

The Persistence of Runs—The Directional Movement of Index Returns

Zhen Li*

University of Notre Dame

I test a stronger version of the random walk hypothesis by examining the hazard ratio of runs and find the directional movement of the CRSP index return predictable. A run is defined as a sequence of returns of the same sign. The persistence level of a run tends to abate in its length but enhance itself in the magnitude of its components (historical return realizations). A positive run lasts longer than a negative run. The impact of a run's components on its persistence level becomes more pronounced when the measuring frequency increases. Size and book-to-market equity ratio have some limited impact on the persistence level of a run. The market interest (T-bill) rate generally decreases the persistence level of a positive run and increases the persistence level of a negative run.

Introduction

Many academic studies in the past twenty years have demonstrated that certain trading strategies, such as the momentum and/or contrarian strategies, generate excess profits over various portfolio holding windows; see DeBondt and Thaler (1985), Jegadeesh (1990), Grinblatt and Titman (1992), Jegadeesh and Titman (1993, 1995), Conrad and Kaul (1998), Moskowitz and Grinblatt (1999), among many others. These trading strategies cast some doubt on the efficient market hypothesis.

These profits usually are driven by a positive or negative autocorrelation, depending on the length of the portfolio holding windows, in time series returns. For example, Fama and French (1988) find a large negative autocorrelation for return horizon beyond a year (long-term); Lo and MacKinlay (1988), using the variance ratio test, show that the weekly (short-term) autocorrelation for index return is positive. Poterba and Summer (1988) find positive autocorrelation over short

* I thank Louis Chan, Ronald Oaxaca, Jaime Zender, and workshop participants at the 11th Global Finance Conference for helpful comments. All errors are mine.

intervals and negative autocorrelation over longer intervals. The positive/negative return autocorrelation implies that certain momentum/contrarian strategies may be profitable.

It is notable that most of the existing studies assume linearity in time-series stock returns and thus are tests of the weakest form of the random walk hypothesis. Some exceptions are Niederhoffer and Osborne (1966) and McQueen and Thorley (1991), both of which examine the directional movement of stock returns using the finite state Markov chain approach and find no support for the random walk hypothesis. These two studies address non-linearity by allowing the transition probabilities to vary depending on a given sequence of prior states. McQueen and Thorley (1994) and Lo, Mamaysky, and Wang (2000) are also examples of non-linear tests.

I propose a non-linear test by studying the hazard ratio of a run over its entire historical path. A run is defined as a sequence of returns of the same sign and the hazard ratio is defined as its probability of ending and reversing direction conditional on its previous return realizations.¹ It is a test of a stronger version of the random walk hypothesis. The purpose is to determine whether the *directional* movement of a stock return is path-dependent and thus predictable. The method I use is the duration model. One advantage of this method over the Markov chain approach is that I allow the length of the return sequence to be flexible, while the Markov chain approach has to specify its order, the exact length of the return sequence.

The results suggest that the directional movement of the CRSP index return is predictable in the following manner. First, the persistence of a run tends to abate in its length, suggesting some mean reversion tendency in returns. At the same time, the persistence level of a run tends to enhance itself in the magnitude of its components (historical return realizations). Second, a positive run lasts longer than a negative run, implying a positive drift term in returns. Third, the impact of a run's components on its persistence level becomes more pronounced when the measuring frequency increases from monthly to weekly and to daily. Fourth, the impact of CRSP equal-weighted index return on the persistence level of a run is more pronounced than that of CRSP value-weighted index return, suggesting that the directional movement of small stocks is possibly more predictable. This may also be true for high book-to-market equity stocks, though in a very limited sense. Fifth, the market interest rate generally decreases the persistence level of a positive run and increases the persistence level of a negative run.

¹ Mood's (1940) general definition of run is: "A run is defined as a succession of similar events preceded and succeeded by different events; the number of elements in a run will be referred to as its length." Fama (1965) is an empirical application of Mood's distributional theory of runs; no evidence of dependence is revealed through runs in his paper. My statistical methodologies are different from those formulated in Mood (1940).

Random Walk Hypothesis of Stock Returns

The simple iid (independently and identically distributed increments) version of random walk hypothesis suggests that the stock price follows this process:

$$P_t = \mu + P_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, \sigma^2), \quad (1)$$

where μ is the drift. In this study, I focus on a binary process, that is, the probability that $P_t \geq P_{t-1}$, which equals the following,

$$\text{prob}[P_t - P_{t-1} = \mu + \varepsilon_t \geq 0] = \text{prob}[\varepsilon_t \geq -\mu] = 1 - F(-\mu), \quad (2)$$

where $F(\cdot)$ is the cdf of ε_t . I obtain the above result because of the iid nature of ε_t . Equation (2) implies that the directional movement of price depends only on its drift, which is constant in this case. If I define return as $R_t = (P_t - P_{t-1})/P_{t-1}$, Equation (2) becomes

$$\text{prob}[R_t \geq 0] = \text{prob}[\varepsilon_t \geq -\mu] = 1 - F(-\mu). \quad (3)$$

This means that the probability of observing a subsequent positive return does not depend on the sequence of historical returns. I have

$$\text{prob}[R_{t+1} \geq 0 | R_t, R_{t-1}, \dots] = \text{prob}[R_{t+1} \geq 0] = 1 - F(-\mu). \quad (4)$$

In the case of a run (defined as a sequence of consecutive returns of the same sign), Equation (4) becomes

$$\text{prob}[R_{t+1} \geq 0 | R_t, R_{t-1}, \dots, R_1] = \text{prob}[R_{t+1} \geq 0] = 1 - F(-\mu), \quad (5)$$

where $T = 1, 2, \dots, t$ indexes each period since the beginning of a run. This suggests that the probability of a subsequent up-tick does not depend on the current length $T = t$ of the run and the magnitude of its components $R_t, R_{t-1}, \dots, R_1$ (historical return realizations). In other words, the random walk property suggests that the conditional probability of a directional return movement is without structure or is constant (equaling its unconditional probability). If the conditional probability is found to be dependent on t and/or $R_t, R_{t-1}, \dots, R_1$, this should constitute a rejection of the random walk hypothesis.

I define the hazard ratio as the conditional probability of a run ending and reversing direction. This conditional probability also can be construed as a prediction of the turning point of a run. Assume that the length of a run T has a pdf of $f(t)$, where t is a realization of T , its cdf is

$$F(t) = \text{prob}(T \leq t) = \int_0^t f(s) ds. \quad (6)$$

This is the probability that a run does not survive t periods. The probability of the run surviving at least t periods is $S(t) = 1 - F(t)$. The hazard function $\lambda(t)$ is defined as

the probability that a run fails to survive the subsequent period conditional on surviving the previous t periods. In its generic form, the hazard function is

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{F[t \leq T < t + \Delta t | T \geq t]}{\Delta t}, \quad (7)$$

which is equivalent to

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{d \ln S(t)}{dt}, \quad (8)$$

since $f(t) = \lambda(t)S(t)$. Note that Equation (8) is also equivalent to

$$S(t) = \exp \left[-\int_0^t \lambda(s) ds \right]. \quad (9)$$

Econometric Model for Analyzing the Hazard Function

In order to perform regression-like parametric analyses and facilitate hypothesis testing, an underlying distribution needs to be specified for the hazard function. Distributions commonly used are the exponential, log-normal, log-logistic, and Weibull. The exponential distribution has a constant hazard rate of λ , the log-normal and log-logistic distributions allow the hazard rate to increase first and then decrease. The Weibull distribution may be ideal for the purpose of this study because of the monotonicity property of its hazard function, which takes the following basic form,

$$\lambda(t) = e^{x\beta} \lambda p (\lambda t)^{p-1}, \quad (10)$$

where λ and p are its parameters.² Note that in equation (10), I also let the Weibull hazard be scaled by a vector of time-varying covariates x , which includes the historical return realizations within a run. The first derivative of the hazard function equation (10) with respect to time t is

$$\frac{d\lambda(t)}{dt} = e^{x\beta} \lambda^2 p(p-1)(\lambda t)^{p-2}. \quad (11)$$

This means that when $p > 1$, the hazard rate is increasing in time t ; when $p = 1$, the hazard rate is constant (thus the Weibull distribution collapses into an exponential distribution); when $p < 1$, the hazard rate is decreasing in time t .

