Journal of Empirical Finance 8 (2001) 573–637

Journal of EMPIRICAL FINANCE

www.elsevier.com/locate/econbase

The specification of conditional expectations ☆

Campbell R. Harvey *

Fuqua School of Business, Duke University, Durham, NC 27708, USA

Abstract

This paper explores different specifications of conditional expectations. The most common specification, linear least squares, is contrasted with nonparametric techniques that make no assumptions about the distribution of the data. Nonparametric regression is successful in capturing some nonlinearities in financial data, in particular, asymmetric responses of security returns to the direction and magnitude of market returns. The technique is ideally suited for empirically modeling returns of securities that have complicated embedded options. The conditional mean and variance of the NYSE market return are also examined. Forecasts of market returns are not improved with the nonparametric techniques which suggests that linear conditional expectations are a reasonable approximation in conditional asset pricing research. However, the linear model produces a disturbing number of negative expected excess returns. My results also indicate that the relation between the conditional mean and variance depends on the specification of the conditional variance. Furthermore, a linear model relating mean to variance is rejected and these tests are not sensitive to the expectation generating mechanism nor the conditioning information. Rejections are driven by the distinct countercyclical variation in the ratio of the conditional mean to variance. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Conditional expectation; Linear least square; Nonparametric regression

Conditional expectations play a central role in finance. These expectations are usually represented by a linear regression. For example, in many asset pricing

☆ This paper was written during 1988–1991. The text follows the final working pages version: August 1991. Part of this paper was written while the author was visiting the Graduate School of Business at the University of Chicago. This paper incorporates some results previously circulated in "Is the Expected Compensation for Market Volatility Constant Through Time?" and "On the Linearity of Conditionally Expected Returns".

* Tel.: +1-919-660-7768; fax: +1-919-660-7971.

E-mail address: cam.harvey@duke.edu (C.R. Harvey).

PII: S0927-5398(01)00036-6

tests, it is commonplace to obtain risk sensitivities through a linear regression model. These sensitivities are used to test the restrictions implied by the model.[1] These risk measures are also critical to the study of firm-specific events. Recent research has focused on the predictability of asset returns. In most applications, the conditionally expected asset returns are represented by a linear regression.[2] In other situations, conditional variances, covariances and betas have been represented by a linear regression.[3] While economic theory tells us how to link conditional expectations with conditional risk and reward, it does not tell us how the conditional expectations are generated.

In order to implement and test asset pricing models, numerous assumptions are made. The assumption of multivariate normality is often invoked. This assumption is sufficient for the linear regression of asset returns on the market return to be well specified. A similar assumption is sufficient to justify a linear model generating forecasted asset returns from a set of predetermined instrumental variables. Of course, these specifications are ad hoc. Indeed, inference is complicated by the existence of these ancillary assumptions. That is, we may incorrectly reject a model's restrictions because we have incorrectly generated the conditional expectations.

This paper compares conditional expectations produced from both linear and nonparametric regression models. The technique of nonparametric density estimation allows us to extract conditional expectations without assuming that the data fall into a particular class of parametric distributions. Both conditional means and volatilities are examined.

Four applications are considered. In the first, simulated portfolio returns are created using two nonlinear models. These models are designed to reflect the portfolio strategies of a market timer. The results show that the nonparametric regression is able to closely replicate the nonlinearities in the data.

The second application focuses on bivariate market model regressions. There are reasons to believe that the stock price response to a market move is not symmetric over the direction of the market move. For example, if the market moves down, the debt–equity ratio of the firm could increase, thereby increasing the equity beta. The returns of a particular industry are examined. The nonparametric regression shows some promise in detecting asymmetries in the returns generating process.

---

[1] See for example, Black et al. (1972), Fama and MacBeth (1973), Gibbons (1982), Stambaugh (1982), Shanken (1985), MacKinlay (1987), and Gibbons et al. (1989) for the single factor CAPM, Breeden et al. (1989) for the consumption CAPM, and Chan et al. (1985), Chen et al. (1986), Shanken and Weinstein (1990) and Ferson and Harvey (1991) for the multiple factor CAPM.

[2] See, for example, Gibbons and Ferson (1985), Keim and Stambaugh (1986), Campbell (1987), Fama and French (1988, 1989), Hodrick (1992) and Harvey (1989, 1991).

[3] Linear variance estimators are used in Hasbrouck (1986) and Campbell (1987). Shanken (1990) assumes that the beta is linear in a set of instruments.

The third application focuses on mimicking portfolios. To mention only one example, hedge ratios are often calculated from a linear regression of the cash price changes on the futures price changes. This method could lead to incorrect hedge ratios if there are nonlinearities in the data. Often these nonlinearities arise because of embedded options in the cash or the futures instrument that may be difficult to price. Nonparametric regression may be helpful in fitting some of these nonlinearities. In an example, the out-of-sample hedging performance of the nonparametric regression and two alternative methods are compared for some mortgage-backed securities.

The final application centers on the conditionally expected returns and volatilities of NYSE value-weighted returns. Many studies assume that conditional expectations are linear in a set of information variables. New evidence is presented that suggests that forecasting performance is not improved with nonparametric models. However, the linear model produces a disturbing number of negative conditionally expected returns. This could be a result of overfitting. The nonparametric fitted values appear more reasonable in this respect. In addition, the same business cycle patterns documented by Fama and French (1989) and Ferson and Harvey (1991) are evident in this alternate method for obtaining conditional expectations.

The relation between the conditional mean return and the conditional variance is also examined. Some find a positive relation between the conditional mean and the conditional variance while others find a negative relation. Results are presented using nine different conditional variance specifications (five nonparametric and four parametric). The relation turns out to be influenced by how information enters the conditional variance estimator. However, a linear model relating conditional mean to conditional variance is strongly rejected. Furthermore, the rejections are caused by distinct countercyclical variation in the ratio of conditional mean to volatility.

This paper is organized as follows. In the first section, the method of nonparametric density estimation is reviewed. The second section presents the empirical applications. Some concluding remarks are offered in the final section.

1. Linear and nonparametric expectations

1.1. Sufficient assumptions for linear conditional expectations

It is often assumed that expectations are linear in a set of information variables. For example, risk is often measured by a linear regression of stock returns on a risk factor. Others have forecasted asset returns by assuming that asset returns are linear in a set of predetermined information variables. What are the sufficient assumptions that must be invoked for a linear model of conditional expectations to be properly specified?

If the data fall into the class of *spherically invariant distributions*, then expectations will be linear in the conditioning information.[4] This class was first investigated by Vershik (1964). Consider a set of random variables, $\{x_1, \dots, x_n\}$, with finite second moments. Let **H** denote a linear manifold spanned by the set of random variables: $\sum_{j=1}^{n} \alpha_j x_j$, where the α_j's are real numbers. If all random variables in the linear manifold **H** which have the same variance have the same distribution, then following Vershik (1964), we call (i) **H** a *spherically invariant space*; (ii) $\{x_1, \dots, x_n\}$ *spherically invariant* and (iii) every distribution function of any variable in **H**, a *spherically invariant distribution*.

For example, normal distributions are spherically invariant. Suppose we have two normal distributions. By subtracting the mean and dividing by the standard deviation, the two distributions will be identical. We know for the normal distribution that the linear manifold generated by any set of normal random variables will be a spherically invariant space.

A disadvantage of Vershik's (1964) definition is that it does not encompass processes like the Cauchy for which the variance is not defined. Blake and Thomas (1968) and Chu (1973) proposed a definition of an elliptical class of distributions that addresses this shortcoming.[5] A random vector $\boldsymbol{x}$ has an elliptical distribution if and only if its probability density function $p(\boldsymbol{x})$ can be expressed as a function of a quadratic form, $p(\boldsymbol{x}) = f(1/2\,\boldsymbol{x}'\mathbf{c}^{-1}\boldsymbol{x})$ where **c** is assumed positive definite. When the variance–covariance matrix exists, **c** is proportional to it and the Vershik (1964), Blake and Thomas (1968) and Chu (1973) definitions are equivalent.[6] However, the quadratic form of the density also applies to processes like the Cauchy whose conditional expectation is still linear with the projection constants depending on the characteristic matrix.

Spherical distributions (not spherically invariant distributions) are a subclass of the elliptical distributions. If the density function is a function of a quadratic form and if **c** is a diagonal positive definite characteristic matrix, then the distribution is said to be spherical. If the **c** is positive definite but not diagonal, then the distribution is elliptical.

All of these distributions have the common feature of symmetry. If the data do not fall into the elliptical class, then a linear model for expectations may not

[4] These are sufficient conditions for linear conditional expectations. Of course, it is possible that investors choose to use a linear filtration on data that are not spherically invariant.

[5] Elliptical distributions have been used in asset pricing theory by Chamberlain (1983), Ingersoll (1987) and Foster and Viswanathan (1993).

[6] Implicit in Chu's (1973) definition is the existence of the density function. Kelker (1970) provides an alternative approach in terms of the characteristic function. Also see Devlin et al. (1976).

generate the true expectations. There are reasons to believe that some of the data in finance do not fit into this class. In particular, many asset returns exhibit skewness. This feature of the data is particularly important for securities that have embedded options.

Models that generate linear conditional expectations are parametric in the sense that some distributional characteristics are either explicitly or implicitly assumed. However, it is possible to obtain conditional expectations without any parametric assumptions. This is the idea of nonparametric density estimation.

1.2. Nonparametric density estimation

With observations $\{y_i,\ x_i\}_{i=1}^{T}$, the regression curve of y on x is

$$E[y|x] = \frac{\int y f(x,y)\,\mathrm{d}y}{\int f(x,y)\,\mathrm{d}y}, \tag{1}$$

where $f(x,y)$ is the joint density of x and y which is assumed to exist. Panel A of Fig. 1 provides an example of a bivariate normal distribution of $\{x,y\}$. In panel B, a particular realization of x is considered. The expectation of y conditioned upon the realization can be obtained by taking the probability weighted average of all y. This is the univariate distribution which is shaded ($\int f(x,y)\mathrm{d}y$). However, this area does not integrate to one. But by dividing by the shaded area (the denominator in Eq. (1)), the probabilities will sum to one. If we assume that these two variables are drawn from a bivariate normal distribution, or more generally a spherically invariant distribution, then the conditional expectation of Y can be written as a linear (regression) function of X.

What if the data are not normal nor even symmetric? The same logic of Eq. (1) can be applied. Panel A of Fig. 2 shows a bivariate distribution that is bimodal and asymmetric. The expectation of y conditioned upon a realization of x can be obtained the same way. In panel B, the probability weighted average of y is obtained from a univariate slice of the joint distribution at x. Probabilities sum to one because we divide by the shaded area. It is obvious from Fig. 2 that y is a nonlinear function of x. A linear function to approximate the conditional expectations is unlikely to generate values close to the true expectations.

The idea of the nonparametric estimation technique of Watson (1964) and Nadaraya (1964) is to go to the data and directly estimate the function $f(x,y)$. An approximation could be obtained by drawing a curve through the midpoints of a histogram. While this may be fairly straightforward with a small amount of data and only one random variable, it is difficult to graphically represent the density beyond the bivariate case. One can think of nonparametric density estimation as a

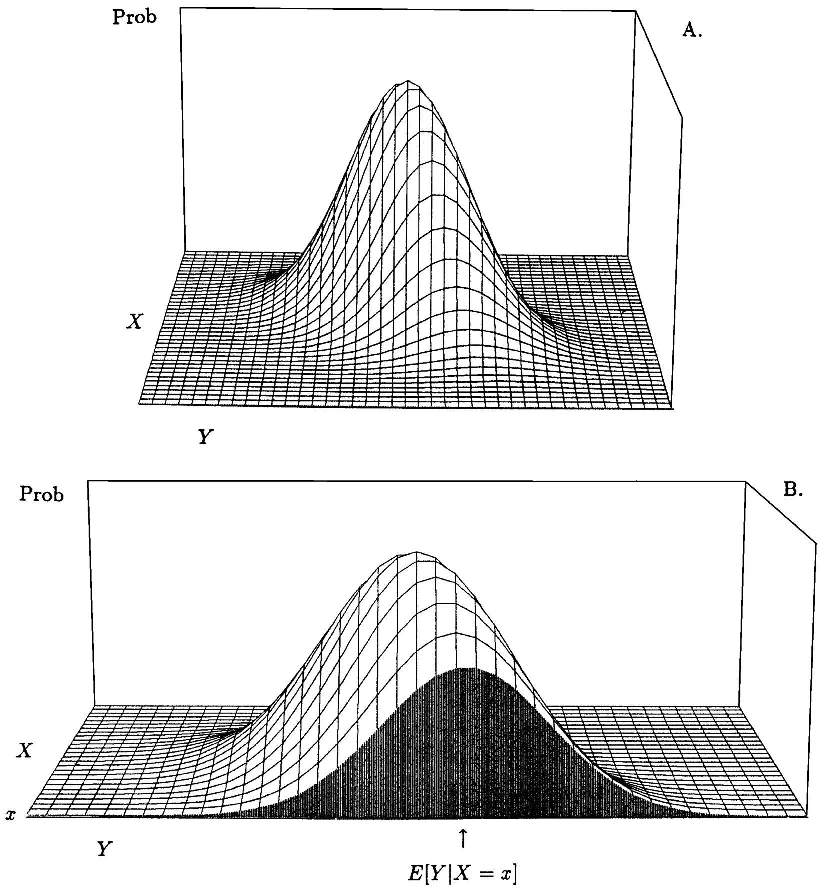


Fig. 1. An illustration of conditional expectations with a bivariate normal distribution. Panel A plots a bivariate normal density of X and Y. In panel B, the conditional expectation $E[Y]|X = x]$ is illustrated.

mechanical way to approximate the density function. Indeed, Watson viewed his paper as contributing "a simple computer method for obtaining a graph from a large number of observations". The empirical distributions are inserted into Eq. (1) to obtain the nonparametric regression estimates.

