JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 42, No. 3, Sept. 2007, pp. 565–594
COPYRIGHT 2007, SCHOOL OF BUSINESS ADMINISTRATION, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

Is IPO Underperformance a Peso Problem?

Andrew Ang, Li Gu, and Yael V. Hochberg*

Abstract

Recent studies suggest that the underperformance of IPOs in the post-1970 sample may be a small sample effect or "Peso problem." That is, IPO underperformance may result from observing too few star performers ex post than were expected ex ante. We develop a model of IPO performance that captures this intuition by allowing returns to be drawn from mixtures of outstanding, benchmark, or poor performing states. We estimate the model under the null of no ex ante average IPO underperformance and construct small sample distributions of various statistics measuring IPO relative performance. We find that small sample biases are extremely unlikely to account for the magnitude of the post-1970 IPO underperformance observed in data.

I. Introduction

Since Ritter's (1991) seminal study, many papers document that firms underperform relative to benchmark indices or to similar stocks following their initial public offerings (IPOs).[1] The study of IPO long-run average underperformance is important as IPO long-run underperformance may indicate a possible informational inefficiency in capital allocation, the influence of behavioral fads in markets, or the existence of trading opportunities that produce superior abnormal returns.

However, the acceptance of the existence of an IPO underperformance effect is far from universal. In a recent paper, Schultz (2003) argues that more IPO activity follows successful IPOs and that measuring the performance of IPOs in event time spuriously induces IPOs to have low average returns even if there are no average abnormal returns ex ante. Schultz claims that there is no underperformance of IPOs in calendar time. Gompers and Lerner (2003) convincingly show that, in an earlier sample from 1935 to 1972, IPOs do not underperform aggregate benchmarks in contrast to the post-1970 sample initially examined by Ritter

*Ang, aa610@columbia.edu, Columbia Business School, 3022 Broadway, 805 Uris, New York, NY 10027; Gu, lgu@fordham.edu, Graduate School of Business 403A, Fordham University, 113 West 60th St., New York, NY 10023; and Hochberg, y-hochberg@kellogg.northwestern.edu, Kellogg School of Management, Northwestern University, 2001 Sheridan Rd., Evanston, IL 60208. We thank Kodjo Apedjinou, Geert Bekaert, Alon Brav, Liz Demers, Michelle Lowry, Stephan Siegel, Larry Dann and Richard Startz (the referees), and seminar participants at Columbia University and the 2000 Western Finance Association Annual Meetings for helpful comments.

[1] Recent summaries of the large IPO literature are provided by Ritter (1995) and Ritter and Welch (2002).

(1991). Gompers and Lerner suggest that the poor performance of offerings in the NASDAQ era could simply arise by chance. The Gompers and Lerner (2003) study implies that the IPO underperformance in the last three decades may be just a small sample effect. That is, there may be no IPO underperformance ex ante, but in the post-1970 period we may have just drawn a small sample where too many IPOs perform very poorly ex post.

A small sample explanation to the IPO underperformance puzzle is initially suggested, but not investigated, by Loughran and Ritter (1995) who propose that IPOs initially have high valuations because investors are betting on long shots—that they have identified the next Microsoft, TCBY, or eBay. If these investors are rational and there is no underperformance ex ante, an IPO underperformance in a small sample will result if the small sample does not contain enough draws of these high performing IPOs. That is, ex post, the sample of IPOs is small enough that there is a marked difference between the small sample distributions of the statistics measuring IPO performance and their long-run asymptotic distributions, where in the population distribution, IPOs do not underperform.

As a simple example, suppose that the true population of IPOs has a small proportion (say 2%) of star performers that have extraordinarily high returns. The majority of IPOs (say 70%) exhibit, on average, zero abnormal returns, while a minority of IPOs (the remaining 28%) display, on average, low abnormal returns. In a small sample, we may over-sample from the distributions representing zero, or low, abnormal returns. This implies that we may easily under-sample star performing firms (say 1% in the sample, as opposed to 2% in the population distribution). In the small sample, when we compute average long-run returns of IPOs, we find an IPO underperformance, but this underperformance arises because the small sample distribution does not match the population distribution of IPOs. Hence, the average underperformance of IPOs may be due to observing too few spectacularly successful IPOs in the data than we expected ex ante from the population distribution.

In this study, we make three main contributions to the IPO literature. First, we show that the post-1970 sample of IPOs exhibits significant underperformance in both event time and calendar time. Schultz (2003) considers calendar-time returns on IPOs less than 60 months old following the offering and finds that the average abnormal calendar-time return is close to zero. Building on Schultz (2003), we also consider well-defined trading strategies of an IPO portfolio consisting of IPOs that have gone public within a particular formation period, but we consider holding period returns of this portfolio over horizons longer than one month. In particular, when we consider holding period returns longer than six months, we find that IPO underperformance reappears. Schultz (2003) misses this calendar-time IPO underperformance by only considering a short holding period horizon. Similarly, we find that IPO underperformance is sensitive to the portfolio formation period. While Schultz (2003) finds no underperformance for a portfolio formation period of 12 months, IPO underperformance reemerges when we expand the portfolio formation period to include IPOs that have gone public over

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

the last two to three years. Hence, we show that the Ritter (1991) finding of low IPO returns remains robust to measurement in both event and calendar time.[2]

Second, we introduce a novel model of returns over time for IPO firms. We build a Markov model that captures the intuition of the distribution of IPO returns being a mixture of star performers, average performers, and firms that underperform. The distribution of star performers has expected returns that are high, but this occurs with low probability. Average or underperforming IPO returns are drawn with much higher probability, and these have average zero and negative excess returns, respectively. At each point in time, an IPO's return is drawn from one of these three distributions, and, following Hamilton (1989), which distribution prevails at each point in time is determined by a Markov variable that is unobserved to the econometrician. The Markov states are persistent, so that a firm that has experienced Microsoft-type draws in the past is more likely to draw Microsoft-type returns in the future.

Our flexible Markov mixture data generating process (DGP) can capture small sample bias or Peso problem effects. As Evans (1996) demonstrates, Markov models are ideal for capturing differences between population distributions and sample realizations because the estimation method permits the implied probabilities of drawing regimes (the Markov process) to be inferred endogenously. This allows the parameters of each distribution (outperforming, average, or underperforming) to be estimated under the null of zero average abnormal returns. Markov models have been previously used to investigate Peso problems in time-series data. For example, Bekaert, Hodrick, and Marshall (2001) examine a Peso problem explanation for the expectations hypothesis in interest rates, Evans and Lewis (1995) examine small sample issues in unbiasedness hypothesis tests with exchange rate data, and Rietz (1988) and Cecchetti, Lam, and Mark (1993) argue that the equity premium in stock market data is high because of rare adverse events. In contrast to these studies, we analyze a small sample explanation in the cross section of IPO returns using event-time returns.

Third, we find that small sample bias is very unlikely to account for the magnitude of IPO underperformance observed in the post-1970 sample. We estimate the Markov switching model using Gibbs sampling, which is a fast and tractable Bayesian estimation technique. Gibbs sampling is an estimation method that is particularly suitable for problems where likelihood functions are difficult to derive or maximize, or where only conditional rather than full likelihood distributions are available.[3] We use the model estimates to generate small sample distributions of IPO long-horizon abnormal returns under the null that there is no ex ante IPO underperformance. We compare the small sample distributions

[2] Dahlquist and de Jong (2003) and Viswanathan and Wei (2003) argue that Schultz's (2003) findings are due to an extreme assumption that the number of IPO events drops to zero after a negative abnormal return. This non-stationarity causes Schultz's abnormal return estimator to be not well defined in large samples. However, Viswanathan and Wei (2004) show that in the Schultz (2003) setting, event-time returns are consistent estimators of the null hypothesis of market efficiency and that event returns asymptotically converge to zero under standard assumptions. In contrast, to these studies, we show IPO underperformance in calendar time is sensitive to the portfolio formation strategy and reappears when longer formation periods or holding periods are considered.

[3] See Kim and Nelson (1999) for an overview of Markov switching models and the Gibbs sampling procedure.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

with the estimated point statistics of IPO long-horizon returns from actual data. We find that the small sample distributions implied by the model do not remotely come close to encompassing the long-horizon point statistics in the data. Hence, we fail to find a small sample explanation for the IPO underperformance effect post-1970, suggesting that the IPO underperformance phenomenon is not "simply an historical accident" (Gompers and Lerner (2003), p. 1931).

Our approach is related to the statistical inference problems in long-horizon returns raised by Conrad and Kaul (1993), Barber and Lyon (1997), Kothari and Warner (1997), and Brav (2000). These authors show that statistics measuring long-run performance relative to a benchmark, such as buy-and-hold and cumulative abnormal returns, are subject to severe small sample biases. However, they do not explicitly consider DGPs that impose the null of no underperformance in a model designed to capture Peso problems.

In our analysis, we concentrate on using broad market-based benchmarks because it is uncertain which risk adjustment is appropriate at the firm level. For example, Brav and Gompers (1997) and Brav, Geczy, and Gompers (2000) argue that if risk adjustments are made to equity returns on the basis of size and book-to-market ratios, then the IPO underperformance effect fails to appear.[4] Eckbo and Norli (2005) argue for additional controls for leverage and liquidity. On the other hand, Loughran and Ritter (2000) show that correcting for abnormal performance using Fama and French (1993) size and book-to-market factors is inappropriate because the Fama-French factors are contaminated by the effects of new firm issues. Because of these issues, Ritter and Welch (2002) stress that the IPO long-run underperformance puzzle is not one of selecting appropriate firm risk adjustments, but rather that IPO firms, or firms with characteristics similar to IPOs, perform poorly compared to market-based benchmarks. Schultz (2003) also employs aggregate benchmarks, and Gompers and Lerner (2003) document that there is no IPO long-run underperformance relative to broad-based indices in the pre-NASDAQ sample.

The remainder of the paper is organized as follows. Section II describes the data used in the paper and presents summary statistics of IPO firm long-run returns for event- and calendar-time portfolios. Section III describes the Markov model underlying our small sample analysis and discusses the estimation results. In Section IV, we apply the model to investigate if the IPO underperformance post-1970 can be explained by small sample bias. Section V concludes.

II. Data

Our data consists of two IPO samples. The first sample, which we refer to as the full sample, consists of firms going public from 1970 to 1996. These firms are drawn from the Securities Data Corporation (SDC) Global New Issues database. To be included in the sample, an IPO firm must have an offer price greater than one dollar and must be subsequently listed on the Center for Research in Securities Prices (CRSP) NYSE-AMEX-NASDAQ tapes within six months of

[4]Similar appropriate benchmarking arguments are made by Eckbo, Masulis, and Norli (2000) for the underperformance of seasoned public offerings.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

the offering date. In line with common practice, we exclude from our sample all unit offerings, REITs, ADRs, limited partnerships, and public offerings of closed-end funds. The full sample consists of 4,843 IPOs taking place during the period 1970 to 1996. The second sample, obtained from Jay Ritter's IPO database at www.iporesources.org comprises 1,524 firms conducting IPOs in the 1975–1984 period.[5] We concentrate primarily on reporting our results for the full sample, and comment on how our methodology fares on Ritter's sample. To describe our data, we first confirm the existence of an IPO underperformance phenomenon in event time in Section A, following the original findings of Ritter (1991), and in calendar time in Section B, contrary to the findings of Schultz (2003).

A. Event-Time IPO Returns

Following standard practice, we construct benchmark-adjusted returns for stock i relative to benchmark m in month t as

$$r_{it}(m) = R_{it} - R_{mt}, \tag{1}$$

where R_{it} is the raw return of firm i in event month t and R_{mt} is the benchmark return in event month t. We compute IPO benchmark-adjusted returns as the raw returns on an IPO minus the benchmark return for the corresponding period. We use three benchmarks: i) the CRSP value-weighted NYSE and AMEX index, ii) the CRSP value-weighted NASDAQ index, and iii) the CRSP smallest decile of NYSE firms. These benchmarks are used in many IPO studies as they represent a set of aggregate indices that are easily investable and represent benchmark alternatives to an IPO investment. We compute returns in equation (1) from the first listing on the CRSP daily return tapes. Event months are defined as successive 21 trading day periods. Thus, returns for the first month comprise the returns on listed days 2–22, the second month of returns comprises the returns of listed days 23–43, and so on.

