Endogenous Events and Long Run Returns

S. Viswanathan
Fuqua School of Business
Duke University
Durham NC 27708
Telephone: 919-660-7784
Email: viswanat@mail.duke.edu

Bin Wei
Fuqua School of Business
Duke University
Durham NC 27708
Telephone: 919-660-7644
Email: bw12@duke.edu

Current Version April 2004
First Version November 2003
ABSTRACT

We analyze event abnormal returns when returns predict events. We show that the expected abnormal return is negative for any fixed sample and this increases with the holding period of returns. However, we prove that under some standard conditions, abnormal returns converge to zero asymptotically. In particular, it suffices that the number of events process is stationary. We prove that the results in Schultz (2003) are due to the assumption of a unit root in the number of events process. We present simulation and small sample evidence for our results to show that sample size and stationarity property of the event process are important in assessing the magnitude of the small sample bias. We also show that the confidence intervals are large when events are endogenous. This dramatically reduces the power of tests based on event abnormal returns.
1 Introduction

There is a long tradition in finance following the work of Fama, Fisher, Jensen and Roll (1969) of using event studies to study announcement effects. Since Ritter’s (1991) work on initial public offerings, variants of this methodology have been used to study the long run performance of events. Such studies include Ritter (1991), Loughran and Ritter (1995), Loughran and Vijh (1997) among others. In particular, this methodology leads to the conclusion that several events that involve managerial actions exhibit long run underperformance. A still unresolved debate has ensued on whether event studies are appropriate methods to study long run performance (see Fama (1998), Loughran and Ritter (1995), Brav (2000), Barber and Lyon (1997) among others).

In this paper we take a deeper look at the the view that events are endogenous. Recently Schultz (2003) argues via an example and simulation that returns predict events and hence event abnormal returns will always be negative (as it is in his example and simulation). Further Schultz (2003) suggests that the magnitude of negative returns he obtains in his simulations explains the long run abnormal performance of IPOs in event studies. We follow Schultz (2003) and assume that returns predict future events. This assumption seems 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. The recent model of Pastor and Veronesi (2003) predicts that managers will time initial public offerings when the stock market is doing well. Similarly Rhodes-Kropf and Viswanathan (2003) show that mergers occur when markets are relatively overvalued (but this is not known to the agents in the model). All these models suggest that we should see corporate events more often when returns are higher. Thus we take as a primitive that returns today are correlated with events in the future. In particular, high stock returns today result in higher number of events in the near and possibly distant future. Using this assumption, we provide a simple proof that the usual buy and hold returns and mean cumulative abnormal returns (hence forth we refer to both as event abnormal returns) computed in event studies have negative expectation. More importantly, we prove that the longer the holding period, the more negative is the expected event abnormal return. Thus we provide a formal proof that when returns predict events, the expected event abnormal return in any fixed sample is negative.

The intuition for the negative event abnormal return is as follows. A priori, we expect that all return histories that are equally likely in the data will be weighted identically in event abnormal returns. However, when 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, we overweight the low returns that happen subsequently. Consequently, the event abnormal return has negative expectation. With a longer holding period, the underweighting of return histories is exacerbated because we now consider 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 implies that the negative expectation of event abnormal returns increases in the holding period.

While this argument proves that the expected event abnormal returns are negative, the asymptotic theory of event abnormal returns is unclear. Schultz (2003) presents an example and suggests via simulation that when market levels predict the number of events, the long run return averaged across all simulations is negative. However, we provide a proof that asymptotically event returns

\footnote{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) find overperformance after open market share repurchases.}
converge to zero under standard assumptions. Thus event returns are consistent estimators of the null hypothesis of market efficiency when there is sufficient data. The assumptions that we require for our argument are quite mild: returns and events are stationary and the cumulative number of events converges to infinity (the assumptions on the number of events can be relaxed even further). The assumptions on returns are standard in the finance and econometrics literature. The assumptions on events seem reasonable since the number of events is weakly positive in each period. Surprisingly, the event process can allow for non-stationarity of certain kinds and can take the value of zero once in a while, i.e., the restriction on the number of events process seems quite minimal. This argument suggests that the negative expected event abnormal return is a small sample problem.

The main intuition for our asymptotic result is as follows. Stationarity in the event process implies that the shock to the event process does not persist. Consequently, the total number of events in a very large sample is not affected by the shock to the number of events today. Hence 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 quite large. The stationarity assumption is important. With a unit root in the number of events process, 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 returns today.

A second (and more technical) intuition for our asymptotic result is as follows. While event abnormal returns are not dynamic trading strategies, there is a dynamic trading strategy (with expected return equal to 0) whose asymptotic behavior determines the asymptotic behavior of event abnormal returns. By applying a version of the martingale difference theorem (see Theorem 3.6 on page 60 in White (2001) for example), we are able to show that this dynamic trading strategy asymptotically converges to zero. Consequently, the asymptotic behavior of event abnormal returns is related to that of dynamic trading strategies, i.e., the abnormal return converges to zero. We believe this result is important for two reasons. First it suggests that the expected negative expectation is a small sample problem. At a deeper level, it suggests that the critique that event returns are not dynamic trading strategies is essentially a small sample critique since asymptotically the behavior of event abnormal returns is similar to that of dynamic trading strategies.

Thus our asymptotic theory differs from Schultz’s simulations. We show that Schultz’s motivating example does have an expected abnormal return that is negative and becomes more negative with holding period. However, we prove that the number of events goes to zero with probability one and thus the example is non-stationary, in fact the log number of events process has a unit root. We do not view this example as realistic because the number of IPOs is unlikely to converge in probability to zero over the long run. Our analysis of this example and our asymptotic theory suggests that non-stationarity in the event process is necessary for the generation of large negative abnormal returns in a large sample. To understand this further we use our asymptotic theory and simulations to study the case where the log of the number of returns is related to the lag of the log of the number of events and the lagged IPO return. In this case we show that when the autoregressive coefficient on the lagged log number of events is less than one 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, the expected value converges to a negative number in the simulation. All this raises the question how close to a unit root the number of events process is and what the small sample properties of event abnormal returns are.

\(^2\)Dejong and Dahlquist (2003) make a similar argument in a contemporaneous paper. They use a different specification for the events process.

\(^3\)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 non-linear function is well known, see Park and Phillips (2001) for more details.
Since the negative expected returns are a small sample problem, we document the importance of this using simulations and an exact small sample calculation. With one lag in the number of events process, we reject the hypothesis that the number of events process has a unit root using both the Augmented Dickey Fuller test and the Elliott Rothenberg Stock (ERS) test. With more lags in the autoregression the evidence is more confusing. Following the extensive work on unit roots (see Ng and Perron (2001) for a synthesis), we use the Ng-Perron optimally chosen lag length and the Elliott Rothenberg Stock test because they have the most power (the power of unit root tests falls with lag length). We find that we do not reject the unit root hypothesis at the 1% level but reject it at the 5% level (here the null is the unit root). Since the null hypothesis is the unit root and the power of unit root tests is low with higher lag lengths, we believe the data cannot discriminate between the unit root hypothesis and the near unit root alternative.

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 an unit root. Even small deviations from the unit root hypothesis lead to sharp drops in the expected bias. With one and three year returns our unit root results are very similar to that obtained in Schultz (2003). With an autoregressive coefficient on 0.98, the magnitude of the bias is around 1/5th of that obtained in Schultz (2003). Thus the bias is very sensitive to the unit root hypothesis.

We also study whether the data has a regime shift in it. Since Ritter (1984), the empirical literature on IPOs has distinguished between hot and cold IPO regimes. Log likelihood tests pick the regime shift model between cold and hot regimes over the autoregressive model. Since regime shift models often cannot be distinguished from unit root models and there is evidence in favor of a regime shift model, we cannot rule out that the data is being generated by a stationary model with regime shifts.

We also consider how confidence intervals are affected by the presence of endogenous events. We compute the exact asymptotic distribution of the event abnormal return under the null (using our VAR estimates for the log of the event process, stock market returns and IPO returns) and show that this distribution has high variance (hence wide confidence intervals) when the persistence in the number of events is high and the relationship between returns and subsequent events is strong. With the fitted parameters from the log of the event process specification, we are unable to reject the hypothesis that event abnormal returns are zero. This suggests to us that event studies have low power when endogenous variation in number of events is correctly accounted for.

Our paper carefully studies the argument that events are endogenous. We prove that in any finite sample the expected event abnormal return is negative and that the longer the holding period the more negative the return. However, we show that asymptotically the event abnormal return converges to zero, i.e. the bias dissapears. While event abnormal returns are not dynamic trading strategies, their asymptotic behavior is determined by a dynamic trading strategy. Hence the negative expectation is a small sample problem. We study via simulation and Stein expansions how important this issue is in small samples and show that the magnitude of the bias depends critically on the sample size and stationarity of the event generating process. Our results also imply that the critique that event studies are not dynamic trading strategies is a small sample critique since the asymptotic behavior of event abnormal returns is determined by a related dynamic trading strategy.

Our paper is organized as follows. Section 2 presents the model, Section 3 proves the results on expected abnormal returns for fixed sample sizes. Section 4 studies the asymptotic theory while Section 5 analyzes Schultz’s example and show why it does not satisfy the assumptions in Section 3. Section 6 studies the convergence of unit root processes while Section 7 studies the small sample properties. Section 8 studies whether the data supports the unit root hypothesis and Section 9 considers confidence intervals and the power of event studies. Finally Section 8 concludes.
2 The Model

Consider the following simple model. Let $r_{m,t}$ be the market return and $r_{IPO,t}$ be the return on an event index (Schultz (2003) defines this by considering events that have occurred in the last three years). While our model holds for a large number of managerial actions like IPOs, SEOs, stock repurchases, mergers and so on, we use IPOs to be concrete. Let $N_t$ be the number of IPOs at time $t$. In most of what we consider the time interval $t$ is a month. Let $I_{t-1}$ be the information of the agent at the end of $t - 1$. We make the following assumptions:

**Assumption 1** $r_{m,t}$, $r_{IPO,t}$ are temporally independent with $E[r_{m,t}] = E[r_{IPO,t}]$; and $N_t$ is independent of $r_{m,t}$, $r_{IPO,t}$.

This is reasonable if $N_t$ is assumed to depend on the market returns until time $t - 1$ and the IPO index returns until time $t - 1$.

**Assumption 2** The event process $\{N_t\}$ is a Markov process. To be more specific,

$$f(N_t|I_{t-1}) = f(N_{t-1}, r_{m,t-1}, r_{IPO,t-1})$$

This is a restrictive assumption that we can relax if needed by including longer histories. That does not change any of the results, it only requires some more state variables and makes the notation more cumbersome.

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

Assumption 3 is the critical assumption that is made in the model. It states that higher values of lagged variables in the information set, market returns, IPO returns and lagged number of events, $(N_{t-1}, 1 + r_{m,t-1}, 1 + r_{IPO,t-1})$ lead to higher number of events in the next period. From Milgrom and Weber (1982) we know that this is the right statistical restriction to embody the idea that events are monotone increasing in lagged IPO returns. Rhodes-Kropf and Viswanathan (2003) use the affiliation assumption in their model of merger waves. In the Appendix we detail standard results on affiliation from Milgrom and Weber (1982). A key advantage of this approach is that we can prove the results on negative expectations of event abnormal returns with great generality.

Distributions that belong to the log-concave class will satisfy the affiliation requirement. Examples of log-concave densities include the multivariate beta, Dirichlet, exponential, gamma, Laplace, normal, uniform, Weibull and Wishart distributions. Thus many standard distributions satisfy the affiliation assumption. The affiliation assumption ensures that higher returns imply higher events and this relationship holds for all histories of returns and events. Since most standard distributions in finance satisfy the assumption, we view Assumption 1 as providing a broad class of distributions for which the results are true.

First, based on the Assumptions 1, 2 and 3 above, it is very easy to prove that $N_t, 1 + r_{m,t}, 1 + r_{IPO,t}$ are affiliated. Furthermore, we have the following theorem that we will use in the next section.

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

**Proof.** See Appendix. □
3 Expected Abnormal Returns for Fixed Sample Sizes

We first define the average cumulative abnormal return and average buy-and-hold abnormal return of $s$ holding periods as

$$\text{CAR}_T(s) = \frac{\sum_{i=0}^{T-1} N_i \left( \sum_{j=1}^{s} ((1 + \text{IPO}_{i+j}) - (1 + E[\text{IPO}_{i+j}])) \right)}{\sum_{i=0}^{T-1} N_i}$$

$$\text{BHAR}_T(s) = \frac{\sum_{i=0}^{T-1} N_i \left( \prod_{j=1}^{s} (1 + \text{IPO}_{i+j}) - \prod_{j=1}^{s} (1 + E[\text{IPO}_{i+j}]) \right)}{\sum_{i=0}^{T-1} N_i}$$

where $E[\text{IPO}_{i,j}]$ is the expected return of IPO index for period $t$ that is subtracted according to some model. These are the most standard definitions used in the literature (see Ritter (1991), Campbell, Lo and MacKinlay (1997), Barber and Lyon (1997), Lyon, Barber, Tsai (1999) and Schultz (2003))\(^4\). Our first theorem shows that the expected cumulative abnormal return is negative.