Two layers of tests of the random walk hypothesis can be performed here. First, if either λ or p is insignificantly different from zero, or p is insignificantly different from 1, the random walk hypothesis is supported; if both λ and p are significantly

² A part from tractability, later in the analysis, I show why the monotonicity property of the Weibull hazard is nice for this study. The nonparametric tests all indicate a hazard ratio that is roughly increasing in time (length of the runs).

different from zero and p is also significantly different from 1, the random walk hypothesis is rejected. Second, if the random walk hypothesis is rejected, I also want to know how the hazard ratio changes as a function of the length of a run. Positive duration dependence is defined as $d\lambda(t)/dt > 0$ while negative duration dependence is defined as $d\lambda(t)/dt < 0$. The Weibull distribution is tractable and parsimonious for the purpose of this study in the sense that when $p < 1$, I can conclude that the longer the length of a run, the less likely that it will end in the subsequent period—the existence of negative duration dependence. When $p > 1$, I can conclude that the longer the length of a run, the more likely that it will end in the next period—the existence of positive duration dependence. When $p = 1$, the conditional probability that a run will end in the next period is not path-dependent.

I also determine whether and how the magnitude of returns within a run and some other factor, such as interest rates, affect its hazard and/or persistence level. To deal with time-varying covariates, I follow Petersen (1986).³ (See Appendix A.)

Table 1—Sample Statistics

Monthly, weekly and daily CRSP value-weighted returns (VWRET) and equal-weighted returns (EWRET) are obtained from the Center for Research in Security Prices (CRSP) for the period between July 1962 and December 1999. The length of a positive run is defined as the number of consecutive positive monthly/weekly/daily returns and the length of a negative run is defined as the number of consecutive negative monthly/weekly/daily returns

Panel A: Return Statistics

	Observations	CRSP Value-Weighted Index			CRSP Equal-Weighted Index		
		Mean	Median	Std.	Mean	Median	Std.
Monthly	450	0.011019	0.013232	0.043557	0.018787	0.020922	0.055759
Weekly	1957	0.002488	0.004056	0.019530	0.004119	0.005657	0.019536
Daily	9442	0.000511	0.000702	0.008208	0.000838	0.001376	0.006820

Panel B: Run Statistics

CRSP Value-Weighted Index

	Positive Runs					Negative Runs				
	Obs.	Max	Min	Mean	Median	Obs.	Max	Min	Mean	Median
Monthly	99	11	1	2.79	2	100	7	1	1.69	1
Weekly	455	16	1	2.49	2	455	8	1	1.80	1
Daily	2033	17	1	2.54	2	2034	12	1	2.10	2

CRSP Equal-Weighted Index

	Positive Runs					Negative Runs				
	Obs.	Max	Min	Mean	Median	Obs.	Max	Min	Mean	Median
Monthly	82	16	1	3.55	3	83	7	1	1.86	1
Weekly	327	21	1	3.86	2	327	10	1	2.09	1
Daily	1608	35	1	3.60	2	1609	15	1	2.27	2

³ For more understanding of the duration model, refer to the following papers/books, among others: Kaplan and Meier (1958), McGinnis (1968), Cox (1972, 1975), Lancaster (1979), Kalbfleisch and Prentice (1980), Heckman and Singer (1982), Allison (1984), Kennan (1985), Petersen (1986), Kiefer (1988), Han and Hausman (1990), etc.

Data and Empirical Results

I perform tests on the market index. Monthly, weekly and daily CRSP value-weighted returns (VWRET) and equal-weighted returns (EWRET) are obtained from the Center for Research in Security Prices (CRSP) for the period between July 1962 and December 1999. See Table 1 for sample statistics of CRSP index returns. A positive (negative) run is defined as a sequence of consecutive positive (negative) returns with observed starting and ending points. By this definition, positive and negative runs alternate within a time-series of CRSP index returns. I test the persistence of positive and negative runs for monthly, weekly and daily CRSP value-weighted and equal-weighted index returns. There are a total of 12 combinations.

Nonparametric Tests

I perform nonparametric tests first. The method used is the Life Table Product-Limit (PL) approach formulated by Kaplan and Meier (1958). For detail, see Appendix B.

Table 2 (Figure 1) shows the nonparametric test results for positive and negative runs for the monthly, weekly and daily CRSP value-weighted index. All six combinations show a hazard function that is generally increasing in the length of the run. This result is qualitatively similar for both positive and negative runs, though the length of an average negative run is considerably shorter than that of an average positive run. This suggests that when the length of a run increases, its conditional probability of reversing direction (hazard ratio) increases; thus, the CRSP index return has some mean-reverting tendency. The other six combinations using CRSP equal-weighted index yield qualitatively similar results. To conserve space, I do not report them here. A hazard ratio that is increasing in the length of a run suggests that the use of the Weibull hazard with its monotonicity property in the parametric specifications is appropriate.

Parametric Tests with Time-Varying CRSP Index Returns

In this section, I perform a series of parametric tests. Parametric tests enable me to test for the significance level of the path-dependency property of the hazard function, if it exists. As discussed earlier, for the Weibull hazard, if $p > 1$, the hazard is increasing in the length of a run; if $p < 1$, the hazard is decreasing in the length of the run; if $p = 1$, the hazard is not path-dependent. I concentrate on testing whether and how the estimated p is different from 1.

The results are shown in Table 3. For the twelve categories of runs (based on the directions of return movement, index types and sampling frequencies), the estimates for p range from a minimum of 1.1900 (weekly positive CRSP equal-weighted runs) to a maximum of 1.7352 (monthly negative CRSP equal-weighted runs) and they are all significantly greater than 1 at less than 1 percent level (one-tail test). This means that when a run lengthens, its probability of discontinuing (hazard ratio) increases.

