

Conflicting Effects of Market Volatility on the Power of Two-Pass OLS Test of the CAPM: A Simulation Analysis

SANDUN FERNANDO

Department of Finance, University of Kelaniya, Sri Lanka

sandunf@kln.ac.lk

P D NIMAL

Department of Finance, University of Kelaniya, Sri Lanka

pdnimal@gmail.com

Abstract

The study focuses on widely used CAPM test, the two-pass test. It is argued and shown by a simulation that power of the test does not continuously decrease when the standard deviation of market premium increases as existing literature reports. Rather it is peaked at the middle indicating the standard deviation of market premium has negative as well as positive effects on the power. The implication of this finding is that rejecting the CAPM is highly possible in lesser volatile market as in a highly volatile market even when the CAPM exists in those markets.

Keywords: CAPM, Two-pass OLS regression, Power determinants, Market volatility

Introduction

Developing the Capital Asset Pricing Model (CAPM) (Sharpe, 1964; Lintner, 1965; and Black, 1972) is a hallmark of modern finance theory. Simply the CAPM explains that the expected return for a risky asset is linearly related to its systematic risk. This can be presented as Equation (1).

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] \quad (1)$$

Where

$$\beta_i = \frac{\sigma(R_i, R_m)}{\sigma^2(R_m)}$$

Here R_i is the return on a asset or portfolio. R_f and R_m are return on risk-free asset and return on market portfolio respectively. β_i is the sensitivity of particular asset's return to market return. $\sigma(R_i, R_m)$ and $\sigma^2(R_m)$ are respectively the covariance between returns of an asset/portfolio and of the market portfolio and the variance of the market portfolio. $E()$ is the expectation operator. The theory has been used for different purposes ranging from explaining risk and return relation, estimating cost of equity capital, measuring performance of fund managers, etc. To test the CAPM, many tests have been developed over the years. In general the empirical evidence for the CAPM is unsupportive (Banz, 1981; Rosenberg, Reid, and Lanstein, 1985; Basu, 1983; Gibbons et al., 1989; Fama and French, 1993) especially after Fama and MacBeth (1973).

The unfavorable empirical evidence does not necessarily mean that the CAPM is invalid, since the tests from which empirical evidence generate are not perfect in detecting the existence of the CAPM, since there is a possibility that a CAPM test can conclude the CAPM does not exist when it really exists. A reliable test has very low possibility of making such a wrong conclusion. Reliability of a statistical test is measured by the power of the test (defined in section 2). Therefore, determining the power of CAPM tests is important for meaningful interpretation of unfavorable empirical evidence on the theory.

Various power studies on CAPM tests have been conducted. However, the power of two-pass Ordinary Least Squares (OLS) regression test of Fama and MacBeth (1973) (two-pass test) has taken considerable attention, probably due to the fact that the two-pass test is one of the most popular CAPM tests among researchers. Further, in Sri Lankan context too the two-pass test is widely used (Samarakoon, 1997; Nimal, 1997). Therefore, this paper focuses on the power of the two-pass test. Such power studies on the two-pass test have been conducted by Afleck-Graves and Bradfield (1993); Elsas et al., (2003); Freeman and Gruemant (2006); Shanken and Zhou (2007); Grauer and Janmaat (2009). They have identified, average market premium (\bar{r}_m), the standard deviation of market risk premium ($\sigma(r_m)$), the standard deviation of error term (the

white noise) ($\sigma(e)$), the duration of testing period, the number of portfolios formed in the test, and the standard deviation of β_i among stocks ($\sigma(\beta_i)$) as power determinants of the test.

However, findings of some of these studies on the power determinants of the two-pass test are not consistent with the arguments that this paper brings forward. Specifically, the argument that $\sigma(r_m)$ has a negative relation with the power of the two-pass test has not studied. Existing literature fails to identify another role played by $\sigma(r_m)$, the positive effects on the power. As explain in sections 2.2 and 2.3, increase in $\sigma(r_m)$ has both positive and negative effects on the power, though only the negative effect has been identified in the studies done so far (Affleck-Graves and Bradfield (1993); Freeman and Gruemant (2006)). Since $\sigma(r_m)$ has two conflicting effects on the power, the power curve in response to changes in $\sigma(r_m)$, may have power peaks and/or troughs somewhere in-between extreme values of $\sigma(r_m)$. Thus, the power curve (changes in power) on changes in $\sigma(r_m)$ may not be continuously downward slopping as in Affleck-Graves and Bradfield (1993). Therefore, this paper explores the conflicting forces of $\sigma(r_m)$ on the power of the two-pass test.

Theoretical Framework

In hypothesis testing, the power of a test is related to its reliability. “The power [of a test] is the probability that the null hypothesis [H_0] will be rejected given that an alternative hypothesis [H_1] is true” (Campbell et al., 1997, p. 204) (Equation (2)).

$$Power = \frac{\text{Number of times false } H_0 \text{ is rejected}}{\text{Total number of tests conducted when } H_0 \text{ is false } (K)} \quad (2)$$

In Fama and MacBeth (1973) type two-pass test, the null hypothesis (H_0) holds that the CAPM does not exist (that is average market risk premium, $\bar{\gamma}_1$, discussed in section 2.1 is equal to zero). Therefore, if the test is powerful it should often reject false null hypothesis when the CAPM exists in the tested market. On the other hand, if the CAPM is rejected in such powerful test, then it is logical to conclude CAPM does not exist in the tested market. However, the problem is the power of the CAPM tests is often not known due to the fact that in actual settings it is unknown which hypothesis (null or alternative) is true. Thus, nothing can be judged about the power of the

test. Therefore, Affleck-Graves and Bradfield (1993); Elsas et al., (2003); Shanken and Zhou (2007); Grauer and Janmaat (2009) use, simulation analysis with a known alternative hypothesis (H_1), that is, a simulated market where the CAPM exists, to determine the power of CAPM tests.

