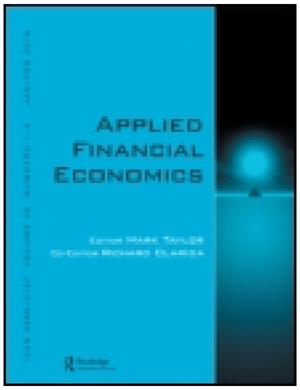

Applied Financial Economics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rafe20

Modelling heavy tails and skewness in film returns

W. D. Walls [a]

[a] Department of Economics, University of Calgary, Calgary, Alberta, Canada, T2N 1N4

Published online: 20 Aug 2006.

To cite this article: W. D. Walls (2005) Modelling heavy tails and skewness in film returns, Applied Financial Economics, 15:17, 1181-1188, DOI: 10.1080/0960310050391040

To link to this article: http://dx.doi.org/10.1080/0960310050391040

Applied Financial Economics, 2005, **15**, 1181–1188

Modelling heavy tails and skewness in film returns

W. D. Walls

Department of Economics, University of Calgary, Calgary, Alberta, Canada T2N 1N4
E-mail: wdwalls@ucalgary.ca

The average of box-office revenue is dominated by extreme outcomes, with most films earning little and most revenues flowing to a few blockbusters. In this paper the skewness and heavy tails of film returns are formally modelled using skew-Normal and skew-*t* distributions. Logarithmic skew-Normal and skew-*t* models of the distribution of box-office revenue are fitted conditional on star actors and directors, budget, release pattern, genre, rating, and year of release. The estimates show significantly more skewness and heavier tails than the log-Normal distribution. It is also found that a wide theatrical release has a much smaller impact on box-office revenue when heavy tails and skewness are explicitly modelled.

I. Introduction

Few products have more commercially uncertain success rates than motion pictures. Each film – whether an original screenplay, an adaptation of a novel, a remake of an earlier film, a sequel, a prequel, or a knockoff of a recently successful film – is a unique combination of creative and technical inputs and there are no known formulas for financial success: high-budget films featuring star actors sometimes fail miserably (e.g., *Waterworld*) and moderate-budget films featuring non-marquee actors sometimes become box-office hits (e.g., *My Big Fat Greek Wedding*). The few successful films earn a disproportionately large share of total box-office revenues as their success in previous weeks seems to propagate further success.

Part of the difficulty in modelling film returns is similar to the problem of modelling asset returns in general: events that are distant from the sample mean occur with a frequency that is improbably large to have been drawn from a Normal distribution. Simply said, the tails of the empirical distribution of returns are 'heavier' – contain more probability mass – than the tails of a Normal (or even a log-Normal) distribution. In addition to the higher-than-Normal probability of extreme events, the distribution of returns is asymmetric with an upper tail that is longer and heavier than the lower tail.

Several recent papers by De Vany and Walls (1999, 2002, 2004) have documented the extreme uncertainty associated with motion-picture success. In their papers the unconditional distribution of movie returns is modelled using the Pareto distribution and most recently the Lévy-stable distribution. The stable distribution – initially proposed as a model of asset returns by Mandelbrot (1963) and Fama (1963) – captures the skewness and heavy tails of movie returns. In addition to providing a good fit to the movie data, the central limit theorem can be

Applied Financial Economics ISSN 0960–3107 print/ISSN 1466–4305 online © 2005 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/0960310050391040

appealled to in applying the the stable distribution. Advances in statistical computation have made it possible to estimate the unconditional stable density as was done by De Vany and Walls (2004). However, estimation of the stable distribution conditional on a set of explanatory variables – the type of regression framework used by applied researchers – poses an overwhelming computational problem. In this paper a sensible alternative way to model skewness and heavy tails in applied work is proposed and applied.

