# Low R square in the cross section of expected returns

Benito Sanchez, Ph. D.
College of Business and Public Management
Kean University
1000 Morris Av.
Union, NJ 07018
Email: bsanchez@kean.edu

## I. Introduction

The capital asset pricing model (CAPM) [Sharpe, 1964; Lintner, 1966] predicts an exactly linear relationship between returns and risk of capital. According to the CAPM, the expected return on any risky security or portfolio of risky securities can be measured by the risk-free rate and the market risk premium multiplied by the beta coefficient (asset risk). This implies that the cross-sectional expected return is linear in beta, being this relationship called the security market line (SML).

However, Fama and French (1992, 1993) find that the cross-sectional variation on expected returns cannot be explained by the beta alone. Particularly, the relation return-beta should yield high cross-sectional $R^2$ if the SML hold. Nevertheless, studies have found that cross sectional R squares ($R^2$) are much lower than time series R squares (for example, Fama-French, 1992; Reinganum, 1981; Lakonishok and Shapiro, 1986). Specifically, the average time series $R^2$ obtained from regressions of excess returns on excess market returns are higher than cross-sectional *R2* obtained from regression of return on betas.

After Fama and French's study, many articles have addressed the relatively low explanatory power of beta in the cross-section of expected return. Roll and Ross (1994) explain this phenomenon happen because of a choice a "wrong" index. They show that there exist some indices that make *true* betas not to have any relationship (zero correlation) with true expected returns. They demonstrate that there are indices that allow calculating true betas having no relationship with true expected returns. Those indices lie "within a set of whose boundaries and can be directly calculated with basis parameters expected returns and covariance of returns)" (Roll and Ross, 1994). They further conclude that "the cross-sectional OLS relation is very sensitive to the choice of an index and indices can be quite close to each others and to the mean-variance frontier and yet still produce significantly different cross-sectional slopes, positive, negative or zero" (Roll and Ross, 1994).

Other possible explanation for the low cross-sectional R2 is error in variable (EIV) problem. Cross sectional regression between betas and returns requires an estimation of betas. The betas are estimated with sampling errors from time series data, which cause the EIV problem. Kim (1995) proposes a correction for EIV for betas and finds that betas do have more explanatory power than size and book-to-market, concluding that "market beta has economically and statistically significant explanatory power for average stock returns both in the presence and absence of the firm size variable".

Kothari, Shanken and Sloan (1995) argue that survivorship bias and beta mis-measurement could explain Fama-French findings of no relationship between betas and expected returns in the cross-section. They claim that the observed explanatory power of book-to market factor (BTM) is likely due to survivorship bias: many of the firms that are excluded from Compustat are firms that failed and it is likely that these firms had high book-to-market ratio and low returns; adding these firms to the database should reduce the explanatory power of book-to-market characteristics. Moreover, they argue that annual betas are more appropriate than monthly betas because the investment horizon for a typical investor is probably closer to a year than a month and show that the relation between beta and return is stronger when betas are estimated using annual returns.

Pettengill, Sundaram, and Mahur (1995) argue that the phenomenon is due to the use of ex-post data to make inference about ex-ante data and therefore, the relationship between beta and realized returns varies from the relationship between beta and expected return, which means researchers should use a time-varying betas in their estimation. Pettengill et al (1995) propose a conditional version of the SLM model, in which a conditional relationship between beta and return exist. During periods of "up" markets where the realized market return exceeds the risk-free return, there should be a positive relationship between beta and return; whereas during periods of "down" markets where the risk-free return exceeds the realized market return there should be negative relationship. After correcting for "ups" and "downs" of the market, they support a linear relation between betas and expected returns.

My explanation to the phenomenon is related to the fact that market variance is higher than the variance of expected return. I derive a simple model that relates cross-sectional $R^2$ with time series $R^2$. The difference between market variance and expected return determines the observed difference between time series $R^2$ and cross-sectional $R^2$. When market variance equals the expected return, cross-sectional $R^2$ will be exactly as the time series $R^2$. However, this is not the case for most of the period January 1927 – December 2004.

I further show that in order for the expected return to have a similar magnitude of the average market variance (that is, cross-sectional $R^2$ equals to time series $R^2$), both high variance of beta and high expected market risk premium is needed. Thus, only when there is low market variance, high variance of beta, and high expected market risk premium, the cross-sectional $R^2$ would be similar (could be also higher) to time series $R^2$.

The remainder of this paper is divided as follow. Section II shows the empirical model I use to explain the relationship between cross sectional $R^2$ and time series $R^2$. It also shows the estimation procedure to obtain the different statistics to be used in the model. Section III describes the data. Section III presents and discusses the empirical results applying the model to both securities and portfolios. Section IV concludes.

## II. An Empirical Model of the Cross-Section of Expected Returns

This section presents a model that relates cross-sectional $R^2$ with time series $R^2$ and summarizes the error in variable problem.

Consider a CAPM-like economy composed of $n$ risky assets, where the market return $r_{m,t}$ at date=$t$ is equally weighted, $r_f$ is the fixed return on a risk free asset, and $r_{it}$ is the date=$t$ return on asset $i$. The empirical model for returns is,

$$r_{i,t} = r_f + \beta_i \left(r_{m,t} - r_f\right) + \varepsilon_{i,t} \text{ where} \tag{1}$$

$$\beta_i \sim id\ N\left(1, \sigma_\beta^2\right), \tag{1a}$$

$$r_{m,t} \sim iid\ N\left(E_m, \sigma_m^2\right), \tag{1b}$$

$$\varepsilon_{i,t} \sim iid\ N\left(0, \sigma_\varepsilon^2\right), \tag{1c}$$

$$\sum_{i=1}^{n} \varepsilon_{i,t} \equiv 0, \text{ and } r_{m,t} \equiv \sum_{i=1}^{n} r_{i,t} \Big/ n \tag{1d, 1e}$$

Beta $\beta_i$ is uncorrelated with other asset betas and has an average beta equal to one. The return on the market $r_{m,t}$ is uncorrelated with its past and is distributed about its average $E_m$. The asset specific term $\varepsilon_{i,t}$ is uncorrelated with its past and other assets idiosyncratic components. By definition, aggregate firm specific risk must vanish, at each date, which guarantees the average of $n$ asset returns is market return. Finally, the random variables $\beta_i$, $\varepsilon_{i,t}$, and $r_{m,t}$ are assumed to be uncorrelated with each other.

*A. Time Series and Cross-Sectional R2s*

I compute overall $R^2$s for the CAPM using average total variances from the cross section and time series. First, I derive the $R^2$s under the true model. For the times series $R^2$ take each firm's time series variance and average that over firms:

$$\sigma_T^2 = \sum_{i-1}^{n} \left(\beta_i^2 \sigma_m^2 + \sigma_\varepsilon^2\right) n^{-1} = \sum_{i-1}^{n} \left(\beta_i^2\right) n^{-1} \sigma_m^2 + \sigma_\varepsilon^2 = \left(\sigma_\beta^2 + 1\right)\sigma_m^2 + \sigma_\varepsilon^2 \tag{2}$$

For the cross-sectional $R^2$ take the cross sectional variance at each date and average over time:

$$\sigma_N^2 = \sum_{t=1}^{T} \left(\left(r_{m,t} - r_f\right)^2 \sigma_\beta^2 + \sigma_\varepsilon^2\right) T^{-1} = \sum_{t=1}^{T} \left(\left(r_{m,t} - r_f\right)^2 T^{-1} \sigma_\beta^2 + \sigma_\varepsilon^2\right) = \left[\sigma_m^2 + \left(E_m - r_f\right)^2\right]\sigma_\beta^2 + \sigma_\varepsilon^2 \tag{3}$$

It is easiest to see the relation between the $R^2$s by noticing,

$$\left(1 - R_T^2\right) = \frac{\sigma_\varepsilon^2}{\sigma_T^2},\ \left(1 - R_N^2\right) = \frac{\sigma_\varepsilon^2}{\sigma_N^2},$$

$$\sigma_T^2 - \sigma_N^2 = \sigma_m^2 - \left(E_m - rf\right)^2 \sigma_\beta^2, \text{ and} \tag{4}$$

$$\frac{1}{\left(1 - R_T^2\right)} - \frac{1}{\left(1 - R_N^2\right)} = \frac{\sigma_T^2 - \sigma_N^2}{\sigma_\varepsilon^2} = \frac{\sigma_m^2 - \left(E_m - rf\right)^2 \sigma_\beta^2}{\sigma_\varepsilon^2}.$$

Solving for the cross sectional $R_N^2$ yields,

$$R_N^2 = 1 - \frac{\left(1 - R_T^2\right)}{1 - \frac{\sigma_m^2 - \left(E_m - rf\right)^2 \sigma_\beta^2}{\sigma_\varepsilon^2}\left(1 - R_T^2\right)} \tag{5}$$

A key assumption of the model is that cross sectional variance of residual is equal to the time series average variance of residual. This variance of residual is the diversifiable risk (unsystematic risk), that in the limit –when the number of stocks in the portfolio increases– will be a constant number. In a balanced panel data the average time series variance residual is exactly equal to the cross-sectional variance residual.

