Estimating the Current Value of Time-Varying Beta

Joseph Cheng
Ithaca College

Elia Kacapyr
Ithaca College

This paper proposes a special type of discounted least squares technique and applies it to the Capital Asset Pricing Model. There is evidence that the value of beta, the measure of risk in the model, varies over time. The technique, entropic least squares, detects differences in the past and present standard error of the model. The rate of change in this standard error is referred to as the entropy rate. Unlike discounted least squares where the discount rate must be assumed in an ad hoc manner, entropic least squares estimates the entropy rate simultaneously with the parameters of the model.

INTRODUCTION

The idea that beta shifts over time in the capital asset pricing model (CAPM) was first suggested by Blume (1975). He found that there is a consistent tendency for a portfolio with an extremely low or high estimated beta in one period to regress toward a less extreme beta in the following periods. Blume discovered a tendency for betas to regress toward unity, the grand mean of all betas, over time. In other words, beta values are not constant. Some analysts estimate an equity's beta of by assigning 2/3 weight to the past data and 1/3 weight to unity. Rosenberg and Guy (1976) estimate beta with past data and adjust that estimate according to a set of individual corporate financials. This suggests that beta may change with corporate characteristics over time.

Fabozzi and Francis (1978) suggest that many stocks' beta coefficients move randomly through time rather than remain stable as the ordinary least squares (OLS) technique presumes. The nonstationarity of beta and the time-varying behavior of equity return co-movements also are suggested by Umstead and Bergstrom (1979), Theobald (1981), McDonald (1985), Lee et al. (1986), Levy (1971), Rosenberg (1985), Kaplanis (1988), and Koch and Koch (1991).

One of the drawbacks of using OLS for time-series estimation is that locally evolving linear trend patterns may be ignored. However, local trends are important when projecting into the immediate future. In this case, it makes more sense to assign greater weight to more recent observations.

We develop a weighted least squares method that amounts to allowing the parameter estimates to change over time. In a scenario where beta values change over time, the objective of financial analysts is to estimate the most current value of beta for making current investment decisions. Recent data should make for better forecasts than older data. Still, the older observations carry some information. As the beta or coefficient value changes over time, the standard error based on the current value of beta would be different for historical periods than for more recent periods. We refer to the rate of change in the standard error over time as the entropy rate. An iterative procedure is used in this paper to estimate the appropriate entropy rate based on the standard error of the residuals as well as to estimate the most current value of beta. This approach leads to alternate estimates of beta in Capital Asset Pricing Model. In the following section, we derive the entropic least squares (ELS) method for estimating parameters in time-variant standard error cases.

DERIVATION OF THE ELS ESTIMATOR

Consider the standard OLS regression equation:

$$Y_t = a + b\, X_t + u_t$$

where Y_t = value of dependent variable at period t
a = constant term
X_t = value of independent variable at period t
b = beta or coefficient of X_t
u_t = error at period t

Before deriving the normalizing factor to be used for rendering the variance of the error term constant, we need first to specify the nature of the change in the standard error of the stochastic error term. In this model, the standard error of the stochastic error term is assumed to grow at a constant rate:

$$S_t = S_0\, e^{rk}$$

where S_t = standard error of u at period t
S_0= base standard error, which is the standard error of u for the latest period.
e = 2.71828
r = rate of entropy (growth rate for the standard error of u)
k = number of time periods retrogressed from the most recent observation, which is zero for the latest time period in the sample.

If we assume the variance grows at a constant rate, the specification changes to

$$S_t = S_0 \sqrt{e^{rk}}$$

Applying OLS to $Y_t = a + b\, X_t + u_t$ would yield inefficient estimates when the variance of u_t is time variant. To correct this problem, we divide both sides of $Y_t = a + b\, X_t + u_t$ by a variance deflator to render the transformed residual homoskedastic. As is typically done to remedy heteroskedasticity, the standard error of the residual term in the original equation is used as the deflator. In this case, the deflator would be the standard error of u_t, which is S_t.

$$Y_t/S_t = a/S_t + b\, X_t/S_t + u_t/S_t$$

The transformed equation has two independent variables (a/S_t and X_t/S_t) and no constant term.
The above transformation essentially assigns a weight of $1/S_t$ to the residual u_t. Since the value of S_t is positive and increases with the number of periods retrogressed from the present, the residuals (u_t/S_t) in earlier periods have higher standard deviations, or higher S_t values, and thus carry lower weights ($1/ S_t$) than residuals of more recent periods.

The variance of the error term (u_t/S_t) in the transformed equation is now constant with respect to time. The estimated beta (b) derived by minimizing the transformed equation is unbiased and efficient.