Recent advances in the statistical theory of non-symmetric density functions have resulted in the skew-Normal and skew-*t* distributions (and their logarithmic versions). These distributions augment the well-known Normal and student-*t* distributions by adding a skewness parameter. It is also feasible to estimate a regression model with skew-normal or skew-*t* random disturbances using standard maximum likelihood techniques. The skew-*t* regression model is particularly appealing in the present application where the data are characterized by heavy tails and skewness, and where the interest is in making statistical inference on the regression coefficients. Unlike some other distributions – such as those in the Lévy-stable family – the skew-*t* model can not appeal to the central limit theorem; in this respect it is an *ad hoc* statistical model. However, the skew-*t* model is intuitively appealing in that it extends the Normal distribution by permitting tails that are heavy and asymmetric. The skew-*t* model can be estimated using standard statistical software, making it a tool within the reach of all applied researchers.

In the following section the skew-Normal and skew-*t* distributions as statistical models of skewness and heavy tails are discussed. In Section III the conditional distribution of film returns is estimated. The results are compared to those obtained by applying standard statistical techniques to the data set and to the results obtained by other researchers. It is concluded in Section IV that the skew-*t* model is a practical tool for use in applications and that is a good approximation to the computationally impractical asymmetric Lévy-stable regression model.

II. Statistical Models of Heavy Tails and Skewness

Several statistical models have been proposed to account for the non-Normal distribution of data from the physical and social sciences. The stable Paretian model was proposed by Mandelbrot (1963) to account for the heavy tails of financial data; other authors have suggested the student-*t* distribution with the appropriate degrees of freedom as a natural alternative to the Normal distribution. The stable Paretian model can account for skewness and heavy tails, but is computationally intractable to condition on a vector of covariates in a regression-like framework.[1] The multivariate student-*t* distribution permits regression analysis of film returns allowing for heavy tails, but it does not allow for asymmetry. Several recent empirical papers in finance have employed various statistical models to capture the skewness and heavy tails in the distribution of financial returns using non-standard distributions that make it difficult for applied researchers to condition the distribution on a vector of explanatory variables.[2] In this section, is proposed the use of the logarithmic versions of the skew-Normal and skew-*t* distributions as models of film returns that explicitly account for skewness, and skewness and heavy tails, respectively, in a multivariate regression-like framework of analysis that is familiar to financial economists and other social scientists.

The skew-normal distribution

Azzalini (1985, 1986) defines a continuous random variable Z to have a skew-Normal distribution, denoted $SN(0, 1, \alpha)$, if it has density function

$$2\phi(z)\Phi(\alpha z) \tag{1}$$

where ϕ and Φ denote the density and distribution functions, respectively, of a standard Normal $N(0, 1)$ variate.[3] The skew-Normal distribution is essentially a Normal distribution that has been augmented by the addition of a shape parameter $\alpha \in (-\infty, +\infty)$ that quantifies the skewness of the distribution; when $\alpha = 0$, the skew-Normal distribution simplifies

[1] Walls (1997) estimates the unconditional distribution of movie revenues using Pareto and parabolic Pareto distributions; Hand (2001) and Maddison (2004) also perform similar calculations on different data sets. De Vany and Walls (2004) model the unconditional distribution of motion-picture profit using the stable Paretian model. Walls (2005) estimates the *conditional* distribution of movie revenues using a *symmetric* stable Paretian regression model that fixes the skewness parameter at zero. No prior work has modelled the conditional distribution of movie returns without imposing symmetry.
[2] For example, Harris and Kucukozmen (2001) use the exponential generalized beta and skew generalized *t* distributions, and Brannas and Nordman (2003) apply the log-generalized gamma and Pearson type IV specifications.
[3] The development in this and the following subsection closely follows the simplified exposition of Azzalini and Kotz (2002).

to the standard Normal distribution. In the empirical application the distribution of the variable $y = \xi + \omega z$ is analysed where ξ is the location parameter and $\omega > 0$ is the scale parameter. Azzalini and Dalla Valle (1996) and Azzalini and Capitanio (1999) formally derive the statistical properties of the multivariate skew-Normal distribution.