Following Ritter (1991), we define a cumulative average benchmark m-adjusted excess return (CAR) to event-month horizon s as

$$\mathrm{CAR}_s(m) = \sum_{t=1}^{s} AR_t(m), \tag{2}$$

where

$$AR_t(m) = \frac{1}{n_t} \sum_{i=1}^{n_t} r_{it}(m)$$

and n_t is the number of stocks in the IPO portfolio in event month t. Thus, $AR_t(m)$ is the average benchmark-adjusted return, where the averaging is done across all IPO firms in event month t. Hence, the $\mathrm{CAR}_s(m)$ statistic cumulates the average abnormal IPO returns across various horizons s. When a firm is delisted during event month t, the return of that IPO is computed until the day of delisting. We use the notation CAR(NYSE/AMEX), CAR(NASDAQ), and CAR(SMALL) to

[5]Ritter's (1991) original sample size is 1,526. Of these, we failed to match two firms (Area Communication and Advanced Semiconductor) to returns in CRSP.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

indicate cumulative abnormal returns calculated using excess returns relative to the NYSE and AMEX index, the NASDAQ index, and the CRSP smallest size decile, respectively.

We also compute cumulative excess holding period returns (CHPs) of stock i relative to benchmark m until the earlier of horizon event month s or its delisting:

$$\text{CHP}_{is}(m) = \prod_{t=1}^{s}(1 + r_{it}(m)) - 1, \tag{3}$$

where $r_{it}(m)$ is the excess return of stock i relative to benchmark m defined in equation (1). The one-period excess return, $r_{it}(m)$, is the return to a zero-cost strategy that goes long an IPO and shorts the benchmark m portfolio. Cumulating these returns provides the long-horizon return to this zero-cost strategy. We report the average CHP across IPO firms:

$$\text{CHP}_{s}(m) = \frac{1}{n_s}\sum_{i=1}^{n_s}\text{CHP}_{is}(m). \tag{4}$$

Since the CHPs are holding period returns, to easily compare CHPs across different horizons we compute annualized CHP statistics using the transformation,

$$\text{CHP}_{s}(m)^{\text{annualized}} = (1 + \text{CHP}_{s}(m))^{\frac{12}{s}} - 1. \tag{5}$$

Similar to the notation for the CARs, we use the notation CHP(NYSE/AMEX), CHP(NASDAQ), and CHP(SMALL) to denote CHPs computed relative to the various benchmarks. We use both the CAR and CHP statistics to measure IPO performance.

Table 1 reports various summary statistics of event-time IPO returns. We turn first to the number and proportion of surviving IPOs, presented at the top of the table. There is remarkable attrition in the number of IPO firms surviving after the date of their IPO. While the majority (over 98%) of IPOs survive their first year, 39% of IPOs delist within five years. This implies that the delisting process is an important part of modeling the distribution of IPO returns, which we explicitly take into account in our empirical framework.[6]

Second, the CARs clearly show the IPO underperformance effect. As can be seen from the mean CARs reported in Panel A of Table 1, IPOs underperform as early as after one year post-issue in event time. For example, at a 12-month horizon, the average CAR is −6% (−5%) relative to small stocks (NASDAQ). After 60 months, the value of the CAR statistic is a dramatic −16% relative to small stocks, −23% relative to NYSE/AMEX, and −31% relative to NASDAQ. Using Ritter's (1991) t-statistics, all the CAR t-statistics corresponding to these very large negative CAR point estimates are highly significant. However, these t-statistics must be interpreted with care because Barber and Lyon (1997) show

[6]Note that not all delistings of IPO firms are necessarily due to bankruptcy or liquidation. A significant proportion of firms delist due to merger or acquisition activity. The event-time CAR and CHP statistics do not need to be adjusted for delisting returns because these statistics take data only up to the delisting date. In contrast, calendar-time returns must be adjusted for a delisting return because they represent investable portfolio returns as we discuss below.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

TABLE 1

Event-Time IPO Returns

Panel A reports summary statistics of the benchmark-adjusted cumulative average returns (CAR) as in equation (2), which following Ritter (1991) are not annualized. Panel B reports the annualized cumulative abnormal excess holding period returns (CHP) in equation (5) of benchmark returns. We report *t*-statistics, computed following Ritter (1991), in parentheses under the corresponding mean. The benchmarks NYSE/AMEX, NASDAQ, and SMALL are the CRSP value-weighted NYSE/AMEX index, the value-weighted NASDAQ index, and the smallest NYSE size decile, respectively. We denote significance at the 5% and 1% levels with * and **, respectively. The sample period is January 1970 to December 1996 with the last five-year period ending at December 2001.

	Event Month					
	1	12	24	36	48	60
Number of surviving firms	4,843	4,767	4,369	3,882	3,411	2,966
Percentage of surviving firms	100%	98.4%	90.2%	80.2%	70.4%	61.2%
Panel A.						
CAR(NYSE/AMEX)	0.008**	−0.071**	−0.164**	−0.221**	−0.231**	−0.227**
	(2.76)	(−7.30)	(−11.5)	(−11.9)	(−10.1)	(−8.26)
CAR(NASDAQ)	0.010**	−0.052**	−0.156**	−0.243**	−0.305**	−0.313**
	(3.66)	(−5.45)	(−11.1)	(−13.3)	(−13.5)	(−11.6)
CAR(SMALL)	0.007*	−0.060**	−0.146**	−0.170**	−0.156**	−0.158**
	(2.46)	(−6.26)	(−10.3)	(−9.24)	(−6.87)	(−5.82)
Panel B.						
CHP(NYSE/AMEX)^{annualized}	0.096**	−0.046**	−0.049**	−0.044**	−0.033**	−0.029**
	(3.16)	(−5.28)	(−7.17)	(−7.72)	(−2.87)	(−5.38)
CHP(NASDAQ)^{annualized}	0.127**	−0.028**	−0.044**	−0.045**	−0.041**	−0.036**
	(4.22)	(−3.18)	(−6.75)	(−8.49)	(−8.32)	(−7.11)
CHP(SMALL)^{annualized}	0.083**	−0.037**	−0.041**	−0.028**	−0.013*	−0.014*
	(2.81)	(−4.32)	(−6.07)	(−4.50)	(−2.10)	(−2.22)

that the small sample distributions for the CAR statistics are severely skewed compared to Ritter's (1991) asymptotic distributions. In our empirical work, we directly construct a small sample distribution under the null of zero IPO underperformance and directly measure the significance of the CAR point estimates.

In Panel B of Table 1, the CHPs display similar patterns to the CARs, showing that the IPO underperformance starts as early as one year in event time. For example, the average CHP relative to NYSE/AMEX is −4.6% per annum at a one-year horizon, and −4.4% per annum at a three-year horizon. The average CHP relative to NYSE/AMEX decreases to −2.9% per annum at the five-year horizon. In summary, these results confirm Ritter's (1991) results that there exists a strong IPO underperformance effect for IPO performance in event time relative to aggregate benchmarks.

B. Calendar-Time IPO Returns

While Table 1 confirms IPO underperformance in event time, Schultz (2003) argues that there is no evidence of IPO underperformance in calendar time. Schultz proposes that higher stock prices result in more equity issuance, and that this pseudo-market timing is behind the underperformance in event-time, equal-weighted abnormal returns. Calendar-time abnormal returns based on weighting each calendar period equally are not affected by psuedo-market timing and Schultz argues that there is no IPO underperformance in calendar time. We show here that IPO underperformance is also seen in calendar time, contrary to Schultz's claims.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

Schultz (2003) concentrates only on one-month holding period returns. To examine calendar-time returns of IPO returns, we generalize the holding period to look at horizons longer than one month, up to a 60-month holding period. In month t, we form an IPO portfolio by placing an equal amount of money in all IPOs that have gone public over the last F months (the portfolio formation period). This portfolio is held from time t to $t+k$. At time $t+k$, the portfolio is rebalanced to only hold IPOs that have gone public over the last F months. Hence, our calendar-time IPO portfolio returns represent the returns on an equal-weighted portfolio of IPOs, each IPO no older than F months. We examine holding period returns over $k = 1$ to $k = 60$ months. After computing the calendar-time IPO raw returns, we subtract the benchmark returns from the IPO portfolio returns to compute benchmark-adjusted holding period returns in calendar time.

Because of the large number of IPO delistings (see Table 1), it is important to adjust for the delisting return. Shumway (1997) recommends assigning a delisting return of -0.3 to an arbitrary firm delisting from CRSP and Shumway and Warther (1999) recommend using a corrected return of -0.55 for a delisting from the NASDAQ stock exchange. In computing their delisting returns, Shumway and Warther track the returns of firms after they delist using data from the Pink Sheets (published by the National Quotations Bureau) up to 100 days post-delisting. In the post-1970 sample, almost all (93%) delisting IPOs delist from NASDAQ, so we assign a delisting return of -0.55 to all delisting IPOs. This correction is likely to be conservative for two reasons. First, IPOs tend to have low event-time returns relative to the average seasoned, listed firm. Second, the final return from a firm that liquidates might not be received for many months after the delisting. Delisting returns are important for investable calendar-time returns because the money returned from investing in the delisting firm is reinvested in the IPO portfolio going forward.

Table 2 reports calendar-time IPO returns over various holding period horizons. The average returns are annualized to make comparison easier. To use all the data, we report the means using overlapping observations, but the point estimates are very similar using non-overlapping observations. To account for the moving average errors induced by the overlapping observations, we compute t-statistics with Newey and West (1987) standard errors, using a lag length of one less than the holding period horizon. Note that the case of $k = 1$ involves no overlapping observations. If we compute the standard errors with simple OLS t-statistics, the magnitude of the OLS t-statistics is approximately four to six times larger than the robust t-statistics reported in the table.

When we consider a portfolio of firms that have gone public over the last year ($F = 12$), Panel A of Table 2 shows that there is no statistically significant underperformance for one-month holding period returns, no matter which benchmark is used. In fact, for a one-month horizon, IPOs that have less than a one-year listing anniversary outperform the NYSE/AMEX index by 1.4% per annum. This is the result reported by Schultz (2003).[7]

[7]Schultz (2003) takes a universe of IPO returns up to 60 months following the offering, and then considers calendar-time returns of these firms. This corresponds to a one-month holding period horizon, but Schultz's formation period interval changes over time and is weighted toward selecting IPOs with short and intermediate histories post-offering.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

TABLE 2

Calendar-Time IPO Returns

Table 2 reports mean holding period returns (unadjusted) and mean adjusted-holding period returns, relative to NYSE/AMEX, NASDAQ, and SMALL benchmarks. We form a portfolio of IPOs that have gone public within the past one year (top panel) or past three years (bottom panel) and hold this portfolio for various holding period horizons. The means are computed with overlapping observations, so we report Newey-West (1987) *t*-statistics in parentheses with a lag length of one less than the holding period horizon in months. All the returns in the table are reported in annualized terms. We denote significance at the 5% and 1% levels with * and **, respectively. The sample period starts from January 1970 for the three-year formation period and from January 1972 for the one-year formation period. Both samples in Panels A and B end at December 2001.