The following are the mechanics of the density estimation procedure. Suppose we have T i.i.d. observations $\{x_1, \ldots, x_T\}$ of a random variable X. Our goal is to estimate the density function of X. The density of X at a point x is

$$f(x) = \lim_{h \to 0} \frac{1}{2h} P(x - h < X < x + h), \tag{2}$$

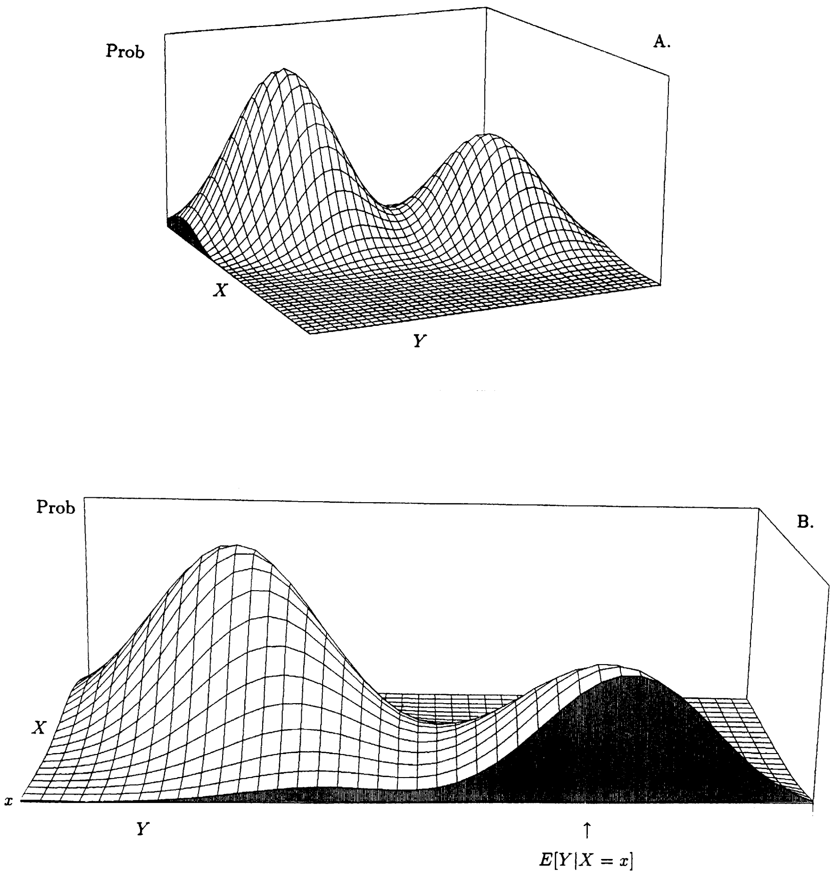


Fig. 2. An illustration of conditional expectations with a general non-symmetric distribution. Panel A plots a bivariate density of X and Y. In panel B, the conditional expectation $E[Y|X=x]$ is illustrated.

where $P(x-h<X<x+h)$ is the probability of the random variable X falling into $(x-h,\ x+h)$. A natural estimator of $f(x)$is

$$\hat{f}(x) = \frac{1}{2h}\left[\text{Frequency of } x_1,\dots,x_T \text{ falling into } (x-h, x+h)\right]$$
$$= \frac{1}{2hT}\left[\text{Number of } x_1,\dots,x_T \text{ falling into } (x-h, x+h)\right]. \tag{3}$$

Define the function $K(x)$ by

$$K(x) = \begin{cases} \frac{1}{2}, & \text{if } |x|<1 \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Then, the above histogram function, $\hat{f}(x)$, can be written as

$$\hat{f}(x) = \frac{1}{hT}\sum_{j=1}^{T} K\left[\frac{x - x_j}{h}\right]. \tag{5}$$

Suppose we have a random variable (say final exam grades) that can take on values between 0 and 100. The parameter h is chosen to be 5. To evaluate $f(70)$, we go through all the data and $K(x)$ assigns a value of one-half to any draw that lies between 65 and 75. Suppose that the sample size is $T = 100$. If 20 grades lie in the interval, then Eq. (5) assigns $\hat{f}(70) = \text{prob}(70) = 2\%$.

The parameter h can be interpreted as half the width of a bar in the histogram. It is also known as the bandwidth. From this simple example, it is clear that h plays a critical role in estimating the density. If the bandwidth is too large, a histogram will be represented by one large rectangle. If the bandwidth is too small, there will be many zeros and many small spikes.

Although the estimator $\hat{f}(x)$ is a density function (non-negative and integral over the whole space equal to 1), it is not a continuous function because it jumps at points $x_j \pm h$. Notice that $K(x)$ in Eq. (4) is also a density function and, if it is replaced by any other density function, the resulting estimator $\hat{f}(x)$ given by Eq. (5) will be a density function as well. So, if we choose a variety of densities a $K(x)$, called *kernels*, we get a variety of *kernel estimators* $\hat{f}(x)$ defined by Eq. (5). If $K(x)$ is continuous, then so is $\hat{f}(x)$. In fact, $\hat{f}(x)$ preserves all the smooth properties of $K(x)$. Generally, if we choose the kernel $K(x)$ such that

$$\int K(x)\,\mathrm{d}x = 1, \quad \int xK(x)\,\mathrm{d}x = 0, \quad \int x^2 K(x)\,\mathrm{d}x < \infty,$$

or under similar conditions, and if we choose the bandwidth h_T such that

$$\lim_{T\to\infty} h_T = 0, \quad \lim_{T\to\infty} Th_T = \infty,$$

then, the estimator $\hat{f}(x)$ given by Eq. (5) will converge to the true density with probability one. Moreover, it is unbiased and has the regular asymptotic normality results. For full statements and proofs of the results, see Robinson (1983), Silverman (1986), Bierens (1987) and Ullah (1988).[7]

The same intuition can be applied to the multivariate case. For the case of one conditioning variable, the nonparametric conditional expectation is

$$\hat{E}[r_t | x_t] = \frac{\sum_{j=1}^{T} r_j K\left[\frac{x_t - x_j}{\gamma_T}\right]}{\sum_{j=1}^{T} K\left[\frac{x_t - x_j}{\gamma_T}\right]} \qquad t = 1,\ldots,T. \tag{6}$$

[7] Many of the asymptotic results have been developed in the context of i.i.d. data. Robinson (1983) studies the case of dependent observations.

This is just the regression curve in Eq. (1) with the kernel estimator, $K(\cdot)$, inserted for the density. The market return is r_t, x_t is the conditioning variable (which can be dated $t-1$), T is the number of observations and γ_T is the bandwidth parameter that controls the amount of smoothing of the data. Additional conditioning variables can be accommodated by letting x_t be a vector.

One can interpret Eq. (6) as a weighted average of r. Define a weighting function

$$w_{jt} = \frac{K\left[\dfrac{x_t - x_j}{\gamma_T}\right]}{\sum_{j=1}^{T} K\left[\dfrac{x_t - x_j}{\gamma_T}\right]}. \tag{7}$$

The conditional expectation of r_t is

$$E[r_t | x_t] = \sum_{j=1}^{T} w_{jt} r_j. \tag{8}$$

This is equivalent to Eq. (6). Notice that the entire series of the returns appears on the right-hand side in Eqs. (6) and (8). This is analogous to using all of the data to estimate a linear regression parameter and the fitted values not being out-of-sample forecasts. However, no fixed parameters are being estimated in Eqs. (6) and (8). The function is evaluated at every point in the time series. Following Robinson (1983), the contemporaneous values of the market return are omitted. Robinson shows that Eq. (6) still provides a consistent estimator of the conditional expectation when the contemporaneous value is omitted.[8] If the conditioning variable is dated at $t-1$, the fitted conditional expectation for the last term, i.e. when $t = T$, is an out-of-sample forecast.

There are many type of kernels in addition to Eq. (4). The Gaussian kernel is

$$K\left[\frac{x_t - x_j}{\gamma_T}\right] = (2\pi)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\frac{x_t - x_j}{\gamma_T}\right)^2\right]. \tag{9}$$

The bandwidth parameter is proportional to $T^{-1/(4+q)}$ where q represents the number of conditioning variables (1, in this case). Silverman (1986) shows that this bandwidth parameter is proportional to the bandwidth that delivers the minimum mean square error.[9] Silverman discusses a number of different kernels and shows with simulation results that "there is very little to choose between the various kernels on the basis of mean integrated square error. It is perfectly legitimate, and indeed desirable, to base the choice of kernel on other considerations, for example, ... the computational effort involved." Throughout the

[8] This also eliminates the possibility that w_{jt} is set equal to 1 and all other values are set to 0.

[9] The multivariate Gaussian kernel is $(2\pi)^{-0.5}\exp[-0.5(\boldsymbol{x}_t - \boldsymbol{x}_j)\boldsymbol{\Gamma}^{-1}(\boldsymbol{x}_t - \boldsymbol{x}_j)']$ where $\boldsymbol{\Gamma}$ is a diagonal matrix with the squared bandwidths γ_{Tk} $k = 1, \ldots, q$ along the diagonal.

empirical section we use the Gaussian kernel (Eq. (9)). This kernel is used in financial applications by Pagan and Ullah (1988), Pagan and Hong (in preparation) and Pagan and Schwert (1990).

Eq. (6) can be implemented once the bandwidth is chosen. Following Silverman (1986), Pagan and Hong (in preparation) and Pagan and Schwert (1990), the bandwidth for the k conditioning variable is $\gamma_{Tk} = \sigma_k T^{-1/(4+q)}$ where σ_k is its standard deviation. This is an objective bandwidth selection that is designed to reflect the volatility of the conditioning variable. In the forecasting applications, the bandwidth is scaled to eliminate any bias in the forecasts.[10]

1.3. Other conditional moments

Since the kernel-based approach estimates the density, other moments can be extracted. We can estimate the conditional variance, $V[r_t|x_t]$, by noting that

$$V[r_t|x_t] = E[r_t^2|x_t] - (E[r_t|x_t])^2. \tag{10}$$

That is, we can get the conditional expectation of r^2 using Eq. (6) and subtract off the square of the conditional expectation of r_t. A similar method can be used to obtain conditional covariances.

It is also straightforward to get a conditional beta function. In the bivariate case, this could be formed by the ratio of the covariance to the variance. More generally, we can differentiate the nonparametric regression function (6) with respect to the variable of interest.[11]

$$\beta(r_t|x_t) = \frac{\partial E[r_t|x_t]}{\partial x_t} = \frac{\sum_{j=1}^{T} K\left[\frac{x_t - x_j}{\gamma_T}\right] r_j K'\left[\frac{x_t - x_j}{\gamma_T}\right] - \sum_{j=1}^{T} r_j K\left[\frac{x_t - x_j}{\gamma_T}\right] \sum_{j=1}^{T} K'\left[\frac{x_t - x_j}{\gamma_T}\right]}{\left(\sum_{j=1}^{T} K\left[\frac{x_t - x_j}{\gamma_T}\right]\right)^2}. \tag{11}$$

[10] Another approach is to search for the scaling factor that minimizes the mean squared error of the regression. However, as Stone (1974) points out, this approach should be executed on only part of the sample. The model can be cross-validated in the hold-out sample. Due to the scarcity of the data, I rely on an objective criterion for bandwidth selection.

[11] The conditional convexity function can be obtained by taking the second derivative of Eq. (6) with respect to x_t.

In the case of the Gaussian kernel in Eq. (9),

$$K'\left[\frac{x_t - x_j}{\gamma_T}\right] = -(2\pi)^{-\frac{1}{2}}\left[\frac{x_t - x_j}{\gamma_T}\right]\exp\left[-\frac{1}{2}\left(\frac{x_t - x_j}{\gamma_T}\right)^2\right]. \tag{12}$$

The beta in Eq. (11) is a function of x_t. That is, r_t may be very sensitive to changes in x_t in a certain range of x_t and insensitive in other ranges. Of course, if the data is multivariate normal, then the response of r_t to x_t is a fixed coefficient. This is a special case of Eq. (11).

1.4. Alternative nonparametric and seminonparametric methods

Many different nonparametric methods have been proposed and are reviewed in Silverman (1986) and Ullah (1988). One popular modification of Eq. (6) is to introduce a bandwidth that depends on the location in the data. The *variable-kernel estimator* replaces γ_T with the distance between x_j and its kth nearest point in the remaining $T-1$ observations. k is chosen to be a positive integer which is usually close to the square root of the sample size. In regions with few data points, this method increases the smoothing of the density.

A related method is the *nearest-neighbor estimator*. This methods replaces γ_T with the distance between x_t and its kth nearest neighbor.[12] In the nearest-neighbor estimator, the bandwidth depends on the point where the density is being estimated. In contrast, the bandwidth in the variable-kernel estimator is independent of x_t. For further details, see Ullah (1988).

Seminonparametric models reflect both parametric and nonparametric component and are reviewed in Robinson (1988). A parametric model can be fit to the data and the residuals can be run through a nonparametric program to investigate whether there are dependencies missed by the parametric model. Gallant et al. (in preparation) use this strategy in modeling the time-series behavior of the British Pound. Gallant and Tauchen (1989) use seminonparametric method to model time-varying conditional distributions. Seminonparametric methods have also been applied by Gallant et al. (1990) to evaluate bounds implied by asset pricing models.

1.5. The information environment

All of the statistical methods for obtaining conditional expectations assume that the information environment has been properly specified. If the econometrician omits information that investors consider important, then the fitted conditional expectations will not necessarily be close to the true conditional expectation. In

[12] This method is used by Diebold and Nason (1989) to forecast exchange rates.

least squares, this is the problem of omitted regressors. The nonparametric method will deliver a consistent estimate of the expectation conditioned on the specified information. However, if the true information environment is different from the measured one, there is no guarantee that the nonparametric estimates are close to the true conditional expectations.

Getting the information environment right is at least as important as properly specifying the mechanism that generates the conditional expectations. That is, even if we have correctly postulated a linear form for the conditional expectations, these expectations will be incorrect if a subset of the true conditioning information is used in the estimation. In asset pricing tests, restrictions may be incorrectly rejected if the tests are based on a subset of the information.[13] Later in the paper, I show that inference on the relation between conditional moments is very sensitive to the specification of information environment.

Unfortunately, there is no obvious solution to the problem of omitted information. Since true conditional expectations are unobservable, it is difficult to calibrate both the general functional form of conditional expectations and the information used in this function. However, one logical step is to conduct a sensitivity analysis. The analysis, which is pursued in this paper, involves changing both the conditioning information and the models that generate the conditional expectations.

2. Empirical applications

2.1. The simulated returns of a market timer

Successful market timing introduces a nonlinearity into the relation between portfolio returns and market returns. The market timer changes the mix of the portfolio toward market-sensitive stocks when the market move upward. In a period of a down market, the market timer will allocate towards market-insensitive securities.

The performance evaluation of the market timer is complicated. Standard techniques are inappropriate because they assume that the risk of the portfolio is fixed over the evaluation period. For example, Jensen's (1969) method would involve estimating a regression of the portfolio excess returns on the market excess returns. If the intercept is significantly different from zero, then the portfolio manager has 'outperformed' the market on a risk-adjusted basis-according to the model. This method assumes that the slope coefficient is the risk of the

[13] See Dybvig and Ross (1985) and Hansen and Richard (1987).

portfolio and that the risk is constant over the evaluation period. This market model assumes a fixed response to movements in the market. However, the idea of market timing is to have a large response when market returns are high and a low response when market returns are low. The slope coefficient is a function of the magnitude of the market return.

The linear regression of the portfolio return on the market return will produce biased estimates of the coefficients. Indeed, the intercept coefficient could appear to be significantly negative in a scenario when the portfolio manager always had higher risk adjusted returns than the market. The evaluation could incorrectly conclude that the portfolio manager "underperformed" the market.

Table 1 presents the simulated returns of two market timers. In the first panel, returns are generated according to the model

$$\mathrm{Simret}_t = xvw_t + 20\,xvw_t^2 + u_t, \tag{13}$$

where xvw represents the value-weighted NYSE portfolio in excess of the 1-month Treasury bill rate. The equity data are from the Center for Research in Security Prices (CRSP) and the bill data are from Ibbotson Associates. The simulated returns are generated with $u \sim N$ (0, 0.04). The market timer's returns are quadratic in the market return. This will guarantee larger returns than the market when the market moves up, neutral returns when the market moves down in small amounts and positive returns when the market moves sharply lower.

Three different evaluation periods are considered: 1964:5 to 1986:12 (276 observations), 1975:1 to 1986:12 (144 observations) and 1981:1 to 1986:12 (72 observations). Although these periods are arbitrary, the goal is to evaluate the sensitivity of the nonparametric regression performance to the number observations used to estimate the density.