The skew-t distribution

The skew-Normal distribution can naturally be extended to obtain the skew-*t* distribution to account for heavier-than-Normal tails. The connection between the skew-Normal and skew-*t* distributions is analogous to that between the Normal and student-*t* distributions.[4] The standard skew-*t* distribution is obtained by considering the transformed variable

$$\tilde{z} = z/(v/df)^{1/2} \tag{2}$$

where v is distributed χ^2 with df degrees of freedom and is statistically independent of z. In the empirical application the distribution of the variable $\tilde{y} = \xi + \omega\tilde{z}$ is analysed. Azzalini and Capitanio (2003) provide a formal detailed treatment of the properties of the skew-*t* distribution, including its multivariate extension.

The skew-Normal and skew-*t* distributions can be fit using their log-transformed versions. These are referred to as log-skew-Normal and log-skew-*t* distributions and they are related to the distributions described above in the same way that the log-Normal distribution corresponds to the Normal distribution. In the empirical analysis that follows, unconditional log-skew-Normal and log-skew-*t* distributions are fitted to motion-picture revenue data.

Skew-normal and skew-t regression analysis

Using the skew-Normal and skew-*t* distributions, the distribution of movie revenues are fitted *conditional* on a vector of movie attributes in a regression-like framework. The distribution of motion-picture revenue is quantified conditional on budget, opening screens, whether or not the movie is a sequel (or prequel), whether or not the director or the actors are stars, the genre, the rating category, and the year of release in the form of a logarithmic linear regression

$$\begin{aligned} \log \text{Revenue}_i = \beta_0 + \beta_1 \log \text{Budget}_i \\ + \beta_2 \log \text{Opening Screens}_i + \beta_3 \text{ Sequel}_i \\ + \beta_4 \text{ Star}_i + \gamma_1' \text{ Genre}_i + \gamma_2' \text{ Rating}_i \\ + \gamma_3' \text{ Year}_i + \mu_i \end{aligned} \tag{3}$$

where i indexes individual movies; Star and Sequel are indicator variables equal to unity when a movie contains a star or is a sequel, respectively, and zero otherwise; the γs are column vectors of coefficients conformable to the sets of indicator variables denoting particular genres, ratings, and release years; and the random disturbance μ_i follows a Normal, skew-Normal, or skew-*t* distribution depending on the model being estimated. This basic linear regression equation for cinema box-office revenue has been employed by several previous researchers.[5] Estimation of the log-linear specification allows the parameters on Budget and Opening Screens to be interpreted as elasticities. Also, since film revenues must be positive numbers, modelling their logarithm is the natural transformation to make. However, as will be shown below, the logarithmic transformation of revenues alone – i.e., the log-Normal distribution – does not capture the skewness and heavy tails of box-office revenue.

III. Data Description and Estimation Results

The data are drawn from the population of movies released domestically from 1985 to 1996.[6] The data were extracted from ACNielson EDI, Inc.'s historical database. The EDI data are compiled from the North American distributor-reported box-office figures and are widely regarded as the standard industry source for published information on motion picture theatrical revenues. The EDI figures are cited and republished by many major industry publications including *Daily Variety* and *Weekly Variety*. From the EDI database, all films for which data the variables of interest were available were selected. The resulting sample of complete cases contained 1989 movies. A detailed description and cross-tabulation of the entire sample of EDI data is contained in De Vany and Walls (1999).

[4] Any good mathematical statistics book will have a thorough treatment of the log-Normal distribution. See, for example, the discussion in Hogg and Craig (1978).
[5] A partial listing of papers would include Smith and Smith (1986), Prag and Cassavant (1994), Litman and Ahn (1998) and Ravid (1999).
[6] In standard industry parlance 'domestic' refers to the USA and Canada.