	Holding Period Horizon (months)				
	1	6	12	36	60
Panel A. Formation Period = 1 Year					
Mean return(unadjusted)	0.159**	0.129*	0.110**	0.120**	0.139*
	(2.81)	(2.56)	(2.59)	(2.72)	(2.17)
Mean return(NYSE/AMEX)	0.014	−0.014	−0.035	−0.039	−0.023
	(0.41)	(−0.42)	(−1.02)	(−0.66)	(−0.23)
Mean return(NASDAQ)	−0.015	−0.053*	−0.072*	−0.075	−0.069
	(−0.65)	(−2.21)	(−2.58)	(−1.42)	(−0.72)
Mean return(SMALL)	−0.024	−0.057*	−0.075**	−0.102*	−0.124
	(−0.74)	(−2.05)	(−2.92)	(−2.62)	(−1.80)
Panel B. Formation Period = 3 Years					
Mean return(unadjusted)	0.146**	0.110*	0.110*	0.121*	0.114*
	(2.80)	(2.33)	(2.57)	(2.59)	(2.31)
Mean return(NYSE/AMEX)	0.002	−0.031	−0.035	−0.036	−0.071
	(0.07)	(−0.95)	(−1.01)	(−0.58)	(−0.85)
Mean return(NASDAQ)	−0.027	−0.069**	−0.072*	−0.073	−0.130
	(−1.30)	(−2.97)	(−2.60)	(−1.33)	(−1.71)
Mean return(SMALL)	−0.035	−0.073**	−0.074**	−0.099*	−0.206**
	(−1.39)	(−3.03)	(−3.18)	(−2.44)	(−3.21)

However, as we increase the holding period from one month to 60 months, the IPO underperformance puzzle reemerges. Beginning with a holding period horizon of six months, the point estimates in Table 2 are negative relative to all three benchmarks. At a one-year horizon, there is an average performance of −7.2%, and −7.5% per annum relative to the NASDAQ and small stock indices, respectively. This underperformance is significant at the 5% level. The average performance of IPOs relative to the total NYSE/AMEX benchmark is −3.5% per annum at the one-year horizon, but is not statistically significant. Although the IPO performance relative to the NYSE/AMEX and NASDAQ indices is statistically insignificant at the 5% level at the 60-month horizon, the magnitude of underperformance is 2.3% per annum for the NYSE/AMEX benchmark and around 7% per annum for the NASDAQ benchmark. For a 60-month holding period horizon, the IPO portfolio underperforms small stocks by an economically very large 12.4% per annum, but this is only statistically significant at the 10% level.

For the three-year formation period reported in Panel B, the evidence of calendar-time underperformance is even stronger. At the six-month horizon, the performance point estimates are already large and negative, and statistically significant relative to the NASDAQ (−6.9% per annum) and small stock (−7.3% per annum) indices at the 1% level. The performance at a 12-month horizon is −3.5%, −7.2%, and −7.4% per annum relative to NYSE/AMEX, NASDAQ, and small stocks, respectively. While underperformance relative to the broad NYSE/AMEX index is not statistically significant, the underperformance is statis-

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

tically significant at the 5% level for the NASDAQ benchmark and significant at the 1% level for the small stock benchmark. The negative performance estimates increase in magnitude at the 60-month holding period horizon, where the IPO performance point estimates are −7.1%, −13.0%, and −20.6% per annum relative to NYSE/AMEX, NASDAQ, and small stocks, respectively. The underperformance in calendar time relative to small stocks is particularly large and highly statistically significant for all holding period horizons greater than six months. Holding a portfolio of IPOs that go public over the last three years produces stronger evidence of long-term underperformance than using a formation period of just one year, because the longer formation period selects more seasoned IPO firms. Note that underperformance is greater for more seasoned firms in event time: IPOs actually tend to outperform benchmarks in the first six months of event time, but tend to underperform significantly over three to five years post-issue in event time.[8]

Why does the IPO underperformance show up in calendar time only for long holding period horizons? First, by focusing only on one-month holding periods, Schultz's method does not capture the long-term performance of IPOs. Table 2 focuses on the effects of changing the formation period of the IPO portfolio and the holding period horizon. As Schultz holds the IPO portfolio only for one month, the average return is heavily weighted toward short-run and intermediate-term event-month returns. Once we consider different holding period horizons and different formation periods, the IPO underperformance reemerges. Longer holding period horizons, or longer portfolio formation periods, allow the portfolio to contain more seasoned IPOs, which have relatively low average returns. Finally, Schultz also ignores the returns of delisting IPO firms. As shown in Table 1, there is a remarkable proportion of IPO firms that delist from CRSP within five years of issue.

Having made the case for long-run IPO underperformance in both event time and calendar time in the post-1970 sample, we now examine the hypothesis that the IPO underperformance may be due to small sample or Peso problem effects.

III. The Model

A. Capturing a Small Sample Problem

The essence of a small sample explanation for long-run IPO underperformance is that the data we observe may not contain the same number of high flying IPOs that we expect from the population distribution. We illustrate this intuition in Figure 1. Suppose that IPO returns are drawn from one of three states: i) an extraordinary state earning 70% over benchmark, ii) an average state where the

[8]We comment on but do not report the results if we make the extreme assumption that a delisting IPO returns all of its money back to an investor immediately (so there is a delisting return of 0.0%). In this case, with a three-year portfolio formation period, the IPO portfolio performance is −0.2%, −4.4%, and −9.2% per annum relative to NYSE/AMEX, NASDAQ, and small stocks, respectively. These averages are statistically insignificant at the 5% level. However, this scenario is extremely unrealistic because many IPOs that delist go bankrupt, and the remainder of any invested money is only realized with a long lag. In fact, we view the Shumway and Warther (1999) delisting correction of −0.55 as conservative, given the very low event-time IPO returns reported in Table 1.

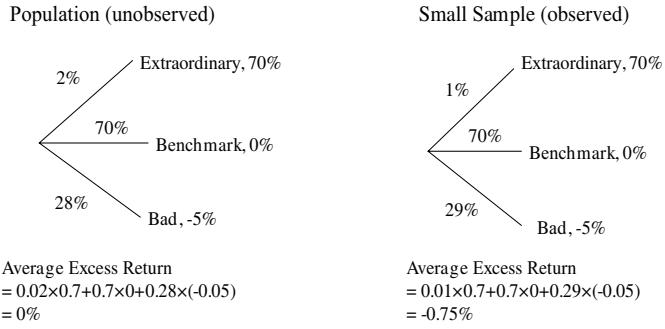
https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

IPO earns the benchmark return, and iii) an underperforming state where the IPO underperforms the benchmark by 5%. In population, the extraordinary state occurs 2% of the time, the average state occurs 70% of the time, and the bad state occurs 28% of the time. The extraordinary state has a very high mean, but occurs rarely so it represents a draw of a highly successful IPO. The average abnormal return for the population is then $0.02 \times 0.7 + 0.7 \times 0 + 0.28 \times (-0.05) = 0\%$.

FIGURE 1

A Simple Example of the Small Sample Bias

Figure 1 shows a population versus small sample distribution, illustrating how ex post underperformance might be realized in a small sample generated from a population distribution with no ex ante underperformance.



In a small sample, we may not observe the same frequency of extraordinary, average, or underperforming states as the population frequency. Suppose that in a small sample we observe that extraordinary returns constitute only 1% of the returns instead of the 2% frequency of extraordinary states in the population. If the proportion of the benchmark returns remains the same as the population, at 70%, then we over-sample low return states. In this case, the average abnormal return for the sample is then negative at $0.01 \times 0.7 + 0.7 \times 0 + 0.29 \times (-0.05) = -0.75\%$. Hence, we observe an average underperformance in the sample, but this is because the population distribution and the small sample population are dissimilar. If we were to observe the same frequency of extraordinary returns in the sample as the proportion of extraordinary returns in the population, then there would be no average underperformance in the small sample.

For this type of Peso problem intuition to be reasonable, we would hope that the distribution of IPO returns in data already contains some large observations—the Peso explanation requires that we have not observed enough similar large observations in a small sample. Panel A of Table 3 shows that the right-hand tail of the distribution of monthly IPO returns encompasses some spectacular one-month returns. The magnitude of the top 10 monthly returns is large enough that firms can easily increase their value by 3–6 times within one month, and the highest one-month IPO return is over 2,500% (which corresponds to Club-Theatre Network in its 21st event month). For comparison, the average monthly return for an IPO in our sample is 0.83% per month. Panel B shows that the top 10 IPOs in the five years after their issue date approximately double in price

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

every year.[9] Clearly, Table 3 shows that the observed distribution of IPO returns includes some impressive returns. According to a small sample explanation, the population distribution of IPO returns must contain a higher frequency of these types of returns, or even more spectacular returns.

TABLE 3

Top Performing IPOs

Table 3 reports the highest 10 one-month IPO returns in Panel A. In Panel B, we report the 10 IPOs with the largest 5-year post-issue cumulative returns. The returns in the table are in levels, not percentages. The sample period is January 1970 to December 2001.

Panel A. Top 10 Highest Monthly Returns

Name	Monthly Return	Event Month	Issue Date
Club-Theatre Network Inc	25.0	21	05/15/90
SoloPoint Inc	5.91	28	06/06/96
Smith Micro Software Inc	5.64	54	09/18/95
Viisage Technology Inc	5.61	59	11/08/96
Exploration Co	5.60	38	11/19/79
Humascan Inc	4.50	30	08/12/96
ECO2 Inc	4.36	33	10/22/92
SkyMall Inc	3.99	25	12/11/96
Western Power & Equipment Corp	3.88	59	06/13/95
Hungarian Broadcasting Corp	3.58	36	12/20/95

Panel B. Top 10 Largest 5-Year Cumulative Returns

Name	5-Year Cumulative Return	5-Year Annualized Cumulative Return	Issue Date
CMG Information Services Inc	147	1.72	01/25/94
American Power Conversion	70.3	1.35	07/22/88
Network Appliance Inc	55.8	1.24	11/21/95
Ascend Communications Inc	50.3	1.20	05/12/94
SDL Inc	42.4	1.13	03/15/95
Ryan's Family Steak Houses	40.3	1.10	07/13/82
Zoltek Cos Inc	34.6	1.04	11/06/92
StrataCom Inc	31.7	1.01	07/21/92
Cisco Systems Inc	24.2	0.91	02/16/90
Liz Claiborne Inc	24.1	0.91	06/09/81

It would be tempting to construct a population distribution of IPOs by just sampling repeatedly from the extreme IPO returns in Table 3. However, we cannot be sure that these returns represent the true distribution of star performers, particularly under the null hypothesis of no ex ante IPO underperformance. The IPO data have an overall average underperformance and the data may more correctly represent the appropriate distribution under the alternative hypothesis that there exists long-run IPO underperformance. The true distribution of the outstanding performers under the null of no ex ante IPO underperformance is not directly observable. However, we now describe how the distribution of IPO returns under the null of no ex ante underperformance can be inferred from a rigorous model that captures the simple intuition of the picture in Figure 1.

[9] Interestingly, Microsoft is not among these firms. For comparison, the cumulative five-year annualized post-IPO return of Microsoft is 72%.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

B. A Markov Switching Model

Extending the simple intuition of Figure 1 into an econometric model requires several steps. First, instead of discrete possible outcomes (for example, outperforming, benchmark, and underperforming) for an IPO's return in excess of benchmark r_{it}, we specify a series of distributions that depend on a state s_{it} prevailing at time t. If the prevailing state corresponds to an outperformance state, then the IPO's return is drawn from the corresponding outperformance distribution. We specify these state-dependent distributions to be normal. Second, we specify the states s_{it} to be persistent so if a firm has been an outperformer in the past it is more likely to be an outperformer next period. Finally, we observe the draws of actual IPO returns, but the econometrician does not observe the sequence of states so the estimation method must infer the states from the data.

Formally, this is a Markov switching model of the type introduced by Hamilton (1989), where the states s_{it} follow a Markov chain, and the IPO draws are from time-varying mixtures of normals.[10] As Bekaert, Erb, Harvey, and Viskanta (1998) and Timmermann (2000), among others, comment, mixtures of normal distributions are easily able to capture heteroskedasticity, fat tails, and other features of equity returns. At each point in time, conditional on no delisting, the abnormal IPO return r_{it} follows the process,

$$r_{it} = \mu(s_{it}) + \sigma(s_{it})\varepsilon_{it}, \tag{6}$$

where ε_{it} is IID $N(0,1)$ and the state s_{it} follows a Markov chain that can take values $s_{it} = 1, \ldots, K$ states. For simplicity, we specify that the draws of ε_{it} and s_{it} are uncorrelated across firms in event time.