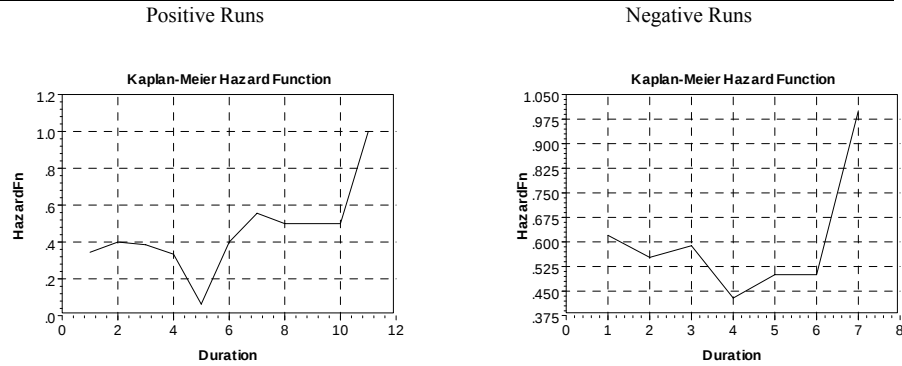
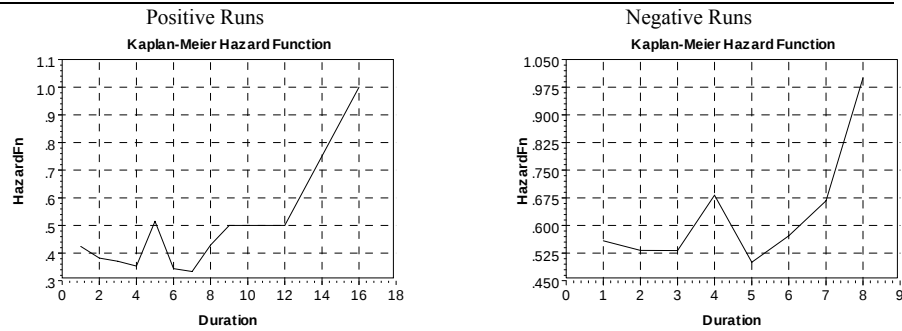
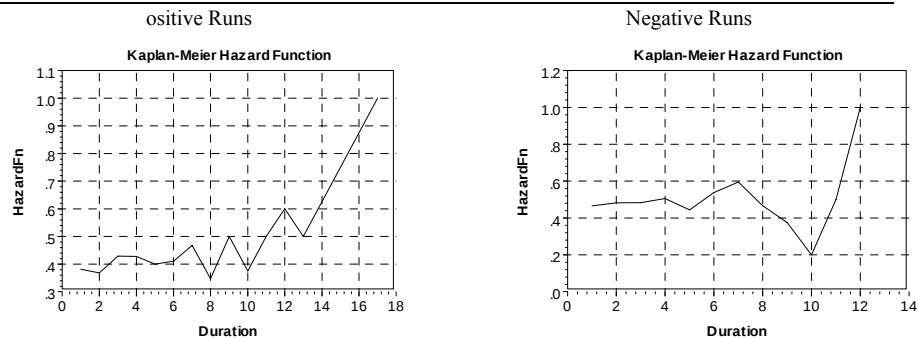
Monthly, weekly, and daily CRSP value-weighted returns (VWRET) are obtained from the Center for Research in Security Prices (CRSP) for the period between July 1962 and December 1999. The length of a positive run is defined as the number of consecutive positive monthly/weekly/daily returns, and the length of a negative run is defined as the number of consecutive negative monthly/weekly/daily returns. "Risk" is the number of runs that potentially can reverse direction during the period; "Exit" is the number of runs that actually reverse direction during the period. The numbers in the parentheses are p-values

Panel A: Positive Runs									
	Monthly Runs			Weekly Runs			Daily Runs		
	Survival	Risk	Exit	Survival Rate	Hazard Rate	Risk	Exit	Survival Rate	Hazard Rate
0-1	99	0	1.0000 (0.000)	0.0000 (0.000)	0.0000 (0.000)	455	0	1.0000 (0.000)	0.0000 (0.000)
1-2	99	34	1.0000 (0.000)	0.4146 (0.070)	0.5384 (0.037)	455	193	1.0000 (0.000)	0.4710 (0.016)
2-3	65	26	0.6566 (0.048)	0.5000 (0.095)	0.4717 (0.046)	262	100	0.5758 (0.023)	0.4510 (0.020)
3-4	39	15	0.3939 (0.049)	0.4762 (0.119)	0.4545 (0.057)	162	60	0.3560 (0.022)	0.5460 (0.028)
4-5	24	8	0.2424 (0.043)	0.4000 (0.139)	0.4286 (0.070)	102	36	0.2242 (0.020)	0.5434 (0.038)
5-6	16	1	0.1616 (0.037)	0.0645 (0.064)	0.6939 (0.112)	66	34	0.1451 (0.017)	0.5000 (0.047)
6-7	15	6	0.1515 (0.036)	0.5000 (0.198)	0.4151 (0.122)	32	11	0.0703 (0.012)	0.5161 (0.062)
7-8	9	5	0.0909 (0.029)	0.7692 (0.318)	0.4000 (0.148)	21	7	0.0462 (0.010)	0.6099 (0.089)
8-9	4	2	0.0404 (0.020)	0.6667 (0.444)	0.5455 (0.214)	14	6	0.0308 (0.008)	0.4198 (0.157)
9-10	2	0	0.0202 (0.014)	0.0000 (0.000)	0.6667 (0.314)	8	4	0.0176 (0.006)	0.6667 (0.157)
10-11	2	2	0.0202 (0.014)	2.0000 (0.000)	0.6667 (0.444)	4	2	0.0088 (0.004)	0.4615 (0.183)
11-12					0.0000 (0.000)	2	0	0.0044 (0.003)	0.6667 (0.281)
12-13					0.6667 (0.629)	2	1	0.0044 (0.003)	0.8571 (0.447)
13-14					0.0000 (0.000)	1	0	0.0022 (0.002)	0.6667 (0.629)
14-15					0.0000 (0.000)	1	0	0.0022 (0.002)	0.0000 (0.000)
15-16					2.0000 (0.000)	1	1	0.0022 (0.002)	0.0000 (0.000)
16-17						1	1	0.0005 (0.000)	2.0000 (0.000)

	Monthly Runs			Weekly Runs			Daily Runs						
	Survival	Risk	Exit	Survival Rate	Hazard Rate	Risk	Exit	Survival Rate	Hazard Rate	Risk	Exit	Survival Rate	Hazard Rate
0-1	100	62	100	1.0000 (0.000)	1.2836 (0.146)	455	254	1.0000 (0.000)	0.9680 (0.056)	2034	0	1.0000 (0.000)	0.0000 (0.000)
1-2	38	21	38	0.3800 (0.049)	1.0909 (0.220)	201	107	0.4418 (0.023)	0.9068 (0.082)	2034	947	1.0000 (0.000)	0.6069 (0.019)
2-3	17	10	17	0.1700 (0.038)	1.1905 (0.342)	94	50	0.2066 (0.019)	0.9058 (0.119)	1087	524	0.5344 (0.011)	0.6352 (0.026)
3-4	7	3	7	0.0700 (0.026)	0.7792 (0.433)	44	30	0.0967 (0.014)	1.2931 (0.202)	563	272	0.2768 (0.010)	0.6370 (0.037)
4-5	4	2	4	0.0400 (0.020)	0.9524 (0.635)	14	7	0.0308 (0.008)	0.8333 (0.297)	291	147	0.1431 (0.008)	0.6759 (0.052)
5-6	2	1	2	0.0200 (0.014)	0.9524 (0.898)	7	4	0.0154 (0.006)	1.0000 (0.458)	144	64	0.0708 (0.006)	0.5714 (0.068)
6-7	1	1	1	0.0100 (0.010)	2.8571 (0.000)	3	2	0.0066 (0.004)	1.2500 (0.765)	80	43	0.0393 (0.004)	0.7350 (0.104)
7-8						1	1	0.0022 (0.002)	2.5000 (0.000)	37	22	0.0182 (0.003)	0.8462 (0.163)
8-9										15	7	0.0074 (0.002)	0.6087 (0.219)
9-10										8	3	0.0039 (0.001)	0.4615 (0.259)
10-11										5	1	0.0025 (0.001)	0.2222 (0.221)
11-12										4	4	0.0020 (0.001)	2.0000 (0.000)

Figure 1—Nonparametric Hazard and Survival Functions for Positive and Negative Runs

Monthly, weekly and daily CRSP value-weighted returns are obtained for the period from July 1, 1962 to December 31, 1999. The length of a positive run is defined as the number of consecutive positive monthly/weekly/daily returns and the length of a negative run is defined as the number of consecutive negative monthly/weekly/daily returns. Positive and negative run alternates so that there are a total of 99 positive and 100 negative monthly runs; 455 positive and 455 negative weekly runs and 2033 positive and 2034 negative daily runs.