The performance of the linear and nonparametric regressions are compared by regressing the portfolio returns on the fitted values from each technique. The true regression function was also estimated. The bandwidth in the nonparametric regression was set equal to the $\sigma_m T^{-1/5}$ where σ_m is the standard deviation of the market excess return and T is the number of observations. The results in Table 1 indicate that the nonparametric regression is able to closely approximate the true function. The first panel shows that the R^2 from the misspecified linear regression is 29% whereas the nonparametric regression delivers a R^2 of 80%. The true regression R^2 is also 80%. There is some evidence that the nonparametric regression is biased. However, this bias could be easily eliminated by varying the bandwidth.[14]

[14] By construction in the second stage regression reported in Table 1, the non-parametric fitted values are corrected for bias. The uncorrected fitted values produce slightly smaller R^2's.

Table 1
A comparison of nonparametric regression and ordinary least squares in fitting the simulated returns of two market timers

Model	α	β	R^2	Autocorrelation	Observations
Simulated returns $= xvw_t + 20xvw_t^2 + u_t$					
Linear OLS	0.000 (0.006)	1.000 (0.094)	0.292	0.111	276
Nonparametric	−0.001 (0.003)	1.091 (0.033)	0.803	0.008	276
True OLS	0.000 (0.003)	1.000 (0.030)	0.799	0.020	276
Linear OLS	0.000 (0.007)	1.000 (0.094)	0.441	−0.105	144
Nonparametric	−0.001 (0.004)	1.109 (0.043)	0.828	0.056	144
True OLS	0.000 (0.003)	1.000 (0.038)	0.833	0.029	144
Linear OLS	0.000 (0.008)	1.000 (0.096)	0.608	0.073	72
Nonparametric	0.000 (0.004)	1.127 (0.056)	0.853	−0.028	72
True OLS	0.000 (0.003)	1.000 (0.048)	0.828	0.000	72
Simulated returns $= 2xvw_t + 2Sign(xvw_t)xvw_t + u_t$					
Linear OLS	0.000 (0.003)	1.000 (0.022)	0.886	0.055	276
Nonparametric	−0.003 (0.002)	1.095 (0.020)	0.919	0.029	276
True OLS	0.000 (0.002)	1.000 (0.018)	0.919	0.033	276
Linear OLS	0.000 (0.004)	1.000 (0.026)	0.911	−0.007	144
Nonparametric	−0.004 (0.003)	1.109 (0.024)	0.936	0.035	144
True OLS	0.000 (0.003)	1.000 (0.022)	0.937	0.041	144
Linear OLS	0.000 (0.054)	1.000 (0.037)	0.913	0.072	72
Nonparametric	−0.003 (0.004)	1.137 (0.033)	0.944	−0.062	72
True OLS	0.000 (0.005)	1.000 (0.032)	0.934	−0.018	72

Two models are considered. In the first, the market timer's returns are quadratic in the market return. In the second, the returns are influenced by a slope dummy on the market returns. The analysis uses the excess return on the value-weighted NYSE portfolio, *xvw*. The disturbance, *u*, is a random normal variable with zero mean and standard deviation of 0.04. Diagnostics are performed by running the following regression:

$$Simret_t = \alpha + \beta Fit_t + \epsilon_t,$$

where Fit represents the fitted values from linear OLS regression, correctly specified OLS regression and nonparametric regression. Heteroskedasticity consistent standard errors are in parentheses.

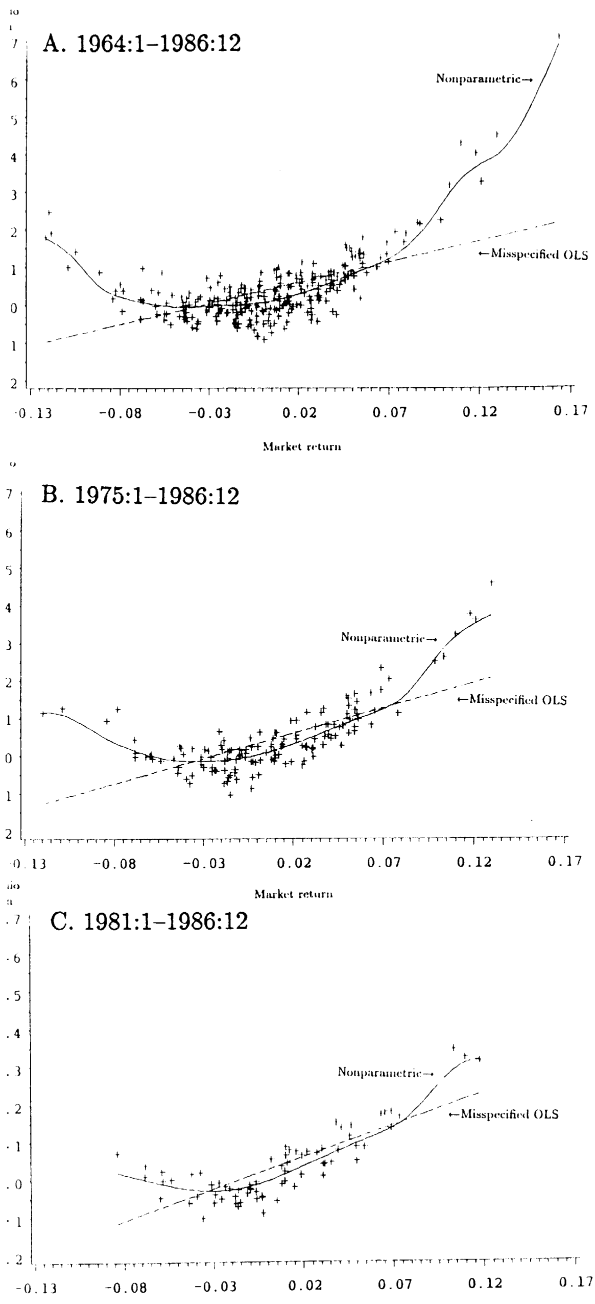
Table 1 also shows that the performance of the nonparametric regression does not appear to be affected by the smaller sample size. This can also be seen in Fig. 3 which plots the actual data, the misspecified OLS fitted values and the nonparametric fitted values. Panels A–C show the poor performance of the linear

---

Fig. 3. A comparison of nonparametric regression and ordinary least squares in fitted the simulated returns of a market timer. Simulated returns are generated from the model

$$Simret_t = R_{mt} + 20R_{mt}^2 + u_t,$$

where R_{mt} represents the value-weighted NYSE return and u_t is a normal variable with a standard deviation of 0.04.



A. 1964:1–1986:12
Nonparametric→
←Misspecified OLS
-0.13
-0.08
-0.03
0.02
0.07
0.12
0.17
Market return
B. 1975:1–1986:12
Nonparametric→
←Misspecified OLS
Market return
C. 1981:1–1986:12
Nonparametric→
←Misspecified OLS
Market return

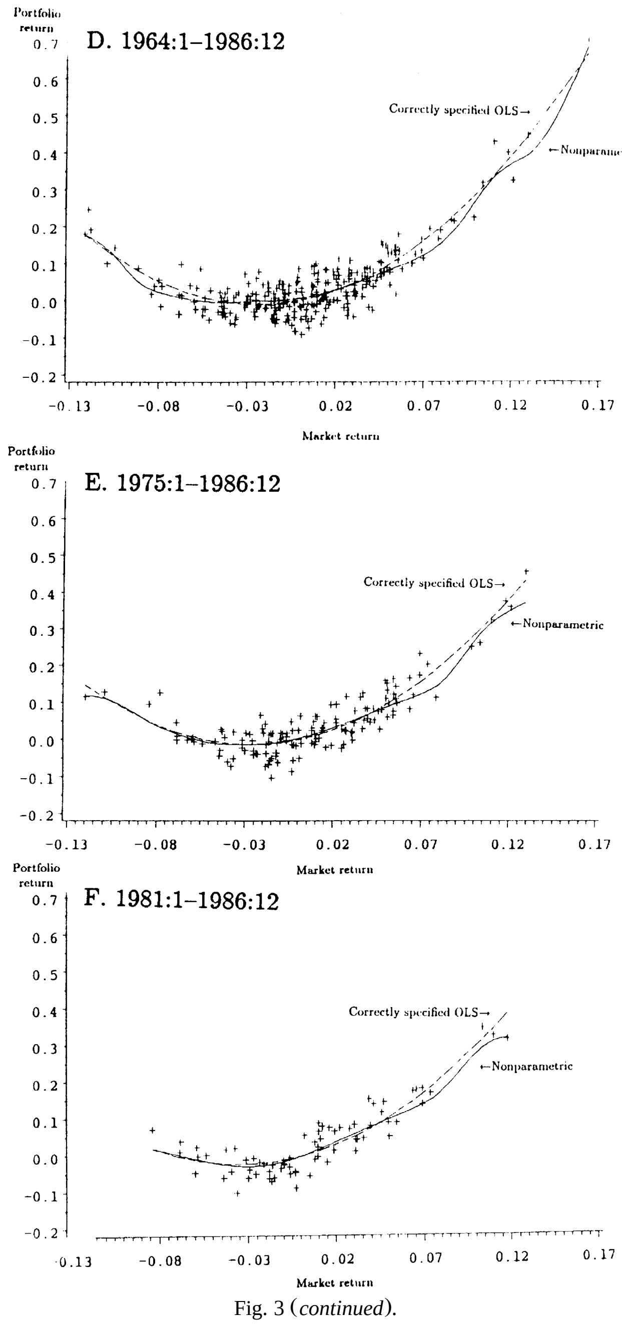


Fig. 3 (*continued*).

OLS estimator. Panels D–F show that the nonparametric regression closely approximates the true regression—especially where the data are concentrated. The fit deteriorates in regions of large market moves. This is to be expected. The large returns are rare events and it is difficult to estimate the empirical density with very few observations. The fit could be potentially improved with a variable kernel estimator which would increase the bandwidth in this region of sparse data.

The second panel of Table 1 and Fig. 4 presents the simulated returns of a second market timer. Returns are generated according to

$$\mathrm{Simret}_t = 2\,xvw_t + \mathrm{Sign}(\,xvw_t)\,xvw_t + u_t. \tag{14}$$

This model is piecewise linear. Compared to Eq. (13), this model is much more linear. In this example, the linear model gives a reasonable (albeit incorrect) approximation. However, the results in Table 1 suggest that the nonparametric model delivers a fit that is closer to the true regression.

Panels A–C of Fig. 4 shows that the misspecified OLS provides fitted values that are too low when the market return is both very high and very low. Panel D presents the nonparametric fitted values for the full sample and the true regression fitted values. The nonparametric regression picks up the 'kink' in the true regression. Where the data is dense, both the nonparametric and the true regression fitted values are virtually identical. This is also true for the other subperiods. However, similar to the previous example, the nonparametric estimator deteriorates at extreme points in the data—but so does the misspecified OLS.

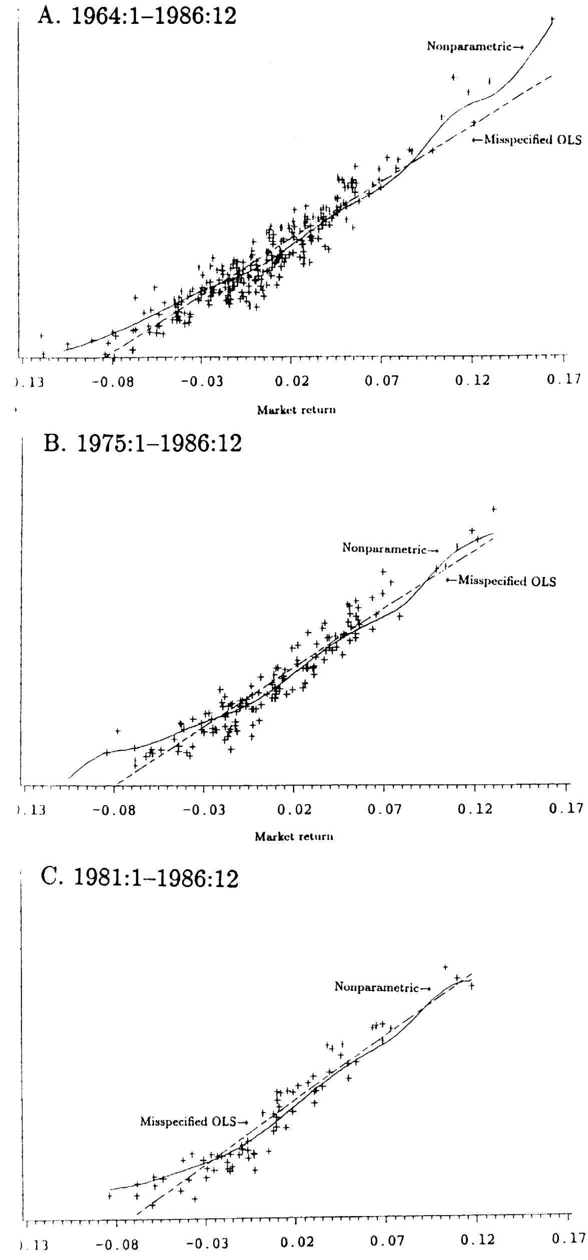
Both simulations suggests that the nonparametric density technique could be useful in detecting nonlinearities in the returns of portfolio managers. The nonparametric regression closely approximated the true regression in various sample sizes. However, the nonparametric fit deteriorates in regions where there are few observations.

2.2. Market model regressions

Risk is usually measured by a linear regression of portfolio returns on the market returns. In the market timing example, the linear model did not do a good job in fitting the data. There are reasons to believe that nonlinearities exist in other portfolio returns. For example, the leverage hypothesis suggests that in up markets the equity beta of the firm should decrease because the firm's market debt–equity ratio will decrease. In down markets, the firm's equity beta should increase.

The nonparametric regression may help us identify some of these patterns. Table 2 presents three different 'market model' regressions for a value-weighted portfolio of firms in industries categorized as leisure industries (SIC 27, 58, 70, 78, 79).[15] Braun et al. (1990) show that the betas in this industry are highly

[15] This portfolio is studied in Breeden et al. (1989), Harvey and Zhou (1990), and Ferson and Harvey (1991). The sample corresponds to the one used by Ferson and Harvey.