In equation (6), the IPO return is normally distributed conditional on the state s_{it}. However, as the prevailing state s_{it} changes across time, the IPO returns are drawn from different distributions. This causes the unconditional IPO return to be nonnormal and heteroskedastic. We can regard each IPO as a new draw from the DGP in equation (6). The states s_{it} are drawn from a common Markov chain that causes the individual IPO returns to be correlated.

We estimate models with $K = 2$ and $K = 3$ states. In the case of two regimes, the Markov transition probability matrix takes the form,

$$\Pi = \begin{bmatrix} P_{11} & 1 - P_{11} \\ 1 - P_{22} & P_{22} \end{bmatrix}, \tag{7}$$

where $P_{11} = Pr(s_{it} = 1 | s_{i,t-1} = 1)$ and $P_{22} = Pr(s_{it} = 2 | s_{i,t-1} = 2)$ are constants and are the same across firms. The stable probability $\pi_1 = Pr(s_{it} = 1)$ corresponding to the system is given by

$$\pi_1 = \frac{1 - P_{22}}{2 - P_{11} - P_{22}},$$

[10] Markov switching models have been used to model equity returns by, among others, Turner, Startz, and Nelson (1989), Hamilton and Susmel (1994), Hamilton and Lin (1996), Ramchand and Susmel (1998), Perez-Quiros and Timmermann (2001), and Ang and Bekaert (2002).

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

which satisfies the relation $\pi = \Pi\pi$, where $\pi = (\pi_1\ \pi_2)'$.[11]

If $K = 2$, we can think of the two distributions corresponding to $s_{it} = 1, 2$ as corresponding to an outperformance distribution and an underperformance distribution. While each IPO may be in a different state, we restrict the transition probabilities of the IPO states to be the same across IPOs. The transition probabilities are persistent and capture the notion that a firm that has outstanding returns in the past is more likely to be an outperforming firm in the future.

We estimate the model under the null of zero expected abnormal outperformance, so we place restrictions on the model parameters such that $\mathrm{E}(r_{it}) = 0$. This involves the restriction:

$$\mathrm{E}[r_{it}] \quad = \quad \pi_1\mu_1 + \pi_2\mu_2 \quad = \quad 0. \tag{8}$$

Hence, there is a restriction involving the conditional means of the distributions of each state,

$$\mu_2 \quad = \quad -\frac{\pi_1\mu_1}{\pi_2}, \tag{9}$$

which we impose in the estimation. Note that this restriction only involves the means of the state-dependent distributions, but not the volatility parameters.

For $K = 3$ states, we can interpret the states as representing periods of outperformance, benchmark performance, or underperformance. In this case, we specify the transition probability matrix of all IPO returns to be:

$$\Pi \quad = \quad \begin{bmatrix} P_{11} & 1 - P_{11} & 0 \\ P_{21} & P_{22} & 1 - P_{21} - P_{22} \\ 0 & 1 - P_{33} & P_{33} \end{bmatrix}, \tag{10}$$

where $P_{11} = Pr(s_{it} = 1|s_{i,t-1} = 1)$, $P_{21} = Pr(s_{it} = 1|s_{i,t-1} = 2)$, $P_{22} = Pr(s_{it} = 2|s_{i,t-1} = 2)$, and $P_{33} = Pr(s_{it} = 3|s_{i,t-1} = 3)$. With the specification in equation (10), firms transit through the benchmark performance state on their way from the outperforming state to the underperforming state, and vice versa. This means that we do not allow a firm to jump immediately from outstanding performance today to underperformance next period. However, we also consider the case of an unrestricted three-state Π matrix where a direct transition from outperformance to underperformance can occur.

To impose the null of zero expected abnormal performance in the case of $K = 3$ states, we impose the restriction,

$$\mathrm{E}[r_{it}] \quad = \quad \pi_1\mu_1 + \pi_2\mu_2 + \pi_3\mu_3 \quad = \quad 0, \tag{11}$$

where $\pi_j = Pr(s_{it} = j)$ are the stable probabilities of the system. Rearranging, we can write μ_3 as a function of μ_2 and μ_1:

$$\mu_3 \quad = \quad \frac{1}{\pi_3}(-\pi_2\mu_2 - \pi_1\mu_1). \tag{12}$$

[11] Estimating the model under time-varying transition probabilities is not computationally feasible and cannot be done with conjugate draws (see the Appendix). However, because the rejection of the null of no underperformance is so strong and the fact that the IPO draws in event time are correlated through the stable distribution of the states s_{it}, we believe that generalizing the model to include time-varying probabilities would not significantly change our results or conclusions.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

Furthermore, we identify the abnormal return in state $s_{it} = 2$ as the average performing state with an expected abnormal return of zero, so we set $\mu_2 = 0$. This yields the restriction:

$$\mu_3 = -\frac{\pi_1 \mu_1}{\pi_3}. \tag{13}$$

How does this model capture the Peso problem intuition? We impose the null of expected benchmark performance through equations (9) or (13). However, we allow a state where an IPO may potentially spectacularly outperform the benchmark return. The estimation reconciles the apparent IPO underperformance in data by estimating the parameters of the outperformance distribution such that the null is satisfied. We do not directly observe the outperformance distribution in data, but the model is able to capture the Peso effect through the transition probabilities and the state-dependent mean parameters.

To complete the model, we specify the attrition process of IPOs. This is important, because as Table 1 shows, 40% of IPOs delist within five years after their issue date.[12] We first model the delisting process and then, conditional on no delisting, apply the Markov switching model of IPO returns in equation (6). We draw the delisting time T_i of the ith IPO from a geometric distribution with probability p,

$$Pr(T_i = k) = (1-p)^{k-1} p, \tag{14}$$

where T_i is in months. If $T_i < 60$, then the IPO delists within five event years post-issue, whereas if $T_i \geq 60$, we observe a full five-year event history of that IPO's returns. Hence, equation (14) represents a truncated geometric distribution. We assume that the probability p is the same across all firms, and the delisting time of each IPO is IID. Conditional on T_i, the IPO's returns are drawn from the Markov switching process in equation (6) for T_i observations. A more common specification for a point process such as T_i is a Poisson distribution, but we show that a Poisson distribution cannot fit the persistent decay pattern of the IPO attrition rates observed in the data.

The model estimation is non-trivial because of the large cross section of firms (4,843 IPOs over the full sample). Recent advances in Bayesian methods allow us to estimate the model using Gibbs sampling techniques following Albert and Chib (1993). We provide details of the estimation method in the Appendix. The Gibbs sampler has several advantages. First, because we model the delisting process and the return of an IPO is specified conditional on no delisting, direct construction of the likelihood function is highly complex. The Gibbs sampler operates on a series of conditional distributions, which are well specified in our model. For example, the distribution of the delisting time, T_i, is a geometric distribution (equation (14)) and the distribution of the IPO return conditional on T_i and each state s_{it} is a normal distribution (equation (6)).

[12] An alternative way to model IPO delisting is to include a fourth, absorbing state into the transition probability matrix. However, the algorithms used to estimate regime switching models require that the transition matrix be ergodic to filter the states that are unobserved to the econometrician, particularly if the initial state is set to be the stable probabilities of the Markov process. See, for example, Hamilton (1989).

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

Second, instead of performing a complex optimization, we construct the posterior distribution of the parameters by simulating from each conditional distribution in turn. This is much easier than maximizing a highly nonlinear likelihood function. Finally, the Gibbs sampler accounts for parameter uncertainty, which we take into account by analyzing the small sample performance of the CAR and CHP statistics. We estimate the models using monthly IPO returns adjusted by the CRSP value-weighted NYSE/AMEX index of post-event time and use a 60-month event horizon. We also estimate (but do not report) the models for Ritter's (1991) original data sample, where the event horizon is 36 months.

Naturally, our statistical inference depends on the ability of our DGP to match data and to successfully capture a Peso problem. We believe that our model is more than flexible enough for this purpose because the superstar state can potentially have a very large expected return, the superstar state is potentially persistent or very fleeting, and there are no restrictions on the fraction of returns belonging to each state.

C. Estimation Results

We first comment on the results of the delisting process and then describe the estimated parameters of the Markov switching process. The point estimate of the geometric probability p in equation (14) is 0.008. This parameter is very precisely estimated with a standard error of 0.001 because of the large number of IPOs in the sample. Table 4 reports the actual and average numbers of surviving firms for various event months for our geometric distribution in equation (14) and a Poisson distribution for comparison. The fit of the geometric distribution is very good, but slightly underestimates the actual number of surviving firms after one year (4,433 versus 4,767 in the data), while matching almost exactly the actual number of surviving firms after five years (3,015 versus 2,966 in the data). In contrast, a Poisson distribution fitted to data has an extremely poor fit, predicting that all firms should delist within five event years because it cannot match the observed slow attrition rate of IPOs in data.

TABLE 4

Actual and Expected Number of Surviving Firms

Table 4 reports the actual number of IPOs surviving k months from issue in the second column (which is the same as the first row of Table 1). The column labeled Geometric reports the expected number of surviving firms from the geometric distribution in equation (14), while the last column reports the expected number of surviving firms from a Poisson distribution fitted to the data. The sample period is January 1970 to December 2001.

Event Month	Actual	Geometric	Poisson
1	4,843	4,843	4,843
6	4,830	4,652	4,843
12	4,767	4,433	4,843
24	4,369	4,026	4,713
36	3,882	3,656	2,050
48	3,411	3,320	89
60	2,966	3,015	0

We report the parameter estimates of the Markov switching part of the DGP in Table 5. We report the mean of the posterior distribution of each parameter together with the standard deviation of the posterior distribution in parentheses.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

Panel A reports results for the two-state model. We can interpret state 1 as the overperformance state, which has a mean of 4.2% per month and a large monthly volatility of 40%, and state 2 represents an underperformance state, which has a mean of −0.8% per month with a monthly volatility of 12%. The estimates of the transition matrix Π in equation (7) show that the underperformance state is very persistent with a half-life of 22 months.[13] In contrast, a firm has a 25% probability of moving from the outperformance state to the underperformance state each month. The transition probabilities in Π correspond to stable probabilities of 16% (84%) for state 1 (2).[14]

TABLE 5

Parameter Estimates

We report parameter estimates of the two- and three-state models in equation (6). For the two-state model in Panel A, the state-dependent means $\mu(s_{it})$ are estimated subject to the restriction in equation (9) with the transition matrix given by equation (7). For the three-state models in Panel B, the means $\mu(s_{it})$ are estimated subject to the restriction in equation (13). The model labeled Constrained Π is estimated with the transition probability matrix Π described in equation (10), while the unconstrained estimation imposes no constraints on the transition probability matrix. The vector π represents the stable probability of the system. The sample period is January 1970 to December 2001.

	State	μ	σ	Transition Matrix Π	π
Panel A. Two-State Model					
	1	0.042 (0.002)	0.396 (0.003)	0.750 0.250 (0.005) –	0.157
	2	−0.008 (0.000)	0.122 (0.001)	0.047 0.953 – (0.001)	0.843
Panel B. Three-State Model					
Constrained Π	1	0.227 (0.014)	0.731 (0.011)	0.684 0.316 0.000 (0.005) – –	0.021
	2	0.000 –	0.102 (0.000)	0.011 0.943 0.046 (0.001) (0.001) –	0.603
	3	−0.013 (0.001)	0.223 (0.001)	0.000 0.073 0.927 – (0.002) –	0.376
Unconstrained Π	1	0.264 (0.044)	0.605 (0.010)	0.289 0.576 0.136 (0.027) (0.022) –	0.018
	2	0.000 –	0.214 (0.002)	0.070 0.930 0.000 (0.000) (0.004) –	0.158
	3	−0.006 (0.000)	0.141 (0.001)	0.002 0.001 0.997 – (0.000) (0.001)	0.824

Panel B of Table 5 reports the results of the three-state model. We report two estimates, one with a transition probability matrix Π following equation (7), where the IPO cannot directly transition from being an outperformer to an underperformer and the other estimate with an unconstrained Π matrix. In the constrained Π estimation, the distribution of IPO returns in the outperformance state 1 has a mean of 22.7% per month with a very high monthly volatility of 73.1%.