Part A: Monthly Runs**Part B: Weekly Runs****Part C: Daily Runs**

This is evidence of positive duration dependence in runs. Thus, there is a high pressure that a run will end and reverse direction when it lengthens. This suggests that a stock return tends to move towards its permanent or trend level in a non-random or a predictable way. This result is consistent with McQueen and Thorley (1991) who show that excess returns exhibit nonrandom walk tendencies in the sense that low (high) returns tend to follow runs of high (low) returns. In their tests, the reversal of stock return movement is reflected in both the first-order and the second-order Markov transition probability matrix. An earlier study by Niederhoffer and Osborne (1966) produces similar results, but only for the first-order Markov transition matrix, not for the second-order.⁴ Bremer and Sweeney (1991) also find reversal by showing that extremely large negative returns are followed on average by larger than expected positive returns over the following dates. On the other hand, McQueen and Thorley (1994), using a logit model in their study of bubbles, show a negative relation between the probability of a run ending and its length.⁵

I also observe that the p estimates are all higher for negative runs than for positive runs, suggesting that negative runs tend to end more quickly than positive runs, a fact consistent with CRSP returns having a positive drift component. Another general observation is that the p estimates are higher for the CRSP value-weighted index than for the CRSP equal-weighted index, except for the monthly negative runs.⁶ This suggests that runs based on CRSP equal-weighted index in general last longer than those based on CRSP value-weighted index.

Heckman and Singer (1982) point out that in a duration model, the hazard function in general tends to have a downward bias in duration time—the length of a run in my case. This is basically a heterogeneity issue. Individuals with higher hazard ratios, due to their individual-specific characteristics, experience events early and are eliminated from the risk set. As time goes on, this selection process yields risk sets that contain individuals with predominantly low risk. Because of this, it is extremely difficult to distinguish hazard ratios that are truly declining in time from simple variation in hazard ratios across individuals. On the other hand, if one observes an increasing hazard function, this always can be regarded as evidence that the hazard ratio really increases in time (Allison, 1984). Based on this argument, I

⁴ Niederhoffer and Osborne (1966) find that the probabilities of (+|-) and (-|+) are about three times those of (-|-) and (+|+), but the probabilities of (++++) and (----) are about two times those of (+|-) [(+|-)] and (-|+) [-|+]; where “+” and “-” represent a price rise and a price decline, respectively.

⁵ McQueen and Thorley (1994) define duration dependence opposite to mine. They consider the negative relation between the probability of a run ending and its length as evidence of bubbles (p. 379).

⁶ The p estimate (1.6993) for the monthly negative CRSP value-weighted runs is lower than that (1.7352) for the monthly negative CRSP equal-weighted runs.

Table 3—Parametric Analysis (Weibull Density) of Positive and Negative Runs Considering the Magnitudes of Returns

Monthly, weekly and daily CRSP value-weighted returns (VWRET) and equal-weighted returns (EWRET) are obtained for the period from July 1, 1962 to December 31, 1999. The length of a positive run is defined as the number of consecutive positive monthly/weekly/daily returns, and the length of a negative run is defined as the number of consecutive negative monthly/weekly/daily returns. The hazard function for the Weibull density is $\lambda(t) = e^{\lambda t} \lambda t^{\lambda-1}$. * indicates significance level at 1 percent (two-tail test), ** indicates significance level at 5 percent (two-tail test) and *** indicates significance level at 10 percent (two-tail test) except for the estimate for p which involves a one-tail test (testing $p > 1$)

Panel A: CRSP Value-Weighted Index									
		Positive Runs			Negative Runs				
		Constant	VWRET	λ	p	Constant	VWRET	λ	p
Monthly	Estimate	-1.1344*	0.0992	0.3227*	1.4251*	-0.5548*	3.0242	0.5219*	1.6993*
Runs	Standard Error	[0.1572]	[3.2643]	[0.0313]	[0.1801]	[0.1253]	[2.5630]	[0.0526]	[0.2212]
Weekly	Estimate	-1.0530*	2.2430	0.3603*	1.4302*	-0.6439*	4.7360**	0.4900*	1.7345*
Runs	Standard Error	[0.0705]	[3.4335]	[0.0168]	[0.0827]	[0.0552]	[2.5141]	[0.0209]	[0.1059]
Daily	Estimate	-1.0054*	-7.1039**	0.3522*	1.4996*	-0.7809*	14.3197*	0.4232*	1.6120*
Runs	Standard Error	[0.0281]	[3.3466]	[0.0072]	[0.0396]	[0.0245]	[2.2810]	[0.0086]	[0.0426]
Panel B: CRSP Equal-Weighted Index									
		Positive Runs			Negative Runs				
		Constant	VWRET	λ	p	Constant	VWRET	λ	p
Monthly	Estimate	-1.0585*	-7.1036**	0.2653*	1.3350*	-0.6034*	3.8304***	0.4775*	1.7352*
Runs	Standard Error	[0.1835]	[3.2413]	[0.0292]	[0.1671]	[0.1294]	[2.1878]	[0.0469]	[0.2278]
Weekly	Estimate	-1.1004*	-24.9980*	0.2512*	1.1900*	-0.6722*	13.0687*	0.4294*	1.6263*
Runs	Standard Error	[0.0971]	[5.3248]	[0.0158]	[0.0767]	[0.0668]	[2.9635]	[0.0217]	[0.1140]
Daily	Estimate	-1.0654*	-72.6748*	0.2662*	1.2188*	-0.7470*	41.2119*	0.3999*	1.5071*
Runs	Standard Error	[0.0406]	[6.7238]	[0.0075]	[0.0350]	[0.0306]	[3.4281]	[0.0100]	[0.0449]

have some confidence in claiming that the hazard ratio of a run increases in its length; that is, a prolonged run has a high tendency of reversing direction.

As for the time-varying CRSP value-weighted index return that covaries with the hazard function, it does not have a significant impact on the hazard function for the monthly and weekly positive CRSP value-weighted runs. But when I increase the measuring frequency to daily, it starts to impact the hazard function. The coefficient on daily return is -7.1039 at less than 5 percent significance level. As for positive CRSP equal-weighted runs, the coefficients on the equal-weighted index are -7.1036, -24.9980, and -72.6748 at less than 5 percent, 1 percent, and 1 percent significance levels for monthly, weekly, and daily runs, respectively. A negative coefficient on the index return for the positive runs suggests that the time-varying index return contributes negatively to the hazard ratio of a run. This means that even though the persistence level of a positive run tends to abate in its length, it tends to strengthen itself in the magnitude of its components.

