

Endogenous Events and Long-Run Returns

S. Viswanathan

Duke University

Bin Wei

Zicklin School of Business, Baruch College, City University of New York

We analyze event abnormal returns when returns predict events. In fixed samples, we show that the expected abnormal return is negative and becomes more negative as the holding period increases. Asymptotically, abnormal returns converge to zero provided that the process of the number of events is stationary. Nonstationarity in the process of the number of events is needed to generate a large negative bias. We present theory and simulations for the specific case of a lognormal model to characterize the magnitude of the small-sample bias. We illustrate the theory by analyzing long-term returns after initial public offerings (IPOs) and seasoned equity offerings (SEOs). (*JEL* G14, G32)

Since the pioneering research of Fama *et al.* (1969), event studies have been used to conclude that markets are semistrong efficient and deviations from market efficiency are small. In an influential paper, Ritter (1991) challenges this view by focusing on the long-run performance of initial public offerings (IPOs) and argues that event returns following IPOs over longer horizons are large and negative on a risk-adjusted basis. Subsequently, Loughran and Ritter (1995) present results on the long-run underperformance of seasoned equity offerings (SEOs) and IPOs; Loughran and Vijh (1997) examine underperformance in stock returns following mergers; Michaely, Thaler, and Womack (1995) investigate reactions of stock prices to dividend omissions; Ikenberry, Lakonishok, and Vermaelen (1995) examine overperformance after open market share repurchases. Fama (1998) surveys the large literature on long-run returns and argues that the power of these tests of market efficiency is low (see also Brav, 2000; Barber and Lyon, 1997; and Mitchell and Stafford, 2000).

More recently, Schultz (2003) has argued that this long-run performance is spurious when returns predict events. The assumption that returns predict events differs from the traditional assumption in event studies. For example, in the standard textbook discussion of event studies (see Campbell, Lo, and

Comments from an anonymous referee, Joel Hasbrouck, and Ravi Jagannathan have substantially improved the paper. We also thank Federico Bandi, Ravi Bansal, Geert Bekaert, Michael Brandt, Alon Brav, Michael Brennan, Ron Gallant, Larry Harris, Raymond Kan, Pete Kyle, Kai Li, Nagpurnanand Prabhala, Ehud Ronn, Ronnie Sadka, Paul Schultz, Ken Singleton, Bhaskar Swaminathan, and seminar participants at Boston College, Duke, Northwestern, Notre Dame, the 2004 EFA meeting at Maastricht, the 2004 FEA conference at USC, and the 2005 WFA conference at Portland, Oregon, for their comments and discussions. All errors are our own. Address correspondence to S. Viswanathan, Fuqua School of Business, Duke University, Durham, NC 27708; telephone: 919-660-7784; e-mail: viswanat@mail.duke.edu.

MacKinlay, 1997, p. 157) this issue is recognized: “Thus the methodology implicitly assumes that the event is exogenous with respect to the change in the market value of the security. . . . There are examples where an event is triggered by a change in the market are ellipses set on the same level as other text? they look higher value of the security, in which case the event is endogenous. For these cases, the usual interpretation is incorrect.” Following Schultz (2003), we believe that for a large class of events, the event-generating process depends on the past history of event returns and thus events are endogenous. One main result of this paper is to provide a fixed sample and asymptotic theory for event studies with endogenous events.

This assumption of event endogeneity is reasonable given many theoretical models in corporate finance. For example, Lucas and McDonald (1990) show that seasoned equity issues are more likely to be preceded by stock price increases. Pastor and Veronesi (2005) predict that managers will time IPOs when the stock market is doing well. Rhodes-Kropf and Viswanathan (2004) show that mergers occur when markets are relatively overvalued. All these models suggest that corporate events occur more often when event returns are higher.

Our small-sample theory shows that when returns predict events, the long-run event-abnormal return will be negative. The intuition for the negative event-abnormal return is as follows. A priori, we expect that all event return histories that are equally likely in the data will be weighted identically in event-abnormal returns. While calculating long-run event-abnormal returns, we show that we overweight histories with a higher number of events and we underweight histories with a lower number of events. When event returns are high, subsequently the number of events is greater, hence the denominator of the event-abnormal return, which is the total number of events, is higher. This implies that we underweight the high returns. The opposite argument holds when the number of events is lower; here, we overweight the subsequent low returns. Consequently, the event-abnormal return has negative expectation. With a longer holding period, the underweighting of return histories is exacerbated because long-term event returns involve a sequence of returns; a sequence of high returns implies even more events in the future. Thus a sequence of high returns is underweighted much more compared to a sequence of low returns. This yields that the negative expectation of event-abnormal returns increases in absolute magnitude with the holding period used to measure event returns, and the negative bias is larger in absolute magnitude with long-run event returns.

While this argument proves that the expected long-run event-abnormal returns are negative in a fixed sample, it says little about the asymptotic theory of long-run event-abnormal returns. Schultz (2003) suggests via an example that when market levels predict the number of events, the long-run return averaged across all simulations is negative. Because Schultz’s example consists only of a simulation, the condition under which his results can be obtained is unclear. We fill in this gap by showing that asymptotically, post-event returns converge to zero under the following sufficient condition: returns and event process are

stationary and the cumulative number of events converges to infinity. This argument suggests that the negative expected long-run event-abnormal return is a small-sample problem unless one believes that the process of the number of events is nonstationary.

The main intuition for our asymptotic result is as follows. Stationarity in the event process implies that the shocks to the event process do not persist forever. Consequently, the total number of events in a very large sample is not affected by the shock to the number of events today. A higher return today implies more events in the near future, but the long-run average number of events is not affected. Asymptotically, we do not underweight high returns and overweight low returns. Of course, this argument is a large-sample argument, and in small samples the bias could be large. The stationarity assumption is important. With a nonstationary process of the number of events, a shock to the number of events today would persist forever, hence the total number of events in a large sample would be affected by the shock to returns today. Our asymptotic theory explains the simulations in Schultz (2003). We show that Schultz's motivating example implicitly assumes nonstationarity in the process of the number of events and his empirical work is based on a unit root specification.

To understand the importance of the nonstationarity assumption, we study a special case in which the log of the number of events is a linear function of the log of the number of events in the previous period and the lagged excess return. In this case, we show that when the autoregressive coefficient on the lagged log number of events is less than 1 (i.e., the process is stationary), convergence to zero occurs in theory and in simulation. In contrast, when the autoregressive coefficient on the lagged log number of events is equal to 1 (i.e., the process contains a unit root), the expected value converges to a negative number in theory and in simulation.¹

Our exact small-sample expected bias calculation shows that at the usual sample size (400 observations), the negative bias is very sensitive to the presence of a unit root. Even small deviations from the unit root hypothesis lead to sharp drops in the bias. With three-year returns, our unit root results are similar to that obtained in Schultz (2003). With an autoregressive coefficient of 0.95, the magnitude of the bias is around one-eighth of that obtained in Schultz (2003). Thus the bias is very sensitive to the unit root hypothesis. We consider alternative weighting schemes that could potentially reduce the bias. These schemes involve weighting the number of events in a given period so as to make the adjusted number of events a more stationary process. Our results suggest that these approaches do substantially reduce the bias but increase the standard error, i.e., there is a trade-off.

We study whether the data generating process for the number of IPOs and SEOs is nonstationary. With one lag in the autoregression, we can reject the unit

¹ The asymptotic theory of functions of unit-root processes depends on the exact function of the integrated process that is considered. The sensitivity of the asymptotic theory to the nature of the nonlinear function is well known; see Park and Phillips (2001) for more details.

root hypothesis for both IPOs and SEOs. With more lags in the autoregression the evidence is mixed. We are unable to reject the null hypothesis of unit root at the 1% level but reject it at the 5% level. In general, it is more difficult to reject the unit root hypothesis for IPOs. Since the null hypothesis is the unit root and the power of unit root tests is low with higher lag lengths, we believe that the data cannot discriminate between the unit root hypothesis and the near-unit-root alternative.

We also consider how confidence intervals are affected by the presence of endogenous events. We consider an extension of our model that allows for correlations between individual event-abnormal returns. We find that even when the correlation is small, the standard deviation of the long-run post-event-abnormal return increases dramatically. While Mitchell and Stafford (2000) have pointed out that the correlation between event-abnormal returns increases the standard deviation (in the context of calendar time regressions), a second effect occurs with endogenous returns that increases the standard deviation much further. Because the number of events is endogenous, persistence in the number of events increases the standard deviation of long-run event returns further. This suggests that the size of tests assumed in event studies is incorrect and that inferences from long-run event-abnormal returns are very noisy.

Our paper is organized as follows. Section 1 presents the general model, studies the bias for fixed sample sizes, and provides the asymptotic theory. Section 2 presents an application of the theory to a lognormal model and considers an alternative weighting scheme. Section 3 considers the prior work of Schultz (2003) and its relation to our work. Section 4 allows for cross-correlations between individual firm event-abnormal returns and derives the asymptotic standard errors. Section 5 considers related work while Section 6 concludes.

1. General Model

1.1 The setup

Consider the following model. Let $r_{m,t}$ be the market return and $r_{IPO,t}$ be the return on an event (here IPOs) index. While our model is not specific to IPOs, we use IPOs to be concrete. Let N_t be the number of IPOs in the end of period t . In our empirical work, we consider both IPOs and SEOs and our time interval t is a month. Let I_{t-1} be the information of an investor or an econometrician at the end of time $t - 1$. We make the following assumptions.

Assumption 1: $r_{m,t}$, $r_{IPO,t}$ are temporally independent with $E[r_{m,t}] = E[r_{IPO,t}]$; N_t is conditionally independent (given the history) of $r_{m,t}$, $r_{IPO,t}$.

Note that N_t depends only on the market returns until time $t - 1$ and on the IPO index returns until time $t - 1$. Hence N_t has no predictive power for

current or future market or IPO returns. This aspect of Assumption 1 states the null hypothesis of market efficiency.

Assumption 2: The event process $\{N_t\}$ is Markovian. To be more specific, $f(N_t|I_{t-1}) = f(N_t|N_{t-1}, r_{m,t-1}, r_{IPO,t-1})$. $f(N_t|I_{t-1})$ denotes the conditional density function of N_t given I_{t-1} . $f(N_t|N_{t-1}, r_{m,t-1}, r_{IPO,t-1})$ is similarly defined.

Assumptions 1 and 2 can be relaxed substantially. Correlations between market return and the IPO returns can be allowed. More complicated dependence than that considered in Assumption 2 can be allowed. None of this would change our results, and the notation would be more cumbersome.

Assumption 3: $f(N_t|N_{t-1}, r_{m,t-1}, r_{IPO,t-1})$ satisfies the affiliation inequality or generalized monotone likelihood ratio inequality.^{2,3}

Affiliation is a stronger dependence concept than correlation. Affiliation requires not only that two random variables, X and Y , are positively correlated, but it also requires that all positive monotone transformations of X and Y are positively correlated conditional on any history. Affiliation captures the idea of conditional positive dependence in the sense that when we see higher values of a monotone function of X , we will see higher values of monotone functions of Y conditional on any history. Generally, positively correlated random variables need not be affiliated, just as random variables with zero correlation need not be independent. Since event returns are nonlinear transformations of the number of events and IPOs returns, we need a stronger notion of dependence than correlation; affiliation is the appropriate concept of positive dependence. A weaker notion than affiliation is association. As a matter of fact, affiliation implies association conditional on any history.⁴

Assumption 3 states that higher values of lagged variables in the information set (such as market returns, IPO returns, and lagged number of events) lead to a larger number of events in the next period. From Milgrom and Weber (1982), we know that this statistical restriction embodies the idea that the number of events is monotonically increasing in lagged IPO returns. Assumption 3 provides restrictions that do not depend on specific distributional assumptions; these restrictions result in a negative bias in long-run event studies.

Based on these three assumptions, using the decomposition of joint probability distributions into conditional and marginal probability distributions, we prove that any subset of the histories of $N_t, r_{m,t}$ and that of $r_{IPO,t}$ are affiliated.

² With continuous random variables where a joint probability density exists, a vector $\tilde{z}$ is affiliated if $f(z' \vee z)f(z' \wedge z) \geq f(z)f(z')$, where $z' \vee z$ is the componentwise maximum for the vector z and $z' \wedge z$ is the componentwise minimum for z . Affiliation is defined more generally in Milgrom and Weber (1982) to allow for discrete and continuous variables. Since we may have discrete variables (the number of events N_t), the definition needs to be amended slightly. The Appendix provides more details.

³ From Assumptions 1 and 3, it follows that the joint density function $f(N_t, N_{t-1}, r_{m,t-1}, r_{IPO,t-1})$ also satisfies the affiliation inequality.

⁴ Variables $X_1, \dots, X_k$ are called to be associated if $\text{cov}[f(X_1, \dots, X_k), g(X_1, \dots, X_k)] \geq 0$ holds for each pair of bounded Borel measurable nondecreasing functions f and g .

Theorem 1. *The random variables $(N_t, \dots, N_{t+s}, r_{m,t}, \dots, r_{m,t+s}, r_{IPO,t}, \dots, r_{IPO,t+s})$ are affiliated.*

Proof. See Appendix. ■

The theorem states that higher values of returns today not only imply higher events tomorrow but also higher events in the future, i.e., the returns predict events not just tomorrow but also farther into the future. This fact has implications for the bias in expected long-run returns in event studies.