[13] Because the Markov states are persistent, the model endogenously generates persistence of IPO returns. However, the large standard deviations of IPO returns make this autocorrelation small in the model and hard to pin down in the data. In the three-state model with a constrained transition matrix, the implied IPO autocorrelation is −0.0006 with a posterior standard deviation of 0.0004. The slight negative autocorrelation results from star performers being more likely to transform into benchmark-performing or underperforming firms than underperforming IPOs becoming star performers. In the data, the mean autocorrelation across IPOs is −0.0443 with a cross-sectional deviation of 0.1680. Thus, our model and IPO data can shed little light on stock-level reversals or momentum.

[14] The implied monthly standard deviation for IPO returns implied by the two-state model is 19.3% per month, which we can compare to the IPO return volatility of 19.2% in data. The corresponding number for the three-state model with the constrained Π matrix is 19.4%.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

This state occurs, on average, 2.1% of the time. While this state has a very high average return, the very top IPO returns in our sample reported in Table 3 comfortably exceed this average return, which suggests that some of the high IPO returns in data could be drawn from this type of outperforming state. The benchmark performance state 2 has zero expected excess returns by construction and has a standard deviation of 10.2% per month. This is the predominant state for IPOs with a stable probability of 60.3%. The underperformance state 3 has a mean of −1.3% per month and a slightly higher standard deviation of 22.3% per month.

This estimation reflects our intuition of a benchmark performance state and an underperformance state from which the majority of IPOs are drawn, and an extraordinary state that occurs rarely (2.1% on average). The transition matrix Π shows that each state is persistent with IPOs that are benchmark (underperformers) having a 94.3% (92.7%) chance to remain in that same state the following month. The outperformance state is less persistent with a probability of $P_{11} = 0.684$ of remaining an outperformer next period, conditional on being an outperformer this period.

The unconstrained Π estimation also broadly reflects this same intuition except that the stable probability of the benchmark performance state 2 decreases to 15.8% (compared to 60.3% in the constrained Π estimation), and the stable probability of the underperformance state 3 increases to 82.4%. Once we allow firms to immediately switch from being winners to underperformers, only 28.9% of outperformers remain outperformers the next month, while 57.6% and 13.6% of outperformers transition to benchmark performance and underperformance states, respectively. Thus, with an unconstrained Π matrix, the winner state 1 becomes even more extreme having a stable probability of only 1.8% and an expected return of 26.4% per month.

An alternative interpretation of the estimation with three states and an unconstrained Π matrix is a system where the majority of IPOs underperform with an average return of −0.6% per month. A smaller number of IPOs have higher returns in line with the benchmark, while a very small fraction (1.8%) have extremely high returns on average. Once an IPO is drawn into the loser state 3 or transitions into this state, it is very likely to remain a loser with a probability of 99.7% per month. A benchmark firm is likely to continue to perform in line with its benchmark status with a probability of staying a benchmark firm of $P_{22}=0.930$. In contrast, the rare outperforming IPO is unlikely to continue its extraordinarily high returns, with a probability of remaining an outperforming IPO of $P_{11}=0.289$, and very quickly transitions to becoming a benchmark performer or an underperformer. Nevertheless, the unconstrained estimation results maintain the intuition of only a small minority of IPOs enjoying very high average returns.

The Bayesian estimation also allows us to compute Bayes factors, which provide a method of testing the null of the two-state Markov model against the three-state models.[15] The Bayes factor is used to compute the posterior odds ratio,

[15]Traditional maximum likelihood ratio tests are computationally very difficult to calculate because of the presence of nuisance parameters that must be integrated out in the test statistics (see, for example, Davies (1987)).

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

$$\frac{p(H_1|Y)}{p(H_2|Y)} = \frac{p(Y|H_1)p(H_1)}{p(Y|H_2)p(H_2)}, \tag{15}$$

where H_1 and H_2 are the models being tested, Y is the data, and $p(H_1)$ is the prior on model H_1 and $p(H_2)$ is the prior on model H_2. For example, in our setting H_1 would be a two-regime Markov model and H_2 would be a three-regime Markov model. Given non-informative priors on both models, the posterior odds ratio simplifies to the Bayes factor, $B_{12} \equiv p(Y|H_1)/p(Y|H_2)$.

We compute the Bayes factor using the harmonic mean estimator of Kass and Raftery (1995) and find extremely strong evidence in favor of the three-state models against the two-state model. The value of $2 \ln B_{12}$ in testing the two-state model against the three-state model with a constrained Π transition matrix is over 7,000, while the value of $2 \ln B_{12}$ for the two-state model against the unconstrained Π model is over 8,800. Any value above 10 is interpreted as very strong evidence against the null model. These values are consistent with the tight posterior standard error bounds for the parameters in Table 5, which result from the fairly large panel of IPOs used in the estimation. We also find strong evidence of the unrestricted Π three-state model against the constrained Π three-state specification with a value of $2 \ln B_{12}$ above 1,400. Nevertheless, we examine the implied small sample statistics of IPO underperformance for all our Markov models.

IV. Is IPO Underperformance a Statistical Fluke?

In this section, we ask whether it is surprising to measure an average long-run IPO underperformance in a small sample and how likely it is that a small sample contains a lower frequency of outperforming IPOs than the population distribution.

A. Robust Statistical Inference

To conduct robust small sample inference, we generate a small sample distribution of the CAR and CHP statistics (in equations (2) and (4)) measuring IPO performance. We construct the small sample distribution from the estimates of the Markov switching model in Table 5 and the delisting process in equation (14) as follows. First, we draw the delisting time T_i for firm i from the distribution in equation (14). If $T_i > 60$, we simulate a full 60-month event-time series for the IPO. If $T_i < 60$, the firm delists prior to the 60-month horizon. Then, we generate a time series of IPO returns for firm i from the Markov process in equation (6) for the number of surviving periods of the IPO.[16]

We simulate 4,843 IPO firms for each small sample. This corresponds to the number of IPOs in our sample in our post-1970 sample period. In the sample, we compute the CAR and CHP statistics and store their values. We repeat this procedure 10,000 times. In this way, we obtain a distribution of small sample

[16] In the rare instance that a simulated return is less than -1 in equation (6), we assume the firm delists at that time. Note that actual delisting returns are not used in computing the event-time CAR and CHP statistics.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

CAR and CHP statistics to which we can compare the point estimates of the CAR and CHP in the actual IPO data. Note that the small sample distribution of the IPO underperformance statistics is constructed under the null of zero expected abnormal IPO returns because the DGP is estimated under this null. We also take into account parameter uncertainty by drawing from the posterior distribution of the parameters. That is, each one of the 10,000 simulated samples is constructed using a different draw from the posterior distribution of the parameters. However, simulating only from the posterior mean of the parameters produces almost identical results because of the tight posterior standard deviations of all the parameter estimates.

B. Empirical Results

Table 6 reports the small sample distribution of CAR statistics from each model. We report results for the full sample in Panel A. The CAR(NYSE/AMEX) estimated in the data for the 60-month event horizon is −0.227 (see Table 1). In contrast, under the two-state model, the CAR small sample distribution is positively skewed, even though there is no expected abnormal IPO performance under the null. Barber and Lyon (1997), Kothari and Warner (1997), and Brav (2000) also report skewed CAR statistics in small samples. The mean and median of the small sample CAR distribution is 0.041, much higher than the −0.227 CAR estimate in data. Note that while the small sample CAR distribution for the unconstrained Π three-state model, with a mean CAR of 0.002, produces less bias than the two-state or the constrained Π three-state model, the mean and median of the small sample CAR distribution is still much higher than the CAR point statistic of −0.227 in the data. However, the large difference per se does not rule out a small sample explanation for the post-1970 IPO underperformance. In order to do this, we must look at the entire small sample distribution to compute a p-value.

Table 6 reports various percentile values for a more detailed picture of the small sample CAR distribution. The data point estimate of −0.227 falls nowhere in the simulated small sample distribution for either the two-state or the three-state models. For the two-state model, the 0.5% cutoff is −0.026. The 0.5% cutoff for the three-state models for the constrained transition probability estimation and the unconstrained transition probability matrix Π, respectively, are −0.037 and −0.081. Since the effective p-value of the −0.227 CAR estimate under the small sample distributions is zero, we overwhelmingly reject the hypothesis that small sample bias can account for the IPO underperformance in the post-1970 sample.

While a small sample explanation may be very unlikely over the post-1970 data, a valid question is that when Ritter (1991) first raised the question of long-run IPO underperformance his shorter data sample might not have been able to rule out a Peso problem explanation. Perhaps it is only with the addition of the late 1980s and 1990s data that the IPO effect has become statistically robust. Panel B of Table 6 investigates this possibility. We use the original Ritter (1991) sample from 1975 to 1984. Using Ritter's original event horizon of 36 months, the CAR in his data sample is −0.251. To construct the CAR small sample distribution corresponding to Ritter's data, we reestimate the models over Ritter's sample period and simulate small samples of 1,524 firms.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

TABLE 6

CAR Small Sample Distribution

Table 6 reports summary statistics of the small sample distribution of the CAR statistic (equation (2)) by simulating from the posterior distribution of the parameter estimates reported in Table 5 using 10,000 simulations. Panel A corresponds to the sample of IPOs going public from January 1970 to December 1996 with an event horizon of 60 months. Panel B corresponds to Ritter's (1991) original sample, consisting of IPOs going public from 1975 to 1984, with an event horizon of 36 months.

		Three-State Model	
	Two-State Model	Constrained Π	Unconstrained Π
Panel A. Full Sample (CAR in data = −0.227)			
Mean	0.041	0.043	0.002
Std. Dev.	0.026	0.031	0.033
Median	0.041	0.044	0.001
Percentiles			
0.5%	−0.026	−0.037	−0.081
1.0%	−0.020	−0.030	−0.073
2.5%	−0.010	−0.019	−0.061
5.0%	−0.002	−0.008	−0.052
95.0%	0.085	0.094	0.057
97.5%	0.093	0.104	0.068
99.0%	0.102	0.115	0.082
99.5%	0.109	0.122	0.092
Panel B. Ritter's (1991) Data, 1975–1984 (CAR in data = −0.251)			
Mean	0.008	0.010	0.010
Std. Dev.	0.044	0.042	0.069
Median	0.007	0.010	0.008
Percentiles			
0.5%	−0.104	−0.098	−0.161
1.0%	−0.092	−0.088	−0.143
2.5%	−0.077	−0.072	−0.119
5.0%	−0.064	−0.059	−0.099
95.0%	0.078	0.079	0.128
97.5%	0.093	0.092	0.154
99.0%	0.110	0.109	0.182
99.5%	0.122	0.118	0.209

Panel B of Table 6 shows that the −0.251 estimate of Ritter's CAR also overwhelmingly rejects the null hypothesis. The 0.5% percentile values range from −0.098 for the constrained Π three-state model to −0.161 for the unconstrained Π three-state model. Again, this is nowhere close to the −0.251 data estimate. Hence, Ritter's (1991) original sample also strongly rejects the notion that his original IPO underperformance findings are merely due to small sample effects.

In Panel A of Table 7, we compare CHP estimates in data with simulated CHP small sample distributions. Over the full sample, the CHP(NYSE/AMEX) statistic is −0.137 over five years, corresponding to an annualized number of −0.029 per annum in Table 1. Note that the CHP small sample distributions are biased upward, ranging from 0.078 for the two-state model to a very large 0.463 for the constrained Π three-state estimation. The constrained Π has a much larger probability of remaining in the extraordinarily high performing state than the unconstrained Π matrix. This allows for some highly positively skewed draws that result in a strong positive bias for the long-horizon CHP statistics.

Similar to the CAR results in Table 6, the small sample CHP distributions in Table 7 overwhelming reject the null hypothesis of zero expected abnormal IPO performance. The data CHP estimate of −0.137 falls nowhere close to the bottom 0.5% or 1% cutoff of the small sample CHP distributions. In particular, the

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

TABLE 7

CHP Small Sample Distribution

Table 7 reports summary statistics of the small sample distribution of the CHP statistic (equation (5)) by simulating from the posterior distribution of the parameter estimates reported in Table 5 using 10,000 simulations. Panel A corresponds to the sample of IPOs going public from January 1970 to December 1996 with an event horizon of 60 months. Panel B corresponds to Ritter's (1991) original sample, consisting of IPOs going public from 1975 to 1984, with an event horizon of 36 months.