Most important, the *expected* cross-sectional $R^2$ depends on the market volatility, the variance of beta, market risk premium, variance of the residual, and past time-series $R^2$. Note in equation 4 that for the cross-sectional $R^2$ to be equal to the time series $R^2$, the expected squared market risk premium times variance of beta –which is the variance of expected return– must be equal to the variance of market return. However, if the magnitude of the market risk premium squared time variance of beta is low relative to the market variance, the net results would be low cross-sectional $R^2$. Therefore, I investigate whether market volatility plays an important role in explaining the observed low cross-sectional $R^2$.

## *B. Expected return and market return variance*

To explain the marked differences between time series and cross sectional $R^2$s, market returns must be more volatile than average returns. Portfolio theory does not provide a definitive answer to the question of whether $\sigma_m^2$ must be greater than $\sigma_E^2$. Theory can show that market variances will be greater than average returns variance, but only given what seem to be the natural features of asset returns and a particular tangency portfolio choice.

First, for an arbitrary set of portfolio weights, portfolio variance depends on returns covariance not expected returns. This follows immediately from inspection of the portfolio variance formula, $\sigma_p^2 = w'\Sigma w$ where $\Sigma$ is the covariance matrix on *n* assets and *w* is a *n*-dimensional vector of portfolio weights, which does not explicitly depend on expected returns.

Mean-variance optimal portfolio weights, however, are not arbitrary and do depend on expected returns but the position of a candidate tangency portfolio depends on the highest Sharpe ratio for a given reference asset. Tangency portfolio variance is[1],

$$\sigma_c^2 = w_c{}'\Sigma w_c$$

$$\text{where } w_c = \frac{\Sigma^{-1}(E_n - c)}{1'z}$$

is a set of optimal weights maximizing the Sharpe ratio for a given reference point *c*, $E_n$ is a vector of expected returns, $\Sigma$ is the covariance matrix of returns,

$$1'z = 1'\Sigma^{-1}(E_n - c) = \omega\lambda_c^* = \omega\frac{E_c - c}{\sigma_c^2}$$

[1] This version of mean-variance optimization follows Benninga (2000).

is a modified Sharpe ratio, $\lambda_c^*$, at its optimum, $\omega$ is the sum of non-normalized portfolio weights, and $E_c$ is the portfolio expected return. Substituting in the weights we get,

$$\sigma_c^2 = (1'z)^{-2}(E_n)'\Sigma^{-1}(E_n)$$

portfolio variance as a function of expected returns, covariance and reference return. Expected return variance is generically

$$\sigma_E^2 = E_n' E_n n^{-1} - (\overline{E}_n)^2$$

Comparing portfolio and expected return variance we obtain,

$$\sigma_p^2 - \sigma_E^2 = E_n'\left[(1'z)^{-2}\Sigma^{-1} - n^{-1}I\right]E_n + (\overline{E}_n)^2 .$$

The sign of the difference depends on the term in brackets. Considering the nature of asset returns, assume return variances are less than one and expected returns are positive. Then term in bracket would be positive, but for the scaling effect of $(1'z)^2$ which is given by particular choice of tangency portfolio. Thus, there is no obvious relation between portfolio and expected return variance.

The equally weighted portfolio variance is always higher given the nature of return variances (i.e., less than one):

$$\sigma_{equal}^2 = \frac{1}{n^2}1'\Sigma 1 = \frac{1}{n^2}1'\left[E(RR') - E_n E_n'\right]1$$

$$\sigma_E^2 = E_n' E_n n^{-1} - (\overline{E}_n)^2$$

$$\sigma_{eq}^2 - \sigma_E^2 = \frac{1}{n^2}1'\left[E(RR') - E_n E_n'\right]1 - E_n' E_n n^{-1} - \overline{E}^2$$

$$\sigma_{eq}^2 - \sigma_E^2 = \frac{1}{n^2}1' E(RR')1 + \frac{1}{n^2}1' E_n E_n' 1 - E_n' E_n n^{-1} - \overline{E}^2$$

$$\sigma_{eq}^2 - \sigma_E^2 = \frac{1}{n^2}1' E(RR')1 - E_n' E_n n^{-1}$$

$$\sigma_{eq}^2 - \sigma_E^2 = \frac{1}{n^2}1' E\left[(E_n + \varepsilon_n)(E_n + \varepsilon_n)'\right]1 - E_n' E_n n^{-1}$$

$$\sigma_{eq}^2 - \sigma_E^2 = \frac{1}{n^2} 1' \left[ E_n E_n' + \Sigma \right] 1 - E_n' E_n n^{-1}$$

$$-\sigma_E^2 = \frac{1}{n^2} 1' \left[ E_n E_n' 1 \right] - E_n' E_n n^{-1}$$

The intuition is that the position of the tangency portfolio is determined by covariance of returns; not the variance of expected returns.

## *II.2 The error in variable problem*[2]

Error in variable arises when researchers use the traditional Fama-MacBeth two-pass methodology. First researchers estimate betas from time series data for individual securities or portfolios, and then they use the estimated betas in a cross sectional regression to obtain an estimation of coefficient parameters that relates expected return and betas. Specifically, in the spirit of Fama and Macbeth (1973), if *true* beta, $\beta_i$, are used in the cross-sectional regression:

$$r_{i,T} = \lambda_{f,T} + \lambda_{m,T} \beta_i + \varepsilon_{i,T} \tag{6}$$

Then, the correct price for the risk free, $\lambda_{f,T} = r_f$, and the realized market premium, $\lambda_{m,T} = r_{m,T} - r_f$ are expected.

However, an econometrician estimates the model of the form,

$$r_{i,t} = r_f + b_i \left( r_{m,t} - r_f \right) + e_{i,t}, \text{ from } \; t = \{ t-36, t-1 \}, \tag{7}$$

to obtain estimates of $\beta_i$, $b_i$, which are then used in a cross-sectional regression of $b_i$'s on the future returns realized at date=$T$:

$$r_{i,T} = \hat{\lambda}_{f,T} + \hat{\lambda}_{m,T} b_i + v_{i,T}, \qquad i = \{1, n\}. \tag{8}$$

The coefficients $\hat{\lambda}_{f,T}$ and $\hat{\lambda}_{m,T}$ are the estimated prices of market risk and return on the risk free asset respectively. Note that the econometrician uses estimates of beta, $b_i$, to estimate the

[2] The main topic of this paper is $R^2$ in the cross-section of expected return. I do not address the Error in Variable (EIV) problem. However, I summarize the problem to show that EIV does not fully explain why cross sectional $R^2$ is so low

parameters of the cross-sectional regression. The problem with using estimates of betas is that these estimates have sampling error that are passed to the coefficients $\hat{\lambda}_{f,T}$ and $\hat{\lambda}_{m,T}$. This is called error in variable (EIV) problem in the literature.