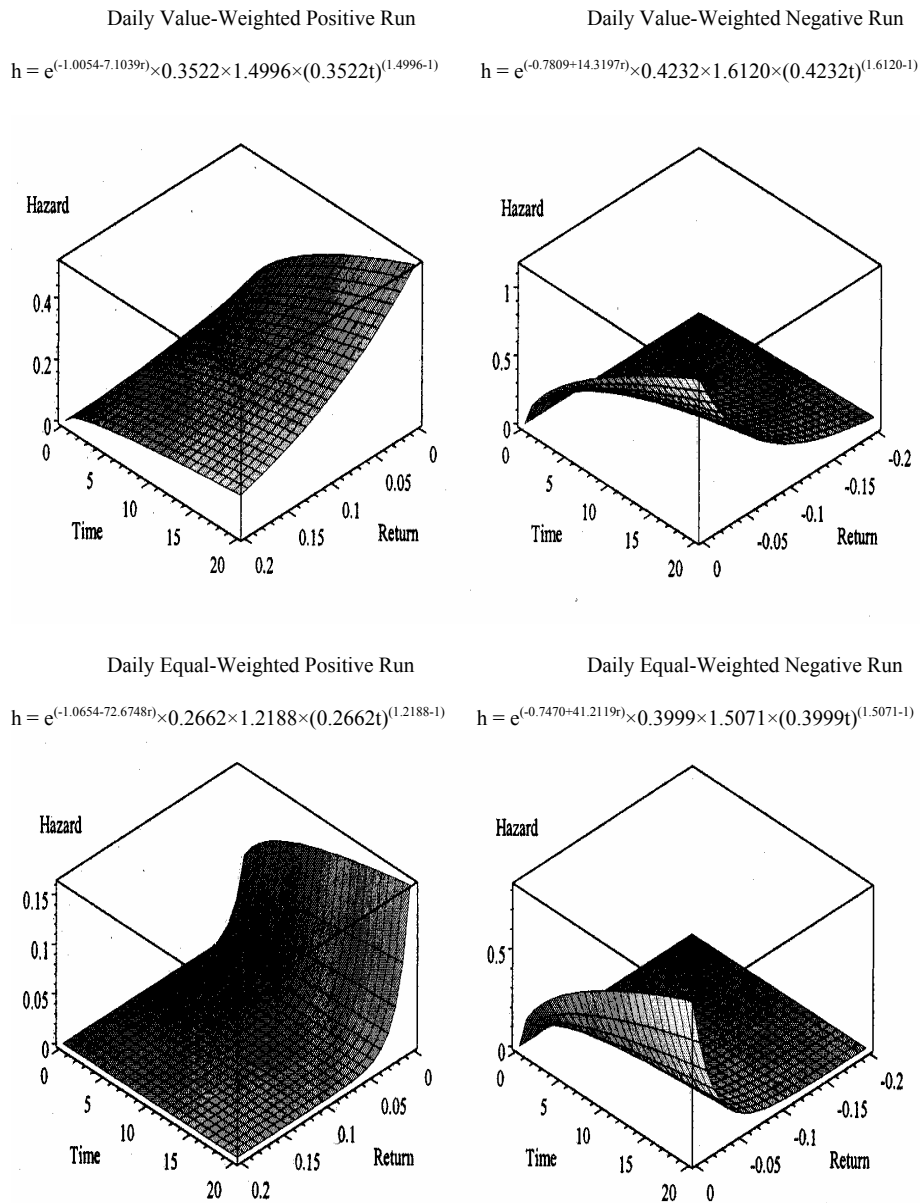
The time-varying index return does not appear to impact significantly the hazard ratio of the monthly negative CRSP value-weighted runs. But when the measuring frequency increases from monthly to weekly and to daily, the significance level of the impact of the time-varying index return increases (to less than 5 percent for the weekly runs to less than 1 percent for the daily runs). This is also true for negative CRSP equal-weighted runs. The coefficients on the CRSP value-weighted and equal-weighted index returns are all positive for negative runs. This means that the more negative the components of a negative run, the lower the hazard ratio and thus the more persistent the run.

The above results suggest that even though the persistence level of a run tends to abate in its length, the magnitude of its components tends to enhance it. See Figure 2 for an illustration based on daily runs. This means that there may be at least two components of stock returns. First, longitudinally, stock return is cyclical and there is a reversal tendency, as in DeBondt and Thaler (1985) and Bremer and Sweeney (1991); second, with the length of a run controlled, there is a positive autocorrelation in stock returns, as in Lo and MacKinlay (1988). This finding may have the potential of explaining why over different holding windows, sometimes momentum trading is profitable and sometimes contrarian trading is profitable.⁷ The question is which

⁷ Momentum trading usually is found to be profitable at intermediate terms, such as holding periods from three months to 12 months; while contrarian trading is usually found to be profitable at very short terms (holding periods of less than three months) or long terms (holding period beyond 12 months). For example, Jegadeesh (1990) finds that contrarian trading based on monthly ranking of stocks generates a monthly profit of 2.49 percent. Jegadeesh and Titman (1993, 2001), Conrad and Kaul (1998), Moskowitz and Grinblatt (1999), among many others, find positive three to 12-month momentum profit. Longer term contrarian profit is documented in De Bondt and Thaler (1985) and also in Jegadeesh and Titman (1993), etc.

Figure 2—Hazard as a Function of CRSP Index Return and Length of a Run

Based on daily runs, the hazard ratio is plotted for value-weighted and equal-weighted positive and negative runs. The estimates for λ and p are taken from Table IV. The hazard is computed as $\lambda(t) = e^{x\beta} \lambda p(\lambda t)^{p-1}$, where x includes a constant term and CRSP index return, t is the length of a run



component dominates the other. For short holding windows, the positive autocorrelation component may dominate mean-reversion component; for long holding windows, the mean-reversion component may dominate the positive autocorrelation component.

Impact of Size and the Book-to-Market Equity Ratio

The impact of the CRSP equal-weighted index return on the hazard function is different from that of the CRSP value-weighted index return. First, start from the lowest measuring frequency, monthly runs, the effect of CRSP equal-weighted index return is significant at less than 5 percent level for positive runs and at less than 10 percent level for negative runs, while that of CRSP value-weighted returns is not significant. When the measuring frequency increases to weekly and daily runs, the impact of the CRSP equal-weighted returns is more significant (with t-statistics at 4.4098 and 12.0217, respectively) than that for the CRSP value-weighted return (with t-statistics at 1.8837 and 6.2778, respectively). Second, the coefficients on CRSP equal-weighted index return (EWRET) are all of greater magnitude than those on the CRSP value-weighted index return (VWRET). Because CRSP equal-weighted return overweighs small capitalization stocks, the above results suggest that the directional movement of small stocks may be more predictable than that of large stocks.

To determine whether this might be the case, I compare the persistence level of runs for large capitalization stocks and small capitalization stocks. If the directional movement of small stocks is more predictable than that of large stocks, then the impact of index return on the hazard function of small stocks should be more significant and perhaps larger in magnitude than that of large stocks. To maximize the effect, I pick only the runs based on the largest quintile (top 20 percent) and smallest quintile (bottom 20 percent) of market capitalization.⁸ Panel A, Table 4 shows the results. For both positive and negative runs, the magnitude and significance level of p are comparable to the results based on the whole universe of CRSP stocks (Table 3). For positive runs, the impact of value-weighted return (VWRET) is significant at less than 10 percent level for small stocks while it is insignificant for large stocks. It is also of greater magnitude for small stocks than for large stocks. I do not observe this difference for negative runs. For the whole universe of CRSP stocks, the impact of monthly VWRET is insignificant for both positive and negative runs. The results based on top and bottom size quintiles offer some limited evidence that the directional movement of small stocks may be more

⁸ Size-based monthly portfolio returns are obtained from Professor Kenneth French. All CRSP stocks are sorted based on their size (market capitalization) for each month during the period from July 1962 to December 1999. Then these stocks are divided into size-based quintiles. Within each quintile, its value-weighted return (VWRET) is computed.