		Three-State Model	
	Two-State Model	Constrained Π	Unconstrained Π
Panel A. Full Sample (CHP in data = −0.137)			
Mean	0.078	0.463	0.080
Std. Dev.	0.064	3.155	0.123
Median	0.070	0.319	0.064
Percentiles			
0.5%	−0.037	0.090	−0.054
1.0%	−0.028	0.106	−0.044
2.5%	−0.014	0.133	−0.028
5.0%	−0.002	0.154	−0.016
95.0%	0.181	0.871	0.211
97.5%	0.214	1.227	0.265
99.0%	0.274	2.120	0.387
99.5%	0.343	3.571	0.520
Panel B. Ritter's (1991) Data, 1975–1984 (CHP in data = −0.127)			
Mean	0.187	0.234	0.124
Std. Dev.	0.224	0.334	1.100
Median	0.169	0.187	0.041
Percentiles			
0.5%	−0.013	−0.006	−0.159
1.0%	0.004	0.008	−0.142
2.5%	0.025	0.030	−0.122
5.0%	0.045	0.050	−0.102
95.0%	0.374	0.514	0.457
97.5%	0.439	0.681	0.687
99.0%	0.548	0.987	1.283
99.5%	0.651	1.499	2.289

most negative 1% cutoff is −0.044 from the unconstrained Π three-state model. Hence, the data again resolutely rejects a small sample explanation of IPO underperformance.

In Panel B, we examine the CHP distributions for the Ritter (1991) sample. In Ritter's data, the CHP statistic is −0.127. While this point estimate lies below the 0.05% tail for the two-state model and the constrained Π three-state model, the left-most tail of the three-state unconstrained Π model does encompass the −0.127 value. However, we still reject that the CHP value is equal to zero at the 5% level using a two-sided test (with the lower 2.5% cutoff equal to −0.122).

In our analysis, we use the NYSE/AMEX benchmark to construct the adjusted IPO returns to estimate the model and to construct the small sample distributions of the IPO performance statistics. We resoundingly reject the null hypothesis that small sample effects could be responsible for the underperformance relative to the NYSE/AMEX benchmark. This benchmark does not produce the largest or most significant point estimates of IPO underperformance in either event time or calendar time from Tables 1 and 2. Hence, other aggregate benchmarks that produce more severe measures of event-time or calendar-time underperformance, such as the NASDAQ and small stock benchmarks, can only

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

result in more overwhelming rejections of a small sample explanation of IPO underperformance.

One possible use of our Markov switching model that we do not examine here is that the model may be able to identify those firms where poor IPO performance is very likely (the underperforming state), or those IPOs whose performance is extremely good (the star performer state). Since these states are persistent, an active investor may be able to infer the probability of each regime for each IPO and form a trading strategy to go long in the most promising IPO firms and short the least promising IPOs. We leave this application of the model to future research.

C. Comparative Statics

In this section, we ask how extreme the distribution of winners must be to fail to reject the null of zero average long-run underperformance in the small sample distribution. Since the structure of the model is able to capture a Peso problem, this is a useful exercise because we can compare the estimates of the model to a distribution of superstar IPO returns (state 1) where we would not be surprised to observe the degree of IPO underperformance present in the actual data. In this exercise, we focus on the original Ritter (1991) CAR statistic.[17] That is, which characteristics of the superstar state are necessary to conclude that the IPO underperformance may be a Peso problem?

Our goal is to determine the value of parameter μ_1, the expected return of the superstar state, where the small sample distribution implied by the model could resemble the observed degree of IPO underperformance in data. To do this, we gradually increase the value of μ_1 from its estimated value in the three-state Markov model. As we change μ_1, we simulate from the point estimates of the parameter values in Table 5, keeping all other parameters the same, except we alter μ_3 so that we maintain the null of zero expected abnormal performance in equation (11). Figure 2 plots the p-value of the CAR point statistic in data, starting from the parameter estimates in Table 5 where the p-value is zero, as a function of μ_1. This is a two-sided p-value and, hence, represents two times the proportion of the small sample distribution lying to the left of the long-horizon CAR point statistic of -0.227. The top (bottom) row of Figure 2 performs this comparative static exercise over the full (Ritter (1991)) sample. The left- (right-) hand column reports the case for the constrained (unconstrained) transition probability three-state model.

Figure 2 shows that in order for a small sample explanation to account for the degree of IPO underperformance in data, the expected return of the winner state has to be truly spectacular. Over the full sample, we must increase μ_1 to over 3.00 or higher per month to produce a p-value higher than 0.05. From the stable probability of state 1 in Table 5, this means that over 2% of all IPOs have to triple their values every month. There are clearly some IPOs that more than triple their value occasionally, such as the top 10 highest monthly returns reported in Panel A of

[17] If we repeat the exercise using the CHP statistic, there is no choice of parameter values for the distribution of superstar IPO returns that can produce small sample distributions where we fail to reject the null at a 95% confidence level.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

FIGURE 2

p-Value of the CAR Statistic as a Function of μ_1

Each plot graphs the *p*-value of the CAR statistic as a function of μ_1, the mean of the IPO return in the outperforming state, in the three-state Markov model. We start with the point marked with a circle, which represents the *p*-value from Table 6, which is zero, corresponding to the estimated value of μ_1 in Table 5. The top row displays the results using the full sample data and the bottom row displays results for the Ritter (1991) sample.

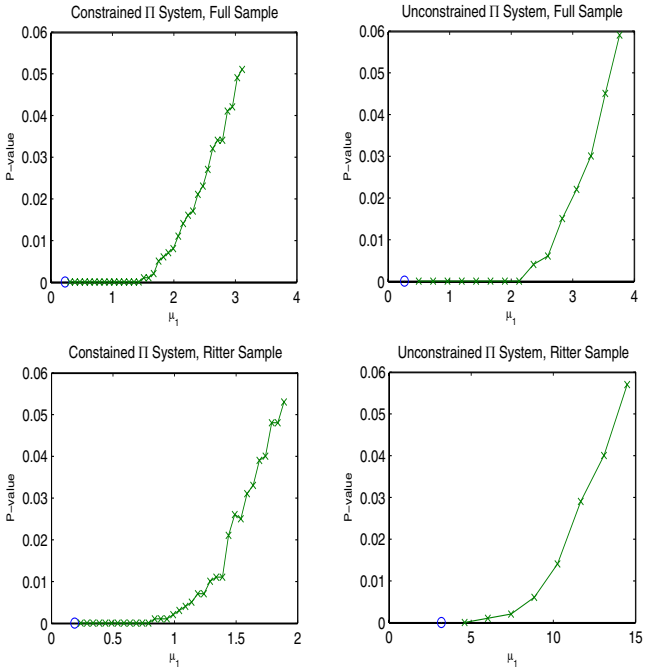


Table 3, but these represent the top 10 among 243,338 total one-month event returns in the full sample of 4,843 firms. Similarly, in the Ritter (1991) sample, the estimate of μ_1 must be approximately 2.00 for the constrained Π estimation and close to 15.00 for the unconstrained Π model. Thus, an absurd number of spectacularly performing firms must be present in the population distribution in order for a small sample explanation to account for the IPO long-run underperformance phenomenon.

The plots in Figure 2 further strengthen the conclusion that a small sample effect is highly unlikely to be driving the IPO long-run underperformance puzzle. Instead, the plots suggest that the low returns of IPO firms over the last three decades are robust and that the IPO long-run underperformance puzzle is not a statistical fluke.

V. Conclusion

The long-run underperformance of IPOs has been an active topic in the IPO literature over the last decade since Ritter (1991). Yet, recent work, most notably by Gompers and Lerner (2003), suggests that the post-1970 long-run IPO underperformance could be simply a statistical fluke. Our study presents new evidence

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

supporting the existence of the IPO underperformance effect. First, we show that IPO underperformance remains robust both in event time and in calendar time. Schultz (2003) fails to find IPO underperformance in calendar time because he considers only a short one-month holding period horizon. Calendar-time IPO underperformance reappears with longer holding period horizons or longer portfolio formation periods.

Second, we present evidence suggesting that IPO underperformance is highly unlikely to be the result of a statistical fluke. We construct a Markov switching model that captures small sample or "Peso problem" effects. At each point in time, IPO returns have the potential to be drawn from superstar states with very high expected returns. A small sample underperformance puzzle may result if we observe too few of these draws ex post, even if the population distribution exhibits no ex ante underperformance. In our estimation, we impose the null of no ex ante IPO underperformance.

We find that the small sample distributions implied by the model for the event-time statistics measuring IPO underperformance do not even come remotely close to encompassing the point statistics in the data. Hence, the null hypothesis that a small sample effect is responsible for IPO underperformance is overwhelmingly rejected. Moreover, the degree of outperformance required for a small sample explanation to hold requires that approximately one in 50 IPOs must at least triple their values every month.

By establishing the robustness of the IPO underperformance puzzle, we lay the groundwork for future research to economically explain this important, statistically robust phenomenon. Some explanations that have been proposed to date include earnings management (Teoh, Welch, and Wong (1998)), constraints on shorting IPOs combined with heterogeneous expectations of investors (Miller (1977)), or behavioral biases (Hirshleifer (2001)), among others.

Appendix

We estimate the model in Section III using Gibbs sampling, adapting the methodology developed by Kim, Nelson, and Startz (1998) and Kim and Nelson (1999). The set of parameters we estimate is

$$\Theta \quad = \quad (\tilde{P}, \tilde{\sigma}^2, \tilde{\mu}),$$

where $\tilde{\sigma}^2 = [\sigma_1^2 \ \sigma_2^2]$, $\tilde{\mu} = [\mu_1 \ \mu_2]$, $\tilde{P} = [P_{11} \ P_{22}]$ for the two-state Markov model, $\tilde{\sigma}^2 = [\sigma_1^2 \ \sigma_2^2 \ \sigma_3^2]$, $\tilde{\mu} = [\mu_1 \ \mu_2 \ \mu_3]$, $\tilde{P} = [P_{11} \ P_{21} \ P_{22} \ P_{32}]$ for the three-state Markov model with restricted transition probabilities, and $\tilde{\sigma}^2 = [\sigma_1^2 \ \sigma_2^2 \ \sigma_3^2]$, $\tilde{\mu} = [\mu_1 \ \mu_2 \ \mu_3]$, $\tilde{P} = [P_{11} \ P_{12} \ P_{21} \ P_{22} \ P_{32} \ P_{33}]$ for the three-state Markov model with unrestricted transition probabilities. For estimation purposes, we parameterize σ_2^2 as $\sigma_2^2 = \sigma_1^2(1+h_2)$ and σ_3^2 as $\sigma_3^2 = \sigma_1^2(1+h_2)(1+h_3)$.

Let T_i denote firm i's surviving period and let n be the number of firms in the sample. We define a Markov state vector of firm i as $\tilde{s}_{T_i} = [s_{i,1} \ s_{i,2} \ \ldots \ s_{i,T_i}]$, for firms $i = 1, 2, \ldots, n$, and a stacked state vector of all firms as $\tilde{s}_T = [s_{1,1} \ldots \ s_{1,T_1} \ s_{2,1} \ldots s_{2,T_2} \ldots \ s_{n,1} \ldots \ s_{n,T_n}]$, where $T = \sum_{i=1}^{n} T_i$. Similarly, we write the vector of returns for the ith IPO as $\tilde{r}_{T_i} = [r_{i,1} \ r_{i,2} \ \ldots \ r_{i,T_i}]$ and denote the collected vector of returns for all IPOs as $\tilde{r}_T = [r_{1,1} \ \ldots r_{1,T_1} \ \ldots \ r_{n,1} \ldots \ r_{n,T_n}]$. Since the states are unobserved to the econometrician, we treat them as parameters and draw them via Gibbs sampling. Hence, the random variables to be drawn are Θ and the stacked Markov states vector $\tilde{s}_T$.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

The Gibbs sampling algorithm iterates successively over the following conditional distributions. Each iteration simulates a drawing from the joint posterior distribution of all the state variables and the model's parameters given the data:

(P1) Generate $\tilde{s}_T$, conditional on $\tilde{\mu}, \tilde{\sigma}^2, \tilde{P}$, and $\tilde{r}_T$,
(P2) Generate $\tilde{\mu}$, conditional on $\tilde{\sigma}^2, \tilde{s}_T$, and $\tilde{r}_T$,
(P3) Generate $\tilde{\sigma}^2$, conditional on $\tilde{\mu}, \tilde{s}_T$, and $\tilde{r}_T$, and
(P4) Generate $\tilde{P}$, conditional on $\tilde{s}_T$.