1.2 Small-sample theory

We first define in Equation (1) the average cumulative abnormal return and average buy-and-hold abnormal return of s holding periods as:

$$\begin{aligned} \overline{CAR}_T(s) &= \frac{\sum_{t=1}^T N_t \left(\sum_{j=1}^s ((1 + r_{IPO,t+j}) - (1 + E[r_{IPO,t+j}])) \right)}{\sum_{t=1}^T N_t} \quad (1) \\ \overline{BHAR}_T(s) &= \frac{\sum_{t=1}^T N_t \left(\prod_{j=1}^s (1 + r_{IPO,t+j}) - \prod_{j=1}^s (1 + E[r_{IPO,t+j}]) \right)}{\sum_{t=1}^T N_t}, \end{aligned}$$

where $E[r_{IPO,t}]$ is the expected return of a benchmark IPO index for period t , CAR is cumulative abnormal returns, and BHAR is buy-and-hold abnormal returns. These are the most standard definitions used in the literature (see Ritter, 1991; Kothari and Warner, 1997; Campbell, Lo, and MacKinlay, 1997; Barber and Lyon, 1997; Lyon, Barber, Tsai, 1999; Schultz, 2003; Li and Prabhala, 2007; Kothari and Warner, 2007). Thus, the specification considered here uses the return on an IPO index.

Our first theorem shows that the expected cumulative abnormal return is negative in a fixed sample. In proving this theorem, we impose no assumption on the stationarity of returns or number of events.

Theorem 2. $E[\overline{CAR}_T(s)] \leq 0$ and $E[\overline{BHAR}_T(s)] \leq 0, \forall s$.

Proof. See Appendix. ■

This theorem makes precise the idea that the cumulative and buy-and-hold abnormal returns have negative expectations even under the null hypotheses that returns are independent or uncorrelated over time. As discussed in the introduction, the intuition involves differential weighting of paths with high- and low-return sequences. The fact that returns predict the number of events leads to the conclusion that, even under the null hypothesis of market efficiency, the expectation of event-abnormal returns is negative. This makes transparent the intuition for the negative long-run expected abnormal returns when returns predict events.

We next explore the effect of different holding periods on the expected cumulative and buy-and-hold abnormal returns. If these expectations become more negative with the length of holding periods, this makes long-run event studies more susceptible to the issue of negative bias. We show, for longer holding periods, that both the expected cumulative and buy-and-hold abnormal returns are more negative.⁵

Theorem 3.

- (1) $E[\overline{CAR}_T(s+1)] \leq E[\overline{CAR}_T(s)], \forall s \geq 1;$
- (2) $E[\overline{BHAR}_{T+1}(s+1)] \leq (1 + E[r_{IPO}])E[\overline{BHAR}_T(s)], \forall s \geq 1.$

Proof. See Appendix. ■

The intuition for this result is as follows. While looking at one-period returns, we have shown that we underweight the high returns and overweight the low returns. With a longer holding period, we are adding more returns to our sequence of returns. From the affiliation assumption, a sequence of high returns is going to lead to even a greater number of events in the future. Thus we will underweight a sequence of high returns even more compared to a sequence of low returns. This leads to the expectation of the event-abnormal return being even more negative as we increase the holding period.

These results show that, if returns predict the future number of events, the cumulative and buy-and-hold average abnormal returns have negative expectations in a fixed sample. Further, these expected returns are more negative the longer the holding period. The next subsection provides the asymptotic theory for post-event-abnormal returns.

1.3 Asymptotic theory

As a first step toward proving the asymptotic theory, we state an intermediate lemma.

Kronecker's Lemma. Let S_t be a sequence converging to infinity (∞). If $\sum_{t=1}^T N_t r_{t+1} / S_t$ converges, then $S_T^{-1} \sum_{t=1}^T N_t r_{t+1}$ converges to zero as T goes to ∞ .

Note that the lemma places no restriction on N_t , which in our framework corresponds to the number of events in each period. In our setup, S_t is the cumulative number of events, i.e., $S_t = \sum_{i=1}^t N_i$; the lemma requires that this cumulative number of events is eventually large and positive. Note that for CAR, $r_{t+s} = \sum_{j=1}^s ((1 + r_{IPO,t+j}) - (1 + E[r_{IPO,t+j}]))$ and for BHAR, $r_{t+s} = (\prod_{j=1}^s (1 + r_{IPO,t+j}) - \prod_{j=1}^s (1 + E[r_{IPO,t+j}]))$. Hence in the future when we refer to returns, we mean excess returns over the appropriate horizon.

⁵ We can prove Theorem 2 and Theorem 3(1) using the weaker notion of association. Some form of association conditional on the event history is needed to prove Theorem 3(2); affiliation is one condition that is sufficient.

Essentially, to prove that long-run event-abnormal returns converge to zero, we use Kronecker's Lemma and note that it suffices that $\sum_{t=1}^T N_t r_{t+1}/S_t$ converges. Note that the sequence $\{N_t r_{t+1}/S_t, G_t = \sigma(N_1, \dots, N_t, S_1, \dots, S_t, r_1, \dots, r_t)\}$ is a martingale difference sequence with respect to the history G_t , i.e., we have a valid dynamic trading strategy. Using standard methods for dealing with martingale difference sequences (see Steele, 2001; Theorem 2.6), the result of the theorem follows.⁶ Based on the intuition given by Kronecker's Lemma, we derive the following theorem:

Theorem 4. *Let $S_T = \sum_{t=1}^T N_t$, if $\sum_{t=1}^\infty E(N_t/S_t)^2 < \infty$, then for any s , $\overline{CAR}_T(s)$ converges to zero almost surely as T goes to infinity.*

Proof. We first provide a proof of the convergence of $\overline{CAR}_T(s)$ for $s = 1$. Note $\overline{CAR}_T = S_T^{-1} \sum_{t=1}^T N_t r_{t+1}$, to show $\overline{CAR}_T \rightarrow 0$, almost surely, by Kronecker's Lemma, we only need to show that $\sum_{t=1}^T N_t r_{t+1}/S_t$ converges, almost surely. Further, note $\{\sum_{t=1}^T N_t r_{t+1}/S_t\}$ is a martingale, by the L^2 -bounded martingale convergence theorem (see Theorem 2.6 in Steele, 2001), it suffices to show: $\exists B < \infty$, such that

$$E \left[\sum_{t=1}^T N_t r_{t+1}/S_t \right]^2 \leq B < \infty \quad \forall t.$$

Since the returns $\{r_t\}_{t=1}^\infty$ are i.i.d., thus $E[\sum_{t=1}^T N_t r_{t+1}/S_t]^2 = \sigma_r^2 \sum_{t=1}^T E(N_t/S_t)^2 < \infty$, which completes the proof of the theorem.

For $s > 1$, $\overline{CAR}_T(s) = \sum_{j=1}^s [\sum_{t=1}^T N_t r_{t+1+j} / \sum_{t=1}^T N_t]$. For each term in the square bracket, it converges to zero almost surely by a similar argument. Since $\overline{CAR}_T(s)$ is the sum of s such terms, it also converges to zero almost surely. ■

The moment condition that is imposed in Theorem 4, $\sum_{t=1}^\infty E(N_t/S_t)^2 < \infty$, is satisfied by most stationary processes that are considered in finance. However, many nonstationary processes will not satisfy this moment condition. We study these issues in greater detail in Section 2.

2. The Lognormal Model

2.1 Asymptotic theory

We now specialize our general model to the following lognormal model:

$$\log N_{t+1} - \mu = \rho(\log N_t - \mu) + \delta r_t + \epsilon_{t+1}, \quad (2)$$

⁶ A related approach is Chow's theorem for martingale difference sequences (see White 2001, p. 60, Theorem 3.76).

where r_t can be considered as some benchmark-adjusted IPO index return, or abnormal return; $\rho > 0$, $\delta > 0$ are assumed to capture the positive effect of previous N_t and r_t . We assume that $\{r_t\}$ and ϵ_{t+1} are i.i.d. white-noise processes with mean zero. We note that in the empirical data, N_t could be discrete. In this section, we are using a continuous specification.⁷

The lognormal model ensures that the number of events is always positive and allows us to consider both stationary and nonstationary models by varying ρ . The specification in Equation (2) allows past returns to persistently impact the number of events.⁸

We show that asymptotically the bias disappears when $\rho < 1$ because the lognormal model satisfies the moment condition imposed in Theorem 4.

Corollary 1. *If $\rho < 1$, the lognormal model for the number of events satisfies the assumptions of Theorem 4 and hence the event-abnormal return converges to zero as the number of observations T goes to infinity.*

Proof. In the Appendix, we show that the lognormal model in Equation (2) satisfies the moment condition in Theorem 4. ■

We confirm the asymptotic theory we have just derived by simulation. Based on the lognormal model in Equation (2), we choose the parameter ρ to be one of 0.2, 0.4, 0.6, 0.8, or 1.0, and the parameter δ as one of 0.2, 0.4, 0.6, or 0.8. Here r_t is assumed to be i.i.d. normally distributed with mean zero and standard deviation of 0.0824. We set the standard deviation to 0.0824 to be consistent with the IPO data, which is described in the next section. Further, we normalize the initial IPO number to one (i.e., $N_0 = 1$) and let r_0 be randomly drawn from its unconditional distribution.

For a given pair, δ and ρ , we run 500 rounds of simulations. At each round, we simulate the data for a period of $T = 100,000$ and save the abnormal return for period of 1000, 2000, ..., 100,000, respectively. Figure 1 presents these results.

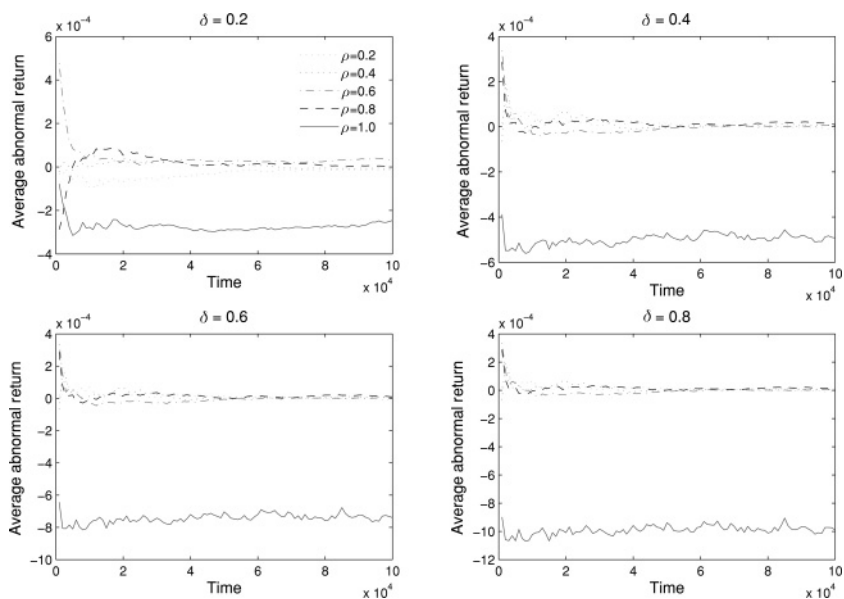
As can be seen from the simulation evidence in Figure 1, for $\rho < 1.0$, the bias goes to zero asymptotically; for $\rho = 1.0$, the negative bias persists asymptotically and gets more negative for bigger δ .

When $\rho = 1$, the moment condition in Theorem 4 is violated. However, we can show that the expected CAR asymptotically converges to a negative constant, as opposed to zero for a stationary event process (see Corollary 1).

Corollary 2. *When $\rho = 1$, $E[\overline{CAR}_T]$ converges to $-\delta\sigma_r^2/2$.*

⁷ Also, N_t could be zero with positive probability in the data. We deal with this when we consider the IPO and SEO data by adding 0.5 to each observation.

⁸ Under this specification, the correlations between $\log(N_t)$ and r_s are positive for any t and s , thus these variables are associated by the theorem in Pitt (1982). Consequently, Theorem 2 and Theorem 3(i) hold for the lognormal model in Equation (2).

**Figure 1****Unit root versus stationary event process**

A graphical illustration of the estimates of the expected *monthly* $\overline{CAR}_T$ for the model without noise: $\log(N_{t+1}) = \rho \log(N_t) + \delta r_t$. r_t is assumed to be i.i.d. normal, with mean zero and standard deviation 0.0824. The standard deviation of r_t is chosen to be 0.0824 to be consistent with our sample. Given a pair of (ρ, δ) , to estimate the expected *monthly* $\overline{CAR}_T$, we run 100 rounds of simulations and use the average of the 100 realizations of $\overline{CAR}_T$ as the estimate of the expectation. For each simulation, we draw 100,000 observations of the IPO return and the number of IPOs, i.e., $T = 100,000$. For each of the four plots, we first fix δ (which is fixed as 0.2, 0.4, 0.6, and 0.8, respectively) and then draw the estimates for $\rho = 0.2, 0.4, 0.6, 0.8$.

Proof. See Appendix. ■

From Corollary 2, the higher δ is, the bias becomes more negative. The intuition is the following: with a higher δ , a shock to the return r_t has a larger impact on the number of events N_t , which makes the covariance between r_t and N_{t-1}/S_T more negative, hence a larger bias. Similarly, with a larger σ_r , the return process is more volatile and extreme values are more likely to occur, which implies a larger impact on N_t and the more negative covariance, as well as the bias.

Overall, the results suggest that, if the process of the number of events is stationary, the negative long-run expected abnormal return is essentially a small-sample problem. Nonstationarity in the process of the number of events is necessary for large long-run negative abnormal returns.⁹

⁹ Our results can be extended to allow for deterministic time trends. Generally, a linear time trend ($N_t = t$) will satisfy the moment condition of Theorem 4 while a geometric time trend ($N_t = e^t$) will not satisfy the moment condition in Theorem 4.

