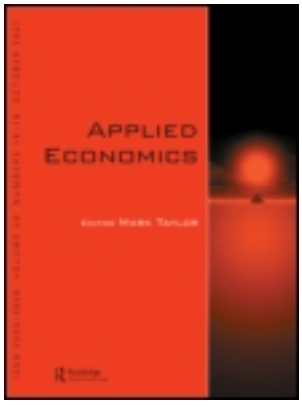

Applied Economics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/raec20

Factors determining UK album success

Caroline Elliott [a] & Rob Simmons [a]
[a] Department of Economics, Lancaster University Management School, Lancaster University, Lancaster LA1 4YX, UK
Version of record first published: 31 Jan 2011.

To cite this article: Caroline Elliott & Rob Simmons (2011): Factors determining UK album success, Applied Economics, 43:30, 4699-4705

To link to this article: http://dx.doi.org/10.1080/00036846.2010.498349

Applied Economics, 2011, **43**, 4699–4705

Routledge
Taylor & Francis Group

Factors determining UK album success

Caroline Elliott* and Rob Simmons

Department of Economics, Lancaster University Management School, Lancaster University, Lancaster LA1 4YX, UK

This article uses a recently compiled dataset on the UK album sales to determine which factors contribute to best-selling album sales success. We control for factors including length of time since release, nationality of artist, artist type and album type, testing the increasing returns to information hypothesis. Information on general public online review scores for the albums in the dataset allows for a strong test of the accuracy of online reviews in predicting music sales, as online reviews are a relatively recent phenomenon, while the release of many of the albums predates the widespread use of Internet.

I. Introduction

Despite the challenges faced by the music industry in recent years, the industry remains important to the UK economy, with the total UK recorded music income reaching £928.8 million in 2009 according to the British Phonographic Industry (2010).[1] Music albums, along with other entertainment goods such as films, can broadly be considered to be experience goods, which have to be 'consumed' for consumers to discover whether the product matches their preferences, although some characteristics of the product can be ascertained, *via* search, prior to purchasing (Nelson, 1970; Shapiro and Varian, 1999). Individual music tracks can be heard on television music channels, online or on the radio, but full information on the album may be difficult to obtain prior to purchase. This article uses a recently compiled dataset on the UK best-selling album sales over the past 50 years to determine which factors contribute to the album sales success. While there is an established literature examining the factors contributing to the sales of films, for example, Elberse and Eliashberg (2003), Moul (2007) and Elliott and Simmons (2008), the literature on the factors contributing to music sales is more limited (see Strobl and Tucker (2000); Connolly and Krueger (2006); Fox and Kochanowski (2007) for recent detailed summaries of developments in the music industry as well as discussion of the limited existing literature that explores the determinants of album success).

In the statistical analysis, we test whether factors such as length of time since release, nationality of artist, type of artist and type of album impact on best-selling album success. We are also able to test the increasing returns to information hypothesis, which suggests that past success of an album gives rise to a greater amount of album information in the public domain, contributing to further album success and sales. Increasing returns to information have previously been identified for a number of experience goods in the entertainment industry, including the US hit singles market in popular music (Giles, 2007); films (Walls (1997) for the US and Hand (2001) for the UK); and Broadway theatre productions (Maddison, 2004). This article extends this literature by confirming the increasing returns to information result for the UK music album sales. The result is

*Corresponding author. E-mail: C.Elliott@lancaster.ac.uk
[1] www.bpi.co.uk.

Applied Economics ISSN 0003–6846 print/ISSN 1466–4283 online © 2011 Taylor & Francis
http://www.informaworld.com
DOI: 10.1080/00036846.2010.498349

particularly strong as it is not only identified for the UK best-selling albums, but replicated using a second dataset of recently released albums.

Potential consumers have an increasing number of sources of information available to them when considering the purchase of experience goods such as music, films and books. Online reviews are increasingly common, and those provided by the general public are an alternative to expert reviews that continue to be published in newspapers and magazines, as well as online. Hence, data were collected on general public online review scores. This allows for a strong test of the accuracy of online reviews in predicting music sales, as online reviews are a relatively recent phenomenon, while the release of many of the albums in the dataset predates the widespread use of Internet. There is a growing empirical literature investigating the impact of third-party reviews in diverse settings such as Broadway shows (Reddy *et al*., 1998) cinema (Eliashberg and Shugan, 1997; Holbrook, 2005; Reinstein and Snyder, 2005; Elliott and Simmons, 2008); books (Chevalier and Mayzlin, 2006). Consequently, reviews, whether offered by 'experts' or increasingly online by members of the public, may be influential in determining the sales. However, we are unaware of any other research into the impact of online reviews in the music industry to date.