**Theorem 2** $E[\text{CAR}_T(s)] \leq 0, \forall s$.

**Proof.**

$$E[\text{CAR}_T(s)] = \sum_{i=0}^{T-1} E \left[ N_i \left( \sum_{j=1}^{s} ((1 + \text{IPO}_{i+j}) - (1 + E[\text{IPO}_{i+j}])) \right) \right]$$

$$= \sum_{i=0}^{T-1} s \sum_{j=1}^{s} E \left[ \frac{N_i \text{IPO}_{i,j}}{\sum_{i=0}^{T-1} N_i} - \frac{N_i E[\text{IPO}_{i,j}]}{\sum_{i=0}^{T-1} N_i} \right]$$

$$= \sum_{i=0}^{T-1} s \sum_{j=1}^{s} E \left[ N_i E \left[ \frac{\text{IPO}_{i,j} I_i}{\sum_{i=0}^{T-1} N_i} \right] \right] - \sum_{i=0}^{T-1} s \sum_{j=1}^{s} \frac{N_i E[\text{IPO}_{i,j}]}{\sum_{i=0}^{T-1} N_i}$$

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

$$E \left[ \frac{\text{IPO}_{i,j}}{\sum_{i=0}^{T-1} N_i} | I_i \right] \leq E \left[ \text{IPO}_{i} | I_i \right] E \left[ \frac{1}{\sum_{i=0}^{T-1} N_i} \right]$$

$$= E \left[ r_{IPO,i} | I_i \right] E \left[ \frac{1}{\sum_{i=0}^{T-1} N_i} \right]$$

Therefore,

$$E[\text{CAR}_T(s)] \leq \sum_{i=0}^{T-1} s \sum_{j=1}^{s} E \left[ N_i E \left[ \frac{\text{IPO}_{i,j}}{\sum_{i=0}^{T-1} N_i} | I_i \right] E \left[ \frac{1}{\sum_{i=0}^{T-1} N_i} \right] \right] - \sum_{i=0}^{T-1} s \sum_{j=1}^{s} \frac{N_i E[\text{IPO}_{i,j}]}{\sum_{i=0}^{T-1} N_i}$$

\(^4\)See Kothari and Warner (1997) for a different approach.
Similarly, we can prove that the expected buy-and-hold abnormal return is negative.

**Theorem 3** $E[\overline{BHAR}_T(s)] \leq 0$, $\forall s$.

**Proof.** See Appendix. 

These two theorems make precise the idea that the usual cumulative abnormal returns and buy and hold returns have negative expectations even under the null hypotheses that returns are independent or uncorrelated over time. The intuition is as follows. Since returns predict events and events are persistent, higher returns today yields higher number of events at every period in the future. This induces a negative correlation between the return on IPOs and one divided by the total number of events, conditional on the information at that point in time. This implies that high returns are underweighted compared to low returns while calculating the expectation, i.e., our prior that event abnormal return will weight two equally likely returns equally is not true. Hence, the fact that returns predict the number of events leads to the conclusion that even under the null hypothesis the expectation of event abnormal returns is negative. This makes transparent the intuition for the negative expected abnormal returns when returns predict events.

We next explore the effect of different holding periods on the expected cumulative abnormal return and expected buy and hold return. 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, the expected cumulative abnormal returns is more negative.

**Theorem 4** $E[\overline{CAR}_T(s+1)] \leq E[\overline{CAR}_T(s)]$, $\forall s \geq 1$.

**Proof.**

\[
E[\overline{CAR}_T(s+1)] - E[\overline{CAR}_T(s)] = \sum_{i=0}^{T-1} E\left[ \frac{N_i(r_{IPO,i+s+1} - E[r_{IPO,i+s+1}])}{\sum_{i=0}^{T-1} N_i} \right] \leq 0
\]

For buy-and-hold abnormal returns, we have a similar result as the theorem above:

**Theorem 5** $E[\overline{BHAR}_{T+1}(s+1)] \leq (1 + E[r_{IPO}])E[\overline{BHAR}_T(s)]$, $\forall s \geq 1$.

**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 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 conclusively that if returns predict the future number of events, the usual buy and hold and cumulative average returns have negative expectation. Further these expected returns are more negative the longer the holding period. Thus these results seem to suggest that event abnormal returns have a bias that is not present with dynamic trading strategies that are studied in calendar time returns. To understand how important this bias is, we study both small sample and asymptotic properties of event abnormal returns.
4 Asymptotic Theory for Stationary Events Generating Process

We consider the asymptotic theory of event abnormal returns. As a starting point we remember the definitions for event abnormal returns.

\[
\begin{align*}
\text{CAR}_T(s) &= \frac{\sum_{t=1}^{T-1} N_t \left( \sum_{j=1}^s \left( (1 + r_{IPO,i+j}) - (1 + E[r_{IPO,i+j}]) \right) \right)}{\sum_{t=0}^{T-1} N_t} \\
\text{BHAR}_T(s) &= \frac{\sum_{t=1}^{T-1} N_t \left( \prod_{j=1}^s \left( (1 + r_{IPO,i+j}) \right) - \prod_{j=1}^s \left( 1 + E[r_{IPO,i+j}] \right) \right)}{\sum_{t=0}^{T-1} N_t}
\end{align*}
\]

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

**Kronecker’s Lemma:** Let \( A_t \) be a sequence converging to infinity \((\infty)\). If \( \sum_{t=1}^{T} \frac{a_t}{A_t} X_t \) converges to zero, then \( A_t^{-1} \sum_{t=1}^{T} a_t X_t \) converges to zero.

Note that the lemma places no restriction on \( a_t \) which corresponds to the number of events in each period (hence we can have zero number of events or even negative number of events). However, the cumulative number of events has to go to \( \infty \). Hence \( A_t \) is eventually large and positive.

The key insight from a finance perspective that Kronecker's lemma yields is that the critique of event studies that they are not dynamic trading strategies is a small sample critique. Asymptotically, the behavior of event abnormal returns is completely determined by a related dynamic trading strategy. Hence if this related dynamic trading strategy converges, event abnormal returns converge. Since we expect dynamic trading strategies to be well behaved, asymptotically event abnormal returns are also well behaved even though they are biased in small samples.

Essentially to prove that event abnormal returns converge to zero, we use Kronecker’s Lemma above and note that it suffices that \( \sum_{t=1}^{T} \frac{a_t}{A_t} X_t \) converges to zero. However the sequence \( \{ \frac{a_t}{A_t} X_t, G_t = \sigma(a_1, \cdots, a_{t+1}, A_1, \cdots, A_{t+1}, X_1, \cdots, X_t) \} \) is a martingale difference sequence with respect to the history \( G_t \), i.e., we have a valid dynamic trading strategy. Thus the asymptotic behavior of event abnormal returns is essentially determined by the asymptotic behavior of a dynamic trading strategy.

Since we have standard methods for dealing with martingale difference sequences, the result of the theorem follows. In fact, we could have followed the argument in Chow’s theorem for martingale difference sequences (see White (2001) page 60, Theorem 3.76), but the theorems we provide impose weaker conditions. Based on the intuition given by Kronecker’s lemma, we derive the following theorem:

**Theorem 6** If \( \sum_{t=1}^{\infty} E \left( \frac{a_t}{A_t} \right)^2 < \infty \), then \( \text{CAR}_T \to 0 \), almost surely.

**Proof.** Note \( \text{CAR}_T = (\sum_{t=1}^{T} a_t r_t) / A_T \), to show \( \text{CAR}_T \to 0 \), almost everywhere, by Kronecker Lemma we only need to show that \( \sum_{t=1}^{\infty} \frac{a_t r_t}{A_t} \) converges, almost surely.
Note that \( \{\sum_{u=1}^{t} a_u r_u / A_u, \sigma(a_1, \cdots, a_t, A_1, \cdots, A_t)\} \) is a martingale, by the \( L^2 \)-Bounded Martingale Convergence Theorem (see Theorem 2.6 in Steele (2001)), we only need to show: \( \exists B < \infty \), such that

\[
E \left[ \sum_{u=1}^{t} \frac{a_u r_u}{A_u} \right]^2 \leq B < \infty \quad \forall t
\]

Since the \( r_u \) are i.i.d.,

\[
E \left[ \sum_{u=1}^{t} \frac{a_u r_u}{A_u} \right]^2 = \sigma^2 \sum_{u=1}^{t} E \left( \frac{a_u}{A_u} \right)^2
\]

which completes the proof of the theorem. \(^5\)

We now apply this theorem to the following lognormal model that we will analyze in some details:

\[
\log N_{t+1} = \theta \log N_t + \delta r_t
\]

This lognormal model ensures positive number of events but allows for both stationary and non-stationary models. Further it allow past returns to persistently impact the number of events. First, we show that asymptotically the bias disappears when \( \theta < 1 \) because when \( \theta < 1 \), the lognormal model satisfies the conditions of the convergence theorem.

**Corollary 7** If \( \theta < 1 \), the lognormal model for the number of events satisfies the assumptions of the Theorem 6 and hence the event abnormal return converges to zero as the number of observations \( T \) goes to infinity.

**Proof.** See Appendix. \( \square \)

We confirm the asymptotic theory we have just derived by simulation. We consider the lognormal model for the number of events

\[
\log N_{t+1} = \theta \log N_t + \delta r_t,
\]

where the parameter \( \theta \) is chosen to be one of 0.2, 0.4, 0.6, 0.8 or 1.0 and the parameter \( \delta \) is chosen to be one of 0.2, 0.4, 0.6 or 0.8. Here \( r_t \) is assumed to be i.i.d. normal, with mean zero and standard deviation of 0.05 (which is very close to actual data). \( N_0 = 1 \), and \( r_0 \) is randomly drawn from \( N(0, 0.05) \).

For a given \( \delta \) and \( \theta \), we do the simulation for 100 rounds. At each round, we simulate the data for a period of \( T = 100,000 \) and save the abnormal return for period of 1000, 2000, 3000, ..., 100000 respectively. Figure 1 presents these results.

**Insert Figure 1 Here**

As can be seen from this simulation evidence from Figure 1 above, for \( \theta < 1.0 \), the bias goes to zero asymptotically; for \( \theta = 1.0 \), the negative bias persists asymptotically and gets more negative for bigger \( \delta \). Our asymptotic theory justifies the simulation behavior for \( \theta < 1.0 \).

We next derive the asymptotic distribution of \( \overline{CAR}_T \) for the case where \( \theta < 1.0 \). This provides a complete asymptotic theory for the lognormal model for the number of events.

---

\(^5\)We note that the asymptotic theorem we present is not necessarily the only possible convergence theorem and other theorems can provide different conditions. We present in the Appendix another theorem due to Taylor and Calhoun (1983) that provides some tradeoff between the restrictions on the event process and restrictions on the return process.
Theorem 8  For \( \log N_{t+1} = \theta \log N_t + \delta r_t \), if \( \theta < 1 \), let \( n_a \triangleq \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \exp(\delta^2 \sigma_r^2 \frac{1-\theta^t}{1-\theta_T}) \), \( \sigma_a^2 \triangleq \lim_{T \to \infty} \frac{\sigma_r^2}{T} \sum_{t=1}^{T} \exp(2\delta^2 \sigma_r^2 \frac{1-\theta^t}{1-\theta_T}) \), then

(1) for the monthly cumulative abnormal returns \( CAR_T(1) \), we can show that

\[
\sqrt{T} CAR_T(1) \xrightarrow{L} N \left( 0, \frac{\sigma_a^2}{n_a^2} \right)
\]

(2) for the \( s \)-month cumulative abnormal returns \( CAR_T(s) \), we can show that

\[
\sqrt{T} CAR_T(s) \xrightarrow{L} N \left( 0, \frac{s^2 \sigma_a^2}{n_a^2} \right)
\]

Proof. See Appendix.

This theorem is important because it provides the asymptotic variance of the event abnormal return. Hence if we can estimate the parameters \( \theta \) and \( \delta \), we can compute the confidence interval for long run returns using this asymptotic distribution. Later in the paper, we use this approach to undertake some hypothesis testing and consider the power of tests with endogenous events.

5 Schultz’s example

In his paper in the Journal of Finance, Schultz (2003) presents an example in Section I that seems to justify his approach. In this example, he shows that event returns are negative and become more negative the longer the horizon he chooses. We consider his example in a formal way and show the following. In his example, abnormal events are always negative. However, the example is flawed in the following sense. 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 events process which we believe is stationary or growing with the economy.

As in Schultz (2003), we assume that each period, excess returns for recent issuers are either positive five percent or negative five percent, and that positive and negative excess returns are equally likely.\(^6\) We assume that the number of offerings each period is determined solely by the price level of potential issuers. If stock prices rise, the number of offerings the next period increases by \( \Delta \% \). If prices fall, the number of offerings decreases by \( \Delta \% \), here we suppose \( \Delta = 5 \% \).

We consider the time series with length \( T \). In each period \( t \in \{1, ..., T\} \), let \( I_t \) indicate the positive or negative excess return by +1 or -1, and \( N_t \) denote the number of new offerings. Thus, \( N_0 \) is set to one and

\[
N_t = N_{t-1}(1 + \Delta I_t), \quad t = 1, ..., T - 1
\]

Moreover, the cumulative abnormal return is the following:

\[
CAR_T(1) = 0.05 * \frac{N_0 I_1 + \ldots + N_{T-1} I_T}{N_0 + \ldots + N_{T-1}}
\]

\[
= \frac{0.05 \Delta}{N_0 + \ldots + N_{T-1}} (\text{note } N_{t-1} I_t = \frac{N_t - N_{t-1}}{\Delta})
\]

\(^6\)We consider the second example that Schultz simulates from (see Section I in his paper). Schultz’s first example can be subsumed in our approach by using asymmetrical up and down jumps in the number of events. We prove results related to the first example later in this section.
Proposition 9 If $\Delta > 0$, then
\[
\Pr(CAR_T(1) \geq 0) < \frac{1}{2}
\]

Proof. See Appendix.

To understand this result, consider the example shown in Figure 2 that shows the evolution of the number of IPOs as in Schultz (2003). From the example it is clear when we have a martingale as assumed above, the number of IPOs will drift downwards. We start with $N_0 = 100$ and with probability of half the number of events goes up to 105 and with probability of half the number of events goes down to 95 in the first period. In the second period, there are three possibilities. With probability $1/4$, $N_2 = 110.25$, with probability $1/2$, $N_2 = 99.75$ and with probability $1/4$, $N_2 = 94.05$. From this it is clear that the number of events is drifting downwards. This implies from Equation (2) that the probability that the cumulative average return over two periods is less than 0 is more than 1/2.