The Two-Pass Test

Fama and MacBeth (1973) test three conditions to ascertain existence of the CAPM. However, they consider detection of condition three (C3), the existence of positive market premium, as critical (Fama and MacBeth, 1973, p. 624). Therefore, subsequent studies on the CAPM as well as power studies like Affleck-Graves and Bradfield (1993) draw attention only on the C3¹. This version of the two-pass test involves three periods, namely formation, estimation, and testing periods to estimate a time series of market premiums ($\hat{\gamma}_1$) to test statistical significance of $\bar{\hat{\gamma}}_1$. The test conducts as follows.

During Formation Period

Estimate $\hat{\beta}_i$ for each stock (i) by regressing monthly premiums (r_{it}) on monthly market (index) premiums (r_{mt}) with Equation (3) during the formation period (Let j be the duration of formation period).

$$r_{it} = \alpha_i + \beta_i r_{mt} + \varepsilon_{it} \quad (3)$$

Where

$$t = 1, 2, \dots, j.$$

$$i = 1, 2, \dots, \text{up to number of stocks in the market.}$$

Create 20 beta-sorted portfolios based on estimated $\hat{\beta}_i$ in step 1.

¹ In addition, testing only the C3 gives maximum power of the two-pass test since satisfying many conditions is less probable to satisfying few conditions.

During Estimation Period

Estimate $\hat{\beta}_i$ again with Equation (3) for $t = j+1, j+2, \dots, j+g$ (here g is the duration of estimation period). Average new $\hat{\beta}_i$ to get portfolio beta (β_p) while keeping the portfolio membership of stocks as in the formation period in step 2.

During Testing Period

Conduct Cross-Sectional Regression (CSR) for each month in the testing period, $t = j+g +1, j+g +2, \dots, j+g +h$, using Equation (4) and estimate a time series of $\hat{\gamma}_{1t}$ (here h is the duration of testing period).

$$r_{ip} = \alpha_t + \gamma_{1t} \hat{\beta}_p + \varepsilon_{ip} \quad (4)$$

Where

$p = 1, 2, \dots$, up to the number of portfolios formed.

Conduct t-test and determine significance of $\bar{\hat{\gamma}}_1$ with Equation (5).

$$t(\hat{\gamma}_1) = \frac{\bar{\hat{\gamma}}_1}{s(\hat{\gamma}_1) / \sqrt{n}} \quad (5)$$

If calculated t-statistic is significant, the two-pass test concludes the CAPM exists.

Also, the power of the two-pass test depends on the sampling distribution of $t(\hat{\gamma}_1)$. For instance, if distribution of $t(\hat{\gamma}_1)$ is concentrated on a large positive value, which is greater than the critical value, the power will be higher. On the other hand if $t(\hat{\gamma}_1)$ is concentrated around smaller value, the power will be lower. Therefore, if any parameter value affects the distribution of $t(\hat{\gamma}_1)$, the power of two-pass test in respect to that population will change. However, it can be observed from the Equation (5) that t-statistic is completely determined by $\bar{\hat{\gamma}}_1$, $s(\hat{\gamma}_1)$, and n . Thus, any effect of changing parameter values on distribution of $t(\hat{\gamma}_1)$ (power) should affect through $\bar{\hat{\gamma}}_1$ and/or $s(\hat{\gamma}_1)$ when n is fixed. It is shown in the next two sub-sections that increase in the

parameter value of $\sigma(r_m)$ affects both positively and negatively on the distribution of $t(\hat{\gamma}_1)$ (power) in contrast to general understanding that increase in $\sigma(r_m)$ affects only negatively on the power.

Negative Effect of $\sigma(r_m)$ on Power of the Two-Pass Test

As discussed above increasing $\sigma(r_m)$ decreases the value on which $t(\hat{\gamma}_1)$ is concentrated. This has a negative effect on the power. The negative effect is already identified in literature (Affleck-Graves and Bradfield, 1993; and Freeman and Guermat (2006). When $\sigma(r_m)$ is higher $t(\hat{\gamma}_1)$ is concentrated on a smaller value as it is determined by $\bar{r}_m / \sigma(r_m)$ (called signal-to-noise ratio) when n is fixed. Therefore an increase in $\sigma(r_m)$ has a negative effect on the power.

Positive Effects of $\sigma(r_m)$ on Power of the Two-Pass Test

Even though direct, more obvious relation between $\sigma(r_m)$ and power is negative, $\sigma(r_m)$ has indirect positive effects on power. In other words $\sigma(r_m)$ is a member of second category of power determinants too. It is shown below that increasing $\sigma(r_m)$ has two positive effects on the power. Since these effects are involved with estimating $\hat{\beta}_i$, both of positive effects together are referred as the estimation effect.

Recall that estimating $\hat{\gamma}_1$ involves three OLS regressions each in formation, estimation, and testing periods. $\hat{\beta}_i$ is estimated two times each in formation and estimation periods and based on estimated $\hat{\beta}_p$, and $\hat{\gamma}_{it}$ is estimated in the testing period. These estimations may differ from their parameter values due to estimation errors and these differences can be explained by studying distribution of OLS slope coefficient (Equation (7)).

$$\hat{b} \sim N(B, SE(\hat{b})). \quad (7)$$

Where \hat{b} is the slope coefficient estimated from a sample drawn from a population with parameter value B . Equation (8) calculates standard error of \hat{b} , denoted $SE(\hat{b})$, in a regression where $y = a + bx + \varepsilon$, the generic form of the OLS regressions run in formation, estimation, and

testing periods. In Equation (8) x_i , N and ε are the independent variable, sample size, and the error term respectively.