Table 4—The Effect of Size (Market Capitalization) and Book-to-Market Equity Ratio

Every month for the period from July 1962 to December 1999 all CRSP stocks are ranked on size—market capitalization (book-to-market equity ratio). These stocks are then dividend into quintiles according to their size (book-to-market equity ratio). Small capitalization (low book-to-market equity ratio) stocks refer to the bottom 20 percent of stocks, and large capitalization (high book-to-market equity ratio) stocks refer to the top 20 percent of stocks. Within both of the bottom and top size quintiles, value-weighted returns (VWRET) are computed for each month during the period. The length of a positive run is defined as the number of consecutive positive monthly/weekly/daily returns, and the length of a negative run is defined as the number of consecutive negative monthly/weekly/daily returns. The hazard function for the Weibull density is $\lambda(t) = e^{-\theta t} \lambda p(\lambda t)^{p-1}$. * indicates significance level at 1 percent (two-tail test), ** indicates significance level at 5 percent (two-tail test) and *** indicates significance level at 10 percent (two-tail test) except for the estimate for p which involves a one-tail test (testing $p > 1$)

Panel A: The Effect of Size (Market Capitalization)

		Positive Runs				Negative Runs			
		Constant	VWRET	λ	p	Constant	VWRET	λ	p
Large Cap Stocks	Estimate	-1.0780*	-0.2710	0.3372*	1.4272*	-0.4939*	2.9102	0.5601*	1.6939*
	Standard Error	[0.1452]	[3.0195]	[0.0324]	[0.1794]	[0.1260]	[2.4955]	[0.0616]	[0.2211]
	Estimate	-0.9935*	-5.0160***	0.3028*	1.4170*	-0.7647*	2.0525	0.4274*	1.7569*
	Standard Error	[0.1602]	[2.6857]	[0.0305]	[0.1842]	[0.1137]	[1.9214]	[0.0375]	[0.2391]

Panel B: The Effect of Book-to-Market Equity Ratio

		Positive Runs				Negative Runs			
		Constant	VWRET	λ	p	Constant	VWRET	λ	p
Low B-to-M Stocks	Estimate	-1.0235*	0.3133	0.3638*	1.3777*	-0.5466*	3.1045	0.5217*	1.5659*
	Standard Error	[0.1569]	[2.6689]	[0.0381]	[0.1744]	[0.1338]	[2.5223]	[0.0582]	[0.1594]
	Estimate	-1.1879*	-1.9554	0.2847*	1.4400*	-0.4793*	3.9998***	0.5435*	1.7306*
	Standard Error	[0.1557]	[3.1977]	[0.0270]	[0.1799]	[0.1364]	[2.3180]	[0.0563]	[0.2183]

predictable than that of large stocks, consistent with the existing literature (such as Fama and French, 1988).

Apart from firm size, the book-to-market equity ratio frequently is found to explain part of the return anomaly in many empirical studies (Fama and French, 1992, 1993, 1995, and 1996). Here, I check whether it has any impact on the directional movement of index return. In the test, low book-to-market equity stocks refer to the bottom 20 percent of stocks, and high book-to-market equity stocks refer to the top 20 percent of stocks.⁹ The results are shown in Panel B, Table 4. The estimates for λ and p are of similar magnitudes as in earlier tests. The impact of value-weighted index return (VWRET) on the hazard function is only significant (at 10 percent level) for negative runs of high book-to-market equity stocks. The role played by book-to-market equity, while consistent with my conjecture, appears to be relatively limited.

Impact of Interest Rate

The interest rate affects stock price. When interest rate goes down (up), stock price goes up (down). It is possible that the interest rate can impact the persistence of a run directly and/or through affecting the impact of its components. Because interest rate is negatively related to stock price, it should have a positive (negative) impact on the hazard ratio of a positive (negative) run. That is, an increase in interest rate should cause a positive run to end sooner and a negative run to last longer. To determine whether this is the case, I include interest rate as an exogenous variable. I obtain three-month T-bill rates from Federal Reserve publications for the period from July 1, 1962 to December 31, 1999 and match them with the run data. I perform tests on monthly and weekly runs.

The results are shown in Table 5. First, based on the significant estimates for λ and p , the presence of positive duration dependence still is unchanged when market interest rate is taken into account. Second, market interest rate has a positive effect on the hazard ratio of a positive run and a negative effect on the hazard ratio of a negative run, consistent with my conjecture; in three occasions (positive weekly value-weighted runs, positive weekly equal-weighted runs, and negative monthly value-weighted runs) the effect is significant. The market interest rate decreases the persistence level of a positive run and increases the persistence level of a negative run. Third, the inclusion of the market interest rate somewhat reduces the impact of the components of runs. In two occasions (negative weekly value-weighted runs and negative monthly equal-weighted runs) the components of runs lose their signifi-

⁹ Book-to-market equity based monthly portfolio returns are obtained from Professor Kenneth French. Every month for the period from July 1962 to December 1999, all CRSP stocks are ranked based on their book-to-market equity ratio. These stocks then are divided into quintiles according to this ratio. Within each quintile, its value-weighted return (VWRET) is computed.

Table 5—The Persistence of Runs and the Market Interest Rate

Monthly and weekly CRSP value-weighted returns (VWRET) and equal-weighted returns (EWRET) are obtained for the period from July 1, 1962 to December 31, 1999. Monthly and weekly T-bill rates (TBILLS) are obtained from the Federal Reserve Publication for the same period. The length of a positive run is defined as the number of consecutive positive monthly/weekly/daily returns, and the length of a negative run is defined as the number of consecutive negative monthly/weekly/daily returns. The hazard function for the Weibull density is $\lambda(t) = e^{\theta_0} \lambda p (\lambda t)^{p-1}$. * indicates significance level at 1 percent (two-tail test), ** indicates significance level at 5 percent (two-tail test) and *** indicates significance level at 10 percent (two-tail test) except for the estimate for p which involves a one-tail test (testing $p > 1$)

Panel A: CRSP Value-Weighted Index											
Positive Runs						Negative Runs					
Monthly Runs	Estimate	Constant	EWRET	TBILLS	λ	p	Constant	EWRET	TBILLS	λ	p
Weekly Runs	Estimate	-1.2969*	-0.3397	36.1248	0.3234*	1.4294*	-0.1742*	2.1745	-78.7090**	0.5262*	1.7628*
	Std. Error	[0.2595]	[3.3076]	[45.6435]	[0.0314]	[0.1775]	[0.2356]	[2.5264]	[39.6345]	[0.0509]	[0.2132]
	Estimate	-1.2645*	1.2513	194.7976**	0.3608*	1.4490*	-0.5347*	4.0817	-101.3009	0.4900*	1.7470*
	Std. Error	[0.1147]	[3.4526]	[87.7531]	[0.0165]	[0.0845]	[0.0940]	[2.5920]	[70.2891]	[0.0206]	[0.1106]
Panel B: CRSP Equal-Weighted Index											
Positive Runs						Negative Runs					
Monthly Runs	Estimate	Constant	EWRET	TBILLS	λ	p	Constant	EWRET	TBILLS	λ	p
Weekly Runs	Estimate	-1.4314*	-7.1341**	74.1592	0.2671*	1.3677*	-0.4265***	3.4117	-36.0198	0.4787*	1.7477*
	Std. Error	[0.2969]	[3.1958]	[47.9803]	[0.0285]	[0.1721]	[0.2399]	[2.2866]	[41.6483]	[0.0470]	[0.2238]
	Estimate	-1.5283*	-25.7191*	372.8017*	0.2533*	1.2186*	-0.5756*	12.6412*	-84.6996	0.4297*	1.6302*
	Std. Error	[0.1542]	[5.2236]	[112.4433]	[0.0155]	[0.0776]	[0.1211]	[2.9780]	[86.6810]	[0.0216]	[0.1138]

cance. So, in general, the market interest rate plays a role that is in the expected direction.