Insert Figure 2 Here

Thus the example suggests that the number of events $N_t$ is drifting downwards. We prove formally next $N_T$ converges to zero in probability.

Theorem 10 $N_T$ converges to zero in probability\(^7\).\n
Proof. See Appendix.

Some intuition for this theorem is as follows.

\[
\begin{align*}
\log N_{t+1} &= \log N_t + \log(1 + \Delta I_t) \\
&= \log N_0 + \sum_{i=0}^{t} \log(1 + \Delta I_i) \\
&= \log N_0 + t \left( \frac{1}{t} \sum_{i=0}^{t} \log(1 + \Delta I_i) \right)
\end{align*}
\]

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 downwards. Further by the SLLN, 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.

Hence we have shown that in Schultz’s example, 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 the with probability close to 0 it takes infinite value and with probability close to 1 it takes zero value in the limit. Further if we start with a positive number of events, the event return is always negative. Our results here are robust to the asymmetrical shocks in the number of events process.

We do not view this example as being an appropriate description of the event generating process for IPOs or any corporate events which we believe are more likely to be stationary or grow over over time. In the log normal model we consider for the number of events, the unit root process that we allow for has the variance increasing over time.

\(^7\)The result is very robust in some sense. For example, when we assume the asymmetric shocks (up $(1 + x)$ with probability $\frac{x}{x+y}$; down $(1 - y)$ with probability $\frac{y}{x+y}$), the number of IPOs still asymptotically goes to zero in probability.
6 Asymptotic Theory For Unit Root Events Process

In a second set of regressions Schultz (2003) runs a regression of events (the number of IPOs) against the level of the IPO index and market index. Since index levels are unit root processes, it seems as if the simulations based on the fitted regression assumes that events follow a unit root process.\(^8\) Strictly speaking, this case does not fall under our convergence theorems because here events are assumed to be normally distributed and hence can be negative, however we can directly prove convergence when the event process is stationary. We analyze the case with normally distributed events and unit roots as there is a well developed theory due to Phillips (1987) and others. Hamilton (1994) summarizes the theory. Then we discuss some generalizations to nonlinear functions of a unit root variable.

We assume that the abnormal returns are distributed according to

\[ N_{t+1} - u = \theta(N_t - u) + \delta r_t \tag{3} \]

This model satisfies the affiliation assumption and using the results in Section 3 has a negative expected bias. Although for a stationary process of \(N_t\) (i.e. \(\theta < 1\)) the bias disappears asymptotically, it is unclear what happens when \(N_t\) follows an unit root process. We analyze the unit root case following Hamilton (1994).

\textbf{Theorem 11} If \(\theta = 1\), then

\[ \sqrt{T} \text{CAR}_T \xrightarrow{L} \sigma_r \left\{ \frac{[W(1)]^2 - 1}{2 \int_0^1 W(s)ds} \right\}. \]

Moreover,

\[ \text{CAR}_T \xrightarrow{P} 0 \]

\textbf{Proof.} See Appendix. \(\blacksquare\)

For a unit root process of \(N_t\), there is no asymptotic bias. Hence even with unit roots we can get asymptotic convergence. However, the number of events can be infinitely negative which seems inconsistent with the data. We do note that the asymptotic distribution in the unit root case does have negative mean suggesting that inference is different in this case. We next consider the lognormal model that we have introduced in the previous section.

For the lognormal model where \(\log(N_{t+1}) - u = \theta(\log(N_t) - u) + \delta r_t\), we show in the following theorem that when \(\theta = 1\), the expected CAR asymptotically converges to a negative constant, as opposed to zero for a stationary event process (see Corollary 7).

\textbf{Theorem 12} when \(\theta = 1\), \(E[\text{CAR}_T]\) converges to \(-\delta \sigma_r^2 / 2\).

\textbf{Proof.} See Appendix. \(\blacksquare\)

Overall our results suggest that even when the event process is non-stationary, asymptotic convergence may or may not occur. Overall, the results strongly suggest that if the number of events process is stationary, the negative expected abnormal return is essentially a small sample problem. Consequently, we study the small sample properties of event abnormal returns more carefully in the next section.

\(\text{\textsuperscript{8}}\)It is possible to have the number of events process to be stationary while the index processes have unit roots if a linear combination of the IPO index and stock price index is co-integrated. Using data simulated from the specification used by Schultz in his simulation we find that it is hard to reject the unit root hypothesis.
7 Small Sample Theory

We provide small sample theory using the following approach. We use Stein’s lemma (1972) for the lognormal model for events:

\[ \log(N_{t+1}) - u = \theta(\log(N_t) - u) + \delta r_t + \epsilon_t \]

where \( r_t \) can be considered as some benchmark-adjusted IPO return, or abnormal return; \( \theta > 0, \delta > 0 \) are assumed to capture the positive effect of previous \( N_t \) and \( r_t \). And \( \{r_t\}_1^\infty \) follows a zero mean random walk process.

Note, we have \( E[\log(N_t)] = u \) and \( \text{Var}[\log(N_t)] = (\delta^2 \sigma^2_t + \sigma^2_t)/(1 - \theta^2) \).

Using Stein’s formula (which is stated in Appendix), we can prove the following result.

**Theorem 13** Under the assumptions above,

\[
E[\text{CAR}_T] = -\delta \sigma^2_T \sum_{t=1}^{T-2} \sum_{s=t+2}^T \theta^{s-t-2} E \left[ \frac{e^{n_t} e^{n_s}}{\left( \sum_{s=1}^T N_s \right)^2} \right] < 0
\]

**Proof.** See Appendix. ■

Hence we have an exact expression for the expected mean cumulative abnormal returns. Similar expressions can be found for longer holding periods. Using these exact expressions we provide the following results.

First, we evaluate the conditional expectation using the approach suggested above for samples of size \( T = 200, 400 \) and \( 600 \) respectively. To be consistent with actual data, we consider the parameter values \( \theta = 0.6, 0.8, 0.85, 0.9, 0.95 \) 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). For each pair of parameters, we evaluate the expectation by simulation. We do 500 rounds of simulations and results are recorded in the tables below\(^9\):

**Insert Table 1 Here**

**Insert Table 2 Here**

From these small sample simulations, we can see that the average abnormal returns are negative and tend to get more negative, as \( \theta \) increases (the persistence of events is higher), or as \( \delta \) increases (the relation between returns and subsequent number of events is stronger). Also it becomes less negative when the number of sample observations \( T \) increases except for the case of \( \theta = 1.0 \). 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 \( \theta < 1.0 \).

At first cut, our approach does not support as large negative expected abnormal returns as Schultz finds in his simulations except in the case of unit roots in the log events process. Schultz (2003) finds magnitudes of -0.12 (-12\%) in Table VI in his paper for 3 year cumulative abnormal

\(^9\)Using Theorem 13, 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 the Stein’s method delivers very accurate estimates of the expected event abnormal return.
returns which is closer to our unit root magnitudes of -0.23(-23%) that we obtain (see Table 2 for this). When we consider the case with $T = 400$ observations and assume that $\theta = 0.95$ and $\delta = 1.75$ we obtain magnitudes of -0.029 (-2.9%) for the expected event abnormal return which are much smaller. With $T = 200$ observations and same parameters we obtain magnitudes of -0.035 (-3.5%) while the unit root magnitudes are -0.16 (-16%). Figure 3 illustrates these numbers graphically and shows that the exact small sample bias is very sensitive to the assumption of an unit root such that even a small deviation from the unit root hypothesis leads to a much smaller sample bias. Hence sample size and the stationarity of the log number of events process play an important role in determining the expected bias.

Insert Figure 3 Here

For the average buy-and-hold abnormal return, we do 300,000 rounds of direct simulation of all the relevant forcing processes. Because of the multiplicative nature of buy and hold returns, we cannot use the simplification obtained from Stein’s formula. The results we obtain are shown in the following table:

Insert Table 3 Here

The results are consistent in magnitude with that obtained for CARs. With $T = 400$ observations, $\theta = 0.95$ and $\delta = 1.75$, we obtain magnitudes of -0.0206 (-2.06%) versus -0.166 (-16.6%). With $T = 200$ observations we obtain magnitudes of -0.0355 (-3.55%) versus -0.151 (-15.1%). Again the magnitudes are smaller by a factor of $1/5^{th}$. This suggests that unless we have a unit root in the log events process or are very close to it ($\theta = 0.99$), 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 4, 5 and 6). In particular, we study the following Two-Lag model:

$$\log(N_t) = \theta_1 \log(N_{t-1}) + \theta_2 \log(N_{t-2}) + \delta r_{IPO,t-1}$$

Insert Table 4, 5 & 6 Here

What matters here is the sum of the coefficients $\theta_1$ and $\theta_2$. When the sum is unit, we obtain significant negative returns. Away from unit, the expected returns are negative but the magnitudes are not as large. The small sample simulations show conclusively that one’s priors on 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.

8 Does the Data Support the Unit Root Hypothesis?

8.1 Testing For Unit Roots

We now examine whether the data supports the unit root hypothesis for the number of events process. First we describe the data. The sample is comprised of 9,190 initial public offerings ranging from February 1973 to December 2002. The selection criteria are the same as in Ritter (1991): (1)
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). Table 7 shows the distribution of the number of offerings each month.

Insert Table 7 Here

We conduct both Augmented Dickey-Fuller and Phillips-Perron tests on the number of IPO offerings against the following two nulls with or without time trend.

\[ H_0: \text{constant term, without time trend, unit root} \]
\[ H'_0: \text{constant term, time trend, unit root} \]

Insert Table 8 Here

These tests suggest that with one lag one can reject the unit root hypothesis, however we need to check their robustness to more lags. We consider next what happens when we allow for more lags in the autoregression. 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 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 9A & 9B. In discussing our results we focus on the ERS test (the local root to unity approach) as this has the highest power.

Insert Table 9A&9B Here

Therefore, from the two tests above in Table 8A and 8B, when considering only one lag we can reject the null hypothesis of unit root process for both \( N_t \) and \( \log(N_t) \), the \( p \)-values are much less than 1%. However from Table 9A and 9B, with more lag lengths we tend to not reject the unit root hypothesis at the 1% level and reject it (with no time trend, ERS test) at the 5% level.10 Since the null hypothesis is the unit root and it is well known that the power of these tests is low at higher lag lengths, 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). Hence one’s prior and economic arguments would determine whether one believes there is an unit root or not.11

### 8.2 VAR specifications

We consider the following VAR for returns and \( \log(\text{IPO numbers}) \). To adjust for continuity, we add 1/2 to IPO numbers.12

---

10 the 5% and sometimes even at the 10% level.
11 Adding more lags to our model will not change the bias calculations that we undertake. This is confirmed by our results for 300,000 rounds of simulations (see Table 4, 5 and 6) in previous section.
12 We have tried other specifications including a Poisson model and a negative binomial model. Those specifications account for the discrete nature of the number of events. The current specification is closest to that used in Schultz (2003).
\[
\begin{pmatrix}
    r_{m,t} \\
    r_{IPO,t} \\
    \log(N_t + \frac{1}{2})
\end{pmatrix}
= U + B
\begin{pmatrix}
    r_{m,t-1} \\
    r_{IPO,t-1} \\
    \log(N_{t-1} + \frac{1}{2})
\end{pmatrix}
+ \begin{pmatrix}
    \epsilon_1 \\
    \epsilon_2 \\
    \epsilon_3
\end{pmatrix}
\] (4)

where

\[ U = (u_1, u_2, u_3)^T, \quad B = \begin{pmatrix}
    b_{11} & b_{12} & b_{13} \\
    b_{21} & b_{22} & b_{23} \\
    b_{31} & b_{32} & b_{33}
\end{pmatrix}. \]

We use this representation rather than that in Schultz (2003) for two reasons. First, level variables with unit roots on the right hand side as used in Schultz lead to spurious t-statistics (see Granger and Newbold (1974)). Second a VAR accounts for all the relationships that we desire to look at in a parsimonious way. We consider two specifications, one in which we restrict some coefficients using a priori theory and another where there are no restrictions. The restriction that we impose a priori is that events do not predict returns which is true under the null hypothesis.

For the restricted VAR with \(b_{12} = b_{13} = b_{23} = 0\), we find that

Insert Table 10 Here

For the unrestricted VAR, we have very similar results as the restricted one:

Insert Table 11 Here

These results support our use of \(\theta = 0.9\) and \(\delta = 1.75\) in the small sample computations that we did in the previous section. With these computations, we find biases that are at most 1/8th of that in the unit root case. This suggests that the sample size and the degree of persistence play an important role on how large the bias is. The bias found in Schultz is most consistent with a unit root in the log events process, an autoregressive coefficient of 0.9 yields much smaller numbers.

Our specification can be contrasted to that used in Schultz (2003). Our specification seems to fit the data better because we use the lagged number of events (see Figure 4). If we simulate 5000 times from the specification used by Schultz (2003) and fit the data using our log number of events specification, we find that the median autoregressive coefficient is close to 1.0, i.e., it is a unit root\(^{13}\). Using 1000 trials from the simulation, we conduct unit root tests on the number of events (IPOs) process. We cannot reject the null hypothesis of an unit root using either the ADF test or the Phillips-Perron test in 98% of the trials. For 238 simulations, where the number of events is always positive (in Schultz’s approach the number of events can be negative), we also test for a unit root in the log number of events process. We find that we cannot reject the null hypothesis of an unit root using either the ADF test or the Phillips-Perron test in 98% of these trials with positive simulated IPO numbers. Thus Schultz’s simulation data is consistent with a unit root in the number of events process. We believe this unit root behavior occurs because Schultz uses level variables (the level of the IPO index, the level of the market index) which have unit roots on the right hand side of his regression when the dependent variable (the number of IPOs) is a stationary process.\(^{14}\)

Insert Figure 4 Here

\(^{13}\) Some descriptive statistics of the autoregressive coefficient are the following: min=0.895, max=1.060, median=1.001, mean=1.000.