$$SE(\hat{b}) = \sqrt{\frac{\sum_{i=1}^N \hat{\varepsilon}_i^2 / (N - 2)}{\sum_{i=1}^N (x_i - \bar{x})^2}}. \quad (8)$$

As observable from the Equation (8) variables x_i , N and ε determine value of $SE(\hat{b})$. Since \hat{b} corresponds to β_i of Equation (3) and to γ_{it} of Equation (4), corresponding variables for x_i , N and ε in the equations for standard errors of $\hat{\beta}_i$ and $\hat{\gamma}_i$ (not presented) can be used to analyse effects of estimations on power of the two-pass test.

r_{mt} , the independent variable in Equation (3) that is run in both formation and estimation periods, corresponds to x_i in Equation (8). Therefore, the denominator of Equation (8) divided by $N-1$ corresponds to the point estimation of $\sigma(r_m)$. Thus, an increase in $\sigma(r_m)$ decreases the standard error of $\hat{\beta}_i$, denoted $SE(\hat{\beta}_i)$, when other determinants of standard error, $\sum_{i=1}^N \hat{\varepsilon}_i^2$ and N , held constant. In summary, any increase of $\sigma(r_m)$ make $\hat{\beta}_i$ more precise in both formation and estimation periods. This tends to increase the power of the two-pass test due to two chains of reactions.

First, lower the $SE(\hat{\beta}_i)$ in both formation and estimation periods, higher the standard deviation of $\hat{\beta}_p$ among portfolios ($\sigma(\beta_p)$) in the estimation period. This is because when the $SE(\hat{\beta}_i)$ is lower, the $\hat{\beta}_i$ distributes more closely to its true value in both formation period and estimation periods. Hence, assignment of portfolio membership on beta-sorted basis is more precise. The enhanced precision preserves the purpose of forming beta-sorted portfolios (increasing $\sigma(\beta_p)$) to estimation period. This can be elaborated as follows.

$$\hat{\beta}_i = \beta_i + e_i \quad (9)$$

Where $e_i \sim N(0, SE(\hat{\beta}_i))$

$\hat{\beta}_i$ can be high when β_i is high or/and e_i is positive. Therefore, if $\hat{\beta}_i$ are sorted high to low and formed beta-sorted portfolios, high-beta (low-beta) portfolios have both high-beta stocks (low-beta) and stocks with positive e_i (negative e_i). However e_i is random and has expected value of zero. If $\hat{\beta}_i$ are re-estimated in the estimation period, when the portfolio membership of stocks is fixed, high-beta (low-beta) portfolio's beta ($\hat{\beta}_p$) decrease (increase) and range of $\hat{\beta}_p$ in the estimation period is lesser than that could be achieved by forming portfolios sorting on β_i . In other words, the standard deviation among $\hat{\beta}_p$, denoted (σ_{β_p}), is lesser than that among portfolios formed sorting on β_i . In addition, larger the $SE(\hat{\beta}_i)$ higher is the possibility of assigning lower-beta (higher-beta) stocks into high-beta (low-beta) portfolios in the formation period, making the σ_{β_p} further smaller. Accordingly higher the $SE(\hat{\beta}_i)$, smaller is $\sigma(\beta_p)$ in the estimation period. As $\hat{\beta}_p$ is the independent variable of the CSR run with Equation (5) in Step 4, the increased $\sigma(\beta_p)$ in the estimation period decreases standard error of estimated market premium in the testing period, denoted ($SE(\hat{\gamma}_1)$). Since $s(\hat{\gamma}_1)$ is a function of both σ_m and an estimation error, reduced $SE(\hat{\gamma}_1)$ decreases $s(\hat{\gamma}_1)$ leading to higher power by increasing $t(\hat{\gamma}_1)$.

In summary $\uparrow \sigma(r_m) \rightarrow \downarrow SE(\hat{\beta}_i) \rightarrow \uparrow \sigma(\beta_p) \rightarrow \downarrow SE(\hat{\gamma}_1) \rightarrow \downarrow s(\hat{\gamma}_1) \rightarrow \uparrow t(\hat{\gamma}_1) \rightarrow \uparrow \text{power}$.

Second positive effect of increasing $\sigma(r_m)$ is fading out the errors-in-variables (EIV) problem which underestimates $\bar{\gamma}_1$. Higher the $\bar{\gamma}_1$, higher is the $t(\gamma_1)$, ultimately leading to increase power (Equation 5). The EIV is a regression incident in which the slope coefficient is underestimated.

In the OLS regression, the dependent variable (y) distributes normally for any given value of independent variable (x). Thus the observed value of y (denoted as y') can differ from the actual value of y for given x due to random reasons (v). The EIV problem occurs when the estimated value of x is used instead of its actual value as in the case of CSR which is run in the testing

period. In this case, in addition to v , another random error term (u) the estimation error of x , adds to the regression model. If x' is the estimated value of x , the model becomes Equation (10).

$$y' = B(x' - u) + v \quad (10)$$

Where

$$u \sim N(0, \sigma^2(u))$$

$$v \sim N(0, \sigma^2(v))$$

Greene (2003) proves when this happens, the estimated value of B , the \hat{b} converge in probability to Equation (11).

$$\text{plim } \hat{b} = B \left(\frac{Q}{Q + \sigma^2(u)} \right) \quad (11)$$

Where

$$Q = \text{plim}(1/n) \sum_i^n x_i^2$$

So estimator \hat{b} is biased towards zero. As observable in Equation (11), the EIV problem can be reduced by either increasing Q (the variance or standard deviation of independent variable, x or decreasing the measurement error of x , $\sigma^2(u)$). In the CSR of the two-pass test, $\hat{\beta}_p$ is the independent variable and $SE(\hat{\beta}_p)$ is the estimation error of the independent variable, $\hat{\beta}_p$. As elaborated above, since $\sigma(r_m)$ and $SE(\hat{\beta}_i)$ have a negative relation, the EIV problem reduces as $\sigma(r_m)$ increases. As a result the $\bar{\gamma}_1$ (which corresponds to \hat{b} in Equation (11)) rises improving the t-statistic. Therefore, this affects positively on the power of the test in addition to the above positive effect.