A simple derivation of the EIV model is presented in appendix I. The estimated $b_i$ is measured with error $\eta_i$

$$b_i = \beta_i + \eta_i,\ \eta_i \sim id\left(0, s_\eta^2\right), \text{ and } E\left(\beta_i \eta_i\right) = 0,\ \text{ where } s_\eta = \frac{s_e}{\sigma_m \sqrt{(T-1)}}$$

The sampling error $\eta_i$ is mean zero, uncorrelated across assets, and orthogonal to the *true* beta. The variance of the sampling error depends on the standard error of the estimated coefficients. Also, note that $s_b^2 = \sigma_\beta^2 + s_\eta^2$. As sample size increases, estimates converge to their true values.

Fama and MacBeth (1973) explicitly recognize the existence of the EIV problem in their procedure. They proposed to use portfolios (rather than individual securities) to minimize the effect of measurement error.[3] However, Fama and French (1993) demonstrate that even using portfolios, cross sectional $R^2$ are lower than time-series $R^2$, showing that the phenomenon is also present when using diversified portfolios.

## *C. Data description and estimation procedure*

I analyze the phenomenon of low cross-sectional $R^2$ compared to time series $R^2$ at security level and at portfolio level. At portfolio level, I use the Fama and French's 25 portfolios formed in size and book to market research portfolios, which are commonly used in academic research. Those portfolios are available in Kenneth French's website.[4] A complete set of portfolio returns are not available until July 1931, which means that cross-sectional estimates are computed starting in July 1934.

The risk-free rate is the Ibbotson 30-day Treasury bill rate which is also available on Kenneth French's website. Equally weighted and value-weighted market indices are from the Center for Research in Securities Prices (CRSP).

---

[3] Portfolios have lower residual variance than securities and therefore the variance of sampling error will be lower.

[4] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/. The explanation for the construction of portfolios and factors can be found in the website. The data is provided by K. French for academic research.

At security level, I analyze a large sample of CRSP NYSE, AMEX and NASDAQ exchange listed ordinary equity returns (share codes 10 and 11) from January 1926 through December 2004. The initial CSRP population is restricted to those securities having at least 37 months of consecutive returns. Thirty-six months is required to obtain parameter estimates, and the 37th month to run cross-sectional regression as well as to estimate cross-sectional statistics, which mean that cross-sectional estimates are computed starting in January 1929.

I compute cross-sectional $R^2$ and other cross-sectional statistics in three different ways. First, I use the Fama and Macbeth (1973) procedure outlined above. For each firm, I obtain betas from time-series return data and then run a regression of firm's return on estimated betas at each month *t*. Second, I run a pooled regression of returns on all firm-month betas estimated with the past 36-month returns (overall sample). Finally, I run a regression of firm average returns on firm average betas (firm level sample). All time series estimates are calculated using past 36 months of return whereas cross-sectional estimates are computed at each month t (cross-sectional monthly), at pooled data, and at firm level.

Specifically, for each firm *i* and at month *t*, I obtain betas from the following regression

$$r_{i,t} - r_f = a_i + b_i\left(r_{m,t} - r_f\right) + e_{i,t}, \qquad t = \{t-36, t-1\},$$

and then I use the estimated betas ($b_i$) to run the regression for each month *T*:

$$r_{i,T} = \hat{\lambda}_{f,T} + \hat{\lambda}_{m,T} b_i + v_{i,T}, \qquad i = \{1, n\}$$

The overall sample considers the pooled regression of return on firm-month betas:

$$r_{i,t} = \hat{\lambda}_f + \hat{\lambda}_m b_{,ti} + \xi_{i,t}, \qquad i = \{1, nt\}$$

Whereas the firm level average returns on average betas considers the following regression:

$$\bar{r}_i = \hat{\lambda}_f + \hat{\lambda}_m \bar{b}_i + \varsigma_i, \qquad i = \{1, n\}$$

## III. Empirical Results

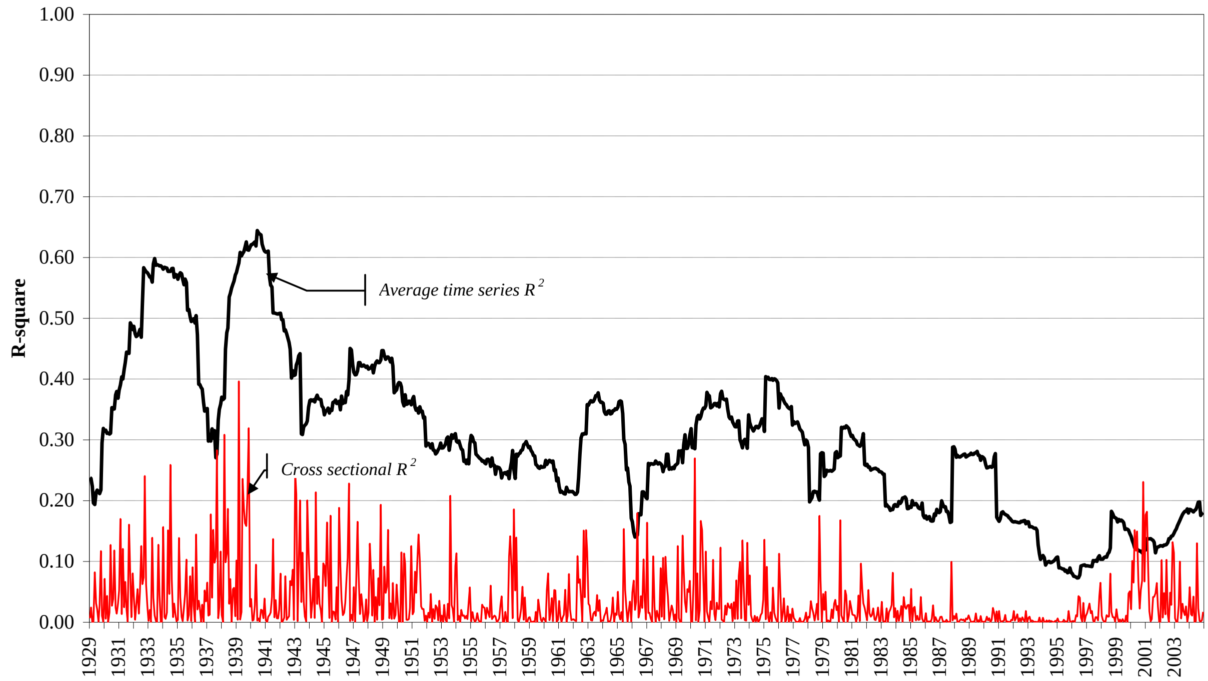
This section reports results and shows that market volatility plays an important role in explaining the observed low cross-sectional $R^2$. Figure 1 shows that cross sectional $R^2$ –for securities– is consistently lower than time series $R^2$ over time since January 1929, with some exceptions in 1937, 1966, 2000 and 2001. The phenomenon is also observed for portfolios, although –as shown later in the paper– portfolios present higher time series $R^2$. The result has been reported by Fama and French (1992), who find that no cross-sectional return and beta relationship, by Reinganum (1981), who uses two market indexes and finds no relationship between cross-sectional returns and betas, and by Lakonishok and Shapiro (1986), who after a series of empirical tests find no relationship between cross-sectional expected returns and betas.

The low cross-sectional $R^2$ is because the market variance is consistently higher than the variance of expected return. In order to facilitate the explanation, I re-write the key formula that explains the observed low cross-sectional $R^2$:

$$\sigma_T^2 - \sigma_N^2 = \sigma_m^2 - \left(E_m - rf\right)^2 \sigma_\beta^2 \tag{9}$$

Remember that the closer the magnitude of market variance, $\sigma_m^2$ , and variance of expected return, $\left(E_m - rf\right)^2 \sigma_\beta^2$, the lower the difference between time series $R^2$ and cross-sectional $R^2$. However, as seen in Figure 2, market variance has been consistently higher than the average return variance, which is a proxy for variance of expected return during the period. This higher market volatility may explain why $R^2$ is low in the cross-section. Market variance has is highest value in the 1930s. In more recent decades, the 1960s and the 2000s are periods characterized by high volatility. Particularly, the period 1965-1989 was characterized by high volatility. Not surprising, many research studies that use data for this period report low cross-sectional $R^2$ (Fama-French, 1992, Chell, Ross and Roll, 1986, Reinganum, 1981, Lakonishok and Shapiro, 1986).