Table 1
Average cumulative abnormal return

	$\delta = 0.5$	$\delta = 1.0$	$\delta = 1.5$	$\delta = 1.75$
Panel A: Average Monthly CAR of holding period $T = 400$				
$\rho = 0.6$	-0.000004	-0.000050	-0.000075	-0.000080
$\rho = 0.8$	-0.000054	-0.000105	-0.000163	-0.000220
$\rho = 0.85$	-0.000094	-0.000191	-0.000264	-0.000294
$\rho = 0.9$	-0.000157	-0.000280	-0.000498	-0.000570
$\rho = 0.95$	-0.000376	-0.000711	-0.001091	-0.001271
$\rho = 1.0$	-0.001522	-0.002971	-0.004493	-0.005246
Panel B: Average 3-year CAR of holding period $T = 400$				
$\rho = 0.6$	-0.000466	-0.001281	-0.001971	-0.002492
$\rho = 0.8$	-0.001233	-0.002933	-0.004205	-0.005113
$\rho = 0.85$	-0.002192	-0.003753	-0.005406	-0.006827
$\rho = 0.9$	-0.003145	-0.005884	-0.009383	-0.011469
$\rho = 0.95$	-0.006708	-0.012655	-0.019601	-0.022620
$\rho = 1.0$	-0.029250	-0.057739	-0.086117	-0.099253

This table reports the estimates of the expected *monthly* $\overline{CAR}_T$ (panel A) and the estimates of the expected *three-year* $\overline{CAR}_T$ (panel B) for $T = 400$ using 500 rounds of simulations based on Theorem 5. All simulations are based on the model used: $\log(N_{t+1}) = \rho \log(N_t) + \delta r_t + \epsilon_{t+1}$, where $\rho = 0.6, 0.8, 0.85, 0.9, 0.95$, or 1.0 and $\delta = 0.5, 1.0, 1.5$, or 1.75 , $r_t \stackrel{iid}{\sim} N(0, 0.0824)$. The standard deviation of r_t is chosen to be 0.0824 to be consistent with our sample.

2.2 Small-sample theory

We provide small-sample theory for the lognormal model. Our approach is to use Stein's Lemma (Stein, 1972) for the lognormal model we introduced in Equation (2). Using Stein's Lemma, we prove the following result that holds for all ρ .¹⁰

Theorem 5. *Under the lognormal specification in Equation (2),*

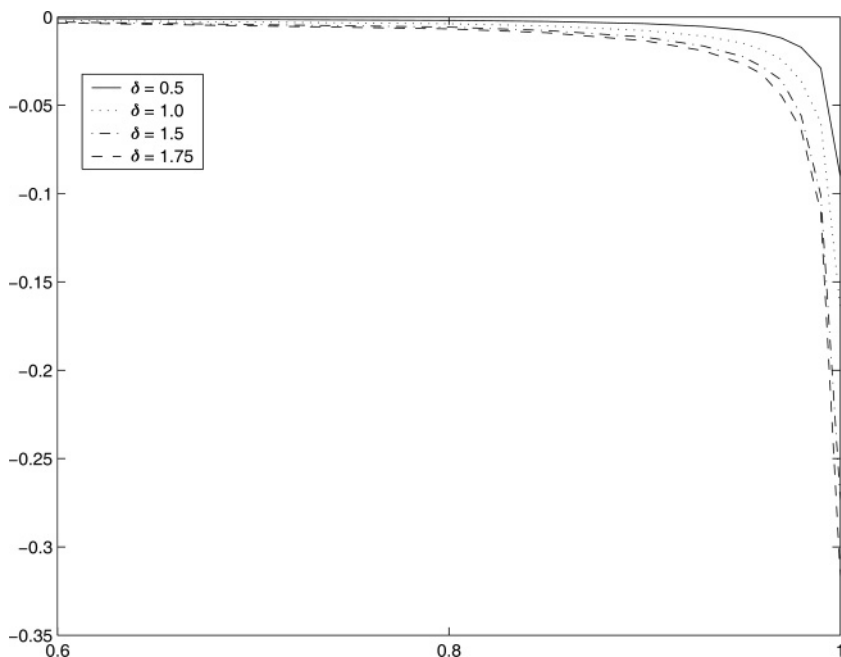
$$E[\overline{CAR}_T] = -\delta \sigma_r^2 \sum_{t=1}^{T-2} \sum_{s=t+2}^T \rho^{s-t-2} E \left[N_t N_s / \left(\sum_{s=1}^T N_s \right)^2 \right] < 0. \quad (3)$$

Proof. See Appendix. ■

We use simulation to evaluate the conditional expectation in Equation (3) using the approach suggested here for a sample of size $T = 400$. In unreported works, we considered $T = 200$ and 600 also. The results are very similar. We consider parameter values $\rho = 0.6, 0.8, 0.85, 0.9, 0.95$, and 1.0 , and $\delta = 0.5, 1.0, 1.5$, and 1.75 (we discuss our choice of parameters in greater detail in the next section). We run 500 rounds of simulations and record the results in Table 1.¹¹

¹⁰ We can only use Stein's Lemma for CARs as they are arithmetic. For BHARs, we cannot obtain such a simple characterization.

¹¹ Using Theorem 5, simulated results converge to the true value very quickly. In fact, the results based on 100 rounds of simulations are very close to the one from 10 rounds of simulations. Thus, Stein's method delivers very accurate estimates of the expected event-abnormal return.

**Figure 2****Small-sample simulation**

This figure plots the estimates of expected $\overline{CAR}_T$ by simulations as ρ goes from 0.6 to 1.0 for a given δ (which is chosen to be one of 0.5, 1.0, 1.5, or 1.75). Based on the model without noise, $\log(N_{t+1}) = \rho \log(N_t) + \delta r_t$, we conduct 500 rounds of simulations assuming r_t is to be i.i.d. normal, with mean zero and standard deviation 0.0824. The standard deviation of r_t is chosen to be 0.0824 to be consistent with our sample of IPOs.

From these small-sample simulations, we can see that the average abnormal returns are negative and tend to get more negative, as ρ increases (the persistence of events is higher), or as δ increases (the relation between returns and subsequent number of events is stronger).¹²

At first cut, our approach does not support as large negative expected abnormal returns as Schultz (2003) finds in his simulations, except in the case where $\rho = 1$ (the unit root case). Schultz finds magnitudes of -0.12 (-12%) in Table VI in his paper for three-year CARs, which is closer to our unit-root magnitude of -0.09 (-9%) that we obtain with 400 observations (see Table 1, panel B). For CARs when $\rho = 0.95$ and $\delta = 1.75$, we obtain a magnitude of -0.015 (-1.5%) for the expected event-abnormal return, which is much smaller.¹³ Figure 2 illustrates these numbers graphically for the model without noise and shows that the exact small-sample bias is very sensitive to the assumption of a unit-root; even a small deviation from the unit-root hypothesis

¹² In earlier drafts, we found that the bias becomes less negative when the number of sample observations T increases except for the case of $\rho = 1.0$. This observation is consistent with our asymptotic theory and our large sample simulation, from which the abnormal returns go to zero as the sample size T goes to infinity for $\rho < 1.0$.

¹³ We obtain higher numbers without the noise ϵ_t in the specification.

Table 2
Average buy-and-hold abnormal return

	Average 3-year BHAR of holding period $T = 400$			
	$\delta = 0.5$	$\delta = 1.0$	$\delta = 1.5$	$\delta = 1.75$
$\rho = 0.6$	-0.000451	-0.001307	-0.002006	-0.002576
$\rho = 0.8$	-0.001246	-0.002947	-0.004242	-0.005171
$\rho = 0.85$	-0.002222	-0.003876	-0.005556	-0.007169
$\rho = 0.9$	-0.003110	-0.006002	-0.009682	-0.011769
$\rho = 0.95$	-0.006828	-0.012753	-0.019636	-0.022888
$\rho = 1.0$	-0.028800	-0.056493	-0.083469	-0.095988

This table reports the estimates of the expected *three-year BHAR* _{T} for $T = 400$ using 500 rounds of simulations based on Theorem 5. All simulations are based on the model used: $\log(N_{t+1}) = \rho \log(N_t) + \delta r_t + \epsilon_{t+1}$, where $\rho = 0.6, 0.8, 0.85, 0.9, 0.95$, or 1.0 and $\delta = 0.5, 1.0, 1.5$ or 1.75 , $r_t \stackrel{iid}{\sim} N(0, 0.0824)$. The standard deviation of r_t is chosen to be 0.0824 to be consistent with our sample.

leads to a dramatic decrease in the bias (making it much less negative). Hence sample size and the stationarity of the process of the number of events play an important role in determining the expected bias.

Next, we provide similar results for the average buy-and-hold abnormal return. Because of the multiplicative nature of buy-and-hold returns, we cannot use the simplification obtained from Stein's formula. Instead, we simulate 300,000 times to find the average buy-and-hold return. The results we obtain are shown in Table 2.

The results are consistent in magnitude with that obtained for CARs. With $T = 400$ observations, $\rho = 0.90$, and $\delta = 1.75$, we obtain magnitudes of -0.011 (-1.1%) versus -0.095 (-9.55%), which is obtained when $\rho = 1.0$ and $\delta = 1.75$ (see Table 2). Again, the magnitudes are much smaller. This suggests that, unless we have a unit root ($\rho = 1$), the expected bias will not be of the same magnitude as that obtained by Schultz (2003).

We also consider what happens when we add more lags to our model. We find that it does not change the bias very much. We confirm this by our results of 300,000 rounds of simulations below (see Table 3). In particular, we study the following two-lag model (with no noise):

$$\log(N_t) = \rho_1 \log(N_{t-1}) + \rho_2 \log(N_{t-2}) + \delta r_{IPO,t-1}. \quad (4)$$

What matters here is the sum of the coefficients ρ_1 and ρ_2 . When the sum is unity, we obtain significant negative returns. Away from unity, the expected returns are negative but the magnitudes are not as large.¹⁴ The small-sample simulations show conclusively that the stationarity of the log number of events process and the sample size play a large role in determining how large the magnitude of the small-sample bias is.

¹⁴ Adding more lags to our model will not change the bias calculations that we undertake. In earlier drafts, we also looked at the simulated three-year cumulative and buy-and-hold abnormal returns. The results are similar to those presented in Table 3 and are not reported here.

Table 3
Average monthly cumulative abnormal return

	$\delta = 2.0$	$\delta = 2.3$	$\delta = 2.5$
Average Monthly CAR of holding period $T = 200$			
$(\rho_1, \rho_2) = (0.6, 0.25)$	-0.000437	-0.000509	-0.000557
$(\rho_1, \rho_2) = (0.6, 0.3)$	-0.000649	-0.000756	-0.000828
$(\rho_1, \rho_2) = (0.7, 0.25)$	-0.001241	-0.001452	-0.001597
$(\rho_1, \rho_2) = (0.7, 0.3)$	-0.005111	-0.005866	-0.006368
Average Monthly CAR of holding period $T = 400$			
$(\rho_1, \rho_2) = (0.6, 0.25)$	-0.000234	-0.000271	-0.000296
$(\rho_1, \rho_2) = (0.6, 0.3)$	-0.000349	-0.000406	-0.000445
$(\rho_1, \rho_2) = (0.7, 0.25)$	-0.000695	-0.000817	-0.000903
$(\rho_1, \rho_2) = (0.7, 0.3)$	-0.005161	-0.005923	-0.006429
Average Monthly CAR of holding period $T = 600$			
$(\rho_1, \rho_2) = (0.6, 0.25)$	-0.000159	-0.000184	-0.000201
$(\rho_1, \rho_2) = (0.6, 0.3)$	-0.000238	-0.000277	-0.000303
$(\rho_1, \rho_2) = (0.7, 0.25)$	-0.000480	-0.000566	-0.000627
$(\rho_1, \rho_2) = (0.7, 0.3)$	-0.005181	-0.005947	-0.006456

This table reports the estimates of monthly $\overline{CAR}_T$, for $T = 200, 400$, and 600 by 300,000 rounds of simulations using $\overline{CAR}_T$'s definition after including one more lag into our model (without noise): $\log(N_t) = \rho_1 \log(N_{t-1}) + \rho_2 \log(N_{t-2}) + \delta r_{IPO,t-1}$, where $\rho_1 = 0.6$ or 0.7 , $\rho_2 = 0.25$ or 0.3 and $\delta = 2.0, 2.3$ or 2.5 , $r_t \stackrel{iid}{\sim} N(0, 0.0824)$. The standard deviation of r_t is chosen to be 0.0824 to be consistent with our sample.

2.3 Empirical tests for unit roots

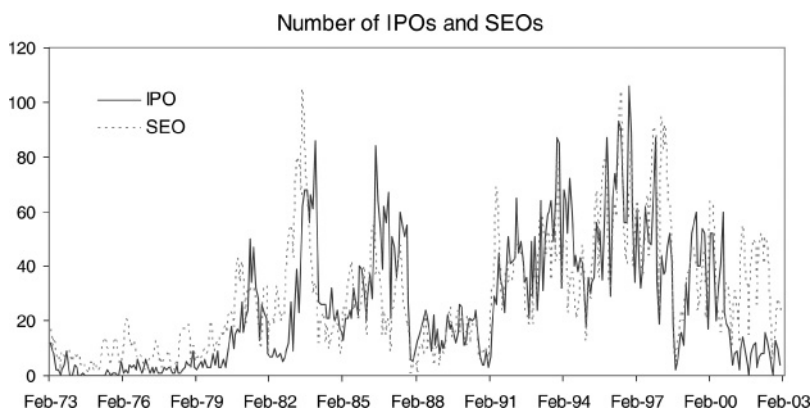
Until now, we have theoretically established the sensitivity of the negative bias to the presence of a unit root in the lognormal autoregressive specification. We now conduct an empirical analysis to see whether the data support the hypothesis of a unit root; this kind of nonstationarity would violate the moment condition required in Theorem 4. Toward that end, we conduct both Augmented Dickey-Fuller and Phillips-Perron tests on the number of IPO and SEO offerings against the following two nulls with or without time trend:

H_0 : constant term, without time trend, unit root

H'_0 : constant term, time trend, unit root.

