time series and cross-sectional behavior of equity market returns. Perhaps one of the most robust findings in this area is stock market momentum, which is the tendency of stocks' relative performance over the next three to twelve months to be predictable from their relative performance in the past three to twelve months. Following Jegadeesh and Titman's (1993) discovery of the momentum phenomenon, copious amounts of theoretical and empirical research have been devoted to try and understand this pattern in stock returns. The focus of this review article is to consider alternative explanations for momentum that have been proposed in the literature and to evaluate the progress that has been made in terms of ascertaining causes for this phenomenon.

The phenomenon of momentum is intriguing because it directly contradicts the notion of a fairly weak form of capital market efficiency; that markets do not have memory with respect to past prices. That such a straightforward source of paper profits presents itself in markets is naturally something that has attracted the interest of several scholars. Indeed, I was able to document several potential explanations for momentum considered in the literature. The rationales proposed are motivated principally in one of three ways. These are:

- Pure underreaction to information
- Continuing overreaction to information
- Required (expected) returns that vary with past returns

In addition to theoretical work, a large number of empirical studies analyze facets of momentum in domestic and international equities, as well as other asset classes.

In reviewing this literature I maintain as a central theme that our learning about momentum is hampered when so many explanations accumulate without any attempt to test for mutual exclusivity of the predictions. I argue that effort should be focused on ruling out alternative explanations for momentum and trying to hone in on the “true” explanation(s) rather than allowing the finding
to get “over-identified” via multiple stories for the same phenomenon.

My choice of papers to review is admittedly subjective. Nonetheless I believe the paper takes first things first and can potentially be a useful starting point to focus future research on momentum. This paper is organized as follows. Section 2 reviews the basic findings. Section 3 summarizes the prevailing explanations for momentum. Section 4 considers empirical work. Section 5 provides brief concluding remarks.

The basic findings

Jegadeesh and Titman (1993) demonstrate a momentum effect (prediction from three to twelve months of past returns). The basic finding is “winners” (the top decile of performers) over the past three to twelve months continue to outperform “losers” (the bottom decile) over the next three to twelve months as well. The phenomenon is documented by doing the following. First, sort stocks each month into deciles based on performance over the past J months (with J ranging from three to 12 months). Then form overlapping portfolios that hold these stocks for the next K months (the range of K is identical to that of J). Thus, in each month t, the position taken in month t-K is closed out. The documented profits are about 12% per year when J and K range between six and 12 months and are thus substantial. This basic momentum phenomenon also obtains in monthly cross-sectional Fama-MacBeth regressions for gross returns ranging from the past three to 12 months as explanatory variables (see, for example, Brennan et al., 1998). An interesting result documented by Jegadeesh and Titman (1993) is that momentum profits tend to reverse when the holding period K is raised to 24 months or higher. Fama and French (1996) as well as Pukthuanthong et al. (2018) find that factor models (such as those based on size and book/market, or on principal components extracted from covariance matrices) are not able to explain momentum profits.

In other work, Rouwenhorst (1998) finds out-of-sample evidence of a momentum effect in many European countries. Grundy and Martin (2001) show that momentum has both firm-specific and systematic-factor-related components. Asness et al. (2013) find that momentum effects are pervasive not only in international equities but also in markets for other assets such as government bonds, commodity futures, and foreign currencies. More recently, Chang et al. (2018) document a curious phenomenon in Japan: While there is no evidence of “standard” momentum, a version of momentum that uses only residual returns does obtain in Japan. This finding deserves further investigation, but overall, the evidence of momentum is robust and strong, in- and out-of-sample, and in cross-country settings.

Explanations for momentum

Why should such a strong return pattern such as momentum persist? A natural possibility is investor misreaction due to behavioral biases. Prominent behavioral attempts to explain the phenomenon are Daniel et al. (1998) (DHS), Barberis et al. (1998) (BSV), and Hong and Stein (1999) (HS). DHS attempts to explain patterns using overconfidence and self-attribution. Overconfidence about private signals causes overreaction. However, agents assess arriving public signals with a self-serving bias that overweight private signals and underweights public signals, indicating a continuing overreaction, followed by a slow correction. These latter phenomena lead to momentum and the long-term reversals accord with the findings of Jegadeesh and Titman (1993).