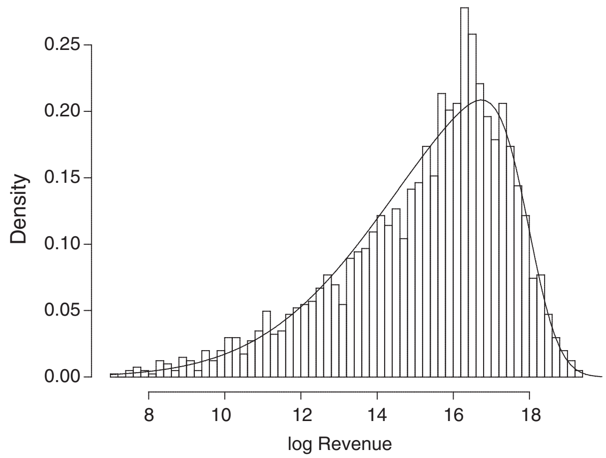

Fig. 1. Skew-*t* fit to the data

In the sample, box-office revenue ranged from a low of 1304 to a high of 245 million, with a mean of 17.2 million. The average budget for these films was 11.9 million with a low of 4801 and a high of 114 million. The largest-grossing movies in the sample include *Batman, The Lion King, Home Alone*, *Forrest Gump*, and *Jurassic Park*. The largest-budget films in the sample include *Space Jam*, *Eraser*, *True Lies*, *Terminator 2: Judgment Day*, and *Waterworld*. Release patterns in the sample also vary widely, with some movies opening on a single theatre screen and the widest release opening on 3012 screens, with the average being 844 screens.

First the unconditional log-skew-Normal and log-skew-*t* distributions were fitted to motion-picture revenue data.[7] The log-skew-Normal and log-skew-*t* distributions gave similar fits to the revenue data, with their respective location parameters being 17.997 and 17.967, and their respective skewness parameters being -5.483 and -5.403.[8] The estimated value of the degrees-of-freedom of the skew-*t* distribution was 108.167 with an estimated standard error of 279.919. With such a large value of degrees-of-freedom, there is little difference between the two fitted distributions. The fitted skew-*t* distribution is plotted over a histogram of log revenue in Fig. 1. A plot of the fitted skew-Normal distribution is nearly identical.

The analysis proceeds by estimating the conditional distribution of film revenues as set out in Equation 3. Estimates of the skew-Normal and skew-*t* regression models are displayed in the columns of Table 1. Of primary methodological interest are the coefficients on skewness (α) and tail thickness (df). In both models it is found that the skewness coefficient differs statistically from zero at a marginal significance level much less than 0.01. Further, from the skew-*t* model it is found that the tails of the distribution are substantially heavier than Normal, with the estimated degrees-of-freedom parameter equal to approximately six with an estimated standard error of about one. The parameter estimates in Table 1 clearly indicate that the Normal model and the skew-Normal model are both statistically rejected in favour of the more general skew-*t* model.

Now some simple diagnostic plots are examined to check the appropriateness of the skew-*t* distribution. In Fig. 2 the fit of the theoretical distribution is visually checked by examining the probability–probability (PP) plot. In the PP-plots the empirical cumulative distribution function is plotted against the theoretical Normal cumulative distribution function in the left panel and against the theoretical skew-*t* cumulative distribution function in the right panel for log revenue. If the theoretical cumulative distribution approximates the observed distribution well, then all points in the PP-plots should fall onto the diagonal line. It is seen that there is noticeable divergence from the diagonal line for the Normal PP-plot. However, the skew-*t* distribution approximates the empirical distribution quite well as evidenced by the diagonal plot in the right panel of the figure.

Now the substantive coefficients reported in Table 1 are discussed. The coefficient on log Budget represents the elasticity of box-office revenue with respect to the film's production budget. In the skew-*t* model the estimate is 0.74 with a standard error of about 0.047. The corresponding elasticity estimate for the skew-Normal model was about 0.65, a substantially smaller elasticity that is nearly two standard errors less than the estimate from the skew-*t* model.