The sample is comprised of 9190 IPOs ranging from February 1973 to December 2002. The selection criteria are the same as in Ritter (1991): (1) an offer price of \$1.00 per share or more; (2) gross proceeds, measured in terms of 1984 purchasing power, of \$1,000,000 or more; (3) the offering involved common stock only (unit offers are excluded); (4) the company is listed on CRSP daily Amex-NYSE or NASDAQ tapes within six months of the offer date; and (5) an investment banker took the company public.

The numbers of IPOs and SEOs are retrieved from Securities Data Corporation (SDC). Figure 3 depicts the time series of IPO and SEO numbers in the sample period. To be consistent with Schultz (2003), we exclude all offerings by funds, investment companies, and real estate investment trusts (SIC codes 6722, 6726, and 6792), as well as offerings by utilities (SIC codes 4911 through 4941) and banks (6000 through 6081). Table 4 shows the distribution of the number of offerings each month.

**Figure 3****Number of IPOs and SEOs**

The time series of IPO and SEO numbers are from February 1973 to December 2002. The solid (blue) line shows the number of IPOs while the dotted (red) line shows the number of SEOs.

Table 4**The distribution of the number of offerings per month**

	Monthly number of initial public offerings	Monthly number of seasoned equity offerings
Mean	25.60	27.25
Median	20	22
Minimum	0	1
Maximum	106	105
First-order autocorrelation	0.85	0.81

The numbers of IPOs and SEOs are retrieved from Securities Data Corporation (SDC). To be consistent with Schultz (2003), we exclude all offerings by funds, investment companies, and REITs (SIC codes 6722, 6726, and 6792), as well as offerings by utilities (SIC codes 4911 through 4941) and banks (6000 through 6081). The following table shows the distribution of the number of offerings each month.

These tests suggest that with one lag, one can reject the unit-root hypothesis; however, we need to check their robustness to more lags. It is well known in the literature that the power of tests falls with lag length, i.e., we are less likely to reject the null. The Schwert (1989) criterion suggests the maximum lag of 16. Recent work appearing in the unit-root literature suggests that the most powerful test is to use the Elliott, Rotheberg, and Stock (ERS) test (1996) for the unit-root hypothesis with the lag length chosen by the Ng and Perron approach (see Ng and Perron, 2001, for a comprehensive discussion). The Ng-Perron test suggests an optimal lag of 14. We conduct ADF and Elliott-Rothenberg Stock (1996), tests and the results are listed in Tables 5 and 6. In discussing our results, we focus on the ERS test (which is based on the local-root-to-unity approach) as this has the highest power.

From Tables 5 and 6, we can see that the test results for both IPOs and SEOs are very similar, although it is harder to reject the unit-root hypothesis

Table 5
Unit-root testing of H_0 without time trend

Lag	Augmented Dickey-Fuller test					Elliott-Rothenberg-Stock test				
	Test statistics		Critical value			Test statistics		Critical value		
	IPO	SEO	1%	5%	10%	IPO	SEO	1%	5%	10%
16	-2.223	-2.497	-3.451	-2.870	-2.571	-2.152	-2.390	-2.580	-1.951	-1.637
15	-2.322	-2.380	-3.451	-2.870	-2.571	-2.189	-2.245	-2.580	-1.955	-1.641
14	-2.208	-2.415	-3.451	-2.870	-2.571	-2.148	-2.346	-2.580	-1.958	-1.644
13	-2.302	-2.415	-3.451	-2.870	-2.571	-2.398	-2.354	-2.580	-1.962	-1.647
12	-2.430	-2.299	-3.451	-2.870	-2.571	-2.482	-2.246	-2.580	-1.965	-1.650
11	-1.989	-2.062	-3.451	-2.870	-2.571	-2.057	-2.035	-2.580	-1.968	-1.653
10	-2.069	-2.204	-3.451	-2.870	-2.571	-2.222	-2.190	-2.580	-1.972	-1.656
9	-2.172	-2.520	-3.451	-2.870	-2.571	-2.491	-2.567	-2.580	-1.975	-1.659
8	-2.056	-2.855	-3.451	-2.870	-2.571	-2.395	-2.834	-2.580	-1.978	-1.661
7	-2.068	-2.781	-3.451	-2.870	-2.571	-2.335	-2.785	-2.580	-1.981	-1.664
6	-2.147	-3.295	-3.451	-2.870	-2.571	-2.384	-3.300	-2.580	-1.984	-1.667
5	-2.212	-3.373	-3.451	-2.870	-2.571	-2.438	-3.282	-2.580	-1.986	-1.669
4	-2.159	-3.583	-3.451	-2.870	-2.571	-2.532	-3.591	-2.580	-1.989	-1.672
3	-2.081	-3.798	-3.451	-2.870	-2.571	-2.444	-3.816	-2.580	-1.992	-1.674
2	-2.701	-4.106	-3.451	-2.870	-2.571	-2.861	-4.130	-2.580	-1.994	-1.676
1	-3.571	-4.612	-3.451	-2.870	-2.571	-3.483	-4.663	-2.580	-1.997	-1.679

This table reports results of both Augmented Dickey-Fuller and Elliott-Rothenberg-Stock unit-root tests on the number of IPO and SEO offerings up to 16 lags against the null H_0 : unit-root process without time trend. The test statistics for IPOs and SEOs as well as critical values (1%, 5%, 10%) are reported. The left half of the table reports the test results using the Augmented Dickey-Fuller test while the right half is for the Elliott-Rothenberg-Stock test.

Table 6
Unit-root testing of H'_0 with time trend

Lag	Augmented Dickey-Fuller test					Elliott-Rothenberg-Stock test				
	Test statistics		Critical value			Test statistics		Critical value		
	IPO	SEO	1%	5%	10%	IPO	SEO	1%	5%	10%
16	-1.676	-3.121	-3.988	-3.424	-3.135	-2.236	-2.861	-3.48	-2.817	-2.535
15	-1.749	-2.890	-3.988	-3.424	-3.135	-2.276	-2.685	-3.48	-2.823	-2.541
14	-1.671	-3.043	-3.988	-3.424	-3.135	-2.222	-2.800	-3.48	-2.829	-2.546
13	-2.138	-3.025	-3.988	-3.424	-3.135	-2.512	-2.803	-3.48	-2.834	-2.552
12	-2.267	-2.848	-3.988	-3.424	-3.135	-2.606	-2.672	-3.48	-2.840	-2.557
11	-1.330	-2.479	-3.988	-3.424	-3.135	-2.101	-2.423	-3.48	-2.846	-2.562
10	-1.700	-2.759	-3.988	-3.424	-3.135	-2.305	-2.606	-3.48	-2.851	-2.567
9	-2.174	-3.295	-3.988	-3.424	-3.135	-2.615	-3.040	-3.48	-2.856	-2.571
8	-2.057	-3.679	-3.988	-3.424	-3.135	-2.502	-3.345	-3.48	-2.861	-2.576
7	-1.886	-3.569	-3.988	-3.424	-3.135	-2.431	-3.275	-3.48	-2.866	-2.581
6	-1.903	-4.280	-3.988	-3.424	-3.135	-2.486	-3.867	-3.48	-2.871	-2.585
5	-1.896	-4.182	-3.988	-3.424	-3.135	-2.548	-3.825	-3.48	-2.876	-2.589
4	-2.169	-4.658	-3.988	-3.424	-3.135	-2.657	-4.176	-3.48	-2.880	-2.593
3	-2.132	-4.907	-3.988	-3.424	-3.135	-2.547	-4.419	-3.48	-2.884	-2.597
2	-3.211	-5.373	-3.988	-3.424	-3.135	-3.038	-4.759	-3.48	-2.889	-2.601
1	-4.325	-5.912	-3.988	-3.424	-3.135	-3.730	-5.342	-3.48	-2.893	-2.604

This table reports results of both Augmented Dickey-Fuller and Elliott-Rothenberg-Stock tests on the number of IPO and SEO offerings up to 16 lags against the null H'_0 : unit-root process with time trend. The test statistics for IPOs and SEOs, as well as critical values (1%, 5%, 10%) are reported. The left half of the table reports the test results using the Augmented Dickey-Fuller test while the right half is for the Elliott-Rothenberg-Stock test.

for the number of IPOs compared to the number of SEOs. When considering only one lag, we can reject the null hypothesis of unit-root process for $\log(N_t)$; the p -values are much less than 1%. However, with more lag lengths, the test results are mixed—we cannot reject the unit-root hypothesis at the 1% level but reject it (with or without time trend under the ERS test) at the 5% level. Since the null hypothesis is the unit-root and it is well known that the power of these tests becomes lower with more lags, we believe that the unit-root tests at higher lag lengths cannot discriminate between the unit-root hypothesis and its alternative (close to unit root).

2.4 An alternative weighting scheme

An alternative weighting scheme is to scale the event number in each event time by the total number of all events that have happened. For example, in the case of IPOs or SEOs, we can scale the number of IPOs or SEOs in any specific month by the total number of all IPOs or SEOs until that point in time. With a highly persistent process of numbers of events, such scaling should help reduce bias; we analyze this intuition.

The traditional CAR, which weights events equally, is defined as:

$$\overline{CAR}_T = \frac{\sum_{t=1}^T N_t r_{t+1}}{\sum_{t=1}^T N_t}. \quad (5)$$

Being aware of potential concerns about the nonstationarity of IPO volume, some authors have deflated the number of IPOs by the total number of firms in their IPO studies, such as Pastor and Veronesi (2005). Using a similar scaled weighting scheme, we can define CAR in an alternative way:

$$\widehat{CAR}_T = \frac{\sum_{t=1}^T \widehat{N}_t r_{t+1}}{\sum_{t=1}^T \widehat{N}_t}, \quad (6)$$

where $\widehat{N}_t$ is defined as the ratio of the number of IPOs in time t , N_t to the total number of IPOs until that point in time $S_t \equiv \sum_{u=1}^t N_u$. That is, $\widehat{N}_t \equiv N_t/S_t$.

Using Stein's Lemma, we can derive the expression of the expectations of $\widehat{CAR}_T$, which parallels the one in Theorem 5 for the traditional measure $\overline{CAR}_T$. The proof is in the Appendix.

Theorem 6. *Under the assumption that $\log(N_{t+1}) - u = \rho(\log(N_t) - u) + \delta r_t + \epsilon_{t+1}$,*

$$E[\widehat{CAR}_T] = -\delta \sigma_r^2 \sum_{t=1}^{T-2} \sum_{s=t+2}^T \rho^{s-t-2} E \left[f(s) \widehat{N}_t \widehat{N}_s / \left(\sum_{s=1}^T N_s \right)^2 \right], \quad (7)$$

Table 7
Average monthly $\overline{CAR}_T$ under equally weighting scheme

	$\delta = 0.5$	$\delta = 1.0$	$\delta = 1.5$	$\delta = 1.75$
Panel A: Mean of $\overline{CAR}_T$ ($\times 10^{-4}$) with $T = 400$				
$\rho = 0.6$	-0.2001	-0.2984	-0.7568	-0.7673
$\rho = 0.8$	-0.5224	-1.0743	-1.7521	-1.9906
$\rho = 0.85$	-0.5777	-1.8624	-2.6184	-3.3653
$\rho = 0.9$	-1.7879	-3.1278	-5.0328	-5.4113
$\rho = 0.95$	-3.5192	-7.1808	-11.0574	-13.1639
$\rho = 1.0$	-15.0864	-30.2697	-44.5558	-51.8022
Panel B: Std. Dev. of $\overline{CAR}_T$ ($\times 10^{-4}$) with $T = 400$				
$\rho = 0.6$	54.9066	55.1574	55.3252	55.5463
$\rho = 0.8$	66.9126	67.3492	67.9923	68.4603
$\rho = 0.85$	74.8013	75.1170	76.2416	76.9925
$\rho = 0.9$	89.4798	90.2106	91.2480	91.9962
$\rho = 0.95$	117.0015	117.5599	118.7614	118.9192
$\rho = 1.0$	171.1819	172.0112	173.2287	173.7028
Panel C: Mean of $\widehat{CAR}_T$ ($\times 10^{-4}$) with $T = 400$				
$\rho = 0.6$	-0.0025	0.5826	0.2684	-0.2312
$\rho = 0.8$	0.6706	0.5095	0.8015	-0.1072
$\rho = 0.85$	0.4510	-0.0333	-0.0517	0.4218
$\rho = 0.9$	0.4604	0.2896	0.1904	0.6934
$\rho = 0.95$	-0.3192	0.2648	0.0476	-0.1847
$\rho = 1.0$	-2.2153	-5.6278	-8.4805	-9.9124
Panel D: Std. Dev. of $\widehat{CAR}_T$ ($\times 10^{-4}$) with $T = 400$				
$\rho = 0.6$	166.5494	166.7364	166.1989	166.7117
$\rho = 0.8$	165.7027	165.8637	165.9608	166.0384
$\rho = 0.85$	164.9011	164.8834	165.2495	165.0750
$\rho = 0.9$	162.6115	162.4133	162.9946	163.2824
$\rho = 0.95$	157.4607	157.4294	157.2405	157.4461
$\rho = 1.0$	172.1102	172.7707	174.1883	173.4290

