Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Empirical Finance 11 (2004) 509–536
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/econbase

Momentum strategies: some bootstrap tests

G. Andrew Karolyi [a,*], Bong-Chan Kho [b]

[a] *Fisher College of Business, Ohio State University, 2100 Neil Avenue, Columbus OH 43210-1144, USA*
[b] *College of Business Administration, Seoul National University, South Korea*

Available online 8 April 2004

Abstract

This study introduces a new estimation-based bootstrap simulation procedure to test whether different returns-generating models can explain the profitability of momentum strategies first documented by Jegadeesh and Titman [J. Finance 48 (1993)]. We incorporate simple random walk and multifactor models and allow for autocorrelation, cross-correlation, conditional heteroskedasticity and predictability through conditioning information variables. We also evaluate alternative sampling procedures for the bootstrap simulations. None of the models, however, are able to generate simulated profits as large as the actual profits. We do find, however, that accounting for time-varying expected returns with market-wide and macroeconomic instrumental variables can explain 75–80% of the profits.

JEL classification: G12; G14
Keywords: Momentum investment strategies; Bootstrap tests; Asset-pricing models

1. Introduction

This study examines the role of time-varying expected returns in the profitability of momentum-based trading strategies, first documented by Jegadeesh and Titman (1993), using a new bootstrap simulation testing procedure. Momentum strategies have been particularly intriguing as they are very simple (buying stocks with high returns, or “winners,” and selling stocks with low returns, or “losers,” over the preceding three to twelve months) and as their profitability remains probably the most difficult CAPM-related anomaly unexplained by the three-factor model of Fama and French (1996).

* Corresponding author. Tel.: +1-614-292-0229; fax: +1-614-292-2418.
E-mail address: karolyi@cob.osu.edu (G.A. Karolyi).

0927-5398/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jempfin.2004.02.004

Recently, Jegadeesh and Titman (2001) confirm that momentum profits have continued in the 1990s, suggesting that their initial results were not due to data mining. Indeed, they suggest that the robustness of the momentum returns appears to be in conflict with standard asset-pricing models and may be driven by investor cognitive biases (Daniel et al., 1998; Barberis et al., 1998) or by investor underreaction to information, such as around earnings announcements (Chan et al., 1996). They point to evidence that favors their interpretation such as the concentration of momentum profits among stocks with low analyst coverage (Hong and Stein, 1999; Hong et al., 2000) and high trading volume (Lee and Swaminathan, 2000).

The momentum anomaly is not without its share of efficient-markets-based explanations, however. Conrad and Kaul (1998) and Berk et al. (1999) have suggested that high (low) realized returns will arise for those with high (low) expected returns, suggesting that the profitability of momentum strategies stems from cross-sectional differences in expected returns. Grundy and Martin (2001) and Wu (2002) counter that expected returns from a Fama–French model, even in a conditional form with time-varying risks and expected returns, cannot explain the phenomenon. Chordia and Shivakumar (2002) demonstrate that standard macroeconomic variables are able to predict momentum payoffs, but Griffin et al. (2003) show that these macroeconomic variables do not explain momentum profits uncovered in markets around the world (Rouwenhorst, 1998, 1999; Chan et al., 2000; Hameed and Kusnadi, 2002).

The goal of this paper is to devise a new test methodology to evaluate how much of the momentum profits are explainable by models of time-varying expected returns. We specifically introduce an estimation-based bootstrap simulation procedure to assess the ability of different returns-generating models that allow for time-varying expected returns, factor risks and a rich variety of cross-sectional/time-series error structures to generate momentum profits that are at least as large as those realized over the past 40 years in the US. In our experiments, we examine simple random walk models and multifactor models (Fama–French three-factor), models with and without conditioning instrumental variables, with and without autocorrelated, cross-autocorrelated and conditionally heteroskedastic errors. We also employ different resampling techniques (with and without replacement) and a new decomposition analysis in the application of the bootstrap approach. Overall, the results indicate that none of these models are able to generate simulated momentum profits as large as the actual profits over the 1964–2000 sample period. However, accounting for time-varying expected returns with market-wide and macroeconomic instrumental variables in our new decomposition analysis can explain up to 75% or 80% of the winner–loser spread returns. This "cup-is-half-empty" finding clearly lends support to those who advocate behavioral explanations of the phenomenon, but we offer those who propose efficient-markets-based explanations a "cup-is-half-full" result with hope for conventional models of time-varying expected returns.

In addition to this new finding, the study makes two important contributions to the existing research on the momentum phenomenon. First, it is the most comprehensive application to momentum profits of a bootstrap simulation approach to date. Conrad and Kaul (1998, Section 3) were the first to publish a simulation experiment of the profitability of momentum returns.[1] Their goal was to show that the main determinant of the profits stemmed from the cross-sectional variation in mean returns. As a result, they employed only a random walk model with drift for the simulations. More importantly, they examine only one type of resampling approach—bootstrapping with replacement—which is the basis of criticism for the follow-up critique by Jegadeesh and Titman (2002, Section 3). They argue for (and evaluate only) a bootstrap procedure that resamples without replacement using a random walk model with drift. They argue for without-replacement simulations on account of the dangers of small sample bias. Our study sheds light on this debate by encompassing both resampling techniques, their advantages and disadvantages. We also examine the effectiveness of these techniques for a variety of returns-generating models and residual error structures.

Our second important contribution is that we provide a methodological framework within which to understand the recent debate over the role of conditioning information in various asset-pricing models that have been proposed to capture return momentum. Chordia and Shivakumar (2002) provide empirical evidence that links momentum profits to business cycles and macroeconomic instrumental variables and show that momentum is explainable by business cycle risk. Wu (2002) and Griffin et al. (2003) show that the conditioning information variables do not explain momentum; Wu employs the Fama–French three-factor model, Griffin et al. applies Chordia and Shivakumar's approach to 39 markets around the world. None of these studies employs a robust, bootstrap-simulation procedure and, furthermore, none provides a decomposition analysis that helps to determine the component of the total momentum returns that is due to market risks, to other extra-market factor risks and, especially, to the conditioning information variables they focus on. Our new parametric bootstrap methodology provides the bootstrap and the decomposition analysis.

The next section establishes the momentum phenomenon for the US over the past 30 years, which is the laboratory for our experiments. Section 3 details the new bootstrap methodology and discusses the advantages of alternative resampling procedures. We describe our asset-pricing models in Section 4, outlining the different error structures and a summary of estimation results. Section 5 provides the results from the bootstrap simulations including the decomposition analysis. Concluding remarks follow.

2. The momentum phenomenon

The momentum strategies that we consider are based on the recent past 3- to 12-month returns and the subsequent holding periods vary from 3 to 12 months. The sample includes a total of 9807 stocks available in the Center for Research in Security Prices (CRSP) database from the New York (NYSE), American (Amex) and Nasdaq stock exchanges

[1] They cite an unpublished working paper, Karolyi and Kho (1993), the predecessor of the current study by the same authors, as the first study to employ a bootstrap experiment and as their primary motivation. See their footnote 11, p. 506. Jegadeesh and Titman (2002) also cite this study, but as "Karolyi and Kho (1993)," with both names misspelled.

Table 1
Momentum portfolio returns

Panel A. Monthly returns on the momentum portfolios

Portfolios	1965–1989					1965–2000				
	J	*K* = 3	*K* = 6	*K* = 9	*K* = 12	*J*	*K* = 3	*K* = 6	*K* = 9	*K* = 12
Loser	3	1.16 (2.72)	0.91 (2.18)	0.86 (2.06)	0.82 (2.00)	3	1.28 (3.65)	0.92 (2.71)	0.85 (2.52)	0.81 (2.46)
Winner	3	1.83 (4.75)	1.75 (4.52)	1.74 (4.46)	1.71 (4.37)	3	2.06 (5.91)	1.92 (5.57)	1.88 (5.51)	1.80 (5.27)
Winner–loser	3	0.67 (2.98)	0.83 (4.19)	0.88 (5.02)	0.89 (5.76)	3	0.78 (3.42)	0.99 (5.06)	1.03 (6.17)	0.99 (6.69)
Loser	6	1.00 (2.26)	0.78 (1.82)	0.70 (1.66)	0.75 (1.81)	6	1.08 (3.00)	0.77 (2.21)	0.68 (2.03)	0.75 (2.25)
Winner	6	2.09 (5.25)	2.03 (5.06)	1.97 (4.91)	1.84 (4.60)	6	2.34 (6.39)	2.23 (6.15)	2.11 (5.90)	1.92 (5.42)
Winner–loser	6	1.09 (4.02)	1.25 (5.07)	1.28 (5.88)	1.09 (5.38)	6	1.27 (4.72)	1.46 (6.15)	1.43 (6.90)	1.16 (6.14)
Loser	9	0.92 (2.08)	0.68 (1.60)	0.68 (1.62)	0.76 (1.81)	9	1.01 (2.84)	0.70 (2.05)	0.71 (2.10)	0.80 (2.38)
Winner	9	2.24 (5.47)	2.14 (5.20)	1.98 (4.83)	1.83 (4.48)	9	2.47 (6.59)	2.29 (6.17)	2.06 (5.66)	1.85 (5.16)
Winner–loser	9	1.32 (4.62)	1.46 (5.65)	1.30 (5.42)	1.07 (4.78)	9	1.46 (5.34)	1.58 (6.46)	1.35 (6.02)	1.05 (5.10)
Loser	12	0.85 (1.95)	0.72 (1.68)	0.75 (1.75)	0.84 (1.97)	12	0.97 (2.81)	0.78 (2.29)	0.81 (2.39)	0.89 (2.64)
Winner	12	2.28 (5.48)	2.07 (4.99)	1.91 (4.62)	1.75 (4.26)	12	2.43 (6.40)	2.16 (5.77)	1.94 (5.29)	1.74 (4.82)
Winner–loser	12	1.44 (5.02)	1.35 (5.06)	1.17 (4.72)	0.92 (3.92)	12	1.46 (5.36)	1.38 (5.48)	1.13 (4.93)	0.85 (4.00)

Panel B. Monthly returns and other statistics on the 6-month/6-month momentum portfolio