The structure of the remainder of the article is as follows: in the following section, the methodology and dataset adopted are discussed. Results are reported in Section III, with conclusions suggested in Section IV.

II. Methodology and Data

Empirical methodology

Pareto's law suggests the following relationship between size and rank:

$$SR^{\beta} = A \tag{1}$$

or, in a natural logarithmic form

$$\ln(S) = \ln(A) - \beta \ln(R) \tag{2}$$

where S is the size, R the rank and A, β the coefficients.

The relationship has been applied in a number of settings, including investigations into the relationship between firm size and rank (see, e.g. Steindl, 1965). However, it has been argued that (sales) growth of a firm, album, film, etc. may reflect its recent growth, such that 'success breeds success', as discussed in De Vany and Walls (1996). For example, this may be due to the greater propensity for word-of-mouth recommendations, which may be particularly valuable for experience good purchases. This has led a number of researchers, including Walls (1997), Hand (2001), Maddison (2004) and Giles (2007), to test for a relationship between size and rank for entertainment goods that departs from the simple Pareto law, with the model estimated typically taking the following form:

$$\ln(S) = \ln(A) + \beta \ln(R) + \gamma \ln(R)^2 \tag{3}$$

Of crucial importance is the sign associated with the coefficient on the squared rank variable. A negative, significantly different from zero coefficient, indicates a departure from Pareto's Law, with autocorrelated growth and increasing returns to information.

Initial Ordinary Least Squares (OLS) regression results suffered from heteroscedasticity. Hence, Huber's robust regression method was adopted in our sales regressions. Robust regression is an estimator designed to eliminate gross outliers. The procedure is to perform an initial screening of the data based on Cook's distance > 1. Once any gross outliers are removed, iterations are performed, using Huber weights and biweights sequentially, until convergence is achieved. The resulting SEs are then robust to heteroscedasticity. This procedure can be applied where a distribution exhibits excess kurtosis due to the presence of outliers, for example, for film revenues where a big blockbuster massively outperforms ordinary films. In our case, there are a few albums with extremely high sales that may appear as outliers.

Data collection

In November 2006, the UK Chart Company published the total sales of the 100 best-selling albums in the UK for the first time, with data collated from the 1950s onwards. The data includes vinyl as well as CD sales, and includes greatest hits, live albums and soundtracks. This allows researchers the opportunity to estimate the factors impacting on album success for the most popular albums. The analysis below focuses on the number of albums sold (*sales*) as a measure of size, with Maddison (2004) similarly using a count of number of Broadway performances. However, note that revenue has been used in Walls (1997) and Hand (2001), while Giles uses the number of weeks in the top position in the US singles chart. The rank and squared rank variables, the *rank* and *rank2* variables, respectively, refer to the position of the albums in the list of the 100 best-selling albums in

Table 1. Online review descriptive statistics

Variable	*N*	Mean	SD	Minimum	Maximum
UKamazonscore	100	4.48	0.61	0	5
UKamazoncount	100	66.66	81.48	0	437
USamazonscore	100	4.44	0.31	3.5	5
USamazoncount	100	327.34	445.52	1	2056
UKamazonrank	100	11 090.88	30 266.46	38	208 837

Notes: Amazonscore variables record the mean online scores received by albums. The UK data are correct as on 21 December 2006, and the US data are correct as on 22 December 2006. Amazoncount variables record the number of online scores submitted for an album, collected at the same time as the mean Amazon scores. There was only one album that had not received any online reviews on the UK Amazon website; this was 'Robson and Jerome' by Robson and Jerome, and so this album was excluded from the regression modelling.

the UK as released by the UK Chart Company. See Appendix for a list of variables used.