A. 1964:1–1986:12
Nonparametric→
←Misspecified OLS
-0.08
-0.03
0.02
0.07
0.12
0.17
Market return
B. 1975:1–1986:12
Nonparametric→
←Misspecified OLS
-0.08
-0.03
0.02
0.07
0.12
0.17
Market return
C. 1981:1–1986:12
Nonparametric→
Misspecified OLS→
-0.08
-0.03
0.02
0.07
0.12
0.17
Market return

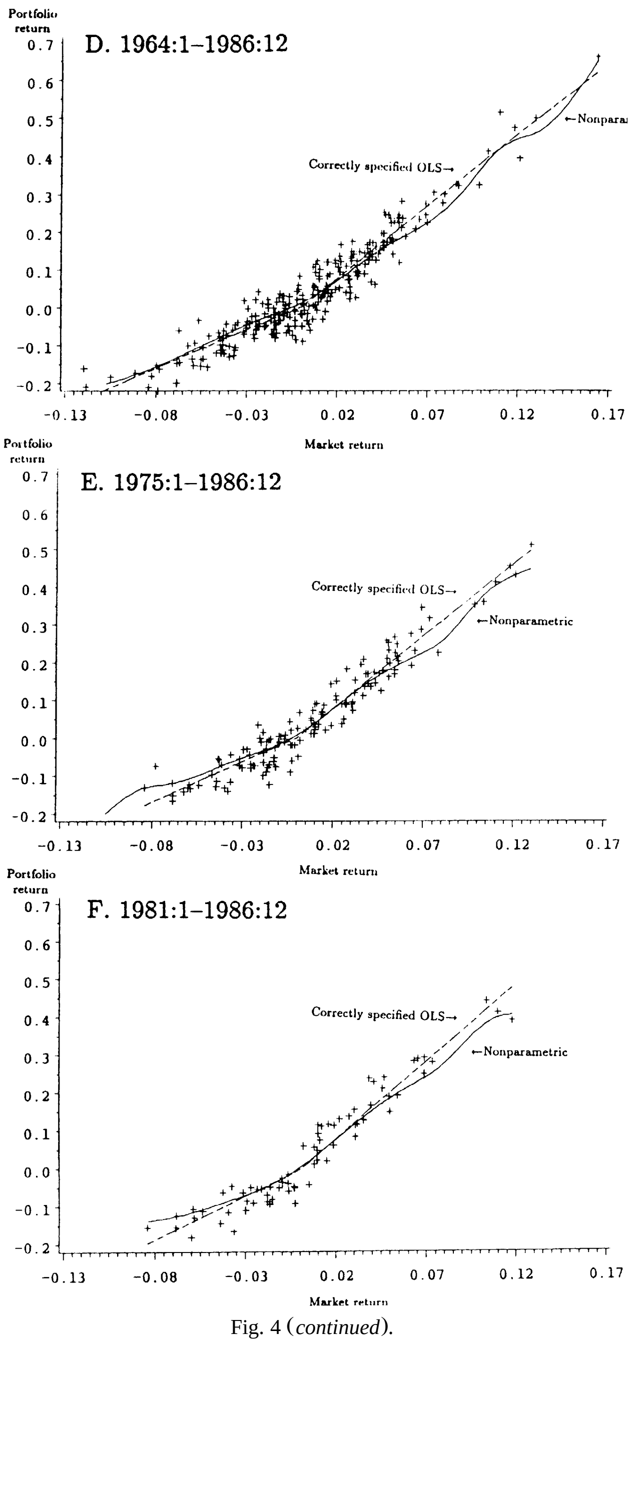


Fig. 4 (*continued*).

variable. The first model presented is a linear regression of the portfolio return on the market return. The second model includes a slope dummy variable. Finally, the nonparametric regression model is presented.

Comparing the linear model with the model with the slope dummy reveals that the beta is higher when market returns are low. The results in Table 2 indicate that the slope dummy enters the full sample and all of the subperiod regressions with a negative coefficient. In the full sample, the coefficient on the slope dummy is almost four standard errors from zero. However, the slope dummy specification is at best approximate. The leverage argument would predict higher betas in low return markets; however, it would not predict a piecewise linear relation.

Given that the true regression function is unknown, the nonparametric approach may give us some insight about potential nonlinearities in the data. The results in Table 2 indicate that fitted values from the nonparametric regression generally explain more of the variation of this industry's returns than the linear market model regression. The increment in explanatory power is of the magnitude observed in Table 1 where a piecewise linear model was compared to the nonparametric model.

Although there is an increase in explanatory power, the increase is fairly small. Another way of assessing the predictions of the two models is to run a regression of the portfolio returns on the fitted values from both the linear and nonparametric regressions. The last two columns report these coefficients. The weights observed for the predictions from the nonparametric model are always close to one. On the other hand, the weights on the linear regression estimates are small and negative. These results suggest that the nonparametric approach has some incremental ability to fit the portfolio returns.

Given the incremental explanatory power of the nonparametric regression, what can we learn about the data? The four panels of Fig. 5 provide plots of the nonparametric response function (11) against the market return. The OLS beta with a slope dummy is also plotted. In each plot, the OLS beta shifts downward when the market return is positive. In the simulation in Table 1, the nonparametric regression was able to closely approximate a piecewise linear relation—when a piecewise linear model was the true returns generating process. The beta function in Fig. 5 suggests that a piecewise linear model is probably not the correct model.

Four different samples are presented in Fig. 5. The first panel plots the betas against the market return for the full sample 1964:5 through 1986:12. The next

Fig. 4. A comparison of nonparametric regression and ordinary least squares in fitted the simulated returns of a market timer. Simulated returns are generated from the model

$$Simret_t = 2R_{mt} + sign(R_{mt})R_{mt} + u_t,$$

where R_{mt} represents the value-weighted NYSE return and u_t is a normal variable with a standard deviation of 0.04.

Table 2
A comparision of three market models for the leisure industry portfolio (SIC 27, 58, 70, 78, 79)

Sample	OLS β (market)	OLS with slope dummy β (market)	OLS with slope dummy β (dummy)	R^2 Market model	R^2 Nonparametric	Weight[a] market model	Weight[a] nonparametric
64:5–86:12	1.298 (0.044)	1.548 (0.084)	−0.473 (0.137)	0.762	0.787	−0.111	1.224
64:5–74:4	1.545 (0.072)	1.796 (0.126)	−0.559 (0.232)	0.795	0.809	−0.014	1.117
74:5–84:4	1.184 (0.063)	1.347 (0.128)	−0.291 (0.199)	0.749	0.778	−0.360	1.515
64:5–69:4	1.287 (0.097)	1.576 (0.186)	−0.363 (0.308)	0.780	0.785	−0.170	0.950
69:5–74:4	1.604 (0.104)	1.819 (0.182)	−0.506 (0.355)	0.805	0.802	−0.033	1.201
74:5–79:4	1.299 (0.088)	1.570 (0.180)	−0.472 (1.274)	0.787	0.833	−0.418	1.606
79:5–84:4	1.029 (0.087)	1.084 (0.178)	−0.102 (0.287)	0.706	0.699	−0.211	1.406

The first model is the standard linear market model. The second model includes a slope dummy variable which takes the value of one when market returns are positive. The third model results from a nonparametric regression. R^2's are for OLS and nonparametric regressions of the returns on a value weighted portfolio composed of leisure industry equities on the value weighted NYSE portfolio. Heteroskedasticity consistent standard errors are in parentheses.

[a]The weights are coefficients from a regression of the industry returns on fitted values from the linear market model and the nonparametric regressions.

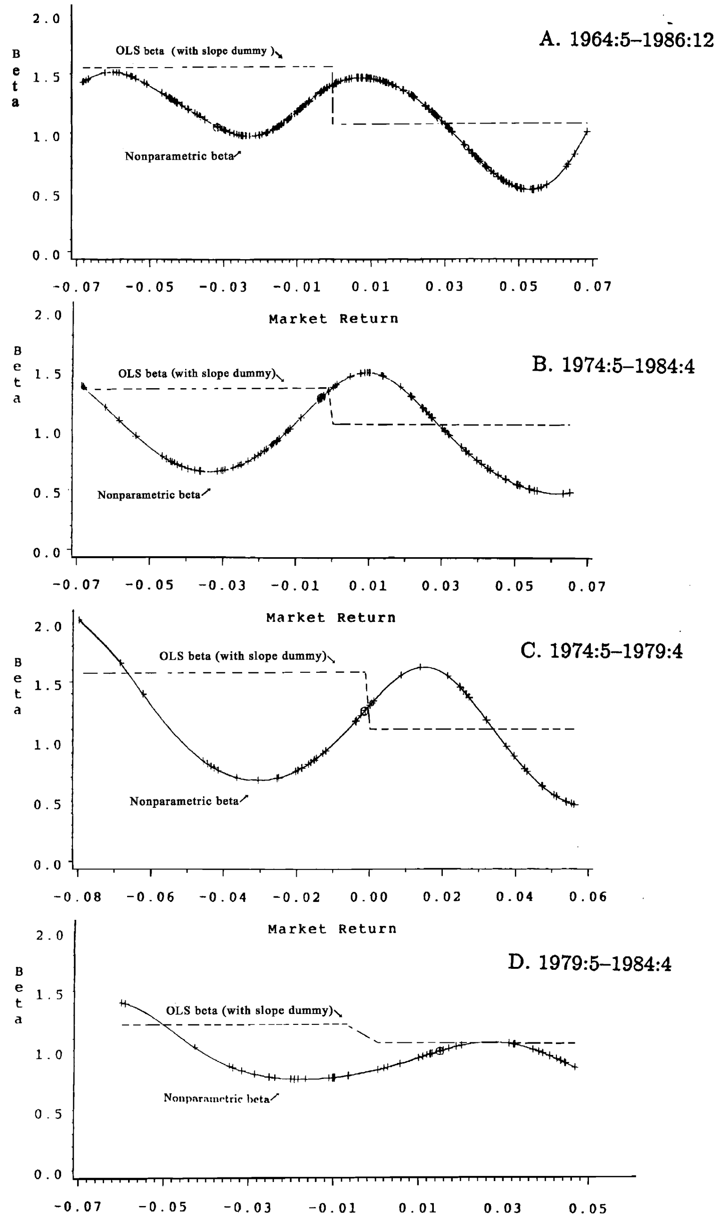


Fig. 5. A comparison of OLS and nonparametric beta coefficients for the leisure industry (SIC 27, 58, 70, 78, 79) from 1964:5 to 1986:12.

three panels consider smaller sub-samples. The range of betas over this period corresponds to the ones presented in Braun et al. (1990) using rolling regression and EGARCH methods. Although not reported, the average nonparametric standard error of the beta coefficients is 0.17 and ranges from 0.13 to 0.33.

Panel A indicates that the betas are generally lower in down markets—however the relation is not a simple one. The betas are highest when the market moves down sharply (which is consistent with the leverage hypothesis) and when the market does not move by much. Indeed, there is a hump in the beta plot at the level of the average market return (1%). The beta drops when the market return is lower than the average and when the market is higher than the average. This pattern in the betas is not consistent with the leverage hypothesis. However, it may be consistent with an infrequent trading explanation.

Potential nonlinearities in the market model may be important if the betas are estimated over a long horizon. To illustrate Bayesian tests of asset pricing models, Harvey and Zhou (1990) report market model betas for 12 industry portfolios estimated with monthly data over 1926–1987 period. However, if a slope dummy is included, it enters six of their regressions with a *t*-statistic greater than 2.0. If a quadratic term is included, it enters eight of their 12 regressions with a *t*-statistic greater than 2.

Although it is not common to use such long intervals for beta estimation, a number of papers studying the abnormal returns of "winners" and "losers" estimate betas over periods of up to 52 years.[16] These studies obtain abnormal returns from a market model regression. However, there is evidence that these market model regressions are misspecified. Chopra et al. (1992) find dramatic differences between betas depending upon the direction of the market return. For example, in up markets, they report average betas for the winner portfolio to be 1.7 in the 4 years prior to ranking. Over the same period, the betas average only 1.0 in down markets. For their loser portfolios, the average beta is 1.8 in the post ranking period in up markets and 1.1 in down markets. Although these results suggest that the linear market model is not properly specified, they (and others) proceed with their analysis assuming that the betas are constant. This is a situation where the nonparametric regression technique could help us understand the nature of the nonlinearities in the data.

2.3. Mimicking portfolios

In many applications in finance, portfolios are required that have maximum correlation with some variable. For example, portfolios are proposed by Breeden (1979) and used in Breeden et al. (1989) that mimic the behavior of aggregate

[16] See DeBondt and Thaler (1985), Chan (1988), Ball and Kothari (1989) and Chopra et al. (1992).

consumption. In other applications, mimicking portfolios are constructed to hedge the price variability of certain securities. The nonparametric regression can be used in both of these applications.

A hedge portfolio is constructed to have maximum correlation with the instrument being hedged. In the linear regression model, the price changes of the cash instrument are projected on to the price changes of the hedging instrument which is usually a futures contract. The regression slope coefficient is the hedge ratio. It approximates the sensitivity of the cash price to a change in the price of the hedge instrument. It also reveals the recommended position in the futures market. The goal of this position is to minimize the conditional variance of the portfolio of the cash and futures instrument.

One of the more challenging instrument to hedge is a mortgage-backed security such as a Government National Mortgage Association (GNMA) pass-through. This security is difficult to hedge because it has an embedded option which is difficult to price. When interest rates fall, borrowers have the option of prepaying their mortgages and refinancing at lower rates. The stated life of the mortgage might be 30 years, however, the effective life is much shorter.

The borrower has a portfolio that resembles a short bond plus a call option with an exercise price at par. This would appear to be straightforward to hedge. However, there are numerous complications. Borrowers prepay for other reasons. If you move, the mortgage is refinanced. The rate of prepayment will be affected not just by interest rates but also by economic activity. To further complicate the problem, some borrowers prepay when it does not appear optimal and other borrowers fail to prepay very high interest rate mortgages when rates are low. As such, the mortgage-backed security is very difficult to model.[17] Most rely on an empirically-based model.

Fig. 6 presents plots of three GNMA coupon securities: 8%, 11.5% and 15% against the nearest-to-delivery Treasury bond futures contract over the 1982–1986 sample. The monthly data are from Breeden and Giarla (1989). The option feature of the GNMAs is evident when the Treasury bond price is high (rates are low). It is also evident from these graphs that some of these mortgage-backed securities trade above par.

The linear regression of the GNMA on the Treasury bond as well as the nonparametric regression fitted values are also plotted in Fig. 6. The linear regression is unable to fit the nonlinearities in these data induced by the embedded prepayment option. The nonparametric regression shows some promise in replicating the option feature even though there are only 60 data points used to estimate the bivariate density.

[17] Stanton (1990) proposes a theoretical model that incorporates some of these features.

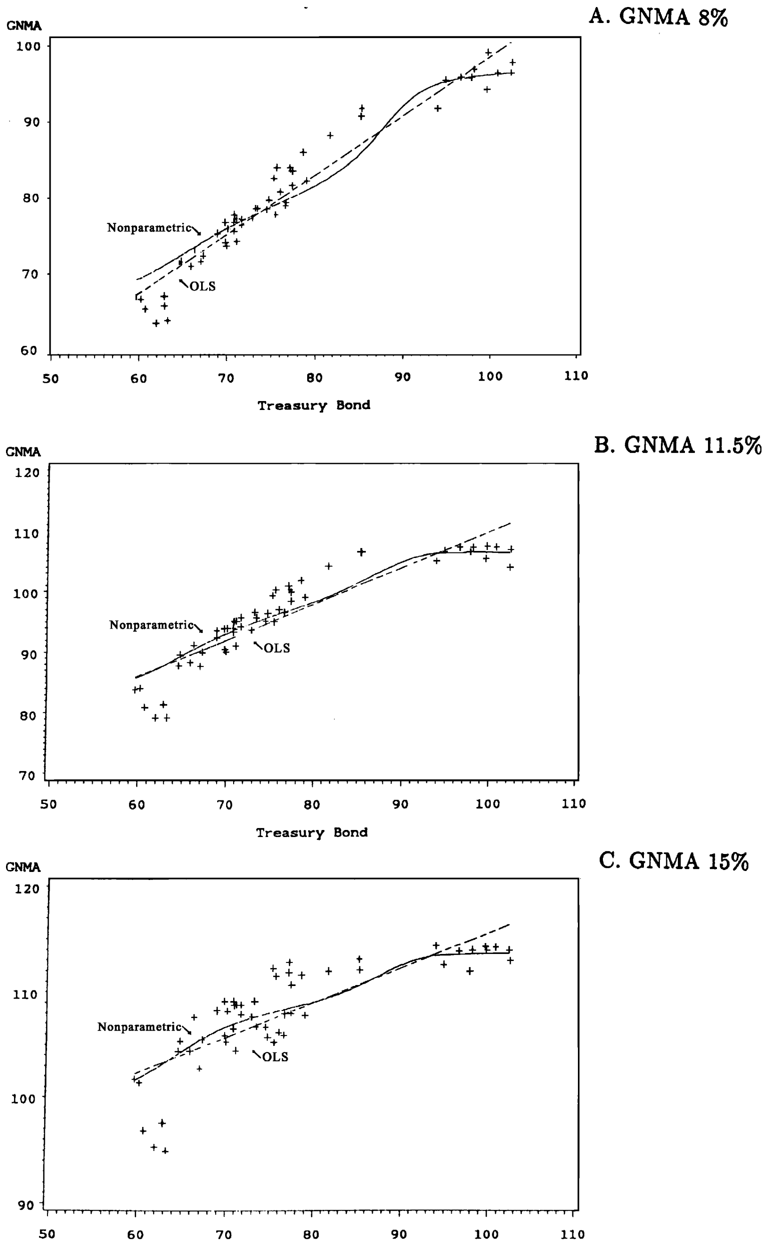


Fig. 6. OLS and nonparametric regressions for three mortgage-backed securities: 1982–1986.