The coefficient on Star in the skew-*t* regression is about 0.73 as compared to 0.84 for the skew-Normal model; in each model the estimated standard error is less than 0.05 so the estimates differ by more than two standard errors. But to give an economic interpretation to coefficients on dummy variables in logarithmic regression, one must be subtracted from the exponentiated coefficient on the dummy variable (Halvorsen and Palmquist, 1980). This results in

[7] The statistical models were estimated in the R language (Ihaka and Gentleman, 1996) using the 'sn' library developed by Adelchi Azzalini.
[8] The differences between the skew-Normal and skew-*t* estimates of location and skewness are small relative to their estimated standard errors of 0.051 and 0.038, respectively.

Table 1. Skew-Normal and skew-*t* regression estimates

	Skew-Normal model		Skew-*t* model	
Variable	Coefficient	Std error	Coefficient	Std error
Constant	5.8713	0.8656	3.9201	0.8731
log Budget	0.6473	0.0427	0.7435	0.0472
Star	0.8428	0.0472	0.7296	0.0413
Sequel	0.4146	0.5385	0.2759	0.4654
log Opening Screens	0.2725	0.0193	0.2728	0.0204
Genre				
Action	−0.1802	0.5362	−0.2899	0.4642
Adventure	−0.1038	0.5662	−0.2110	0.4971
Animated	0.5007	0.6147	0.3025	0.5530
Black comedy	−0.2055	0.5985	−0.2518	0.5312
Comedy	0.2072	0.5307	0.0886	0.4596
Drama	0.2158	0.5284	0.0992	0.4566
Fantasy	0.3212	0.5875	0.1153	0.5159
Horror	0.1716	0.5546	0.1052	0.4795
Musical	0.0493	0.5881	−0.1419	0.5231
Romantic comedy	0.2797	0.5414	0.1177	0.4716
Sci-Fi	−0.0373	0.5776	−0.1486	0.5058
Suspense	−0.3101	0.5435	−0.4109	0.4724
Western	0.2224	0.6577	−0.0728	0.5845
Rating category				
PG	−0.3961	0.2627	−0.2378	0.2524
PG-13	−0.5852	0.2649	−0.4411	0.2524
R	−0.5570	0.2636	−0.4154	0.2513
Year of release				
1986	−0.3300	0.2160	−0.3525	0.1975
1987	−0.3407	0.2132	−0.3395	0.1969
1988	−0.5998	0.2092	−0.5299	0.1931
1989	−1.0960	0.2038	−1.0189	0.1888
1990	−0.7099	0.2094	−0.6895	0.1917
1991	−0.4959	0.2089	−0.4761	0.1911
1992	−0.6121	0.2141	−0.6214	0.1966
1993	−0.3961	0.2192	−0.4641	0.2027
1994	−0.4519	0.2212	−0.5326	0.2033
1995	−0.5176	0.2169	−0.5421	0.1986
1996	−0.6401	0.2129	−0.7070	0.1962
α (skewness)	−1.9198	0.1504	−1.2131	0.1503
df (tail weight)			6.0525	0.9740
log-Likelihood	−3660.041		−3636.552	
Observations	1989		1989	

Stars increasing revenues by a factor of $\exp(0.73) - 1 = 1.075$ or 7.5% as compared to the factor of $\exp(0.84) - 1 = 1.316$ or 31.6% in the skew-Normal model. In practical term, this enormous difference in 'star power' would erroneously be attributed to stars by ignoring the heavy tails of the skew-*t* model. The coefficient on log Opening Screens in both models is estimated to be about 0.27 with a standard error of about 0.02. This coefficient is the elasticity of total box-office revenue with respect to opening screens. While each coefficient differs statistically from zero, the estimates from the two regression model are nearly identical in comparison to their standard errors.

In both the skew-Normal and skew-*t* models, the coefficient on Sequel is positive but not statistically different from zero. The coefficients on individual genres and years of release generally do not differ statistically from zero. However the coefficients on the R and PG-13 ratings indicate that these categories of films have statistically lower returns than G rated movies, a result that confirms the earlier finding of De Vany and Walls (2002) based on the Pareto distribution.