**Figure 1. Securities average time series $R^2$ and cross sectional $R^2$**
***Jan 1929 - Dec 2004***

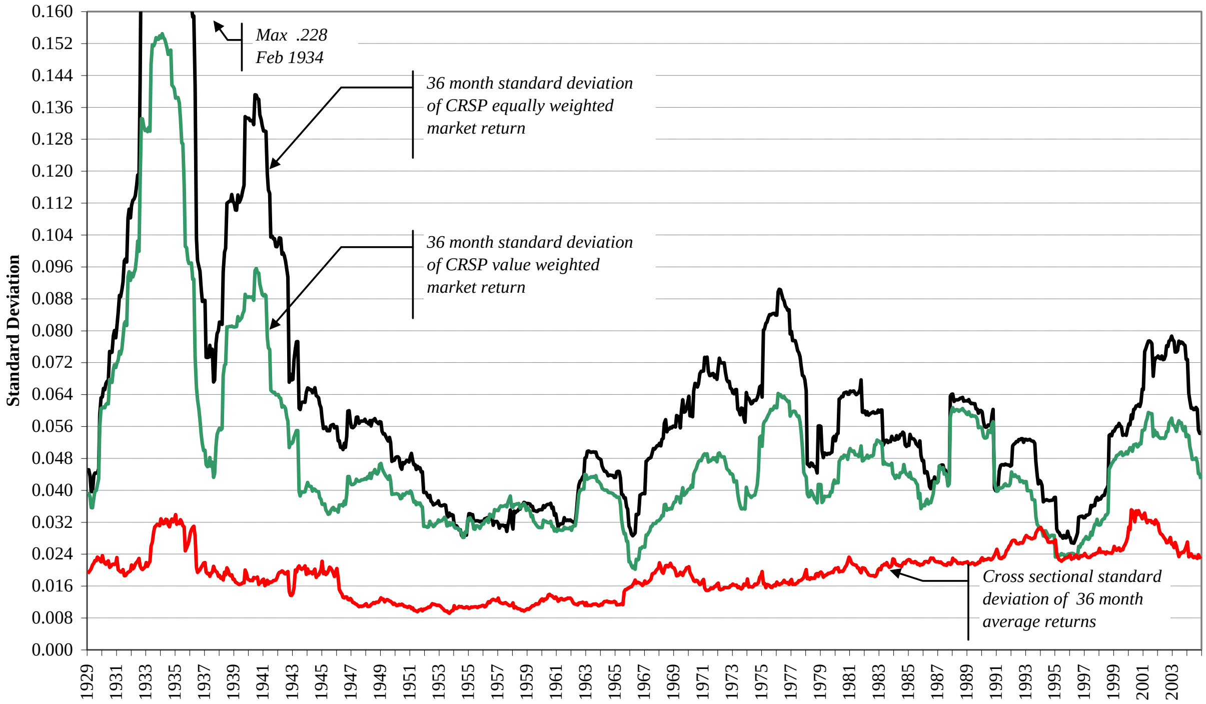



The population of CRSP securities from January 1926 to December 2004 is considered. To be selected into a sample month=*t*, the first stage OLS market model, $r_i - r_f = a_i + b_i(r_m - r_f) + e_i$, requires firm *i* to have non-missing returns for the prior 36 months. The second stage month=*t* cross sectional SML regression, $r_{i,t} = \hat{\lambda}_f + \hat{\lambda}_m b_i + v_{i,t}$, requires firms to have a non-missing month=*t* return. Thus, the sample spans January 1929 to December 2004 (*T*=912 months). The average time series *R2* from the first stage OLS as well as the cross-sectional *R2* from the second stage OLS is plotted.

Table 1 presents monthly statistics as well as the Fama and Macbeth regression results. As was shown in Figure 1, the average market variance, $s_m^2$, is higher than the average variance of firms' 36 month average return, $s_{\bar{r},T}^2$; this higher variance of the market is causing the observed low cross-sectional $R^2$. Moreover, time series residual variance is not significantly different than cross-sectional residual variance statistically. I consider that the small difference between the time series residual variance and cross-sectional variance is due to the fact that my dataset is an unbalanced panel data.

The average variance of beta, $s_b^2$, is 0.417. However its volatility as well as the mean has increased during the last 15 years. The average for the period 1989-2004 is 0.90, more than twice, the long term average.

**Figure 2. Market and security average return volaility**
***Jan 1929 - Dec 2004***




The figure plots the 36-month standard deviation of both CRSP equally weighted and value weighted return as well as the cross-sectional standard deviation of 36-month average return. The 36-month average return is calculated for each firm *i* at time *t*, and then the cross-sectional standard deviation is computed using all available firms at time *t*.

The estimated intercepts and slopes from the cross-sectional regression are significantly different than the average risk free rate and the average market risk premium. Those results are consistent with Fama and Macbeth (1973), who find significant difference using a sample of securities from 1963-1970.

Table 1 also reports that the average time series $R_T^2$ is around 10 times the average cross-sectional $R_N^2$. The next section further studies why cross-sectional $R^2$ is relatively low compared with the market model and why it changes through time.

Table 2 presents the results for overall sample and firm level sample. Two important conclusions can be derived. The estimated cross-sectional variance of residual is different depending of the method we use and therefore the estimated cross-sectional $R^2$. Moreover, there are significance difference between the estimated intercept ( $\lambda_f$ ) and the risk free rate, and the estimated slope ( $\lambda_m$ ) and the average risk premium.

**Table 1. Monthly statistics and Fama MacBeth cross-sectional regression results (January 1929 – December 2004, T=912)**

| Sym | Variable ( at month=t ) | Mean | Min | Max | 5th | 95th |
|---|---|---|---|---|---|---|
| $s_N^2$ | *Cross sectional variance of returns* | 0.0190 | 0.0018 | 0.3195 | 0.0033 | 0.0522 |
| $s_T^2$ | *Average of firms' 36 month returns variance* | 0.0236 | 0.0040 | 0.1128 | 0.0053 | 0.0577 |
| $s_{\bar{r},T}^2$ | *Variance of firms' 36 month average return* | 0.0004 | 0.0001 | 0.0012 | 0.0001 | 0.0010 |
| $s_m^2$ | *Variance of CRSP equally weighted 36 month return* | 0.0057 | 0.0007 | 0.0522 | 0.0010 | 0.0199 |
| $s_{m-vw}^2$ | *Variance of CRSP value weighted 36 month return* | 0.0031 | 0.0004 | 0.0239 | 0.0008 | 0.0094 |
| $s_b^2$ | *Cross sectional variance of betas* | 0.4116 | 0.1180 | 2.2673 | 0.1528 | 1.0662 |
| $s_\eta^2$ | *Average variance of firms' beta estimates* | 0.1525 | 0.0219 | 0.8830 | 0.0283 | 0.5024 |
| $s_{e,T}^2$ | *Average variance of firms' market model residuals* | 0.0165 | 0.0029 | 0.0507 | 0.0040 | 0.0428 |
| $s_{v,N}^2$ | *Variance of SML regression residuals* | 0.0183 | 0.0017 | 0.3072 | 0.0032 | 0.0500 |
| $R_T^2$ | *Average $R^2$ from firms' market model regressions* | 0.2995 | 0.0726 | 0.6441 | 0.1061 | 0.5767 |
| $R_N^2$ | *Cross sectional $R^2$ from SML regressions* | 0.0341 | 0.0000 | 0.3960 | 0.0001 | 0.1509 |
| $r_m$ | *CRSP equally weighted return* | 0.0130 | -0.3118 | 0.6551 | -0.0950 | 0.1085 |
| $r_f$ | *One month Ibbotson Treasury Bill rate* | 0.0031 | -0.0003 | 0.0152 | 0.0001 | 0.0079 |
| $\hat{\lambda}_f$ | *Intercept from SML regressions**** | 0.0114 | -0.2114 | 0.3558 | -0.0541 | 0.0652 |
| $rp_m$ | *Market premia: $r_m - r_f$* | 0.0099 | -0.3123 | 0.6548 | -0.0997 | 0.1071 |
| $\hat{\lambda}_m$ | *Slope estimate from SML regressions*** | 0.0025 | -0.1692 | 0.5197 | -0.0635 | 0.0846 |
| $N$ | *Number of securities per month* | 2,223 | 420 | 5,044 | 554 | 4,820 |

