A Nonparametric Approach to Multifactor Modeling

Michael J. Gallagher

Fordham University, Economics Department

(Working Paper)

Abstract

Recent literature has started to explore the use of nonparametric methods to estimate alphas and betas in the conditional CAPM and conditional multifactor models. This paper explores two of the most recent contributions and proposes a third method. Nonparametric estimation of factor modeling involves choosing techniques which are different both technically and in application, but common in the nonparametric literature. The methodology does not impose any functional form on how alphas (pricing errors) or betas (factor loadings) evolve over time. Local data is used in estimations, but of crucial significance is the bandwidth selection or optimal window size. Clearly, observations further away from time t are less relevant in estimating time t alphas and betas, so if we are too far away from time t we potentially have a very large bias. However, if too small a bandwidth is selected, the estimate could be quite noisy, leading to a large variance. A popular technique in the literature is the leave-one-out-cross-validation method which is completely data driven. The researcher may use simulations to illustrate how the optimal window size varies with changes in the underlying unobservable state variables. Another bandwidth selection procedure often used in the literature is the plug-in method. The plug-in method however, relies on choosing an unknown parameter in estimating the optimal window size, while the leave-one-out method is completely data driven. This paper explores the effectiveness of these two methods and proposes a nonparametric estimation of multifactor models using a cross validated local polynomial regression method. Local polynomial regression has emerged as a leading approach to nonparametric estimations of regression functions. Using a completely data driven approach to the joint determination of polynomial order and bandwidth eliminates the ad-hoc approach to determining polynomial order which may not be optimal.

Motivation

Conditional factor models such as the Capital Asset Pricing Model and the Fama French three factor model have provided fertile research material for several decades. The results of much of this research has been ambiguous. Generally, depending on the conditioning information used, models can be constructed to fit the empirical data. Researchers use conditioning information to fit the data, but this information is unobservable ex ante. Nonparametric regression does not require the researcher to specify the functional form of the data being estimated, therefore it seems intuitive to me that factors should be data driven and not predetermined state variables. The aim of this paper is to test conditional factor models using a new nonparametric methodology developed by Hall and Racine (Hall and Racine forthcoming) and compare the results to the two most recent and prominent studies. (Li and Yang 2012) and (Ang and Kristensen 2011).

I examine two different approaches to nonparametric multifactor models currently at the forefront of the most recent literature and propose a third. Li & Yang and Ang & Kristensen test conditional factor models utilizing different but familiar techniques in the nonparametric literature. The problem with conventional tests of conditional models is determining the actual information set with which investors make decisions. As Cochrane (2005 p.143) points out, "Models such as the CAPM imply a conditional linear factor model with respect to investors' information sets. However, the best we can hope to do is to test implications conditioned down on variables that we can observe and include in a test. Thus, a conditional linear factor model is not *testable!*" (his emphasis). Nonparametric estimation, however, attempts to eliminate the need to know the investors information set, and allows the data generating process to *reveal* to us the data with which we should estimate the model. I explore two of the most recent techniques utilized and compare these to a third. I hypothesize the third newly developed nonparametric technique will deliver a better fit with lower pricing errors of the CAPM and the FF models to the data.

- A leave-one-out cross validation method is implemented. Li & Yang utilize this method and their approach is detailed below. According to Li & Yang, this technique generates the optimal data window with which they estimate conditional alphas and betas. I provide results of this method for comparison.
- A global plug-in method is also tested. Ang & Kristensen use a global plug in method to determine optimal bandwidth, also detailed below. Both papers overwhelmingly find that both the CAPM and the Fama French (1993), the two most important asset pricing models in use today, fail to explain the return variations of the portfolios sorted by value premium or past returns. Both models also show factors are indeed time varying.
- I utilize cross validated local polynomial regression, a new nonparametric technique developed by Peter Hall and Jeffery Racine, hoping to improve on the results of Li & Yang and Ang & Kristensen. Cross validated local polynomial regression is emerging as a data driven approach to joint determination of polynomial order and bandwidth, which promises improvements in finite sample efficiency and rates of convergence. I hypothesize the results obtained by this new technique will provide a better fit with lower pricing errors.

Portfolio Modeling

In this chapter I would like to explore briefly the history of portfolio construction and risk assessment. It seems natural a concise recap of financial theory over the last several decades provides a segue into the direction of current research.

Single Index Models

In 1952 Harry Markowitz developed 'Modern Portfolio Theory' for which he was awarded the Nobel Prize. In a nutshell, Markowitz embraced the notion of diversification. He developed the minimum variance frontier; a graph of the lowest possible variance that can be attained for a given portfolio expected return. Obviously a huge contribution to portfolio management, however computationally unwieldy, as the analysis required huge numbers of covariances between assets to be calculated. This is a good place to introduce the simple regression equation for the single index model, the simple relationship which determines the construction details of later models.

$$R_i(t) = \alpha_i + \beta_i R_M(t) + e_i(t) \tag{1}$$

Where a market index is M, excess returns are $R_M = r_M - r_f$, and standard deviation is σ_M. Alpha in the above equation, the intercept, is the security's expected excess return when the market excess return is zero. β, is the security's sensitivity to the index. To understand the power of diversification, consider the following:

$$R_p = \alpha_p + \beta_p R_M + e_p \tag{2}$$

Simply the counterpart to the above equation for portfolios. As the number of assets in the portfolio increase, the risk attributable to any one asset becomes smaller. Consider an equally weighted portfolio, where the weight of each portfolio is $w_i = \frac{1}{n}$ and the excess returns are:

$$\begin{aligned} R_p = \sum_{i=1}^{n} w_i R_i = \frac{1}{n}\sum_{i=1}^{n} R_i = \frac{1}{n}\sum_{i=1}^{n}(\alpha_i + \beta_i R_m = e_i) \\ = \frac{1}{n}\sum_{i=1}^{n}\alpha_i + \left(\frac{1}{n}\sum_{i=1}^{n}\beta_i\right) R_M + \frac{1}{n}\sum_{i=1}^{n} e_i \end{aligned} \tag{3}$$

One can see the sensitivity to the market, given by β_p is:

$$\beta_p = \frac{1}{n}\sum_{i=1}^{n}\beta_i \tag{4}$$

And analogously the alpha;

$$\alpha_p = \frac{1}{n}\sum_{i=1}^{n}\alpha_i \tag{5}$$

and finally an error term;

$$e_p = \frac{1}{n}\sum_{i=1}^{n} {}_i \tag{6}$$

Having partitioned the various components of the excess return, we can say the portfolio variance is

$$\sigma_p^2 = \beta_p^2\sigma_M^2 + \sigma^2(e_p) \tag{7}$$

The systematic risk, or market risk, of the portfolio is $\beta_p^2\sigma_M^2$, and is not diversifiable. The firm specific risk, which we are trying to diversify away, is denoted by $\sigma^2(e_p)$, and the equal weighted portfolio variance is given by:

$$\sigma^2(e_p) = \sum_{i=1}^{n}\left(\frac{1}{n}\right)^2 \sigma^2(e_i) = \frac{1}{n}\bar{\sigma}^2(e) \tag{8}$$

The single index model therefore relegates risk to market risk or firm specific risk. The market is a broad based selection of securities such as the S&P 500. The regression estimate provides both a Beta coefficient which is the asset sensitivity to the broad based index that has been chosen, and an Alpha value. Since the Beta coefficient is commonly agreed upon in investment circles, and well publicised, the Alpha is the key variable which denotes a good or a bad investment. Additionally, Alpha, the premium value above or below a given risk free rate, is a macroeconomic or a security analysis variable, it is the variable which distinguishes between good and mediocre portfolio managers. Hence the often touted mainstream journal title, 'The search for Alpha'. The single index model does lay the foundation for the following, the CAPM.

Capital Asset Pricing Model

The Capital Asset Pricing Model (CAPM) was developed in 1964 by William Sharpe, Sharpe (1964), John Lintner, Lintner (1965), and Jan Mossin, Mossin (1966). This venerable model of asset pricing, developed in 1964, was built on the work of Harry Markowitz from 1952, and the single index model. I mention Markowitz, Sharp, Lintner, as a background to the development of modern asset pricing and risk management. We have indeed come a long way since the keen insights of these giants, but we still have mysteries to solve in the financial world. The equity premium puzzle springs to mind, along with asset pricing bubbles. The CAPM is the cornerstone of many undergraduate and even graduate school courses. The CAPM provides a prediction of the relationship between the risk of an asset and its expected return. Without elaborating on the simplifying assumptions, and perhaps too succinctly, I will derive the CAPM.

1. All investors will hold a portfolio of risky assets in proportions which mimic assets in the **MARKET PORTFOLIO (M)**. The proportion of each stock in the market portfolio equals the market value of the stock divided by the total market value.

2. The Market Portfolio will be on the efficient frontier, in other words, mean/variance efficient. The Market portfolio will also be the tangent portfolio to the optimal capital allocation line. Therefore, the line through the risk free rate, and the market portfolio is the best attainable capital allocation line. Contrast this Capital Market Line (CML) to a Security Market Line (SML), which graphs an individual asset's risk premium as a function of risk. The SML is the benchmark for an asset's performance, a fairly priced asset will plot exactly on the SML. I will return to the discussion of the SML when I discuss Arbitrage Pricing Theory. All investors hold **(M)**, differing only in the amount invested in it versus the market portfolio. Bodie (2009)

3. The premium of the market portfolio will be in proportion to it's risk and the degree of risk aversion of the representative investor.

$$E(r_M) - r_f = \bar{A}\sigma_M^2 \tag{9}$$

   Where σ_M^2 is the variance of the market, and also because M is the optimal portfolio, σ_M^2 is the systematic risk in this universe. $\bar{A}$ is the average degree of risk aversion across investors.

4. The risk premium on any individual asset will be proportional to risk on the market, and the *beta coefficient* of the security relative to the market. Therefore, *beta* measures the extent to which the stock and the market move together. Formally:

$$\beta_i = \frac{Cov(r_i, r_M)}{\sigma_M^2} \tag{10}$$

This brings us to a point where we can say the risk premium on individual securities is:

$$E(r_i) - r_f = \frac{Cov(r_i, r_M)}{\sigma_M^2}[E(r_M) - r_f] = \beta_i[E(r_M) - r_f] \tag{11}$$

Combining the index model with the Beta-Return relationship in expectations, we arrive at the following:

$$E(r_M) - r_f = \alpha_i + \beta_i[E(r_M) - r_f] \tag{12}$$

It is easy to see, that if the Beta-Return relationship is correctly calculated, then α will be equal to zero. In fact, this is the benchmark test whether a Beta or any other factor, or factors, accurately models the returns of an asset.

Therefore, the CAPM very elegantly states the excess return of a stock over the risk free return, is equal to the β of that stock times the market portfolio. This is the starting point for all future research on asset returns, β has been transformed to factors, and a individual assets returns may be modeled with multiple factors according to the whims of the portfolio manager; for example, business cycle risk, interest or inflation rate risk, energy price risk, to name a few. The original CAPM was completely unconditional; the investors information set was not adjusted for time or any other knowable (or guessable) contributions to variation in factors.

Factor and Multifactor Models

The progression of portfolio analysis detailed in the above sections brings me to introduce multifactor models. Very simply, instead of calculating a beta relationship to a market proxy, we calculate a relationship to broader sources of risk. More specifically, we quantify the relationship of a asset to broader macroeconomic variables, variables which are pertinent in particular to the class of assets we are valuing. One can easily see the intuitive appeal of such a powerful tool. Consider the model:

$$r_i = E(r_i) + \beta_{iGDP} + \beta_{iIR} + e_i \tag{13}$$

The return of an asset is equal to the expected return plus the asset's sensitivity to the factors; in this case GDP and the Interest Rate. Surely any factors could be considered, and indeed powerful algorithms are employed to determine appropriate factors, and choice of factors and their Betas nowadays distinguishes good managers from the average. Multifactor models have become the norm in portfolio and risk management, therefore it is important to assess their validity and seek improvements.