Portfolios	1965–1989	1990–2000	1965–2000						
	Return (%)	Return (%)	Return (%)	Return Jan	Return non-Jan	Number of stocks	Price (US$)	Mkt. cap. (millions US$)	Size decile
P1 (loser)	0.78	0.74	0.77	6.46	0.25	282	15.8	347	3.8
P2	1.15	0.97	1.10	5.60	0.69	282	21.4	635	4.2
P3	1.24	1.17	1.22	4.88	0.88	282	26.7	809	4.4
P4	1.27	1.27	1.27	4.43	0.99	282	26.6	928	4.6
P5	1.30	1.27	1.29	4.08	1.04	282	30.6	1042	4.6
P6	1.39	1.29	1.36	3.84	1.13	282	31.8	1131	4.7
P7	1.40	1.37	1.39	3.62	1.18	282	33.3	1221	4.6
P8	1.51	1.55	1.52	3.67	1.33	282	32.0	1219	4.4
P9	1.63	1.83	1.69	3.76	1.50	281	32.8	1025	4.1
P10 (winner)	2.03	2.68	2.23	4.45	2.03	281	29.8	634	3.2
Winner–loser	1.25	1.95	1.46	−2.01	1.77	−1	13.9	287	−0.6
(*t* Statistic)	(5.07)	(3.61)	(6.15)	(−2.07)	(3.64)				
[% Positive]	[68.7]	[68.9]	[68.8]	[30.6]	[72.2]				

This table reports the monthly returns for momentum decile portfolios formed based on past *J*-month returns and held for subsequent *K* months. Loser (P1) is the equal-weighted portfolio of 10% of the stocks with the lowest past *J*-month returns, and winner (P10) is for those with the highest past *J*-month returns. The sample includes a total of 9807 stocks traded on the NYSE/AMEX/NASDAQ during the period of January 1963 through December 2000 (456 months) with at least 24 observations of monthly returns, excluding stocks priced less than US$5 at the beginning of the holding period and stocks in the smallest market cap decile (NYSE size decile cutoff). "Jan" and "non-Jan" denote average returns for only the month of January and only non-January months, respectively. "Mkt. cap." denotes average market capitalization and "Size decile" is the average decile rank by market capitalization for firms over the available sample period. *t* statistics of the monthly returns are in parentheses.

during the period of January 1963 to December 2000 (456 months). The stocks are excluded if they have fewer than 24 monthly returns available, if their traded price is below US$5 per share at the beginning of the holding period, and if they fall into the lowest market capitalization decile (based on the decile cutoffs defined by the NYSE sample).[2] At the end of each month, all eligible stocks are ranked in ascending order based on their returns in the previous J months and assigned to one of 10 decile groups. The strategy buys the decile 10 (winner) stocks and sells the decile 1 (loser) stocks, holding this position for the next K months. We define a strategy that evaluates returns over the past J months and holds the position for the next K months as the "J-month/K-month" strategy. The strategy is applied to overlapping J-month/K-month horizons. The portfolio returns are constructed as an equal-weighted average of the component stocks and rebalanced monthly, so that firms which drop out of the sample due to delisting, suspension, merger or other reasons cannot affect the returns beyond the month in which they drop out.[3]

Table 1, Panel A reports the average returns of the winner and loser portfolios as well as of the zero-cost, winner–loser spread portfolio for all combinations of 3-, 6-, 9- and 12-month J-month/K-month horizons. For each strategy, we report the average monthly return and its associated t statistic. We also report separately, the results for the full sample period (1965–2000) as well as Jegadeesh and Titman's (1993) original sample period (1965–1989). The results are generally very similar to Jegadeesh and Titman (1993, 2001). The highest average winner–loser return spread (1.58% per month) arises for the 9-month/6-month strategy and the lowest average arises for the 3-month/3-month strategy (0.78%). These relative rankings are similar for the original Jegadeesh and Titman sample period and for the updated sample period through 2000. The t statistics indicate that these mean return spreads are significant at any reasonable levels.

Panel B presents summary statistics for the firms within all deciles of the 6-month/6-month strategy. The monotonic pattern in momentum profits increasing from loser to winner deciles is clear in the Jegadeesh and Titman sample period as well as the full sample period. The sensitivity of the momentum returns to Januarys is also discernible (Jegadeesh and Titman, 2001, Table 2). Finally, the small market capitalization stocks are concentrated in the extreme loser and extreme winner deciles, as shown by Jegadeesh and Titman (2001, Table 3).[4]

[2] This decision to exclude the lowest decile of small market capitalization stocks follows Jegadeesh and Titman (2001). Their objective is to exclude stocks for which the execution of the strategy is too costly.

[3] Korajczyk and Sadka (in press) demonstrate that the profitability of momentum trading strategies can be significantly diminished if proper account is taken of the price impact induced by large trades and by the portfolio turnover activity required to sustain the strategy.

[4] It should be noted that for all of our analysis, the evaluation period includes the ranking month (e.g., J equals 6 months implies past returns for months -5 to 0) and the holding period follows the ranking month (e.g., K equals months $+1$ to $+6$). We replicated all of our analysis with a 1-month skip after the ranking month. In fact, the 3- to 6-month holding period return spreads were even larger with the skip-month approach and the 9- to 12-month holding period return spreads were smaller. This debate is featured in both Moskowitz and Grinblatt (1999) and Grundy and Martin (2001).

3. A new bootstrap approach

The bootstrap, in general, is a non-parametric approach that allows researchers to augment standard statistical tests and inference procedures. It is a method for estimating the distribution of an estimator or test statistic by resampling one's data. It amounts to treating the data as if they were the population for the purpose of evaluating the distribution of interest. It provides a way to substitute computation for mathematical analysis if calculating the asymptotic distribution of an estimator or statistic is difficult.

In finance applications, we usually use standard t statistics which assume Normal, stationary and independent distributions for stock returns, even though research has shown well-known deviations from this property, such as autocorrelation, cross-autocorrelations, conditional heteroskedasticity, skewness and leptokurtosis. Although formalized test statistics incorporating these attributes are available, the bootstrap method is more general in that it does not rely on a particular distributional assumption and specifically uses empirical distributions generated from a proposed returns-generating model of interest. Bootstrap methods have been the object of much research in statistics since its introduction by Efron (1979) and have been adopted in finance in a number of studies examining investment strategies.[5] The key advantage of bootstrapping for these and our purposes is that we can calibrate the returns of buy, sell and zero-cost buy–sell spread portfolios in the actual data by examining their empirical distributions simulated under the null hypothesis of a particular model.

3.1. Estimation-based bootstrap

The distributions of the conditional moments under various null hypotheses for stock returns movements will be estimated using the bootstrap methodology inspired by Efron (1982), Freedman (1984), Freedman and Peters (1984a,b) and Efron and Tibshirani (1993). Computer simulations of the time series designed to capture the properties of the various null returns-generating models are performed using the estimation-based bootstrap of Freedman and Peters (1984a,b). In this procedure, each model is fit to the original series to obtain parameters and residuals. We standardize the residuals using estimated standard deviations for the error process. The estimated residuals are then sampled to form a scrambled residuals series which is then used with the estimated parameters to form a new representative series for the given null model.[6] The standardized residuals are not restricted to a particular distribution, such as Normal, by this procedure.

Each of the simulations is based on 500 replications of the null model which is estimated separately for each of the 9807 stocks in the CRSP universe. This should give a good approximation of the return distribution under the null model. We will outline the various returns-generating models in the next section, but they will include a random walk

[5] See, for example, Brock et al. (1992), Levich and Thomas (1993), Kho (1996), Conrad and Kaul (1998), Sullivan et al. (1999), Jegadeesh and Titman (2001) and Kosowi et al. (2001).

[6] Freedman (1984) provides theoretical arguments as well as simulation evidence that the estimation-based bootstrap gives good estimates of standard errors for a class of linear models driven by error processes with unknown variance matrices that must be estimated from the data.

with drift as well as Fama–French's three-factor model and will allow autocorrelation (AR(1) form), cross-autocorrelation (lagged market return), lagged instrumental variables as conditioning information and a generalized autoregressive conditional heteroskedasticity (GARCH) in the returns. For the simplest random walk model with drift, we estimate the drift constant for each stock and simulate the new returns series by "scrambling" the returns from the original CRSP series. The "scrambling" procedure forms a new time series by randomly drawing from the original series with or without replacement. The scrambled series will have the same drift in prices, the same volatility and the same unconditional distribution; however, by construction, the returns are independent and identically distributed.

Unlike previous Finance applications (e.g., Brock et al., 1992; Kho, 1996), which examine stock indexes or currency futures, a number of special considerations must be given to estimation-based bootstrap simulations with individuals stocks. First, firms enter and exit the CRSP database regularly. In our applications, we choose to preserve these entry/exit points for each firm and scramble the standardized residuals within these given periods of analysis. It is important to note that we require a minimum of 24 monthly returns in order to proceed with the estimation. Second, the standardized residuals are scrambled only within the time series available for each individual stock and not across stocks. For any hypothesized returns-generating model, this approach is limiting in that it may destroy important cross-sectional dependence structure in returns in addition to time-series dependence that exists.[7] Third, we scramble only the standardized residuals for each stock and not the macroeconomic and market-wide factors and instruments that are components of the returns-generating models. There are a few studies that have performed a more general residual-and-factor resampling approach (Pesaran and Timmermann, 1995; Kosowi et al., 2001), but the dimensionality of our exercise with almost 10,000 stocks over 30 years precludes this possibility for now. Finally, there is a potential look-ahead bias in the estimation-based bootstrap approach. The parameters of a given null returns-generating model are estimated using the full sample period for which the stocks have available time series. This creates a bias in favor of a time-series structure (e.g., long-term positive drift) that may not have been known at the time of the investment decision. It also creates a potential measurement error bias that afflicts all simulated returns.[8]

The tests in this study will present simulated p values, which count the number of replications for each null model that generates returns at least as large as in the real data. We also average the simulated t statistics for the 500 replications. We will report a

[7] We thank Marno Verbeek and Peter Schotman for pointing out the limitations of our approach. They suggested an alternative block-bootstrapping procedure in which "blocks" of stocks grouped by industry or market capitalization class are scrambled jointly. One procedural difficulty with block-bootstrapping is the unbalanced panel data (with firms entering and exiting at different points in time). A "wild bootstrapping" alternative may be able to address this challenge. We have not implemented such procedures to date.