The population of CRSP securities from January 1926 to December 2004 is considered. To be selected into a sample month=*t*, the first stage OLS market model, $r_i - r_f = a_i + b_i(r_m - r_f) + e_i$, requires firm *i* to have non-missing returns for the prior 36 months. The second stage month=*t* cross sectional SML regression, $r_{i,t} = \hat{\lambda}_f + \hat{\lambda}_m b_i + v_{i,t}$, requires firms to have a non-missing month=*t* return. Thus, the sample spans January 1929 to December 2004 (*T*=912 months). The mean, minimum, maximum, 5th percentile, and 95th percentile of months are reported. Statistics labeled 36 months use the same sample of prior returns as do the market model regressions. A description of each statistic is included as part of the Table. The *, **, *** represent *10, 5*, and *1%* statistical significance of SML hypotheses: $\hat{\lambda}_f = E(r_f)$ and $\hat{\lambda}_m = E(r_m - r_f)$, respectively.

## Table 2. Overall and firm level sample statistics with cross-sectional regression results (January 1929 – December 2004)

***Panel A. Overall sample (No. of firm- months =2.03 million)***

| Sym | Variable | Mean | Var | 5th | 95th |
|---|---|---|---|---|---|
| $r, s_N^2$ | *Return* | 0.0149 | 0.0277 | -0.1964 | 0.2500 |
| $\bar{r}_{36}, s_{\bar{r},T}^2$ | *Average return last 36 months* | 0.0147 | 0.0006 | -0.0207 | 0.0541 |
| $s_T^2$ | *Returns variance past 36 months* | 0.0258 | 0.0082 | 0.0025 | 0.0827 |
| $s_{e,T}^2$ | *Variance of market model residuals* | 0.0207 | 0.0056 | 0.0019 | 0.0677 |
| $R_T^2$ | *$R^2$ from market model regressions* | 0.2324 | 0.0316 | 0.0063 | 0.5639 |
| $b, s_b^2$ | *Market model beta* | 1.0001 | 0.5605 | 0.0796 | 2.2214 |
| $s_\eta^2$ | *Variance market model beta estimates* | 0.2284 | 1.0979 | 0.0165 | 0.7847 |

***Panel B. Firm level sample (No. of firms = 16,137)***

| Sym | Variable | Mean | Var | 5th | 95th |
|---|---|---|---|---|---|
| $\bar{r}, s_{\bar{r}}^2$ | *Unconditional average return* | 0.0127 | 0.0003 | -0.0205 | 0.0385 |
| $\bar{s}_T^2$ | *Average returns variance last 36 months* | 0.0405 | 0.0065 | 0.0046 | 0.1264 |
| $\bar{s}_{e,T}^2$ | *Average market model residual variance* | 0.0343 | 0.0054 | 0.0038 | 0.1057 |
| $R_T^2$ | *Average $R^2$ from market model regressions* | 0.1888 | 0.0149 | 0.0224 | 0.4141 |
| $\bar{b}, s_b^2$ | *Average market model beta* | 1.1177 | 0.5432 | 0.1818 | 2.4537 |
| $\bar{s}_\eta^2$ | *Average variance of market model beta estimates* | 0.3697 | 0.7698 | 0.0357 | 1.1739 |

***Panel C. SML overall and firm level cross-sectional regressions***

| | | | *Overall (2.03 mil)* | | *Firm Level (16,137)* | |
|---|---|---|---|---|---|---|
| **Sym** | **Variable** | **Hypothesis ($H_o$)** | **Estimate** | ***t*-stat ($H_o$)** | **Estimate** | ***t*-stat ($H_o$)** |
| $\hat{\lambda}_f$ | *Intercept* | $\hat{\lambda}_f = E(r_f)$ | 0.0146 | 51.38 | 0.0082 | 12.42 |
| $\hat{\lambda}_m$ | *Slope* | $\hat{\lambda}_m = E(r_m - r_f)$ | 0.0003 | -40.83 | 0.0041 | -15.22 |
| $s_{e,N}^2$ | *Residual variance* | | 0.0277 | | 0.0003 | |
| $R_N^2$ | *Cross-sectional $R^2$* | | 0.0000 | | 0.0268 | |

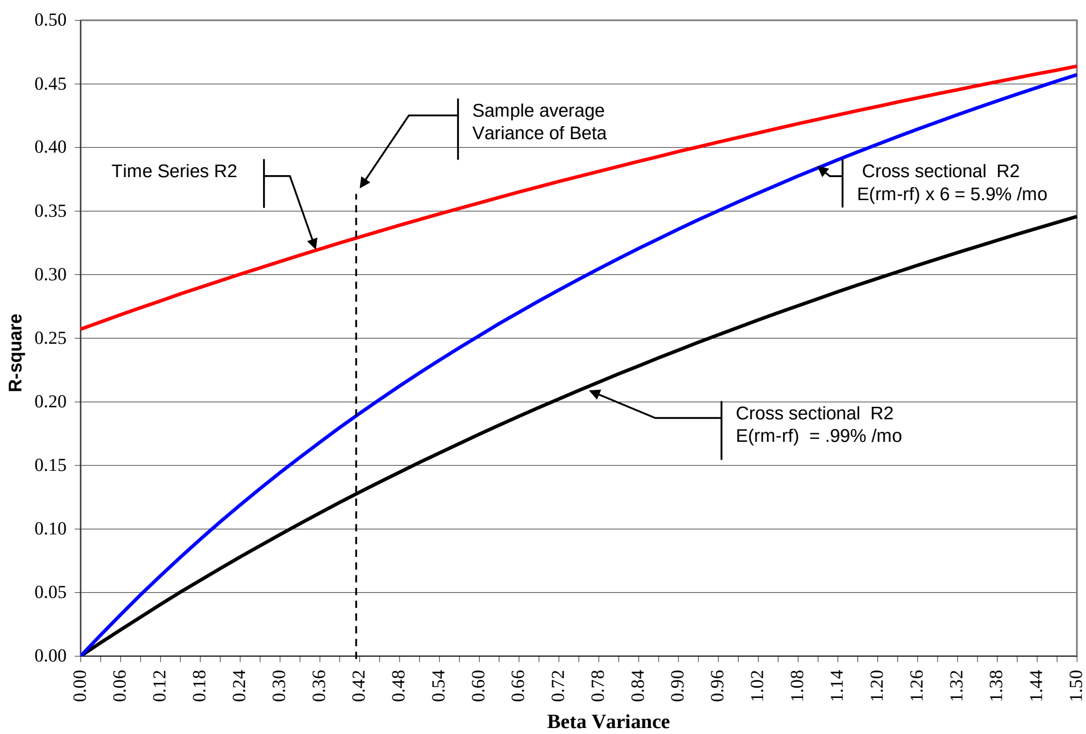
From the population of CRSP securities spanning January 1926 to December 2004, the sample used in Panels A-C require firm i for any month t have a non-missing return for that month and a 36 month valid returns history. OLS market model, $r_i - r_f = a_i + b_i(r_m - r_f) + e_i$, derived variables for firm i in month t use 36 month returns prior to month t and variables labeled 36 months use the same sample in their computation. Variables descriptions are part of this table. The mean, variance, 5th percentile, and 95th percentile of firm months or of firms are reported. Panel A reports overall sample statistics of all firm months and Panel B reports statistics based on unconditional firm averages. Panel C reports SML cross sectional regression results using all firm months, $r_{i,t} = \hat{\lambda}_f + \hat{\lambda}_m b_{i,t-1} + v_{i,t}$, and from using firm unconditional averages $\bar{r}_i = \hat{\lambda}_f + \hat{\lambda}_m \bar{b}_i + v_i$. The regression hypotheses and their associated heteroskedasticity adjusted t-stats are shown.