Next, consider the hedging performance of the nonparametric regression. One criticism of nonparametric techniques is that they perform poorly on an out-of-sample basis. Hedging performance must be evaluated on an out-of-sample basis. The nonparametric regressions will be rolled through time to obtain hedge weights for the each month.

The in-sample analysis of linear and nonparametric rolling regressions is presented in Table 3. The models are estimated 12 times through November 1986. This evaluation period is chosen because it is considered one of the most difficult periods to hedge. Wall Street investment banks with mortgage-backed portfolios experienced large losses during this period because their hedges performed poorly. Most of these hedges were based on linear regressions.

The first two columns of Table 3 show that the OLS and nonparametric R^2 statistics are fairly similar throughout 1986 for the GNMA 8%. The nonparametric technique delivers in-sample R^2's that are about 3% larger than the linear regression. However, the difference in the fit is much more dramatic for the 11.5%. By mid-1986, there is a difference of more than 20% in the linear and nonparametric R^2's. A similar pattern is found in the GNMA 15%. In the first few months, the R^2 from the linear and nonparametric regressions are similar. However, by mid-1986 the nonparametric R^2 is almost twice the size of the linear regression R^2. These results indicate that the nonparametric technique is able to more closely mimic the returns of the mortgage-backed security. However, these statistics represent the in-sample fit.

Table 4 presents the out-of-sample hedging performance of three models: rolling linear regression, rolling nonparametric, and the Breeden and Giarla (1989) dynamic hedging technique. The dynamic hedging technique is nonlinear in that the hedge weights depend on the price of the mortgage. Breeden and Giarla estimate the elasticities of various coupon mortgages in 1984. These mortgage elasticities are used throughout the evaluation period.[18] The mortgage elasticities combined with the elasticity of the Treasury bond futures produces a hedge weight.

Table 4 provides a comparison of the various techniques in 1986 for the three mortgage backed securities. At the beginning of each month, the portfolio consists of US$100 million in par value mortgages. In the first column represents the gain (or loss) for holding this portfolio for 1 month. There are a number of component to this cash flow. The mortgage appreciates or depreciates depending upon economic conditions. However, some of the mortgage pool prepays. That is, if the mortgage started out at par and appreciates in value by 10%, the holder doe not necessarily make 10%. If 50% of the pool prepays, then the capital appreciation is only 5%. In addition, GNMAs pay a monthly coupon equal to one-twelfth of the

[18] These elasticities are reported in Fig. 21 of Breeden and Giarla (1989).

Table 3
In-sample analysis of mortgage price changes regressed on Treasury bond futures price changes: rolling OLS and nonparametric regressions, 1982–1986

Estimated through	R^2 OLS GNMA 8.0%	R^2 nonparametric GNMA 8.0%	R^2 OLS GNMA 11.5%	R^2 nonparametric GNMA 11.5%	R^2 OLS GNMA 15.0%	R^2 nonparametric GNMA 15.0%
123185	0.871	0.859	0.830	0.826	0.502	0.540
13186	0.866	0.855	0.828	0.825	0.497	0.540
22886	0.848	0.854	0.668	0.812	0.354	0.530
33186	0.772	0.808	0.511	0.723	0.351	0.527
43086	0.769	0.806	0.466	0.676	0.349	0.522
53086	0.790	0.820	0.477	0.671	0.320	0.501
63086	0.787	0.815	0.456	0.638	0.295	0.471
73186	0.765	0.797	0.445	0.628	0.302	0.475
82986	0.763	0.792	0.432	0.601	0.304	0.474
93086	0.760	0.782	0.405	0.557	0.268	0.440
103186	0.760	0.782	0.403	0.555	0.268	0.438
112886	0.759	0.781	0.401	0.553	0.260	0.435

R^2's are for OLS and nonparametric regressions of the mortgage price change on change in the nearest to delivery Treasury bond futures contract. Regressions are estimated up to the date in the first column. The full sample contains 59 monthly observations.

Table 4
Summary of out-of-sample mortgage-backed securities hedging performance in 1986: nonparametric density estimation versus rolling OLS regression and dynamic hedging technique

Date	Change in mortgage value	T-bonds short rolling OLS	Futures gain rolling OLS	Net gain rolling OLS	T-bonds short dynamic	Futures gain dynamic	Net gain dynamic	T-bonds short nonparametric	Futures gain nonparametric	Net gain nonparametric
GNMA 8.0%										
13186	−1 010 792	837	52 313	−958 479	705	44 063	−966 729	794	49 625	−961 167
22886	4 775 417	830	−8 014 688	−3 239 271	698	−6 740 063	−1 964 646	539	−5 204 719	−429 302
33186	849 250	721	−5 407 500	−4 558 250	558	−4 185 000	−3 335 750	0	0	849 250
43086	167 583	612	956 250	1 123 833	502	784 375	951 958	0	0	167 583
53086	−4 473 438	617	4 241 875	−231 563	514	3 533 750	−939 688	467	3 210 625	−1 262 813
63086	2 556 031	635	−3 611 563	−1 055 531	608	−3 458 000	−901 969	371	−2 110 063	445 969
73186	1 718 562	602	1 053 500	2 772 062	570	997 500	2 716 062	211	369 250	2 087 812
82986	2 152 375	613	−2 815 969	−663 594	534	−2 453 063	−300 688	450	−2 067 188	85 188
93086	−1 647 938	591	3 453 656	1 805 719	508	2 968 625	1 320 687	422	2 466 063	818 125
103186	1 195 896	598	−915 688	280 208	544	−833 000	362 896	349	−534 406	661 490
112886	2 373 036	594	−965 250	1 407 786	538	−874 250	1 498 786	547	−888 875	1 484 161
123186	−259 115	600	918 750	659 635	536	820 750	561 635	544	833 000	573 885
Total	8 396 870		−11 054 313	−2 657 443		−9 394 313	−997 443		−3 876 688	4 520 182
GNMA 11.5%										
13186	−401 250	1015	63 438	−337 813	395	24 688	−376 563	1015	63 438	−337 813
22886	1 261 917	1012	−9 772 125	−8 510 208	393	−3 794 906	−2 532 990	697	−6 730 406	−5 468 490
33186	−2 418 490	751	−5 632 500	−8 050 990	336	−2 520 000	−4 938 490	0	0	−2 418 490
43086	3 644 031	567	885 938	4 529 969	451	704 688	4 348 719	0	0	3 644 031
53086	−2 091 083	572	3 932 500	1 841 417	309	2 124 375	33 292	521	3 581 875	1 490 792
63086	729 604	581	−3 304 438	−2 574 833	520	−2 957 500	−2 227 896	0	0	729 604

73186	1 350 469	526	920 500	2 270 969	309	540 750	1 891 219	0	0	1 350 469
82986	583 594	533	− 2 448 469	− 1 864 875	321	− 1 474 594	− 891 000	522	− 2 397 938	− 1 814 344
93086	615 578	496	2 898 500	3 514 078	300	1 753 125	2 368 703	284	1 659 625	2 275 203
103186	219 073	493	− 754 906	− 535 833	330	− 505 313	− 286 240	0	0	219 073
112886	384 250	485	− 788 125	− 403 875	322	− 523 250	− 139 000	548	− 890 500	− 506 250
123186	481 615	475	727 344	1 208 958	314	480 813	962 427	524	802 375	1 283 990
Total	4 359 307		− 13 272 344	− 8 913 036		− 6 147 125	− 1 787 818		− 3 911 531	447 776
GNMA 15.0%										
13186	1 317 604	524	32 750	1 350 354	219	13 688	1 331 292	499	31 188	1 348 792
22886	− 179 156	527	− 5 088 844	− 5 268 000	222	− 2 143 688	− 2 322 844	522	− 5 040 563	− 5 219 719
33186	1 507 000	359	− 2 692 500	− 1 185 500	187	− 1 402 500	104 500	0	0	1 507 000
43086	432 292	320	500 000	932 292	169	264 063	696 354	0	0	432 292
53086	289 010	323	2 220 625	2 509 635	173	1 189 375	1 478 385	515	3 540 625	3 829 635
63086	103 359	313	− 1 780 188	− 1 676 828	194	− 1 103 375	− 1 000 016	0	0	103 359
73186	− 2 060 333	278	486 500	− 1 573 833	176	308 000	− 1 752 333	0	0	− 2 060 333
82986	1 213 667	287	− 1 318 406	− 104 740	178	− 817 688	395 979	502	− 2 306 063	− 1 092 396
93086	1 120 760	275	1 607 031	2 727 792	167	975 906	2 096 667	147	859 031	1 979 792
103186	360 167	309	− 473 156	− 112 990	184	− 281 750	78 417	0	0	360 167
112886	361 333	263	− 427 375	− 66 042	180	− 292 500	68 833	201	− 326 625	34 708
123186	280 833	257	393 531	674 365	175	267 969	548 802	179	274 094	554 927
Total	4 746 536		− 6 540 031	− 1 793 495		− 3 022 500	1 724 036		− 2 968 313	1 778 224

The change in the mortgage value calculated as the change in value of the principal in the pool, plus the prepayments at par, plus the monthly coupon minus the Eurodollar financing cost. US$100 million in face-value is invested each month. All hedge ratios are estimated 1 month prior. T-bonds short represents the number of Treasury bond contracts sold on the Chicago Board of Trade. The nearest to delivery contract is used. The rolling OLS represents the hedge position suggested by a linear regression of mortgage price changes on the Treasury bond price changes. Dynamic represents the hedge position implied by the ratio of the elasticity of the mortgage to the elasticity of the Treasury bond futures. Nonparametric represents the hedge position from the beta function (11). Negative nonparametric hedge ratios are set equal to 0.

stated coupon rate. Finally, I subtract one-twelfth of the stated Eurodollar rate. This approximates the cost of funds for purchasing the mortgage-backed security.

Mortgage hedging received considerable attention in March 1986 when many investment banks reported large losses. To hedge the mortgage, the investment banks shorted the Treasury bond futures. The problem can be illustrated by examining the linear OLS hedge performance during that month. In March, interest rates dropped dramatically. The GNMA 8% appreciated by a small amount because an acceleration in the prepayments was anticipated. However, the lower rates caused large losses in the futures. The linear strategy suffered a US$5.4 million loss in the futures that month which was not offset by a gain of the same magnitude in the mortgage. The net losses for the 11.5% were more severe; a US$5.6 million loss was registered in the futures while the mortgage lost US$2.4 million as a result of increased prepayments. In this month, the hedge did not provide offsetting cash flows.

The results in Table 4 and in panel A for Figs. 7–9 show that the OLS based hedges performed poorly during 1986. The net losses for the 8% GNMA are US$2.6 million, US$8.9 million for the GNMA 11.5% and US$1.7 million for the GNMA 15%. The dynamic hedging technique fares much better. After interest rates decreased in February, the mortgage elasticities decreased (with higher prices) and the number of short Treasury bond contracts were reduced. However, the number of contracts is not reduced enough to spare large losses in March. During that month, the dynamic hedge lost US$3.3 million in the GNMA 8%, US$4.9 million in the GNMA 11.5% and broke even on the GNMA 15%. However, over the full 12-month period, the technique performed adequately. The results in Table 4 and panel B of Figs. 7–9 show that the dynamic technique lost US$1 million in the GNMA 8%, and US$1.8 in the GNMA 11.5%. For the GNMA 15%, there was a US$1.7 million profit.

Finally, the results of the nonparametric regressions are presented. Hedge ratios are calculated from the beta function given by Eq. (11). These hedge ratios are highly variable which is probably due to the scarcity of data used to estimate the joint density. The hedge ratios actually go negative in some months. This would imply a long position in the Treasury bond futures. Such a position is unusual for a hedge. However, it is not unreasonable. For example, in March of 1986 mortgage prices decreased and Treasury bond prices increased. A long position in the Treasury futures would have been highly profitable. To be conservative, I set the hedge ratios equal to zero if the response function drops below zero.

The results in Table 4 and panel C of Figs. 7–9 indicate that the nonparametric hedging appears to work. For the GNMA 8%, there is a profit of US$4.5 million. A smaller gain of US$0.5 million is registered for the GNMA 11.5% and there is a profit of US$1.8 million for the GNMA 15%.

The performance of the nonparametric hedging can be attributed to the technique's ability to pick up the nonlinearity induced by the prepayment option. When rates increased in February 1986, the nonparametric response function

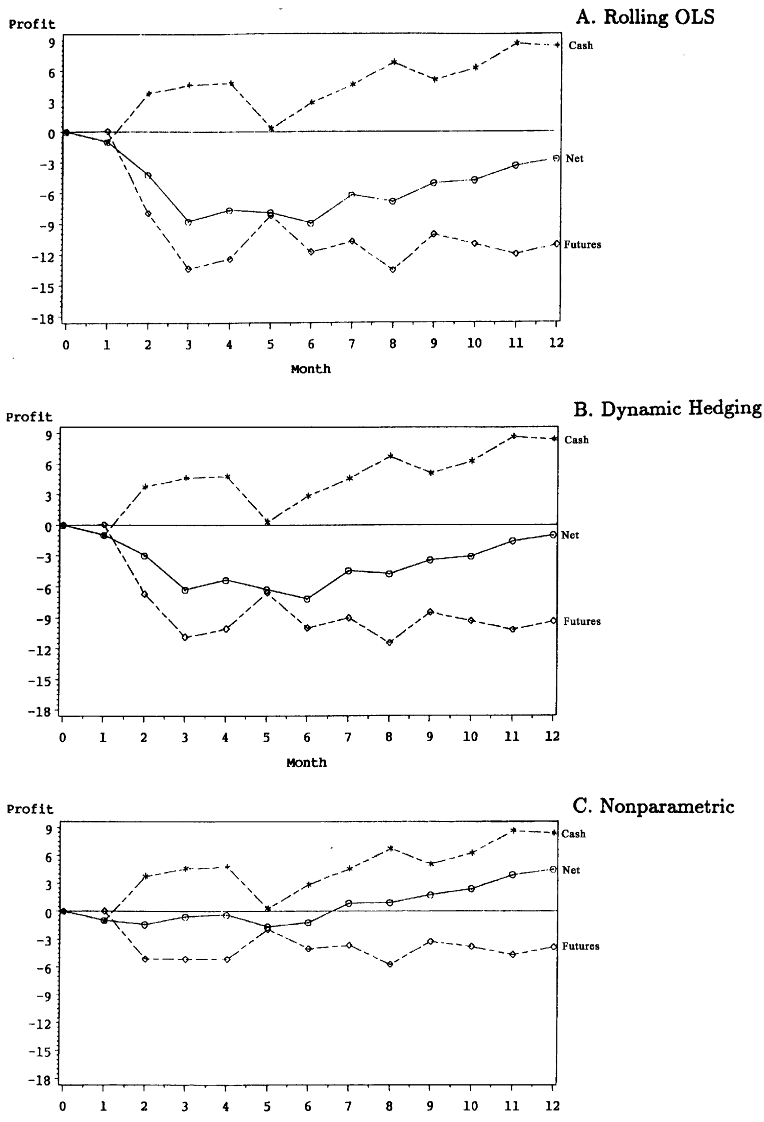


Fig. 7. GNMA 8% out-of-sample cumulative futures hedge performance in 1986 for a US$100 million portfolio using three models: rolling OLS regression, Breeden–Giarla dynamic hedging and nonparametric regression.

sharply decreases indicating that the mortgage is relatively insensitive to changes in the Treasury bond futures price. This flexibility reduces the number of short Treasury bond contracts thereby improving the performance of the hedge.