[8] The exact direction of this bias is not known for our momentum experiment. An alternative procedure would be to move the estimation window (e.g., the past 60 months) with the momentum ranking month and scramble the standardized residuals within the next 60-month time block. We thank Marno Verbeek for this important point and note its similarity to the "small sample bias" critique of Jegadeesh and Titman (2002, Section IV.C), which they direct at the resampling procedure of Karolyi and Kho (1993) and Conrad and Kaul (1998).

summary of parameter estimates for the various models as well as of the residual diagnostics for the respective models.

One important advantage of the estimation-based bootstrap approach of Freedman and Peters (1984a,b) is that it allows a decomposition analysis of the simulated momentum returns. Consider the Fama–French three-factor model as our returns-generating model of interest. This model provides three potential components for the simulation analysis: the long-term drift, or intercept, the factor risk loadings, or betas (market, size (SMB) and book-to-market (HML)), and the residuals. When we resample the standardized residuals (with or without replacement), we can generate simulated momentum returns using only the intercept or only the intercept plus the factor risk loadings in order to determine which component contributes to the overall simulated momentum returns. This decomposition analysis is helpful in diagnosing the relative importance of the drift component (a key point of debate between Conrad and Kaul, 1998 and Jegadeesh and Titman, 2002), of the factor risks and of the role of conditioning information variables (central to the debate between Chordia and Shivakumar, 2002 and Griffin et al., 2003).

3.2. Alternative resampling methods

Since the original publication of Efron's seminal paper (Efron, 1979), much attention has been directed to the potential scope of the bootstrap as a procedure to tackle many statistical problems. In this time, many have explored the wide variety of bootstrap applications and, more importantly, the many limitations of the procedure (see surveys by Efron and Tibshirani, 1993; Young, 1994; Horowitz, 2001). Young (1994) states:

> "Research has shown, not unexpectedly, that there is no specific implementation of the bootstrap paradigm which is universally superior to others. Even if the simplest bootstrap sampling scheme, which resamples from the empirical distribution function of the given data, is asymptotically valid, it may not be the only feasible approach, or the best." (p. 385)

One of the most important debates between the two existing bootstrap simulation approaches on the momentum phenomenon stems from the resampling procedures. The Conrad and Kaul (1998) experiment scrambles the monthly returns (or residuals from the random walk with drift model) by randomly drawing a return for each month with replacement from the observed distribution of returns. As they generate the same momentum returns from the scrambled series as in the real data, Conrad and Kaul conclude that the momentum profits arise because of the cross-sectional dispersion in drift estimates. Jegadeesh and Titman (2002, Section IV.C and Appendix) argue, however, that, by drawing returns with replacement, a bias is introduced that arises from the likelihood that the same observation for any stock is drawn in both the evaluation/ranking and holding periods. They advocate that the experiment should involve resampling without replacement and they show that the simulated momentum returns are never as large as the actual returns.

An important counterargument to resampling returns without replacement is also based on a small-sample bias. That is, given that the scrambling procedure restricts the residuals to the time series of availability for each stock and for that stock only, sampling without replacement necessarily forces each standardized residual to be used in a particular replication. This occurs even if the probability of that particular observation being drawn (for the empirical distribution) is unconditionally very low. This approach, in some sense, forces an error structure with potentially extremely large outliers to play an influential role in the simulated series unless handled explicitly in the null model.

In the Finance literature, it is worth noting that among the various applications of bootstrap simulations, the majority employ sampling with replacement (Brock et al., 1992; Levich and Thomas, 1993; Pesaran and Timmermann, 1995; Kho, 1996; Conrad and Kaul, 1998; Sullivan et al., 1999; Kosowi et al., 2001). The statistics literature does not have a similar consensus. A number of studies have examined the advantages and disadvantages of resampling with and without replacement. The tradeoffs revolve around the stringency of the conditions necessary for consistency and on the rates of convergence for different estimators and their accuracy. Swanepoel (1986) and Bickel et al. (1997) list the conditions to ensure consistency of certain estimators, such as for the distribution of the square of a sample average or for the maximum of a sample. Theorem 2.1 in Horowitz (2001) summarizes the conditions as related to the continuity of the estimator function and to the domain and range of permitted distribution functions.

The method of sampling without replacement has been investigated in detail by Politis and Romano (1994), who show that it consistently estimates the distribution of a statistic under very weak conditions, namely, a well-behaved asymptotic distribution for the estimator. Politis et al. (1997) extend the analysis to non-replacement sampling methods for heteroskedastic time series. Horowitz (2001, Section 2) surveys this literature and states that the conditions for consistency in replacement sampling and, therefore, for greater accuracy "are not difficult to establish in settings that arise frequently in applications in econometrics." However, non-replacement sampling will be useful when the replacement bootstrap is known to be inconsistent or where checking its consistency is difficult.

An important objective of our study is to understand these tradeoffs better in the context of the debate over simulation methods applied to the momentum phenomenon. The statistics literature offers only limited guidance on the right choice for our application, so it is left as an empirical question. Our study will, therefore, employ both bootstrap sampling with and without replacement.

4. Alternative models of the returns-generating process

We examine several alternative models of the returns-generating process to test which model is most plausible in explaining momentum profits. We begin with a naïve random walk model with drift, which is used to test whether the cross-sectional dispersion of mean returns (drift) alone is sufficient to explain the momentum profits (Conrad and Kaul, 1998). We also consider the Fama–French three-factor model that includes a market factor (proxied by the value-weighted CRSP index in excess of the risk-free rate) and two factor-mimicking portfolios that capture the size effect (SMB) and the book-to-market effect (HML).

We also evaluate several different error structures for each of the two models. The first incorporates simply a first-order autocorrelation in the returns. Returns have been shown to be positively autocorrelated over short horizons, even for value-weighted portfolios of the largest capitalization companies (Conrad and Kaul, 1989). We, therefore, extend the random walk with drift into an AR(1) model and simply add a one-month lagged own-stock return to the Fama–French three-factor model. A second feature of covariance structure of returns arises from the cross-autocorrelation structure in returns as it relates to market capitalization (Lo and MacKinlay, 1990; Mech, 1993; Conrad et al., 1991; Kroner and Ng, 1998). That is, returns of large cap stocks tend to lead the returns of smaller stocks. To both the random walk model and the Fama–French three-factor model, we add a 1-month lagged market return to capture the cross-autocorrelation structure in returns.

The third extension for estimation is a GARCH structure. Under this returns-generating process, volatility can change over time as a function of past returns shocks and past volatility (Engle, 1982; Bollerslev, 1986). This is a rich specification that is popular in Finance, but it is operationally difficult to estimate a stock-specific GARCH model for each of the 9807 stocks. Moreover, established multivariate GARCH models (Kroner and Ng, 1998) cannot be easily adapted to such a large cross-section. As a result, for both the random walk and Fama–French three-factor model, we estimate a univariate GARCH(1,1) for the CRSP value-weighted index over the 1965–2000 period. We then deflate the individual stock returns by the conditional volatility from that model for the estimation and simulation steps.[9]

Finally, we incorporate conditioning information into our returns-generating models. Chordia and Shivakumar (2002), Wu (2002) and Griffin et al. (2003) motivate a conditional version of these models since, in a dynamic market, risk exposures, prices of risk and thus expected returns are likely to vary through time. We follow their approach and propose a number of instrumental variables, including lagged one-month Treasury bill yield, the CRSP value-weighted index's dividend yield, the US bond default and term premiums (Ibbotson and Associates) and a January dummy. The instruments are demeaned over the period for which the time series of the stock is available.[10] Note that we do not scramble the instrumental variables, only the residuals from the returns-generating models (Pesaran and Timmermann, 1995; Kosowi et al., 2001).

The estimation results are presented in Table 2. The first panel presents the mean, median and standard deviation of coefficient estimates, average t statistics, adjusted R^2 and number of monthly observations for the random walk model with drift, and the second panel presents those for the Fama–French model. On average, 153 monthly observations are used in the estimation with a minimum of 24 months (as required by our screening criteria) and a maximum of 456, for the full period of analysis. For the random walk model, the adjusted R^2 are low, as expected, whereas for the Fama–French model, the time-series regressions have adjusted R^2 that average 19.4% and are as high as 77.1% for some stocks.

[9] This specification is simple but feasible. In earlier versions of this paper, we explored a number of alternative GARCH estimation procedures. For example, we estimated time-varying volatility structure for portfolios sorted by market capitalization and then adapted that for specific stocks according to the group to which they belong in a particular year.

[10] In earlier versions of this paper that examined only the single-factor market model. We also allowed the factor risk loadings to interact linearly with these instrumental variables, thus allowing the conditional betas to vary over time, following Shanken (1990) and Ferson and Schadt (1996). With a three-factor model and six instruments, the number of parameters became too cumbersome for the estimation requirements (e.g., minimum 24 months of returns per stock). These earlier results are available from the authors.