In summary $\uparrow \sigma(r_m) \rightarrow \downarrow SE(\hat{\beta}_i) \rightarrow \downarrow SE(\hat{\beta}_p) \rightarrow \uparrow \bar{\gamma}_1 \rightarrow \uparrow t(\hat{\gamma}_1) \rightarrow \uparrow \text{power}$.

To sum up the discussion, $\sigma(r_m)$ has a both positive and negative effect on the power of the two-pass test. As $\sigma(r_m)$ increases signal-to-noise ratio effect (negative effect) increases and the

estimation effect (positive effect) decreases. Therefore it is difficult to conclude an increase in $\sigma(r_m)$ reduces or improves the power. If the positive effects overruns the negative effect of increasing $\sigma(r_m)$, the power increases or vice-versa. If at all levels of $\sigma(r_m)$, the positive and negative effects exactly offset, there is no consequence of increasing $\sigma(r_m)$ on the power. However if one effect (either positive or negative) overruns the other after some level of $\sigma(r_m)$, the power curve in response to changes in $\sigma(r_m)$ shows a peak or trough. However, findings of Affleck-Graves and Bradfield (1993) are not consistent with this argument and they report the power reduces continuously when the standard deviation of market premium increases. Therefore this paper explores conflicting (positive and negative) forces of $\sigma(r_m)$.

Methodology

To explore behavior of power in respect to changes in $\sigma(r_m)$ the study creates an artificial stock market where the CAPM holds true.

The artificial market is created based on Affleck-Graves and Bradfield (1993) methodology where the mean market premium (\bar{r}_m), the standard deviation market premium ($\sigma(r_m)$), the standard deviation of beta across stocks ($\sigma(\beta_i)$), the standard deviation of error term (unsystematic risk) ($\sigma(e)$) are used as parameters on which the market is defined. To measure the power, the artificial market is regenerated (iterated) for 10000 times on a given set of parameters and count number of times the CAPM is detected by the two-pass test out of 10000 (see appendix I for the simulation steps). Then value for $\sigma(r_m)$ in the simulated market is changed for a range of possible values and power of the two-pass test is measured at each level of $\sigma(r_m)$. The power curve is derived by plotting power levels against corresponding values of $\sigma(r_m)$. Final result of this process is a set of paired observations on $\sigma(r_m)$ and power level of the two-pass test. For reference purposes, the 10000 iterations run on a given set of parameters is called a simulation run and a set of simulation runs which differ from each other only on their value for $\sigma(r_m)$ is called a series of simulation. In other words the power curve is drawn on the results of a series of simulations.

Matlab® 2006 version is used to run the simulation. The scripts (M-files) written for executing the simulation steps are available from authors on request. The simulations use an arbitrarily selected fixed base (a random seed) as starting point for random number generation. The use of common random seed enhances comparability of results as the change of power in any two simulation runs is strictly due to the difference in $\sigma(r_m)$ as random numbers are the same for both simulation runs. Further, the use of random seed enables a researcher regenerate findings using the scripts, if required.

The results of simulations are reported as tables. A table presents the average of estimated average market premiums ($\bar{\hat{\gamma}}_1$), the average of t-statistics ($\overline{t(\hat{\gamma}_1)}$), the average of coefficients of determination ($\overline{R^2}$), and the power of the test at 5% and 1% significant levels for each simulation run.

If the power curve in response to changes in $\sigma(r_m)$ is flat, it means the positive and negative effects are same in magnitude. If positive and negative effects do not change in the same magnitude as $\sigma(r_m)$ change, then the power curve on $\sigma(r_m)$ has peak(s) and/or trough(s). On the other hand, if the curve is continuously decreasing, the existing conception that $\sigma(r_m)$ has only a negative effect is reassured. Therefore, any shape of the power curve other than that of continuously decreasing supports the argument that $\sigma(r_m)$ has both positive and negative effects on the power of the two-pass test. To further strengthen this argument, the study conducts two more series of supportive simulations in addition to the main simulation explained above. Details of these two simulations are discussed after presenting results of the main simulation.

Tests for Detecting Peaks and Troughs of the Power Curve

To test the existence of peaks and troughs, the z-tests for the difference of population proportions are conducted using Equation (12).

$$z = \frac{P_A - P_B}{\sqrt{P_{A+B}(1 - P_{A+B})\left(\frac{1}{K_A} + \frac{1}{K_B}\right)}} \quad (12)$$

Here P_A and P_B are powers at two levels of $\sigma(r_m)$. P_{A+B} is the average power of two levels. K denotes the number of iterations (it is 10000 in this study). The rest of this section discusses the process of detecting peaks and troughs using Equation (12).