## *A. Explaining the cross-sectional R2 volatility*

Cross-sectional $R^2$ are consistently lower than time series $R^2$. Furthermore, the observed $R^2$ fluctuate over time. The reason for this fluctuation depends on changes in market variances, change in expected market risk premium, or change in beta variance. I study the combined effect of changing beta variance and expected market risk premium. If the interaction of these two variables is enough to make the variance of expected return equal to the variance of market, the resulting cross-sectional $R^2$ would equal to the time series $R^2$.

Figure 3 plots $R^2$ as a function of beta variance and expected return, using securities' residual variance, given the market variance for the period 1927-2004 ($\sigma_m^2 = 0.00570$). The graph shows that higher variance of beta is associated with higher cross-sectional $R^2$. However, this is also true for time series $R^2$. Thus, change in variance of betas will not be enough to produce the variance needed to reach the variance of the market. The variance of beta is lower than 1.06, 95 percent of the time (see Table 1), which imposes a bound to the maximum cross-sectional $R^2$ that we may see in practice.

**Figure 3. Behavior of Time Series and Cross-sectional R2 (firms)**



The Figure plots sensitivity to computed time series, $R_T^2$, and cross-sectional, $R_N^2$, to changes in variance of betas and changes in expected market risk premium. Time series $R^2$ is calculated as one minus the residual variance, $\sigma_e^2$, divided by the estimated time series variance, $\sigma_T^2 = (\sigma_\beta^2 + 1) * \sigma_m^2 + \sigma_e^2$. Similarly, cross-sectional $R^2$ is calculated as one minus the residual variance, $\sigma_e^2$, divided by the estimated cross-sectional variance, $\sigma_N^2 = (\sigma_m^2 + E(r_m - r_f)^2) * \sigma_\beta^2 + \sigma_e^2$.

Moreover, not only high variance of beta is needed, but also high expected market risk premium. If the expected risk premium is 0.99% per month, which is the observed average during the period, it would not be enough to produce the high variance of expected return needed to reach the time series $R^2$, even in the presence of high variance of beta. In the example, only when the variance of beta is 1.50 and the market risk is premium is 6%, the cross-sectional $R^2$ would be equal to the time series $R^2$! Thus, in periods where market variance is low, and both expected risk market risk premium and variance of beta are high, the cross-sectional $R^2$ would be similar (or even higher) than the time series $R^2$. Nevertheless, the cross-sectional $R^2$ will vary according to the fluctuation in expected market risk premium, variance of betas, and market volatility. The observed volatility of the cross-sectional $R^2$ is due to the fact that market variance change, expected market risk premium, and variance of beta change through time.

Given the average market risk premium, market variance, and estimated variance of the residual and variance of beta, Table 3 reports computed and estimated $R^2$s, and compares the different approaches under study. All approaches predict a low cross-sectional $R^2$ relative to time series $R^2$. For example, the expected time series $R^2$ is 0.33, whereas the expected cross-sectional $R^2$ is 0.13, when using Fama-Macbeth regressions (cross-section by month); the results are similar under pooled sample. However, the average regression estimated cross-sectional $R^2$ for pooled sample is extremely low because there is no time variation when pooling data.

**Table 3. R squares comparison for individual firms**

| | Sym | Variable name | | | |
|---|---|---|---|---|---|
| ***Market Statistics*** | | | | | |
| | $r_f$ | Risk-free rate | 0.00305 | | |
| | $E_m$ | Expected Market Return | 0.01298 | | |
| | $\sigma_m^2$ | Market Variance | 0.00570 | | |
| | | | | | |
| | | | Cross-sections by month | Overall | Firm level |
| ***Securities Statistics*** | | | | | |
| | $\sigma_\beta^2$ | Cross-sectional Variance of Beta | 0.41162 | 0.56052 | 0.54318 |
| | $\sigma_\varepsilon^2$ | Time-series Residual Variance | 0.01647 | 0.02074 | 0.03428 |
| | $\sigma_T^2$ | Time-series Return Variance | 0.02452 | 0.02963 | 0.04308 |
| | $\sigma_N^2$ | Cross-sectional Return Variance | 0.01886 | 0.02399 | 0.03743 |
| ***R squares*** | | | | | |
| | $R_T^2$ | Over all Time-series R2 | 0.32819 | 0.30017 | 0.20421 |
| | $R_N^2$ | Computed cross-sectional R2 | 0.12657 | 0.13549 | 0.08415 |
| | $R_N^2$ | Reg-estimated cross-sectional R2 | 0.03409 | 0.00000 | 0.02682 |
| | $R_N^2$ | EIV Corrected cross-sectional R2 | 0.07543 | 0.00023 | 0.06044 |

The population of CRSP securities from January 1926 to December 2004 is considered. To be selected into a sample month=*t*, the first stage OLS market model, $r_i - r_f = a_i + b_i(r_m - r_f) + e_i$, requires firm *i* to have non-missing returns for the prior 36 months. The second stage month=*t* cross sectional SML regression, $r_{i,t} = \hat{\lambda}_f + \hat{\lambda}_m b_i + v_{i,t}$, requires firms to have a non-missing month=*t* return. The Table reports the computed and estimated cross-sectional R2 using the different approaches to estimate parameters: cross-section by month (Fama-Macbeth type regressions), pooled sample (overall sample), and firm level sample. Time series $R^2$ is computed as one minus the residual variance, $\sigma_e^2$, divided by the estimated time series variance, $\sigma_T^2 = (\sigma_\beta^2 + 1) * \sigma_m^2 + \sigma_e^2$. Similarly, cross-sectional $R^2$ is computed as one minus the residual variance, $\sigma_e^2$, divided by the estimated cross-sectional variance, $\sigma_N^2 = (\sigma_m^2 + E(r_m - r_f)^2) * \sigma_\beta^2 + \sigma_e^2$.

*B. Time series and cross-sectional R2 in portfolios*

This section reports findings when portfolios are used in the analysis (instead of securities). The error in variable problems is mitigated through using portfolios because the variance of the estimate, $s_{\eta}^2$, is lower when using diversified portfolios. Furthermore, time series $R^2$ are higher when using portfolios because the variance of the residual is lower than the variance of residual.

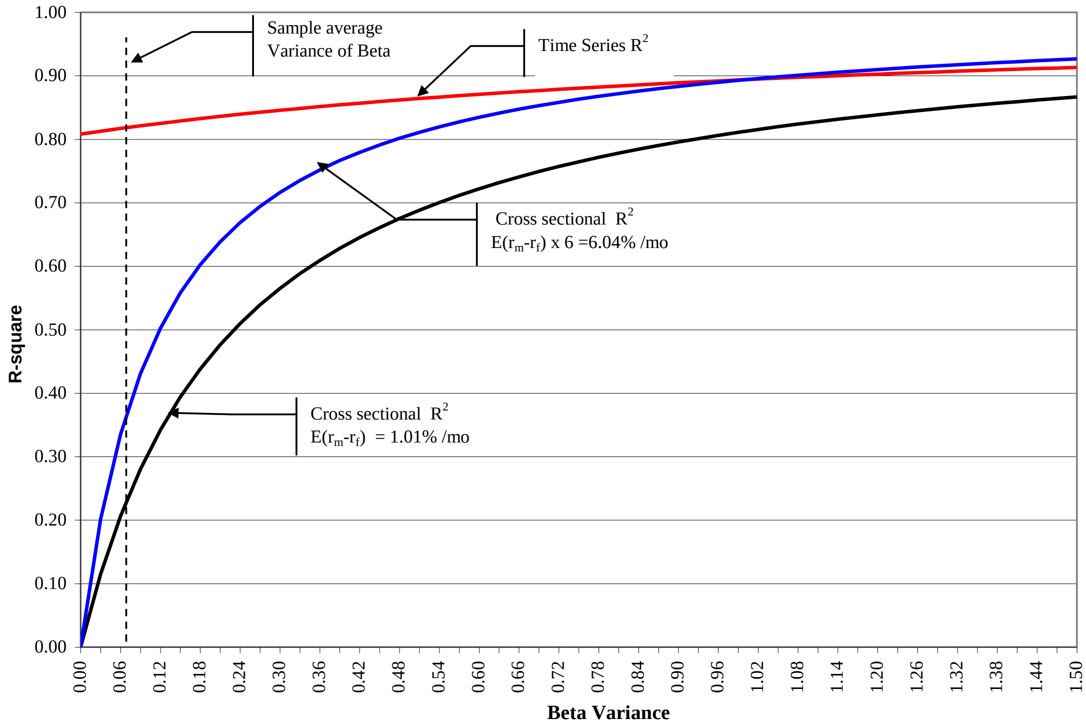
The results are similar to the ones reported for securities. The variance of portfolio expected return, $s_{\bar{r},T}^2$, is lower than the variance of the market, suggesting the same conclusion outline before. The lower cross-sectional $R^2$ observed when using portfolios due to the expected return does not have enough volatility to equal the market volatility. The intercept and slope are significantly different than the risk free rate and expected market risk premium, respectively.