In BSV, extrapolation from random sequences, in which investors expect patterns in small samples to continue, results in overreaction and reversals in response to small-sample patterns, whereas conservatism, wherein agents attach insufficient weight to new information, creates momentum through underreaction. In HS there is momentum because a class of “news-watchers” get informed sequentially, and fail to condition on market prices. Further, momentum traders, who trade mechanically in the direction of past price changes, create overreaction because they continue trading even after news is fully incorporated into prices, followed by reversals when they close out their positions. Hong and Stein (2007) note that when newswatchers observe different signals but do not condition on the market price, they create volume, which is suggestive of the view that momentum should be strong when volume is high.

Brav and Heaton (2002) consider uncertainty about the economy’s parameters such as the asset value’s mean and rational Bayesian learning to explain predictable return patterns. Thus, if agents are unsure whether mean shifts have occurred and they have, there is underreaction and hence, momentum. Equally, there can be overreaction to recent data if agents believe a structural shift to occur and they do not.

Hong et al. (2007) suggest that investors use overly-simplified models to evaluate stocks, ignoring the true, more complex model. For example, agents may think stock prices are simple functions of some macro variables when they are not. An investor who believes in a particular model uses this model to make persistent forecast errors while ignoring a persistent but pertinent information signal; this leads to momentum. Further, an investor who naively uses an extrapolation model, can materially change beliefs after seeing a break in a sequence of positive earnings, leading to large price reactions.

Da et al. (2014) argue that momentum arises because investors underreact to information arriving in small bits much like the proverbial frog in a pan that underreacts as the water is slowly brought to boil. They show that stocks where past returns accumulate gradually exhibit more momentum that stocks where returns are accumulated in a lumpy fashion.

In yet another behavioral explanation for momentum, Grinblatt and Han (2005) argue that the disposition effect causes this pattern. Specifically, a reluctance to sell losers and an eagerness to sell winners (the disposition effect) both cause price to underreact to true fundamental news for losing stocks. They show that momentum is related to unrealized capital gains in their setting, as their model predicts.

Antoniou et al. (2013) argue that momentum arises because of cognitive dissonance. Investors react properly to news which confirms their beliefs but underreact to news that disconfirms their beliefs. On average, therefore, they underreact, which gives rise to
momentum. They show that momentum arises principally in optimistic periods because investors underreact to bad news. More recently, Luo et al. (2018) argue that momentum arises because overconfident investors are skeptical that other investors have received information, which causes underreaction to information. This paper takes the view that overconfidence takes a dual form: overestimating the precision of one’s own information (which causes reversals), and underestimating the precision of others’ information (which causes continuations).

Johnson (2002) provides a model of a rational momentum effect. The intuition is that the relation between equity prices and growth rates is highly non-linear, which means that growth rate risk rises with growth rates. Thus high past returns imply higher growth rate risk, which leads to higher required returns in the future, implying momentum. He argues that persistent technology shocks that affect growth rates do hit the economy, albeit infrequently, thus potentially causing momentum.

In another rational story for momentum, Holden and Subrahmanyam (2002) argue that momentum arises because uncertainty due to news resolves slowly, causing risk premia to decrease in an autocorrelated fashion. In their model, agents tend to become informed about a stock sequentially prior to the revelation of a public signal. As more and people get informed, risk premia persistently decline, which gives rise to momentum.

As can be seen from the above, we have at least ten explanations for momentum in the literature. It seems unlikely that the phenomenon can be driven by all of these rationales, and yet, as we see in the next section, we have not made much progress in understanding which, if any, of these explanations are more relevant as true drivers of the phenomenon.

Empirical issues surrounding momentum

Hong et al. (2000) show that momentum profits increase in size and analyst coverage. This supports the notion that neglected stocks have slower news diffusion and thus greater momentum (according with the HS story). Doukas and McKnight (2005) confirm that HLM’s results also apply in Europe. Cooper et al. (2004) show that momentum profits are larger after positive market returns than after negative ones and view this as confirming self-attribution bias (i.e., increasing returns when public information conveyed by the market is positive).