Arbitrage Pricing Theory

Arbitrage is the simultaneous purchase and sale of securities to profit from discrepancies in their prices. I would be amiss if I didn't give a brief mention of the contribution of Arbitrage Pricing Theory (APT) to asset pricing. The law of one price states that identical assets should be worth the same. If they were not, one could simultaneously purchase the underpriced and sell the overpriced for a riskless profit. Developed in 1976 by Stephen Ross, Ross (1973) APT uses a single or multifactor model to depict a factor Beta, the relationship between a factor and an asset's expected return. Of course the Security Market Line, generated from the particular model, using predetermined factors, is the line of fairly priced assets. Any assets over or under the line are over or under priced, and will immediately be bought and sold by the market to bring prices back in line.

Fama French 1993

Lastly in this brief history of the development of asset pricing models is the famed Fama-French three factor model. Fama (1993) Developed in 1993 by Eugene Fama and Kenneth French, the model has dominated both empirical research and application in business for the past twenty years. Essentially, rather than choosing macroeconomic variables as factors for market risk, Fama and French determined that based on past evidence, certain variables seemed to predict average returns well, and hence are capturing risk premiums. The Fama-French three factor model:

$$r_{it} = \alpha_i + \beta_{iM} R_{Mt} + \beta_{iSMB} SMB_t + \beta_{iHML} HML_t + e_i \tag{14}$$

where the factors are:

SMB: Small minus Big. The return on a portfolio of small stocks in excess of the return on a portfolio of large stocks.

HML: High minus Low. The return on a portfolio of stocks with high book to market ratio in excess of the return on a portfolio of stocks with a low book to market ratio. Immediately one must think, these factors are not really proxies for any kind of macroeconomic risk, Fama and French would argue that they do. They would say firms with high book to market ratios are more likely to be in financial distress, and small size stocks are more sensitive to business cycle risk. In any event, whether you agree with the factors or not, the Fama and French model is taught in Finance courses and is in use today all over the world.

Nonparametric Regression

Nonparametric econometrics do not require the researcher to know the specific functional form of the objects being estimated. Rather, a nonparametric approach allows the data to determine the functional form. These methods are becoming increasingly popular when standard parametric econometrics have failed to adequately describe a model. The revered Capital Asset Pricing Model (CAPM) and the Fama-French (FF) models are the cornerstone of many undergraduate and graduate school finance courses. However, there is a wealth of literature which questions the validity of both of these models. It seems to me exploring models in which the data actually determines the functional form could possibly yield fruitful results in improving the way we look at asset pricing and risk management.

Central to nonparametric techniques is the use of a kernel density estimator. Consider the use of a histogram to model a distribution; the researcher must select an origin and bin width, and the results are particularly sensitive to both choices. A kernel density estimator was constructed to alleviate these limitations. We illustrate with the following, often termed the Rosenblatt-Parzen estimator. Rosenblatt (1956), Parzen (1962).

$$\hat{f} = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right) \tag{15}$$

Common to all methods of nonparametric estimation are the problems of *bias* and *variance*. *Bias* occurs if too large a data window is used in an estimate, resulting in points far from a time t affecting the estimate. *Variance* occurs if too few points are used around time t, resulting a noisy estimate. Pointwise, one can utilize mean square error to access properties of the kernel. Consider:

$$mse\hat{f}(x) = E\{\hat{f}(x) - f(x)\}^2 = var\hat{f}(x) + \{bias\hat{f}(x)\}^2 \tag{16}$$

With a Taylor expansion, we can approximate the bias and variance: Pagan and Ullah (1999)

$$bias\hat{f}(x) \approx \frac{h^2}{2} f''(x) k_2 \tag{17}$$

$$var\hat{f}(x) \approx \frac{f(x)}{nh}^2 (z)dz \tag{18}$$

One can easily see the bias and variance are dependent upon the bandwidth; bias is smaller with a smaller h and variance is larger with a smaller h. Herein is the crux of the estimation problem; choosing the "optimal" bandwidth!

The above equations provide pointwise estimates, when in fact we are after a aggregate mean square error which will provide a global error measure. The integrated mean square error provides such a measure. Li and Racine (2007)

$$\begin{aligned}
IMSE\hat{f}(x) &= \hat{f}(x)dx \\
&= \int var\hat{f}(x) + \int \left\{bias\hat{f}(x)\right\}^2 dx \\
&\approx \int \left[\frac{f(x)}{nh} \int K^2(z)dz + \left\{\frac{h^2}{2} f''(x) k_2\right\}^2\right] dx \\
&= \frac{1}{nh} \int K^2(z)dz \int f(x)dx + \left\{\frac{h^2}{2} k_2\right\}^2 \int \{f''(x)\}^2 dx \\
&= \frac{\Phi_0}{nh} + \frac{h^4}{4} k_2^2 \Phi_1
\end{aligned} \tag{19}$$

Minimizing Equation 19 with respect to h obtains a global bandwidth which will balance bias and variance:

$$\begin{aligned}
h_{opt} &= \Phi_0^{1/5} k_2^{-2/5} \Phi_1^{-1/5} n^{-1/5} \\
&= \left\{\frac{\int K(z)dz}{(\int z^2 K(z)dz)^2 \int\{f''(x)\}^2 dx}\right\}^{1/5} n^{-1/5} = cn^{-1/5}
\end{aligned} \tag{20}$$

The reader will note I have neglected to discuss optimal choice of kernel, and instead have solely focused on bandwidth optimization. This is because while bandwidth selection is crucial to correctly specify a model, kernel selection is not. It is well known in the nonparametric literature various choices of kernel provide efficient density estimation, and can be chosen based on computational considerations. Silverman (1986) In fact one of the models I delve into deeper, the Yang and Li model, use an Epanechnikov kernel, while the other research I explore, the Ang and Kristensen model, use a Gaussian Kernel.

There are four general approaches to bandwidth selection; (1) reference rule of thumb, (2) plug-in, (3) cross-validation, and (4) bootstrap.

Reference rule of thumb, as the name suggests, uses a standard family of distributions to assign a value to the unknown constant $\int\{f''(x)\}^2 dx$. For a normal family of distributions, σ is typically used.

The Plug-in method involves plugging in an estimate of $\int\{f''(x)\}^2 dx$ to the optimal bandwidth formula based on a preconceived notion of the bandwidth such as is done with reference rule of thumb. Loader Loader (1999) has this to say about plug in methods. "When the results are analyzed carefully, the much touted plug-in approaches have fared rather poorly, being tuned largely by arbitrary specification of pilot bandwidths and being heavily biased when this specification is wrong."[1]

Cross Validation Methods include least squares cross validation, and likelihood cross validation. Least squares has intuitive appeal in that it is fully data driven. A bandwidth is selected which minimizes the IMSE, the difference between $\hat{f}(x)$ and $f(x)$.

$$\int \left\{ \hat{f}(x) - f(x) \right\}^2 dx = \int f(x)^2 dx - 2 \int \hat{f}(x) f(x) dx + \int f(x)^2 dx \tag{21}$$

One of the problems associated with least squares is it is sensitive to discretized data, and small scale effects in the data. Indeed, Ang & Kristensen discuss this: they abandoned their test of cross validation because the data windows were too small to make intuitive economic sense. Likelihood cross validation chooses a bandwidth h to maximize the (leave one out) log likelihood function:

$$L = logL = \sum_{i=1}^{n} log \hat{f}_{-i}(x) \tag{22}$$

In other words, $\hat{f}_{-i}(x)$ is the leave one out kernel estimator of $f(X_i)$ that uses all points except X_i to construct the density estimate.

$$\hat{f}_{-i}(x) = \frac{1}{(n-1)h} \sum_{j=1,j}^{n} K\left(\frac{X_j - x}{h} \right) \tag{23}$$

A shortcoming of the likelihood cross validation is that it can over smooth fat tail distributions.

Lastly, I'll introduce bootstrap methods. Basically,the method is to estimate the IMSE in equation 19, and then minimize it over all bandwidths. Needless to say this method is computationally demanding, and also when the objective function is stochastic creates numerical minimization issues.

[1]See Loader 1999 p.495.

Fama and Macbeth, 1973

Although some time ago, this paper deserves a brief mention. Fama and MacBeth tested the two parameter model in the early seventies, and variations of their cross sectional regression testing has become a popular choice in testing empirical financial models. The Fama MacBeth procedure is a two step process, first, asset betas are computed for every asset and period using time series regressions with a number of previous periods, usually three to five years. Secondly, cross sectional regressions of returns are run at every period, generating a time series of estimated slope coefficients. The sample mean of this series provides the constant slope estimator. Fama (1973)

Gibbons, Shanken, and Ross, 1989

Gibbons, Shanken and Ross (GSR) develop a widely used method of testing the efficiency of a given portfolio. They draw our attention to the fact that the CAPM has provided bountiful research material, but then call into question whether the tests have actually been useful. They note the numerous tests of the CAPM using various adaptations of regression techniques. A notable example is the Fama MacBeth (1973) study I looked at earlier. Usually, cross-sectional regression are run of asset returns on estimated beta coefficients, and estimates of the slope are reported. The question they raise is if these tests are even tests of the Sharpe-Lintner theory. In order to truly test the CAPM, one should use the "market portfolio", but many researchers have used proxy portfolios, and furthermore sometimes data are grouped together to reduce measurement error. This grouping of the proxy portfolios reduces the power of the regression tests, so even if the portfolios reflect the "market", the power of the tests is unknown. Basically (GSR) claim to develop a canonical example of a test using multivariate statistical methods. They test whether any portfolio is ex ante mean variance efficient. In other words, they have developed a tractable multivariate statistic for testing whether a portfolio is mean variance efficient.

The null of any factor model is of course that alpha equals zero. GSR ask us to consider the following:

$$\tilde{r}_{it} = \alpha_{ip} + \beta_{ip}\tilde{r}_{pt} + \tilde{\epsilon_{it}} \qquad \forall i = 1, ..., N \tag{24}$$

where $\tilde{r}_{it}$ is defined as the excess return on asset i in period t, $\tilde{r}pt$ is defined as the excess return on the portfolio whose efficiency is being tested, and $\tilde{\epsilon}_{it}$ is an independent normally distributed disturbance term with mean zero and nonsingular covariance matrix Σ conditional on excess returns for portfolio p. When a portfolio is mean variance efficient, then the following first order condition is satisfied:

$$\xi(\tilde{r}_{it}) = \beta_{ip}\xi(\tilde{r}_{pt}) \tag{25}$$

The above first order condition along with the distribution assumption allow for the following null hypothesis:

$$H_0 : \alpha_{ip} = 0, \qquad \forall_i = 1, ..., N \tag{26}$$

This is essentially the same as testing the CAPM except $\tilde{r}_{pt}$ replaces the market portfolio. Additionally, instead of reporting the significance of alpha across all N equations, they report N univariate t statistics for each equation.[2] If regressions are run using ordinary least squares for each equation, the intercepts will have a multivariate normal distribution conditional on

[2] Given the normality assumption, the null hypothesis can be tested using "Hotelling's T^2 test" a a multivariate generalization of the univariate t-test. Gibbons, Shanken, and Ross (1989)

$\tilde{r}_{pt}$ $(\forall t = 1, ..., T)$) with:

$$\sqrt{T/(1+\hat{\theta}_p^2)\hat{\alpha}_p} \sim N\ \sqrt{T/(1+\hat{\theta}_p^2)\alpha_p ; \Sigma} \tag{27}$$

Where T is the number of time series, $\hat{\alpha}_p' \equiv (\hat{\alpha}_{1p}, \hat{\alpha}_{2p}, ..., \hat{\alpha}_{Np})$; $\hat{\theta}_p \equiv \bar{r}_p / s_p$; $\bar{r}_p \equiv$ the sample mean of $\tilde{r}_{pt}$ and $s_p^2 \equiv$ sample variance of $\tilde{r}_{pt}$[3] Given the distribution:

$$W_u \equiv \hat{\alpha}_p' \hat{\Sigma}^{-1} \hat{\alpha}_p / (1 + \hat{\theta}_p^2) \tag{28}$$

where $\hat{\Sigma}$ is an unbiased residual covariance matrix, and a noncentrality parameter λ given by:

$$\lambda = \left[T/(1+\hat{\theta}_p^2) \right] \alpha_p' \Sigma^{-1} \alpha_p \tag{29}$$