Figure 4 shows the behavior of time series and cross-sectional $R^2$ when portfolios are used. The curve is steeper because the variance of the residual is lower. However, as in the security case, increasing variance of beta is not enough to have similar variances (variance of the market close to variance of expected return) that would make time series and cross sectional $R^2$ equal. A market risk premium of 6 percent per month is needed in order to have similar time-series and cross-sectional $R^2$!

Moreover, the sample average variance of beta is 0.06 and therefore the expected time series $R^2$ is .82 whereas the expected cross-sectional $R^2$ is 0.22. The results are summarized in

Table 4 for different approaches. In all the cases, the computed cross-sectional $R^2$ is lower than the time-series $R^2$.

**Figure 4. Behavior of Time Series and Cross-sectional $R^2$ (portfolios)**




The Figure plots sensitivity to computed time series, $R_T^2$, and cross-sectional, $R_N^2$, to changes in variance of betas and changes in expected market risk premium. Time series $R^2$ is calculated as one minus the residual variance, $\sigma_e^2$, divided by the estimated time series variance, $\sigma_T^2 = \left(\sigma_\beta^2 + 1\right) * \sigma_m^2 + \sigma_e^2$. Similarly, cross-sectional $R^2$ is calculated as one minus the residual variance, $\sigma_e^2$, divided by the estimated cross-sectional variance, $\sigma_N^2 = \left(\sigma_m^2 + E\left(r_m - r_f\right)^2\right) * \sigma_\beta^2 + \sigma_e^2$.

**Table 4. R squares comparison for portfolios**

| | Sym | Variable name | | | |
|---|---|---|---|---|---|
| ***Market Statistics*** | | | | | |
| | $r_f$ | Risk-free rate | 0.00318 | | |
| | $E_m$ | Expected Market Return | 0.01325 | | |
| | $\sigma_m^2$ | Market Variance | 0.00367 | | |
| | | | | | |
| | | | Cross-sections by month | Overall | Firm level |
| ***Portfolio Statistics*** | | | | | |
| | $\sigma_\beta^2$ | Cross-sectional Variance of Beta | 0.06500 | 0.07150 | 0.03311 |
| | $\sigma_\varepsilon^2$ | Time-series Residual Variance | 0.00087 | 0.00087 | 0.00087 |
| | $\sigma_T^2$ | Time-series Return Variance | 0.00478 | 0.00480 | 0.00466 |
| | $\sigma_N^2$ | Time-series Return Variance | 0.00112 | 0.00114 | 0.00100 |
| ***R squares*** | | | | | |
| | $R_T^2$ | Over all Time-series R2 | 0.81775 | 0.81865 | 0.81317 |
| | $R_N^2$ | Computed cross-sectional R2 | 0.21962 | 0.23640 | 0.12538 |
| | $R_N^2$ | Reg-estimated cross-sectional R2 | 0.26417 | 0.00025 | 0.10303 |
| | $R_N^2$ | EIV Corrected cross-sectional R2 | 0.30329 | 0.00034 | 0.16459 |

Fama and French's 25 size and book to market portfolios is considered. To be selected into a sample month=*t*, the first stage OLS market model, $r_i - r_f = a_i + b_i(r_m - r_f) + e_i$, requires portfolio *i* to have non-missing returns for the prior 36 months. The second stage month=*t* cross sectional SML regression, $r_{i,t} = \hat{\lambda}_f + \hat{\lambda}_m b_i + v_{i,t}$, requires portfolios to have a non-missing month=*t* return. Thus, the sample period that has valid return is July 1934 to December 2004. The Table reports the computed and estimated cross-sectional R2 using the different approaches to estimate parameters: cross-section by month (Fama-Macbeth type regressions), pooled sample (overall sample), and firm level sample. Time series $R^2$ is computed as one minus the residual variance, $\sigma_e^2$, divided by the estimated time series variance, $\sigma_T^2 = (\sigma_\beta^2 + 1) * \sigma_m^2 + \sigma_e^2$. Similarly, cross-sectional $R^2$ is computed as one minus the residual variance, $\sigma_e^2$, divided by the estimated cross-sectional variance, $\sigma_N^2 = (\sigma_m^2 + E(r_m - r_f)^2) * \sigma_\beta^2 + \sigma_e^2$.

## IV. Discussion

The security market line implies a perfect linear relationship between firm's beta and returns. However the literature finds that this relationship is very weak. Specifically, the cross-sectional $R^2$ from regression of returns on betas are much lower than average time-series $R^2$ from regression of excess return on market risk premium.

This paper provides an explanation to this phenomenon. I derive a simple model that relates cross-sectional $R^2$ with time series $R^2$ and find that market variance is key determinant of the observed low cross-sectional $R^2$. The difference between market variance and expected return determines the observed difference between time series $R^2$ and cross-sectional $R^2$. When market variance equals the expected return, cross-sectional $R^2$ will be exactly as the time series $R^2$. I further show that in order for the expected return to have a similar magnitude of the average market variance (that is, cross-sectional $R^2$ equals to time series $R^2$), both high variance of beta and high expected market risk premium is needed. Thus, only when there is low market variance, high variance of beta, and high expected market risk premium, the cross-sectional $R^2$ would be similar (could be also higher) to time series $R^2$.

**References**


Fama, E., French, K., 1992. The cross-section of expected returns. *Journal of Finance* 47, 427–466.

Fama, E., Macbeth, J. (1973) Risk, return and equilibrium: empirical test. *Journal of Political Economy* 81, 607-636

Kim D. 1995. The error in the variables problem in the cross-section of expected return. *The Journal of Financial*, Vol. 50, No. 5, pp 1605 – 1634

Kothari S.P., Shanken, J., and Sloan R., 1995. Another look at the cross-section of expected stock returns. The Journal of Finance, Vol. 50, No. 1, pp. 185 – 224.

Lakonishock, J. and Shapiro, A. 1986. Systematic risk, total risk and size as determinant of stock market returns. *Journal of Banking and Finance*, 10, pp 115 – 132.

Lintner, John, 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budget. *Review of Economics and Statistics*. 47, pp. 13 – 57

Pentengill, G., Sundaram S., and Mathur, I., 1995. The conditional relation between beta and return. The journal of financial and quantitative analysis. Vol 30., No. 1.

Reinganum, Marc R., 1981. A new empirical perspective of the CAPM. *Journal of Financial and Quantitative Analysis* 16, 439 – 462.

Roll R. and Ross S. 1994. On the cross-sectional relationship between expected return and betas. *The Journal of Finance*, Vol. 49. No. 1. pp 101 – 121.

Sharpe, William F., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance* 19, pp. 425 – 442.