Table 2
Cross-sectional distribution of parameter estimates for alternative models

Parameter estimates	Mean	t Statistic	Median	Standard deviation
Panel A: Random walk model with drift (RW)				
1. RW model: $R_{it}=a_i+\varepsilon_{it}$				
a_i	0.956	(44.13)	0.958	2.144
Adj. R^2	0.000	(32.85)	0.000	0.000
2. RW model with auto- and cross-correlations: $R_{it}=a_i+\rho_{0i}R_{it-1}+\rho_{1i}R_{mt-1}+\varepsilon_{it}$				
a_i	0.807	(34.90)	0.863	2.289
$\rho_{0,i}$	−0.043	(−32.88)	−0.047	0.131
$\rho_{1,i}$	0.281	(47.90)	0.253	0.581
Adj. R^2	0.009	(23.77)	0.003	0.037
3. RW model with auto-, cross-correlations and heteroskedasticity: $R_{it}^*=a_i^*+\rho_{0i}^*R_{it-1}^*+\rho_{1i}^*R_{mt-1}^*+\varepsilon_{it}^*$				
a_i^*	0.157	(29.85)	0.174	0.522
$\rho_{0,i}^*$	−0.037	(−28.19)	−0.041	0.131
$\rho_{1,i}^*$	0.296	(51.98)	0.260	0.564
Adj. R^2	0.006	(16.76)	0.001	0.035
4. RW model with auto-, cross-correlations, heteroskedasticity and instruments: $R_{it}^*=a_i^*(Z_{t-1})+\rho_{0i}^*R_{it-1}^*+\rho_{1i}^*R_{mt-1}^*+\varepsilon_{it}^*$				
$a_{i,0}^*$	0.165	(30.36)	0.182	0.537
$a_{i,1}^*$ (TYLD_{t-1})	−3.509	(−6.92)	−1.504	50.186
$a_{i,2}^*$ (DEF_{t-1})	0.587	(12.91)	0.720	4.504
$a_{i,3}^*$ (TERM_{t-1})	0.227	(13.68)	0.367	1.642
$a_{i,4}^*$ (DIV_{t-1})	3.863	(4.33)	1.195	88.453
$a_{i,5}^*$ (JAN_t)	3.926	(39.90)	2.874	9.743
$\rho_{0,i}^*$	−0.048	(−35.35)	−0.052	0.135
$\rho_{1,i}^*$	0.236	(34.36)	0.184	0.680
Adj. R^2	0.026	(36.05)	0.020	0.071
Panel B: Fama–French three-factor model (FF)				
5. FF model: $R_{it}=a_i+b_iR_{mt}+c_i\mathrm{SMB}_t+d_i\mathrm{HML}_t+\varepsilon_{it}$				
a_i	0.176	(7.92)	0.173	2.207
b_i	1.059	(164.25)	1.043	0.639
c_i	0.906	(88.07)	0.761	1.019
d_i	0.197	(16.57)	0.326	1.175
Adj. R^2	0.194	(149.90)	0.186	0.128
6. FF model with auto- and cross-correlations: $R_{it}=a_i+b_iR_{mt}+c_i\mathrm{SMB}_t+d_i\mathrm{HML}_t+\rho_{0i}R_{it-0}+\rho_{1i}R_{mt-1}+\varepsilon_{it}$				
a_i	0.096	(4.01)	0.153	2.375
b_i	1.064	(161.70)	1.046	0.652
c_i	0.890	(83.74)	0.743	1.052
d_i	0.186	(15.10)	0.317	1.217
$\rho_{0,i}$	−0.062	(−48.90)	−0.066	0.126
$\rho_{1,i}$	0.164	(27.04)	0.125	0.600
Adj. R^2	0.204	(152.74)	0.198	0.132

Table 2 (*continued*)

Panel B: Fama–French three-factor model (FF)

Parameter estimates	Mean	t Statistic	Median	Standard deviation
7. FF model with auto-, cross-correlations and heteroskedasticity:				
$R_{it}^* = a_i^* + b_i^* R_{mt}^* + c_i^* \text{SMB}_t^* + d_i^* \text{HML}_t^* + \rho_{0i}^* R_{it-1}^* + \rho_{1i}^* R_{mt-1}^* + \varepsilon_{it}^*$				
a_i^*	0.021	(3.78)	0.035	0.545
b_i^*	1.056	(165.92)	1.040	0.631
c_i^*	0.894	(85.65)	0.750	1.033
d_i^*	0.161	(13.44)	0.296	1.189
$\rho_{0,i}^*$	−0.058	(−45.82)	−0.061	0.125
$\rho_{1,i}^*$	0.160	(27.33)	0.124	0.579
Adj. R^2	0.198	(152.06)	0.191	0.129
8. FF model with auto-, cross-correlations, heteroskedasticity and instruments:				
$R_{it}^* = a_i^*(Z_{t-1}) + b_i^* R_{mt}^* + c_i^* \text{SMB}_t^* + d_i^* \text{HML}_t^* + \rho_{0i}^* R_{it-1}^* + \rho_{1i}^* R_{mt-1}^* + \varepsilon_{it}^*$				
$a_{i,0}^*$	0.016	(2.89)	0.033	0.563
$a_{i,1}^*$ (TYLD_{t-1})	1.118	(2.37)	0.091	46.681
$a_{i,2}^*$ (DEF_{t-1})	−0.161	(−3.51)	−0.031	4.538
$a_{i,3}^*$ (TERM_{t-1})	−0.152	(−8.95)	−0.047	1.683
$a_{i,4}^*$ (DIV_{t-1})	−1.250	(−1.49)	0.400	83.253
$a_{i,5}^*$ (JAN_t)	1.456	(14.63)	0.354	9.860
b_i^*	1.061	(155.84)	1.039	0.674
c_i^*	0.870	(81.16)	0.730	1.062
d_i^*	0.149	(11.61)	0.295	1.267
$\rho_{0,i}^*$	−0.066	(−50.73)	−0.069	0.129
$\rho_{1,i}^*$	0.207	(28.73)	0.146	0.713
Adj. R^2	0.208	(150.55)	0.204	0.137

Reported are the cross-sectional mean, median, standard deviation and mean t statistic of the estimated parameters for the random walk with drift (RW) and Fama–French three-factor (FF) models with different error structures allowing for auto-correlation, cross-correlations, heteroskedasticity and lagged instrumental variables. The autocorrelation extension represents an AR(1) and the cross-correlation extension includes the lagged CRSP value-weighted market index return. The heteroskedasticity extension deflates the individual stock returns and the lagged own-stock, market and other variables by the conditional volatility of the CRSP value-weighted market index return, estimated by a GARCH(1,1) model. *denotes the adjustment for conditional heteroskedasticity with this approach. The instrumental variables Z_{t-1} include the lagged 1-month Treasury bill yield (TYLD_{t-1}), the bond default (DEF_{t-1}) and term (TERM_{t-1}) premiums, both from Ibbotson and Associates, the CRSP value-weighted index dividend yield (DIV_{t-1}) and a January dummy (JAN_t). "Adj. R^2" is the average adjusted R^2 across the stocks and the average number of monthly observations per stock in the estimation period is 153. The models are estimated for 9807 stocks traded on the NYSE/AMEX/NASDAQ during the period of January 1963 through December 2000 (456 months) with at least 24 observations of monthly returns, excluding stocks priced less than US$5 at the beginning of the holding period and stocks in the smallest market cap decile (NYSE size decile cutoff).

The drift coefficients for the random walk model (a_i) average 0.96% per month with significant cross-sectional standard deviation (2.14%). The extreme outliers among stocks have drift coefficients as low as −14.49% and as high as 26.15%. These coefficients are typically highly significant with an average t statistic of 44.13. When the lagged own-stock return is added, the AR(1) coefficient ($\rho_{0,i}$) is negative and significant, on average, with values around −0.04. The coefficient on the lagged market return ($\rho_{1,i}$) is positive with a value around 0.28 and is consistent with the findings in Lo and MacKinlay (1990) and

Table 3
Residual diagnostics of the alternative returns generating models

Cross-sectional statistics	Statistics of the estimated residuals							
	No. of months	Mean	Median	Standard deviation	Skewness	Excess kurtosis	ρ_1	$t(\rho_1)$
1. RW model								
Mean	153	0.000	−1.494	15.910	0.885	4.357	−0.016	(−0.17)
Median	112	0.000	−1.051	14.110	0.665	2.085	−0.016	(−0.18)
Standard deviation	115	0.000	2.112	8.708	1.101	9.107	0.120	(1.15)
2. RW model with auto- and cross-correlations								
Mean	153	0.000	−1.471	15.623	0.853	4.165	−0.003	(−0.02)
Median	112	0.000	−1.022	13.884	0.645	1.986	−0.001	(−0.01)
Standard deviation	115	0.000	2.011	8.493	1.070	8.777	0.032	(0.24)
3. RW model with auto-, cross-correlations and heteroskedasticity								
Mean	153	0.000	−0.314	3.648	0.758	3.983	−0.004	(−0.04)
Median	112	0.000	−0.221	3.209	0.545	1.854	−0.002	(−0.03)
Standard deviation	115	0.000	0.447	1.918	1.079	8.921	0.033	(0.25)
4. RW model with auto-, cross-correlations, heteroskedasticity and instruments								
Mean	153	0.000	−0.262	3.485	0.694	3.611	−0.009	(−0.09)
Median	112	0.000	−0.184	3.090	0.501	1.658	−0.006	(−0.08)
Standard deviation	115	0.000	0.402	1.815	1.003	8.297	0.048	(0.36)
5. FF model								
Mean	153	0.000	−1.221	14.045	0.900	4.457	−0.069	(−0.79)
Median	112	0.000	−0.858	12.290	0.691	2.020	−0.073	(−0.78)
Standard deviation	115	0.000	1.751	8.027	1.092	9.321	0.121	(1.24)
6. FF model with auto- and cross-correlations								
Mean	153	0.000	−1.159	13.758	0.862	4.247	−0.014	(−0.13)
Median	112	0.000	−0.829	12.087	0.663	1.909	−0.010	(−0.11)
Standard deviation	115	0.000	1.692	7.803	1.064	9.011	0.051	(0.41)
7. FF model with auto-, cross-correlations and heteroskedasticity								
Mean	153	0.000	−0.268	3.226	0.853	4.254	−0.014	(−0.13)
Median	112	0.000	−0.196	2.818	0.653	1.898	−0.009	(−0.11)
Standard deviation	115	0.000	0.387	1.789	1.077	9.309	0.049	(0.40)
8. FF model with auto-, cross-correlations, heteroskedasticity and instruments								
Mean	153	0.000	−0.225	3.084	0.778	3.870	−0.015	(−0.14)
Median	112	0.000	−0.169	2.702	0.590	1.702	−0.011	(−0.13)
Standard deviation	115	0.000	0.343	1.690	1.010	8.715	0.059	(0.45)

Reported is the residual diagnostics for the mean, median, standard deviation, skewness, excess kurtosis, first-order autocorrelation (ρ_1) and its t statistic "$t(\rho_1)$" of the estimated residuals for the returns-generating models described in Table 2. The models are estimated for 9807 stocks traded on the NYSE/AMEX/NASDAQ during the period of January 1963 through December 2000 (456 months) with at least 24 observations of monthly returns, excluding stocks priced less than US$5 at the beginning of the holding period and stocks in the smallest market cap decile (NYSE size decile cutoff).