\(^{14}\) We can have a stationary dependent variable and non-stationary independent variables if the right hand side variables are co-integrated. Such an approach is different from that used in Schultz (2003) and seems much more cumbersome than our approach that directly uses the number of events and returns.
8.3 Regime Shifts

In determining the data generating process, we are restricting ourselves to the class of autoregressive processes. One view of the data is that there are two regimes, a "hot" market regime and a "cold" market regime. This distinction between “hot” IPO markets and “cold” IPO markets is often used in the empirical literature (see Ritter (1984) for example). It is known since Perron (1989) and further confirmed in the more recent work of Nelson, Piger and Zivot (2001) that tests of unit roots will accept the unit root too often against an alternative that has regime switching between stationary AR (1) processes (both autoregression coefficient and variance switch between regimes). Hence we study whether the data has regime shifts between a cold regime and a hot regime.

To answer the question of whether the data is better described by a regime shifting model, we study the following Markov regime switching model:

\[
\begin{align*}
\log(N_{t+1} + 0.5) &= a_1 + b_1 \log(N_t + 0.5) + cR_{IPO,t} + \epsilon_1, \text{ Regime 1} \\
\log(N_{t+1} + 0.5) &= a_2 + b_2 \log(N_t + 0.5) + cR_{IPO,t} + \epsilon_2, \text{ Regime 2}
\end{align*}
\]

We allow state-contingent volatility in the model, i.e., \(\text{var}(\epsilon_1)\) and \(\text{var}(\epsilon_2)\) could differ. For simplicity, we make the coefficient of the return of IPO index constant cross regimes. Maximum likelihood estimates of this model are given in Table 12A.\(^1\)

\textbf{Insert Table 12A Here}

Here regime 1 corresponds to the cold regime (it has low number of IPOs) while regime 2 corresponds to the hot regime (it has a high number of IPOs). Not surprisingly the hot regime is more persistent and has a much higher autoregressive coefficient. Importantly, regime switching attenuates the autoregressive coefficients a lot, they are clearly different from one. Further the results of the likelihood ratio test imply that this Markov regime switching model fits the data better than the one with only one regime.

\textbf{Insert Table 12B Here}

We plot the regime probabilities are plotted in Figure 5. Clearly, regime probabilities track the cold regimes in the early 1980s and late 1990s/early 2000s quite well. The cold regime in the early 1990s that was short is barely picked up.\(^2\) Also we plot the predicted logarithm of IPO numbers based on our regime switching model as in Figure 6.

\textbf{Insert Figure 5&6 Here}

We also simulate from the regime shift model that we estimate to see whether the simulated data looks like the actual IPO data. While we do not present these results, it is clear from the unit root tests that we reject the unit root hypothesis at one lag but at higher lags the evidence is very mixed. This is consistent with the actual data and thus we have some confidence that with regime shifts one could believe that there is an unit root when there is really a stationary regime shift process. All this suggests to us that a stationary regime shift model is a distinct possibility that cannot be ruled out relative to the hypothesis of a close to unit root process in a single regime.

\(^1\)To ensure that we have found a global maxima we used a simulated annealing algorithm.

\(^2\)We also looked at the specification without the stock return. This tends to pick up the short cold regime in the early 1990s better.
9 Confidence Intervals and the Power of Event Studies

While the magnitude of the expected bias has been the focus of most of this paper, the sampling variation in the event abnormal return is also important. Since the weights in the abnormal returns are random and determined by the event process, the usual asymptotic variance that is used in event studies is incorrect. In particular, event studies ignore the variation in the number of events over time.

In Section 4, we determined the asymptotic variance as a function of the parameters $\theta$ and $\delta$. In Figure 7, we show how this varies with the parameters $\theta$ and $\delta$. It is clear as these parameters increase (the persistence in the number of IPOs goes up and the relationship between returns and subsequent number of IPOs is stronger), the variance of the event abnormal return goes up dramatically. Hence exact computation of the variance is important.

We find confidence intervals for the event abnormal return estimator as follows. We substitute our estimated values of $\theta = 0.88991$, $\delta = 1.85763$, $\sigma^2_r = 0.006790$ we can determine the asymptotic variance using Theorem 8. We derive using Theorem 8 that $n_a \approx 1.0579$ and $\sigma^2_a \approx 0.008506$. The observed average monthly cumulative abnormal return is: $\text{CAR}_{T}(1) = -0.97\%$ with $T = 357$. By Theorem 8, the theoretical 95% confidence interval is: $[-0.46\%, 0.46\%]$. Hence, we reject the null hypothesis that $E[\text{CAR}_{T}(1)] = 0$. Moreover, for 3-year cumulative abnormal return $\text{CAR}_{T}(36)$, the observed value is: $-9.44\%$ with $T = 322$ and the associated theoretical 95% confidence interval is: $[-17.49\%, 17.49\%]$. In this case, the null is accepted.

One can ask how these confidence intervals compare with that traditionally used in event studies. Traditionally, event studies use confidence intervals based on the idea that events are deterministic. We compare the traditional approach (deterministic events) with the confidence interval with endogenous events using the our model that

$$\log N_{t+1} = \theta \log N_t + \delta r_t$$

Hence the randomness in events is generated by the noise $r_t$ and if the coefficient $\delta$ is not equal to zero, the events will be stochastic. Hence a comparison of stochastic events versus deterministic events can be made by comparing what happens to the variance of event abnormal returns when the coefficient $\delta$ is changed. Some simple calculation shows that the variance of event abnormal returns with stochastic events is much higher than that due to deterministic events (the usual confidence intervals are understated). Figure 8 shows the ratio of the two variances and the asymptotic variance when $\theta = 0.95$, $\delta = 1.75$ and $\text{var}[r_t] = 0.679$. From the figure it is clear that the asymptotic confidence intervals are much larger with stochastic events as the ratio increases as the sample size increases (though it eventually converges). This makes very clear that one cannot use deterministic confidence intervals in event studies when events are stochastic.

This computation suggests that endogenous events increase the asymptotic variance considerably and consequently event studies have low power. This issue of the low power of event studies has not received prior attention in part because prior studies have failed to account for the variation in the number of events and the correlation between the returns and subsequent number of events.
10 Conclusions

We present a theory of event abnormal returns when returns predict events. We prove that expected abnormal returns are negative and become more negative the longer the holding period. This suggests that there is a small sample bias in the use of event returns. However, asymptotically we show that the bias dissappears because event abnormal returns converge in probability to zero under some assumptions on the return generating process and the number of events process. The intuition for this is as follows. Since the number of events process is stationary, the long run average number of events is not affected by the current shock to returns, i.e., the effect of the shock dissipates over time. Thus the stationarity assumption on the number of events process is sufficient to generate consistency of event abnormal returns.

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. 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 non-stationary case convergence may or may not occur depending on the specific process that is used. Using this lognormal model, we look at the data. We show that the number of events process does not have a unit root using standard unit root tests with one lag in the autoregression. At higher lags, the power of the unit root test falls and we are less able to discriminate between the unit root hypothesis and its alternatives. To further analyze this, we use Stein's method and compute the small sample expected bias in the lognormal model using the parameters from the vector autoregression. We show that the sample size and the degree of persistence in the log events process determines whether the expected small sample bias is large or not — a small deviation from the unit root hypothesis reduces the expected bias a lot. We also deduce the asymptotic sampling variance and show that under this variance the usually observed negative abnormal returns are not significantly different from zero. This is because the dependence of events on returns increases the confidence interval and lowers the power of the test. We believe that the issue of power merits more attention as it suggests that event studies cannot resolve the issue of whether IPOs underperform in the long run.
REFERENCES

References


Appendix On Affiliation and Stein’s Formula

Affiliation
Let \( I_t \) denote the information available at the beginning of period \( t \), i.e., \( I_t = \{ r_{m,t}, \cdots, r_{m,1}, r_{IPO,t}, \cdots, r_{IPO,1} \} \), and \( I_0 = \emptyset \) (empty set).

For notation’s simplicity, we use \( f(\cdot) \) to denote various density functions.

**Definition 14** A subset \( A \) of \( R^k \) is called **increasing** if its indicator function \( 1_A \) is nondecreasing.

**Definition 15** A subset \( S \) of \( R^k \) is a **sublattice** if its indicator function \( 1_S \) is affiliated, i.e., if \( z \wedge z' \) and \( z \vee z' \) are in \( S \) whenever \( z \) and \( z' \) are.

**Definition 16** \( Z_1, \cdots, Z_k \) are **affiliated** if for all increasing sets \( A \) and \( B \) and every sublattice \( S \),
\[ P(AB|S) \geq P(A|S)P(B|S), \]
i.e., if the variables are associated conditional on any sublattice.

For affiliated random variables, we have the following established theorems (please refer to Milgrom and Weber (1982)):

**Theorem 17** If \( Z_1, \cdots, Z_k \) are affiliated and \( g_1, \cdots, g_k \) are all nondecreasing functions (or all non-increasing functions), then \( g_1(Z_1), \cdots, g_k(Z_k) \) are affiliated.

**Theorem 18** If \( Z_1, \cdots, Z_k \) are affiliated, then any subset of \( Z_1, \cdots, Z_k \) are affiliated.

**Theorem 19** *(The Key Theorem)* The following statements are equivalent.

(i) \( Z_1, \cdots, Z_k \) are affiliated.

(ii) For every pair of nondecreasing functions \( g \) and \( h \) and every sublattice \( S \),
\[ E[g(Z)h(Z)|S] \geq E[g(Z)|S]E[h(Z)|S] \]

(iii) For every nondecreasing functions \( g \), increasing set \( A \), and every sublattice \( S \),
\[ E[g(Z)|A_S] \geq E[g(Z)|S] \geq E[g(Z)|\neg A_S] \]

**Stein’s Lemma**
Let \( X = (X_1, \cdots, X_n) \) be multivariate normally distributed with arbitrary mean vector \( u \) and covariance matrix \( \Sigma \). For any function \( h(x_1, \cdots, x_n) \) such that \( \partial h / \partial x_i \) exists almost everywhere and \( E \left| \frac{\partial}{\partial x_i} h(X) \right| < \infty, i = 1, \cdots, n \), we write \( \nabla h(X) = \left( \frac{\partial}{\partial x_1} h(X), \cdots, \frac{\partial}{\partial x_n} h(X) \right)^T \). Then the following identity is true:
\[ \text{cov}[X, h(X)] = \Sigma E[\nabla h(X)] \]

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

Please refer to Stein (1972) or Liu (1994).
Appendix On Proofs

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

\[
f(N_{t+1}, N_t, 1 + r_{m,t+1}, 1 + r_{m,t}, 1 + r_{IPO,t+1}, 1 + r_{IPO,t})
= f(1 + r_{m,t+1}, 1 + r_{IPO,t+1})f(N_{t+1}, N_t, 1 + r_{m,t}, 1 + r_{IPO,t}) \quad \text{(by assumptions 1 & 2)}
= f(1 + r_{m,t+1}, 1 + r_{IPO,t+1})f(N_{t+1})f(N_t, 1 + r_{m,t}, 1 + r_{IPO,t})
\]

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

\[
f(N_{t+s+1}, \ldots, N_t, 1 + r_{m,t+s+1}, \ldots, 1 + r_{m,t}, 1 + r_{IPO,t+s+1}, \ldots, 1 + r_{IPO,t})
= f(N_{t+s+1}, \ldots, N_t, 1 + r_{m,t+s}, \ldots, 1 + r_{m,t}, 1 + r_{IPO,t+s}, \ldots, 1 + r_{IPO,t})
= f(1 + r_{m,t+s+1}, 1 + r_{IPO,t+s+1})
\]

Hence, \( (N_{t+s+1}, \ldots, N_t, 1 + r_{m,t+s+1}, \ldots, 1 + r_{m,t}, 1 + r_{IPO,t+s+1}, \ldots, 1 + r_{IPO,t}) \) are also affiliated.

Proof of Theorem 3.

\[
E[BHAR_T(s)] = E\left[\frac{\sum_{i=0}^{T-1} N_i \left( \prod_{j=1}^{s} (1 + r_{IPO,i+j}) - \prod_{j=1}^{s} (1 + E[r_{IPO,i+j}]) \right)}{\sum_{i=0}^{T-1} N_i}\right]
\]

Note, by affiliation, we have

\[
E\left[\frac{N_i \prod_{j=1}^{s} (1 + r_{IPO,i+j})}{\sum_{i=0}^{T-1} N_i}\right] = E\left[\frac{N_i \prod_{j=1}^{s-1} (1 + r_{IPO,i+j}) E^{\left(1 + r_{IPO,i+s}\right)}_i}{\sum_{i=0}^{T-1} N_i}\right]
\]

\[
\leq E\left[\frac{N_i \prod_{j=1}^{s-1} (1 + r_{IPO,i+j}) (1 + E[r_{IPO,i+j}])}{\sum_{i=0}^{T-1} N_i}\right] \leq \cdots \leq E\left[\frac{N_i \prod_{j=1}^{s} (1 + E[r_{IPO,i+j}])}{\sum_{i=0}^{T-1} N_i}\right]
\]

Therefore, \( E[BHAR_T(s)] \leq 0 \)
Proof of Theorem 5.

\[ E[BHAR_{T+1}(s+1)] \]

\[ = \sum_{i=0}^{T-1} E \left[ \frac{N_i \prod_{j=1}^{s+1} (1 + r\text{IPO},i+j)}{\sum_{i=0}^{T-1} N_i} \right] - \sum_{i=0}^{T-1} E \left[ \frac{N_i \prod_{j=1}^{s+1} E[1 + r\text{IPO},i+j]}{\sum_{i=0}^{T-1} N_i} \right] \]

\[ \leq \sum_{i=0}^{T-s} E[1 + r\text{IPO},i+s+1] E \left[ \frac{N_i \prod_{j=1}^{s+1} (1 + r\text{IPO},i+j)}{\sum_{i=0}^{T-1} N_i} \right] - \sum_{i=0}^{T-1} E \left[ \frac{N_i \prod_{j=1}^{s+1} E[1 + r\text{IPO},i+j]}{\sum_{i=0}^{T-1} N_i} \right] \]

\[ = (1 + E[r\text{IPO}]) E[BHAR_T(s)] \]

A Second Asymptotic Theorem:

For the sake of completeness we state a theorem due to Taylor and Calhoun (1983). This provides a different theorem that allows us to analyze weighted random sums of the kind used in event studies.