Dealing with the Problem of Unknown Base Standard Error and Growth Rate

Before OLS can be applied to the transformed equation, the values of S_t must first be computed so that the appropriate weights can be assigned. The value of S_t is dependent on S_0 and r. The problem encountered at this stage is that neither S_0 nor r is observable. In the method to be proposed, this problem is solved without *a priori* knowledge of either S_0 or r.

Since S_0 is unknown at this stage, it is convenient to use it as a divisor for S_t. And because S_0 is a constant, the new ratio (F_t) will be proportionate to S_t:

$$F_t = S_t/S_0 = e^{rk}$$

Then F_t can be used as the divisor for $Y_t = a + b X_t + u_t$ since F_t combines the two unknowns (S_0 and r) into a single variable. The newly transformed equation is:

$$Y_t/F_t = a/F_t + b X_t/F_t + u_t/F_t$$

The newly transformed equation will exhibit homoskedasticity since it is divided through by F_t which is proportionate to S_t. In addition, b in the newly transformed equation will equal b in $Y_t = a + b X_t + u_t$ since Y_t and X_t have been scaled by the same factor, F_t.

F_t (= e^{rk}) is unknown only because r is unknown. Therefore, a reiteration procedure is devised to estimate $Y_t/F_t = a/F_t + b X_t/F_t + u_t/F_t$ and r simultaneously using a stepwise algorithm.

If there is no change in the standard error of the residuals (S_t) over time, then the rate of entropy (r) would be zero. In this case, the divisor F_t simply reduces to 1 for all observations and thus Y_t/F_t would be same as Y_t and X_t/F_t the same as X_t. Consequently, the results of regression on $Y_t/F_t = a/F_t + b X_t/F_t + u_t/F_t$ are equivalent to OLS, which entails the implicit assumption of r equal to zero. However, if r is positive, then entropy occurs in the data and the magnitude of the error worsens with k. With a positive value for r, F_t would equal unity only for the most recent observation. In such a case, dividing the u_t by F_t would mean that the weight assigned to the error terms is $1/F_t$. The value of this weight equals 1 for the most recent observation where k = 0. The weight on the error term will decline progressively for observations going back further into the past. The question, however, is what specific value for r, which determines F and thus the weights, should we assume. Since the value of r is unknown, the reiterative procedure begins by setting r to zero, then increasing it by a small increment with each successive iteration. In the first iteration where r is set at zero, the value of F is one and OLS is applied to $Y_t/F_t = a/F_t + b X_t/F_t + u_t/F_t$.

For the second iteration, r is increased by .001 (or some other increment deemed appropriate to the case at hand) to generate a new set of F_t's. Once again, OLS is applied to $Y_t/F_t = a/F_t + b X_t/F_t + u_t/F_t$. For the third iteration, r is set at .002. The process continues until the value of r in the particular iteration is essentially equal to the implied value of r.

Estimating the Implied Entropy Rate

The implied entropy rate is obtained by taking the natural log of both sides of $S_t = S_0\, e^{rk}$:

$$\ln S_t = \ln S_0 + r^* k$$

where r is replaced with r*, the implied entropy rate

Multiply both sides by 2 to obtain:

$$2 \ln S_t = 2 \ln S_0 + 2 r^* k$$

or

$$\ln (S_t^2) = \ln (S_0^2) + 2 r^* k$$

Since we do not have data on $\ln(S_t^2)$, we use each individual observation of estimated u_t^2 as a proxy for S_t^2. In addition, a stochastic term (v_t) is added to the deterministic equation above in order to create a stochastic regression equation as follows:

$$\ln(u_t^2) = \ln(S_0^2) + 2r^*k + v_t$$

According to the equation above, the size of the error term squared in log form ($\ln(u_t^2)$) is expected to equal the size of the base error term ($\ln(S_0^2)$) plus the number of periods away from current period times 2r*. The stochastic error term v_t is assumed to be white noise. Note that $\ln(u_t^2)$ may be viewed as the dependent variable of a simple regression equation and $\ln(S_0^2)$ as the constant term. The independent variable is 2k with r* as the coefficient. OLS is applied to obtain the estimated value for r*, the implied entropy rate.

To summarize the procedure to this point: An arbitrary value close to zero is selected for r and used to run the following regression:

$$Y_t/F_t = a/F_t + b\,X_t/F_t + u_t/F_t$$

Recall that $F_t = S_t/S_0 = e^{rk}$. The error terms (u_t/F_t) squared from this regression are used as the u_t^2 in

$$\ln(u_t^2) = \ln(S_0^2) + 2r^*k + v_t$$

OLS is applied to the regression above to obtain an estimate for r*, the implied entropy rate. When the difference between r and r* is essentially zero the process stops. The value of b in this final regression is the entropic least squares estimate. It can be decided beforehand how close r* must be to r before considering them equal.