Conrad et al. (1991). Each of these coefficients is rescaled to some extent when the adjustment for market-wide conditional heteroskedasticity is incorporated, but the most dramatic change occurs for the drift coefficient (a_i*). It is still positive and significant but with values averaging around 0.16% per month instead of 0.96% per month. This adjustment is similar in its approach to and in the result observed for the weighted-least-squares market risk premium calculation in French et al. (1987, Table 4). Finally, the lagged macroeconomic and market-wide instruments are jointly significant and have a large impact on the random walk model. Overall, the adjusted R^2 increase to 2.60% on average across stocks and can be as large as 62.2% for some. It is difficult to interpret economically the sign of these lagged information variables, but there are similarities with previous work. The positive coefficients on the default premium ($a_{i,2}$*), CRSP dividend yield ($a_{i,4}$*) and negative coefficients on the Treasury yield ($a_{i,1}$*) are consistent with Ferson and Harvey (1999), but the positive coefficient for the term premium ($a_{i,3}$*) is not.

The factor risk loadings for the three factors of the Fama–French model are all significant with positive coefficients, on average. The market betas (b_i) average, as expected, between 1.05 and 1.12. The SMB coefficient (c_i) is higher than expected, on average, around 0.87–0.90. This implies a substantial small-cap bias in the sample of 9807 stocks, which is surprising given the market capitalization cutoff criteria. Finally, there is a significant bias towards value stocks in our sample with a large positive HML coefficient (d_i) averaging around 0.15–0.20.

Table 3 presents the residual diagnostics corresponding to the various models. We report summary statistics for the cross-section of stocks of their respective mean, median, standard deviation, skewness, excess kurtosis, first-order autocorrelation and associated t statistic for the unstandardized residuals of each model. Overall, the residuals average around zero although there are influential positive outliers, as evidence by the negative median residuals and positive skewness. As expected, the residuals display positive excess kurtosis and negative autocorrelation. For the random walk model with drift, it is worth noting the extent to which the negative autocorrelation diminishes with the model extension that allows for an AR(1). In a similar manner, the standard deviations of the residuals are significantly dampened across most stocks when we apply our market-wide correction for conditional heteroskedasticity. Nevertheless, the decrease in the excess positive kurtosis with this correction is not very large.

5. Bootstrap tests

5.1. Random walk model

Table 4 presents the simulation results for the 500 replications of the random walk model (Panel A) and the Fama–French three-factor model (Panel B). For each replication of the simulated returns, we evaluate the momentum returns for the 6-month/6-month strategy. The simulations are run separately for each of the four different possible error structures and using with replacement and without replacement sampling procedures. In each model, we report the average simulated return for the 10 momentum portfolios and the winner–loser returns spread, the associated average t statistics and the simulated p value, which simply

Table 4
Bootstrap tests for the 6-month/6-month momentum strategy under alternative models

Panel A: Random walk model with drift (RW)

Portfolios	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P10 − P1
Actual return	0.769	1.098	1.216	1.272	1.290	1.358	1.388	1.523	1.691	2.228	1.460
(*t*-value)	(2.21)	(3.87)	(4.69)	(5.19)	(5.42)	(5.77)	(5.81)	(6.05)	(6.00)	(6.15)	(6.15)
Bootstrap simulations with replacement											
1. RW model											
Sim. return	1.251	1.393	1.405	1.411	1.422	1.448	1.486	1.544	1.639	1.892	0.640
(*t*-value)	(30.60)	(52.88)	(61.30)	(65.94)	(68.61)	(70.03)	(70.09)	(68.49)	(63.39)	(45.23)	(11.74)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.85]	[0.01]	[0.00]	[0.00]
2. RW model with auto- and cross-correlations											
Sim. return	1.253	1.390	1.400	1.406	1.419	1.441	1.479	1.533	1.627	1.878	0.625
(*t*-value)	(17.73)	(27.46)	(32.16)	(35.55)	(37.79)	(38.96)	(39.02)	(37.98)	(35.90)	(29.82)	(11.06)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.70]	[0.00]	[0.00]	[0.00]
3. RW model with auto-, cross-correlations and heteroskedasticity											
Sim. return	1.160	1.302	1.323	1.335	1.349	1.373	1.410	1.462	1.556	1.804	0.644
(*t*-value)	(15.86)	(24.45)	(28.60)	(31.56)	(33.56)	(34.60)	(34.85)	(34.06)	(32.43)	(27.43)	(11.18)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.82]	[0.86]	[0.00]	[0.00]	[0.00]	[0.00]
4. RW model with auto-, cross-correlations, heteroskedasticity and instruments											
Sim. return	0.910	1.183	1.250	1.292	1.330	1.377	1.434	1.513	1.644	2.007	1.097
(*t*-value)	(7.82)	(13.36)	(15.77)	(17.46)	(18.75)	(19.78)	(20.55)	(21.15)	(21.50)	(20.55)	(14.19)
[Sim. *p*-value]	[1.00]	[1.00]	[0.95]	[0.84]	[0.98]	[0.87]	[1.00]	[0.31]	[0.03]	[0.00]	[0.00]
Bootstrap simulations without replacement											
1. RW model											
Sim. return	1.646	1.564	1.506	1.469	1.446	1.439	1.440	1.451	1.470	1.462	− 0.184
(*t*-value)	(38.21)	(57.90)	(64.72)	(68.17)	(69.62)	(69.58)	(68.33)	(65.26)	(58.96)	(40.41)	(−3.62)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.00]	[0.00]	[0.00]	[0.00]
2. RW model with auto- and cross-correlations											
Sim. return	1.631	1.552	1.500	1.464	1.443	1.434	1.433	1.443	1.461	1.460	−0.171
(*t*-value)	(22.64)	(30.51)	(34.33)	(36.93)	(38.40)	(38.66)	(37.79)	(35.90)	(32.56)	(24.53)	(−3.18)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.00]	[0.00]	[0.00]	[0.00]
Bootstrap simulations without replacement											
3. RW model with auto-, cross-correlations and heteroskedasticity											
Sim. return	1.540	1.471	1.421	1.389	1.370	1.360	1.360	1.370	1.385	1.381	−0.160
(*t*-value)	(20.68)	(27.42)	(30.55)	(32.76)	(33.99)	(34.34)	(33.74)	(32.18)	(29.36)	(22.24)	(−2.90)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.57]	[0.05]	[0.00]	[0.00]	[0.00]	[0.00]
4. RW model with auto-, cross-correlations, heteroskedasticity and instruments											
Sim. return	1.255	1.336	1.341	1.341	1.349	1.364	1.387	1.425	1.487	1.619	0.364
(*t*-value)	(10.73)	(15.05)	(16.87)	(18.09)	(18.94)	(19.57)	(19.89)	(19.96)	(19.56)	(17.10)	(5.01)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.65]	[0.46]	[0.00]	[0.00]	[0.00]	[0.00]

Table 4 (*continued*)

Panel B: Fama–French three-factor model (FF)

Portfolios	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P10-P1
Bootstrap simulations with replacement											
5. FF model											
Sim. return	1.111	1.305	1.343	1.366	1.389	1.419	1.458	1.517	1.617	1.893	0.782
(*t*-value)	(3.43)	(4.75)	(5.28)	(5.62)	(5.85)	(6.01)	(6.09)	(6.09)	(5.99)	(5.76)	(5.17)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.35]	[0.00]	[0.00]	[0.00]
6. FF model with auto- and cross-correlations											
Sim. return	1.131	1.314	1.346	1.369	1.390	1.420	1.455	1.512	1.608	1.890	0.759
(*t*-value)	(3.47)	(4.77)	(5.28)	(5.62)	(5.85)	(6.01)	(6.08)	(6.09)	(5.99)	(5.76)	(5.05)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.25]	[0.00]	[0.00]	[0.00]
7. FF model with auto-, cross-correlations and heteroskedasticity											
Sim. return	1.126	1.300	1.334	1.357	1.379	1.408	1.450	1.510	1.615	1.910	0.784
(*t*-value)	(3.49)	(4.75)	(5.26)	(5.59)	(5.82)	(5.97)	(6.07)	(6.08)	(6.02)	(5.85)	(5.63)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.23]	[0.00]	[0.00]	[0.00]
8. FF model with auto-, cross-correlations, heteroskedasticity and instruments											
Sim. return	0.932	1.211	1.285	1.329	1.371	1.420	1.482	1.568	1.715	2.113	1.181
(*t*-value)	(2.88)	(4.44)	(5.08)	(5.49)	(5.79)	(6.02)	(6.20)	(6.30)	(6.35)	(6.38)	(7.63)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.99]	[0.87]	[0.00]	[0.00]
Bootstrap simulations without replacement											
5. FF model											
Sim. return	1.427	1.442	1.423	1.412	1.408	1.410	1.422	1.442	1.481	1.540	0.113
(*t*-value)	(4.40)	(5.25)	(5.59)	(5.80)	(5.92)	(5.97)	(5.93)	(5.79)	(5.50)	(4.69)	(0.75)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.99]	[0.00]	[0.00]	[0.00]	[0.00]
6. FF model with auto- and cross-correlations											
Sim. return	1.438	1.445	1.425	1.413	1.408	1.409	1.419	1.438	1.475	1.546	0.108
(*t*-value)	(4.42)	(5.25)	(5.59)	(5.80)	(5.92)	(5.96)	(5.93)	(5.79)	(5.50)	(4.72)	(0.72)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.99]	[0.00]	[0.00]	[0.00]	[0.00]
7. FF model with auto-, cross-correlations and heteroskedasticity											
Sim. return	1.434	1.435	1.415	1.401	1.396	1.401	1.413	1.436	1.478	1.557	0.123
(*t*-value)	(4.43)	(5.23)	(5.57)	(5.77)	(5.89)	(5.94)	(5.91)	(5.79)	(5.52)	(4.78)	(0.89)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.95]	[0.00]	[0.00]	[0.00]	[0.00]
8. FF model with auto-, cross-correlations, heteroskedasticity and instruments											
Sim. return	1.217	1.333	1.355	1.370	1.388	1.412	1.447	1.500	1.591	1.797	0.580
(*t*-value)	(3.76)	(4.88)	(5.36)	(5.66)	(5.86)	(5.99)	(6.05)	(6.04)	(5.90)	(5.44)	(3.78)
[Sim. *p*-value]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[1.00]	[0.06]	[0.00]	[0.00]	[0.00]

The bootstrap tests for the 6-month/6-month momentum strategy are performed with or without replacement for a sample of 9807 stocks traded in the NYSE/AMEX/NASDAQ during the period of January 1963 to December 2000 (456 months) with at least 24 observations of monthly returns, excluding stocks priced less than US$5 at the beginning of the holding period and stocks in the smallest market cap decile (NYSE size decile cutoff). P1denotes the lowest past 6-month returns portfolio and P10 is the highest. The simulated return (Sim. return) is an average of the 500 simulated returns, and the simulated *p*-value (Sim. *p*-val.) is the fraction of 500 replications that generate simulated returns larger than the actual one. *t*-values for mean returns are presented in parentheses and simulated *p*-values are in brackets. Table 2 details the different returns-generating models and specifications used, and Table 3 their respective residual diagnostics.

computes the fraction of the 500 replications that deliver simulated returns for that portfolio at least as large as in the real data.