Trading Strategies and Economic Significance

The above analyses suggest that when the length of a run increases, it tends to reverse and change direction. Based on this result, it is possible to devise some trading strategies to exploit this empirical regularity and make trading profits. Some simple and intuitive strategies would be the following: When a positive run is observed to have lasted n periods, an investor should short the CRSP index and long the one-month T-bills at the beginning of period $n + 1$; when a negative run is observed to have lasted n periods, an investor should long the CRSP index and short the one-month T-bills at the beginning of period $n + 1$. Note that these two strategies are both hedged strategies because the net investment outlay at the beginning of period $n + 1$ is always zero.

I examine whether the above simple strategies can be profitable using out-of-sample post-examination daily run data from January 1, 2000 to September 30, 2004. Because this post-examination period is relatively short, I do not use low frequency monthly and weekly runs. If I use monthly and weekly runs, the number of observations may be too low to offer reasonable statistical tests for economic significance. I try run length up to five days for positive and negative CRSP value-weighted and equal-weighted daily runs. I find one strategy that is significantly profitable. If a CRSP value-weighted negative run has lasted for five days, an investor longing the CRSP value-weighted index and shorting one-month T-bills at the beginning of the sixth day will make a significant profit of 1.11 percent on that day (t -value = 2.36) because a negative five-day CRSP value-weighted run tends to reverse direction on the sixth day. There are only 13 negative daily runs that last at least five days during the 1,192-trading-day period between January 1, 2000 and September 30, 2004, however. Assume that there are 252 trading days during a year, annual return based on this strategy is $(1.11 \text{ percent} \times 13 / 1192) \times 252 = 3.05$ percent. This hedged annual profit is moderately economically significant.

The main purpose of this paper is to test the random walk hypothesis. I provide evidence that CRSP index returns are path-dependent and reject the random walk hypothesis. It is not my main intention to show how the observed path-dependency can be translated into profitable trading strategies. While the above analysis suggests such possibilities, a great deal of caution should be used in applying such strategies.

Summary and Conclusion

I test the random walk hypothesis of stock returns by studying the hazard ratio of runs based on CRSP index returns. A run is defined as a sequence of consecutive returns of the same sign with observed starting and ending points. If the stock price

follows a random walk process, then the hazard ratio or the conditional probability of a run ending and reversing direction should not be path-dependent.

The results reject the random walk hypothesis and suggest that the directional movement of the market index return is predictable (path-dependent) in the following manners. First, the persistence of a run tends to abate in its length, suggesting some mean reversion tendency in returns. This is evidence of positive duration dependence in runs. At the same time, the persistence of a run tends to enhance itself in the magnitude of its components (historical return realizations). This suggests that there may be two components of stock returns. Longitudinally, stock return is cyclical and there is a reversal tendency; with the length of the run controlled, there is a positive autocorrelation in stock return. This finding may have the potential of explaining why over different holding windows momentum trading is profitable sometimes while sometimes contrarian trading is profitable. The question is which component dominates the other.

Second, a positive run generally lasts longer than a negative run, implying a positive drift term in returns. Third, the impact of CRSP index return on the persistence of a run becomes more pronounced when the measuring frequency increases. Fourth, the impact of CRSP equal-weighted index return on the persistence of a run is more pronounced than that of CRSP value-weighted index return, suggesting that the directional movement of small stocks is possibly more predictable. This also may be true for high book-to-market equity stocks, though in a very limited sense. Fifth, the market interest rate decreases the persistence level of a positive run and increases the persistence level of a negative run. While my additional analysis suggests that there might be some moderately profitable trading strategies based on the empirical regularities presented in this paper, a great deal of caution should be used in applying these strategies.

References

1. Allison, P.D., "Event History Analysis: Regression for Longitudinal Event Data," *Sage University Papers Series: Quantitative Applications in the Social Sciences* (1984)
2. Bremer, M., and R.J. Sweeney, "The Reversal of Large Stock-Price Decreases," *Journal of Finance*, 46 (1991), pp. 747-754.
3. Conrad, J., and G. Kaul, "An Anatomy of Trading Strategies," *Review of Financial Studies*, 11 (1998), pp. 489-519.
4. Cox, D. R., "Regression Models and Life-Tables," *Journal of the Royal Statistical Society, Series B* 34 (1972), pp. 187-220.
5. Cox, D.R., "Partial Likelihood," *Biometrika*, 62 (1975), pp. 369-276.
6. DeBondt, W.F.M., and R. Thaler, "Does the Stock Market Overreact?" *Journal of Finance*, 40 (1985), pp. 793-805.
7. Diebold, F.X., J. Lee, and G.C. Weinbach, "Regime Switching with Time-varying Transition Probabilities," in C. Hargreaves (ed.), *Nonstationary Time-Series Analysis and Cointegration* (Oxford University Press, 1994), pp. 283-302.

8. Diebold, F.X., and G.D. Rudebusch, "Measuring Business Cycles: A Modern Perspective," *Review of Economics and Statistics*, 78 (1996), pp. 67-77.
9. Fama, E.F., "The Behavior of Stock-Market Prices," *Journal of Business* 38, (1965), pp. 34-105.
10. Fama, E.F., "Efficient Capital Market: A Review of Theory and Empirical Work," *Journal of Finance*, 25 (1970), pp. 383-417.
11. Fama, E.F., and K.R. French, "Permanent and Temporary Components of Stock Prices," *Journal of Political Economy*, 96 (1988), pp. 246-273.
12. Fama, E.F., and K.R. French, "The cross-section of expected stock returns," *Journal of Finance*, 47 (1992), pp. 427-465.
13. Fama, E.F., and K.R. French, "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33 (1993), pp. 3-56.
14. Fama, E.F., and K.R. French, "Size and Book-to-market Factors in Earnings and returns," *Journal Finance*, 50 (1995), pp. 131-155.
15. Fama, E.F., and K.R. French, "Multifactor Explanations of Asset Pricing Anomalies," *Journal of Finance*, 51 (1996), pp. 55-84.
16. Filardo, A.J., "Business-cycle Phases and their Transition Dynamics," *Journal of Business and Economic Statistics*, 12 (1994), pp. 299-308.
17. Grinblatt, M., and S. Titman, "The Persistence of Mutual Fund Performance," *Journal of Finance*, 47 (1992), pp. 1977-1984.
18. Han, A., and J.A. Hausman, "Flexible Parametric Estimation of Duration and Competing Risk Models," *Journal of Applied Econometrics*, 5 (1990), pp. 1-28.
19. Heckman, J.J., and B. Singer, "The Identification Problem in Econometric Models for Duration Data," in Werner Hildenbrand (ed.), *Advances in Econometrics* (Cambridge University Press, 1982).
20. Jegadeesh, N., "Evidence of Predictable Behavior of Security Returns," *Journal of Finance*, 45 (1990), pp. 881-898.
21. Jegadeesh, N., and S. Titman, "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *Journal of Finance*, 48 (1993), pp. 65-91.
22. Jegadeesh, N., and S. Titman, "Overreaction, Delayed Reaction, and Contrarian Profits," *Review of Financial Studies*, 8 (1995), pp. 973-993.
23. Jegadeesh, N., and S. Titman, "Profitability of Momentum Strategies: An Evaluation of Alternative Explanations," *Journal of Finance*, 56 (2001), pp. 699-720.
24. Kalbfleisch, J.D., and R.L. Prentice, *The Statistical Analysis of Failure Time Data* (New York: John Wiley, 1980).
25. Kaplan, E.L., and P. Meier, "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association*, 53 (1958), pp. 457-481.
26. Kennan, J., "The Duration of Contract Strikes in U.S. Manufacturing," *Journal of Econometrics*, 28 (1985), pp. 5-28.
27. Kiefer, N.M., "Economic Duration Data and Hazard Functions," *Journal of Economic Literature*, 26 (1988), pp. 646-679.
28. Lancaster, T., "Econometric Methods for the Duration of Unemployment," *Econometrica*, 47 (1979), pp. 939-956.
29. Lo, A.W., and A.C. MacKinlay, "Stock Market Prices Do Not Follow Random Walks: Evidence From a Simple Specification Test," *Review of Financial Studies*, 1 (1988), pp. 41-66.