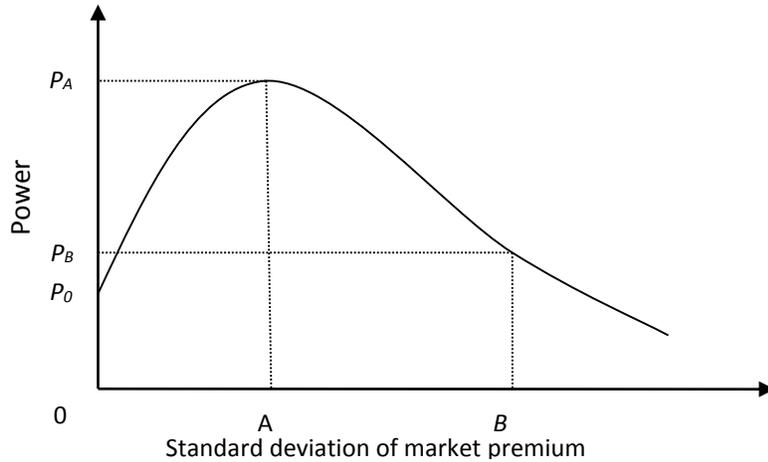


Figure01: Test for peaked power curve

The horizontal axis in the Fig. 1 and Fig. 2 is $\sigma(r_m)$. If the power curve is peaked like in the Fig. 1, the power levels in between the values 0 and A of $\sigma(r_m)$ are compared with power at the peak (P_A). The purpose is to ascertain whether there is a significant increment of power from point 0 through A. Similarly, the power on points after A is compared with P_A . That is the power levels like P_B are compared with P_A to detect any significant power drop on the way to B from A. If the power curve behave like in Fig. 2, that is a trough is visible, the power levels at values between 0 and A, A and B are compared with P_A to determine whether the trough has significant power drop.

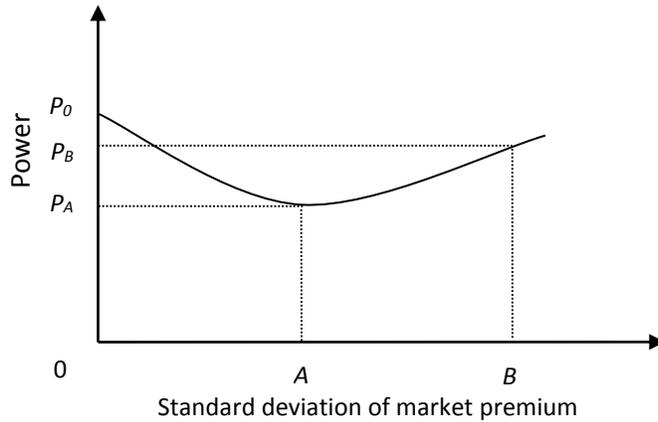


Figure 02: Test for trough power curve

Initial Parameter Values

Affleck-Graver and Bradfield (1993) use parameter values from existing literature to create simulated market. The main disadvantage of their approach is that, not all parameters correspond to a single period. The situation becomes worse as power of two-pass test is sensitive to changes in parameter values like \bar{r}_m , which change from time to time. Another approach is found in Elsas et al., (2003), where they use 15-year monthly returns to calculate parameters eliminating the above problem. Therefore, this research uses an adopted version of Elsas et al., (2003) to estimate initial parameter values.

The study draws a sample from the Colombo Stock Exchange (CSE) to exemplify the real world context. Trading statistics in Data Library of CSE are used to calculate return for each stock from 1st of July 1996 to the 30th of June 2006 and returns on All Share Price Index for the same period are used as a proxy for market return. The mean market premium (\bar{r}_m) is calculated as simple average of returns on the All Share Price Index in excess of three months treasury bill rate. $\sigma(r_m)$ is calculated as the standard deviation of premiums on the All Share Price Index. Beta coefficient for each stock is estimated by running the market model, in Equation (13).

$$r_{it} = \alpha_{it} + \beta_{it}r_{mt} + e_{it} \text{ Where } t = 1, 2, \dots, 120 \quad (13)$$

Then the standard deviation of beta across stocks ($\sigma(\beta_i)$) is estimated by calculating the standard deviation of betas estimated using Equation (13). The standard deviation of error term (unsystematic risk) ($\sigma(e)$) is the standard deviation of e_{it} components generated in running the market model (Equation 13). Further the mean beta ($\bar{\beta}$) of the simulated market is considered as one (1) in accordance with the CAPM. Table 1 presents the parameter values estimated.

Table 01: Initial parameter values extracted from the CSE

\bar{r}_m	$\sigma(r_m)$	$\sigma(\beta_i)$	$\bar{\beta}$	$\sigma(e)$
0.00423	0.07509	0.35990	1	0.14098

Source: Calculate based on CSE trading statistics 1996-2006

In addition to these parameters, duration of the portfolio formation, beta estimation and testing periods are set at 4, 5, and 5 years in length respectively, guided by Affleck-Graves and Bradfield (1993). Further, following Affleck-Graves and Bradfield (1993); and Elsas et al., (2003) the artificial stock market consists of 100 stocks.

Details of Simulations Conducted and Findings

To derive a power curve in response to changes in $\sigma(r_m)$, seven separate simulation runs are conducted by changing value of $\sigma(r_m)$ around its value in the base simulation (that is $\sigma(r_m) = 0.0751$). The range of values used is arbitrary. However, a wider range is preferred as to observe the behavior of the power curve for a wider range of $\sigma(r_m)$. The values used in the seven simulation runs (including the value in the base simulation) are 0.0100, 0.0250, 0.0500, 0.0751, 0.1250, 0.1750, and 0.2250.

Power and Standard Deviation of Marker Premium

Table 2 and Fig. 3 present results of the simulation. In Table 2 and Fig. 3 power of the two-pass test is very low at near-zero values of $\sigma(r_m)$ and starts climbing as $\sigma(r_m)$ increases until it reaches a peak. Even though, the observed behavior of the power curve is in accordance with the arguments of this paper, to conclude the existence of statistically significant peak, that is low

power at both extremes of $\sigma(r_m)$, power levels at different values of $\sigma(r_m)$ are compared with the peak power using the z-test.