This constitutes the test developed by (GSR), under the null hypothesis, $\alpha_p = 0$ and $\lambda = 0$ and a central F distribution is constructed. Moreover, the distribution provides a way to study the power of the test. Gibbons et al. (1989)

Jagannathan and Wang, 1996

Jagannathan and Wang (JW) are one of the first to test the CAPM in a conditional sense. In other words, they allow betas and the market risk premium to vary over time. (JW) qualify their research by noting that in the twenty years prior to their work, a number of studies have empirically examined the performance of the static version of the CAPM. The results of these studies fail to support the view that the CAPM explains the cross-sectional variation in average returns. (JW) highlight the importance of the findings of Fama and French just a few years previous. Of course as mentioned earlier, the findings of Fama and French that the static CAPM is unable to explain the cross section of average returns are so contrary to the CAPM that they shake the foundations on which MBA and other managerial course materials in finance are built. Jagannathan (1996) The significant contribution of the (JW) work is to let betas vary over time. In the traditional CAPM model investors live for one period, not a terribly realistic assumption. (JW) adapt the stance that it is very likely the relative risk of a firms cash flow is going to vary over time, therefore the beta and expected return will depend on the information available at any point in time and will vary over time. Therefore, (JW) assume the conditional version of the CAPM holds, that is the expected return on an asset conditioned on the available information set at the time, and it is linear in its conditional beta. Following the numerous studies which have shown the inability of the unconditional CAPM to predict returns, (JW) elaborate on the necessary contributions of the conditional CAPM. Clearly it has become imperative, and in fact it has become common practice, to allow time variability in betas. The direction (JW) take is to derive an unconditional asset pricing model starting with a conditional version of the CAPM. Given a conditional CAPM:

$$E[R_{it} \mid I_{t-1}] = \gamma_{0t-1} + \gamma_{1t-1} \beta_{it-1} \tag{30}$$

where β_{it-1} is the conditional beta of asset i, or:

$$\beta_{it-1} = \frac{Cov(R_{it}, R_{mt} \mid I_{t-1})}{Var(R_{mt} \mid I_{t-1})} \tag{31}$$

[3] Furthermore $\hat{\alpha}_p$ and $\hat{\Sigma}$ are independent with $(T-2)\hat{\Sigma}$ having a Wishart distribution with parameters $(T-2)$ and Σ. This implies that $(T(T-N-1)/N(T-2))W_u$ has a noncentral F distribution with degrees of freedom N and $(T-N-1)$. Gibbons, Shanken, and Ross (1989)

They take the unconditional expectation of both sides of equation 3.1:

$$E[R_{it}] = \gamma_0 + \gamma_1 \bar{\beta}_i + Cov(\gamma_{1t-1}, \beta_{it-1}) \tag{32}$$

Jagannathan and Wang then decompose the expected beta $\bar{\beta}_i = E[\beta_{it-1}]$ into two orthogonal components by projecting the conditional beta on the market risk premium. In other words, (JW) have constructed a two factor model. Not a two factor model in the spirit of Ross (1976), but a first factor, a beta which measures average returns in the traditional sense, and a second factor, a "beta instability" coefficient which reports the sensitivity of the beta to market returns. For every asset they define a $beta - premsensitivity$:

$$\zeta_i = Cov(\beta_{it-1}, \gamma_{1t-1})/Var(\gamma_{1t-1}) \tag{33}$$

and a residual beta:

$$\eta_{it-1} = \beta_{it-1} - \bar{\beta}_i - \zeta_i(\gamma_{1t-1} - \gamma_1) \tag{34}$$

Hence using the beta premium sensitivity, (JW) are able to take a conditional model with varying betas, and using a further breakdown of the beta coefficient, fit the model in an unconditional sense. In other words, the expected return on any asset is a linear function of its expected beta and its beta-prem sensitivity. As I mentioned, the contribution of Jagannathan and Wang is a significant departure from a two beta linear factor model proposed by Ross, firstly, they do not assume returns have a linear factor structure as is commonly understood, and second, γ_{1t-1} is a predetermined variable not a factor as in the traditional sense. Hence it is a significant contribution to moving research towards the more modern versions of multifactor models.

Lewellen and Nagel, 2006

Jonathan Lewellen and Stefan Nagel, (LN) Lewellen (2006) test the conditional CAPM using short window regressions. They challenge the findings of Jagannathan and Wang, 1996, and others (e.g., Jensen, 1968; Dybvig and Ross, 1985) who found a stock's conditional alpha might be zero, while it's unconditional alpha is not, if it's beta is allowed to change through time and is correlated with the equity premium or with market volatility. Lewellen (2006) Using short window regressions they estimate time series of conditional alphas and betas for size, Book/Market (B/M), and momentum portfolios from 1964-2001. The alpha estimates are a direct test of the conditional CAPM, average conditional alphas should be zero if the conditional CAPM holds, but their research shows alphas are statistically significant and close to the unconditional alphas. The findings of Lewellen and Nagel counter the findings of Jagannathan and Wang. In other words, Lewellen and Nagel (LN) find that a conditional CAPM which allows betas to vary over time performs nearly as poorly as the unconditional CAPM. One caveat is that Jagannathan and Wang used cross sectional regressions while LN used time series regressions.

LN focus on the asset pricing anomalies uncovered by work of Fama and French in the early nineties. Namely that small stocks outperform large stocks (size), firms with high B/M ratios outperform those with low B/M ratios (value premium), and stocks with high prior returns continue to outperform those with low prior returns (momentum). They argue that theoretically the conditional CAPM could hold perfectly every period even though the assets are mispriced by the unconditional CAPM. A stocks conditional alpha could be zero, with the unconditional alpha not equal to zero, if the beta is allowed to change through time and it is correlated with the market premium or with market volatility. In other words, the market portfolio might be conditionally mean-variance efficient in each period, but at the same time not on the unconditional mean-variance efficient frontier. Lewellen (2006) The point is illustrated as follows, assume market volatility is constant, if the conditional CAPM holds, the stock's unconditional alpha depends primarily on the covariance between its beta and the market risk premium. $\alpha^u \approx cov(\beta_t, \gamma_t)$ If a stock's monthly beta has a standard deviation of 0.3, and the monthly risk premium has a standard deviation of 0.5% then the upper bound if β_i and γ_i were perfectly correlated would be at most 15%. However, the B/M strategy has an alpha of 0.59% monthly with a standard error of 0.14% and the momentum strategy has an alpha of 1.01% with a standard error of 0.28%. Both would be much too large to be explained by time varying betas according to Lewellen and Nagel. Lewellen and Nagel are not contesting that betas vary over time, they just believe that betas do not vary enough to explain the size of the pricing errors in the unconditional models.

Lewellen and Nagel use short window regressions; every month, quarter, half-year, or year using daily weekly or monthly returns. They estimate *time series* of conditional alphas and betas. Of course the alphas should be zero, but instead are large and statistically significant. The betas do vary over time, generally with variables used to measure business conditions. LN found no evidence that betas co-vary with market risk premium in a way which might explain the unconditional alphas. Additionally they conclude that betas do not vary in such a way as to account for the asset pricing anomalies mentioned above. Also as mentioned previously, Jaganathan and Wang implement cross sectional regressions, while Lewellen and Nagel utilize time series intercept tests, and they question the validity of the cross sectional regressions JW actually used.

Ferreira, Gil-Bazo, and Orbe, 2011

Ferreira, Gil-Bazo, and Orbe, (FGO), further the quest to develop beta pricing models which move away from the traditional assumption of constant betas and towards the idea of time varying risk premia. They have developed a nonparametric method to estimate conditional beta pricing models that allows for flexibility in dynamics of betas. Ferreira (2011) Their test is an extension of the Fama-Macbeth procedure summarized earlier. As discussed, the Fama MacBeth procedure is a two step process, first, asset betas are computed for every asset and period using time series regressions with a number of previous periods, usually three to five years. Secondly, cross sectional regressions of returns are run at every period, generating a time series of estimated slope coefficients. The sample mean of this series provides the constant slope estimator. FGO adapt this procedure in their research, however, conditional betas are assumed to be smooth functions of state variables. Then instead of using cross sections of returns and betas, they use the entire sample. The trick is they use a Seemingly Unrelated Regression Equations model where each equation in the model correspond to one asset. The slope coefficients are free parameters which are allowed to vary freely through time but constrain the slopes to be equal across assets. Therefore, they are able to estimate betas under no specific parametric structure. Results from FGO indicate that the nonparametric estimator clearly outperforms the traditional rolling window estimator under all specifications.

Li & Yang (2011) and Ang & Kristensen (2012)

In addition to the above referenced work, there are two recent studies which I will primarily use in my research. Yan Li at Temple University and Liyan Yang at University of Toronto have done extensive research. Testing Conditional Factor Models: A nonparametric Approach, (2011). Yan Li and Liyan Yang use a leave-one-out cross-validation method to nonparametrically determine alpha and betas in conditional factor models. In a contemporaneous study, Testing Conditional Factor Models. Ang and Kristensen (2012) use a plug in method, also a common nonparametric methodology. Both studies use similar data. In the following sections I explore these studies in depth, as the basis for my research is to compare the results from these two researchers to the results I obtain from the nonparametric procedure developed by Hall and Racine.

Data & Methodology

The methodology I would like to utilize to run models is detailed in the forthcoming work by Jeffery Racine and Peter Hall. (Hall and Racine forthcoming) Hall and Racine propose a Data Generating Process (DGP) which is a leave one out cross validation model which jointly determines bandwith and polynomial order. This new methodology in nonparametrics has not been applied to factor models, and I think it should be very interesting indeed to compare the results using this new methodology to the existing studies.

The Data is readily available on Dr. Kenneth French's website, and I am choosing to also use data used in the most recent studies to facilitate comparison. From Kenneth French's data library, I'll use 25 Size-Book to Market portfolios. I'll form six size and B/M portfolios. S is the average of the five portfolios in the lowest size quintile, B is the average of the five portfolios in the highest size quintile, and S-B is the difference. G is the average of the five portfolios in the lowest B/M quintile, V is the average of the five portfolios in the highest B/M quintile, and V-G is the difference. The three momentum portfolios are constructed based on the ten momentum portfolios, where I let W stand for the winner portfolio, the portfolio with the highest return among the 10 momentum portfolios; L for the loser portfolio, the portfolio with the lowest return among the 10 momentum portfolios; and W-L for their difference. I will use both value weighted and equally weighted portfolios using daily data from July 1963 to December 2007.

Nonsynchronous Trading

It is well known that some stocks, whether they are thinly traded or illiquid stocks, tend to react with a week or more delay to common news, so a daily beta will miss the small stock covariance with market returns. Lo and Craig MacKinlay (1990) This delayed reaction can affect beta's and should be addressed. One of the methods used in the literature is to use Dimson regressions. Dimson (1979) Dimson regressions include both current and lagged returns and betas are just the sum of all lags. This mechanism allows returns to be included for stocks which trade infrequently. As per Dimson, for daily returns, four lags of market returns are included.

$$R_{i,t} = \alpha_i + \beta_{i0} R_{M,t} + \beta_{i1} R_{M,t-1} + \beta_{i2} \frac{[(R_{M,t-2} + R_{M,t-3} + R_{M,t-4}]}{3} + \epsilon_{i,t} \qquad (35)$$

The daily beta is then: $\beta_i = \beta_{i0} + \beta_{i1} + \beta_{i2}$ Weekly returns similarly are:

$$R_{i,t} = \alpha_i + \beta_{i0} R_{M,t} + \beta_{i1} R_{M,t-1} + \beta_{i2} R_{M,t-2} + \epsilon_{i,t} \qquad (36)$$

Finally, monthly betas are estimated with just one lag of market return.

$$R_{i,t} = \alpha_i + \beta_{i0} R_{M,t} + \beta_{i1} R_{M,t-1} + \epsilon_{i,t} \qquad (37)$$

I compare the alphas and betas generated with Dimson regressions and with compounding returns to alphas and betas generated with a simple ordinary least squares regression. Table 1 and Table 2 along with Figure 2 and Figure 3 report the results. The remainder of my empirical tests do not use Dimson regressions or compounding, however I felt it was necessary to explore the difference.