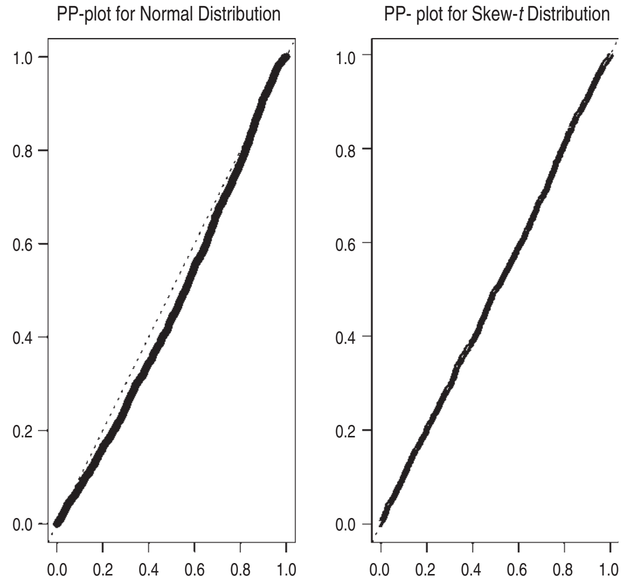

Fig. 2. Probability plot of skew-*t* regression residuals

For the purpose of direct comparison, estimates obtained using least-squares regression and minimum absolute deviation (MAD) regression are reported in Table 2. The least-squares regression corresponds to maximum likelihood estimation of a standard log-Normal regression and the MAD regression corresponds to modelling the median, a technique often used by practitioners when data have influential outlying observations.[9]

The coefficient estimate on log Budget is about 0.71 in the least-squares regression and 0.73 in the median regression; both of these estimates lie in the interval between the skew-Normal and skew-*t* estimates. The coefficient on Star in the least-squares model is 0.84, approximately the same as the estimate from the skew-Normal model. But the estimate on Star from the median regression is about 0.64, substantially smaller than any of the other estimates. The coefficient on log Opening Screens in the least-squares regression and the median regression is about 0.35, which is substantially larger and statistically different than the estimate obtained form the skew-Normal and skew-*t* models.

The estimates differ substantially from those obtained by some other researchers, most notably Litman and Ahn (1998). Their estimates imply an elasticity of box-office gross with respect to budget of 0.23 which is substantially – and statistically – smaller than the present estimate of 0.74.[10] Their estimates also imply an elasticity of box-office gross with respect to opening screens of 0.65 as compared to the estimate of 0.27.[11]

Litman and Ahn (1998) report that Budget and Screens are statistically different from zero in their model; however, they can not reject the hypothesis that the coefficients on Star and the Rating variables are zero. It is difficult to speculate on the reasons for their particular results, though it could be related to their specification search[12] or their truncated sample of data that includes only 241 films.[13]

IV. Conclusions

The motion-picture market has a winner-take-all property where a small proportion of successful films earns the majority of box-office revenue. The average return across films is dominated by extreme events, namely those few films that populate the long upper tail of the distribution of returns. To be useful in practice a statistical model of film returns should capture (1) the asymmetry implied by the winner-take-all property, (2) the heavy tails implied by the importance of extreme events, and (3) allow returns to be conditioned on a vector of explanatory variables.

Recent advances in the statistical theory of non-symmetric density functions make it feasible to estimate a regression model with skew-Normal or skew-*t* random disturbances using standard maximum likelihood techniques. The skew-*t* regression model is particularly appealing in economics and finance where the data are characterized by heavy tails and skewness, and where interest is in analysing conditional distributions. However, the skew-*t* model is intuitively appealing in that it extends the Normal distribution by permitting tails that are heavy and asymmetric. Also, the skew-*t*

[9] The MAD estimator is also a maximum likelihood estimator when the disturbances follow a two-tailed exponential distribution. See Judge *et al.* (1985, pp. 836–37) and the references cited therein for further discussion.
[10] They do not explicitly report elasticities in their paper. The elasticity has been calculated using their estimate regression coefficient on budget of 0.38254, their reported average budget of 31.38 million and average box-office revenue of 51.24.
[11] Again, Litman and Ahn (1998) do not explicitly report elasticity estimates in their paper. Point elasticity has been calculated using their estimated regression coefficient of 0.01982, and their reported average screens of 1669.9 and average box-office gross of 31.38 million.
[12] Litman and Ahn (1998) are more honest than most empirical researchers in stating that their regression results, 'represent the final 'best fit' after initial screening of different groups of independent variables through correlation analysis' (p. 188).
[13] Their sample of data includes only the films listed in *Variety's Top-100* chart for the years 1993–1995 inclusive.