Further explanatory variables considered include dummy variables that identify albums by bands, male and female solo artists, artists of UK origin, greatest hits and soundtrack albums (typically to films or musicals). Information on release date and number of months since release, number of albums previously released by the artist(s), number of weeks in the UK album chart and number of weeks spent at number one in the UK album chart were obtained from Betts (2005). Note that only the number of albums previously released by a band or artist in their current recording guise is counted. Hence, for example, Robbie Williams's or George Michael's album releases when members of the bands Take That and Wham, respectively, are not counted as albums previously released by them. It was felt that participation in previous bands may increase artist recognition when they move on and, for example, embark on solo careers, but this may help or hinder the later album sales, particularly if an artist's style or genre of music changes. Some correlation between the UK and US album sales may be expected, and so a dummy variable was created taking the value unity when an album had obtained sales of at least one million in the US.

As highlighted in Section I, a number of papers have investigated the impact of online reviews on sales in a variety of industries. It is, maybe, surprisingly difficult to find online reviews of a broad range of albums in the UK, particularly if reviews are required of albums that were released more than a few years previously. The only suitable UK site was found to be www.amazon.co.uk, to which members of the public are able to submit online reviews, including a score between 0 and 5 (5 being the maximum score available). However, we were initially concerned about the representativeness of the average online scores awarded as some albums had only received a small number of reviews and associated scores.[2] Consequently, the online reviews on the US Amazon site www.amazon.com were also collected. As shown in Table 1, the SD of scores awarded both on the UK and US websites was very low, and it is hypothesized that respondents typically only give reviews and scores to albums that they have appreciated. If this is true, then the number of reviews submitted may also be indicative of the general public's views of albums. Hence, data were collected on the number of individuals reviewing albums and submitting online scores, on both the UK and US Amazon sites. Finally, data were collected from the UK Amazon site on the ranking (as of 3 January 2007) of albums according to Amazon sales (*UKamazonrank*). A further explanatory variable was then created, using the squared UK Amazon sales rankings (*UKamazonrank2*).

Typically, natural logarithms of continuous variables were taken, with the missing values resulting from taking the log of a zero value recorded as zeros.

There is unavoidable sample selection bias in the dataset as a result of collecting data for only the 100 greatest selling albums in the UK. However, the UK Chart Company was unable to provide sales data for a greater number of albums. Sample selection bias seems to be an unfortunate feature of many papers that examine the factors determining sales success for a number of entertainment products. Reflecting this, the current analysis is only able to indicate factors that contribute to sales success of highly successful albums, rather than albums in the UK market more generally.

[2] In fact one album, Robson and Jerome by Robson and Jerome, had not received any online review or rating score. Consequently, this album was deleted from the analysis that follows.

Table 2. Robust regression results

Variable	Regression 1	Regression 2	Regression 3
lrank	−0.0334 (0.00)	−0.0409 (0.00)	−0.0381 (0.00)
lrank2	−0.0396 (0.00)	−0.0388 (0.00)	−0.0393 (0.00)
Months × *lrank*	-2.47^{e-6} (0.27)		
UK	−0.0036 (0.30)		
lmonths	0.0016 (0.83)		
lwksinukchart	0.0018 (0.62)		
lwks@1inukchart	0.0006 (0.72)		
2000interact	−0.0008 (0.53)		
lUKamazonrank	−0.0164 (0.03)	−0.0120 (0.05)	
lUKamazonrank2	0.0010 (0.03)	0.0008 (0.05)	
Constant	15.3794 (0.00)	15.3870 (0.00)	15.3387 (0.00)
N	99	99	99
F	2838.29	7116.68	13 750.39
$\bar{R}^2$	0.9966	0.9966	0.9964

Note: *p*-values in parentheses.

III. Results

Increasing returns to information

Results in Table 2 are reported using the log of the number of album sales as the dependent variable, estimated using Huber's robust regression method. A general-to-specific methodology was used to select the explanatory variables in the reported regressions, with the final iterations reported. The final column of Table 2 reports the regression results with only the rank and squared rank explanatory variables to aid comparison with the previous papers published. We recognize that as in all of the literature on increasing returns to information, there is a problem that the explanatory rank variables will be endogenously determined. Unfortunately, this is an unavoidable weakness facing all studies of this nature, including those of Walls (1997), Hand (2001), Maddison (2004) and Giles (2007). Further, as identified by Hand (2001), using rank and squared rank variables as explanatory variables may be problematic as they will be closely related.