Compounding Returns

The effect of compounding across differing time periods also needs to be addressed. Alphas and betas for different time periods should be slightly different because of compounding. Lewellen and Nagel address this issue with the following microstructure: If daily returns are I.I.D. then expected N-day returns are $e[1 + R_i]^N - 1$ and N-day beta is:

$$\beta_i(N) = \frac{E[(1 + R_i)(1 + R_M)]^N - E[1 + R_i]^N E[1 + R_M]^N}{E[(1 + R_M)^2]^N - E[1 + R_M]^{2N}} \tag{38}$$

One can see the betas will spread out as the time horizon lengthens: $\beta_i(N)$ increases in N if $\beta_i(1) > 1$ and decreases in N if $\beta_i(1) < 1$. In either event, again, the results are reported in Table 1 and Table 2 along with Figure 2 and Figure 3, and one can see the differences are negligible.

The Ordinary Least Squares Unconditional Model

This section presents the results from a traditional unconditional parametrically specified CAPM and FF model. The data is of course the same, the purpose is to establish a benchmark against which the recent work and my research can be compared. Table 3 reports the ordinary least squares unconditional results.

A Leave one out Cross Validation Approach

This section discusses the contributions of Li & Yang, one of the methods I explore and compare to the model I'm using. In a conditional model, the traditional approach is to allow factor loadings to depend on observable state variables, however estimates of factor loadings are very sensitive to instrumental state variables. Li & Yang propose a nonparametric method, no functional form is imposed on how alphas and betas evolve over time. A leave-one-out-cross-validation method gives the optimal window size for estimation, and windows overlap allowing gradual changes in betas. They propose two tests to evaluate performance, the first on individual pricing errors, the second on average pricing errors. They test the ability of both the CAPM and the FF to explain return variations on portfolios sorted by size, book to market, and past returns. Unlike Lewellen and Nagel (2006) who use arbitrarily divided windows of months, quarters, half-years, and years, Li and Yang (2011) provide an optimal window for estimating, and find that estimates are very sensitive to the choice of window size. In a contemporaneous work, Ang and Kristensen (2011) also use a nonparametric method to test factor models. On the technical side, compared to Ang and Kristensen, first, Li and Yang use simulations to illustrate how optimal window size varies with changes in underlying unobservable state variables. The simulations show that if the underlying state variable is more volatile; alphas and betas change more quickly thereby requiring a smaller window size for estimation. Second, Ang and Kristensen use a plug in method to determine window size. The plug in method, while popular in the nonparametric literature, requires the estimation of some unknown parameter. The leave one out cross validation method is completely data driven. Third, Li and Yang use a fitness measure test to determine if alphas and betas are indeed time varying, while Ang and Kristensen use a constancy test looking at the generalized likelihood ratio. On the application side, Ang and Kristensen look at only value weighted portfolios, while Li and Yang test both value weighted and equally weighted.

Li & Yang use the following Econometric Framework in the research I analyze. Consider the following:

$$R_{i,t+1} = \alpha_{i,t+1} + \beta_{i,t}\, f_{t+1} + \epsilon_{i,t+1}\,, t = 1, 2, ...T \tag{39}$$

Where $R_{i,t+1}$ is excess return for asset i at time $t+1$, and $f_{t+1} = (f_{1,t+1}\,, f_2, t+1, ...f_{K,t+1}\,)$ is excess returns for $K \geq 1$ portfolios. For the conditional CAPM, $f_{t+1} = R_{m,t+1}$, where $R_{m,t+1}$ is the market excess return. For the conditional FF,$f_{t+1} = (R_{m,t+1}\,, SMB + t + 1, HML_{t+1})$ where SMB and HML are returns on portfolios for size and book-to-market factors. To allow for heteroskedastic errors assume $E_t(\epsilon_{i,t+1}\,|f_{t+1}) = 0$ and that $E_t(\epsilon^2_i{}_{,t+1}\,|f_{t+1}) = \sigma^2_i(f_t + 1)$

Asset $i's$ alpha and beta at time t are $\alpha_{i,t} \in R$ and $\beta_{i,t} \in R^K$ which are contingent on the investors information set at time t. Unfortunately, investors information set is not known at time t, so Li and Yang, like (Lewellen and Nagel 2006) do not specify conditioning variables but assume betas are relatively stable within adjoining periods. Therefore, alphas and betas are estimated directly without preconceived conditioning variables and they are allowed to vary over time.

Estimation of the model: Using an appropriate window size, discussed in the next section, the data in each window estimates the conditional alpha and beta corresponding to time t: $\alpha_{i,t}\, and \beta_{i,t}\, (t = 1, 2, ...T)$ Assume the sequences of alphas and betas are generated from two smoothing functions $\alpha_i : [0,1] \mapsto R$ and $\beta_i : [0,1] \mapsto R^K$ which gives us:

$$\alpha_{it} = \alpha_i(\frac{t}{T})\; and\; \beta_{i,t} = \beta_i(\frac{t}{T}) \tag{40}$$

For a sample size T, every observation is mapped to an interval between 0 and 1 by doing a transformation of t. Then choose $\alpha_i(.)$ and $\beta(.)$ to minimize the following local sum of squared residuals.

$$\min_{\alpha_i(.),\beta_i(.)} \sum\nolimits_{s=t-[Th]}^{t+[Th]} [R_{i,s+1} - \alpha_i(s/T) - \beta_i(s/T) f_{s+1}]^2 k_{st} \tag{41}$$

Where $k_{st} = \frac{1}{h} k(\frac{s-t}{Th})$, s is a particular data point within the window, $R^{i,s+1}$ is asset i's excess return at time $s+1$, and f_{s+1} is the factor return at time $s+1$ With T the sample size, and h the 'bandwidth', $(0 \leq 1)$. The distribution around point t from $t-[Th]$ to $t+[Th]$ is used to estimate $\alpha_i(.)$ and $\beta_i(.)$ using the following Epanechnikov kernel.

$$k(u) = \frac{3}{4}(1-u^2), for\; |u| \leq 1 \tag{42}$$

The two choice variables in equation (41) , $\alpha_i(.)$ and $\beta_i(.)$, are estimated using local data non-parametrically, ie., there are no restrictions placed on their functional form. Researchers typically use two methods to further simplify these functions, the "local linear smoothing method," or "the Nadaraya-Watson method." The local linear smoothing method uses a first order taylor expansion around time t within window $(t-[Th], t+[Th])$ so that $\alpha_i(\frac{s}{T}) = \alpha_0 + \alpha_1.\frac{s-t}{T}$ and $\beta_i(\frac{s}{T}) = \beta_0 + \beta_1.\frac{s-t}{T}$. Additionally if $\alpha_1 = \beta_1 = 0$ in other words, if $\alpha_i(.)$ and $\beta_i(.)$ are constant within the estimation window, then it is called the the Nadaraya-Watson or the local constant smoothing method. Li and Yang report only the local constant smoothing method. Therefore equation (41) becomes

$$\min_{\alpha_0,\beta_0} \sum\nolimits_{s=t-[Th]}^{t+[Th]} (R_{i,s+1} - \alpha_0 - \beta_0 f_{s+1}]^2 k_{st} \tag{43}$$

Using this local constant smoothing method, the estimates for constants α_0 and β_0 become the estimates for time t conditional alpha and conditional beta. The minimization problem equation (43) provides the estimates of alpha and beta, but it is contingent on the optimal 'window' or bandwidth size. Now we take a look at the choice of window size.

Choice of window size: As was mentioned in the introduction, choice of bandwidth or window size is critical in nonparametric estimation. If the window is too large, data far away from time t which is of little relevance is being used in the estimate, this can create substantial bias in the estimate. Conversely, if the window is too small, there are too few data points used in the estimation, which can create substantial variance in the estimate. Therefore the optimal window size serves to balance bias and variance. Local estimates of alpha and betas are measured by Mean Square Error (MSE), however, common to nonparametric techniques, the global performance of the estimated alphas and betas is measured by the Integrated Mean Square Error (IMSE) which is a sum of time-t MSE across all time periods.

$$IMSE(h) \sum_{t=1}^{T} E[(\alpha_{i,t} + \beta_{i,t} f_{t+1} - \hat{\alpha}_{i,t} - \hat{\beta}_{i,t} f_{t+1})^2] \tag{44}$$

Li and Yang use the familiar "leave-one-out-cross-validation" method to estimate the optimal h which minimizes IMSE(h). One should define leave one out estimators $\hat{\alpha}_{0-i,t}$ and $\hat{\beta}_{0-i,t}$ from this regression.

$$\min_{\alpha_0,\beta_0} \sum\nolimits_{s=t-[Th],s}^{t+[Th]} (R_{i,s+1} - \alpha_0 - \beta_0 f_{s+1}]^2 k_{st} \tag{45}$$

where $\hat{\alpha}_{0-i,t}$ and $\hat{\beta}_{0-i,t}$ are the estimates for asset i at time t. Define:

$$CV(h)\sum_{t=1}^{T}(R_{i,s+1}-\hat{\alpha}_{0-i,t}-\hat{\beta}_{0-i,t}f_{s+1})^2 \tag{46}$$

The optimal window size h^* is chosen to minimize equation (46).

A Plug-in Approach

This section is specifically concerned with the work of Ang and Kristensen. Ang and Kristensen also test conditional factor models, however they use a slightly different methodology.

The null of the factor model is that an asset's excess return should be zero after controlling for that assets systematic exposure. Most of the literature assume factor loadings are constant, however there is overwhelming evidence that factor loadings vary over time, particularly for the standard CAPM and Fama French (1993) models. Ang and Kristensen estimate alphas and betas nonparametrically using techniques similar to those found in the literature on realized volatility. Ang and Kristensen also develop estimators of long run alphas and betas which are the averages of conditional alphas or factor loadings across time. The estimators are analyzed in a continuous time setting however all the estimations are of discrete time models. This work builds on the literature using short windows of high frequency data to estimate time-varying betas. Similar to Lewellen & Nagel, Ang & Kristensen use local information to obtain estimates of conditional alphas and betas, and then extend this literature in several important ways. First, they provide a formal distribution theory for conditional and long run estimators and develop data-driven methods for choosing optimal window widths used in estimation. Secondly, they use kernel methods to estimate time varying betas and are able to use all the data efficiently in estimating conditional alphas and betas. Previous studies have used truncated backward looking windows to estimate, and therefore have larger MSE's compared with the two sided kernel. Lastly, Ang and Kristensen develop tests for the significance of conditional and long run alphas across assets in the presence of time varying betas. Prior literature restricts attention to only single assets, whereas the ability to test whether alphas are equal to zero across assets is useful for determining if a strong relationship exists between conditional assets and firm characteristics.

Ang & Kristensen utilize the following econometric framework: Let $R = (R_1, ..., R_M)'$ denote a vector of excess returns of M assets observed at n time points, $0 < t_1 < t_2 < ... < t_n < T$, within a time span $T > 0$. The returns should be explained through a set of J common factors, $f = (f_1, ..., f_j)'$ observed at the same time points. The following conditional model should explain the returns of stock $k(k = 1, ..., M)$ at time $t_i(i = 1, ..., n)$:

$$R_{k,j} = \alpha_k(t_i) + \beta_k(t_i)'f_i + \omega_{kk}(t_i)z_{k,j} \tag{47}$$

where $R_{k,j}$ and f_i are the observed return and factors respectively at time t_i. Similarly, in matrix notation:

$$R_i = \alpha(t_i) + \beta(t_i)'f_i + \Omega^{1/2}(t_i)z_i \tag{48}$$

where $\alpha(t) = (\alpha_1(t), ..., \alpha_M(t))' \in R^M$ is the vector of conditional alphas across stocks $k = 1, ..., M$ and $\beta(t) = (\beta_1(t), ..., \beta_m(t))' \in R^{JxM}$ is the corresponding matrix of conditional betas. Ang & Kristensen produce time series estimates of the realized conditional alphas, $\alpha(t)$ and conditional factor loadings, $\beta(t)$, and their respective standard errors.