This table reports the estimates of the expected *monthly* $\overline{CAR}_T$ and the expected *monthly* $\widehat{CAR}_T$ for $T = 400$ using 200,000 rounds of simulations. $\overline{CAR}_T$ is defined by weighting events equally as in Equation (5). $\widehat{CAR}_T$ is defined under scaled weighting scheme as in Equation (6). All simulations are based on the model $\log(N_{t+1}) = \rho \log(N_t) + \delta r_t + \epsilon_{t+1}$, where $\rho = 0.6, 0.8, 0.85, 0.9, 0.95$ or 1.0 and $\delta = 0.5, 1.0, 1.5$ or 1.75 . The simulation-based estimates of mean and standard deviation of $\overline{CAR}_T$ ($\widehat{CAR}_T$) are reported in panels A and B (C and D), respectively.

where $f(\tau) = 1 - \sum_{u=t+1}^{\tau} \rho^{u-\tau} N_u / S_{\tau}$. In particular, when $\rho = 1$, then $f(\tau) = 1 - \sum_{u=t+1}^{\tau} N_u / S_{\tau} > 0$, and thus:

$$E[\overline{CAR}_T] < 0.$$

Proof. See Appendix. ■

From Theorem 6, we can see that when $\rho < 1$, the sign of $f(s)$ is undetermined, and so is $E[\overline{CAR}_T]$. On the contrary, when $\rho = 1$, $f(s)$ is always positive, which leads to the negative expectation that $E[\widehat{CAR}_T] < 0$. The simulation results reported in Table 7 confirm it. The intuition for these results is the following. Without weighting, when N_t goes up, so does N_{t+1} . With weighting, an increase in N_t results in a smaller increase in $\widehat{N}_{t+1}$ because the denominator of $\widehat{N}_{t+1} = N_{t+1} / S_{t+1}$ also goes up; in fact, for low values of the persistence parameter ρ , $\widehat{N}_{t+1}$ may even decrease. This attenuates the negative bias; how-

ever, the addition of another endogenous parameter S_t generally increases the standard errors.

Further, from Table 7, we can see that under the alternative weighting scheme, $E[\widehat{CAR}_T]$ is much smaller in the absolute term than $E[\overline{CAR}_T]$ for most parameter specifications. However, for a less persistent event process (such as $\rho = 0.6$ or 0.8), its standard deviation becomes two or three times bigger than the latter, which overshadows the benefit from reducing the bias. On the contrary, for a very persistent event process (such as $\rho = 0.95$ or 1.0), its standard deviation increases only a bit because the scaled number of events is more stationary. Therefore, the scaled weighting scheme is better for a very persistent event process, but worse for a less persistent process.¹⁵

3. Relation to Schultz (2003)

Our analytical results can be used to understand better the examples and simulation in Schultz (2003). Schultz presents an example showing that long-run event-abnormal returns are negative and become more negative the longer the horizon chosen. We present an example that is in the same spirit and is consistent with the lognormal model we have presented.

In his example, Schultz (2003) considers a binomial-tree process for the number of events: $N_{t+1} = N_t(1 + \Delta I_t)$. In this setup, with probability 1, the number of events goes to 0. Further, the event process is a martingale that has expectation one, but in the limit with probability close to 1, there are zero events, and with probability very close to 0, there are infinitely many events. This suggests that this example is not an appropriate description of the event process. A heuristic proof of this fact is as follows.¹⁶ After taking logs of the equation determining the number of events, we can rewrite:

$$\begin{aligned}\log N_{t+1} &= \log N_t + \log(1 + \Delta I_t) \\ &= \log N_0 + t \left(\frac{1}{t} \sum_{i=0}^t \log(1 + \Delta I_i) \right).\end{aligned}\quad (8)$$

From Jensen's inequality, we know that $E[\log(1 + \Delta I_t)] = k < \log E[(1 + \Delta I_t)] = 0$, so we know that $E[\log N_{t+1}]$ drifts downward. Further, by the strong law of large numbers, we can prove that $\frac{1}{t} \sum_{i=0}^t \log(1 + \Delta I_i) \xrightarrow{a.s.} k < 0$. Hence it follows that $\log N_{t+1} \xrightarrow{a.s.} -\infty$, from which the result follows that the number of events goes to 0 with probability 1, though the mean number of events is always 1. Thus the event process has fixed mean, but with probability zero it

¹⁵ We also considered a version of the lognormal model where the parameter δ varies across industries, i.e., the relationship between returns and events is industry specific. Now, accounting for this variation across industries and weighting, as we have done in this section, reduces the bias further. These results are available from the authors.

¹⁶ A prior draft provided a more formal proof that is available from the authors upon request.

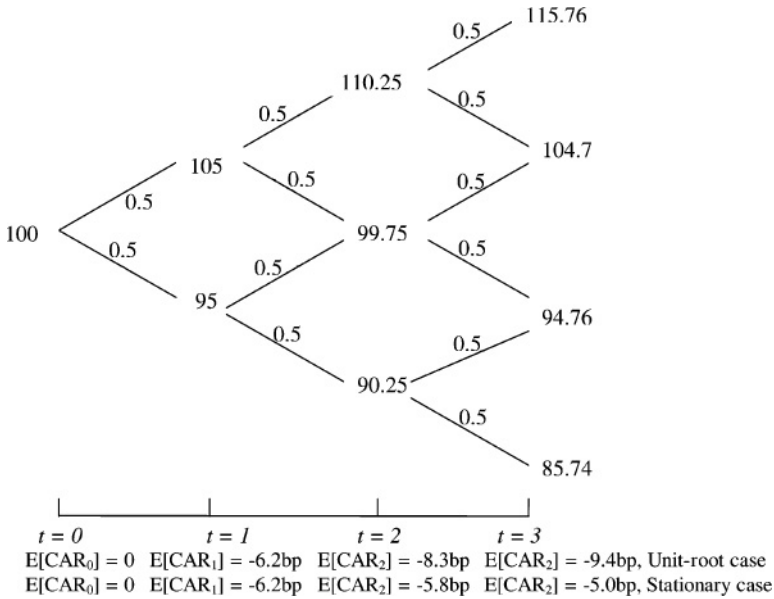


Figure 4
Three-period example

This figure shows a three-period example. In the i_{th} period, the return r_i can be either 5% or -5% with equal probabilities, $i = 1, 2, 3$. At time 0, there is only one new issue with price 100, $N_0 = 1$. The total number of IPOs at time t follows the model: $\log N_t = \rho \log N_{t-1} + r_t$, $t = 1, 2, 3$. We consider both the unit-root case where $\rho = 1.0$ and the stationary case where $\rho = 0.1$. $E[CAR_t]$ is reported for both cases at the bottom of the figure.

takes infinite value and with probability one it takes zero value in the limit. Further, if we start with a positive number of events, the event return is always negative.¹⁷

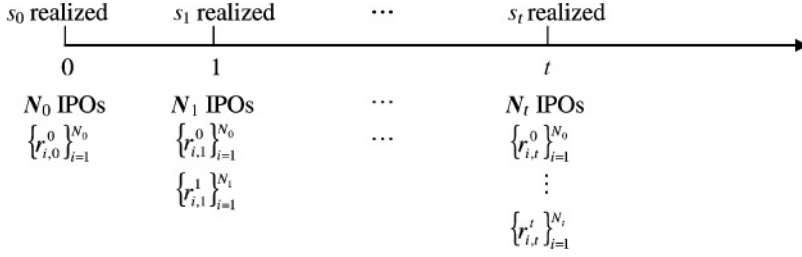
As in Schultz (2003), we consider an example where returns can go up or down each period by 5%. The current stock price P_0 is \$100. Hence at the end of the first period, the stock price can be \$105 or \$95 with equal probability. We allow the number of events to be determined by the process

$$\log(N_{t+1}) - u = \rho(\log(N_t) - u) + \delta r_t. \quad (9)$$

Our example is closely related to Schultz's binomial example. The main difference is that we replace his binomial variable I_t by a normal variable r_t . As a result, ours is a ARMA-type model, which is analytically tractable. Furthermore, by the central limit theorem, our model (when $\rho = 1$) has similar asymptotic behavior.

We consider two examples: one where there is a unit-root ($\rho = 1.0$) and one where there is stationarity ($\rho = 0.1$). Figure 4 shows the example and computes the expected CARs as we go forward in time. For the stationary case, the expected CAR declines over time in absolute magnitude, while for the

¹⁷ Our results here are robust to having asymmetrical shocks in the number of events process.

**Figure 5****Timeline of IPOs and returns**

This figure shows the timeline and explains the notations in the presence of cross-sectional dispersion. Here $r_{i,t}^k$ is the return at time t of the i th firm that went through an IPO at time k and s_t is the common factor at time t , such as the market return. In the t th period, $r_{i,t}^k = \beta s_t + u_{i,t}^k$, $i = 1, 2, \dots, N_k$, $k = 0, 1, \dots, t$, where $s_t \stackrel{iid}{\sim} N(0, \sigma_s^2)$ and $u_{i,j}^k \stackrel{iid}{\sim} N(0, \sigma_u^2)$.

unit-root case, the expected CAR increases over time in absolute magnitude. The examples and simulation of Schultz (2003) are closer to the unit-root case, hence his results.

4. Extension: Cross-Sectional Dispersion

Until this section, we have focused on the bias in mean returns caused by the endogenous nature of events. We now consider the effect of endogenous events on inference, especially on the size of statistical tests. Toward this end, we analyze the effect of event endogeneity on the standard deviation of CARs. To find the asymptotic standard deviation, we present a more general model that allows for cross-sectional dispersion in event returns. Cross-sectional correlation between returns in calendar time implies correlations across periods in event time—for example, the first month event return on this month's events is correlated with the second month event return on yesterday's events. This important aspect has been ignored in the prior literature on long-run event studies such as Barber and Lyon (1997) (Mitchell and Stafford, 2000, do account for this in the context of calendar time regressions). Figure 5 shows the timeline and explains the notation in the presence of cross-sectional dispersion.

In the t th period, event excess returns are given by

$$\begin{aligned} r_{i,t}^0 &= \beta s_t + u_{i,t}^0, i = 1, 2, \dots, N_0 \\ &\vdots \\ r_{i,t}^t &= \beta s_t + u_{i,t}^t, i = 1, 2, \dots, N_t, \end{aligned} \quad (10)$$

where $s_t \stackrel{iid}{\sim} N(0, \sigma_s^2)$ and $u_{i,j}^k \stackrel{iid}{\sim} N(0, \sigma_u^2)$. Here $r_{i,t}^k$ is the excess return at time t of the i th firm that went through an IPO at time k , and s_t is the common factor at time t , like an industry common component or other common factor

that affects all IPO returns in that period.¹⁸ Our specification allows for a single cross-section parameter β ; this can be relaxed. Therefore, $\sigma_r^2 \triangleq \text{Var}(r_{i,j}^k) = \beta^2 \sigma_s^2 + \sigma_u^2$.

We also maintain this assumption: $\log(N_{t+1}) - \mu = \rho(\log(N_t) - \mu) + \delta s_t + \epsilon_{t+1}$. Hence the number of IPOs is determined by the common factor in the prior period, the lagged number of IPOs, and an error term. By definition,

$$\overline{\text{CAR}}_T(S) = \frac{\sum_{t=1}^T \sum_{i=1}^{N_t} \sum_{j=1}^S r_{i,t+j}^t}{\sum_{t=1}^T N_t}. \quad (11)$$

Theorem 7. For $\log(N_{t+1}) - \mu = \rho(\log(N_t) - \mu) + \delta s_t + \epsilon_{t+1}$ and $\rho < 1$, then

(1) for the monthly CAR $\overline{\text{CAR}}_T$ (i.e., $\overline{\text{CAR}}_T(1)$), we can show that

$$\sqrt{T} \overline{\text{CAR}}_T \xrightarrow{L} N\left(0, \frac{\sigma_a^2}{n_a^2}\right), \quad (12)$$

(2) for the S -month CAR $\overline{\text{CAR}}_T(S)$, we can show that

$$\sqrt{T} \overline{\text{CAR}}_T(S) \xrightarrow{L} N\left(0, \frac{\sigma_b^2}{n_a^2}\right), \quad (13)$$

where $\sigma_a^2 \equiv \sigma_r^2 n_a + \beta^2 \sigma_s^2 (n_a^4 - n_a)$, $\sigma_b^2 \equiv S \sigma_a^2 + 2 \sum_{i=1}^{S-1} \sum_{k=1}^{S-i} \exp\left((\delta^2 \sigma_s^2 + \sigma_\epsilon^2) \frac{(1+\rho^k)}{(1-\rho^2)}\right)$, and $n_a \equiv \exp\left(\frac{\delta^2 \sigma_s^2 + \sigma_\epsilon^2}{2(1-\rho^2)}\right)$.

Proof. See Appendix. ■

To evaluate how the cross-sectional dispersion affects the asymptotic variance, we use the following parameter specification based on data: $\delta = 1.75$, $\rho = 0.95$, $\sigma_s = 0.0824$, $\sigma_\epsilon = 0.6117$. To find out σ_r , which is the standard deviation of the historical IPO returns, we use $\sigma_r = 0.2110$ from the squared “average cross-sectional variance” reported in Table II in Ritter (1991). We choose β to be 0, 0.1, 0.2, 0.3, 0.4, 0.5, or 1, and set the sample size to 200. The results are shown in Table 8.