Table02: Power under different levels of the standard deviation of the market risk premium

An artificial stock market is created as to hold the CAPM by extracting parameter values from the CSE. Then, the two-pass test is performed on the market. This process is iterated 10 000 times and the power is defined as the number of times the CAPM is detected divide by 10 000. To ascertain the shape of the power curve in response to changes in the standard deviation of market premium (σ_m), the value for parameter σ_m is swung around its estimated value ($\sigma_m=0.0751$) and the simulation is repeated. Powers at different values of σ_m are compared with the peak power of each power column using z-test for the difference of population proportions (power) and the significance of differences are reported.

$\sigma(r_m)$	$\hat{\gamma}_1$	$t(\hat{\gamma}_1)$	$\overline{R^2}$	Power ^b	
				0.01	0.05
0.0100	0.0002	0.0460	0.0526	0.0090**	0.0495**
0.0250	0.0012	0.2090	0.0536	0.0117	0.0527**
0.0500	0.0030	0.3724	0.0764	0.0131	0.0645 ^d
0.0751 ^a	0.0037	0.3489	0.1285	0.0147 ^c	0.0614
0.1250	0.0039	0.2361	0.2504	0.0111*	0.0561*
0.1750	0.0039	0.1710	0.3543	0.0098**	0.0534**
0.2250	0.0038	0.1324	0.4362	0.0096**	0.0513**

$\hat{\gamma}_1$, $t(\hat{\gamma}_1)$, and $\overline{R^2}$ are averages of 10 000 iterations of estimated risk premiums, t statistics in the t-test, and average coefficients of determination of CSRs in each iteration respectively.

^a Estimated level of the standard deviation of market premium in the CSE (the base simulation).

^b The power column reports the proportion of times the CAPM is detected when it is true out of 10 000 iterations.

^c The peak power at 1%

^d The peak power at 5%

** significantly different from the peak power at 1% (two tailed z-test)

* significantly different from the peak power at 5% (two tailed z-test)

According to Table 2, a significant power drop is evident at very low and very high volatility of market premium. This relationship is clearer at 5% of the two-pass test in Fig. 3. At 5% the peak power is 0.0645, which corresponds to 0.05 of $\sigma(r_m)$ and the peak at 1% is 0.0147 which is at 0.0751 of $\sigma(r_m)$. Further as observable from the Table 2 both power columns, at 5% and 1%, have significant power drops above and below the peak clearly showing existence of conflicting forces of $\sigma(r_m)$ on power of the two-pass test. Previous researchers like Affleck-Graves and Bradfield (1993) do not identify this peaked shape of the power curve. In their study, power level

continuously decreases as $\sigma(r_m)$ increases. A possible explanation on why existing studies do not find peaked power curve on $\sigma(r_m)$, is that they may start their analysis after the turning (peak) point (for example $\sigma(r_m) = 0.05$ of 5% line), thus they may be unable to identify the dual effect of $\sigma(r_m)$. However, the implication of this finding is that uncovering a higher possibility of rejecting the CAPM in a low volatile market even when it exists, in contrary to common understanding that the two-pass test is more robust in a less volatile market.

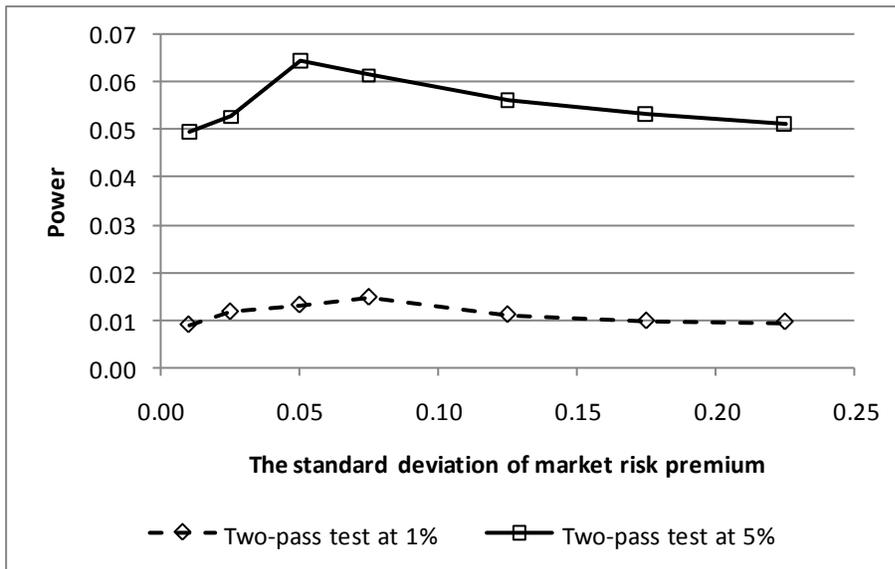


Figure03: Power results for different standard deviation of marker risk premium

Supportive Simulations

Previous section presented the shape of the power curve in response to changes in $\sigma(r_m)$. However, it does not elaborate how the shape emerged. Recall that $\sigma(r_m)$ has conflicting effects on the power; the effect of estimation (the positive effects of σ_m) and the signal-to-noise ratio effect (the negative effect of $\sigma(r_m)$). Thus, the power levels derived in the main series of simulations are the net effect of these two forces, estimation (positive) effect and signal-to-noise ratio (negative) effect. The signal-to-noise ratio effect can be isolated by using actual β_i in the step 2 of simulation procedure instead of $\hat{\beta}_i$ (estimated beta). Since the use of actual β_i

eliminates all effects associated with beta estimation (the estimation/positive effect) leaving only the signal-to-noise ratio/negative effect.

The test, which performs with actual beta, is called two-pass test with actual beta (the first series of supportive simulations). In this series of simulations, beta estimation does not take place, thus the results presented in Table 3, depicted in Fig. 4 show how changes in $\sigma(r_m)$ affect power when the estimation problem does not exist, and only the signal-to-noise ratio effect (the negative effect of $\sigma(r_m)$) is prevailed. Therefore, continuously downward sloping power curve is expected in this series of simulation.