Note that in (P2), by the null hypothesis that $\tilde{\mu}'\tilde{\pi} = 0$ in equation (11), only μ_1 is generated and the value of the other conditional mean parameter is inferred. In (P4), $\tilde{P}$ only requires knowledge of the states $\tilde{s}_T$. All the conditional distributions (P1) to (P4) are also conditional on the delisting times, which are exogenously given by a geometric distribution. Conditional on observing the delisting time, we can separate the estimation of the geometric distribution and the Markov switching DGP because we assume that the Markov switching process holds conditional on the delisting time, and the delisting process is unaffected by the parameters of the Markov DGP. We describe drawing parameters from each conditional distribution in turn.

(P1) Drawing $\tilde{s}_T$, Conditional on $\tilde{\mu}, \tilde{\sigma}^2, \tilde{P}$, and $\tilde{r}_{T_i}$. We assume independence of returns across firms. Thus, we can generate $\tilde{s}_{T_i}$, conditional on firm i surviving to T_i. We then stack $\tilde{s}_{T_i}$ to construct $\tilde{s}_T$. To generate $\tilde{s}_{T_i}$, we first run the Hamilton (1989) filter to obtain the conditional distributions $g(s_{i,t}|\tilde{r}_{i,t})$ for all t.

Let $\psi_{i,t-1}$ denote the information set up to time $t-1$ for firm i, which represents lagged firm i returns. Given $Pr[s_{i,t-1} = k|\psi_{i,t-1}]$ at the beginning of time t, and using Bayes' rule,

$$\text{(A-1)} \quad Pr[s_{i,t} = j, s_{i,t-1} = k|\psi_{i,t-1}] = Pr[s_{i,t} = j|s_{i,t-1} = k]Pr[s_{i,t-1} = k|\psi_{i,t-1}],$$

where $Pr[s_{i,t} = j|s_{i,t-1}] = Pr[s_t = j|s_{t-1} = k]$ are the transition probabilities $\tilde{P}$, which are firm invariant.

We update the probability $Pr[s_{i,t} = j, s_{i,t-1} = k|\psi_{i,t-1}]$ using

$$\text{(A-2)} \quad Pr[s_{i,t} = j, s_{i,t-1} = k|\psi_{i,t}] = \frac{f(r_{i,t}|s_{i,t} = j, s_{i,t-1} = k, \psi_{i,t-1})Pr[s_{i,t} = j, s_{i,t-1} = k|\psi_{i,t-1}]}{\sum_{s_{i,t}=1}^{K}\sum_{s_{i,t-1}=1}^{K} f(r_{i,t}|s_{i,t} = j, s_{i,t-1} = k, \psi_{i,t-1})Pr[s_{i,t} = j, s_{i,t-1} = k|\psi_{i,t-1}]},$$

where K is the number of states and

$$f(r_{i,t}|s_{i,t} = j, s_{i,t-1} = k, \psi_{i,t-1}) = \frac{1}{\sqrt{2\pi\sigma_{s_t}^2}} \exp\left(-\frac{(r_{i,t} - \mu_{s_t})^2}{2\sigma_{s_t}^2}\right),$$

since the state-dependent parameters $\mu_{s_t} \equiv \mu(s_t)$ and $\sigma_{s_t} \equiv \sigma(s_t)$ are the same across firms.

We can obtain $g(s_{i,T_i}|\tilde{r}_{T_i})$ by summing over the K states using the standard Hamilton (1989) updating recursion,

$$\text{(A-3)} \quad Pr[s_{i,t}|\psi_{i,t}] = \sum_{s_{i,t-1}=1}^{K} Pr[s_{i,t} = j, s_{i,t-1} = k|\psi_{i,t}].$$

The last run of the Hamilton (1989) filter provides us with $g(s_{i,T_i}|\tilde{r}_{T_i})$, from which we can generate s_{T_i}. We then backward generate $s_{i,t}$, conditional on $\tilde{r}_{i,t}$ and $s_{i,t+1}, t = T_i - 1, T_i - 2, \ldots, 1$, using the multi-move Carter and Kohn (1994) algorithm adapted by Kim and Nelson (1999). This uses the following result:

$$\text{(A-4)} \quad g(s_{i,t}|\tilde{r}_{i,t}, s_{i,t+1}) \propto g(s_{i,t+1}|s_{i,t})g(s_{i,t}|\tilde{r}_{i,t}),$$

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

where $g(s_{i,t+1}|s_{i,t}) = g(s_{t+1}|s_t)$ is the transition probability. We calculate $Pr[s_{i,t}=k|s_{i,t+1}, \tilde{r}_{i,t}]$ using

$$Pr[s_{i,t}=k|s_{i,t+1}, \tilde{r}_{i,t}] = \frac{g(s_{t+1}|s_t=1)g(s_{i,t}=k|\tilde{r}_{i,t})}{\sum_{j=1}^{K} g(s_{t+1}|s_t=j)g(s_{i,t}=j|\tilde{r}_{i,t})}. \tag{A-5}$$

The Carter-Kohn (1994) algorithm simulates $s_{i,t}$, $t = 1, 2, \ldots, T_i$ as a block from the joint distribution $g(\tilde{s}_{T_i}|\tilde{\mu}, \tilde{\sigma}^2, \tilde{P}, \tilde{r}_{T_i})$. The algorithm is run for each firm i separately.

Next, we describe draws from (P2) through (P4) for a three-state Markov model with unrestricted transition probabilities as a general case. Our other models are special cases (and involve a reduced number of parameters).

(P2) Drawing $\tilde{\mu}$, Conditional on $\tilde{\sigma}^2, \tilde{s}_T$, and $\tilde{r}_T$. Equation (6) can be rewritten in the following form,

$$r_{it} = \mu_1 \bar{s}_{1it} + \mu_2 \bar{s}_{2it} + \mu_3 \bar{s}_{3it} + \sigma_{s_{it}} \varepsilon_{it}, \tag{A-6}$$

if we let $k = 1, 2, 3$, and define:

$$\bar{s}_{kit} = \begin{cases} 1 & \text{if } s_{it} = k \\ 0 & \text{otherwise.} \end{cases} \tag{A-7}$$

Dividing both sides of equation (A-6) by $\sigma_{s_{it}}$, we obtain

$$r^{\dagger}_{it} = \mu_1 x_{1it} + \mu_2 x_{2it} + \mu_3 x_{3it} + \varepsilon_{it}, \qquad \varepsilon_{it} \sim \text{IID}\, N(0,1), \tag{A-8}$$

where $r^{\dagger}_{it} = r_{it}/\sigma_{s_{it}}$ and $x_{kit} = \bar{s}_{kit}/\sigma_{s_{it}}$. Note that $\sigma_{s_{it}}$ takes only the values of σ_{s_t}, which are common to all firms, but the actual value $\sigma_{s_{it}}$ depends on firm i and time t. Using this notation, we can rewrite equation (A-8) in matrix notation as

$$R^{\dagger} = X\tilde{\mu} + V, \qquad V \sim N(0, I_T), \tag{A-9}$$

where $R^{\dagger}$ is the stacked vector of transformed returns $R^{\dagger} = [r^{\dagger}_{1,1} \ldots r^{\dagger}_{1,T_1} \ldots r^{\dagger}_{n,1} \ldots r^{\dagger}_{n,T_n}]$ for all n firms and $X = [x_{1t}\ x_{2t}\ x_{3t}]$ stacks the values of x_{kit} for the different values of the regime k across the rows and all the time series of firm returns across the columns.

If we assume a normal prior for $\tilde{\mu}$ of $\tilde{\mu}|\tilde{\sigma}^2 \sim N(b_0, B_0)$, the posterior distribution is given by $\tilde{\mu}|\tilde{\sigma}^2, \tilde{s}_T, \tilde{r}_t \sim N(b_1, B_1)$, where $b_1 = (B_0^{-1} + X'X)^{-1}(B_0^{-1}b_0 + X'R^{\dagger})$ and $B_1 = (B_0^{-1} + X'X)^{-1}$. We assign a value of zero to b_0 and B_0^{-1}, which effectively represent a non-informative prior.

(P3) Drawing $\tilde{\sigma}^2$, Conditional on $\tilde{\mu}, \tilde{s}_T$, and $\tilde{r}_T$. By definition of $\bar{s}_{kit}$ in equation (A-7), we can write

$$\sigma^2_{it} = \sigma^2_1 \bar{s}_{1it} + \sigma^2_2 \bar{s}_{2it} + \sigma^2_3 \bar{s}_{3it},$$

for the vector of realizations of conditional variances corresponding to the stacked returns $\tilde{r}_{T_i}$ of firm i, conditional on the stacked regime realizations $\tilde{s}_{T_i}$ for firm i. We can redefine this as

$$\sigma^2_{it} = \sigma^2_1 (1 + \bar{s}_{2it}h_2)(1 + \bar{s}_{3it}h_2)(1 + \bar{s}_{3it}h_3), \tag{A-10}$$

where $\sigma^2_2 = \sigma^2_1(1 + h_2)$ and $\sigma^2_3 = \sigma^2_1(1 + h_2)(1 + h_3)$. Using this specification of σ_1, h_2, and h_3, we first generate σ^2_1, and then generate $h_2 = (1 + h_2)$ and $h_3 = (1 + h_3)$ to indirectly generate σ^2_2 and σ^2_3.

To draw σ^2_1, conditional on h_2 and h_3, we divide both sides of equation (6) by

$$\sqrt{(1 + \bar{s}_{2it}h_2)(1 + \bar{s}_{3it}h_2)(1 + \bar{s}_{3it}h_3)}$$

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

to obtain

$$r_{it}^{\star} = \mu_1 x_{1it}^{\star} + \mu_2 x_{2it}^{\star} + \mu_3 x_{3it}^{\star} + v_{it}^{\star}, \qquad v_{it}^{\star} \sim \text{IID}\, N(0, \sigma_1^2), \tag{A-11}$$

where

$$\begin{aligned} r_{it}^{\star} &= \frac{r_{it}}{\sqrt{(1+\bar{s}_{2it}h_2)(1+\bar{s}_{3it}h_2)(1+\bar{s}_{3it}h_3)}} \\ x_{kit}^{\star} &= \frac{\bar{s}_{kit}}{\sqrt{(1+\bar{s}_{2it}h_2)(1+\bar{s}_{3it}h_2)(1+\bar{s}_{3it}h_3)}} \\ v_{it}^{\star} &= \frac{\varepsilon_{it}}{\sqrt{(1+\bar{s}_{2it}h_2)(1+\bar{s}_{3it}h_2)(1+\bar{s}_{3it}h_3)}}. \end{aligned}$$

We stack each firm's returns in equation (A-11) to write

$$r_t^{\star} = \mu_1 x_{1t}^{\star} + \mu_2 x_{2t}^{\star} + \mu_3 x_{3t}^{\star} + v_t^{\star}.$$

We choose an inverted gamma distribution as a conjugate prior for σ_1^2, so $\sigma_1^2|h_2, h_3, \tilde{\mu} \sim IG((\nu_0/2), (\delta_0/2))$. Kim and Nelson (1999) show that the posterior distribution is also an inverted gamma distribution given by $\sigma_1^2|h_2, h_3, \tilde{\mu}, \tilde{s}_T, \tilde{r}_T \sim IG((\nu_1/2), (\delta_1/2))$, where $\nu_1 = \nu_0 + T$ and $\delta_1 = \delta_0 + \sum_{t=1}^{T}(r_t^{\star} - \mu_1 x_{1t}^{\star} - \mu_2 x_{2t}^{\star} - \mu_3 x_{3t}^{\star})$. We assign a value of zero to ν_0 and δ_0 to represent a non-informative prior.