APPLICATION TO THE CAPITAL ASSET PRICING MODEL

As an illustration, the ELS procedure is applied to estimating beta for Apple Inc. Since the structure and product mix of this firm has changed markedly over the sample period (1993-2008) it might be expected that the beta value changed over time. In this situation, the ELS procedure may be particularly effective.

The dependent variable of the CAPM in this case is APPLE RETURN, the monthly returns for Apple, Inc. from February 1993 through July 2008. The independent variable is S&P RETURN, the monthly returns for the Standard and Poor's 500 Index over the same time period. Table 1 compares the results of the OLS and ELS procedures. We estimate beta for Apple by regressing the monthly return of Apple against that of the S&P, using both ELS and OLS.

Convergence is achieved by the ELS reiteration procedure at r =.0040689. The entropy rate, r, represents the monthly rate of change in the standard error of the ELS model. Higher entropy rates indicate firms whose beta is changing more rapidly over time. The hypothesis that the entropy rate is equal to zero is rejected at the 5 percent critical level in this case.

The ELS estimate of beta (1.65324) suggests that Apple, Inc. is a riskier stock than the OLS estimate (1.47020). However, the two estimates of Apple's beta are not significantly different at the 5 percent critical level. Given a longer time-series, the estimates would diverge. The fact that the entropy rate (r) is significant suggests that the estimate of Apple's beta in the CAPM is changing over time. This in turn implies that Apple is a dynamic company in the sense that one of its essential characteristics, riskiness, is changing in relation to the market average.

The intercept of the ELS regression (-6.19427) represents $\ln(S_0^2)$. From this we can derive the value of S_0 to be 0.045178.

TABLE 1
ESTIMATED RESULTS FOR APPLE, INC.

APPLE $RETURN_t = a + b$ S&P $RETURN_t + u_t$

	OLS	**ELS**
Beta	1.47020 (5.71681)	1.65324 (6.71794)
R-squared	0.150829	0.207383
r	N.A.	0.0040689 (2.41958)
Standard Error of (S_0)	0.138156	0.045178 (-17.0813)
Standard Error of (S_t)	0.138156	$0.045178*e^{0.0040689\ k}$

*Figures in parentheses represent t-statistics.

Application to Industry Indexes

In this section, the ELS procedure is used to estimate the CAPM for two industry classifications, pharmaceuticals (DRG) and technology (TXX). The results are contrasted with OLS estimation. Monthly rates of return in each industry are regressed on the S&P rates of return. The estimation period is June 1996 through September 2008. Table 2 gives the results for the pharmaceutical industry.

TABLE 2
ESTIMATED RESULTS FOR PHARMACEUTICALS (DRG)

$DRG_t = a + b$ S&P $RETURN_t + u_t$

	OLS	**ELS**
Beta	0.585990 (7.40749)	0.556135 (7.21914)
R-squared	0.273165	0.26019
r	N.A.	0.0055000 (2.58576)
Standard Error of (S_0)	.040883	0.013372 (-23.6293)
Standard Error of (S_t)	.040883	$0.013372*e^{0.0055\ k}$

*Figures in parentheses represent t-statistics.

The entropy rate for this industry is 0.0055000, which is statistically different from zero at the 5 percent critical level.

Table 3 shows the results for the technology sector (TXX).

TABLE 3
ESTIMATED RESULTS FOR TECHNOLOGY (TXX)

$TXX_t = a + b\ S\&P\ RETURN_t + u_t$

	OLS	**ELS**
Beta	1.81913 (17.4602)	1.75849 (18.8645)
R-squared	0.676173	0.707389
r	N.A.	0.0079 (4.007)
Standard Error of (S_0)	0.053845	0.013948 (24.9911)
Standard Error of (S_t)	0.053845	$0.013948 * e^{0.0079\, k}$

*Figures in parentheses represent t-statistics.

The entropy rate for this industry is 0.0079, which is statistically different from zero at the 5 percent critical level. The above results indicate that pharmaceuticals and technology may be considered dynamic industries because their entropy rates are significantly different than zero.

As a final application, the ELS procedure is used to estimate the CAPM for several industries using the Fama-French classifications and data. The data are monthly from January 1970 through December 2005. First, the return on the market portfolio is calculated as the average return for all 48 Fama-French industries. We then estimated beta for each industry by regressing the monthly return of the respective industry against the market portfolio return using both OLS and ELS. Table 4 presents the results.