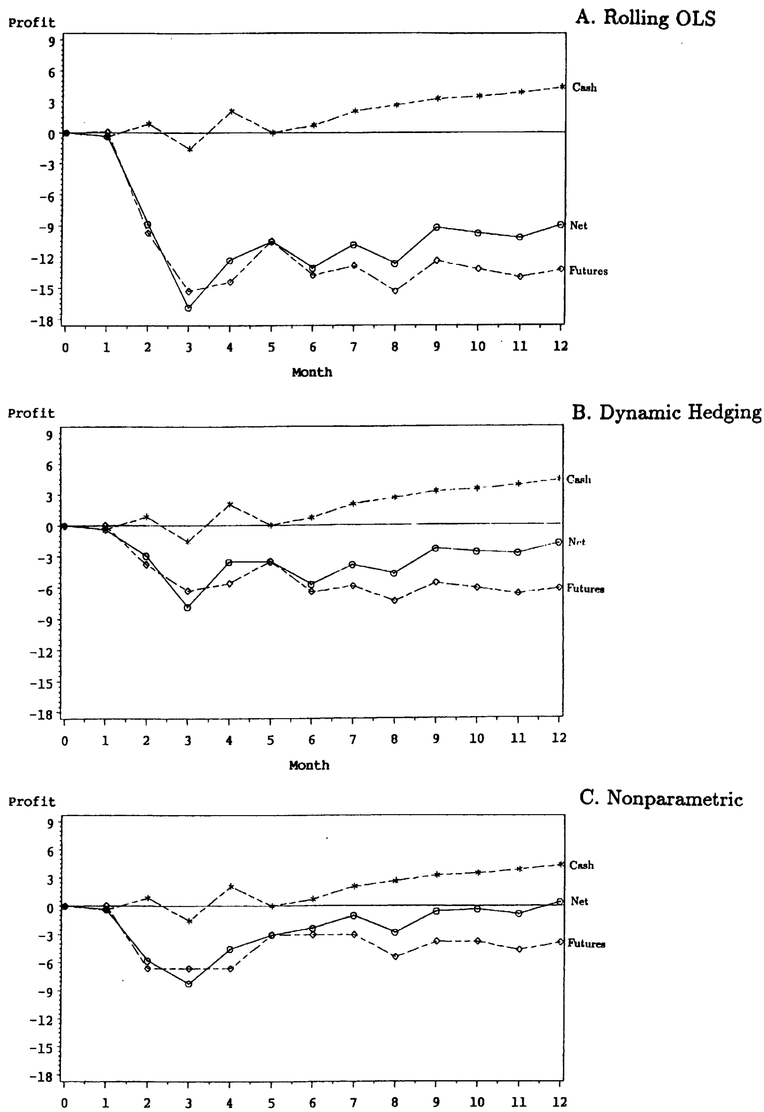


Fig. 8. GNMA 11.5% out-of-sample cumulative futures hedge performance in 1986 for a US$100 million portfolio using three models: rolling OLS regression, Breeden–Giarla dynamic hedging and nonparametric regression.

2.4. Conditionally expected returns and volatility

Most would agree that there is predictable time-variation in measures of expected market returns. For example, Keim and Stambaugh (1986) and Fama and

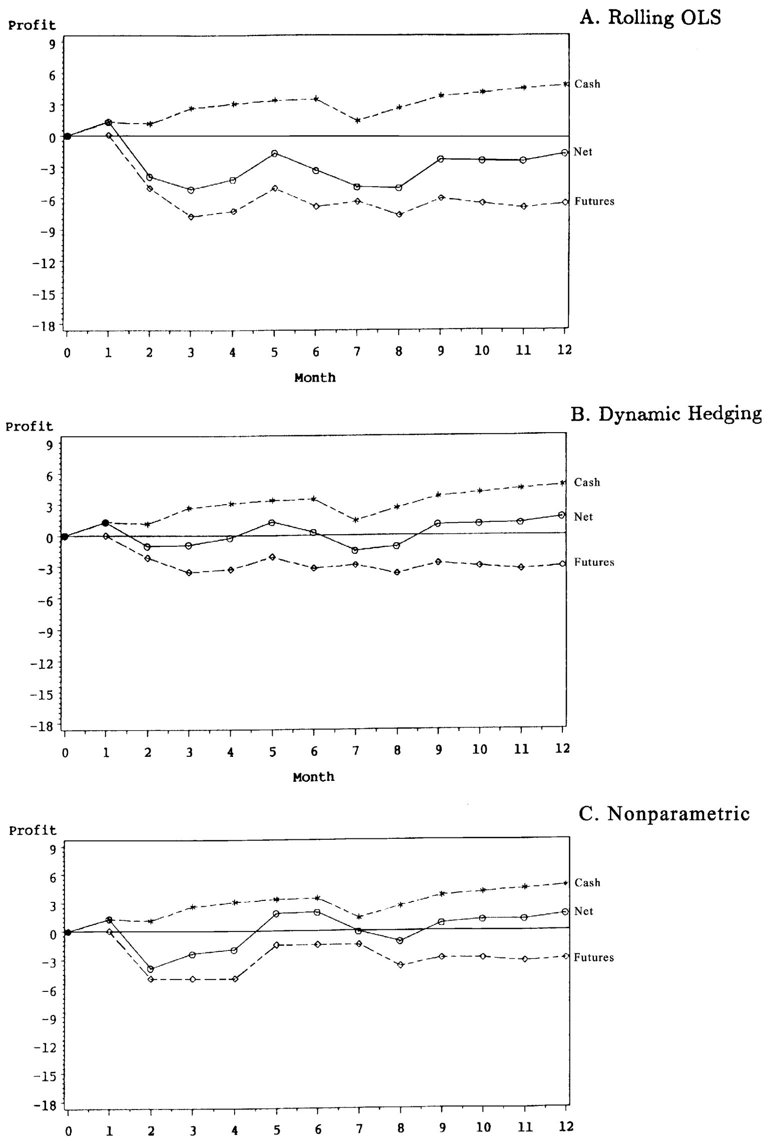


Fig. 9. GNMA 15% out-of-sample cumulative futures hedge performance in 1986 for a US$100 million portfolio using three models: rolling OLS regression, Breeden–Giarla dynamic hedging and nonparametric regression.

French (1988, 1989) find that a number of information variables are able to predict US stock returns. Harvey (1991) finds that a similar set of information variables are able to predict stock returns in many different countries. Even stronger is the agreement that conditional volatility changes through time. For example, the work of Officer (1973), Black (1976), Merton (1980), Christie (1982), French et al. (1987), Bollerstev et al. (1988), Schwert (1989a,b), Pagan and Schwert (1990), Gallant et al. (1992, 1993), Li (1990), Cao and Tsay (1992), Nelson (1991), and Harvey and Whaley (1991, 1992) provide convincing evidence that conditional volatility changes predictably through time.

There is considerable disagreement about the relation between the conditional mean and the conditional variance. French et al. (1987) "find evidence of a positive relation between the expected risk premium on common stocks and the predictable level of volatility." Using a different methodology, Pagan and Hong (in preparation) find "striking" evidence of a negative relation between the conditional mean and conditional volatility which is in "contrast to the positive effects claimed by ... French, Schwert and Stambaugh."

In addition, some asset pricing studies assume that the ratio of the conditional mean to variance is fixed. Merton (1980) shows the conditions whereby this ratio can be linked to the representative agent's relative risk aversion. Although there the evidence presented in Campbell (1987) and Harvey (1989, 1991) suggests that this ratio is time-varying, tests have proceeded with the constancy assumption. For a recent example, see Chan et al. (1992).

However, there is a complication in these analyses; the true conditional mean and the true conditional variance are unobservable. The finding of a negative or positive relation between the conditional mean and variance could be a result of the model for the means and variances being wrong. Similarly, evidence that the ratio of mean to variance changing through time could also be spuriously caused by incorrect modeling of the numerator and/or the denominator of the ratio. For example, consider the role of omitted information. Harvey (1989, 1991) examined the constancy of the ratio of conditional mean to the conditional variance of the market. Conditional on the true information set, $\boldsymbol{\Omega}_{t-1}$, this implies that

$$E\left[r_t|\Omega_{t-1}\right] = \lambda \mathrm{Var}\left[r_t|\Omega_{t-1}\right], \tag{15}$$

where λ is the constant of proportionality. Suppose the econometrician uses a subset of the information, $\boldsymbol{Z}_{t-1}$. A test of whether $E[r_t|\boldsymbol{Z}_{t-1}]$ is proportional to $\mathrm{Var}[r_t|\boldsymbol{Z}_{t-1}]$ could be misleading because

$$E\left[r_t|Z_{t-1}\right] = \lambda \mathrm{Var}\left[r_t|\mathbf{Z}_{t-1}\right] - \lambda \mathrm{Var}\left[E\left[r_t|\Omega_{t-1}\right]|\mathbf{Z}_{t-1}\right]. \tag{16}$$

The second term on the right-hand side of Eq. (16) is omitted and could complicate the inference.

Two forces could cause incorrect rejection of asset pricing restrictions: the omission of information and the misspecification of the expectation generating mechanism. It is important to conduct a sensitivity analysis of how conclusions are affected by the specification of conditional expectations. In this section, such a sensitivity analysis is conducted over the both the structure of the expectation generating mechanism and the information environment.

2.4.1. The conditional mean

The variable of interest is the excess return on the Center for Research in Security Prices (CRSP) value-weighted NYSE return in excess of the 30-day Treasury bill. The bill data are an updated version of the data used in Fama (1984) and are also available from CRSP. Some summary statistics on this market return and the information variables used in the analysis are presented in Table 5.

The information variables include: the spread in yields between Moody's Baa and Aaa rated bonds, the dividend yield on the S&P 500 in excess of the 30-day Treasury bill rate, and the excess holding period return on the 3-month Treasury bill. A number of studies have used information variables similar to these. For example, Keim and Stambaugh find that the Moody's 'junk' bond yield spread is able to predict equity returns. Fama and French (1988) document the predictive power of the dividend yield for longer horizon returns, however, they show that the dividend yield has little power to predict monthly returns. Harvey (1989, 1991) uses the dividend yield in excess of the 30-day Treasury bill rate. Such a definition makes this information variable look like the negative of the short-term interest rate. Fama and Schwert (1977) show that stock returns can be predicted with the nominal interest rate. The final variable is the excess return on the 3-month Treasury bill. Campbell (1987) shows that short-term premiums have the ability to predict stock returns.

Of course, there are many other variables that have been proposed to predict stock returns. The number of conditioning variables is kept small because the nonparametric technique will only perform well on low dimensional densities given the relatively small sample. One variable that is included in a number of studies that is not included in this one is the lagged market return. This variable does not appear to be important for the conditional mean. Linear regressions of the market return on its lag reveal no significant relation. Nonparametric regression also suggests that lagged returns do not significantly influence the mean. Conrad and Kaul (1989) argue that the last week of the month might be a more important instrument than the return over the entire month. No evidence was found to support this idea. When the return on the S&P 500 during the last week of the lagged month included as a regressor, it does not help explain any variation in the returns.

The first panel of Table 5 provides means, standard deviations, minimum, maximum and autocorrelations of the excess market return and the information variables. Consistent with other studies, there is little autocorrelation in the excess market return. The short-term premium has significant first-order autocorrelation. The excess dividend yield and the junk yield spread are far more persistent time series. Fig. 10 provides time-series plots of these information variables. All three

Table 5

The conditional mean of the value weighted NYSE return in excess of the one month bill rate, 1947:9–1988:12 (495 observations)

(A) Summary statistics

Variable	Mean	Std. dev.	Minimum	Maximum	ρ_1	ρ_2	ρ_3	ρ_4	ρ_{12}	ρ_{24}
Market excess return	0.00634	0.04188	−0.22171	0.16285	0.058	−0.035	0.013	0.043	0.055	−0.022
Excess three month bill return	0.00053	0.00106	−0.00305	0.00880	0.315	0.080	0.071	0.070	−0.011	0.056
Baa−Aaa yield spread	0.00079	0.00038	0.00027	0.00220	0.974	0.941	0.915	0.895	0.703	0.519
Excess S&P 500 dividend yield	−0.00034	0.00275	−0.00935	0.00490	0.973	0.952	0.931	0.913	0.826	0.721

(B) Correlations of the market return and instruments

Variable	*xvw*	l1p3	l1junk	l1xdiv
Market excess return	1.000	0.175	0.067	0.159
Excess three month bill return		1.000	0.393	−0.256
Baa−Aaa yield spread			1.000	−0.599
Excess S&P 500 dividend yield				1.000

(C) Conditional mean of the market return

Model	Instruments	$\bar{R}_2$
OLS	l1p3, l1junk, l1xdiv	0.089
OLS	l1junk, l1xdiv	0.062
Nonparametric-kernel	l1p3, l1junk, l1xdiv	0.027
Nonparametric-kernel	l1junk, l1xdiv	0.026
Nonparametric-Fourier	l1p3, l1junk, l1xdiv squares and trigonometric terms	0.084
Nonparametric-Fourier	l1junk, l1xdiv squares and trigonometric terms	0.053

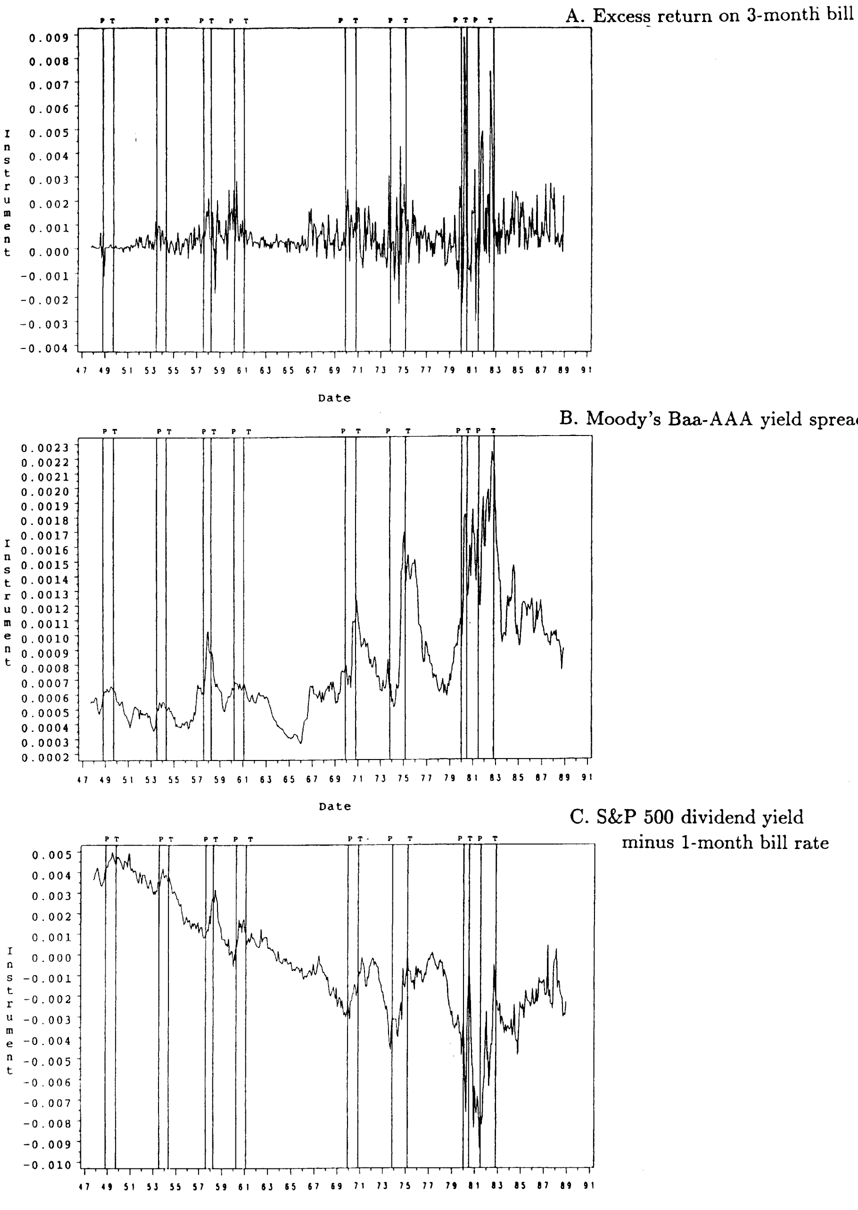


Fig. 10. Business-cycle variation in the instrumental variables.

variables have distinct business-cycle patterns. The level of the junk yield spread and the excess dividend move with the business cycle. The pattern in the short-term premium is different; its volatility, rather than level, has a business-cycle pattern.