From comparing the first and fourth columns in Table 8, it is clear that with mild correlation (say a correlation of 0.03, which corresponds to a beta of 0.5), the standard deviation increases sharply relative to the case where the number of events is random (from 0.132 to 0.424). The persistence in the number of events interacts with the correlation to increase the standard deviation and

¹⁸ Note that we suppress the IPO subscript and that $r_{IPO,t} = \frac{1}{N_t - 1} \sum_{i=1}^{N_t - 1} r_{i,t}^{t-1}$.

Table 8
Asymptotic standard deviation of three-year $\overline{CAR}_T$

β	Asymptotic standard deviation of 3-year $\overline{CAR}_T$ ($T = 200$)			
	$(\delta = 1.75, \rho = 0.95)$	$(\delta = 1.75, \rho = 0)$	$(\delta = 0, \rho = 0.95)$	$(\delta = 0, \rho = 0)$
0.0	0.03251	0.08110	0.03429	0.08152
0.1	0.09062	0.08374	0.08511	0.08415
0.2	0.17227	0.09122	0.15953	0.09158
0.3	0.25584	0.10247	0.23620	0.10279
0.4	0.33991	0.11640	0.31348	0.11667
0.5	0.42419	0.13219	0.39101	0.13240
1.0	0.84651	0.22397	0.77977	0.22402

This table reports the asymptotic standard deviation of three-year $\overline{CAR}_T$ for $T = 200$ based on the asymptotic distribution stated in Theorem 7(2). To derive the asymptotic distribution, it is assumed that $\log(N_{t+1}) - \mu = \rho(\log(N_t) - \mu) + \delta s_t + \epsilon_{t+1}$. To be consistent with our data, we set the variance of the common factor s_t , $\sigma_s^2 = 0.0824$, and the variance of ϵ_{t+1} , $\sigma_\epsilon^2 = 0.6117$. Further, we use $\sigma_r = 0.211$, the cross-sectional standard deviation of the historical IPO returns from Table II Ritter (1991). For each of four pairs of specification of (δ, ρ) , we calculate the asymptotic standard deviation of three-year $\overline{CAR}_T$ for various values of β , which are 0.0, ..., 0.5, and 1.0. By Theorem 7, the asymptotic standard deviation is: $\sigma_b(n_a \sqrt{T})^{-1}$. The explicit expressions of σ_a and σ_b are given in Theorem 7. Note that the cross-correlation of IPO returns equals $\beta^2 \sigma_s^2 / \sigma_r^2$.

hence lowers the t -statistic.¹⁹ While Mitchell and Stafford (2000) have pointed out that the correlation between event-abnormal returns increases the standard deviation and reduces t -statistics in calendar time, we present results with event-abnormal returns and show that this effect is compounded by the presence of a persistent event process. This suggests that the size of tests in event studies is incorrect and that correct inferences based on event studies do not suggest significant long-run event-abnormal returns.

5. Related Work

Our paper is also related to recent work by Butler, Grullon, and Weston (2005); Baker, Taliaferro, and Wurgler (2006); and Dejong and Dahlquist (2007). The paper by Dejong and Dahlquist (2007) also studies the bias in long-run event-abnormal returns. Dejong and Dahlquist study the bias using a different model of events; they also show that the bias disappears asymptotically with a stationary event process. Our paper differs from theirs along several important dimensions: first, for a very general class of models, we give a formal proof of the existence of negative bias in the small-sample case, and also derive explicit expressions of the bias and asymptotic variance in a more specific model. We also provide these calculations for alternative approaches to weighting the number of events. Second, and most importantly, we point out that the extent of the bias differs considerably for stationary and nonstationary event process. Especially, we show that the negative bias persists when the event process is

¹⁹ The fourth column of Table 8 shows that correlation between event-abnormal returns increases the standard deviation from 0.081 to 0.132; persistence in the number of events further increases this to 0.424. With the usual mean abnormal return of -19% or 0.19, the t -statistics are insignificant.

nonstationary and thus we highlight the important role of the nature of the event process in event studies when events are endogenous.

Also related are papers on market timing of aggregate variables, for example, Butler, Grullon, and Weston (2005) and Baker, Taliaferro, and Wurgler (2006). These studies ask whether market timing of aggregate variables by managers can explain why some aggregate managerial decision variables (such as the equity share in new equity and debt issues, aggregate insider trading, corporate investment plans, etc.) predict stock returns in sample. The approach in this literature is to run a time-series regression of returns on lagged aggregate managerial decision variables (that are often highly persistent). If innovations in returns predict innovations in managerial decisions and current managerial decisions are highly correlated to prior managerial decisions, a small-sample bias arises even though there is no relationship between lagged managerial decisions and returns. While the intuition for this bias has similarities to our paper (both use the relationship between returns and managerial decisions), the time-series regressions used in these papers differ considerably from event studies and the exact intuition is very different.

6. Conclusion

Schultz (2003) has recently argued via simulation that, when returns predict events, long-run event returns are downward biased. We provide the fixed sample and asymptotic theory for long-run event studies when returns predict events. In fixed samples, we prove that expected abnormal returns are negative and become more negative the longer the holding period. This implies that there is a small-sample bias in the use of long-run event returns. Asymptotically, we show that the bias disappears because long-run event-abnormal returns converge in probability to zero when the process of the number of events is stationary. Thus the stationarity assumption on the process of the number of events is sufficient to generate consistency of event-abnormal returns in large samples.

We consider a model where the log number of events follows an autoregressive specification. In the stationary case, we show that convergence occurs while in the nonstationary case, convergence does not occur. To analyze the small-sample bias further, we use Stein's Lemma and compute the small-sample expected bias in the lognormal model. We show that the sample size and the degree of persistence in the event process determines whether the expected small-sample bias is large or not—a small deviation from the unit-root hypothesis reduces the bias a lot. We consider an alternative weighting scheme for the number of events that makes the modified number of events process more stationary and show that this leads to lower bias but higher standard errors. We then prove that the motivating example in Schultz (2003) does have negative expected returns but the number of events converges to zero. We show that the example does not satisfy our convergence theorem since it is not stationary.

Our analysis of IPO and SEO data shows that unit-root tests cannot discriminate between unit root and the near-unit-root alternative. We also derive the asymptotic distribution of long-run event-abnormal returns and show that small correlations between event-abnormal returns interact with persistence in the process of the number of events to increase the standard deviation. Our analysis suggests that with event endogeneity, inference in long-run event studies is more complicated than generally believed.

Appendix

We state Stein's Lemma, which is to be used in some proofs later on. Please refer to Stein (1972) or Liu (1994).

Stein's Lemma

Let $X = (X_1, \dots, X_n)$ be multivariate normally distributed with arbitrary mean vector μ and covariance matrix Σ . For any function $h(x_1, \dots, x_n)$ such that $\partial h / \partial x_i$ exists almost everywhere and $E|\frac{\partial}{\partial x_i} h(X)| < \infty, i = 1, \dots, n$, we write $\nabla h(X) = (\frac{\partial}{\partial x_1} h(X), \dots, \frac{\partial}{\partial x_n} h(X))^T$. Then the following identity is true:

$$\text{cov}[X, h(X)] = \Sigma E[\nabla h(X)].$$

Specifically,

$$\text{cov}[X_1, h(X_1, \dots, X_n)] = \sum_{i=1}^n \text{cov}(X_1, X_i) E \left[\frac{\partial}{\partial x_i} h(X_1, \dots, X_n) \right].$$

We discuss the definition of affiliation when some random variables are discrete.

Affiliation

While affiliation, as defined in Milgrom and Weber (1982), applies to both discrete and continuous random variables, the usual definition of affiliation uses the existence of a probability density function and is

$$f(z' \vee z) f(z' \wedge z) \geq f(z') f(z), \quad (\text{A1})$$

where z is a vector of random variables. Here $z' \vee z$ is the componentwise maximum of the two random variables and $z' \wedge z$ is the componentwise minimum. We can extend this definition to the case in which some of the variables in the vector z are discrete (the vector y) and the rest are continuous (the vector x). Then the appropriate definition is

$$\begin{aligned} p(y' \vee y | x' \vee x) f(x' \vee x) p(y' \wedge y | x' \wedge x) f(x' \wedge x) \\ \geq p(y' | x') f(x') p(y | x) f(x); \end{aligned} \quad (\text{A2})$$

here $p(y|x)$ is the probability of the discrete event y given the continuous variable x and $f(x)$ being the probability density of x .

Clearly, Equation (A2) reduces to Equation (A1) when a joint probability density exists. With this, all the relevant theorems on affiliation go through.

Proof of Theorem 1. First, when $s = 0$, we already know that $N_t, 1 + r_{m,t}, 1 + r_{IPO,t}$ are affiliated. And when $s = 1$, $N_{t+1}, N_t, 1 + r_{m,t+1}, 1 + r_{m,t}, 1 + r_{IPO,t+1}, 1 + r_{IPO,t}$ are affiliated, because

$$\begin{aligned} & f(N_{t+1}, N_t, 1 + r_{m,t+1}, 1 + r_{m,t}, 1 + r_{IPO,t+1}, 1 + r_{IPO,t}) \\ &= f(1 + r_{m,t+1}, 1 + r_{IPO,t+1}) f(N_{t+1}, N_t, 1 + r_{m,t}, 1 + r_{IPO,t}) \\ & \quad \text{(by assumptions 1 and 2)} \\ &= f(1 + r_{m,t+1}, 1 + r_{IPO,t+1}) f(N_{t+1}|N_t, 1 + r_{m,t}, 1 + r_{IPO,t}) \\ & \quad \times f(N_t, 1 + r_{m,t}, 1 + r_{IPO,t}). \end{aligned}$$

Suppose that $(N_{t+s}, \dots, N_t, 1 + r_{m,t+s}, \dots, 1 + r_{m,t}, 1 + r_{IPO,t+s}, \dots, 1 + r_{IPO,t})$ are affiliated. For $(N_{t+s+1}, \dots, N_t, 1 + r_{m,t+s+1}, \dots, 1 + r_{m,t}, 1 + r_{IPO,t+s+1}, \dots, 1 + r_{IPO,t})$, we have

$$\begin{aligned} & f(N_{t+s+1}, \dots, N_t, 1 + r_{m,t+s+1}, \dots, 1 + r_{m,t}, 1 + r_{IPO,t+s+1}, \dots, 1 + r_{IPO,t}) \\ &= f(N_{t+s+1}, \dots, N_t, 1 + r_{m,t+s}, \dots, 1 + r_{m,t}, 1 + r_{IPO,t+s}, \dots, 1 + r_{IPO,t}) \\ & f(1 + r_{m,t+s+1}, 1 + r_{IPO,t+s+1}) \\ &= f(1 + r_{m,t+s+1}, 1 + r_{IPO,t+s+1}) \\ & f(N_{t+s+1}|N_{t+s}, \dots, N_t, 1 + r_{m,t+s}, \dots, 1 + r_{m,t}, 1 + r_{IPO,t+s}, \dots, 1 + r_{IPO,t}) \\ & f(N_{t+s}, \dots, N_t, 1 + r_{m,t+s}, \dots, 1 + r_{m,t}, 1 + r_{IPO,t+s}, \dots, 1 + r_{IPO,t}). \end{aligned}$$

Hence, by induction, $(N_{t+s+1}, \dots, N_t, 1 + r_{m,t+s+1}, \dots, 1 + r_{m,t}, 1 + r_{IPO,t+s+1}, \dots, 1 + r_{IPO,t})$ are also affiliated. ■

Proof of Theorem 2. Throughout the proof below, let $E_t(\cdot)$ be the expectation conditional on information set I_t , $E(\cdot | I_t)$. For the CAR $\overline{CAR}_T(s)$, we have:

$$\begin{aligned} E[\overline{CAR}_T(s)] &= \sum_{t=0}^{T-1} \sum_{j=1}^s E \left[N_t r_{IPO,t+j} / \sum_{t=0}^{T-1} N_t \right] \\ & \quad - E \left[\sum_{t=0}^{T-1} \sum_{j=1}^s N_t E[r_{IPO,t+j}] / \sum_{t=0}^{T-1} N_t \right] \\ &= \sum_{t=0}^{T-1} \sum_{j=1}^s E \left[N_t E_t \left[r_{IPO,t+j} / \sum_{t=0}^{T-1} N_t \right] \right] - \sum_{j=1}^s E[r_{IPO,t+j}]. \end{aligned}$$

Since $(N_1, \dots, N_T, 1 + r_{m,1}, \dots, 1 + r_{m,T}, 1 + r_{IPO,1}, \dots, 1 + r_{IPO,T})$ are affiliated, we use the key implication of affiliation that monotone increasing functions of affiliated variables have positive covariance conditional on any history (see the theorems in Milgrom and Weber, 1982) and obtain that

$$\begin{aligned} E_t \left[r_{IPO,t+j} / \sum_{t=0}^{T-1} N_t \right] &\leq E_t[r_{IPO,t+j}] E_t \left[1 / \sum_{t=0}^{T-1} N_t \right] \\ &= E[r_{IPO,t+j}] E_t \left[1 / \sum_{t=0}^{T-1} N_t \right]. \end{aligned}$$

Therefore,

$$\begin{aligned}
E[\overline{CAR}_T(s)] &\leq \sum_{t=0}^{T-1} \sum_{j=1}^s E \left[N_t E(r_{IPO,t+j}) E_t \left(1 / \sum_{t=0}^{T-1} N_t \right) \right] - \sum_{j=1}^s E[r_{IPO,t+j}] \\
&= \sum_{t=0}^{T-1} \sum_{j=1}^s E(r_{IPO,t+j}) E \left[E_t \left(N_t / \sum_{t=0}^{T-1} N_t \right) \right] - \sum_{j=1}^s E[r_{IPO,t+j}] \\
&= \sum_{j=1}^s E(r_{IPO,t+j}) \sum_{t=0}^{T-1} E \left[N_t / \sum_{t=0}^{T-1} N_t \right] - \sum_{j=1}^s E[r_{IPO,t+j}] = 0.
\end{aligned}$$