Ang & Kristensen propose the following local least squares estimators of $\alpha_k(t)$ and $\beta_k(t)$ for asset k at any time $0 \leq t \leq T$:

$$[\hat{\alpha}_k(t), \hat{\beta}_k(t)']' = \arg\min_{(\alpha,\beta)} \sum_{i=1}^{n} k_{h_t T}(t_i - t)(R_{k,i} - \alpha - \beta' f_i)^2, \tag{49}$$

In the above, $K_{h_k T}(t_i - t)(z) \equiv K(z/(h_k T))/(h_k T)$ where $K(.)$ is the kernel and $h_k > 0$ the bandwidth. Solving equation (49) results in the optimal estimators which are kernel weighted least squares.

$$[\hat{\alpha}_k(t), \hat{\beta}_k(t)']' = \left[\sum_{i=1}^{n} K_{h_k} T(t_i - t) X_i X_i'\right]^{-1} \left[\sum_{i=1}^{n} K_{h_k} T(t_i - t) X_i R_{k,i}\right] \tag{50}$$

The above estimators give weights to individual observations according to how close they are in time to point t. The shape of the kernel determines how the observations are weighted. Ang and Kristensen report their findings using the Gaussian density as kernel:

$$K(z) = \frac{1}{\sqrt{2\pi}} e^{(-\frac{z^2}{2})} \tag{51}$$

As is commonly understood in the nonparametric literature, the shape of the kernel is not so important as the choice of bandwidth. As has been mentioned, a large bandwidth includes less influential data far from time t, creating a bias. Too small a bandwidth around time t provides not enough information resulting in significant variance. A simple multivariate OLS regression is run to determine estimates.

$$[\hat{\alpha}_k(t), \hat{\beta}_k(t)']' = \left[\sum_{i=1}^{n} K_h T(t_i - t) X_i X_i'\right]^{-1} \left[\sum_{i=1}^{n} K_h T(t_i - t) X_i R_i\right] \tag{52}$$

Table (5) reports the results of the plug-in method.

A Cross-Validated Local Polynomial Regression Approach

This is directly from Jeffery Racine and Peter Hall's forthcoming work. It is the model they use to estimate, I hope to use it to calculate the bandwidth I will use in the conditional model.

$$Y_i = g(X_i) + \epsilon_i \tag{53}$$

$$S(c) = \frac{1}{nh} \sum_{i=1}^{n} \left\{ Y_i - q\left(\frac{x - X_i}{h}\right) |c \right\}^2 K\left(\frac{x - X_i}{h}\right) \tag{54}$$

$$-\frac{1}{2}\frac{\partial}{\partial c_j} S(c) = \frac{1}{nh} \sum_{i=1}^{n} \left\{ Y_i - q\left(\frac{x - X_i}{h}\right) |c \right\} \left(\frac{x - X_i}{h}\right)^j K\left(\frac{x - X_i}{h}\right) \tag{55}$$

$$= \{V(x) - \widehat{M}(x)c\}_j$$

$$V_j(x) = \frac{1}{nh} \sum_{i=1}^{n} Y_i \left(\frac{x - X_i}{h}\right)^j K\left(\frac{x - X_i}{h}\right) \tag{56}$$

$$\widehat{m}_{jk}(x) = \frac{1}{nh}\sum_{i=1}^{n}\left(\frac{x-X_i}{h}\right)^{j+k} K\left(\frac{x-X_i}{h}\right) \tag{57}$$

$$\hat{c}(x) = (\hat{c}_0(x), ..., \hat{c}_p(x))^T = \widehat{M}(x)^{-1}V(x) \tag{58}$$

$$\hat{g}(x) = \hat{c}_0(x) \tag{59}$$

$$ISE(h,p) = \int_x (\hat{g}-g)^2 fx \tag{60}$$

$$CV(h,p) = \frac{1}{n}\sum_{i=1}^{n}\left\{Y_i - \hat{g}_{-i}(X_i)\right\}^2 \tag{61}$$

$$CV(h,p) = \frac{1}{n}\sum_{i=1}^{n}\left\{g(X_i) - \hat{g}_{-i}(X_i)\right\}^2 + \frac{2}{n}\sum_{i=1}^{n}\left\{g(X_i) - \hat{g}_{-i}(X_i)\right\}\epsilon_i + \frac{1}{n}\sum_{i=1}^{n}\epsilon_i^2 \tag{62}$$

Unfortunately the application is still in beta status, and it has not been possible to generate results with this approach. The Cross-Validation Local Polynomial Regression Approach relies on a file transfer protocol with NOMAD, Nonlinear Optimization by Mesh Adaptive Direct Search (NOMAD), and there are currently issues calling this optimization software in the application. Notwithstanding this unforeseen delay in the research for this paper, I consider the work done thus far to be fruitful if only for the number of questions which arise and the potential for further research. I shall address some of these areas in the conclusion.

Conclusion

This paper has compared three approaches to multifactor modeling, the Ordinary Least Squares Regression approach, a Leave One Out Cross-Validation approach, and a Plug-in approach. The OLS method is completely non-conditioned, while both the Cross-Validation and Plug-in method are conditioned on time by minimizing an Integrated Mean Square Error to determine an optimal bandwidth specification. No doubt the Cross-Validation approach emerges as the winner, at least in terms of returning parameters with the lowest standard errors. This result is significant in and of itself, however it almost prompts more questions than it has answered. This opens the door to a host of further research opportunities.

Clearly the research in Nonparametric factor modeling thus far has been time conditioned and with good reason, the empirical evidence thus far has shown us that factor loadings are indeed time varying. In that sense, much of the work done so far has not been nonparametric in the purest sense, but rather semi-parametric or even parametric.

Firstly, what would it look like if we chose a bandwidth not as a function of time, but as a function of the data? In other words, adapt a truer nonparametric approach by letting the data completely guide us to the specification of the distribution. How different would the bandwidth be if the Integrated Mean Square Error was solely a function of the data and not time dependent. Figure (7) explores the relationship between IMSE, alphas, betas, and time. Should we question the minimum integrated mean square error methodology used to determine a *global* bandwidth?

The Capital Asset Pricing Model and the Fama French imply a linear relationship between risk and return. Perhaps the assumptions behind the CAPM need to be revisited. Perhaps estimating factors in an asset pricing model completely nonparametrically will shed some light on the linearity assumption. The CAPM also assumes a risk adverse individual in a rational expectations framework. Perhaps, uncovering a functional form through nonparametric estimation will result in an application of a model similar to CAPM with varying investor utility functions.

In any event, it would seem there are numerous opportunities for further research, and hopefully this paper has provided a platform from which some of this research may continue.

List of Tables

List of Figures

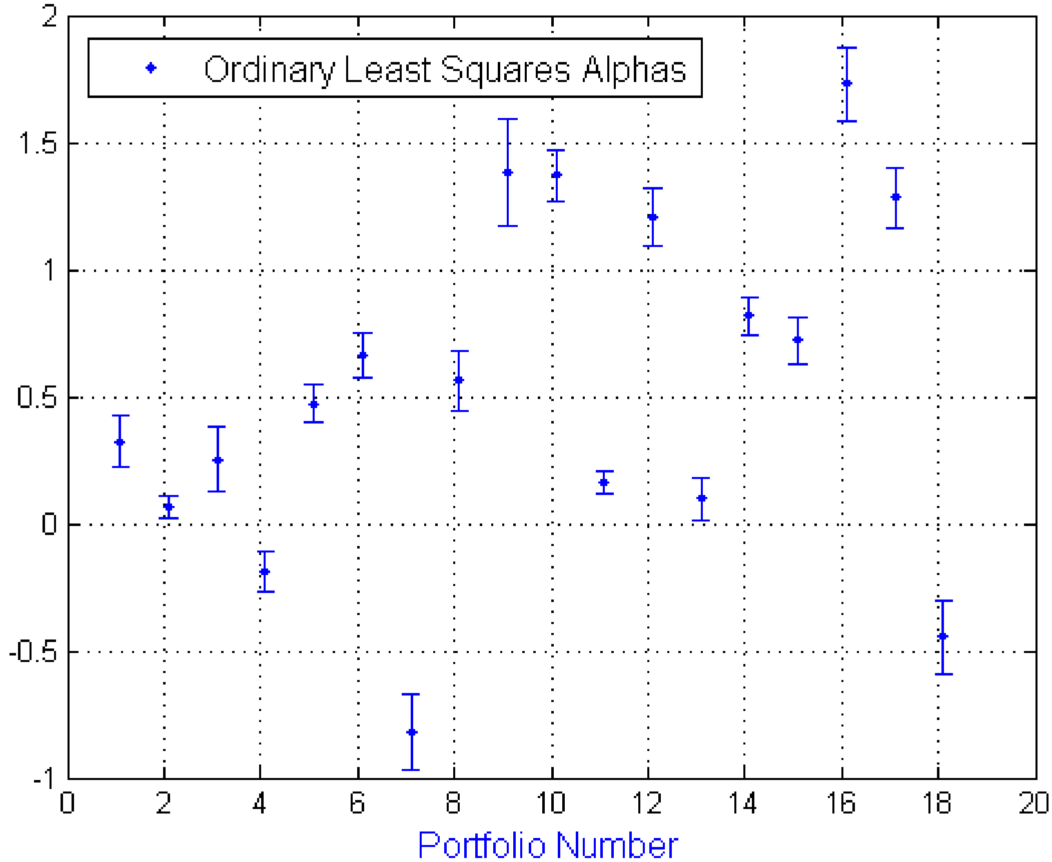

Figure 1: Ordinary Least Squares Unconditional Alphas

Table 1: **Dimson Regressions:**
Summary of Excess Returns, Unconditional Alphas and Unconditional Betas, and the associated standard errors. Daily returns and weekly returns all expressed in percent monthly.

Panel A: Value Weighted portfolios

	S	B	S-B	G	V	V-G	L	W	W-L
Ave. Ex. Rtns.									
Daily	0.657	0.511	0.146	0.315	0.843	0.528	-0.283	1.133	1.417
Weekly	0.696	0.502	0.158	0.352	0.860	0.543	-0.240	1.161	1.470
Monthly	0.765	0.501	0.196	0.380	0.891	0.560	-0.178	1.156	1.512
Std. Error									
Daily	0.163	0.177	0.131	0.208	0.159	0.104	0.254	0.249	0.212
Weekly	0.212	0.181	0.145	0.247	0.188	0.129	0.292	0.274	0.253
Monthly	0.265	0.175	0.193	0.266	0.212	0.151	0.329	0.267	0.281
Unconditional Alphas									
Daily	0.183	0.092	0.090	-0.263	0.411	0.674	-0.896	0.568	1.464
Weekly	0.139	0.089	0.007	-0.258	0.389	0.686	-0.912	0.601	1.580
Monthly	0.136	0.094	-0.034	-0.257	0.386	0.697	-0.887	0.606	1.680
Std. Error									
Daily	0.095	0.046	0.116	0.072	0.072	0.089	0.148	0.117	0.211
Weekly	0.027	0.011	0.032	0.020	0.021	0.026	0.042	0.031	0.060
Monthly	0.154	0.055	0.183	0.106	0.109	0.139	0.192	0.148	0.280
Unconditional Betas									
Daily	1.101	0.892	0.121	1.242	0.920	-0.322	1.319	1.216	-0.103
Weekly	0.929	0.907	0.262	1.207	0.869	0.347	1.216	1.257	0.240
Monthly	1.135	0.881	0.350	1.285	0.961	0.333	1.417	1.181	0.279
Std. Error									
Daily	0.023	0.011	0.028	0.017	0.017	0.021	0.035	0.028	0.050
Weekly	0.022	0.009	0.026	0.016	0.016	0.021	0.033	0.025	0.048
Monthly	0.049	0.018	0.059	0.034	0.035	0.045	0.061	0.048	0.090

Panel B: Equal Weighted Portfolios

Ave. Ex. Rtns.									
Daily	1.653	0.602	1.051	0.589	1.172	0.583	2.084	1.721	-0.363
Weekly	1.699	0.602	1.066	0.631	1.191	0.598	2.180	1.770	-0.318
Monthly	1.804	0.606	1.114	0.672	1.231	0.616	2.387	1.831	-0.246
Std. Error									
Daily	0.146	0.174	0.128	0.206	0.154	0.106	0.197	0.206	0.148
Weekly	0.205	0.187	0.148	0.250	0.187	0.131	0.281	0.252	0.197
Monthly	0.278	0.187	0.200	0.277	0.217	0.154	0.403	0.285	0.249
Unconditional Alphas									
Daily	1.209	0.163	1.046	0.005	0.745	0.741	1.537	1.170	-0.367
Weekly	1.150	0.162	0.953	0.001	0.718	0.757	1.479	1.181	-0.213
Monthly	1.151	0.171	0.892	0.009	0.716	0.770	1.547	1.194	-0.073
Std. Error									
Daily	0.092	0.047	0.104	0.079	0.073	0.090	0.138	0.155	0.144
Weekly	0.029	0.012	0.032	0.022	0.021	0.027	0.046	0.032	0.045
Monthly	0.178	0.062	0.191	0.118	0.115	0.140	0.297	0.171	0.247
Unconditional Betas									
Daily	0.945	0.935	0.009	1.255	0.906	-0.349	1.169	1.181	0.012
Weekly	0.852	0.933	0.312	1.202	0.850	0.359	1.055	1.106	0.274
Monthly	1.131	0.936	0.348	1.323	0.970	0.366	1.450	1.212	0.268
Std. Error									
Daily	0.022	0.011	0.022	0.019	0.017	0.021	0.033	0.027	0.034
Weekly	0.023	0.010	0.026	0.018	0.017	0.021	0.037	0.026	0.036
Monthly	0.057	0.020	0.061	0.038	0.037	0.045	0.095	0.055	0.079

Table 2: **Ordinary Least Squares Regressions:**
Summary of Excess Returns, Unconditional Alphas and Unconditional Betas, and the associated standard errors. Daily returns and weekly returns all expressed in percent monthly.