A summary of the analysis of the conditional mean is presented in the final panel of Table 5. With all three information variables, the linear OLS model can explain 8.9% of the unconditional variance of returns from 1947:08 to 1988:12 (495 observations). Although not reported, all variables enter the regressions with positive coefficients that are reliably different from zero at standard significance levels. When the short-term premium is dropped as an information variable, the explanatory power drops by about one third.

The second part shows the results of the nonparametric density estimation. Using three information variables, the nonparametric predictions account for 2.7% of the unconditional variance. In contrast to the linear model, when the short-term premium is dropped, the explanatory power is virtually unaffected. As a result, in latter tables, only two information variables will be retained in the nonparametric density estimation.

Finally, Table 5 also presents the results of another nonparametric model: Gallant's (1981) Fourier Flexible Form. This method is a series expansion. The linear regression is augmented by the squares of the information variables plus some trigonometric terms. The trigonometric terms involve scaling the information variables to fall into the range of $(0, 2\pi)$, and adding the sine and cosine of the lagged transformed variables as well as the sine and cosine of twice the transformed variables. Interestingly, the addition of the squares and trigonometric terms does not improve the explanatory power.

Given the earlier results, the nonparametric regression should exhibit higher explanatory power if there is a nonlinear relation between the market return and the information variables. However, the results in Table 1 indicate that the explanatory power of the nonparametric density estimation method is less than the linear model. These results appear to support the specification of a linear conditional mean model.[19]

However, there is an important qualification. The criteria for model selection is usually R^2—the ratio of the unconditional variance of the fitted values to the unconditional variance of the actual market returns. Ideally, we would like a measure of how close the fitted values are to the true conditionally expected returns. Given that the true conditionally expected returns are unobservable, such a metric is not possible.

The difference between the nonparametric and the linear OLS R^2 is obvious from Fig. 11 which plots the two models' fitted values. Both series exhibit similar

[19] Using a different test, Hollifield (1991) also shows that there is little evidence of nonlinearity in equity portfolio returns.

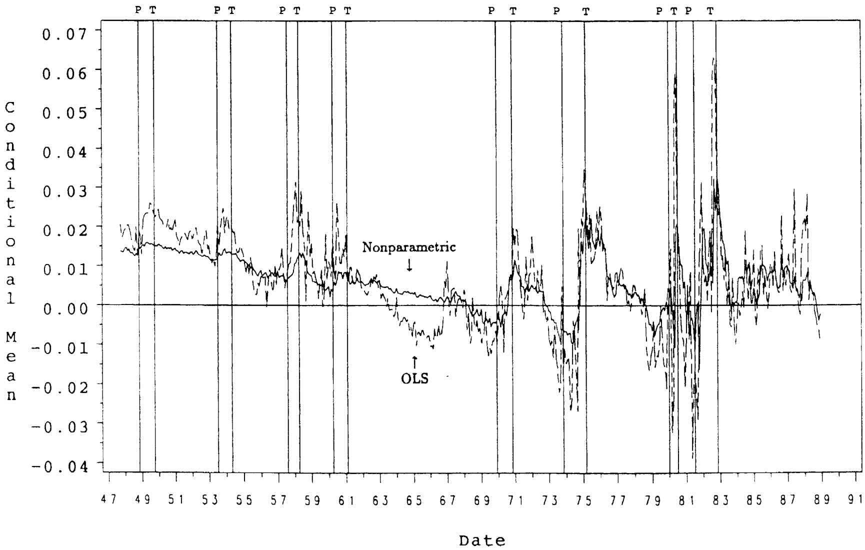


Fig. 11. A comparison of the conditional mean excess return on the value-weighted NYSE index using two different expectation models. The OLS model is based on a linear regression of the excess market return on three instrumental variables: the excess return on the 3-month bill, the Moody's Baa–Aaa yield spread, and the S&P 500 dividend yield minus the 1-month bill rate. The nonparametric-kernel model is conditioning on the Moody's Baa–Aaa yield spread and the S&P 500 dividend yield minus 1-month bill rate.

business-cycle patterns: expected returns are highest at business-cycle troughs and lowest just before peaks. However, the nonparametric fitted values have a smaller range and are much smoother than the OLS fitted values. This smoothness reduces the unconditional variance of the fitted values and accounts for the lower R^2.

It is not clear that the nonparametric fitted values are less reasonable than the OLS fitted values. In particular, the OLS fitted values may be *too* variable to be plausible. For example, is it reasonable to believe that investors expected 5.8%, 6.3% and 6.1% monthly excess returns in August through October 1982? Is it reasonable that the conditionally expected monthly return was -4.0% and -2.0% in May and June of 1981?

Another troubling aspect about the OLS fitted values is the number of negative expected excess returns. This implies that investors expected a lower return on the stock market than the conditionally riskless Treasury bill return. The OLS model suggests that the conditionally expected excess return is negative 144 of 495 months (29%). The average conditionally expected return in these months is -0.84% and the standard deviation of the expected returns is 0.70%. Between June 1963 and October 1966 almost all the expected returns are negative.

Negative expected excess returns are not inconsistent with asset pricing theory. Investor may be willing to purchase assets that have expected returns that are less than the risk-free rate if the asset will provide a hedge of their consumption. Asness (1991) measures the time-varying consumption betas of the quarterly value-weighted NYSE return and finds that there are period of negative conditional consumption betas. These periods are brief and coincide with the business cycle peaks of December 1969, November 1973, January 1980, and July 1981. However, the OLS model suggests that expected excess returns are negative during many more episodes.

The nonparametric conditional expectations are negative 82 of 495 months (17%). The negative values coincide with the four episodes of negative consumption betas. Asness (1991) also reports that the magnitude of negative conditional betas match the magnitude of the nonparametric conditionally expected returns. That is, the consumption betas become more negative moving from the 1969 peak to the 1981 peak. The same pattern is found in the nonparametric conditionally expected returns.

The nonparametric analysis of the conditional mean does not reveal any obvious nonlinearities in the relation between the market return and a set of pre-determined information variables. This result is reassuring for researchers that have specified linear conditional expectations in asset pricing tests. However, the superior fit of the OLS model might be a result of data snooping rather than a closer approximation to the true conditional expectation. The nonparametric conditional expectation appears to be a reasonable alternative. The sensitivity of inferences should be assessed by using both the linear and the nonparametric conditional expectations.

2.4.2. The conditional variance

The true conditional variance, like the conditional mean, is unobservable. I will present nine different models of the conditional variance: four parametric models and five nonparametric models. The focus will be on both the structure of the expectations and the role of the conditioning information.

To start, a definition of the conditional mean is necessary. Many studies have not paid much attention to the conditional mean. For example, French et al. (1987) let the mean return follow a first-order moving average process. Pagan and Schwert (1990) project the monthly stock returns on seasonal dummies and let the mean return follow a first-order moving average process. Pagan and Hong (in preparation) condition on the lagged market return—which has limited or no ability to track time-variation in the expected returns. Indeed, Braun et al. (1990) conjecture that "ignoring conditional means should have a minor effect in estimating the second moment matrix." It is potentially interesting to assess the role of the conditional mean in the conditional variance estimator.

For each variance estimator, some diagnostic experiments are provided in Table 6. In the first panel, the following regression is run

$$\hat{e}_t^2 = \left(xvw_t - E\left[xvw_{t-1}^{\mathrm{OLS}} \right] \right)^2 = \alpha_i + \beta_i \hat{\sigma}_{it}^2 + u_{it}, \tag{17}$$

where $\hat{e}_t$ is the deviation from OLS conditional mean and $\hat{\sigma}_i^2$ is the model i variance estimator. The β coefficient represents the bias. An adjusted R^2 measure is also reported as well as a tests of whether u_{it} or $\hat{e}_{it}^2 - \hat{\sigma}_{it}^2$ are correlated to the financial instruments.

The conditional expectation of $\hat{e}_{it}^2$ is the measure of conditional variance. The next two panels of Table 5 change the definition of conditional variance by changing the definition of the conditional mean. In the second panel, the left-hand-side variable is the squared deviation from the nonparametric conditional mean. The third panel shows regressions using squared deviation from the unconditional mean return. These exercises provide some information on the relation between the fitted variance estimator and different definitions of conditional variance.

The two variance models falls into the class of two-step variance estimator proposed by Davidian and Carroll (1987). In the first step, deviations from the conditional mean are calculated with the OLS model in Table 5. The second step involves regressing the squared deviations are regressed on some information variables. The fitted values are measures of the conditional variance. These models will be referred to a 'OLS' models. The first specification follows Pagan and Schwert (1990). The squared innovations are regressed on eight lags of the squared deviation. However, Pagan and Schwert's (1990) model for the conditional mean is different. They allow only allow for a first-order moving average process in the conditional mean.

The conditioning information is altered in the second OLS variance model. Instead of eight lags of the squared innovation in the conditional mean, only one lag is used along with the three financial variables used in the estimation of the conditional mean.

The next set of models are based on the nonparametric density estimation technique and will be referred to a 'nonparametric-kernel'. The first model is a mixture of parametric and nonparametric models. The squared deviation from the conditional mean model (parametric) is forecasted using the nonparametric Eq. (6). Following Pagan and Schwert (1990), only one conditioning variable is used, the lagged innovation in the conditional mean.

The second nonparametric model is purely nonparametric. Eq. (6) is used to forecast both the market return and the squared market return. The expectations are conditioned on two financial variables: the junk yield spread and excess dividend yield. The conditional variance is formed by Eq. (10).

The third set of models use Gallant's (1981) Fourier Flexible Form and will be referred to as 'nonparametric-Fourier'. Three versions are presented. In each of these, the dependent variable is the square of the conditional mean from the OLS estimation. The first conditions on the lagged innovation in the conditional mean, the lagged squared innovation and the trigonometric terms. The second version uses two lags of the innovation in the conditional mean. The final version uses one lag of the innovation in the conditional mean plus the three financial instruments.

Table 6
The conditional variance of the value weighted NYSE return, 1947:9–1988:12 (495 observations)

Conditional variance	Conditioning information	β_i	$\bar{R}_2$	χ^2 u_{it} on information	χ^2 $\hat{e}_t^2 - \sigma_{it}^2$ on information
(A) $\hat{e}_t^2 = (xvw_t - E_{t-1}^{OLS}[xvw_t])^2 = \alpha_i + \beta_i \hat{\sigma}_{it}^2 + u_{it}$					
OLS	$\hat{e}_{t-1}^2, \ldots, \hat{e}_{t-8}^2$	1.000 (0.374)	0.015	6.747 [0.080]	6.747 [0.080]
OLS	$\hat{e}_{t-1}^2$, l1p3, l1junk, l1xdiv	1.000 (0.276)	0.019	0.000 [1.000]	0.000 [1.000]
Nonparametric-kernel	$\hat{e}_{t-1}$	1.003 (0.314)	0.013	9.331 [0.025]	9.331 [0.025]
Nonparametric-kernel	l1junk, l1xdiv	0.936 (0.374)	0.005	1.030 [0.794]	0.780 [0.854]
Nonparametric-Fourier	$\hat{e}_{t-1}$, square and trigonometric terms	1.000 (0.210)	.039	8.846 [0.031]	8.846 [0.031]
Nonparametric-Fourier	$\hat{e}_{t-1}$, $\hat{e}_{t-2}$, squares and trigonometric terms	1.000 (0.190)	0.062	7.387 [0.064]	7.387 [0.064]
Nonparametric-Fourier	$\hat{e}_{t-1}$, l1p3, l1junk, l1xdiv squares and trigonometric terms	1.000 (0.216)	0.089	0.000 [1.000]	0.000 [1.000]
French, Schwert and Stambaugh	monthly standard deviation of daily S&P 500 returns	0.548 (0.165)	0.050	5.221 [0.156]	12.923 [0.005]
EGARCH	$\hat{e}_{t-1}^2$, $\hat{e}_{t-1}/\hat{\sigma}_{t-1}$	0.484 (0.169)	0.029	4.288 [0.232]	11.108 [0.011]
(B) $\hat{\epsilon}_t^2 = (xvw_t - E_{t-1}^{Kernel}[xvw_t])^2 = \alpha_i + \beta_{\hat{\sigma}\,it}^2 + u_{it}$					
OLS	$\hat{e}_{t-1}^2, \ldots, \hat{e}_{t-8}^2$	1.122 (0.418)	0.015	9.676 [0.022]	9.942 [0.019]
OLS	$\hat{e}_{t-1}^2$, l1p3, l1junk, l1xdiv	1.194 (0.305)	0.022	0.293 [0.961]	0.887 [0.829]

Nonparametric-kernel	$\hat{e}_{t-1}$	1.307 (0.388)	0.019	12.862 [0.005]	12.786 [0.005]
Nonparametric-kernel	l1junk, l1xdiv	1.193 (0.414)	0.007	1.757 [0.624]	2.767 [0.429]
Nonparametric-Fourier	$\hat{e}_{t-1}$, square and trigonometric terms	1.117 (0.290)	0.039	12.182 [0.007]	12.246 [0.007]
Nonparametric-Fourier	$\hat{e}_{t-1}$, $\hat{e}_{t-2}$, squares and trigonometric terms	1.160 (0.274)	0.068	10.659 [0.014]	10.990 [0.012]
Nonparametric-Fourier	$\hat{e}_{t-1}$, l1p3, l1junk, l1xdiv squares and trigonometric terms	1.156 (0.315)	0.096	0.388 [0.943]	0.919 [0.840]
French, Schwert and Stambaugh	monthly standard deviation of daily S&P 500 returns	0.676 (0.221)	0.062	5.314 [0.150]	7.437 [0.059]
EGARCH	$\hat{e}^2_{t-1}$, $\hat{e}_{t-1}/\hat{\sigma}_{t-1}$	0.629 (0.221)	0.041	4.687 [0.196]	5.866 [0.118]
(C) $\hat{\eta}_t^2 = (xvw_t - E[xvw_t])^2 = \alpha_i + \beta^2_{\hat{\sigma}\,it} + u_{it}$					
OLS	$\hat{e}^2_{t-1}, \ldots, \hat{e}^2_{t-8}$	1.107 (0.402)	0.014	9.062 [0.028]	9.292 [0.026]
OLS	$\hat{e}^2_{t-1}$, l1p3, l1junk, l1xdiv	1.207 (0.294)	0.021	0.110 [0.991]	0.699 [0.873]
Nonparametric-kernel	$\hat{e}_{t-1}$	1.351 (0.381)	0.019	11.770 [0.008]	11.733 [0.008]
Nonparametric-kernel	l1junk, l1xdiv	1.163 (0.421)	0.006	1.516 [0.679]	2.440 [0.486]
Nonparametric-Fourier	$\hat{e}_{t-1}$, square and trigonometric terms	1.163 (0.269)	0.040	11.120 [0.011]	11.222 [0.011]
Nonparametric-Fourier	$\hat{e}_{t-1}$, $\hat{e}_{t-2}$, squares and trigonometric terms	1.123 (0.251)	0.060	9.697 [0.021]	9.950 [0.019]
Nonparametric-Fourier	$\hat{e}_{t-1}$, l1p3, l1junk, l1xdiv squares and trigonometric terms	1.168 (0.280)	0.093	0.181 [0.981]	0.709 [0.871]
French, Schwert and Stambaugh	monthly standard deviation of daily S&P 500 returns	0.700 (0.199)	0.063	4.554 [0.207]	5.919 [0.116]
EGARCH	$\hat{e}^2_{t-1}$, $\hat{e}_{t-1}/\hat{\sigma}_{t-1}$	0.619 (0.204)	0.037	4.119 [0.249]	4.510 [0.211]

Heteroskedasticity consistent standard errors in parentheses. The chi-square statistic has three degrees of freedom. *P*-values are in brackets.