Nevertheless, the results reported confirm the result found in previous analyses, including that of Giles (2007) for the US music industry, namely, that increasing returns to information exist, as shown by the highly significant negative coefficient always associated with the squared logged sales rank variable. This result has also been identified in the film industry by Walls (1997) and Hand (2001), and for Broadway productions by Maddison (2004).

The finding of increasing returns to information in Giles (2007) comes from hit (number one) singles in the US popular music market. It is worth considering just how general this result is. We can also show evidence from a sample of 315 recent albums released in the UK that appeared in the top 40 over the period 2002 to 2004. This is a broader category, as it is not restricted to number one albums and albums themselves tend often to occupy niche markets and enjoy less mass popularity than the hit singles analysed by Giles. For this second sample, we do not know the sales figures, just the number of weeks in the UK top 40 album chart. We construct a rank measure based on weeks of appearance over the sample period. Where number of weeks of appearance is the same for more than one album, we base the rank measure on the highest chart position obtained.

Using Huber's robust regression estimates to facilitate comparison with the results reported above and again adopting a double-log functional form our estimates are as follows:

$$\begin{aligned} \text{LOG WEEKS} &= 3.366 + 1.191\ \text{LOG RANK} \\ &\quad - 0.287\ (\text{LOG RANK})^2 \qquad (4) \\ \bar{R}^2 &= 0.96 \end{aligned}$$

where both coefficients have *p*-values of 0.00. Hence, the results of Giles for US hit singles, and our own findings for the most popular albums in the UK are supported by this additional dataset that comprises some less popular albums. Broadening the music category away from the most popular singles or albums seems not to damage the increasing returns result.

Similar Huber's robust regressions using the 100 best-selling albums dataset were then estimated, with either the number of weeks in the chart (to ensure comparability with the previous regression) or the number of weeks at number one in the albums chart

Table 3. Additional increasing returns regression results

Variable	Regression 1 Dependent variable *lwksinukchart*	Regression 2 Dependent variable *lwks@1inukchart*
lrank	−0.0203 (0.40)	2.6213 (0.00)
lrank2	−0.1006 (0.00)	−0.5459 (0.00)
Constant	6.1904 (0.00)	−0.5369 (0.01)
N	99	99
F	6994.89	898.81
$\bar{R}^2$	0.9930	0.9487

Note: *p*-values in parentheses.

(to aid comparability with Giles (2007)) as the dependent variable, with the rank variables similarly constructed according to either weeks in the chart or the number of weeks at the top chart position. Results are reported in Table 3.

Yet again, the results support the hypothesis of increasing returns to information. Consequently, this result appears particularly robust. Both Giles (2007) and this article use datasets spanning approximately 50 years, and so the relationship between measures of success and rank may have been predicted to be weaker, as numerous factors may influence single or album sales over such a long period of time. This prediction is not borne out. A comparable increasing returns result emerges for albums and singles, and the same result is identified in the dataset of additional albums reported in this article even though the data period is much shorter and some of the albums appeal to more niche audiences.

Determinants of album sales

Determining the factors that influence (logged) sales is complicated by the strength of the relationships between the dependent variable and the logged rank and squared rank variables, −0.9767 and −0.9785, respectively. The regression results in Table 2 indicate that the only other coefficients significantly different from zero are those associated with the UK Amazon sales rankings. This is, maybe, not surprising given the diversity of albums that appear in the list of the 100 best-selling albums in the UK. However, it was expected that stronger relationships between album sales and the length of time since the album's release and the weeks in the UK album chart may have been identified. Following Giles (2007), an interaction variable was created by multiplying the log of sales rank by the number of months since release. Yet, unlike the results discussed in the 'Increasing returns to information', our results differ from those of Giles (2007) as the coefficient on this variable was not found to be significantly different from zero.

Considering the impact of online reviews, in regressions including the logged average UK or US online review score, the coefficient on these variables was consistently found to be insignificantly different from zero, although this was not unexpected given the very low SDs associated with these data, reported in Table 1. It is hypothesized that typically it is individuals who have enjoyed an album who make the effort to submit a review score for the album online. Hence, not only was the SD of the scores low, but Table 1 reports the high simple arithmetic means of scores submitted to both the UK and US Amazon sites.[3] More surprisingly, when the log of the number of scores recorded in either the UK or US was included, again the coefficient on these variables was found to be insignificantly different from zero.