Panel A: Value Weighted portfolios

	S	B	S-B	G	V	V-G	L	W	W-L
Ave. Ex. Rtns.									
Daily	0.657	0.511	0.146	0.315	0.843	0.528	-0.283	1.133	1.417
Weekly	0.696	0.502	0.158	0.352	0.860	0.543	0.-240	1.161	1.470
Monthly	0.765	0.501	0.196	0.380	0.891	0.560	-0.178	1.156	1.512
Std. Error									
Daily	0.163	0.177	0.131	0.208	0.159	0.104	0.254	0.249	0.212
Weekly	0.212	0.181	0.145	0.247	0.188	0.129	0.292	0.274	0.253
Monthly	0.265	0.175	0.193	0.266	0.212	0.151	0.329	0.267	0.281
Unconditional Alphas									
Daily	0.321	0.066	0.254	-0.188	0.474	0.662	-0.817	0.565	1.38
Weekly	0.273	0.078	0.154	-0.212	0.455	0.705	-0.804	0.573	1.445
Monthly	0.239	0.088	0.075	-0.222	0.439	0.714	-0.841	0.603	1.631
Std. Error									
Daily	0.198	0.091	0.243	0.147	0.144	0.174	0.291	0.230	0.415
Weekly	0.122	0.047	0.145	0.085	0.089	0.111	0.178	0.130	0.254
Monthly	0.159	0.054	0.187	0.106	0.110	0.138	0.191	0.147	0.278
Unconditional Betas									
Daily	0.711	0.948	-0.238	1.078	0.782	-0.296	1.144	1.222	0.078
Weekly	0.904	0.906	0.007	1.205	0.864	-0.346	1.205	1.256	0.054
Monthly	1.124	0.880	0.261	1.287	0.962	-0.332	1.418	1.180	-0.256
Std. Error									
Daily	0.010	0.005	0.013	0.008	0.008	0.009	0.015	0.012	0.022
Weekly	0.013	0.005	0.016	0.009	0.010	0.012	0.020	0.014	0.028
Monthly	0.036	0.012	0.043	0.024	0.025	0.031	0.044	0.033	0.063

Panel B: Equal Weighted Portfolios

Ave. Ex. Rtns.									
Daily	1.653	0.602	1.051	0.589	1.172	0.583	2.084	1.721	-0.363
Weekly	1.699	0.602	1.066	0.631	1.191	0.598	2.180	1.770	-0.318
Monthly	1.804	0.606	1.114	0.672	1.231	0.616	2.387	1.831	-0.246
Std. Error									
Daily	0.146	0.174	0.128	0.206	0.154	0.106	0.197	0.206	0.148
Weekly	0.205	0.187	0.148	0.250	0.187	0.131	0.281	0.252	0.197
Monthly	0.278	0.187	0.200	0.277	0.217	0.154	0.403	0.285	0.249
Unconditional Alphas									
Daily	1.372	0.164	1.207	0.097	0.818	0.720	1.730	1.283	-0.447
Weekly	1.320	0.165	1.122	0.070	0.797	0.766	1.717	1.256	-0.373
Monthly	1.285	0.167	1.030	0.053	0.776	0.785	1.726	1.264	-0.172
Std. Error									
Daily	0.197	0.091	0.219	0.162	0.147	0.178	0.284	0.234	0.284
Weekly	0.134	0.052	0.147	0.097	0.093	0.111	0.207	0.137	0.196
Monthly	0.185	0.061	0.198	0.119	0.116	0.138	0.304	0.171	0.248
Unconditional Betas									
Daily	0.588	0.934	-0.345	1.051	0.747	-0.304	0.746	0.93263	0.186
Weekly	0.809	0.933	-0.119	1.197	0.842	-0.348	0.989	1.099	0.118
Monthly	1.107	0.935	0.180	1.324	0.969	-0.365	1.415	1.210	-0.159
Std. Error									
Daily	0.010	0.005	0.012	0.009	0.008	0.009	0.015	0.012	0.015
Weekly	0.015	0.006	0.016	0.011	0.010	0.012	0.023	0.015	0.022
Monthly	0.042	0.014	0.045	0.027	0.026	0.032	0.069	0.039	0.056

Table 3: **Ordinary Least Squares Unconditional Results:**
Summary Statistics for Size, B/M, and Momentum Portfolios. This table reports average excess returns, unconditional CAPM alphas, unconditional FF alphas, and the associated standard errors. All expressed in percent monthly. The Unconditional CAPM Alphas are obtained by letting $f = R_m$, and the Unconditional FF Alphas are obtained by letting $f = (R_m, SMB, HML)$

Panel A: Value Weighted portfolios

	S	B	S-B	G	V	V-G	L	W	W-L
Ave. Ex. Rtns.	0.657	0.511	0.146	0.315	0.843	0.528	-0.283	1.133	1.417
Std. Error	0.163	0.177	0.131	0.208	0.159	0.104	0.254	0.249	0.212
CAPM									
Alpha	0.321	0.066	0.254	-0.188	0.474	0.662	-0.817	0.565	1.38
Std.Err.	0.198	0.091	0.243	0.147	0.144	0.174	0.291	0.230	0.415
Beta	0.711	0.948	-0.238	1.078	0.782	-0.296	1.144	1.222	0.078
Std. Err.	0.010	0.005	0.013	0.008	0.008	0.009	0.015	0.012	0.022
FF									
Alpha	-0.034	-0.074	0.040	-0.154	-0.046	0.108	-0.835	0.564	1.399
Std.Err.	0.090	0.054	0.085	0.068	0.067	0.070	0.287	0.217	0.426
MKT Beta	0.921	1.016	-0.095	1.083	1.075	-0.008	1.162	1.234	0.072
Std. Err.	0.006	0.004	0.006	0.005	0.005	0.005	0.020	0.015	0.029
SMB Beta	0.976	-0.239	1.216	0.621	0.498	-0.122	0.411	0.437	0.026
Std. Err.	0.009	0.005	0.008	0.007	0.007	0.007	0.029	0.022	0.043
HML Beta	0.317	0.342	-0.024	-0.265	0.797	1.062	-0.114	-0.144	-0.029
Std. Err.	0.012	0.007	0.011	0.009	0.009	0.009	0.038	0.029	0.057

Panel B: Equal Weighted Portfolios

	S	B	S-B	G	V	V-G	L	W	W-L
Ave. Ex. Rtns.	1.653	0.602	1.051	0.589	1.172	0.583	2.084	1.721	-0.363
Std. Error	0.146	0.174	0.128	0.206	0.154	0.106	0.197	0.206	0.148
CAPM									
Alpha	1.372	0.164	1.207	0.097	0.818	0.720	1.730	1.283	-0.447
Std.Err.	0.197	0.091	0.219	0.162	0.147	0.178	0.284	0.234	0.284
Beta	0.588	0.934	-0.345	1.051	0.747	-0.304	0.746	0.93263	0.186
Std. Err.	0.010	0.005	0.012	0.009	0.008	0.009	0.015	0.012	0.015
FF									
Alpha	1.043	-0.018	1.061	0.109	0.316	0.207	1.462	1.033	-0.429
Std.Err.	0.123	0.073	0.131	0.095	0.079	0.093	0.233	0.152	0.291
MKT Beta	0.792	1.028	-0.236	1.070	1.030	-0.038	0.920	1.094	0.174
Std. Err.	0.008	0.005	0.009	0.006	0.005	0.006	0.002	0.010	0.020
SMB Beta	0.865	-0.082	0.948	0.650	0.500	-0.149	0.940	0.991	0.052
Std. Err.	0.012	0.007	0.013	0.009	0.008	0.009	0.024	0.015	0.023
HML Beta	0.323	0.364	-0.041	-0.232	0.762	0.994	0.182	0.131	-0.050
Std. Err.	0.016	0.009	0.017	0.013	0.015	0.012	0.030	0.020	0.038

Table 4: **Cross-Validation Leave One Out Results:**
Summary Statistics for Size, B/M, and Momentum Portfolios. This table reports average excess returns, unconditional CAPM alphas, unconditional FF alphas, and the associated standard errors. All expressed in percent monthly. The Unconditional CAPM Alphas are obtained by letting $f = R_m$, and the Unconditional FF Alphas are obtained by letting $f = (R_m, SMB, HML)$

Panel A: Value Weighted portfolios

	S	B	S-B	G	V	V-G	L	W	W-L
FF									
MKT Beta	0.927	1.011	-0.084	1.091	1.064	-0.028	1.164	1.241	.0739
Std. Err.	0.023	0.006	0.0024	0.001	0.002	0.001	0.005	0.005	0.010
SMB Beta	0.957	-0.221	1.178	0.625	0.500	-0.125	0.411	0.382	-0.029
Std. Err.	0.003	0.001	0.003	0.001	0.001	0.001	0.009	0.007	0.014
HML Beta	0.239	0.252	-0.021	-0.283	0.689	0.971	-0.039	-0.130	-0.096
Std. Err.	0.003	0.002	0.002	0.002	0.003	0.002	0.011	0.011	0.019

Panel B: Equal Weighted Portfolios

	S	B	S-B	G	V	V-G	L	W	W-L
FF									
MKT Beta	0.762	1.018	-0.257	1.072	1.013	-0.059	0.883	1.098	0.211
Std. Err.	0.004	0.001	0.005	0.002	0.002	0.002	0.006	0.004	0.009
SMB Beta	0.843	-0.018	0.853	0.687	0.517	-0.169	0.915	0.939	0.021
Std. Err.	0.005	0.002	0.005	0.002	0.002	0.002	0.008	0.006	0.011
HML Beta	0.232	0.269	-0.040	-0.194	0.646	0.854	0.216	0144	-0.068
Std. Err.	0.004	0.003	0.005	0.004	0.003	0.003	0.009	0.006	0.012

Table 5: **Plug-in Method Results:**
Summary Statistics for Size, B/M, and Momentum Portfolios. This table reports average excess returns, conditional Alphas and conditional Betas, and the associated standard errors. The conditional Alphas are obtained by letting $f = R_m$, and the conditional Betas are obtained by letting $f = (R_m, SMB, HML)$