Table 2. Least-squares and median regression estimates

Variable	Least-squares		Median regression	
	Coefficient	Std error	Coefficient	Std error
Constant	3.5093	0.9139	3.2511	0.8347
log Budget	0.7086	0.0468	0.7291	0.0437
Star	0.8441	0.1039	0.6367	0.0967
Sequel	−0.0150	0.5701	−0.1461	0.5014
log Opening Screens	0.3501	0.0176	0.3565	0.0163
Genre				
Action	−0.6770	0.5668	−0.7407	0.4987
Adventure	−0.6832	0.5961	−0.7539	0.5275
Animated	−0.1592	0.6418	−0.3235	0.5700
Black comedy	−0.4392	0.6339	−0.7055	0.5635
Comedy	−0.3057	0.5598	−0.4786	0.4917
Drama	−0.2124	0.5589	−0.3336	0.4913
Fantasy	−0.0562	0.6215	−0.3802	0.5513
Horror	−0.2196	0.5885	−0.2926	0.5199
Musical	−0.4183	0.6196	−0.7681	0.5495
Romantic comedy	−0.2814	0.5705	−0.4463	0.5022
Sci-Fi	−0.5799	0.6097	−0.6769	0.5402
Suspense	−0.8769	0.5722	−0.8932	0.5043
Western	−0.2838	0.6879	−0.6292	0.6107
Rating category				
PG	−0.4447	0.2722	−0.1676	0.2513
PG-13	−0.6273	0.2736	−0.3698	0.2523
R	−0.5583	0.2725	−0.3028	0.2512
Year of release				
1986	−0.3619	0.2282	−0.5156	0.2129
1987	−0.3766	0.2253	−0.4835	0.2098
1988	−0.6852	0.2215	−0.7148	0.2062
1989	−1.2642	0.2151	−1.1188	0.2003
1990	−0.7818	0.2220	−0.7957	0.2066
1991	−0.4818	0.2212	−0.5822	0.2058
1992	−0.6480	0.2263	−0.7968	0.2105
1993	−0.4613	0.2309	−0.6920	0.2151
1994	−0.5086	0.2333	−0.7641	0.2169
1995	−0.6122	0.2286	−0.6588	0.2131
1996	−0.6903	0.2245	−0.9380	0.2093
R^2	0.5106		0.3188	
Observations	1989		1989	

model is computationally straightforward and estimable using standard statistical software that is freely available.[14] In this respect, the skew-t model appears to be a practical approximation to the computationally overwhelming asymmetric Lévy-stable regression model.

In the empirical analysis logarithmic skew-Normal and skew-t models of the distribution of box-office revenue have been fitted conditional on star actors and directors, budget, release pattern, genre, rating, and year of release. Statistical evidence is found of skewness and heavy tails that leads to a clear rejection of the log-Normal model implicit in log-linear regression analysis. It is also found that a wide theatrical release has a much smaller impact on box-office revenue when heavy tails and skewness are modelled explicitly.

Acknowledgements

This research was supported by a grant from the Committee on Research and Conference Grants of the University of Hong Kong. The author would like to thank M. C. Auld for his most generous

[14] Readers are referred to the web page for The R Project for Statistical Computing at www.R-project.org.

donation of the computing resources used in the production of this research, and Adelchi Azzalini for helpful comments on an earlier version of this paper.

References

Azzalini, A. (1985) A class of distribution which includes the normal ones, *Scandanavian Journal of Statistics*, **12**, 171–8.