Lee and Swaminathan (2000) argue that trading volume helps shed light on momentum and reversals. They argue that stocks go through a life cycle wherein they are alternately desirable and undesirable. A high volume winner is in the desirable phase and is shown to have greater momentum. Further, a low volume loser is in the undesirable phase and also has greater momentum as it is expected to continue to be undesirable.

Hvidkjaer (2006) explores the link between momentum and small traders. He shows that small traders buy loser stocks for up to an year after these stocks are identified as losers, so that they delay selling them. In a variant on momentum, Heston and Sadka (2009) show that well-performing stocks during a month outperform poor performers in that particular month for more than fifteen years. This puzzling finding deserves attention in future research.

Jegadeesh and Titman (2001) show that momentum profits reverse in the long-term during the 1961–1985 period, but not in later years. The former finding is consistent with momentum being a delayed overreaction but why the delayed overreaction disappears in recent years is an open question. A recent empirical paper by Conrad and Yavuz (2017) provides evidence that the individual stocks that contribute to momentum profits do not experience long-run reversals. This evidence indicates that it is the underreaction channel rather than the delayed overreaction channel that drives the momentum effect. It is worth noting that the point estimates of the long-term return following momentum portfolio formation are negative in Conrad and Yavuz (2017), though not significant. This suggests that more power (from a longer time series) may help shed better light on momentum and reversals. More work also needs to be done to reconcile how the Conrad and Yavuz (2017) paper first in with earlier work by Lee and Swaminathan (2000), which documents long-run reversals as well as momentum.

Overall, the above work which indirectly or directly tests behavioral models still leaves several questions unanswered. First, it seems that the HS and DHS stories explain the exact same phenomenon but which, if any, of these stories are more consistent with the data? Is it possible to separate out the trend-chasers and the momentum traders and see if they actually trade in the manner suggested by the theory? Do traders exhibit self-attribution bias over the long horizons suggested by DHS? What is the role of conservatism and extrapolation in momentum? How does the “frog-in-the-pan” theory of Da et al. (2014) fit in with DHS, BSV, and HS, if at all? Can one test the cognitive dissonance setting of Antoniou et al. (2013) against the other models? The answers to these questions are clearly relevant, but as yet, have not received much attention in the literature.

In an interesting paper, Lewellen (2002) shows a striking result. He shows that portfolios sorted on criteria such as size and book/market ratios demonstrate momentum in a relative sense (relatively better performing portfolios continue to outperform relatively worse performing ones), but autocorrelations of these portfolios are negative at the horizons over which momentum obtains. He argues that this finding is inconsistent with underreaction, which should also lead to positive autocorrelations. He shows that the finding is consistent with “excess covariance,” wherein prices covary too much relative to the covariation in fundamentals. For example, if investors mistakenly believe that information about one firm has implications for another firm, then the excess positive covariation can lead to momentum, but also own-price corrections lead to negative autocorrelations. Lewellen’s (2002) findings have largely been ignored by researchers, but how they fit in with behavioral theories would be useful to know.

In another interesting finding, Grinblatt and Moskowitz (2004) demonstrate that return consistency is important for momentum profits. Returns accumulated gradually exhibit much more momentum than returns accumulated in a lumpy fashion. They attribute this finding to the disposition effect in that as stocks rise slowly, the reluctance to realize unrealized capital gains leads to underreaction and momentum (an analogous argument applies to losers). This finding is closely related to the “frog-in-the-pan” theory of Da et al. (2014). However, the latter authors argue that the effect of return consistency on momentum does not arise from the
disposition effect but from slow reaction to consistent, modest news. The argument most convincingly relies on the notion that analysts’ forecasts (with no direct link to the disposition effect) also underreact to stocks with slow accumulation of news.

Novy-Marx (2012) argues that momentum profits arise not because recent winners continue to outperform recent losers, but because good performers over the past seven to 12 months continue to outperform bad performers over that same period (so that momentum is really an “echo”). Goyal and Wahal (2015) demonstrate that no such “echo” obtains in an international setting comprising more than thirty countries. They show that the U.S.-based result of Novy-Marx (2012) arises primarily due to a reversal in the second month prior to portfolio formation, a finding which is itself an interesting puzzle.