For simulations of the simple random walk model with replacement, the simulated returns for the loser portfolios (P1, P2 or P3) are much higher than the actual returns, whereas those for the winner portfolios (P9 and P10) are much lower. In fact, not even one replication generates the low portfolio returns for the P1 to P5 (loser) portfolios. As a result, the simulated winner–loser return spreads average 0.64% per month, well below the 1.46% in the actual data. None of the 500 replications generate momentum returns as large as the actual 1.46%. When the random walk model is extended to include autocorrelation and cross-correlations, there is little impact on the average returns by momentum decile portfolio or for the winner–loser spreads. However, there is considerably greater dispersion in the simulated returns and the t statistics are accordingly much lower. Adding conditional heteroskedasticity to the random walk model interestingly generates lower simulated returns on average across all portfolios, but the winner–loser spreads do not change much (0.64%).

A significant change in the simulated returns occurs once we incorporate the lagged macroeconomic and market-wide information variables. The average simulated returns on the loser portfolios decrease by 20–30 basis points and those of the winner portfolio increase by a similar magnitude. The resulting simulated momentum spread reaches 1.09% per month. While the instruments have a large impact on the spread, the simulated p values still indicate that none of the replications generate momentum spread returns as large as the real data. It is interesting to note that the time-series variation of the simulated portfolio returns increase by another order of magnitude compared to the previous models, as witnessed by the decrease in the t statistics.

The simulated results for the random walk model without replacement provide a very different result from the simulations with replacement. In three of the four different structures for the random walk model (unconditional, with autocorrelation, cross-correlations and heteroskedasticity), the simulated returns for the loser portfolios are larger on average than those of the winner portfolios resulting in a negative momentum spread. For the unconditional model, the loser portfolio simulated returns average 1.65% while those of the winner portfolio average 1.46%, implying a -0.18% spread. Only for the model with conditional information variables do we see a positive simulated spread of 0.36%. But this magnitude is well below the actual momentum spread (simulated p value of 0.00).

Three important results arise. First, the bootstrap simulations of the random walk model with replacement are unable to yield momentum spreads as large as in the real data. This result contrasts with that in Conrad and Kaul (1998, Section 3). They do employ a different methodology (using weighted relative-strength profits) and study a different sample period (1964–1985), but they show that the random walk model (dispersion of mean returns) can, in fact, capture the actual momentum profits. Second, we can compare our results to the findings of Chordia and Shivakumar (2002) which shows that time-varying expected returns modeled with lagged information variables, like bond term and default premiums, dividend yields, can explain the momentum spreads. Our bootstrap simulation results find that they have important predictive power, but not enough to capture the total momentum returns. Finally, simulations using a sampling procedure without replacement obtain much weaker results, even with the use of information variables. This result is very similar to that shown in Jegadeesh and Titman (2002, Table 3) in which their results without replacement for a weighted relative-strength profits simulation can generate real profits only 0.54% of the replications.

5.2. Fama–French three-factor model

Panel B of Table 4 presents the equivalent results for the Fama–French model. The with-replacement simulation results for the unconditional model generate a higher momentum returns spread than that of the random walk model. On average, the simulated returns are 0.78%, almost 14 basis points higher than for the random walk. This higher momentum spread stems almost exclusively from the lower returns for the loser portfolio (1.11% versus 1.25% for the random walk model). Nevertheless, the simulated p values indicate that none of these loser portfolio returns are as low as the actual loser portfolio returns of 0.77% (simulated p value of 1.00). The addition of additional structure for autocorrelations, cross-correlations and heteroskedasticity in returns leads to negligible differences. As before with the random walk model, the inclusion of the lagged information variables do generate simulated returns with a larger momentum return spread. The simulated loser portfolio returns decrease by almost 20 basis points (0.93%) and those of the winner portfolio increase by 20 basis points (2.11%), leading to a 1.18% spread. This result is 10 basis points closer than that generated by the random walk model simulations. However, again, none of the 500 replications are able to generate momentum returns spreads as large as the actual 1.46%.

Bootstrap simulations of the Fama–French model without replacement never generate negative momentum returns like those for the random walk model in Panel A, but they are still much lower than those of the simulations with replacement. The unconditional model and those with autocorrelation, cross-correlation and heteroskedasticity provide a momentum spread around 0.12%. The loser portfolio's simulated returns average around 1.43% while those of the winner portfolio average 1.55%. When the predictive power of the lagged information variables are added to the model, the simulated momentum spread widens significantly to 0.58%. The simulated loser portfolio returns decline to 1.22% similar to those of the random walk model, but the increase in the simulated winner portfolio returns to 1.80% is much greater than the effect the information variables had for the random walk. For the simulations of the Fama–French model without replacement, none of the replications can generate returns as large as the actual 1.46% momentum spread.

These results compare favorably to those in Grundy and Martin (2001) and Fama and French (1996) studies which show that the three-factor model is unable to explain the momentum returns. Our results do not corroborate the findings of Wu (2002) which shows that the Fama–French model in its conditional form with lagged information variables can capture short-term momentum returns. An important difference with the Wu study is that the information variables are allowed to influence not only time-varying expected returns, but also the three factor loadings as well as the factor prices of risk. Of course, the Wu study implements this analysis with six portfolios of extreme winner and loser portfolios only. This extension is not feasible with the current formulation for 9807 individual stocks, but some ways of operationalizing a richer model structure may be worthy of further study.

5.3. Decomposition of simulation returns

Table 5 provides a decomposition of the simulation returns for the random walk (Panel A) and Fama–French (Panel B) models with different error structures and with different resampling procedures. The advantage of the decomposition is that it allows us to determine the components of the simulated returns spreads for a particular model. It is useful, for example, to know what fraction of the 1.09% simulated momentum return spread for the random walk model with autocorrelation, cross-correlation, heteroskedasticity and information variables can be attributed to the role of the drift alone or the information variables alone. The table presents the actual returns and for the simulations with replacement (denoted WR in the table) and without replacement (WOR) for each relevant component.

For the random walk model in Panel A, the actual momentum spread of 1.46% arises from 0.58% for the constant or drift (C1 in the table) and 0.88% from the residuals (defined as C6). One interpretation of this result is that there is important structure in actual momentum returns that cannot be captured by the differences in drift or mean returns. Indeed, the simulated returns with replacement provide an average difference in drift of 0.64% (very close to that of the actual returns) as the base component and add nothing from the scrambled residuals (which average zero on average, by construction). By contrast, for the simulations without replacement, important structure in actual momentum residuals is scrambled and is necessarily used for each replication. This results in a negative residual component (− 0.84%) to add to the difference in drift component of 0.66%. The simulated spread for sampling without replacement of − 0.18% obtains as a result.

This pattern continues for each of the random walk model variations. The additional contribution from the autocorrelation (C2) or autocorrelation and cross-correlation (C3) components to the base drift component (C1) is negligible. The impact of scrambling the residuals (C6) with replacement is neutral and without replacement is negative in each of these variations. When the model with lagged information variables is simulated, they contribute an important component (C5) to the drift. On average, the simulated returns are almost identical to those for the actual data (1.10%). That the simulations with replacement and without replacement are unable to generate momentum returns as large as in the actual data stems from the residual component of 0.35%.

For the simulations of the Fama–French model (Panel B), an important contribution arises from the factor risks (C4). For each of the variations of the model, the contribution of the factors to the momentum returns from the base component due to just the differences in drift alone averages 15 basis points. For the unconditional Fama–French model, the simulations begin with a difference in drift of 0.70% to which the factors contribute 10 basis points to 0.79%. For the Fama–French model with drift, auto- and cross-correlations and heteroskedasticity as the base (C3), the addition of the factors contribute 17 basis points from 0.49% to 0.66%, on average. The lagged information variables add another 50 basis points, on average, as with the random walk model. Nevertheless, the decomposition analysis shows that there is still 0.29% remaining in the residuals of the actual data that the simulations with or without replacement cannot meet, which is why these models are unable to deliver simulated momentum returns.

For all variations of the random walk and Fama–French three-factor models, the simulations with replacement generate simulated residuals (C6) very close to zero for winners, losers and their spreads. However, the simulations without replacement produces simulated residuals much lower than zero, ranging from − 0.63% to − 0.84%. This indicates that there is a tendency to force higher residuals to be drawn during the holding period for losers that have exhibited lower residuals in the ranking periods, and vice versa for winners. This tendency could lead to an upward bias in the holding period returns for losers and a downward bias for winners, resulting in a downward bias in the winner–loser spreads.

The result of a downward bias in simulations without replacement is important to better understand the arguments of Jegadeesh and Titman (2002). They state that the simulations with replacement allow the same returns for any stock to be drawn in both ranking and holding periods causing an upward bias in the winner–loser spreads in a small sample, whereas there is no bias in the simulation without replacement. Their argument is largely based on the observation that the simulation without replacement produces a winner–loser spread closer to zero whereas the simulation with replacement generates a substantially positive spread. They interpret the difference between the two experiments as associated with the upward bias arising from the resampling with replacement. Though our simulation results also provide the similar observation that the simulated winner–loser spreads are much smaller in the case of simulations without replacement, it is important to note that the smaller magnitude of the simulated spreads in simulations without replacement is more likely to have been caused by the downward bias in the without-replacement simulated residuals. After all, in simulations with replacement, the simulated residuals behave like white noise with a mean close to zero and thus do not contribute to the simulated spreads in any of the models.