Table03: Power under different levels of the standard deviation of the market premium using actual beta

An artificial stock market is created as to hold the CAPM, by extracting parameter values from the CSE. Then, the two-pass test is performed using actual beta of each stock, instead of estimated beta. Here actual beta means, the beta values generated in the step 2 of the simulation procedure. This process is iterated 10 000 times and the power is defined as number of times the CAPM is detected divide by 10 000. To ascertain the shape of the power curve in response to changes in the standard deviation of market premium (σ_m), when actual beta is used, the value for parameter σ_m is swung around its estimated value ($\sigma_m=0.0751$) and the simulation is repeated. Thus, the table presents the power when the estimation effect is not present.

$\sigma(r_m)$	$\hat{\gamma}_1$	$t(\hat{\gamma}_1)$	\bar{R}^2	Power ^b	
				0.01	0.05
0.0100	0.0041	0.7935	0.0559	0.0342	0.1201
0.0250	0.0041	0.6891	0.0701	0.0269	0.1042
0.0500	0.0041	0.5029	0.1130	0.0195	0.0763
0.0751 ^a	0.0041	0.3753	0.1679	0.0143	0.0660
0.1250	0.0040	0.2393	0.2784	0.0111	0.0567
0.1750	0.0039	0.1720	0.3727	0.0099	0.0536
0.2250	0.0039	0.1328	0.4486	0.0095	0.0519

$\hat{\gamma}_1$, $t(\hat{\gamma}_1)$, and \bar{R}^2 are averages of 10 000 iterations of estimated risk premiums, t statistics in the t-test, and average coefficients of determination of CSRs in each iteration respectively.

^a Estimated level of the standard deviation of market premium in the CSE.

^b The power column reports the proportion of times the CAPM is detected when it is true out of 10 000 iterations at 1% and 5% significance levels.

From Table 3, it is observable that power peaks at the beginning ($\sigma(r_m)=0.01$) and continuously decreases as $\sigma(r_m)$ increases confirming negative relation between σ_m and the power as expected. This is because, as $\sigma(r_m)$ increases, the t-statistic decreases (as observable in the Table

3), reducing the possibility of rejecting null hypothesis that the CAPM does not exist. The drop of power in the test with actual beta becomes gentler as $\sigma(r_m)$ continues to increase diminishing the gap between two curves. By the point 0.0751 power curve of two tests become almost identical. That is, there is no advantage of using actual beta when $\sigma(r_m)$ is large. In other words, the estimation effect, including widely cited EIV problem, becomes less significant issue in the two-pass test of the CAPM in a market where market premium is highly volatile.

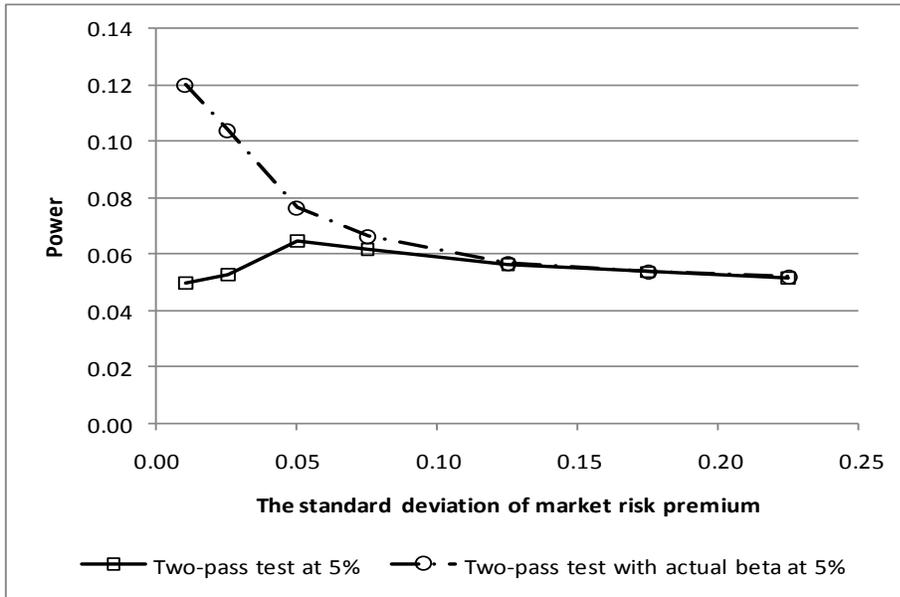


Figure 04: Comparison of power under different levels of standard deviations of market premium with estimated beta and actual beta

However, $\hat{\gamma}_1$ values in Table 2 are underestimated and the underestimation becomes severe as decreasing σ_m , indicating the presence of the EIV problem. Since the EIV problem is not present in the two-pass test with actual beta (in Table 3), $\hat{\gamma}_1$ in Table 3 should be equal to its parameter value 0.0043. However in the Table 3, $\hat{\gamma}_1$ are still lesser than 0.0043 though the EIV problem is not there. The only way this can be happened is the bias introduced by the fixed random seed. The bias may be positive or negative, in this case, the bias is negative as Table 3 underestimates

$\hat{\gamma}_1$. To show validity of this argument, Table 4 re-generates the results of Table 3 without a fixed random seed for generating random numbers (the second series of supportive simulations).