30. Lo, A.W., H. Mamaysky, and J. Wang, "Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation," *Journal of Finance*, 55 (2000), pp. 1705-1765.
31. McGinnis, R., "A Stochastic Model of Social Mobility," *American Sociological Review*, 33 (1968), pp. 712-722.
32. McQueen, G., and S. Thorley, "Are Stock Returns Predictable?" A Test Using Markov Chains, *Journal of Finance*, 46 (1991), pp. 239-263.
33. McQueen, G., and S. Thorley, "Bubbles, Stock Returns, and Duration Dependence," *Journal of Financial and Quantitative Analysis*, 29 (1994), pp. 379-401.
34. Mood, A.M., "The Distribution Theory of Runs," *Annals of Mathematical Statistics*, 11 (1940), pp. 367-392.
35. Moskowitz, T.J., and M. Grinblatt, "Do Industries Explain Momentum?" *Journal of Finance*, 54 (1999), pp. 1249-1290.
36. Neftci, S.N., "Are Economic Time Series Asymmetric Over the Business Cycle?" *Journal of Political Economy*, 92 (1984), pp. 307-328.
37. Niederhoffer, V., and M.F.M. Osborne, "Market Making and Reversal on the Stock Exchange," *Journal of the American Statistical Association*, 61 (1966), pp. 897-916.
38. Petersen, T., "Fitting Parametric Survival Models with Time-dependent Covariates," *Applied Statistics*, 35 (1986), pp. 281-288.
39. Poterba, J.M., and L.H. Summers, "Mean Reversion in Stock Prices: Evidence and Implications," *Journal of Financial Economics*, 22 (1988), pp. 27-59.
40. Sichel, D.E., "Business Cycle Duration Dependence: A Parametric Approach," *Review of Economics and Statistics*, 73 (1991), pp. 254-260.

Appendix A—Dealing with Time-Varying Covariates

Following Petersen (1986) whose procedure is more general, I use a special case of his methodology. Because I use equal-interval return data, the length of a run is an integer (in months, weeks, or days) and its increment is exactly one period (a month, a week or a day, depending on the measuring frequency). Suppose the length of a run is t periods (months, weeks, or days), let $j = 1, \dots, k$ index these periods, I have $t_1 = 1$ (starting point), $t_k = t$ (ending point), and $t_j - t_{j-1} = 1$ (unit increment for $j > 1$). From equation (10), the conditional survival function (for surviving between t_{j-1} and t_j) is

$$S[t_j | x(t_j)] = F[T \geq t_j | T \geq t_{j-1}, x(t_j)] = \exp\left\{-\int_{t_{j-1}}^{t_j} \lambda[s | x(t_j)] ds\right\}. \quad (A.1)$$

Because $t_j - t_{j-1} = 1$ and $x(t_j)$ is constant between t_{j-1} and t_j (period-end return data), the integration sign can be taken away, the conditional survival function (A.1) becomes

$$S[t_j | x(t_j)] = F[T \geq t_j | T \geq t_{j-1}, x(t_j)] = \exp\{-\lambda[t_j | x(t_j)]\}. \quad (A.2)$$

So, the survival function (for surviving beyond t periods) is

$$S[t | x(t_1), \dots, x(t_k)] = \prod_{j=1}^k \exp\{-\lambda[t_j | x(t_j)]\} = \exp\left\{-\sum_{j=1}^k \lambda[t_j | x(t_j)]\right\}. \quad (A.3)$$

From equation (9), the Weibull distribution can be written as

$$f_i[t_{ik} | x(t_{i1}), \dots, x(t_{ik})] = [\lambda(t_{ik} | x(t_{ik}))] S[t_{ik} | x(t_{i1}), \dots, x(t_{ik})]. \quad (A.4)$$

The maximum likelihood function is thus

$$ML = \prod_{i=1}^n f_i[t_{ik} | x(t_{i1}), \dots, x(t_{ik})] = \prod_{i=1}^n [\lambda(t_{ik} | x(t_{ik}))] S[t_{ik} | x(t_{i1}), \dots, x(t_{ik})], \quad (A.5)$$

where $i = 1, \dots, n$ indexes runs (positive or negative) in the sample. Substitute (A.3) into (A.5), the logarithm of the maximum likelihood function is

$$\ln(ML) = \sum_{i=1}^n \ln[\lambda(t_{ik} | x(t_{ik}))] - \sum_{i=1}^n \sum_{j=1}^k \lambda[t_{ij} | x_i(t_{ij})]. \quad (A.6)$$

Maximize the above logarithm of the maximum likelihood function with respect to λ , p , and β ; their estimated values then can be obtained.¹⁰

¹⁰ For more understanding of the duration model, refer to Kaplan and Meier (1958), McGinnis (1968), Cox (1972, 1975), Lancaster (1979), Kalbfleisch and Prentice (1980), Heckman and Singer (1982), Allison (1984), Kennan (1985), Petersen (1986), Kiefer (1988), Han and Hausman (1990), etc.

Appendix B—Life Table Product-Limit (PL) Approach (Kaplan and Meier, 1958)

I first categorize runs according to their length and use j to index each category based on length of runs [$j = 1, \dots, \max(\text{length of runs})$]. The size of the risk set (sequences of equal-signed index returns that have not reversed their original directions) of length j is denoted r_j and the number of sequences that actually exit the risk set j (those sequences that reverse directions) in the subsequent period is denoted m_j . The proportion of sequences in the risk set that exits (runs that end and reverse direction) is

$$q_j = \frac{m_j}{r_j} . \quad (\text{B.1})$$

The survival function (the cumulative proportion of observations surviving to the beginning of the interval j) is

$$p_j = (1 - q_{j-1})p_{j-1} , \quad (\text{B.2})$$

where $p_1 = 1$. The hazard rate is

$$\lambda_j = \frac{2q_j}{2 - q_j} . \quad (\text{B.3})$$

The variance for the survival function at time interval j is

$$\text{Var}(p_j) = p_j^2 \sum_{i=1}^j \frac{q_i}{r_i(1 - q_i)} , \quad (\text{B.4})$$

where $i = 1, \dots, j$ indexes time from 1 to j . The variance for the hazard function at time interval j is

$$\text{Var}(\lambda_j) = \lambda_j^2 \frac{1 - \lambda_j^2/4}{r_j q_j} . \quad (\text{B.5})$$

Copyright of Quarterly Journal of Business & Economics is the property of College of Business Administration/Nebraska. The copyright in an individual article may be maintained by the author in certain cases. Content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.