Azzalini, A. (1986) Further results on a class of distribution which includes the normal ones, *Statistica*, **46**, 199–208.

Azzalini, A. and Capitanio, A. (1999) Statistical applications of the multivariate skew-normal distribution, *Journal of the Royal Statistical Society*, **B61**, 579–602.

Azzalini, A. and Capitanio, A. (2003) Distributions generated by perturbation of symmetry with emphasis on a multivariate skew-t distribution, *Journal of the Royal Statistical Society*, **B65**, 367–89.

Azzalini, A. and Dalla Valle, A. (1996) The multivariate skew-normal distribution, *Biometrika*, **83**, 715–26.

Azzalini, A. and Kotz, S. (2002) Log-skew-normal and log-skew-*t* distributions as models for family income data, Working Paper, Department of Statistical Sciences, University of Padua, Italy.

Brannas, K. and Nordman, N. (2003) Conditional skewness modelling for stock returns, *Applied Economics Letters*, **10**, 725–28.

De Vany, A. S. and Walls, W. D. (1999) Uncertainty in the movie industry, does star power reduce the terror of the box office?, *Journal of Cultural Economics*, **23**, 285–318.

De Vany, A. S. and Walls, W. D. (2002) Does Hollywood make too many R-rated movies?: risk, stochastic dominance, and the illusion of expectation, *Journal of Business*, **75**, 425–51.

De Vany, A. S. and Walls, W. D. (2004) Motion picture profit, the stable Paretian hypothesis, and the curse of the superstar, *Journal of Economic Dynamics and Control*, **28**, 1035–57.

Fama, E. (1963) Mandelbrot and the stable Paretian hypothesis, *Journal of Business*, **36**, 420–29.

Halvorsen, R. and Palmquist, R. (1980) The interpretation of dummy variables in semilogarithmic equations, *American Economic Review*, **70**, 474–75.

Hand, C. (2001) Increasing returns to information: further evidence from the UK film market, *Applied Economics Letters*, **8**, 419–21.

Harris, R. D. F. and Kucukozmen, C. C. (2001) The empirical distribution of stock returns: evidence from an emerging European market, *Applied Economics Letters*, **8**, 367–71.

Hogg, R. V. and Craig, A. T. (1978) *Introduction to Mathematical Statistics*, 4th edn, Macmillan, New York.

Ihaka, R. and Gentleman, R. (1996) R: a language for data analysis and graphics, *Journal of Computational and Graphical Statistics*, **5**, 299–314.

Judge, G., Griffiths, W., Hill, R., Lutkepohl, H. and Lee, T. (1985) *The Theory and Practice of Econometrics*, 2nd edn, Wiley, New York.

Litman, B. R. and Ahn, H. (1998) Predicting financial success of motion pictures: the early 90s experience, in *The Motion Picture Mega-Industry*, chapter 10 (Ed.) B. R. Litman, Allyn and Bacon, Needham Heights, MA, pp. 172–97.

Maddison, D. (2004) Increasing returns to information and the survival of Broadway theatre productions, *Applied Economics Letters*, **11**, 639–43.

Mandelbrot, B. (1963) New methods in statistical economics, *Journal of Political Economy*, **71**, 421–40.

Prag, J. and Cassavant, J. (1994) An empirical study of determinants of revenues and marketing expenditures in the motion picture industry, *Journal of Cultural Economics*, **18**, 217–35.

Ravid, S. A. (1999) Information, blockbusters and stars: a study of the film industry, *Journal of Business*, **72**, 463–86.

Smith, S. P. and Smith, V. K. (1986) Successful movies: a preliminary empirical analysis, *Applied Economics*, **18**, 501–07.

Walls, W. D. (1997) Increasing returns to information: evidence from the Hong Kong movie market, *Applied Economics Letters*, **4**, 187–90.

Walls, W. D. (2005) Modeling movie success when nobody knows anything: conditional stable-distribution analysis of film returns, *Journal of Cultural Economics*, **29**(3), 177–90.