An inspection of the data suggested that the majority of online reviews relate to albums released relatively recently. A dummy variable taking the value of unity for albums released in the UK since 1 January 2000 was created, with this variable multiplied by the log of the number of UK online scores recorded, to create an interaction variable *2000interact*. The choice of 1 January 2000 was not arbitrary. Rather, Gopal *et al*. (2004) suggest that during the 1990s, and particularly in the late 1990s, the music industry changed due to the availability of downloadable music online.[4] Nevertheless, when this variable was included, the coefficient associated with it was again found to be insignificantly different from zero.

This suggests that online reviews do not have an impact on sales, even though online sales are related to album sales more generally, as evidenced by the coefficients on the UK Amazon rank variables. We can only speculate as to whether this result will continue to hold in the future as consumers continue to post reviews online for an increasing number of products. Again a nonlinear relationship was identified between the album sales and the UK Amazon

[3] To confirm that the high average scores and low SD of scores is not just a feature of the albums in the dataset, the mean and SD of the UK Amazon scores for a second sample of albums were calculated. Using the top 30 albums sold in the UK in the week commencing 5 November 2006 (just before the 100 best-selling UK albums sales figures were published) and the Amazon scores from December 2007, the mean score was 4.18, with a SD of 0.56.

[4] Nevertheless, the analysis was repeated with 1 January 2002 the date from which the dummy variable took the value unity. Very similar results were obtained, although the explanatory power of the model was lower. Results are available upon request.

sales rankings, as shown by the negative coefficient on *lUKamazonrank* and a positive coefficient on *lUKamazonrank2*, both of these coefficients being significantly different from zero at least at the 5% level. Note that the analysis benefits from the consistency to be gained from using sales rank and sales rank squared explanatory variables, and similarly Amazon sales rankings and squared Amazon sales rankings.

It should be noted that the coefficients on the dummy variables representing type of album, artist (other than the UK based artists) and the US sales in excess of 1 million were never found to be significantly different from zero, at at least the 10% significance level. These findings contrast with those reported by Fox and Kochanowski (2007) who found, amongst other results, that solo artists and male artists achieved more gold albums in the US market. A dummy explanatory variable was also created taking the value unity when a leading member of a group or solo artist is deceased.[5] However, the coefficient on this explanatory variable was never found to be significantly different from zero, indicating that album sales do not increase significantly after the death of a band member. Similarly, the coefficient on the logged number of albums previously released was never significantly different from zero, so this variable was omitted in the final reported regressions.[6]

IV. Conclusions

This article contributes to a limited literature exploring the factors determining music album success. Results support strongly the hypothesis of increasing returns to information found previously to hold in the music industry and more generally, for a number of entertainment products. However, the research also explored whether other factors could be identified as contributing to the best-selling album sales success, concluding that it is difficult to identify such factors.

This article offers a further original contribution to the existing literature investigating album success as it tests whether online reviews can predict music success. Online reviews are a relatively recent phenomenon, while the release of many of the albums in the dataset predates the widespread use of Internet. Results suggest that online reviews do not predict album sales neither generally, nor just for albums released relatively recently (since 2000). Interestingly, there was a very limited distribution of scores given to albums, both in the UK and the US, suggesting that individuals typically were only likely to submit online reviews and scores for the albums that they had enjoyed. Yet even when the number of individuals submitting an online score for an album in the UK was adopted as an explanatory variable rather than the mean of the scores submitted, the same result emerges.

Ultimately, it appears that success really does breed further success to the exclusion of other factors in the UK album industry. Nevertheless, a drawback of the analysis arises from data limitations. Data were only available for the 100 best-selling albums in the UK, and so the analysis can only offer insights into the factors influencing sales of the best-selling albums. In addition, as online reviews continue to be posted, the richness of online review data available is expected to increase. Further research is still required into factors influencing album sales more generally, but this future research will benefit from the availability of greater online review data.

Acknowledgements

The authors thank Owen Rodwell for providing excellent research assistance, the two anonymous referees, as well as the participants of the 2007 Scottish Economic Society and Western Economic Association Conferences for their helpful comments. All errors, of course, remain the responsibility of the authors.

References

Betts, G. (2005) *Complete UK Hit Albums 1956–2005*, Collins, London.