6. Conclusions

The finance literature has continued to struggle to understand the profitability of momentum-based trading strategies, first documented by Jegadeesh and Titman (1993). These momentum returns have been consistently documented in international markets and continue to exist in US markets. The main contribution of this paper is to devise a new test methodology to evaluate how much of the momentum profits can be captured by models of time-varying expected returns. Specifically, we employ an estimation-based bootstrap simulation procedure to assess the ability of different returns-generating models that allow for time-varying expected returns, factor risks and a rich variety of cross-sectional/time-series error structures to generate momentum profits at least as large as those realized over the past forty years. In our experiments, we examine simple random walk models and the Fama–French three-factor model, with and without conditioning information variables and allowing for autocorrelated, cross-correlated and conditionally heteroskedastic errors.

Table 5
Decomposition of actual and simulated returns to the 6/6 momentum strategy under alternative models

Panel A: Random walk model with drift (RW)

Portfolio	C1 = constant	C2 = C1 + autocorrelation	C3 = C2 + cross-correlation	C4 = C3 + factors	C5 = C4 + instruments	C6 = residuals	C5 + C6 + R_{ft} = R_{it}
1. RW model							
Actual	0.577 (22.75)					0.882 (3.74)	1.460 (6.15)
WR	0.641 [1.00]					0.000 [0.00]	0.640 [0.00]
WOR	0.655 [1.00]					– 0.839 [0.00]	– 0.184 [0.00]
2. RW model with auto- and cross-correlations							
Actual	0.605 (23.83)	0.530 (17.63)	0.500 (14.27)			0.958 (4.15)	1.458 (6.18)
WR	0.685 [1.00]	0.671 [1.00]	0.625 [1.00]			0.000 [0.00]	0.625 [0.00]
WOR	0.698 [1.00]	0.684 [1.00]	0.638 [1.00]			– 0.809 [0.00]	– 0.171 [0.00]
3. RW model with auto-, cross-correlations and heteroskedasticity							
Actual	0.601 (24.37)	0.552 (18.84)	0.534 (15.55)			0.925 (3.98)	1.460 (6.15)
WR	0.699 [1.00]	0.687 [1.00]	0.651 [1.00]			– 0.007 [0.00]	0.644 [0.00]
WOR	0.713 [1.00]	0.701 [1.00]	0.665 [1.00]			– 0.824 [0.00]	– 0.160 [0.00]
4. RW model with auto-, cross-correlations, heteroskedasticity and instruments							
Actual	0.618 (22.83)	0.535 (18.71)	0.476 (14.27)		1.107 (15.30)	0.354 (1.60)	1.461 (6.16)
WR	0.678 [1.00]	0.639 [1.00]	0.562 [1.00]		1.100 [0.33]	– 0.003 [0.00]	1.097 [0.00]
WOR	0.690 [1.00]	0.651 [1.00]	0.572 [1.00]		1.119 [0.80]	– 0.755 [0.00]	0.364 [0.00]

Panel B: Fama–French three-factor model (FF)

5. FF model							
Actual	0.582 (20.57)			0.728 (4.40)		0.733 (5.08)	1.461 (6.16)
WR	0.696 [1.00]			0.782 [1.00]		0.000 [0.00]	0.782 [0.00]
WOR	0.709 [1.00]			0.797 [1.00]		– 0.684 [0.00]	0.113 [0.00]
6. FF model with auto- and cross-correlations							
Actual	0.619 (21.52)	0.472 (14.27)	0.461 (13.48)	0.615 (3.69)		0.845 (5.93)	1.460 (6.16)
WR	0.752 [1.00]	0.716 [1.00]	0.683 [1.00]	0.759 [1.00]		0.001 [0.00]	0.759 [0.00]
WOR	0.764 [1.00]	0.727 [1.00]	0.693 [1.00]	0.772 [1.00]		– 0.663 [0.00]	0.108 [0.00]
7. FF model with auto-, cross-correlations and heteroskedasticity							
Actual	0.620 (21.00)	0.493 (14.94)	0.493 (14.28)	0.655 (4.02)		0.804 (5.56)	1.459 (6.15)
WR	0.749 [1.00]	0.714 [1.00]	0.691 [1.00]	0.791 [1.00]		– 0.007 [0.00]	0.784 [0.00]
WOR	0.761 [1.00]	0.725 [1.00]	0.702 [1.00]	0.806 [1.00]		– 0.683 [0.00]	0.123 [0.00]
8. FF model with auto-, cross-correlations, heteroskedasticity and instruments							
Actual	0.624 (20.94)	0.462 (14.35)	0.429 (12.97)	0.528 (3.39)	1.170 (6.75)	0.291 (2.05)	1.461 (6.16)
WR	0.711 [1.00]	0.637 [1.00]	0.585 [1.00]	0.632 [1.00]	1.187 [0.82]	– 0.007 [0.00]	1.181 [0.00]
WOR	0.722 [1.00]	0.648 [1.00]	0.594 [1.00]	0.643 [1.00]	1.205 [0.97]	– 0.626 [0.00]	0.580 [0.00]

Monthly excess returns in actual data are decomposed into components computed as the parameter estimate times the predetermined value of the regressors in the particular model, and each component is averaged across stocks within each of the past-return decile portfolio determined by the 6/6 momentum strategy. *t*-statistics for the means are reported in parentheses. The momentum strategy is also applied to the simulated returns series that are constructed for each stock by scrambling the estimated residuals with or without replacement while maintaining the other estimated components the same as in the actual data. The average value of each component is calculated across stocks within each of the simulated past-return decile portfolios determined by the momentum strategy. Reported are their averages over 500 replications ("WR" for the simulations with replacement, and "WOR" for those without replacement). The simulated p-value shown in brackets is the fraction of 500 replications that generate a simulated component larger than the actual one.

We also investigate alternative sampling techniques in the bootstrap simulations—namely, with replacement and without replacement—and a new decomposition analysis.

Overall, we find that none of the models are able to generate simulated momentum profits as large as the actual profits over the 1964–2000 sample period. However, accounting for time-varying expected returns with macroeconomic and market-wide information variables can generate simulated momentum returns that are 75–80% of the actual winner–loser spreads. We offer these results as important new evidence for the current debate between those that advocate behavioral explanations of the phenomenon (investor cognitive biases, investor underreaction to news) and those that rationalize momentum as evidence of time-varying expected returns. For the former group, our evidence suggests that the simulated momentum profits consistent with well-known models even implemented with rich structure cannot explain the actual results. For the latter group, we show that the simulated profits can yield momentum profits that are a large fraction of the actual profits.

It is important to recognize what the paper does not contribute and to define a number of limitations and shortcomings. First, the analysis in this paper focuses on only the 6-month/6-month investment strategy, though the literature has evaluated the momentum profits in many varied forms. For example, in the original Jegadeesh and Titman (1993) study and in their follow-up (2001) study, they study the momentum portfolio returns up to 3 years after the portfolio formation with the idea that even if markets are not fully efficient the effect of any information will likely be impounded in prices within this time frame. They show a monotonic increase in momentum returns to the end of 12 months following formation and then a dramatic reversal until 5 years following formation by which time the initial positive returns are completely eliminated. They claim that this post-holding period evidence is further evidence against the Conrad–Kaul hypothesis that the momentum profits are due to unconditional differences in winner and loser stocks' drift or average returns.

We investigate this claim with our sample and simulated returns. In Fig. 1, we replicate the finding in Jegadeesh and Titman (2001) for the actual returns and simulate the returns, both with and without replacement, using the Fama–French three-factor model with conditioning information variables (reporting the 5th and 95th quantiles of the 500 replications). The results show that the simulated returns are never able to reach the actual momentum returns to month 12, especially those sampling without replacement, which we knew already. However, the figure also shows that the simulated returns with replacement continue to extrapolate positive momentum returns to month 60 and does not capture the reversal pattern in the actual data. By contrast, those returns simulated without replacement sampling actually do provide a reversal in the long term.[11] This finding is only preliminary but, in our opinion, calls for a bootstrap simulation analysis of a broader array of the momentum/contrarian returns strategies that others have uncovered.

Another important limitation of the current study is in terms of the procedures for operationalizing the bootstrap simulation. The dimensionality of our experiment for 456 months for 9807 stocks precludes a number of creative and likely richer approaches. For

[11] We explicitly thank Louis Chan and Werner DeBondt for this excellent suggestion.

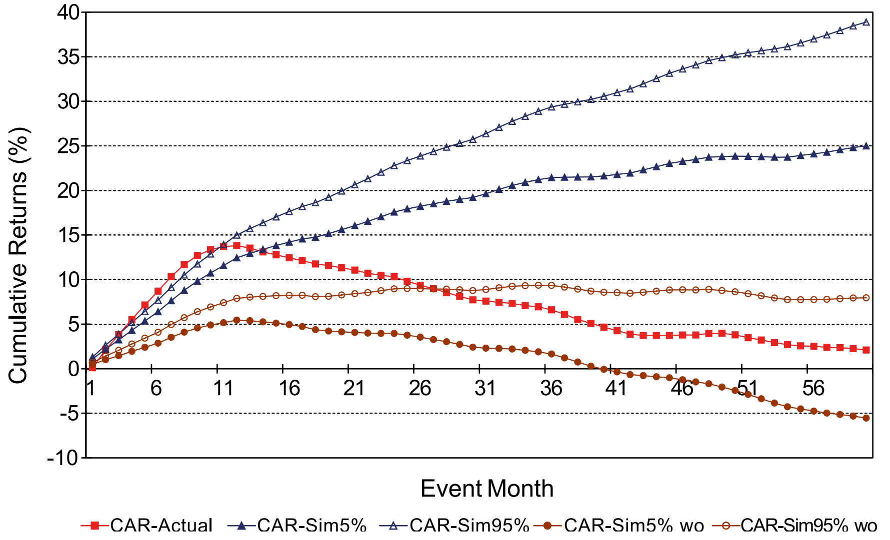

Fig. 1. Bootstrap simulations of long-horizon momentum returns using the Fama–French three-factor model using with and without replacement procedures. The figure represents the actual and simulated cumulative momentum portfolio returns for 1965–2000 with a sample of stocks traded on the NYSE, Amex and Nasdaq. The sample comprises all stocks that are larger than the smallest NYSE market cap decile at the beginning of the event period. Stocks priced less than US$5 at the beginning of each event month are excluded. The simulated returns are generated by scrambling residuals with and without replacement from the Fama–French three-factor model with autocorrelation, cross-correlations, heteroskedasticity and including lagged instrumental variables. See Table 2 for estimation details. CAR denotes the cumulative momentum returns, 5% and 95% represent the 5th and 95th quantiles of the simulated returns with replacement (wo denotes simulations without replacement). Event month denotes the number of months since the ranking of the stocks into momentum portfolios.

example, our estimation-based bootstrap uses the full time series available for a stock for estimation of parameters which can create a look-ahead bias in the simulations. We only scramble residuals within the time-series domain available for a stock and not across stocks, which could be explored by "wild bootstrapping" techniques. In so doing, we also likely destroy important cross-sectional dependence structure in the returns, which might be alleviated by a "block-bootstrapping" procedure. We caution readers about these limitations and look forward to exploring these other possible extensions in future work.