Stochastic Boundedness: A sequence of random elements \( \{X_t\} \) is said to be stochastically bounded by a positive random variable \( X \) (we state this as \( \{X_t\} \leq X \)) if \( \sup_t P(|X_t > x|) \leq P[X > x] \) for each \( x > 0 \).

The theorem that we state below consider weighted sums of the kind \( A_T^{-1} \sum_{t=1}^{T} a_t X_t \) where \( A_t \) and \( a_t \) are positive random variables and \( A_t/a_t \rightarrow \infty \).

Further if \( \sup_t A_T^{-1} \sum_{t=1}^{T} a_t < M \), then it follows that \( A_T^{-1} \sum_{t=1}^{T} a_t X_t \rightarrow 0 \) almost surely.

Theorem 20 (Taylor-Calhoun): Let \( E \) be a \( p \)-smoothable space, \( 1 < p \leq 2 \). Let \( \{X_t\} \) be a sequence of random elements in \( E \) such that \( \{X_t\} \leq X \) (in the sense of stochastic boundedness) with \( EX < \infty \) and \( EX_t = 0 \) for each \( t \). Assume that \( X_{t+1} \) is independent of \( G_t = \sigma\{a_1/A_1, ..., a_{t+1}/A_{t+1}, X_1, ..., X_t\} \) for each \( t \). If

\[ EN(X) = \sum_{t=1}^{\infty} E[I_{\{X \geq A_t/a_t\}}] < \infty \]

and

\[ \int_0^{\infty} t^{p-1} P[X > t] \int_t^{\infty} \frac{EN(y)}{y^{p+1}} dy dt < \infty \]

then

\[ A_T^{-1} \sum_{t=1}^{T} a_t (X_t - V_t) \rightarrow 0 \) almost surely.

First we show that abnormal returns fit the requirements of the above theorem. In our setup \( X_t = (1 + r\text{IPO},t+1) - (1 + E[r\text{IPO},t+1]) \), which is independent and identically distributed, so we can
choose $X = |X_1|$ and satisfy the stochastic boundedness requirement that $\{X_t\} \leq X$. Further we have that $a_t = N_t$ and $A_T = \sum_{t=1}^{T} a_t$. If we assume the stationary event process case, it immediately follows that $A_t/a_t \to \infty$. In fact, we can show that a stationary event process is sufficient to satisfy $EX^{1/r} = E||X_1||^{1/r} < \infty$ where $r > 1/p$. Essentially all we require is that returns and the event process be stationary and have first absolute moments of order smaller than one, thus variances have to exist.

**Proof of Corollary 7:**

Since $a_t = \exp((\delta r_{t-2} + \theta r_{t-3} + \cdots + \theta^{t-2} r_0))$, then

$$A_t = \sum_{i=1}^{t} a_i \geq t \left( \prod_{i=1}^{T} a_i \right)^{1/t} = t \exp \left( \sum_{i=2}^{t} \sum_{j=0}^{i-2} \theta^{i-j-2} r_j \right)^{1/t}$$

For $\theta < 1$,

$$A_t = t \exp \left( \sum_{j=0}^{t-2} \frac{1 - \theta^{t-j-1}}{1 - \theta} r_j \right)$$

For $\theta = 1$,

$$A_t = t \exp \left( \sum_{j=0}^{t-2} (t - j - 1) r_j \right)$$

Therefore, for $\theta < 1$,

$$E \left[ \frac{a_t}{A_t} \right]^2 \leq \frac{1}{t^2} E \left[ \exp \left( 2 \sum_{j=0}^{t-2} \frac{1 - \theta^{t-j-1}}{1 - \theta} r_j \right) \right]$$

$$= \frac{1}{t^2} \exp \left( 2 \delta^2 \sigma_r^2 \frac{1 - \theta^{(t-1)}}{1 - \theta} - \frac{2}{t(1 - \theta)} \left( \frac{1 - \theta^{t-1}}{1 - \theta} - \theta(1 - \theta^{2(t-2)}) \right) \right)$$

$$\times \exp \left( \frac{4 \delta^2 \sigma_r^2}{t^2 (1 - \theta)^2} \left( t - 1 - \frac{2\theta(1 - \theta^{t-1})}{1 - \theta} + \frac{\theta^2(1 - \theta^{2(t-2)})}{1 - \theta^2} \right) \right)$$

$$\leq \frac{1}{t^2}$$

Hence

$$\sum_{t=1}^{\infty} E \left( \frac{a_t}{A_t} \right)^2 < \infty$$

Hence, it satisfies the condition in the theorem above, therefore the asymptotic bias will be zero, for $\theta < 1$.

**Proof of Theorem 8** First, we prove that $\{N_t; \Omega_t\}_{t=1}^{\infty}$ is L-mixingale, where $\Omega_t \triangleq \sigma\{N_t, \cdots, N_1\} = \sigma\{r_{t-1}, \cdots, N_0\}$ for $t \geq 1$; $\Omega_0 \triangleq \Phi$, for $t = 0$. 

26
**Proof.** In fact, let $Y_t = N_t - E[N_t] = N_t - \exp\left(\frac{\delta^2 \sigma^2_r 1 - \theta^{2m}}{2} \right)$, then

$$E[Y_t|\Omega_{t-m}] = \exp\left(\frac{\delta^2 \sigma^2_r 1 - \theta^{2m}}{2} \right) \left[ \exp\left(\delta \sum_{i=0}^{t-m-1} \theta^{t-i} r_i\right) - \exp\left(\frac{\delta^2 \sigma^2_r \theta^{2m} - \theta^{2t}}{2} \right) \right]$$

Thus,

$$E[E[Y_t|\Omega_{t-m}]] = \exp\left(\frac{\delta^2 \sigma^2_r 1 - \theta^{2m}}{2} \right) \left[ \exp\left(\delta \sum_{i=0}^{t-m-1} \theta^{t-i} r_i\right) - \exp\left(\frac{\delta^2 \sigma^2_r \theta^{2m} - \theta^{2t}}{2} \right) \right]$$

$$\leq \exp\left(\frac{\delta^2 \sigma^2_r 1 - \theta^{2t}}{2} \right) \left[ \exp\left(\delta \sum_{i=0}^{t-m-1} \theta^{t-i} r_i\right) - \exp\left(\frac{\delta^2 \sigma^2_r \theta^{2m} - \theta^{2t}}{2} \right) \right]^{1/2}$$

$$= \exp\left(\frac{\delta^2 \sigma^2_r 1 - \theta^{2t}}{2} \right) \left[ \exp\left(\delta \sum_{i=0}^{t-m-1} \theta^{t-i} r_i\right) - \exp\left(\frac{\delta^2 \sigma^2_r \theta^{2m} - \theta^{2t}}{2} \right) \right]^{1/2}$$

Note further that $\lim_{m \to \infty} \left[ \exp\left(\delta^2 \sigma^2_r (t - m)\theta^{2m} - \theta^{2t}\right) - \exp\left(\delta^2 \sigma^2_r \theta^{2m} - \theta^{2t}\right) \right] = 0$, hence $\{N_t; \Omega_t\}_{t=1}^\infty$ is L-mixingale.

Similarly, we can prove that $\{N_t^2; \Omega_t\}_{t=1}^\infty$ is L-mixingale.

Note further that $\lim_{T \to \infty} \frac{1}{T} \sum_{i=1}^{T} \exp\left(\frac{\delta^2 \sigma^2_r 1 - \theta^{2t}}{2} \right) < \infty$, which is denoted as $n_a$. By Theorem in Hamilton (1994), we have

$$\frac{1}{T} \sum_{i=1}^{T} N_t \overset{p}{\to} n_a$$

Similarly, we can prove that

$$\frac{1}{T} \sum_{i=1}^{T} N_t^2 \overset{p}{\to} \sigma_a^2$$

Then by Theorem in Hamilton (1994), we have

$$\frac{1}{\sqrt{T}} \sum_{i=1}^{T} N_t r_{t+1} \overset{p}{\to} N(0, \sigma_a^2)$$

Finally, $\sqrt{CAR_T} \overset{L}{\to} N\left(0, \frac{\sigma_a^2}{T}\right)$.

For $s$-month cumulative abnormal returns, we can similarly show that

$$\frac{1}{T} \sum_{t=1}^{T} N_t \left(\sum_{i=1}^{s} r_{t+i}\right)^2 \overset{p}{\to} \tilde{\sigma}_a^2$$

where $\tilde{\sigma}_a^2 = \lim_{T \to \infty} \frac{s^2 \sigma^2_r}{2} T \sum_{t=1}^{T} \exp(2\delta^2 \sigma^2_r 1 - \theta^{2t}) = s^2 \sigma_a^2$. This completes the proof of the theorem. 

**Proof of Proposition 9.** Suppose among $I_1, \cdots, I_T$ there are $i$ "+1" and $(T - i)$ "-1", then $N_T = (1 + \Delta)^i (1 - \Delta)^{T-i}$. And

$$(1 + \Delta)^i (1 - \Delta)^{T-i} \geq 1 \iff \left(\frac{1 + \Delta}{1 - \Delta}\right)^i \geq \left(\frac{1}{1 - \Delta}\right)^T \iff i \geq \frac{T \log\left(\frac{1}{1 - \Delta}\right)}{\log\left(\frac{1 + \Delta}{1 - \Delta}\right)}$$
Let $i_0 = \left\lceil \frac{T \log(\frac{1}{\Delta})}{\log(\frac{1}{1-\Delta})} \right\rceil$, which is the largest integer not exceeding $\frac{T \log(\frac{1}{\Delta})}{\log(\frac{1}{1-\Delta})}$. Because $CAR_T(1) = \frac{0.05 * N_T - N_0}{N_T - N_{T-1}}$, we have

$$\Pr(CAR_T(1) \geq 0) = \Pr(N_T - N_0 \geq 0) = \sum_{i=i_0}^{T} \Pr(i^\ast + 1^\ast, (T - i)^\ast - 1^\ast) = \sum_{i=i_0}^{T} \left( \frac{T}{i} \right) \left( \frac{1}{2} \right)^T.$$

Note, if $\Delta > 0$, then $1 > 1 - \Delta^2$ or equivalently $\log(\frac{1}{1-\Delta}) > \frac{1}{T} \log(\frac{1+\Delta}{\Delta})$. Thus, $i_0 > \frac{T}{2}$.

Because $\sum_{i=0}^{T} \left( \frac{T}{i} \right) = 2^T$ and $i_0 > \frac{T}{2}$, we have $\sum_{i=i_0}^{T} \left( \frac{T}{i} \right) < 2^{T-1}$. Therefore, $\Pr(CAR_T(1) \geq 0) < \frac{1}{2}$.

To prove the next theorem we need the following submartingale convergence theorem (see Billingsley 1995):

**Theorem 21** Let $X_1, X_2, \cdots$ be a submartingale. If $K = \sup_i E[|X_i|] < \infty$, then $X_t \to X$ with probability 1, where $X$ is a random variable satisfying $E[|X|] \leq K$.

**Proof of Theorem 10.** To show that, we need to show $\Pr(N_T \geq \epsilon) \to 0$, $\forall \epsilon > 0$. Suppose among $I_1, \cdots, I_T$ there are $i^\ast + 1^\ast$ and $(T - i)^\ast - 1^\ast$, then $N_T = (1 + \Delta)^{i(1 - \Delta)}$. Hence

$$N_T \geq \epsilon \Leftrightarrow (1 + \Delta)^{i(1 - \Delta)} \geq \epsilon \Leftrightarrow i \geq \frac{\log(\epsilon)}{\log(1-\Delta)} + \frac{T \log(\frac{1}{1-\Delta})}{\log(\frac{1}{1-\Delta})}$$

Let $i_0 = \left\lceil \frac{T \log(\epsilon)}{\log(1-\Delta)} + \frac{T \log(\frac{1}{1-\Delta})}{\log(\frac{1}{1-\Delta})} \right\rceil$, which is the nearest integer no less than $\frac{\log(\epsilon)}{\log(1-\Delta)} + \frac{T \log(\frac{1}{1-\Delta})}{\log(\frac{1}{1-\Delta})}$. Therefore

$$\Pr(N_T \geq \epsilon) = \sum_{i=i_0}^{T} \Pr(i^\ast + 1^\ast, (T - i)^\ast - 1^\ast) = \sum_{i=i_0}^{T} \left( \frac{T}{i} \right) \left( \frac{1}{2} \right)^T.$$

Let $a = \frac{T \log(\frac{1}{1-a})}{\log(\frac{1}{1-a})} \approx 0.5503$, when $\Delta = 0.2$. Note $\forall \epsilon > 0$, we can find $a'$ for large enough $T$, such that, $1 < a' < a$ and $\frac{\log(\epsilon)}{\log(1-a')} + \frac{T \log(\frac{1}{1-a'})}{\log(\frac{1}{1-a'})} = \frac{\log(\epsilon)}{\log(1-a')} + aT \geq a'T$. Thus without loss of generality, we can just let $i_0 = \left\lceil \frac{T \log(\frac{1}{1-a})}{\log(\frac{1}{1-a})} \right\rceil$.