Panel A: Value Weighted portfolios

	S	B	S-B	G	V	V-G	L	W	W-L
FF									
Alpha	-0.001	-0.007	-0.049	-0.016	-0.001	-0.041	-0.056	0.036	0.043
Std.Err.	0.034	0.021	0.033	0.026	0.026	0.027	0.110	0.083	0.164
MKT Beta	0.920	1.015	-0.094	1.083	1.075	-0.007	1.162	1.234	0.072
Std. Err.	0.003	0.001	0.003	0.002	0.002	0.002	0.010	0.007	0.014
SMB Beta	0.976	-0.239	1.216	0.620	0.498	-0.121	0.411	0.436	0.026
Std. Err.	0.004	0.002	0.004	0.003	0.003	0.003	0.014	0.011	0.022
HML Beta	.0317	0.341	-0.024	-0.264	0.797	1.061	-0.114	-0.144	-0.029
Std. Err.	0.006	0.003	0.005	0.004	0.004	0.004	0.019	0.014	0.028

Panel B: Equal Weighted Portfolios

	S	B	S-B	G	V	V-G	L	W	W-L
FF									
Alpha	0.1420	0.0059	0.0826	0.0265	0.0509	-0.030	0.1837	0.1217	-0.118
Std.Err.	0.0475	0.0281	0.0506	0.0367	0.0304	0.0360	0.0899	0.0587	0.1123
MKT Beta	0.7919	1.0284	-0.236	1.0696	1.0302	-0.038	0.9199	1.0940	0.1745
Std. Err.	0.0043	0.0026	0.0046	0.0033	0.0028	0.0033	0.0082	0.0053	0.0102
SMB Beta	0.8647	-0.082	0.9479	0.6503	0.5006	-0.149	0.9402	0.9912	0.0518
Std. Err.	0.0064	0.0038	0.0068	0.0049	0.0041	0.0049	0.0121	0.0079	0.0151
HML Beta	0.3227	0.3643	-0.041	-0.232	0.7619	0.9942	0.1821	0.1313	-0.050
Std. Err.	0.0083	0.0049	0.0089	0.0064	0.0053	0.0063	0.0158	0.0103	0.0197