Finally, we explore only a few possible returns-generating models for the estimation-based bootstrap simulations. Moreover, we add additional structure, such as autocorrelation, cross-correlation and heteroskedasticity in a crude manner. Our choice again stems from the dimensionality of the estimation and simulation experiment. It would be worthwhile to explore richer approaches, such as stock-specific autocorrelation and cross-correlation structures and stock-specific GARCH specifications (e.g., Factor ARCH models, Kroner and Ng, 1998). One important issue, however, is the issue of model mining. Our objective in this study has not been to search for the "best" model for the momentum phenomenon. Rather, we have focused on models that had already been proposed and investigated by existing studies in the momentum literature in order to showcase our new bootstrap simulation experiment. Nevertheless, the risk of model mining is an important one that we caution researchers that plan to extend the approaches in our paper to heed.

Acknowledgements

We are grateful for comments from two anonymous referees, Editor Franz Palm, Marno Verbeek, Louis Chan, Werner DeBondt, David Hirshleifer, Peter Schotman and other participants at Ohio State University, the Korean Finance Association annual meetings and the LIFE Conference on Behavioral Finance in Palma de Mallorca, Spain. Both the Dice Center for Financial Economics at Ohio State University (Karolyi) and the Institute of Management Research at Seoul National University (Kho) provided financial support. All remaining errors are our own. This paper is an extension of a 1996 Ohio State University Dice Center working paper by the same authors, entitled "Time-Varying Risk Premia and the Returns to Buying Winners and Selling Losers: Caveat Emptor et Venditor." Participants in seminars given with the earlier paper at Ohio State, Chicago, Australian Graduate School of Management, Western Ontario, Toronto, and Queen's greatly influenced our thinking on these issues.

References

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of Financial Economics 49 (3), 307–343.

Berk, J.B., Green, R., Naik, V., 1999. Optimal investment, growth options, and security returns. Journal of Finance 54 (5), 1553–1607.

Bickel, P.J., Goetze, F., van Zwet, W.R., 1997. Resampling fewer than *n* observations: gains, losses and remedies for losses. Statistica Sinica 7, 1–32.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31 (3), 307–328.

Brock, W.A., Lakonishok, J., LeBaron, B., 1992. Simple technical rules and the stochastic properties of stock returns. Journal of Finance 24 (1), 66–77.

Chan, L.K.C., Jegadeesh, N., Lakonishok, J., 1996. Momentum strategies. Journal of Finance 51 (5), 1681–1713.

Chan, K., Hameed, A., Tong, W., 2000. Profitability of momentum strategies in the international equity markets. Journal of Financial and Quantitative Analysis 35 (2), 153–172.

Chordia, T., Shivakumar, L., 2002. Momentum, business cycle and time-varying expected returns. Journal of Finance 57 (2), 985–1019.

Conrad, J., Kaul, G., 1989. Mean reversion in short-horizon expected returns. Review of Financial Studies 2 (2), 225–240.

Conrad, J., Kaul, G., 1998. An anatomy of trading strategies. Review of Financial Studies 11 (3), 489–519.

Conrad, J., Gultekin, M.N., Kaul, G., 1991. Asymmetric predictability of conditional variances. Review of Financial Studies 4 (4), 597–622.

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. Journal of Finance 53 (6), 1839–1885.

Efron, B., 1979. Bootstrap methods: another look at the jackknife. Annals of Statistics 7, 1–26.

Efron, B., 1982. The Jackknife, the Bootstrap and Other Resampling Plans. Capital City Press, Montpelier, VT.

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman & Hall, New York, NY.
Engle, R.F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflations. Econometrica 50 (4), 987–1008.
Fama, E.F., French, K.R., 1996. Multifactor explanation of asset pricing anomalies. Journal of Finance 51 (1), 55–84.
Ferson, W.E., Harvey, C.R., 1999. Conditioning variables and the cross section of stock returns. Journal of Finance 54 (4), 1325–1360.
Ferson, W.E., Schadt, R.W., 1996. Measuring fund strategy and performance in changing economic conditions. Journal of Finance 51 (2), 425–461.
Freedman, D.A., 1984. On bootstrapping two-stage least squares estimates in stationary linear models. Annals of Statistics 12, 827–842.
Freedman, D.A., Peters, S.C., 1984a. Bootstrapping an economic model: some empirical results. Journal of Business and Economic Statistics 2 (2), 150–158.
Freedman, D.A., Peters, S.C., 1984b. Some notes on the bootstrap in regression problems. Journal of Business and Economic Statistics 2 (4), 406–409.
French, K.R., Schwert, G.W., Stambaugh, R.F., 1987. Expected stock returns and volatility. Journal of Financial Economics 19 (1), 3–30.
Griffin, J., Ji, S., Martin, J.S., 2003. Momentum investing and business cycle risk: evidence from pole to pole. Journal of Finance 58 (6), 2515–2547.
Grundy, B.D., Martin, J.S., 2001. Understanding the nature of the risks and the source of the rewards to momentum investing. Review of Financial Studies 14 (1), 29–78.
Hameed, A., Kusnadi, Y., 2002. Momentum strategies: evidence from pacific basin stock markets. Journal of Financial Research 45 (3), 383–398.
Hong, H., Stein, J.C., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. Journal of Finance 54 (6), 2143–2184.
Hong, H., Lim, T., Stein, J.C., 2000. Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies. Journal of Finance 55 (1), 265–295.
Horowitz, J.L., 2001. The Bootstrap. In: Heckman, J.J., Leamer, E.E. (Eds.), Handbook of Econometrics, vol. 5. North-Holland, Amsterdam, The Netherlands, pp. 3159–3228 (Chapter 52).
Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48 (1), 65–92.
Jegadeesh, N., Titman, S., 2001. Profitability of momentum strategies: an evaluation of alternative explanations. Journal of Finance 56 (2), 699–720.
Jegadeesh, N., Titman, S., 2002. Cross-sectional and time-series determinants of momentum returns. Review of Financial Studies 15 (1), 143–158.
Karolyi, G.A., Kho, B.-C., 1993. Time-varying risk premia and the returns to buying winners and selling losers: Caveat Emptor et Venditor, Ohio State University working paper.
Kho, B.-C., 1996. Time-varying risk premia, volatility, and technical trading rule profits: Evidence from foreign currency futures markets. Journal of Financial Economics 41 (2), 249–290.
Korajczyk, R., Sadka, R., 2004. Are momentum profits robust to trading costs? Journal of finance (in press).
Kosowi, R., Timmermann, A., White, H., Wermers, R., 2001. Can mutual fund stars really pick stocks? New evidence from a bootstrap analysis. University of Maryland at College Park working paper.
Kroner, K.F., Ng, V.K., 1998. Modeling asymmetric comovements of asset returns. Review of Financial Studies 11 (4), 817–844.
Lee, C.M.C., Swaminathan, B., 2000. Price momentum and trading volume. Journal of Finance 55 (5), 2017–2069.
Levich, R.M., Thomas III, L.R. 1993. The significance of technical trading-rule profits in the foreign exchange market: a bootstrap approach. Journal of International Money and Finance 12 (5), 451–474.
Lo, A.W., MacKinlay, A.C., 1990. Data-snooping biases in tests of financial asset pricing models. Review of Financial Studies 3 (3), 431–468.
Mech, T.S., 1993. Portfolio return autocorrelation. Journal of Financial Economics 34 (3), 307–344.
Moskowitz, T.J., Grinblatt, M., 1999. Do industries explain momentum? Journal of Finance 54 (4), 1249–1290.
Pesaran, M.H., Timmermann, A., 1995. Predictability of stock returns: robustness and economic significance. Journal of Finance 50 (4), 1201–1228.

Politis, D.N., Romano, J.P., 1994. Large sample confidence regions based on subsamples under minimal assumptions. Annals of Statistics 22, 2031–2050.

Politis, D.N., Romano, J.P., Wolf, M., 1997. Subsampling for heteroskedastic time series. Journal of Econometrics 80 (2), 281–317.

Rouwenhorst, K.G., 1998. International momentum strategies. Journal of Finance 53 (1), 267–284.

Rouwenhorst, K.G., 1999. Local return factors and turnover in emerging stock markets. Journal of Finance 54 (4), 1439–1464.

Shanken, J., 1990. Intertemporal asset pricing: an empirical investigation. Journal of Econometrics 45, 99–120.

Sullivan, R., Timmermann, A., White, H., 1999. Data-snooping, technical trading rule performance, and the bootstrap. Journal of Finance 54 (5), 1647–1691.

Swanepoel, J.W.H., 1986. A note on proving that the modified bootstrap works. Communications in Statistics Theory and Methods 15 (2), 3193–3203.

Wu, X., 2002. A conditional multifactor model of return momentum. Journal of Banking and Finance 26 (8), 1675–1696.

Young, G.A., 1994. Bootstrap: more than a stab in the dark. Statistical Science 9 (3), 382–395.