British Phonographic Industry (2010) Available at http://www.bpi.co.uk/press-area/news-amp3b-press-release/article/recorded-music-sales-revenue-stabilise-in-2009 aspx (accessed 5 May 5 2010).

Chevalier, J. and Mayzlin, D. (2006) The effect of word of mouth on sales: online book reviews, *Journal of Marketing Research*, **43**, 345–54.

Connolly, M. and Krueger, A. B. (2006) Rockonomics: the economics of popular music, in *Handbook of the Economics of Art and Culture* (Eds) V. Ginsburgh and D. Throsby, Elsevier, Amsterdam, pp. 667–719.

[5] The dummy variable takes the value unity in the dataset for albums recorded by The Beatles, Bob Marley, Nirvana and Queen.

[6] A zero value was inserted in the place of a missing value when the number of albums previously released was zero. Otherwise, when this variable was logged, 35 missing values would have reduced the size of the dataset considerably.

De Vany, A. and Walls, W. D. (1996) Bose-Einstein dynamics and adaptive contracting in the motion picture industry, *Economic Journal*, **106**, 1493–514.
Elberse, A. and Eliashberg, J. (2003) Demand and supply dynamics for sequentially released products in international markets: the case of motion pictures, *Marketing Science*, **22**, 329–54.
Eliashberg, J. and Shugan, S. (1997) Film critics: influencers or predictors?, *Journal of Marketing*, **61**, 68–78.
Elliott, C. and Simmons, R. (2008) Determinants of UK box office success: the impact of quality signals, *Review of Industrial Organization*, **33**, 93–111.
Fox, M. and Kochanowski, P. (2007) Multi-stage markets in the recording industry, *Popular Music and Society*, **30**, 173–95.
Giles, D. (2007) Increasing returns to information in the US popular music industry, *Applied Economics Letters*, **14**, 327–31.
Gopal, R., Sanders, G., Bhattacharjee, S., Agrawal, M. and Wagner, S. (2004) A behavioral model of digital music piracy, *Journal of Organizational Computing and Electronic Commerce*, **46**, 107–11.
Hand, C. (2001) Increasing returns to information: further evidence from the UK film market, *Applied Economics Letters*, **8**, 419–21.
Holbrook, M. (2005) The role of ordinary evaluations in the market for popular culture: do consumers have 'good taste'?, *Marketing Letters*, **16**, 75–86.
Maddison, D. (2004) Increasing returns to information and the survival of broadway theatre productions, *Applied Economics Letters*, **11**, 639–43.
Moul, C. (2007) Measuring word of mouth's impact on theatrical movie admissions, *Journal of Economics and Management Strategy*, **16**, 859–92.
Nelson, P. (1970) Information and consumer behavior, *Journal of Political Economy*, **78**, 311–29.
Reddy, S., Swaminathan, V. and Motley, C. (1998) Exploring the determinants of broadway show success, *Journal of Marketing Research*, **35**, 370–83.
Reinstein, D. and Snyder, C. (2005) The influence of expert reviews on consumer demand for experience goods: a case study of movie critics, *Journal of Industrial Economics*, **53**, 27–52.
Shapiro, C. and Varian, H. (1999) *Information Rules: A Strategic Guide to the Network Economy*, Harvard Business School Press, Boston.
Steindl, J. (1965) *Random Process and the Growth of Firm: A Study of the Pareto Law*, Hafneer, New York.
Strobl, E. A. and Tucker, C. (2000) The dynamics of chart success in the UK pre-recorded popular music industry, *Journal of Cultural Economics*, **24**, 113–34.
Walls, W. D. (1997) Increasing returns to information: evidence from the Hong Kong movie market, *Applied Economics Letters*, **4**, 287–90.

Appendix: Variable List

Variable name	Descriptor
Rank	Sales rank unless otherwise stated
rank2	Squared rank
UK	Dummy variable equal to unity if artist is from the UK
Months	Number of months since album release
wksinukchart	Count of weeks in the UK album chart
wks@1inukchart	Count of weeks at top of the UK album chart
2000interact	Dummy variable equal to unity for albums released since 1 January 2000, multiplied by the number of UK Amazon scores
UKamazonrank	Rank of sales on the UK Amazon site
UKamazonrank2	Squared rank of the UK Amazon sales
L	This denotes a (natural) logged variable