# Appendix A

**Table 5. Monthly statistics and Fama MacBeth cross-sectional regression results (25 size/btm portfolios. July 1934 – December 2004, T=846)**

| Sym | Variable ( at month=t ) | Mean | Min | Max | 5th | 95th |
|---|---|---|---|---|---|---|
| $s_N^2$ | *Cross sectional variance of returns* | 0.0243 | 0.0055 | 0.2213 | 0.0110 | 0.0519 |
| $s_T^2$ | *Average of portfolios' 36 month returns variance* | 0.0651 | 0.0303 | 0.2607 | 0.0355 | 0.1459 |
| $s_{\bar{r},T}^2$ | *Variance of portfolios' 36 month average return* | 0.0054 | 0.0020 | 0.0195 | 0.0027 | 0.0109 |
| $s_m^2$ | *Variance of CRSP equally weighted 36 month return* | 0.0602 | 0.0267 | 0.2253 | 0.0312 | 0.1299 |
| $s_{m-vw}^2$ | *Variance of CRSP value weighted 36 month return* | 0.0458 | 0.0202 | 0.1503 | 0.0277 | 0.0850 |
| $s_b^2$ | *Cross sectional variance of betas* | 0.2343 | 0.1250 | 0.6060 | 0.1411 | 0.3839 |
| $s_\eta^2$ | *Average variance of portfolios' beta estimates* | 0.0764 | 0.0469 | 0.1215 | 0.0539 | 0.1085 |
| $s_{e,T}^2$ | *Average variance of portfolios' market model residuals* | 0.0260 | 0.0145 | 0.0938 | 0.0163 | 0.0454 |
| $s_{v,N}^2$ | *Variance of SML regression residuals* | 0.0198 | 0.0047 | 0.1437 | 0.0101 | 0.0385 |
| $R_T^2$ | *Average $R^2$ from portfolios' market model regressions* | 0.8142 | 0.4175 | 0.9390 | 0.5952 | 0.9192 |
| $R_N^2$ | *Cross sectional $R^2$ from SML regressions* | 0.2659 | 0.0000 | 0.9156 | 0.0016 | 0.7363 |
| $r_m$ | *CRSP equally weighted return* | 0.0130 | -0.2856 | 0.3927 | -0.0808 | 0.1015 |
| $r_f$ | *One month Ibbotson Treasury Bill rate* | 0.0033 | -0.0002 | 0.0152 | 0.0001 | 0.0080 |
| $\hat{\lambda}_f$ | *Intercept from SML regressions**** | 0.0093 | -0.5563 | 0.2097 | -0.0805 | 0.0902 |
| $rp_m$ | *Market premia:* $r_m - r_f$ | 0.0097 | -0.2856 | 0.3926 | -0.0847 | 0.0982 |
| $\hat{\lambda}_m$ | *Slope estimate from SML regressions*** | 0.0035 | -0.1995 | 0.9215 | -0.1015 | 0.1238 |

Fama and French's 25 size and book to market portfolios from July 1931 to December 2004 is considered. The first stage OLS market model, $r_i - r_f = a_i + b_i(r_m - r_f) + e_i$, is estimated using the prior 36 months to month=*t*. Then, the second stage at month=*t* cross sectional SML regression, $r_{i,t} = \hat{\lambda}_f + \hat{\lambda}_m b_i + v_{i,t}$, is computed using parameters estimates from the first stage. Thus, the sample spans July 1934 to December 2004 (*T*=810 months). The mean, minimum, maximum, 5th percentile, and 95th percentile of months are reported. Statistics labeled 36 months use the same sample of prior returns as do the market model regressions. A description of each statistic is included as part of the Table. The *, **, *** represent *10*, *5*, and *1%* statistical significance of SML hypotheses: $\hat{\lambda}_f = E(r_f)$ and $\hat{\lambda}_m = E(r_m - r_f)$. All variance estimates are scaled by the square root (**e.g. standard deviations are reported**).

## Table 6. Overall and portfolio level sample statistics with cross-sectional regression results (January 1934 – December 2004)

***Panel A. Overall sample (No. of portfolio-months =20,250)***

| Sym | Variable | Mean | S.D. | 5th | 95th |
|---|---|---|---|---|---|
| $r, s_N^2$ | *Return* | 0.0130 | 0.0656 | -0.0829 | 0.1072 |
| $\bar{r}_{36}, s_{\bar{r},T}^2$ | *Average return last 36 months* | 0.0140 | 0.0115 | -0.0022 | 0.0326 |
| $s_T^2$ | *Returns variance past 36 months* | 0.0633 | - | 0.0300 | 0.1376 |
| $s_{e,T}^2$ | *Variance of market model residuals* | 0.0240 | - | 0.0104 | 0.0512 |
| $R_T^2$ | *$R^2$ from market model regressions* | 0.8961 | 0.1059 | 0.7012 | 0.9818 |
| $b, s_b^2$ | *Market model beta* | 0.9469 | 0.2576 | 0.5551 | 1.3697 |
| $s_\eta^2$ | *Variance market model beta estimates* | 0.0710 | - | 0.0340 | 0.1273 |

***Panel B. Firm level sample (No. of portfolios = 25)***

| Sym | Variable | Mean | S.D. | 5th | 95th |
|---|---|---|---|---|---|
| $\bar{r}, s_{\bar{r}}^2$ | *Unconditional average return* | 0.0131 | 0.0030 | 0.0091 | 0.0174 |
| $\bar{s}_T^2$ | *Average returns variance last 36 months* | 0.0732 | - | 0.0530 | 0.1032 |
| $\bar{s}_{e,T}^2$ | *Average market model residual variance* | 0.0276 | - | 0.0187 | 0.0464 |
| $R_T^2$ | *Average $R^2$ from market model regressions* | 0.9014 | 0.0411 | 0.8254 | 0.9482 |
| $\bar{b}, s_{\bar{b}}^2$ | *Average market model beta* | 0.9469 | 0.1768 | 0.6948 | 1.2375 |
| $\bar{s}_\eta^2$ | *Average variance of market model beta estimates* | 0.0760 | - | 0.0530 | 0.0988 |

***Panel C. SML overall and firm level cross-sectional regressions***

| | | | *Overall (20,250)* | | *Portfolio Level (25)* | |
|---|---|---|---|---|---|---|
| **Sym** | **Variable** | **Hypothesis ($H_o$)** | **Estimate** | ***t*-stat ($H_o$)** | **Estimate** | ***t*-stat ($H_o$)** |
| $\hat{\lambda}_f$ | *Intercept* | $\hat{\lambda}_f = E(r_f)$ | 0.0101 | 3.51 | 0.0070 | 0.86 |
| $\hat{\lambda}_m$ | *Slope* | $\hat{\lambda}_m = E(r_m - r_f)$ | 0.0032 | -3.01 | 0.0064 | -0.66 |
| $s_{e,N}^2$ | *Residual variance* | | 0.0656 | | 0.0000 | |
| $R_N^2$ | *Cross-sectional $R^2$* | | 0.0002 | | 0.1473 | |

Fama and French's 25 size and book to market portfolios from July 1931 to December 2004 is considered. The sample used in Panels *A*-*C* requires firm *i* for any month *t* have a non-missing return for that month and a 36 month valid returns history. OLS market model, $r_i - r_f = a_i + b_i(r_m - r_f) + e_i$, derived variables for firm *i* in month *t* use 36 month returns prior to month *t* and variables labeled 36 months use the same sample in their computation. Variables descriptions are part of this table. The mean, variance, 5th percentile, and 95th percentile of firm months or of firms are reported. Panel *A* reports overall sample statistics of all firm months and Panel *B* reports statistics based on unconditional firm averages. Panel *C* reports SML cross sectional regression results using all firm months, $r_{i,t} = \hat{\lambda}_f + \hat{\lambda}_m b_{i,t-1} + v_{i,t}$, and from using firm unconditional averages $\bar{r}_i = \hat{\lambda}_f + \hat{\lambda}_m \bar{b}_i + v_i$. The regression hypotheses and their associated heteroskedasticity adjusted *t*-stats are shown. All variance estimates are scaled by the square root (e.g. standard deviations are reported).