To generate $\bar{h}_2 = (1 + h_2)$, conditional on σ_1^2 and h_3, we divide both sides of equation (6) by $\sqrt{\sigma_1^2(1+\bar{s}_{3it}h_3)}$ and stack all firm returns to write

$$r_t^{\star\star} = \mu_1 x_{1t}^{\star\star} + \mu_2 x_{2t}^{\star\star} + \mu_3 x_{3t}^{\star\star} + v_t^{\star\star}, \qquad v_t^{\star\star} \sim \text{IID}\, N(0, (1+h_2)), \tag{A-12}$$

where $r_t^{\star\star}$ stacks all transformed firm returns $r_{it}^{\star\star} = r_{it}/\sqrt{\sigma_1^2(1+\bar{s}_{3it}h_3)}$ and x_{kt} stacks all transformed realizations $x_{kit}^{\star\star} = \bar{s}_{kit}/\sqrt{\sigma_1^2(1+\bar{s}_{3it}h_3)}$.

By specifying an inverted gamma distribution for the prior as $\underline{h}_2 IG((\nu_2/2), (\delta_2/2))$, the conditional posterior distribution of $\bar{h}_2 = 1 + h_2$ is given by $\bar{h}_2|\sigma_1^2, h_3, \tilde{\mu}, \tilde{s}_T, \tilde{r}_T \sim IG((\nu_3/2), (\delta_3/2))$, where $\nu_3 = \nu_2 + T_2$, and $\delta_3 = \delta_2 + \sum_{N_2}(r_t^{\star\star} - \mu_1 x_{1t}^{\star\star} - \mu_2 x_{2t}^{\star\star} - \mu_3 x_{3t}^{\star\star})$. The set N_2 represents the realizations of states 2 or 3 across firms, $N_2 = \{t : s_{it} = 2 \text{ or } 3\}$ and T_2 is the total number of simulated states 2 or 3 across firms. We assign a value of zero to ν_2 and δ_2 for the prior, which effectively represents agnostic beliefs.

We generate $\bar{h}_3 = 1 + h_3$ in a similar fashion to generating $\bar{h}_2 = (1 + h_2)$.

(P4) Drawing $\tilde{P} = [P_{11}\ P_{12}\ P_{21}\ P_{22}\ P_{32}\ P_{33}]$, Conditional on $\tilde{s}_T$. Conditional on the states $\tilde{s}_T$, the transition probabilities are independent of $\tilde{r}_T$ and the other parameters of the model. Since the transitions of each firm are independent of the transition of each other firm, we can use the combined transitions of all firms to estimate $\tilde{P}$. We introduce the notation $n_{jk}, j, k = 1, 2, 3$ to represent the total number of transitions from state $s_{t-1} = j$ to $s_t = k$, $t = 2, 3, \ldots, T$, where we consider the total transitions of all firms. Define $P_{jj} = Pr[s_t \neq j|s_{t-1} = j], j = 1, 2, 3$, and $P_{jk} = Pr[s_t = k|s_{t-1} = j, s_t \neq j]$ for $k \neq j$. Then, $P_{jk} = Pr[s_t = k|s_{t-1} = j] = P_{jk} \times (1 - P_{jj})$ for $k \neq j$. Similarly, define n_{jj} to be the number of transitions from state $s_{t-1} = j$ to $s_t \neq j$.

By taking the Beta family of distributions as conjugate priors, Kim and Nelson (1999) show that the posterior distributions of P_{jj} are given by Beta distributions $P_{jj}|\tilde{s}_T \sim \text{Beta}(u_{jj} + n_{jj}, \bar{u}_{jj} + \bar{n}_{jj})$, where u_{jj} and $\bar{u}_{jj}$ are the known hyperparameters of the priors. We assign a value of zero to u_{jj} and $\bar{u}_{jj}$.

Drawing the other off-diagonal elements in the Π transition probability matrix is a straightforward generalization of the method used to draw the diagonal transition probabilities P_{jj}. For example, given P_{11}, P_{12} can be computed by $P_{12} = P_{12} \times (1 - P_{11})$, where P_{12} is drawn from the posterior Beta distribution $P_{12}|\tilde{s}_T \sim \text{Beta}(u_{12} + n_{12}, u_{13} + n_{13})$, where u_{12} and u_{13} are the known hyperparameters of the prior Beta distribution. Similar to the draws for P_{jj}, we assign a value of zero to u_{12} and u_{13}. Finally, conditional on P_{11} and P_{12}, the adding up constraint implies $P_{13} = 1 - P_{11} - P_{12}$. The other off-diagonal elements in Π are similarly drawn.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

References

Albert, J. H., and S. Chib. "Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts." *Journal of Business and Economic Statistics*, 11 (1993), 1–15.

Ang, A., and G. Bekaert. "Regime Switching in Interest Rates." *Journal of Business and Economic Statistics*, 20 (2002), 163–182.

Barber, B. M., and J. D. Lyon. "Detecting Long-Run Abnormal Stock Returns: The Empirical Power and Specification of Test Statistics." *Journal of Financial Economics*, 43 (1997), 341–372.

Bekaert, G.; C. Erb; C. R. Harvey; and T. Viskanta. "Distributional Characteristics of Emerging Market Returns and Asset Allocation." *Journal of Portfolio Management*, 24 (1998), 102–116.

Bekaert, G.; R. J. Hodrick; and D. A. Marshall. "Peso Problem Anomalies for Term Structure Anomalies." *Journal of Monetary Economics*, 48 (2001), 241–270.

Brav, A. "Inference in Long-Horizon Event Studies: A Bayesian Approach with Applications to Initial Public Offerings." *Journal of Finance*, 55 (2000), 1979–2016.

Brav, A.; C. Geczy; and P. A. Gompers. "Is the Abnormal Return Following Equity Issuances Anomalous?" *Journal of Financial Economics*, 56 (2000), 209–249.

Brav, A., and P. A. Gompers. "Myth or Reality? The Long-Run Underperformance of Initial Public Offerings: Evidence from Venture and Nonventure Capital-Backed Companies." *Journal of Finance*, 52 (1997), 1791–1821.

Carter, C. K., and R. Kohn. "On Gibbs Sampling for State Space Models." *Biometrika*, 81 (1994), 541–553.

Cecchetti, S. G.; P. Lam; and N. C. Mark. "The Equity Premium and the Risk-Free Rate: Matching the Moments." *Journal of Monetary Economics*, 31 (1993), 21–46.

Conrad, J., and G. Kaul. "Long-Term Market Overreaction or Biases in Computed Returns?" *Journal of Finance*, 48 (1993), 39–63.

Dahlquist, M., and F. de Jong. "Pseudo Market Timing: Fact or Fiction?" Working Paper, Stockholm Institute for Financial Research (2003).

Davies, R. B. "Hypothesis Testing when a Nuisance Parameter is Present only under the Alternative." *Biometrika*, 74 (1987), 1, 33–43.

Eckbo, B. E.; R. W. Masulis; and O. Norli. "Seasoned Public Offerings: Resolution of the 'New Issues Puzzle'." *Journal of Financial Economics*, 56 (2000), 251–291.

Eckbo, B. E., and O. Norli. "Leverage, Liquidity and Long-Run IPO Returns." *Journal of Corporate Finance*, 11 (2005), 1–35.

Evans, M. D. D. "Peso Problems: Their Theoretical and Empirical Implications." In *Handbook of Statistics: Statistical Methods in Finance*, G. S. Maddala and C. R. Rao, eds. Amsterdam, The Netherlands: North Holland (1996).

Evans, M. D. D., and K. K. Lewis. "Do Long-Term Swings in the Dollar Affect Estimates of the Risk Premia?" *Review of Financial Studies*, 8 (1995), 709–742.

Fama, E., and K. French. "Common Risk Factors in the Returns on Stock and Bonds." *Journal of Financial Economics*, 33 (1993), 3–56.

Gompers, P. A., and J. Lerner. "The Really Long-Run Performance of Initial Public Offerings: The Pre-Nasdaq Evidence." *Journal of Finance*, 58 (2003), 1355–1392.

Hamilton, J. D. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle." *Econometrica*, 57 (1989), 357–384.

Hamilton, J. D., and G. Lin. "Stock Market Volatility and the Business Cycle." *Journal of Applied Econometrics*, 11 (1996), 573–593.

Hamilton, J. D., and R. Susmel. "Autoregressive Conditional Heteroskedasticity and Changes in Regime." *Journal of Econometrics*, 64 (1994), 307–333.

Hirshleifer, D. "Investor Psychology and Asset Pricing." *Journal of Finance*, 56 (2001), 1533–1597.

Kass, R. E., and A. E. Raftery. "Bayes Factors." *Journal of the American Statistical Association*, 90 (1995), 773–795.

Kim, C. J., and C. R. Nelson. *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. Cambridge, MA: MIT Press (1999).

Kim, C. J.; C. R. Nelson; and R. Startz. "Testing for Mean Reversion in Heteroskedastic Data Based on Gibbs-Sampling-Augmented Randomization." *Journal of Empirical Finance*, 5 (1998), 131–154.

Kothari, S. P., and J. B. Warner. "Measuring Long-Horizon Security Price Performance." *Journal of Financial Economics*, 43 (1997), 301–339.

Loughran, T., and J. R. Ritter. "The New Issues Puzzle." *Journal of Finance*, 50 (1995), 23–51.

Loughran, T., and J. R. Ritter. "Uniformly Least Powerful Tests of Market Efficiency." *Journal of Financial Economics*, 55 (2000), 361–389.

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press

Miller, M. H. "Debt and Taxes." *Journal of Finance*, 32 (1977), 261–275.
Newey, W. K., and K. D. West. "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55 (1987), 703–708.
Perez-Quiros, G., and A. Timmermann. "Firm Size and Cyclical Variations in Stock Returns." *Journal of Finance*, 55 (2000), 1229–1262.
Ramchand, L., and R. Susmel. "Volatility and Cross Correlation across Major Stock Markets." *Journal of Empirical Finance*, 5 (1998), 397–416.
Rietz, T. A. "The Equity Risk Premium: A Solution." *Journal of Monetary Economics*, 22 (1988), 117–131.
Ritter, J. R. "The Long-Run Performance of Initial Public Offerings." *Journal of Finance*, 46 (1991), 3–27.
Ritter, J. R., and I. Welch. "A Review of IPO Activity, Pricing and Allocations." *Journal of Finance*, 57 (2002), 1795–1828.
Schultz, P. "Pseudo Market Timing and the Long-Run Underperformance of IPOs." *Journal of Finance*, 58 (2003), 483–517.
Shumway, T. "The Delisting Bias in CRSP Data." *Journal of Finance*, 52 (1997), 327–340.
Shumway, T., and V. A. Warther. "The Delisting Bias in CRSP's Nasdaq Data and Its Implications for the Size Effect." *Journal of Finance*, 54 (1999), 2361–2379.
Teoh, S.; I. Welch; and T. Wong. "Earnings Management and the Long-Run Market Performance of Initial Public Offerings." *Journal of Finance*, 53 (1998), 1935–1974.
Timmermann, A. "Moments of Markov Switching Models." *Journal of Econometrics*, 96 (2000), 75–111.
Turner, C. M.; R. Startz; and C. R. Nelson. "A Markov Model of Herteroskedasticity, Risk, and Learning in the Stock Market." *Journal of Financial Economics*, 25 (1989), 3–22.
Viswanathan, S., and B. Wei. "Endogenous Events and Long Run Returns." Working Paper, Duke University (2004).

https://doi.org/10.1017/S0022109000004105 Published online by Cambridge University Press