Similarly, we can show that $E[\overline{BHAR}_T(s)] \leq 0$. ■

Proof of Theorem 3. Similarly, using the properties of affiliated random variables, we have

$$\begin{aligned}
E[\overline{CAR}_T(s+1)] - E[\overline{CAR}_T(s)] \\
= \sum_{t=0}^{T-1} E \left[N_t (r_{IPO,t+s+1} - E[r_{IPO,t+s+1}]) / \sum_{t=0}^{T-1} N_t \right] \leq 0.
\end{aligned}$$

To prove the theorem, note that

$$\begin{aligned}
E[BHAR_{T+1}(s+1)] &= \sum_{t=0}^{T-1} E \left[N_t \prod_{j=1}^{s+1} (1 + r_{IPO,t+j}) / \sum_{t=0}^{T-1} N_t \right] \\
&\quad - \sum_{t=0}^{T-1} E \left[N_t \prod_{j=1}^{s+1} E[1 + r_{IPO,t+j}] / \sum_{t=0}^{T-1} N_t \right] \\
&= \sum_{t=0}^{T-1} E \left[N_t \prod_{j=1}^s (1 + r_{IPO,t+j}) E_{t+s} \left(\frac{1 + r_{IPO,t+s+1}}{\sum_{t=0}^{T-1} N_t} \right) \right] \\
&\quad - \sum_{t=0}^{T-1} E \left[\frac{N_t \prod_{j=1}^{s+1} E(1 + r_{IPO,t+j})}{\sum_{t=0}^{T-1} N_t} \right].
\end{aligned}$$

By affiliation inequality again, we have $E_{t+s}[(1 + r_{IPO,t+s+1})(\sum_{t=0}^{T-1} N_t)^{-1}]$ is less than $E[1 + r_{IPO,t+s+1}]E_{t+s}(\sum_{t=0}^{T-1} N_t)^{-1}$; therefore, from the preceding equations, $E[BHAR_{T+1}(s+1)]$ sat-

ifies the following inequality:

$$\begin{aligned}
 & E[BHAR_{T+1}(s+1)] \\
 & \leq \sum_{t=0}^{T-1} E \left[N_t \prod_{j=1}^s (1+r_{IPO,t+j}) E[1+r_{IPO,t+s+1}] E_{t+s} \left(\sum_{t=0}^{T-1} N_t \right)^{-1} \right] \\
 & \quad - \sum_{t=0}^{T-1} E \left[\frac{N_t}{\sum_{t=0}^{T-1} N_t} \prod_{j=1}^{s+1} E[1+r_{IPO,t+j}] \right] \\
 & = (1 + E[r_{IPO}]) \sum_{t=0}^{T-1} E \\
 & \quad \times \left[\frac{N_t}{\sum_{t=0}^{T-1} N_t} \left(\prod_{j=1}^s (1+r_{IPO,t+j}) - \prod_{j=1}^s E[1+r_{IPO,t+j}] \right) \right] \\
 & = (1 + E[r_{IPO}]) E[BHAR_T(s)]. \quad \blacksquare
 \end{aligned}$$

Proof of Corollary 1. Without loss of generality, we let $\mu = 0$; otherwise, let $N_t^* = N_t \exp(-\mu)$. Note that the scaling constant affects neither cumulative nor buy-and-hold abnormal returns. We make the same assumption from now on in the proofs. Since $N_t = \exp(\sum_{i=0}^{t-2} \rho^{t-i-2} (\delta r_i + \epsilon_{i+1}))$, then:

$$\begin{aligned}
 S_t & \equiv \sum_{i=1}^t N_i \geq t \left(\prod_{i=1}^t N_i \right)^{1/t} \\
 & = t \left[\exp \left(\sum_{i=2}^t \sum_{j=0}^{i-2} \rho^{i-j-2} (\delta r_j + \epsilon_{j+1}) \right) \right]^{1/t} \\
 & = \begin{cases} t \exp \left(1/t \sum_{j=0}^{t-2} \frac{1-\rho^{t-j-1}}{1-\rho} (\delta r_j + \epsilon_{j+1}) \right) & \text{if } \rho < 1 \\ t \exp \left(1/t \sum_{j=0}^{t-2} (t-j-1) (\delta r_j + \epsilon_{j+1}) \right) & \text{if } \rho = 1. \end{cases}
 \end{aligned}$$

Therefore, for $\rho < 1$,

$$\begin{aligned}
 E[N_t/S_t]^2 & \leq \frac{1}{t^2} E \left[\exp \left(2 \sum_{j=0}^{t-2} \left(\rho^{t-j-2} - \frac{1}{t} \frac{1-\rho^{t-j-1}}{1-\rho} \right) (\delta r_j + \epsilon_{j+1}) \right) \right] \\
 & = \frac{1}{t^2} \exp \left(4(\delta^2 \sigma_r^2 + \sigma_\epsilon^2) \right. \\
 & \quad \times \left(\frac{1-\rho^{2(t-1)}}{1-\rho} - \frac{2}{t(1-\rho)} \left(\frac{1-\rho^{t-1}}{1-\rho} - \frac{\rho(1-\rho^{2t-2})}{1-\rho} \right) \right) \\
 & \quad \times \exp \left(\frac{4(\delta^2 \sigma_r^2 + \sigma_\epsilon^2)}{t^2(1-\rho)^2} \left(t-1 - \frac{2\rho(1-\rho^{t-1})}{1-\rho} + \frac{\rho^2(1-\rho^{2t-2})}{1-\rho^2} \right) \right) \\
 & = O(t^{-2}).
 \end{aligned}$$

Hence, $\sum_{t=1}^{\infty} E(N_t/S_t)^2 < \infty$, which satisfies the condition in the preceding theorem. Therefore, the asymptotic bias is zero, for $\rho < 1$.

Proof of Corollary 2. Let $\lambda_t = N_t/(\sum_{s=1}^T N_s)$

$$\begin{aligned} E[\overline{CAR}_T] &= -\delta\sigma_r^2 \sum_{t=1}^{T-2} \sum_{s=t+2}^T E[\lambda_t \lambda_s] \\ &= -\delta\sigma_r^2/2 + \delta\sigma_r^2/2 E\left[\sum_{t=1}^T \lambda_t^2\right] + \delta\sigma_r^2 E\left[\sum_{t=1}^{T-1} \lambda_t \lambda_{t+1}\right] \\ &\leq -\delta\sigma_r^2/2 + 3\delta\sigma_r^2/2 E\left[\sum_{t=1}^T \lambda_t^2\right], \text{ because } 2\lambda_t \lambda_{t+1} \leq \lambda_t^2 + \lambda_{t+1}^2. \end{aligned}$$

We only need to prove that $E[\sum_{t=1}^T \lambda_t^2]$ converges to zero. In fact, first notice that given $N_0 = 0$ and $N_t = \exp(\sum_{i=0}^{t-1} (\delta r_i + \epsilon_{i+1}))$, we have

$$\begin{aligned} E[\lambda_t^2] &= E\left[\left(\sum_{s=1}^{t-1} \exp\left(-\sum_{u=t-s}^{t-1} (\delta r_u + \epsilon_{u+1})\right) + 1\right.\right. \\ &\quad \left.\left. + \sum_{s=t}^T \exp\left(-\sum_{u=t}^s (\delta r_u + \epsilon_{u+1})\right)\right)^{-2}\right] \\ &\leq E\left[\left(1 + \sum_{s=1}^{[T/2]} \exp\left(-\sum_{u=1}^s \bar{r}_u\right)\right)^{-2}\right], \text{ where } \bar{r}_t \stackrel{iid}{\sim} N(0, (\delta^2\sigma_r^2 + \sigma_\epsilon^2)). \end{aligned}$$

Therefore,

$$\begin{aligned} E\left[\sum_{t=1}^T \lambda_t^2\right] &\leq T E\left[\left(1 + \sum_{s=1}^{[T/2]} \exp\left(-\sum_{u=1}^s \bar{r}_u\right)\right)^{-2}\right] \\ &\leq T E\left[\left(1 + \sum_{s=1}^{[T/2]} \exp\left(-\sum_{u=1}^s \bar{r}_u\right)\right)^{-1}\right] \text{ because} \\ &\quad \times \left(1 + \sum_{s=1}^{[T/2]} \exp\left(-\sum_{u=1}^s \bar{r}_u\right)\right)^{-1} < 1. \end{aligned}$$

Without loss of generality, we only need to prove that for $\bar{r}_t \stackrel{iid}{\sim} N(0, 1)$, as T goes to infinity, $T E[(1 + \sum_{s=1}^{[T/2]} e^{-\sum_{u=1}^s \bar{r}_u})^{-1}]$ converges to zero. Let $S_T \triangleq E[(1 + \sum_{s=1}^{[T/2]} \exp(-\sum_{u=1}^s \bar{r}_u))^{-1}]$. We can prove that for a large enough T , say, $T \geq T_0$, $\exists \epsilon_0 > 0$, such that $S_T \leq T^{-1-\epsilon_0}$.

Notice that $f(x) = (1 + 1/x)^{-1}$ is a concave function; by Jensen's inequality, we have for any $\delta > 0, x > 0$,

$$\begin{aligned} S_T &\leq E \left[\frac{1}{1 + \frac{\exp(\bar{r}_1)}{S_{T-1}}} \right] \leq E \left[\frac{1}{1 + \exp(\bar{r}_1)(T-1)^{1+\epsilon_0}} \right] \\ &= \sum_{n=-\infty}^{\infty} \int_{\ln(a^n)}^{\ln(a^{n+1})} (1 + \exp(\bar{r}_1)(T-1)^{1+\epsilon_0})^{-1} \\ &\quad \times \Phi(\bar{r}_1) d\bar{r}_1 (a > 1 \text{ is to be determined}) \\ &\leq \sum_{n=-\infty}^{\infty} (1 + \exp(\ln(a^n))(T-1)^{1+\epsilon_0})^{-1} \Pr(\ln(a^n) \leq \bar{r}_1 \leq \ln(a^{n+1})) \\ &\leq \sum_{n=0}^{\infty} (T-1)^{-1-\epsilon_0} a^{-n} \Pr(\ln(a^n) \leq \bar{r}_1 \leq \ln(a^{n+1})) \\ &\quad + \sum_{n=1}^{\infty} (T-1)^{-1-\epsilon_0} a^n \Pr(\ln(a^{-n}) \leq \bar{r}_1 \leq \ln(a^{-n+1})). \end{aligned}$$

Furthermore, by the mean value theorem, we have $\exists \ln(a^n) \leq x \leq \ln(a^{n+1})$, such that

$$\begin{aligned} \Pr(\ln(a^n) \leq \bar{r}_1 \leq \ln(a^{n+1})) &= (2\pi)^{-1/2} \exp(-x^2/2) (\ln(a^{n+1}) - \ln(a^n)) \\ &\leq (2\pi)^{-1/2} \ln a \exp(-(\ln(a^n))^2/2). \end{aligned}$$

Hence,

$$\begin{aligned} S_T &\leq \sum_{n=0}^{\infty} (T-1)^{-1-\epsilon_0} a^{-n} \Pr(\ln(a^n) \leq r_1 \leq \ln(a^{n+1})) \\ &\quad + \sum_{n=1}^{\infty} (T-1)^{-1-\epsilon_0} a^n \Pr(\ln(a^{-n}) \leq r_1 \leq \ln(a^{-n+1})) \\ &\leq (2\pi)^{-1/2} (T-1)^{-1-\epsilon_0} \ln a \\ &\quad \times \left[\sum_{n=0}^{\infty} a^{-n} \exp(-(\ln(a^n))^2/2) + \sum_{n=1}^{\infty} a^n \exp(-(\ln(a^{-n}))^2/2) \right]. \end{aligned}$$

The inequality above holds for any $a > 1$. In particular, if we choose $a = 1.1$, then

$$\begin{aligned} &(2\pi)^{-1/2} \ln a \left[\sum_{n=0}^{\infty} a^{-n} \exp(-(\ln(a^n))^2/2) + \sum_{n=1}^{\infty} a^n \exp(-(\ln(a^{-n}))^2/2) \right] \\ &\approx 0.52. \end{aligned}$$

Note that $0.6(T-1)^{-1.1} \leq T^{-1.1}$ holds for large enough T and small ϵ_0 (for example $\epsilon_0 = 0.1, T \geq 3$). Hence, $S_3 \leq T^{-1.1}$. ■

Proof of Theorem 5. First note that $E[N_{T-1} r_T / \sum_{s=1}^T N_s] = E[N_T r_{T+1} / \sum_{s=1}^T N_s] = 0$. Hence, we only need to consider $0 \leq t \leq T-2$.