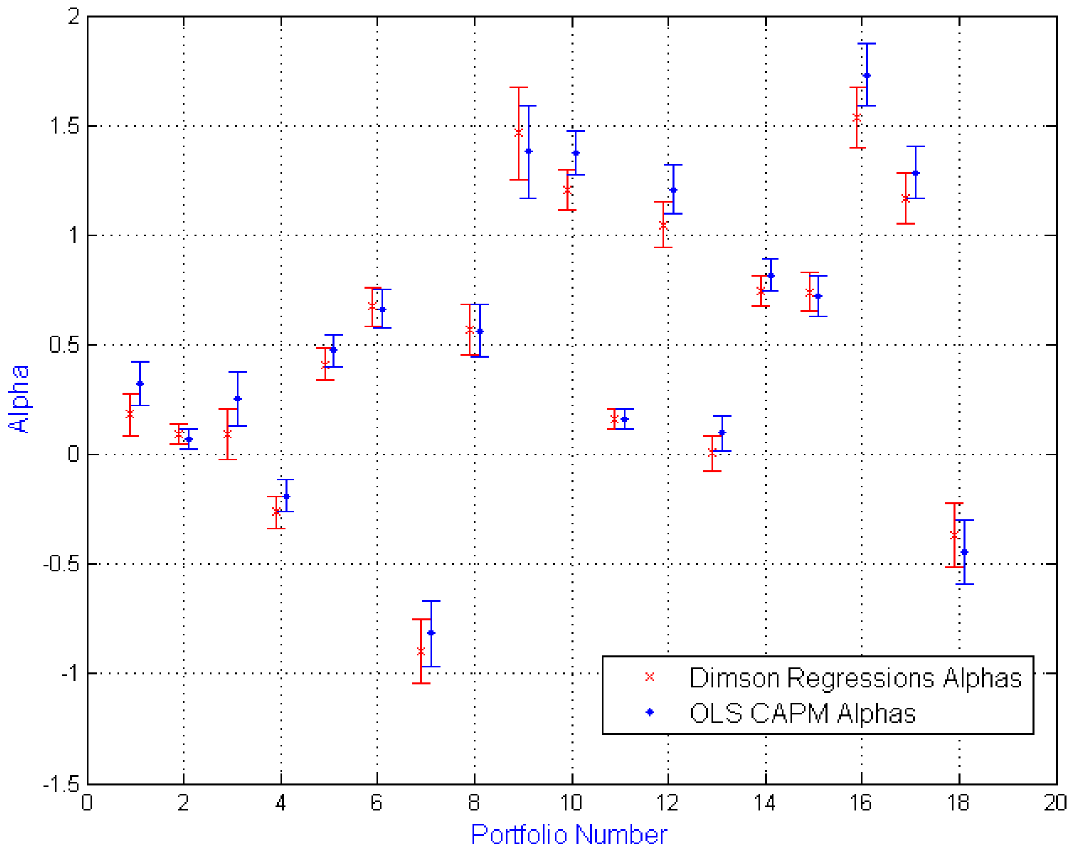

Figure 2: Comparison of Dimson and OLS Alphas

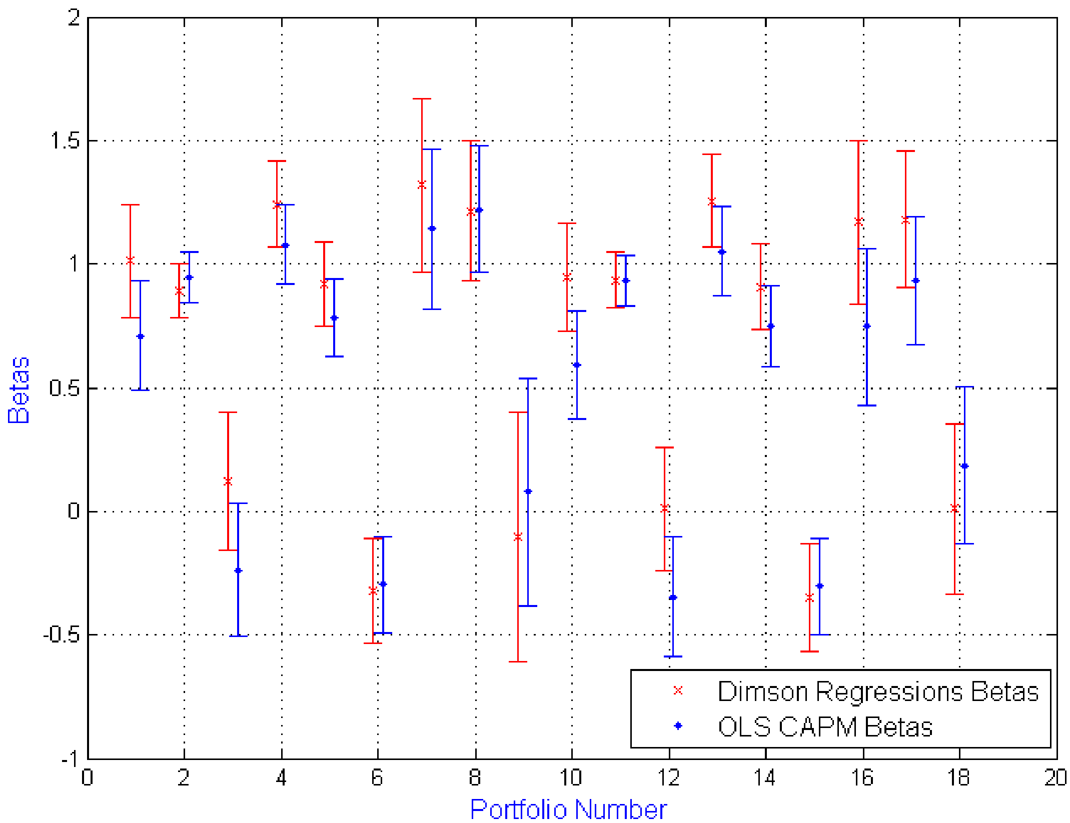

Figure 3: Comparison of Dimson and OLS Betas

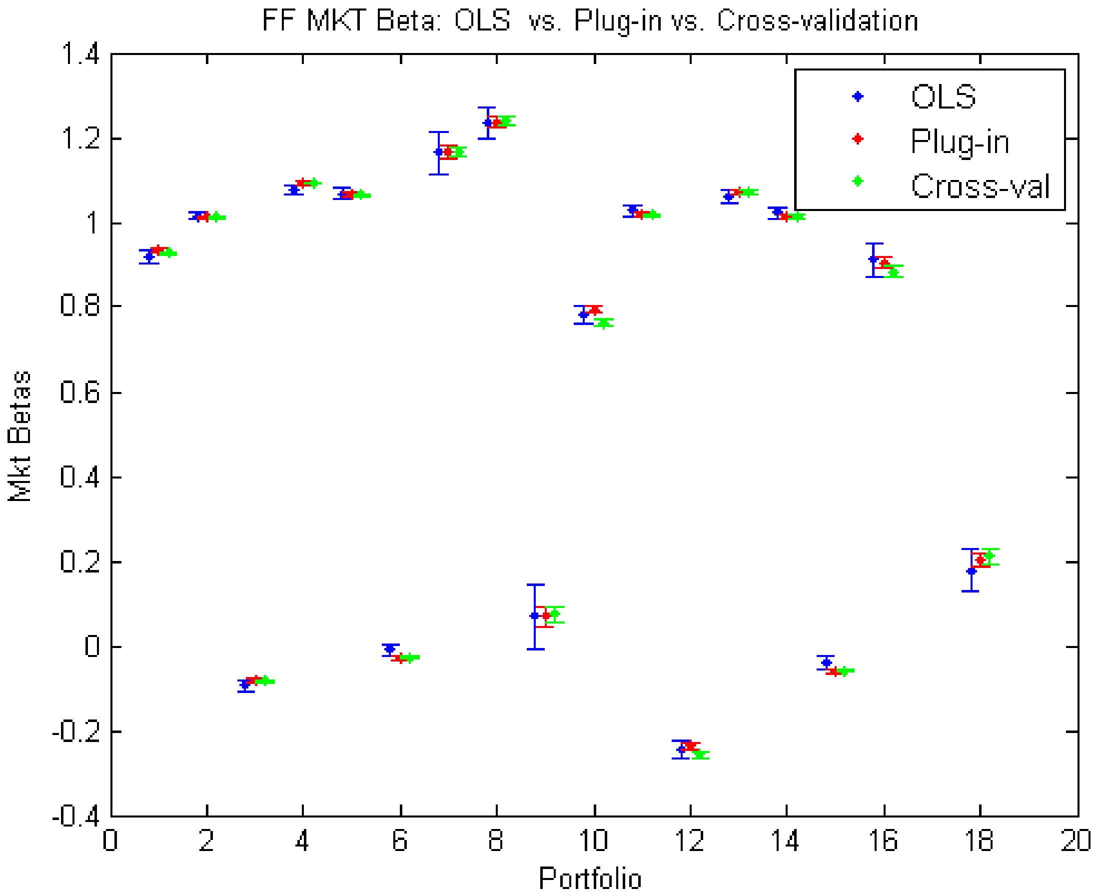

Figure 4: Comparison FF MKT Factor

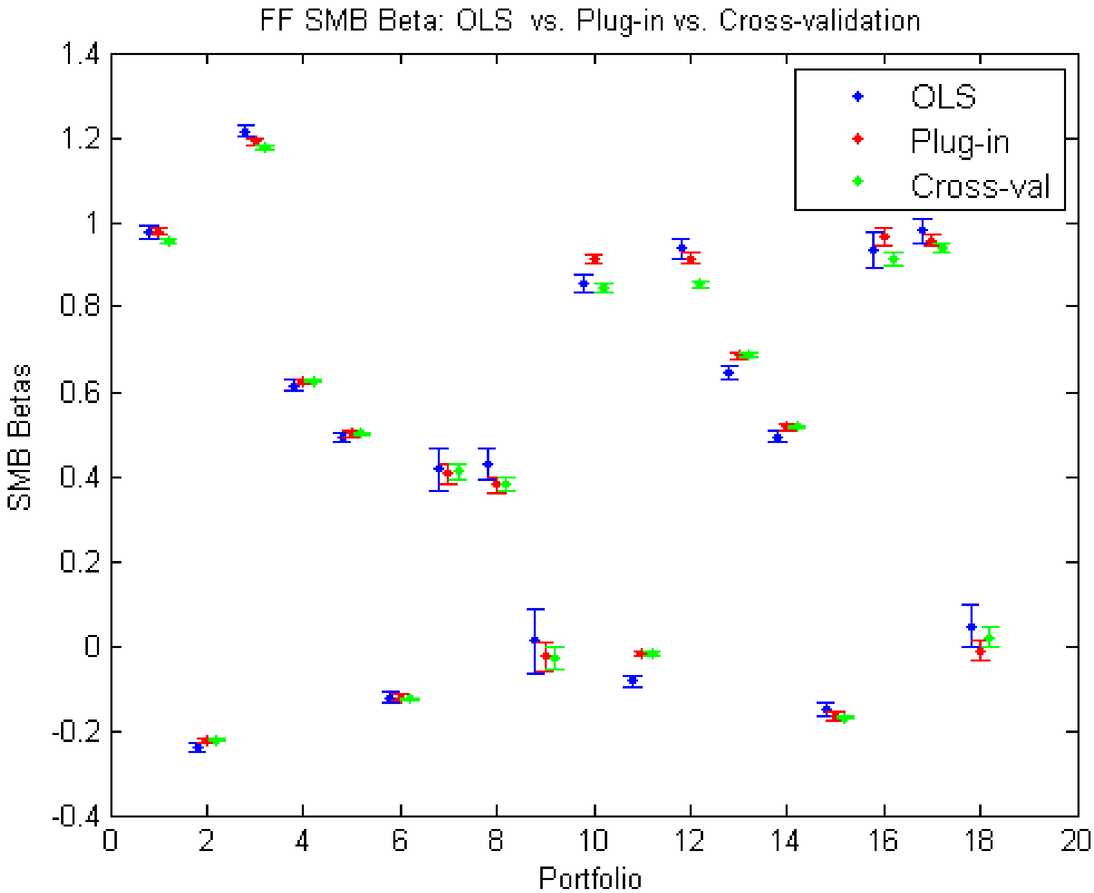

Figure 5: Comparison FF SMB Factor

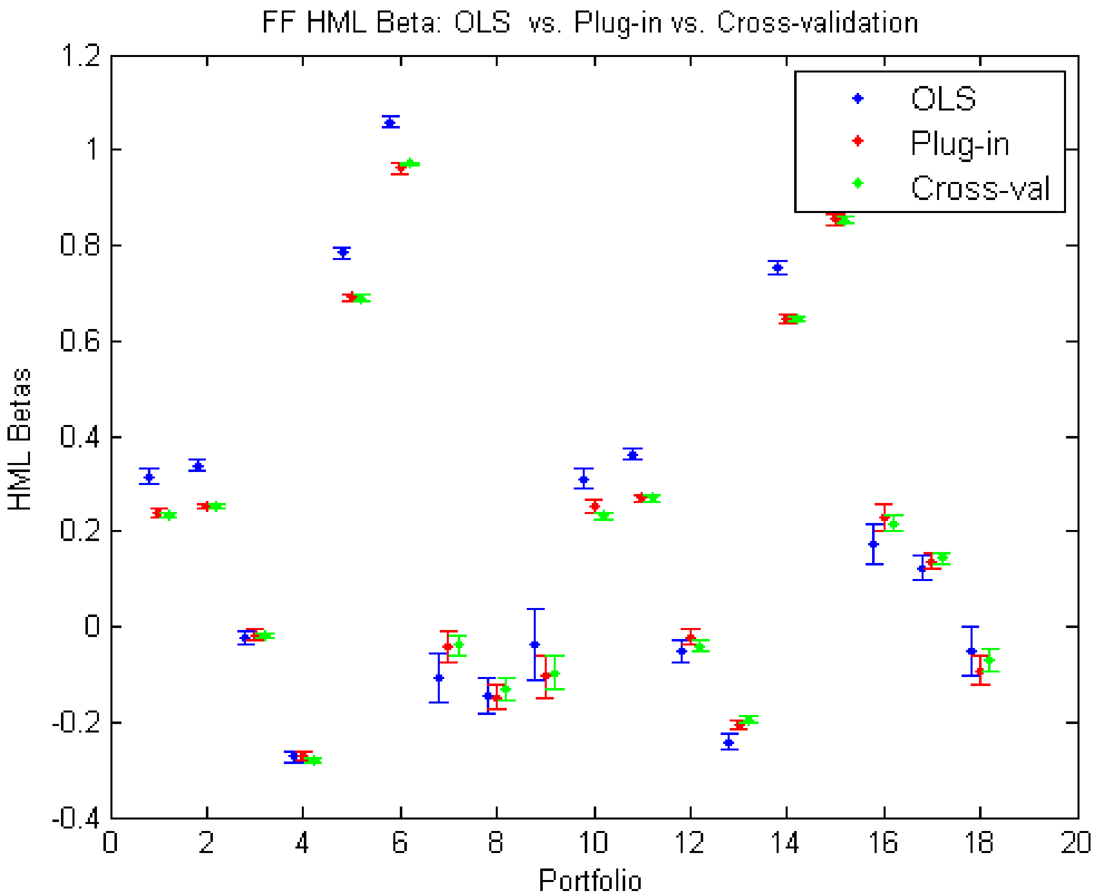

Figure 6: Comparison FF HML Factor

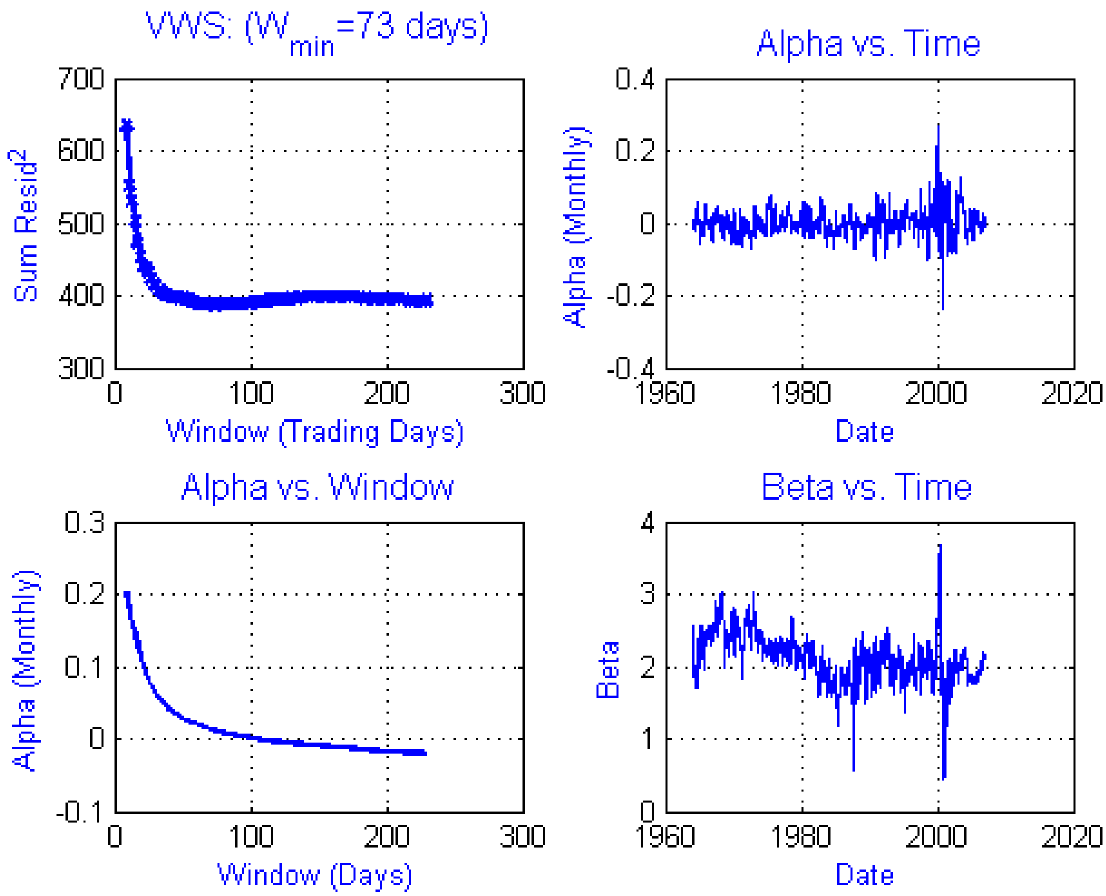

Figure 7: Alphas, Betas, IMSE as a function of time

References

Abramson M, Audet C, Couture G, Dennis, Jr J, Le Digabel S, Tribes C. ???? The NOMAD project. Software available at `http://www.gerad.ca/nomad`, URL `http://www.gerad.ca/nomad`.

Ang A, Kristensen D. 2012. Testing conditional factor models. *Journal of Financial Economics*, **106**: 132–156.

Bodie M Kane. 2009. *Investments*. McGraw Hill, N.Y., USA.

Cai CM Zongwu; Kuan. 2005. Nonparametric Modeling for Conditional Capital Asset Pricing Models. *Working paper*.

Connor MLO Gregory; Hagmann. March 2012, v. 80, iss. 2, pp. 713-54. Efficient Semiparametric Estimation of the Fama-French Model and Extensions. *Econometrica*.

Dimson E. 1979. Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics*, **7**: 197–226.

Dybvig PH, Ross SA. 1985. Differential information and performance measurement using a security market line. *The Journal of Finance*, **40**: 383–399.

Fama EF, French KR. 1992. The Cross-Section of Expected Stock Returns. *The Journal of Finance 47.2*.

Fama KR Eugene F; French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*.

Fama MJD Eugene F. 1973. Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*.

Ferreira JOS Eva; Gil-Bazo. 2011. Conditional beta pricing models: A nonparametric approach. *Working paper*.

Gibbons MR, Ross SA, Shanken J. 1989. A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society*, 1121–1152.

Jagannathan Z Ravi; Wang. 1996. The Conditional CAPM and the Cross-Section of Expected returns. *The Journal of Finance*.

Jensen MC. 1968. The performance of mutual funds in the period 1945–1964. *The Journal of Finance*, **23**: 389–416.

Lewellen S Jonathan; Nagel. 2006. The Conditional CAPM does not Explain Asset-Pricing Anamolies. *National Bureau of Economic Research, Inc, NBER Working Papers: 9974*.

Li L Yan; Yang. 2011. Testing Conditional factor models: A nonparametric approach. *Journal of Empirical Finance*.

Li Q, Racine JS. 2007. *Nonparametric econometrics: Theory and practice*. Princeton University Press.

Lintner J. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The review of economics and statistics*, **47**: 13–37.

Lo AW, Craig MacKinlay A. 1990. An econometric analysis of nonsynchronous trading. *Journal of Econometrics*, **45**: 181–211.

Loader CR. 1999. Bandwidth selection: Classical or plug-in? *Annals of Statistics*, 415–438.

Mossin J. 1966. Equilibrium in a capital asset market. *Econometrica: Journal of the Econometric Society*, 768–783.

Motta CMvSR Giovanni; Hafner. 2011. Locally Stationary Factor Models: Identification and Nonparametric Estimation. *Econometric Theory*.

Pagan A, Ullah A. 1999. *Nonparametric econometrics*. Cambridge University Press.

Parzen E. 1962. On estimation of a probability density function and mode. *The annals of mathematical statistics*, **33**: 1065–1076.

Rosenblatt M. 1956. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, 832–837.

Ross SA. 1973. Return, risk and arbitrage. Tech. rep., Wharton School Rodney L. White Center for Financial Research.

Sharpe WF. 1964. CAPITAL ASSET PRICES: A THEORY OF MARKET EQUILIBRIUM UNDER CONDITIONS OF RISK*. *The journal of finance*, **19**: 425–442.

Silverman BW. 1986. *Density estimation for statistics and data analysis*, vol. 26. CRC press.

Vinod HD. 2008. *Hands-on Intermediate Econometrics Using R: Templates forExtending Dozens of Practical Examples*. World Scientific, Hackensack, NJ, URL `http://www.worldscibooks.com/economics/6895.html`, iSBN 10-981-281-885-5.