TABLE 4
ESTIMATED RESULTS FOR FAMA-FRENCH INDUSTRIES

$\text{INDUSTRY RETURN}_t = a + b \text{ 48-INDUSTRY RETURN}_t + u_t$

Industry	OLS Beta	ELS Beta	r
Agriculture	0.820065 (18.787)	0.810531 (18.404)	0.00022 (0.5549)
Entertainment	1.31528 (31.8457)	1.321 (31.9167)	0.00026 (0.69100)
Health Services	1.23096 (22.429)	1.14446 (21.421)	0.0009 (2.265)
Medical Equipment	0.83456 (26.052)	0.831657 (25.8968)	0.00022 (0.5848)
Chemicals	0.982735 (37.3815)	0.988259 (37.4076)	0.00006 (0.15298)
Personal Services	1.05628 (26.9408)	1.00585 (25.2758)	0.00065 (1.73293)
Paper	0.955395 (31.6414)	0.959361 (31.6497)	0.00080 (2.05085)
Wholesale Trade	1.03128 (46.8354)	1.0246 (46.2189)	0.00011 (.26938)
Meals	1.06911 (32.8892	1.05456 (32.4307	0.00036 (0.92267
Insurance	0.916294 (28.8947)	0.92031 (28.806)	0.00053 (1.382)

*Figures in parentheses represent t-statistics.

For the Fama-French data set, only two industries have significant entropy rates at the 5 percent critical level - health services and paper. The entropy rate for personal services is significant the 10 percent critical level. These three industries are found to have more significant entropy rates than the others which suggests that they are dynamic industries whose beta values vary over time. In such cases, analysts who use ELS to estimate the beta value of industries may very well outperform analysts who use OLS.

CONCLUSIONS

This paper offers an alternative approach for estimating the current value of beta in the Capital Asset Pricing Model (CAPM). Entropic least squares (ELS) is more sophisticated than ordinary least squares (OLS) or discounted least squares (DLS). ELS allows for the standard error of the stochastic disturbance term to vary over time. The rate of growth of the standard error is referred to as the entropy rate, which is to be estimated rather than assumed. Unlike DLS, ELS estimates of the entropy rate are determined by the data themselves with an iterative process.

In this paper, ELS was used to estimate beta for the Capital Asset Pricing Model (CAPM). We found that Apple, Inc. has an entropy rate that is statistically different from zero. This implies that the CAPM beta for Apple, Inc. varies through time. Furthermore, health services, paper, and perhaps the personal services industries may be considered more dynamic than other industries in the sense that their operations or financial structure are changing through time so that reliance on a fixed beta is not appropriate.

ELS is a regression technique that can be useful in many financial time-series applications. When the entropy rate is found to be statistically significant in a particular model it implies that the regression coefficients are time dependent**.** In such cases, ELS should outperform more tradition estimation procedures.

REFERENCES

Blume, M. (1975). Betas and their Regression Tendencies. *Journal of Finance*, 30, (3), 785-796.

Fabozzi, F.J. & Francis, J.C. (1978). Beta as a Random Coefficient. *Journal of Financial and Quantitative Analysis*, 13, (1), 101-116.

Kaplanis, E.C. (1988). Stability and Forecasting of the Co-movement Measures of International Stock Market Returns. *Journal of International Money and Finance*, 7, 63–75.

Koch, P.D. & Koch, T.W. (1991). Evolution in Dynamic Linkages Across National Stock Indexes. *Journal of International Money and Finance*, 10, 231–251.

Lee, C.F., Newbold, P., Finnerty, J.E. & Chu, C.C. (1986). On Accounting-Based, Market-Based and Composite-Based Beta Predictions: Methods and Implications. *The Financial Review*, 21, 51-68.

Levy, R.A. (1971). On the Short Term Stationarity of Beta Coefficient. *Financial Analyst Journal*, 27, 55-62.

McDonald, B. (1985). Making Sense Out of Unstable Alphas and Betas. *Journal of Portfolio Management*, 11, 19-22.

Park, R.E. (1966). Estimation with Heteroscedastic Error Terms. *Econometrica*, 34, (4), 888.

Rosenberg, B. (1985). Prediction of Common Stock Betas. *Journal of Portfolio Management*, 11, 5-14.

Rosenberg, Barr & Guy, J. (1976). Beta and Investment Fundamentals, Part 1. *Financial Analyst Journal*, 32, (3), 60-72.

Rosenberg, Barr & Guy, J. (1976). Beta and Investment Fundamentals, Part 2. *Financial Analyst Journal*, 32, (4), 62-70.

Theobald, M. (1981). Beta Stationarity and Estimation Period: Some Analytical Results. *Journal of Financial and Quantitative Analysis*, 15, (5), 747-757.