Let $X = (r_1, \dots, r_{T-1}, \epsilon_2, \dots, \epsilon_T)$, $n_t = \log(N_t)$, and $h(X) = N_t / \sum_{s=1}^T N_s$, $\forall t$. By Stein's formula, we have

$$\begin{aligned}
 E \left[N_t r_{t+1} / \sum_{s=1}^T N_s \right] &= \text{cov} \left[e^{n_t} / \sum_{s=1}^T e^{n_s}, r_{t+1} \right] \\
 &= \sum_{s=1}^{T-1} \text{cov}(r_{t+1}, r_s) E \left[\frac{\partial}{\partial r_s} h(X) \right] \\
 &= -\text{cov}(r_{t+1}, r_{t+1}) E \left[e^{n_t} \sum_{s=t+2}^T \rho^{s-t-2} \delta e^{n_s} \left(\sum_{s=1}^T e^{n_s} \right)^{-2} \right] \\
 &\quad \times (\text{Since } \text{cov}(r_t, r_s) = 0, \text{ if } t \neq s.) \\
 &= - \sum_{s=t+2}^T \rho^{s-t-2} \delta \sigma_r^2 E \left[e^{n_t} e^{n_s} \left(\sum_{s=1}^T e^{n_s} \right)^{-2} \right] < 0 \\
 &\quad \times \left(\text{Since } E \left[e^{n_t} e^{n_s} \left(\sum_{s=1}^T e^{n_s} \right)^{-2} \right] > 0 \right).
 \end{aligned}$$

Therefore,

$$E[\overline{CAR}_T] = -\delta \sigma_r^2 \sum_{t=1}^{T-2} \sum_{s=t+2}^T \rho^{s-t-2} E \left[e^{n_t} e^{n_s} \left(\sum_{s=1}^T N_s \right)^{-2} \right] < 0. \quad \blacksquare$$

Proof of Theorem 6. First note that $E[\widehat{N}_{T-1} r_T / \sum_{s=1}^T \widehat{N}_s] = E[\widehat{N}_T r_{T+1} / \sum_{s=1}^T \widehat{N}_s] = 0$. Hence, we only need to consider $0 \leq t \leq T-2$. Let $X = (r_1, \dots, r_{T-1}, \epsilon_1, \dots, \epsilon_{T-1})$, and $h(X) = \widehat{N}_t / \sum_{s=1}^T \widehat{N}_s$, $\forall t$. By Stein's formula, we have

$$\begin{aligned}
 E \left[\widehat{N}_t r_{t+1} / \sum_{u=1}^T \widehat{N}_u \right] &= \sum_{s=1}^{T-1} \text{cov}(r_{t+1}, r_s) E \left[\frac{\partial}{\partial r_s} h(X) \right] \\
 &= -\text{cov}(r_{t+1}, r_{t+1}) E \left[-\widehat{N}_t / S_T^2 \sum_{s=t+2}^T \frac{\partial \widehat{N}_s}{\partial r_{t+1}} \right] \\
 &\quad \times (\text{Since } \text{cov}(r_t, r_s) = 0, \text{ if } t \neq s.) \\
 &= - \sum_{s=t+2}^T \rho^{s-t-2} \delta \sigma_r^2 E \left[-f(s) \widehat{N}_t \widehat{N}_s / S_T^2 \right].
 \end{aligned}$$

When $\rho = 1$, then $f(\tau) = 1 - \sum_{u=t+1}^T N_u / S_t > 0$, and thus $E[\overline{CAR}_T] < 0$. \blacksquare

To prove Theorem 7, we are going to use the two lemmas below. The proofs are omitted here, which can be found in Corollary 5.26 and Theorem 3.57, respectively, in White (2001).

Lemma 1. Let $\{Y_t\}_{t=1}^\infty$ be a scalar martingale difference sequence with $\bar{Y}_T = \frac{1}{T} \sum_{t=1}^T Y_t$. Suppose that (a) $E(Y_t^2) = \sigma_t^2 > 0$ with $\frac{1}{T} \sum_{t=1}^T \sigma_t^2 \rightarrow \sigma^2 > 0$, (b) $E|Y_t|^r < \infty$ for some $r > 2$ and all t , and (c) $\frac{1}{T} \sum_{t=1}^T Y_t^2 \xrightarrow{P} \sigma^2$. Then $\sqrt{T} \bar{Y}_T \xrightarrow{L} N(0, \sigma^2)$.

Lemma 2. Let $\{Z_t\}$ be a scalar sequence with asymptotically uncorrelated elements²⁰ with means $\mu_t \equiv E(Z_t)$ and $\sigma_t^2 \equiv \text{var}(Z_t) < \Delta < \infty$. Then $\bar{Z}_n - \bar{\mu}_n \xrightarrow{a.s.} 0$.

Proof of Theorem 7. It is easy to prove that $\{N_t\}$ has asymptotically uncorrelated elements, and thus by Lemma 2, we have that $\frac{1}{T} \sum_{t=1}^T N_t \xrightarrow{a.s.} n_a$.

(1) For the monthly cumulative abnormal returns $\overline{CAR}_T$,

$$\sum_{t=1}^T \sum_{i=1}^{N_t} r_{i,t+1}^t = \sum_{t=1}^T \sum_{i=1}^{N_t} (\beta s_{t+1} + u_{i,t+1}^t) = \sum_{t=1}^T \left(\beta N_t s_{t+1} + \sum_{i=1}^{N_t} u_{i,t+1}^t \right).$$

Let $e_{t+1} \triangleq (\beta N_t s_{t+1} + \sum_{i=1}^{N_t} u_{i,t+1}^t)$, then $\{e_t; \Omega_t\}_{t=1}^\infty$ is a martingale difference process, $\Omega_t \triangleq \sigma\{e_t, \dots, e_1\}$ for $t \geq 1$. It's easy to show $\frac{1}{T} \sum_{t=1}^T e_{t+1}^2 \xrightarrow{a.s.} \sigma_a^2 \triangleq \text{var}(e_{t+1}) = \sigma_r^2 \exp\left(\frac{\delta^2 \sigma_s^2 + \sigma_u^2}{2(1-\rho^2)}\right) + \beta^2 \sigma_s^2 \left(\exp\left(\frac{2(\delta^2 \sigma_s^2 + \sigma_u^2)}{(1-\rho^2)}\right) - \exp\left(\frac{\delta^2 \sigma_s^2 + \sigma_u^2}{2(1-\rho^2)}\right)\right)$, then $\sqrt{T} \overline{CAR}_T \xrightarrow{L} N(0, \sigma_a^2/n_a^2)$.

(2) For the S -month cumulative abnormal returns $\overline{CAR}_T(S)$,

$$\begin{aligned} \sum_{t=1}^T \sum_{i=1}^{N_t} \sum_{j=1}^S r_{i,t+j}^t &= \sum_{t=1}^T \sum_{i=1}^{N_t} \sum_{j=1}^S (\beta s_{t+j} + u_{i,t+j}^t) \\ &= \sum_{t=2}^{T+S} \left[\left(\sum_{k=\max(t-S, 1)}^{\min(t-1, T)} N_k \right) \beta s_t + \left(\sum_{k=\max(t-S, 1)}^{\min(t-1, T)} \sum_{i=1}^{N_k} u_{i,t}^k \right) \right] \\ &\stackrel{asy.equ.}{\sim} \sum_{t=1+S}^{T+S} \left[\beta s_t \sum_{k=t-S}^{t-1} N_k + \sum_{k=t-S}^{t-1} \sum_{i=1}^{N_k} u_{i,t}^k \right] \\ &\triangleq \sum_{t=1+S}^{T+S} e_t. \end{aligned}$$

It is easy to show $\frac{1}{T} \sum_{t=1+S}^{T+S} e_t^2 \xrightarrow{a.s.} \sigma_b^2 \triangleq \text{var}(e_t)$, then $\sqrt{T} \overline{CAR}_T \xrightarrow{L} N(0, \frac{\sigma_b^2}{n_a^2})$. We derive σ_b^2 as follows:

$$\begin{aligned} \sigma_b^2 &= \text{var}(e_t) \\ &= \beta^2 \sigma_s^2 E \left(\sum_{k=t-S}^{t-1} N_k \right)^2 + \sigma_u^2 \sum_{k=t-S}^{t-1} E(N_k) \\ &= S \sigma_a^2 + 2 \sum_{i=1}^{S-1} \sum_{k=1}^{S-i} \exp \left(\left(\delta^2 \sigma_s^2 + \sigma_u^2 \right) \frac{(1+\rho^k)}{(1-\rho^2)} \right). \end{aligned}$$

References

- Baker, M., R. Taliaferro, and J. Wurgler. 2006. Predicting Returns with Managerial Decision Variables: Is There a Small-Sample Bias? *Journal of Finance* 61:1711–30.
- Barber, B. M., and J. D. Lyon. 1997. Detecting Long-Run Abnormal Stock Returns: The Empirical Power and Specification of Test-Statistics. *Journal of Financial Economics* 43:341–72.

²⁰ See Definition 3.55 in White (2001). The scalar sequence $\{Z_t\}$ has asymptotically uncorrelated elements if there exist constants $\{\rho_\tau, \tau \geq 0\}$ such that $0 \leq \rho_\tau \leq 1$, $\sum_{\tau=0}^\infty \rho_\tau < \infty$ and $\text{cov}(Z_t, Z_{t+\tau}) \leq \rho_\tau (\text{var}(Z_t) \text{var}(Z_{t+\tau}))^{1/2}$ for all $\tau > 0$, where $\text{var}(Z_t) < \infty$ for all t .

- Brav A. 2000. Inference in Long-Horizon Event Studies: A Bayesian Approach with Application to Initial Public Offerings. *Journal of Finance* 55:1979–2016.
- Butler A. W., G. Grullon, and J. P. Weston. 2005. Can Managers Forecast Aggregate Market Returns? *Journal of Finance* 60:963–86.
- Campbell, J. Y., W. Lo, and A. C. MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.
- Dejong, F., and M. Dahlquist. 2007. Pseudo Market Timing: A Reappraisal. *Journal of Financial and Quantitative Analysis* (forthcoming).
- Elliott, G., T. J. Rothenberg, and J. H. Stock. 1996. Efficient Tests for an Autoregressive Unit Root. *Econometrica* 64:813–36.
- Fama, E. F. 1998. Market Efficiency, Long-Term Returns, and Behavioral Finance. *Journal of Financial Economics* 49:283–306.
- Fama, E., L. Fisher, M. Jensen, and R. Roll. 1969. The Adjustment of Stock Prices to New Information. *International Economic Review* 10:1–21.
- Ikenberry, D., J. Lakonishok, and T. Vermaelen. 1995. Market Underreaction to Open Market Share Repurchases. *Journal of Financial Economics* 39:181–208.
- Kothari, S. P., and J. B. Warner. 1997. Measuring Long-Horizon Security Price Performance. *Journal of Financial Economics* 43:301–39.
- Kothari, S. P., and J. B. Warner. 2007. Econometrics of Event Studies, in B. Espen Eckbo (ed.), *Handbook of Corporate Finance Empirical Corporate Finance*, ch 1. Amsterdam, The Netherlands: Elsevier/North-Holland.
- Li, K., and N. R. Prabhala. 2007. Self-Selection Models in Corporate Finance, in B. Espen Eckbo (ed.), *Handbook of Corporate Finance Empirical Corporate Finance*, ch 2. Amsterdam, The Netherlands: Elsevier/North-Holland.
- Liu, J. 1994. Siegel's Formula via Stein's Identities. *Statistics and Probability Letters* 21:247–51.
- Loughran, T., and J. R. Ritter. 1995. The New Issues Puzzle. *Journal of Finance* 50:23–52.
- Loughran, T., and A. M. Vijh. 1997. Do Long-Term Shareholders Benefit From Corporate Acquisitions? *Journal of Finance* 52:1765–90.
- Lucas, D., and R. McDonald. 1990. Equity Issues and Stock Price Dynamics. *Journal of Finance* 45:1019–43.
- Lyon, J., B. Barber, and C.-L. Tsai. 1999. Improved Methods for Tests of Long-Run Abnormal Stock Returns. *Journal of Finance* 54:165–201.
- Michaely, R., R. Thaler, and K. Womack. 1995. Price Reactions to Dividend Initiations and Omissions: Overreaction or Drift? *Journal of Finance* 50:573–608.
- Milgrom, P. R., and R. Weber. 1982. A Theory of Auctions and Competitive Bidding. *Econometrica* 50:1089–122.
- Mitchell, M. L., and E. Stafford. 2000. Managerial Decisions and Long-Term Stock Price Performance. *Journal of Business* 73:287–329.
- Ng, S., and P. Perron. 2001. Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power. *Econometrica* 69:1519–54.
- Park, J. Y., and P. C. B. Phillips. 2001. Nonlinear Regressions with Integrated Time Series. *Econometrica* 69:117–61.
- Pastor, L., and P. Veronesi. 2005. Stock Prices and IPO Waves. *Journal of Finance* 60:1713–57.
- Pitt, L. D. 1982. Positively Correlated Normal Variables Are Associated. *Annals of Probability* 10:496–99.
- Ritter, J. R. 1991. The Long-Run Performance of Initial Public Offerings. *Journal of Finance* 46:3–27.

Rhodes-Kropf, M., and S. Viswanathan. 2004. Market Valuation and Merger Waves. *Journal of Finance* 59:2685–718.

Schultz, P. 2003. Pseudo Market Timing and the Long-Run Underperformance of IPOs. *Journal of Finance* 58:483–517.

Schwert, G. 1989. Tests for Unit Roots: A Monte Carlo Investigation. *Journal of Business and Economic Statistics* 7:147–59.

Steele, J. M. 2001. *Stochastic Calculus and Financial Applications*. New York, NY: Springer.

Stein, C. M. 1972. A Bound for the Error in the Normal Approximation to the Distribution of a Sum of Dependent Random Variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* 2:583–602.

White, H. 2001. *Asymptotic Theory for Econometricians*, rev. ed. San Diego, CA: Academic Press.

Copyright of Review of Financial Studies is the property of Society for Financial Studies and its 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.