As shown above, $i_0 > \frac{T}{2}$. As a result, we have $\left( \frac{T}{i_0} \right) \leq \left( \frac{T}{i_0} \right)^{aT} = \left( \frac{T}{aT} \right)^i$, for $i \geq i_0 > \frac{T}{2}$.

By Sterling’s Formula, $n! \sim e^{-n}n^n\sqrt{2\pi n}$, we have

$$\left( \frac{T}{aT} \right)^i = \frac{T!}{(aT)!(1-a)T)!} \sim \frac{1}{\sqrt{2\pi a(1-a)T}} \frac{1}{(a^a(1-a)^{1-a})^T}.$$
Then,

$$\Pr(N_T \geq \epsilon) = \sum_{i=0}^{T} \left( \begin{array}{c} T \\ i \end{array} \right) \left( \frac{1}{2} \right)^T \leq \sum_{i=0}^{T} \left( \begin{array}{c} T \\ i_0 \end{array} \right) \left( \frac{1}{2} \right)^T$$

$$\sim \frac{1}{\sqrt{2\pi a(1-a)T}} \frac{1}{(a^a(1-a)^{1-a})^{T/2}} \left( \frac{1}{2} \right)^T$$

$$= \sqrt{\frac{T}{8\pi a(1-a)(2a^a(1-a)^{1-a})^T}}$$

Because $2a^a(1-a)^{1-a} > 1$, if $a > \frac{1}{2}$. In fact, if $a = 0.5503$, $2a^a(1-a)^{1-a} = 1.0051$. Therefore, as $T \to \infty$, then

$$\sqrt{\frac{T}{8\pi a(1-a)(2a^a(1-a)^{1-a})^T}} \to 0$$

So, $N_T \to 0$. ■

**Proof of Theorem 11.** By (3), $N_t - u = N_0 - u + \delta(r_{t-1} + \cdots + r_0)$, or $N_t = N_0 + \delta R_t$, where $R_t = (r_{t-1} + \cdots + r_0)$.

Then

$$\sum_{s=1}^{T} N_s = TN_0 + \delta \sum_{s=1}^{T} R_s$$

$$= TN_0 + \delta \sum_{s=1}^{T} \sum_{i=0}^{s-1} r_i = TN_0 + \delta \sum_{i=0}^{T-1} (T - i) r_i$$

By Proposition 17.4 (a) & (c) in Hamilton (1994), we have

$$T^{-1/2} \sum_{i=0}^{T-1} r_i \xrightarrow{L} \sigma_r W(1)$$

And

$$T^{-3/2} \sum_{i=0}^{T-1} ir_i \xrightarrow{L} \sigma_r W(1) - \sigma_r \int_0^1 W(s)ds$$

Hence

$$T^{-3/2} \sum_{s=1}^{T} N_s = T^{-1/2} N_0 + \delta T^{-1/2} \sum_{i=0}^{T-1} r_i - \delta T^{-3/2} \sum_{i=0}^{T-1} ir_i$$

$$\xrightarrow{L} \delta \sigma_r \int_0^1 W(s)ds$$

Furthermore,

$$\sum_{i=1}^{T} N_t r_{t+1} = N_0 \sum_{t=1}^{T} r_{t+1} + \delta \sum_{t=1}^{T} R_t r_{t+1}$$
Again, from Proposition 17.4 (b) in Hamilton (1994), we have

\[
T^{-1} \sum_{t=1}^{T} R_t r_{t+1} = T^{-1} \sum_{t=1}^{T} (R_{t+1} - r_t) r_{t+1} = T^{-1} \sum_{t=1}^{T} R_{t+1} r_{t+1} - T^{-1} \sum_{t=1}^{T} r_t r_{t+1} \rightarrow \frac{1}{2} \sigma_r^2 \{ |W(1)|^2 - 1 \}
\]

Therefore,

\[
T^{-1} \sum_{t=1}^{T} N_t r_{t+1} \rightarrow \frac{1}{2} \delta \sigma_r^2 \{ |W(1)|^2 - 1 \}
\]

Finally, we have as \( T \to \infty \),

\[
\sqrt{T} \frac{CAR_T}{\sqrt{\frac{1}{2} \delta \sigma_r^2 \{ |W(1)|^2 - 1 \}}} = \frac{\sigma_r \{ |W(1)|^2 - 1 \}}{2 \int_0^1 W(s)ds}.
\]

Then \( CAR_T = \frac{1}{\sqrt{T}} \sqrt{T} CAR_T = o_p(1)O_p(1) = o_p(1) \), since \( \sqrt{T} CAR_T = O_p(1) \).

**Proof of Theorem 12.** Let \( \lambda_t = N_t / \sum_{s=1}^{T} N_s \)

\[
E[\overline{CAR}_T] = -\delta \sigma_r^2 \frac{T-2}{2} \sum_{t=1}^{T-1} E[\lambda_t \lambda_s]
\]

\[
\leq -\delta \sigma_r^2 \left[ \frac{1 - \sum_{t=1}^{T} \lambda_t^2}{2} \right]
\]

We only need to prove that:

\[
E \left[ \sum_{t=1}^{T} \lambda_t^2 \right] \text{ converges to zero.}
\]

In fact, first notice that given \( N_0 = 0 \) and \( N_t = \exp(\delta (r_0 + r_1 + \cdots + r_{t-1})) \), we have

\[
E[\lambda_t^2] = E \left[ \frac{1}{\left( \sum_{s=1}^{t-1} \exp \left( -\delta \left( \sum_{s=1}^{t-1} r_u \right) \right) \right) + \left( 1 + \sum_{s=t}^{T} \exp \left( -\delta \left( \sum_{s}^{T} r_u \right) \right) \right) } \right]^2 \leq E \left[ \frac{1}{\left( 1 + \sum_{s=1}^{T/2} \exp \left( -\delta \left( \sum_{s}^{T/2} r_u \right) \right) \right)^2} \right] (\tau_t \text{ iid } N(0, 1))
\]
Therefore,
\[
E \left[ \sum_{t=1}^{T} \lambda_t^2 \right] \leq TE \left[ \frac{1}{1 + \sum_{s=1}^{\lceil T/2 \rceil} \exp(-\delta (\sum_s^r r_u))} \right] \leq TE \left[ \frac{1}{1 + \sum_{s=1}^{\lceil T/2 \rceil} \exp(-\delta (\sum_s^r r_u))} \right] \cdot \frac{1}{1 + \sum_{s=1}^{\lceil T/2 \rceil} \exp(-\delta (\sum_s^r r_u))} < 1
\]

We only need to prove that for \( r_i \overset{iid}{\sim} N(0, 1) \),
\[
TE \left[ \frac{1}{1 + \sum_{s=1}^{\lceil T/2 \rceil} \exp(-\delta (\sum_s^r r_u))} \right] \to 0, \text{ as } T \to \infty
\]

Let \( S_T \triangleq E \left[ \frac{1}{1 + \sum_{s=1}^{\lceil T/2 \rceil} \exp(-\delta (\sum_s^r r_u))} \right] \). We can prove that for large enough \( T \), say \( T \geq T_0 \), \( \exists \epsilon_0 > 0 \), such that,
\[
S_T \leq \frac{1}{T^{1+\epsilon_0}}
\]

Notice that, \( f(x) = (1 + \delta/x)^{-1} \) is a concave function, by Jensen’s inequality, we have for any \( \delta > 0, x > 0 \),
\[
S_T \leq E \left[ \frac{1}{1 + \exp(r_1)(T-1)_{1+\epsilon_0}} \right] \\
\leq E \left[ \frac{1}{1 + \exp(r_1)(T-1)_{1+\epsilon_0}} \right] \\
= \sum_{n=-\infty}^{\infty} \int_{\ln(a^n)}^{\ln(a^{n+1})} \frac{1}{1 + \exp(r_1)(T-1)_{1+\epsilon_0}} \Phi(r_1)dr_1 \quad (a > 1 \text{ is to be determined}) \\
\leq \sum_{n=-\infty}^{\infty} \frac{1}{(T-1)_{1+\epsilon_0}} \left( \frac{1}{a} \right)^n \Pr (\ln(a^n) \leq r_1 \leq \ln(a^{n+1})) \\
\leq \sum_{n=0}^{\infty} \frac{1}{(T-1)_{1+\epsilon_0}} \left( \frac{1}{a} \right)^n \Pr (\ln(a^n) \leq r_1 \leq \ln(a^{n+1})) \\
+ \sum_{n=1}^{\infty} \frac{1}{(T-1)_{1+\epsilon_0}} a^n \Pr (\ln(a^{-n}) \leq r_1 \leq \ln(a^{-n+1}))
\]

Furthermore, by Mean Value Theorem, we have, \( \exists \ln(a^n) \leq x \leq \ln(a^{n+1}) \)
\[
\Pr (\ln(a^n) \leq r_1 \leq \ln(a^{n+1})) \\
= \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x^2}{2} \right) (\ln(a^{n+1}) - \ln(a^n)) \\
\leq \frac{\ln a}{\sqrt{2\pi}} \exp \left( -\frac{(\ln(a^n))^2}{2} \right)
\]

31
Proof of Theorem 13.

Let \( X = (r_1, \cdots, r_{T-1}) \), where \( n_t = \log(N_t) \) and \( h(X) = \frac{N_t}{\sum_{t=1}^{T} N_s} \), \( \forall t \). By Stein’s formula, we have

\[
E \left[ \frac{N_T}{\sum_{s=1}^{T} N_s} \right] = \frac{\sum_{s=1}^{T} e_{n_t}^s}{\sum_{s=1}^{T} e_{n_s}^s} = \sum_{s=1}^{T-1} \text{cov}(r_{t+1}, r_s) E \left[ \frac{\partial}{\partial r_s} h(X) \right]
\]

\[
= -\text{cov}(r_{t+1}, r_{t+1}) E \left[ \frac{e_{n_t}^{s+1} e_{n_s}^s}{\left( \sum_{s=1}^{T} e_{n_s}^s \right)^2} \right] (\text{Since } \text{cov}(r_t, r_s) = 0, \text{if } t \neq s.)
\]

\[
= -\sum_{s=t+2}^{T} \theta^{s-t-2} \delta \sigma_t^2 E \left[ \frac{e_{n_t}^{s} e_{n_s}^s}{\left( \sum_{s=1}^{T} e_{n_s}^s \right)^2} \right] < 0 (\text{Since } E \left[ \frac{e_{n_t}^{s} e_{n_s}^s}{\left( \sum_{s=1}^{T} e_{n_s}^s \right)^2} \right] > 0)
\]

Therefore,

\[
E[CAR_T] = -\delta \sigma_t^2 \sum_{t=1}^{T-2} \sum_{s=t+2}^{T} \theta^{s-t-2} E \left[ \frac{e_{n_t}^{s} e_{n_s}^s}{\left( \sum_{s=1}^{T} e_{n_s}^s \right)^2} \right] < 0
\]
Figure 1: Large Sample Simulation

Figure 1 presents the results based on averages of 100 rounds of simulation, for $\theta = 0.2, 0.4, 0.6, 0.8$ and 1.0, $\delta = 0.2, 0.4, 0.6$ and 0.8. For each simulation, we draw 100,000 observations of the IPO return and the number of IPOs, i.e. $T=100,000$. 
Figure 2 shows Schultz’s example. It is clear from this that the number of events drifts downwards over time and the expected CAR is negative.
Figure 3 shows the behavior of the expected cumulative abnormal return as we approach the unit root in the number of events process ($\theta = 1$) for various values of $\delta$ (the effect of current returns on tomorrow's number of events).
Figure 4

Figure 4 compares the actual IPO numbers (solid line) with that predicted by Schultz (2003) (dotted line) and with that predicted by our approach (dashed line).
Figure 5 gives the probability of staying in regime 1 for our Markov regime switching model.
Figure 6 plots the predicted logarithm of IPO numbers based on our regime switching model.
Figure 7 depicts the relationship between asymptotic variance of CAR and θ, δ, where θ is chosen from 0.1 to 0.9 and δ is from 0.5 to 2.5.
The first panel of Figure 8 shows the ratio of the two variances based on the assumptions of deterministic event process and the stochastic one respectively; and the second panel shows the asymptotic variance of CAR_T for T = 1, 2, ..., 200 and θ = 0.95, δ = 1.75 and \( \text{var}[r_t] = 0.679 \).
This table reports the estimates of expectation of monthly $\overline{CAR}_T$ for $T = 200, 400, 600$ respectively from 500 rounds of simulations using Stein’s formula (c.f. Theorem 13). All simulations are based on the model used: $\log(N_{t+1}) = \theta \log(N_t) + \delta r_t$, where $\theta = 0.6, 0.8, 0.85, 0.9, 0.95$ or $1.0$ and $\delta = 0.5, 1.0, 1.5$ or $1.75$, $r_t \sim N(0, 0.0824)$. The standard deviation of $r_t$ is chosen to be $0.0824$ to be consistent with our sample.