There is some research on how momentum relates to fundamental quantities such as macroeconomic variables and earnings announcements. For example, Chordia and Shivakumar (2002) argue that momentum profits can be accounted for by the business cycle, in that they decrease significantly after adjusting returns for predictability based on macroeconomic variables. Griffin et al. (2003), however, find only modest support for the dependence of momentum profits on the business cycle in international markets, and find pervasive momentum across many countries.

Avramov et al. (2007) show that momentum profits are stronger in more distressed companies. It is not clear how this finding relates to the finding that momentum profits are sensitive to business cycles. Chan et al. (1996) show that while both earnings and past returns explain future momentum profits, past returns are not a substitute for past earnings. Chordia and Shivakumar (2006), however, show that earnings momentum substitutes for price momentum, in that the predictive power of past returns is proxied by the predictive power of past earnings surprises. A reconciliation between Chan et al. (1996) and Chordia and Shivakumar (2006) would seem to be desirable.

Fama and French (1996) show that momentum profits cannot be captured by a standard factor model that includes the market, as well as size and book/market-based factors (developed in Fama and French, 1992). The size and book/market factors are return differentials on portfolios with extreme values of size and book/market. Here size and book/market proxy for distress (that is, stocks with higher exposures to size and book/market factors are likely to get distressed and hence command bigger returns. However, Hong, Xue, and Zhang (2015) (HXZ) show that a “q-theoretic” factor model which includes factors based on real investment and return on equity (ROE) accounts for momentum profits. The rationale behind the investment factor is that companies that invest heavily have high net present value projects, implying lower discount rates and low returns. Similarly, profitable companies (that are expected to be profitable in the future), in equilibrium, choose more projects such that the discount rate is higher on incremental projects, implying higher average discount rates and higher expected returns. This latter idea could potentially explain profits if profitability tracks return performance. However, Novy-Marx (2015) argues that the HXZ results are simply due to the notion that transitory earnings surprises influence momentum profits and not due to general persistence in ROE. This is indicative that q-theory may not play a primary role in explaining momentum.

Sagi and Seasholes (2007) consider firm-specific variables that drive momentum. They find that momentum strategies for firms with volatility in revenue growth, lower costs, or material growth options outperform traditional momentum strategies by up to five percentage points per year. This suggests that momentum works better in growth firms. The argument given by Sagi and Seasholes (2007) for this is the following. Firms with good return performance tend to have strong growth options but these options are overrepresented in a firm’s market capitalization, so that such firms are more risky and thus earn higher expected returns. This argument seems reminiscent of Johnson (2002) and more work linking Sagi and Seasholes (2007) with Johnson (2002) would seem desirable.

There is also a literature linking momentum to trading costs. Lesmond et al. (2004) show that momentum strategies require a high degree of turnover. Further, stocks with high momentum also tend to be those with relatively higher trading costs. Accounting for these features makes momentum profits disappear. Korajczyk and Sadka (2004) after carefully estimating intraday price impacts of trades, and examining how these impact momentum strategies, also conclude that momentum profits do not survive transaction costs. Sadka (2006) shows that about half of the time-variation in momentum profits can be explained by the liquidity risk exposure of the momentum portfolios. That is, the sensitivity of the portfolio returns to an aggregate price impact factor are strongly priced in the cross-section. However, using actual data from a large money management firm, Frazzini et al. (2012) show that actual trading costs faced by big institutions may be lower than that suggested by the preceding studies, thus enhancing the potential profitability of momentum strategies.

Two recent papers indicate that as momentum has become more popular, and technology has resulted in trading cost declines, momentum profits have declined. McLean and Pontiff (2016) show that profits from a comprehensive set of more than ninety anomalies (including momentum) are more than halved after publication of the research documenting the anomaly in an academic journal. Chordia et al. (2014) (CST) show that momentum profits markedly diminish after the 2001 introduction of the practice of quoting stocks in decimals, as opposed to quoting in fractions such as 1/8 and 1/16. Decimalization is known to have lowered bid-ask spreads (Chordia et al., 2008), so that the CST finding accords with the view that momentum profits are sensitive to trading costs.