Table 04: Power under different levels of the standard deviation of the market premium using actual beta and without the random seed

An artificial stock market is created as to hold the CAPM, by extracting parameter values from the CSE. Then, the two-pass test is performed using actual beta of each stock, instead of estimated beta. Here actual beta means, the beta values generated in step 2 of simulation procedure. This process is iterated 10 000 times and the power is defined as the number of times the CAPM is detected divide by 10 000. To ascertain shape of the power curve in response to changes in the standard deviation of market premium (σ_m), when actual beta is used, the value for parameter σ_m is swung around its estimated value ($\sigma_m = 0.0751$) and the simulation is repeated. The difference between Table 3 and this table is the usage of random seed. In Table 3 the same random seed, that is, the same set of random numbers are used for all simulation runs at different values of σ_m , in order to enhance the comparability. This may cause a common bias on all simulation results in Table 3. Thus, the premium may be constantly under or over estimated. Hence, to show that the EIV problem does not prevail (so that estimated premium is equal to parameter value, 0.0043) when actual betas are used, the results of this table has been generated by using an arbitrary random seed for each simulation run.

$\sigma(r_m)$	$\hat{\gamma}_1$	$t(\hat{\gamma}_1)$	\bar{R}^2	Power ^b	
				0.01	0.05
0.0100	0.0042	0.8127	0.0559	0.0358	0.1288
0.0250	0.0042	0.7051	0.0701	0.0269	0.1069
0.0500	0.0044	0.5336	0.1134	0.0185	0.0774
0.0751 ^a	0.0044	0.4028	0.1679	0.0154	0.0656
0.1250	0.0039	0.2351	0.2783	0.0117	0.0543
0.1750	0.0043	0.1876	0.3724	0.0109	0.0563
0.2250	0.0043	0.1480	0.4488	0.0111	0.0506

$\hat{\gamma}_1$, $t(\hat{\gamma}_1)$, and \bar{R}^2 are averages of 10 000 iterations of estimated risk premiums, t statistics in the t-test, and average coefficients of determination of CSRs in each iteration, respectively.

^a Estimated level of the standard deviation of market premium in the CSE.

^b The power column reports the proportion of times the CAPM is detected when it is true out of 10 000 iterations.

To isolate the estimation effect the signal-to-noise ratio should be fixed. However, it is impossible to isolate the estimation effect, as the signal-to-noise ratio cannot be controlled without affecting another power determinant, \bar{r}_m , while $\sigma(r_m)$ is changing as the signal-to-noise ratio is defined as $\bar{r}_m / \sigma(r_m)$.

However, the estimation (positive) effect is observable in Fig. 4 where power of the two-pass test with actual beta continuously decreases as the $\sigma(r_m)$ increases, but the power curve of two-pass

tests increases up to a certain point ($\sigma(r_m)=0.05$) even though the power of the two-pass test with actual beta decreases dramatically. Even after this point, negative slope of the power curve of two-pass tests is gentler than that of the two-pass tests with actual beta. Together, both of these observations imply the existence of opposite force that has an increasing (positive) effect on the power of the two-pass test as $\sigma(r_m)$ increases.

Conclusion

Though many tests of the CAPM have been developed, the two-pass test is widely used. Results of the two-pass test are unsupportive of the CAPM in general. However, few attempts have been taken to measure the reliability (power) of the test. Therefore, the evidence (results) generated from a test with unknown power cannot be taken as conclusive as not detecting the CAPM may be due to lack of power in the test rather than non-existence of the CAPM in the tested market. In addition, the attempts to measure the power of the two-pass test reports some contradictory evidence as opposed to theoretical relationships with power determinants. The research addresses such an inadequately addressed power determinant of the two-pass test, the standard deviation of market risk premium. The standard deviation of market premium creates two opposite forces on the power. One, by affecting precision of beta estimation and, the second, is affecting signal-to-noise ratio.

The study uses series of simulations to detect the behavior of the power curve in response to changes in the standard deviation of market risk premium. The results show the power curve is peaked at the middle showing existence of both positive and negative effect by the standard deviation of market premium. However existing literature reports only a continuously decreasing power curve. The increasing power at low level of the standard deviation of market premium is a unique contribution of this study. The implication of this finding is that the higher possibility of rejecting the CAPM in less volatile market even when it exists. This is somewhat contradictory with general understanding that the CAPM can be easily captured in less volatile markets.

References

- Affleck-Graves, J. F., and Bradfield, D.J. (1993). An Examination of the Power of Univariate Tests of the CAPM: A simulation Approach. *Journal of Economics and Business*, 45: 17-33.
- Black, F. (1972). Capital market equilibrium with restricted borrowing. *Journal of Business*, 45(3): 444-455.
- Campbell, J.Y., Lo, A.W., and Mackinlay, A.G. (1997). *The econometrics of financial markets*, New Jersey: Princenton university press.
- Elsas, R., El-Shaer, M., and Theissen, E. (2003). Beta and Return Revisited: Evidence from the German Stock Market. *Journal of International Financial Markets, Institutions and Money*, 13: 1-18.
- Fama, E., and French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33:3-56.
- Fama, E., and MacBeth, J. (1973). Risk, return and equilibrium. *Journal of Political Economy*, 81(3): 607-636.
- Freeman, M.C. and Gruemant, C. (2006). The conditional relationship between beta and return: A reassessment, *Journal of Business Finance and Accounting*, 33(7 & 8): 1213-1239.
- Gibbons, M.R., Ross, S.A., and Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica*, 57(5): 1121-1152.
- Grauer, R. R. and Janmaat, J. A. (2009). On the power of cross-sectional and multivariate tests of the CAPM. *Journal of Banking & Finance*, 33: 775–787.
- Greene, W.H. (2003). *Econometric Analysis*, 5th ed, Pearson Education and Dorling Kindersley (India), New Delhi.
- Lintner, J. (1965). The valuation of risky assets and the selection of risky Investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47(1): 13-37.
- Shanken, J. and Zhou, G. (2007), “Estimating and testing beta pricing models: Alternative methods and their performance in simulations”, *Journal of Financial Economics*, 84(1): 40-86.
- Sharpe, F. W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19 (3): 425-442.