<table>
<thead>
<tr>
<th></th>
<th>Average Monthly CAR of holding period $T = 200$</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$ = 0.6</td>
<td>$\delta = 0.5$ -0.000041 -0.000083 -0.000125 -0.000146</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.8</td>
<td>$\delta = 1.0$ -0.000082 -0.000165 -0.000250 -0.000293</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.85</td>
<td>$\delta = 1.5$ -0.000109 -0.000219 -0.000332 -0.000389</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.9</td>
<td>$\delta = 1.75$ -0.000160 -0.000324 -0.000494 -0.000582</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 1.0</td>
<td>$\delta = 0.5$ -0.000062 -0.000125 -0.000190 -0.000222</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.95</td>
<td>$\delta = 1.0$ -0.000092 -0.000184 -0.000278 -0.000329</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 1.0</td>
<td>$\delta = 1.5$ -0.000110 -0.000223 -0.000336 -0.000389</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.95</td>
<td>$\delta = 1.75$ -0.000160 -0.000324 -0.000494 -0.000582</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 1.0</td>
<td>$\delta = 0.5$ -0.000011 -0.000022 -0.000033 -0.000049</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.95</td>
<td>$\delta = 1.0$ -0.000011 -0.000022 -0.000033 -0.000049</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 1.0</td>
<td>$\delta = 1.5$ -0.000011 -0.000022 -0.000033 -0.000049</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.95</td>
<td>$\delta = 1.75$ -0.000011 -0.000022 -0.000033 -0.000049</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Average Monthly CAR of holding period $T = 400$</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$ = 0.6</td>
<td>$\delta = 0.5$ -0.000031 -0.000063 -0.000094 -0.000110</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.8</td>
<td>$\delta = 1.0$ -0.000062 -0.000125 -0.000189 -0.000222</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.85</td>
<td>$\delta = 1.5$ -0.000082 -0.000166 -0.000253 -0.000297</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.9</td>
<td>$\delta = 1.75$ -0.000122 -0.000248 -0.000379 -0.000445</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.95</td>
<td>$\delta = 0.5$ -0.000031 -0.000063 -0.000094 -0.000110</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 1.0</td>
<td>$\delta = 1.0$ -0.000062 -0.000125 -0.000189 -0.000222</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.95</td>
<td>$\delta = 1.5$ -0.000082 -0.000166 -0.000253 -0.000297</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 1.0</td>
<td>$\delta = 1.75$ -0.000122 -0.000248 -0.000379 -0.000445</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Average Monthly CAR of holding period $T = 600$</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$ = 0.6</td>
<td>$\delta = 0.5$ -0.000028 -0.000056 -0.000084 -0.000098</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.8</td>
<td>$\delta = 1.0$ -0.000055 -0.000112 -0.000169 -0.000199</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.85</td>
<td>$\delta = 1.5$ -0.000074 -0.000149 -0.000226 -0.000265</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.9</td>
<td>$\delta = 1.75$ -0.000110 -0.000223 -0.000341 -0.000401</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.95</td>
<td>$\delta = 0.5$ -0.000028 -0.000056 -0.000084 -0.000098</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 1.0</td>
<td>$\delta = 1.0$ -0.000055 -0.000112 -0.000169 -0.000199</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 0.95</td>
<td>$\delta = 1.5$ -0.000074 -0.000149 -0.000226 -0.000265</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta$ = 1.0</td>
<td>$\delta = 1.75$ -0.000110 -0.000223 -0.000341 -0.000401</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2: Average 3-year Cumulative Abnormal Return

This table reports the estimates of expectation of 3-year $\text{CAR}_T$ for $T = 200, 400, 600$ respectively from 500 rounds of simulations using Stein’s formula (c.f. Theorem 13). All simulations are based on the model used: $\log(N_{t+1}) = \theta \log(N_t) + \delta r_t$, where $\theta = 0.6, 0.8, 0.85, 0.9, 0.95$ or 1.0 and $\delta = 0.5, 1.0, 1.5$ or 1.75, $r_t \overset{iid}{\sim} N(0, 0.0824)$. The standard deviation of $r_t$ is chosen to be 0.0824 to be consistent with our sample.

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$\theta$</th>
<th>Average 3-year CAR of holding period $T = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.6</td>
<td>-0.001367 -0.002734 -0.004102 -0.004786</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>-0.005238 -0.007127 -0.010712 -0.014245</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>-0.009878 -0.019846 -0.029825 -0.034758</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>-0.049794 -0.097194 -0.140433 -0.160823</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$\theta$</th>
<th>Average 3-year CAR of holding period $T = 400$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.6</td>
<td>-0.001065 -0.002131 -0.003198 -0.003734</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>-0.002111 -0.004223 -0.006346 -0.007417</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>-0.002798 -0.005604 -0.008426 -0.009827</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>-0.004144 -0.008321 -0.012526 -0.014570</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$\theta$</th>
<th>Average 3-year CAR of holding period $T = 600$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.6</td>
<td>-0.000965 -0.001930 -0.002896 -0.003380</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>-0.001916 -0.003832 -0.005764 -0.006736</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>-0.002545 -0.005091 -0.007653 -0.008932</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>-0.003785 -0.007585 -0.011466 -0.013329</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>$\theta$</th>
<th>Average 3-year CAR of holding period $T = 600$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.6</td>
<td>-0.008705 -0.176310 -0.260842 -0.305045</td>
</tr>
<tr>
<td></td>
<td>0.8</td>
<td>-0.088705 -0.176310 -0.260842 -0.305045</td>
</tr>
<tr>
<td></td>
<td>0.9</td>
<td>-0.003785 -0.007585 -0.011466 -0.013329</td>
</tr>
<tr>
<td></td>
<td>1.0</td>
<td>-0.008705 -0.176310 -0.260842 -0.305045</td>
</tr>
</tbody>
</table>
Table 3: Average 3-year Buy-and-Hold Abnormal Return

This table reports the estimates of expectation of 3-year BHAR for $T = 200, 400, 600$ respectively from 300,000 rounds of simulations using the definition of buy-hold abnormal returns. All simulations are based on the model used: \( \log(N_{t+1}) = \theta \log(N_t) + \delta r_t \), where \( \theta = 0.6, 0.8, 0.85, 0.9, 0.95 \) or 1.0 and \( \delta = 0.5, 1.0, 1.5 \) or 1.75, \( r_t \overset{iid}{\sim} N(0, 0.0824) \). The standard deviation of \( r_t \) is chosen to be 0.0824 to be consistent with our sample.

<table>
<thead>
<tr>
<th></th>
<th>( \delta = 0.5 )</th>
<th>( \delta = 1.0 )</th>
<th>( \delta = 1.5 )</th>
<th>( \delta = 1.75 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta = 0.6 )</td>
<td>-0.001372</td>
<td>-0.002738</td>
<td>-0.004110</td>
<td>-0.004797</td>
</tr>
<tr>
<td>( \theta = 0.8 )</td>
<td>-0.002702</td>
<td>-0.005418</td>
<td>-0.008148</td>
<td>-0.009524</td>
</tr>
<tr>
<td>( \theta = 0.85 )</td>
<td>-0.003569</td>
<td>-0.007166</td>
<td>-0.010796</td>
<td>-0.012628</td>
</tr>
<tr>
<td>( \theta = 0.9 )</td>
<td>-0.005262</td>
<td>-0.010583</td>
<td>-0.015988</td>
<td>-0.018715</td>
</tr>
<tr>
<td>( \theta = 0.95 )</td>
<td>-0.009930</td>
<td>-0.020368</td>
<td>-0.030298</td>
<td>-0.035519</td>
</tr>
<tr>
<td>( \theta = 1.0 )</td>
<td>-0.048894</td>
<td>-0.093937</td>
<td>-0.133901</td>
<td>-0.151881</td>
</tr>
<tr>
<td>( \theta = 1.0 )</td>
<td>-0.048894</td>
<td>-0.093937</td>
<td>-0.133901</td>
<td>-0.151881</td>
</tr>
</tbody>
</table>

Average 3-year BHAR of holding period $T = 400$

<table>
<thead>
<tr>
<th></th>
<th>( \delta = 0.5 )</th>
<th>( \delta = 1.0 )</th>
<th>( \delta = 1.5 )</th>
<th>( \delta = 1.75 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta = 0.6 )</td>
<td>-0.000745</td>
<td>-0.001473</td>
<td>-0.002204</td>
<td>-0.002571</td>
</tr>
<tr>
<td>( \theta = 0.8 )</td>
<td>-0.001465</td>
<td>-0.002926</td>
<td>-0.004401</td>
<td>-0.005148</td>
</tr>
<tr>
<td>( \theta = 0.85 )</td>
<td>-0.001941</td>
<td>-0.003890</td>
<td>-0.005869</td>
<td>-0.006874</td>
</tr>
<tr>
<td>( \theta = 0.9 )</td>
<td>-0.002886</td>
<td>-0.005809</td>
<td>-0.008810</td>
<td>\textbf{-0.010340}</td>
</tr>
<tr>
<td>( \theta = 0.95 )</td>
<td>-0.005626</td>
<td>-0.011443</td>
<td>-0.017520</td>
<td>-0.020689</td>
</tr>
<tr>
<td>( \theta = 1.0 )</td>
<td>-0.053930</td>
<td>-0.103245</td>
<td>-0.146728</td>
<td>\textbf{-0.166237}</td>
</tr>
</tbody>
</table>

Average 3-year BHAR of holding period $T = 600$

<table>
<thead>
<tr>
<th></th>
<th>( \delta = 0.5 )</th>
<th>( \delta = 1.0 )</th>
<th>( \delta = 1.5 )</th>
<th>( \delta = 1.75 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta = 0.6 )</td>
<td>-0.000752</td>
<td>-0.001248</td>
<td>-0.001747</td>
<td>-0.001998</td>
</tr>
<tr>
<td>( \theta = 0.8 )</td>
<td>-0.001244</td>
<td>-0.002242</td>
<td>-0.003252</td>
<td>-0.003764</td>
</tr>
<tr>
<td>( \theta = 0.85 )</td>
<td>-0.001570</td>
<td>-0.002904</td>
<td>-0.004262</td>
<td>-0.004953</td>
</tr>
<tr>
<td>( \theta = 0.9 )</td>
<td>-0.002221</td>
<td>-0.004229</td>
<td>-0.006299</td>
<td>-0.007359</td>
</tr>
<tr>
<td>( \theta = 0.95 )</td>
<td>-0.004136</td>
<td>-0.008189</td>
<td>-0.012469</td>
<td>-0.014725</td>
</tr>
<tr>
<td>( \theta = 1.0 )</td>
<td>-0.055963</td>
<td>-0.106873</td>
<td>-0.151803</td>
<td>-0.171990</td>
</tr>
</tbody>
</table>
Table 4: Average Monthly Cumulative Abnormal Return

This table reports the estimates of monthly \( \text{CAR}_T \), for \( T = 200, 400, 600 \) by simulation after including one more lag into our model: \( \log(N_t) = \theta_1 \log(N_{t-1}) + \theta_2 \log(N_{t-2}) + \delta r_{IPO,t-1} \), where \( \theta_1 = 0.6 \) or 0.7, \( \theta_2 = 0.25 \) or 0.3 and \( \delta = 2.0, 2.3 \) or 2.5, \( r_t \sim N(0, 0.0824) \). The standard deviation of \( r_t \) is chosen to be 0.0824 to be consistent with our sample. We conduct 300,000 rounds of direct simulations using \( \text{CAR}_T \)'s definition.

<table>
<thead>
<tr>
<th>Average Monthly CAR of holding period ( T = 200 )</th>
<th>( \delta = 2.0 )</th>
<th>( \delta = 2.3 )</th>
<th>( \delta = 2.5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( (\theta_1, \theta_2) = (0.6, 0.25) )</td>
<td>-0.000437</td>
<td>-0.000509</td>
<td>-0.000557</td>
</tr>
<tr>
<td>( (\theta_1, \theta_2) = (0.6, 0.3) )</td>
<td>-0.000649</td>
<td>-0.000756</td>
<td>-0.000828</td>
</tr>
<tr>
<td>( (\theta_1, \theta_2) = (0.7, 0.25) )</td>
<td>-0.001241</td>
<td>-0.001452</td>
<td>-0.001597</td>
</tr>
<tr>
<td>( (\theta_1, \theta_2) = (0.7, 0.3) )</td>
<td>-0.005111</td>
<td>-0.005866</td>
<td>-0.006368</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Average Monthly CAR of holding period ( T = 400 )</th>
<th>( \delta = 2.0 )</th>
<th>( \delta = 2.3 )</th>
<th>( \delta = 2.5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( (\theta_1, \theta_2) = (0.6, 0.25) )</td>
<td>-0.000234</td>
<td>-0.000271</td>
<td>-0.000296</td>
</tr>
<tr>
<td>( (\theta_1, \theta_2) = (0.6, 0.3) )</td>
<td>-0.000349</td>
<td>-0.000406</td>
<td>-0.000445</td>
</tr>
<tr>
<td>( (\theta_1, \theta_2) = (0.7, 0.25) )</td>
<td>-0.000695</td>
<td>-0.000817</td>
<td>-0.000903</td>
</tr>
<tr>
<td>( (\theta_1, \theta_2) = (0.7, 0.3) )</td>
<td>-0.005161</td>
<td>-0.005866</td>
<td>-0.006368</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Average Monthly CAR of holding period ( T = 600 )</th>
<th>( \delta = 2.0 )</th>
<th>( \delta = 2.3 )</th>
<th>( \delta = 2.5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( (\theta_1, \theta_2) = (0.6, 0.25) )</td>
<td>-0.000159</td>
<td>-0.000184</td>
<td>-0.000201</td>
</tr>
<tr>
<td>( (\theta_1, \theta_2) = (0.6, 0.3) )</td>
<td>-0.000238</td>
<td>-0.000277</td>
<td>-0.000303</td>
</tr>
<tr>
<td>( (\theta_1, \theta_2) = (0.7, 0.25) )</td>
<td>-0.000480</td>
<td>-0.000566</td>
<td>-0.000627</td>
</tr>
<tr>
<td>( (\theta_1, \theta_2) = (0.7, 0.3) )</td>
<td>-0.005181</td>
<td>-0.005947</td>
<td>-0.006456</td>
</tr>
</tbody>
</table>
Table 5: Average 3-year Cumulative Abnormal Return

This table reports the estimates of 3-year CAR, for $T = 200, 400, 600$ by simulation after including one more lag into our model: $\log(N_t) = \theta_1 \log(N_{t-1}) + \theta_2 \log(N_{t-2}) + \delta r_{IPO,t-1}$, where $\theta_1 = 0.6$ or $0.7$, $\theta_2 = 0.25$ or $0.3$ and $\delta = 2.0, 2.3$ or $2.5$, $r_t \sim N(0, 0.0824)$. The standard deviation of $r_t$ is chosen to be 0.0824 to be consistent with our sample. We conduct 300,000 rounds of direct simulations using CAR$_T$’s definition.