Israel and Moskowitz (2013) show using a long sample period from 1927 to 2011, however, that momentum profits do not exhibit a time trend. It is possible that use of a long sample period masks the effect of some factor. More work linking Sagi and Seasholes (2007) with Johnson (2002) would seem desirable.

In terms of other asset classes, Jostova et al. (2013) demonstrate momentum in corporate bonds. They also show that momentum profits are stronger in junk bonds. Gebhardt et al. (2005) show that momentum spills over from stocks to investment grade bonds. That is, they show the profitability of momentum strategies in investment grade bonds that are constructed based on past performance in the corresponding stocks (a finding confirmed by Chordia et al., 2017). Bali et al. (2017) show that momentum profits in corporate bonds mainly prevail for the high credit risk sector, and these authors also demonstrate that they are principally confined to the period corresponding to the recent financial crisis.

Beracha and Skiba (2011) demonstrate the profitability of momentum in the US residential real estate market. A strategy that
buys the winning markets and sells losing markets (amongst the 380 markets they consider) earns more than 8% per year. As pointed out in Section 2, Asness et al. (2013) demonstrate that momentum strategies are profitable not just for equities, but across diverse asset classes such as government bonds, currencies, and commodity futures markets. Hence evidence is accumulating that momentum is pervasive across many different categories of assets.

Daniel and Moskowitz (2016) show that while momentum strategies are mostly profitable over time, they fail during some specific periods (see also Barroso and Santa-Clara, 2015). These periods are characterized by those when the economy is recovering from a recession, when volatility is typically high. The phenomenon appears to be driven by losing stocks, which tend to fall sharply during recessions, and slowly recover. The consequent failure of the short-leg of momentum strategies is what leads to momentum crashes.

Moskowitz et al. (2012) uncover time-series momentum, where long-short portfolios are formed solely based on the performance of a stock relative to some benchmark (say zero), rather than in a relative fashion. They show (as do Asness et al., 2013, in a cross-sectional setting) that such strategies work in a variety of asset classes such as equity indices, foreign currencies, commodity futures, and government bonds. However, Goyal and Jegadeesh (2018) argue that such strategies are not market-neutral (unlike standard momentum strategies). This is because the number of stocks whose returns exceed the benchmark (say zero) is typically not equal to the number for which returns fall short of the benchmark. For example, in a bull market many stocks will do well. This means that a standard alpha calculated from a factor model is not readily interpretable as an abnormal return. Goyal and Jegadeesh (2018) accounting for market performance materially mitigates time-series momentum profits. More research that sheds light on whether time-series momentum is indeed a valid construct would be very welcome. As it stands, one is not sure what to make of this important alternative way of constructing momentum strategies.

Concluding remarks

Much work is needed to synthesize the literature on momentum. Does momentum arise due to the simple phenomenon of slow diffusion of news, or due to continuing overreaction, the “frog in the pan” phenomenon, or time varying expected returns? Are momentum profits sensitive to business cycles or not? Does earnings momentum proxy for price momentum? Does time-series momentum exist as a genuine phenomenon? Even more basically, is momentum a behavioral phenomenon arising from investor misreaction, or is it some sort of rational phenomenon? Since every paper tells its own story, I am unable to discern answers to these questions. So it is unclear, at least to me, how to interpret the current state of the literature on momentum.

One other intriguing issue is that there is nothing particularly special about the horizon three to twelve months. Why do momentum profits arise over this specific horizon all across the world? Analysis of this question would seem to be promising. A related issue is why markets exhibit alternating patterns of monthly reversals (Jegadeesh, 1990), quarterly to annual momentum, and three to five year reversals (Debondt and Thaler, 1985) that also form an intriguing problem for future research.

Overall, it seems to me we need more papers that are not overly descriptive nor those that are wedded to particular stories, but those that attempt to flesh out the “true” drivers of momentum by examining competing hypotheses. On the theoretical side, we probably need to take a pause on building theories for momentum while allowing more light to be shed on which, if any, of the existing theories receive more support in the data. I hope that we can make significant progress on the topic of momentum, which is of interest from an academic standpoint as well as from the perspective of the commercial finance industry.

References