<table>
<thead>
<tr>
<th>Average 3-year CAR of holding period $T = 200$</th>
<th>$\delta = 2.0$</th>
<th>$\delta = 2.3$</th>
<th>$\delta = 2.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\theta_1, \theta_2) = (0.6, 0.25)$</td>
<td>-0.013971</td>
<td>-0.016116</td>
<td>-0.017535</td>
</tr>
<tr>
<td>$(\theta_1, \theta_2) = (0.6, 0.3)$</td>
<td>-0.020543</td>
<td>-0.023691</td>
<td>-0.025780</td>
</tr>
<tr>
<td>$(\theta_1, \theta_2) = (0.7, 0.25)$</td>
<td>-0.038431</td>
<td>-0.044323</td>
<td>-0.048227</td>
</tr>
<tr>
<td>$(\theta_1, \theta_2) = (0.7, 0.3)$</td>
<td>-0.144167</td>
<td>-0.162630</td>
<td>-0.174390</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Average 3-year CAR of holding period $T = 400$</th>
<th>$\delta = 2.0$</th>
<th>$\delta = 2.3$</th>
<th>$\delta = 2.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\theta_1, \theta_2) = (0.6, 0.25)$</td>
<td>-0.007780</td>
<td>-0.008951</td>
<td>-0.009729</td>
</tr>
<tr>
<td>$(\theta_1, \theta_2) = (0.6, 0.3)$</td>
<td>-0.011542</td>
<td>-0.013304</td>
<td>-0.014482</td>
</tr>
<tr>
<td>$(\theta_1, \theta_2) = (0.7, 0.25)$</td>
<td>-0.022619</td>
<td>-0.026230</td>
<td>-0.028671</td>
</tr>
<tr>
<td>$(\theta_1, \theta_2) = (0.7, 0.3)$</td>
<td>-0.158705</td>
<td>-0.178936</td>
<td>-0.191828</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Average 3-year CAR of holding period $T = 600$</th>
<th>$\delta = 2.0$</th>
<th>$\delta = 2.3$</th>
<th>$\delta = 2.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\theta_1, \theta_2) = (0.6, 0.25)$</td>
<td>-0.005417</td>
<td>-0.006216</td>
<td>-0.006749</td>
</tr>
<tr>
<td>$(\theta_1, \theta_2) = (0.6, 0.3)$</td>
<td>-0.008031</td>
<td>-0.009246</td>
<td>-0.010061</td>
</tr>
<tr>
<td>$(\theta_1, \theta_2) = (0.7, 0.25)$</td>
<td>-0.015939</td>
<td>-0.018520</td>
<td>-0.020278</td>
</tr>
<tr>
<td>$(\theta_1, \theta_2) = (0.7, 0.3)$</td>
<td>-0.163960</td>
<td>-0.184913</td>
<td>-0.198279</td>
</tr>
</tbody>
</table>
Table 6: Average 3-year Buy-Hold Abnormal Return

This table reports the estimates of 3-year \( BHAR_T \), for \( T = 200, 400, 600 \) by simulation after including one more lag into our model: \( \log(N_t) = \theta_1 \log(N_{t-1}) + \theta_2 \log(N_{t-2}) + \delta r_{IPO,t-1} \), where \( \theta_1 = 0.6 \) or 0.7, \( \theta_2 = 0.25 \) or 0.3 and \( \delta = 2.0, 2.3 \) or 2.5, \( r_t \sim iid N(0,0.0824) \). The standard deviation of \( r_t \) is chosen to be 0.0824 to be consistent with our sample. We conduct 300,000 rounds of direct simulations using 3-year \( BHAR_T \)'s definition.

<table>
<thead>
<tr>
<th>( (\theta_1, \theta_2) )</th>
<th>( \delta = 2.0 )</th>
<th>( \delta = 2.3 )</th>
<th>( \delta = 2.5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.6, 0.25)</td>
<td>-0.014205</td>
<td>-0.016425</td>
<td>-0.017898</td>
</tr>
<tr>
<td>(0.6, 0.3)</td>
<td>-0.020944</td>
<td>-0.024218</td>
<td>-0.026404</td>
</tr>
<tr>
<td>(0.7, 0.25)</td>
<td>-0.039127</td>
<td>-0.045208</td>
<td>-0.049246</td>
</tr>
<tr>
<td>(0.7, 0.3)</td>
<td>-0.137248</td>
<td>-0.153809</td>
<td>-0.164234</td>
</tr>
<tr>
<td>(0.6, 0.25)</td>
<td>-0.008007</td>
<td>-0.009236</td>
<td>-0.010058</td>
</tr>
<tr>
<td>(0.6, 0.3)</td>
<td>-0.011930</td>
<td>-0.013803</td>
<td>-0.015066</td>
</tr>
<tr>
<td>(0.7, 0.25)</td>
<td>-0.023501</td>
<td>-0.027372</td>
<td>-0.030006</td>
</tr>
<tr>
<td>(0.7, 0.3)</td>
<td>-0.150404</td>
<td>-0.168362</td>
<td>-0.179660</td>
</tr>
<tr>
<td>(0.6, 0.25)</td>
<td>-0.005555</td>
<td>-0.006399</td>
<td>-0.006965</td>
</tr>
<tr>
<td>(0.6, 0.3)</td>
<td>-0.008300</td>
<td>-0.009603</td>
<td>-0.010486</td>
</tr>
<tr>
<td>(0.7, 0.25)</td>
<td>-0.016655</td>
<td>-0.019468</td>
<td>-0.021400</td>
</tr>
<tr>
<td>(0.7, 0.3)</td>
<td>-0.155188</td>
<td>-0.173722</td>
<td>-0.185389</td>
</tr>
</tbody>
</table>
Table 7: The Distribution of the Number of Offerings per Month

Descriptive statistics for IPO(SEO) numbers from February 1973 to December 2002. There are 9,190 initial public offerings in this period.

<table>
<thead>
<tr>
<th>Monthly Number of</th>
<th>Monthly Number of</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Public</td>
<td>Seasoned Equity</td>
</tr>
<tr>
<td>Offerings</td>
<td>Offerings</td>
</tr>
<tr>
<td>Mean</td>
<td>25.60</td>
</tr>
<tr>
<td>Median</td>
<td>20</td>
</tr>
<tr>
<td>Minimum</td>
<td>0</td>
</tr>
<tr>
<td>Maximum</td>
<td>106</td>
</tr>
<tr>
<td>First order autocorrelation</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Table 8: Unit Root Test with one lag

This table reports results of both Augmented Dickey-Fuller and Phillips-Perron tests on the number of IPO offerings and the corresponding logarithm against the following two nulls ($H_0$, $H_{00}$) with or without time trend. Here we only consider one lag in testing procedure.

<table>
<thead>
<tr>
<th>Panel A: $H_0$: constant term without time trend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>ADF Test</td>
</tr>
<tr>
<td>Test</td>
</tr>
<tr>
<td>Statistic</td>
</tr>
<tr>
<td>log($N_t + 0.5$)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: $H_{00}$: constant term with time trend</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>ADF Test</td>
</tr>
<tr>
<td>Test</td>
</tr>
<tr>
<td>Statistic</td>
</tr>
<tr>
<td>log($N_t + 0.5$)</td>
</tr>
</tbody>
</table>

*MacKinnon critical values for rejection of hypothesis of a unit root.
Table 9A: Unit Root Testing of $H_0$ for more lags

This table reports results of both Augmented Dickey-Fuller and Elliott-Rothenberg-Stock tests on the number of IPO offerings up to 16 lags against the null $H_0$: unit root process without time trend.

<table>
<thead>
<tr>
<th>Lag</th>
<th>ADF Test</th>
<th>Elliott-Rothenberg-Stock Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Test Statistics</td>
<td>1%</td>
</tr>
<tr>
<td>4</td>
<td>-2.159</td>
<td>-3.451</td>
</tr>
<tr>
<td>3</td>
<td>-2.081</td>
<td>-3.451</td>
</tr>
</tbody>
</table>
Table 9B: Unit Root Testing of $H_0^0$ for more lags

This table reports results of both Augmented Dickey-Fuller and Elliott-Rothenberg-Stock tests on the number of IPO offerings up to 16 lags against the null $H_0^0$: unit root process with time trend.

<table>
<thead>
<tr>
<th>Lag</th>
<th>Test Statistics</th>
<th>Critical Value</th>
<th>Test Statistics</th>
<th>Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1%</td>
<td>5%</td>
<td>10%</td>
<td>1%</td>
</tr>
</tbody>
</table>
### Table 10: Model Parameter Estimates

This table reports the estimations of the restricted Vector Autoregressive model (c.f. Equation 4). The restrictions imposed are: $b_{12} = b_{13} = b_{23} = 0$.

| Equation | Parameter | Estimate | Std Error | T Ratio | Prob>|T| |
|----------|-----------|----------|-----------|---------|----------|
| MKTre(t) | $u_1$ 0.00821 | 0.00257 | 3.19 | 0.0015 |
|          | $b_{11}$ 0.04563 | 0.05314 | 0.86 | 0.3910 |
|          | $b_{12}$ 0 | 0 | NA | NA |
|          | $b_{13}$ 0 | 0 | NA | NA |
| IPOre(t) | $u_2$ 0.00464 | 0.00429 | 1.08 | 0.2808 |
|          | $b_{21}$ 0.23832 | 0.11289 | 2.11 | 0.0355 |
|          | $b_{22}$ 0.08584 | 0.04992 | 1.72 | 0.0864 |
|          | $b_{23}$ 0 | 0 | NA | NA |
| logIPOno(t) | $u_3$ 0.28502 | 0.06931 | 4.11 | 0.0001 |
|          | $b_{31}$ 0.21947 | 1.17524 | 0.19 | 0.8520 |
|          | $b_{32}$ 1.73638 | 0.67825 | 2.56 | 0.0109 |
|          | $b_{33}$ 0.88724 | 0.02343 | 37.87 | 0.0001 |

### Table 11: Model Parameter Estimates

This table reports the estimations of the unrestricted Vector Autoregressive model (c.f. Equation 4).

| Equation | Parameter | Estimate | Std Error | T Ratio | Prob>|T| |
|----------|-----------|----------|-----------|---------|----------|
| MKTre(t) | $u_1$ 0.01027 | 0.00553 | 1.86 | 0.0643 |
|          | $b_{11}$ 0.02605 | 0.09350 | 0.28 | 0.7807 |
|          | $b_{12}$ 0.01522 | 0.05436 | 0.28 | 0.7796 |
|          | $b_{13}$ -0.00075397 | 0.00188 | -0.40 | 0.6886 |
| IPOre(t) | $u_2$ 0.02406 | 0.00922 | 2.61 | 0.0095 |
|          | $b_{21}$ 0.26114 | 0.15585 | 1.68 | 0.0947 |
|          | $b_{22}$ 0.08191 | 0.09060 | 0.90 | 0.3666 |
|          | $b_{23}$ -0.00741 | 0.00313 | -2.37 | 0.0186 |
| logIPOno(t) | $u_3$ 0.27943 | 0.07060 | 3.96 | 0.0001 |
|          | $b_{31}$ 0.13380 | 1.19306 | 0.11 | 0.9108 |
|          | $b_{32}$ 1.79373 | 0.69360 | 2.59 | 0.0101 |
|          | $b_{33}$ 0.88949 | 0.02399 | 37.08 | 0.0001 |
Table 12A: Maximum Likelihood Estimation of Regime Switching Model

This table records the maximum likelihood estimation of the regime switching model: in regime \( i \) (\( i = 1 \) or \( 2 \)), \( \log(N_{t+1} + 0.5) = a_i + b_i \log(N_t + 0.5) + cR_{IPO,t} + \epsilon_i \). \( p_{ii} \) denotes the transition probability that next regime is regime \( i \) given the current one is regime \( i \), \( i = 1, 2 \).

\[
\begin{align*}
a_1 &= 0.38209733023276; \\
b_1 &= 0.58714689381549; \\
c &= 2.07722887525329; \\
a_2 &= 0.61570203204385; \\
b_2 &= 0.81094872780512; \\
p_{11} &= 0.98754466866105; \\
p_{22} &= 0.99461003917911; \\
\sigma_1 &= \text{std}(\epsilon_1) = 0.83465047639259; \\
\sigma_2 &= \text{std}(\epsilon_2) = 0.40485669899499;
\end{align*}
\]

Table 12B: Likelihood Ratio Test of Regime Switching Model

We conduct likelihood ratio test of regime switching model to see whether adding regime switching would fit the IPO event process better. In particular, we are testing against the null \( H_0: a_1 = a_2, b_1 = b_2, \sigma_1 = \sigma_2 \) for regime switching model 5.

<table>
<thead>
<tr>
<th>Model</th>
<th>LogLikelihood</th>
<th>2 log(( \frac{L_0}{L} ))</th>
<th>D.F.</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Markov Regime Switching Model</td>
<td>-271.0561</td>
<td>116.4264</td>
<td>3</td>
<td>0.0000</td>
</tr>
<tr>
<td>AR(1) model without Regime Switching</td>
<td>-329.2693</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>