The French et al. (1987) variance estimator is also investigated. This estimator is based on the monthly variance of the daily S&P 500 returns:

$$\sigma_t^2 = \sum_{i=1}^{N_t} r_{it}^2 + \sum_{i=1}^{N_t-1} r_{it} r_{i+1,t}, \tag{18}$$

where there are N_t daily returns, r_{it} in month t. French et al. fit a third-order moving average process to the first differences in the log of the market standard deviation. Using data from 1947–1989, I estimated the same time-series model to obtain conditional variances.[20] This is a model of the conditional variance of the S&P 500 returns. Unfortunately, the value-weighted NYSE is not available on a daily basis before 1962.

The final model considered is the an EGARCH model proposed by Nelson (1991). The EGARCH model allows for volatility to respond asymmetrically to positive or negative market movements. The model's intuition follows the observation that market volatility tends to rise following negative returns and fall after positive returns—the leverage effect. The EGARCH variance series was obtained from Braun et al. (1990) and represents the fitted values over the 1928–1988 period.[21]

Some summary statistics on the variance models are presented in Table 6.[22] In the first panel, the squared deviation from the OLS conditional mean are regressed on the fitted variance estimators. The ability of the models to explain the variance of the squared innovation is limited. The R^2 values range from 0.5%(nonparametric-kernel with financial conditioning variables) to 8.9% (nonparametric-Fourier with financial conditioning variables).

The addition of the financial conditioning information generally improves the explanatory power of the models. This is evident from the χ^2 test which checks to see if the errors are correlated with the financial information. The disturbances from Eq. (14) of four of the models that do not use the financial conditioning information are related to this conditioning information. The variance errors of six of the models that do not use the conditioning information are related to the financial instruments. Indeed, the highest explanatory power come from the nonparametric-Fourier model that includes the financial information.

[20] The three moving-average parameters (standard errors) are: 0.571 (0.044), 0.056 (0.051) and 0.146 (0.044). There was no significant autocorrelation in the model residuals. The conditional variance is calculated as $\sigma_t^2 = \exp[2\ \ln\hat{\sigma}_t + 2\mathrm{Var}(u_t)]$ where u are the model residuals.

[21] I thank Phillip Braun, Dan Nelson and Alain Sunier for making this data available to me. It is important to note that their model is fit to a longer time series 1928–1988. All of the other methods were fit to data from 1947.

[22] There was one negative conditional variance in the OLS variance that uses eight lags of the squared innovation in the conditional mean. There were 10 negative variances in the nonparametric-Fourier with the financial conditioning information.

The final two panels show the effect of changing the definition of variance for the left-hand-side of Eq. (14). Panel B considers the squares of the deviations from the nonparametric conditional mean and panel C examines deviations from the unconditional mean. Both the explanatory power of the regressions and the bias parameters change as a result of the different definition of the conditional mean. This suggests that the conditional mean may be important for the conditional variance.

I investigate the role of the conditional mean for each of the variance estimators with the decomposition in Eq. (10). If the same information is used for both the mean and variance estimation, then the conditional variance is just the conditionally expected squared return minus the square of the conditionally expected return. If the conditioning information is different, then we can write the unconditional variance of the conditional variance as

$$\begin{aligned}
&\mathrm{Var}\Big[E\Big[\big(r_t - E[r_t|\mathrm{OLS}]\big)^2|\mathbf{Z}_{t-1}\Big]\Big] \\
&\quad = \mathrm{Var}\big[E\big[r_t^2|\mathbf{Z}_{t-1}\big]\big] - \mathrm{Var}\big[E\big[2r_t E[r_t|\mathrm{OLS}]|\mathbf{Z}_{t-1}\big]\big] \\
&\qquad + \mathrm{Var}\Big[E\Big[\big(E[r_t||\mathrm{OLS}]\big)^2|\mathbf{Z}_{t-1}\Big]\Big] + \mathrm{Covariances}, \qquad (19)
\end{aligned}$$

where OLS represents the conditioning information used for the mean and $\boldsymbol{Z}$ is the conditioning information used in the variance estimation. If the conditional mean does not matter, then the variance of the fitted conditional variance should look like the variance of the expected squared returns.

The unconditional variances presented in Table 7 suggest that the conditional mean does matter. This is particularly evident for the variance models that use the financial conditioning information (which is also used to get the conditional mean). For example, in the nonparametric-Fourier model with the full information set, the unconditional variance of the fitted conditional variance is 8.670. The variance of the expected squared returns is 13.647. This alone suggests that there must be a large covariance term between the square of conditional mean and the conditionally expected squared return. However, in the models which do not use the information that the mean is conditioned on, the role of the conditional mean is diminished. This is particularly evident for the OLS variance model that conditions only on lagged squared innovations in the mean and the nonparametric model that conditions only on the lagged innovation. Interestingly, the conditional mean appears to play a role in both the EGARCH and French et al. (1987) estimator.

2.4.3. The relation between mean and variance

Given the estimators of the conditional variance, we can explore the relation between the excess return on the market portfolio and the conditional variance. Merton (1980) shows that if the representative agent have logarithmic utility or if consumption growth follows an i.i.d. process, then the coefficient of proportional-

Table 7
Does the conditional mean matter in the conditional variance? 1947:9–1988:12 (495 observations)

Model	Information $\boldsymbol{Z}$	Unconditional variance of			
		$E[(r - E[r \mid \mathrm{OLS}])^2 \mid \boldsymbol{Z}]$	$E[r^2 \mid \boldsymbol{Z}]$	$E[2rE[r \mid \mathrm{OLS}] \mid \boldsymbol{Z}]$	$E[(E[r \mid \mathrm{OLS}])^2 \mid \boldsymbol{Z}]$
OLS	$\hat{e}^2_{t-1}, \ldots, \hat{e}^2_{t-8}$	1.656	2.358	0.496	0.049
OLS	$\hat{e}^2_{t-1}$, l1p3, l1junk, l1xdiv	1.967	2.590	1.899	0.598
Nonparametric-kernel	$\hat{e}_{t-1}$	1.450	2.838	–	0.001
Nonparametric-kernel	l1junk, l1xdiv	0.750	0.674	–	0.142
Nonparametric-Fourier	$\hat{e}_{t-1}$, square and trig. terms	3.882	5.571	0.662	0.013
Nonparametric-Fourier	$\hat{e}_{t-1}$, $\hat{e}_{t-2}$, squares and trig. terms	6.086	8.298	1.039	0.045
Nonparametric-Fourier	$\hat{e}_{t-1}$, l1p3, l1junk, l1xdiv, squares and trig. terms	8.670	13.647	6.791	1.297
French, Schwert and Stambaugh	monthly std. dev. of daily S&P 500	4.976	8.447	0.932	0.082
EGARCH	$\hat{e}^2_{t-1}$, $\hat{e}_{t-1} / \hat{\sigma}_{t-1}$	2.969	5.367	0.738	0.071

All variances are multiplied by 10 000 000.

Table 8
The relation between the conditional mean and conditional variance of the value-weighted NYSE return, 1947:9–1988:12 (495 observations)

(A) $xvw_t = \beta_i \hat{\sigma}_{it}^2 + u_{it}$

Conditional variance	Conditioning information	β_i	$\bar{R}^2$ u_{it} on information
OLS	$\hat{e}_{t-1}^2, \ldots, \hat{e}_{t-8}^2$	3.825 (1.266)	0.088
OLS	$\hat{e}_{t-1}^2$, l1p3, l1junk, l1xdiv	2.889 (1.271)	0.094
Nonparametric-kernel	$\hat{e}_{t-1}$	4.085 (1.254)	0.091
Nonparametric-kernel	l1junk, l1xdiv	3.330 (1.142)	0.095
Nonparametric-Fourier	$\hat{e}_{t-1}$, square and trigonometric terms	3.274 (1.366)	0.090
Nonparametric-Fourier	$\hat{e}_{t-1}$, $\hat{e}_{t-2}$, squares and trigonometric terms	3.078 (1.383)	0.088
Nonparametric-Fourier	$\hat{e}_{t-1}$, l1p3, l1junk, l1xdiv squares and trigonometric terms	2.310 (1.438)	0.093
French, Schwert and Stambaugh	monthly standard deviation of daily S&P 500 returns	2.679 (1.238)	0.083
EGARCH	$\hat{e}_{t-1}^2$, $\hat{e}_{t-1}/\hat{\sigma}_{t-1}$	2.931 (0.985)	0.084

(B) $xvw_t = \alpha_i + \beta_i \hat{\sigma}_{it}^2 + u_{it}$

Model	Conditioning variables	α_i	β_i	$\bar{R}^2$	$\bar{R}^2$ u_{it} on information
OLS	$\hat{e}_{t-1}^2, \ldots, \hat{e}_{t-8}^2$	0.005 (0.010)	1.136 (6.638)	−0.002	0.089
OLS	$\hat{e}_{t-1}^2$, l1p3, l1junk, l1xdiv	0.024 (0.006)	−11.322 (4.232)	0.012	0.078
Nonparametric-kernel	$\hat{e}_{t-1}$	−0.000 (0.008)	4.122 (5.123)	−0.001	0.091
Nonparametric-kernel	l1junk, l1xdiv	0.033 (0.011)	−15.653 (6.605)	0.008	0.073
Nonparametric-Fourier	$\hat{e}_{t-1}$, square and trigonometric terms	0.009 (0.006)	−1.439 (4.005)	−0.002	0.089
Nonparametric-Fourier	$\hat{e}_{t-1}$, $\hat{e}_{t-2}$, squares and trigonometric terms	0.008 (0.005)	−0.737 (3.353)	−0.002	0.089
Nonparametric-Fourier	$\hat{e}_{t-1}$, l1p3, l1junk, l1xdiv squares and trigonometric terms	0.010 (0.005)	−2.593 (3.254)	0.001	0.085
French, Schwert and Stambaugh	monthly standard deviation of daily S&P 500 returns	0.004 (0.004)	1.035 (2.325)	−0.001	0.087
EGARCH	$\hat{e}_{t-1}^2$, $\hat{e}_{t-1}/\hat{\sigma}_{t-1}$	0.001 (0.005)	2.497 (2.329)	0.002	0.084

Heteroskedasticity consistent standard errors in parentheses.

ity between the conditional mean and the conditional variance can be interpreted as relative risk aversion. However, these are strong assumptions. Under more general assumptions, Backus and Gregory (1989) and Campbell (1993) show that this coefficient does not necessarily represent relative risk aversion and in principal can be positive or negative.

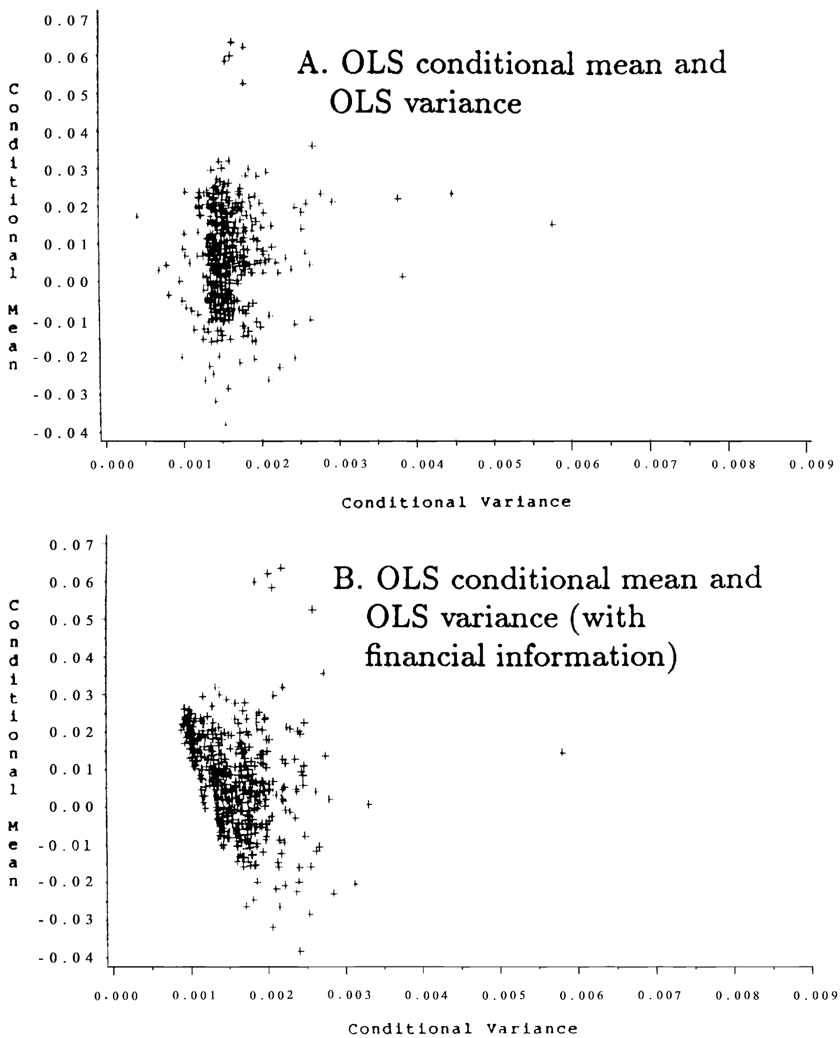
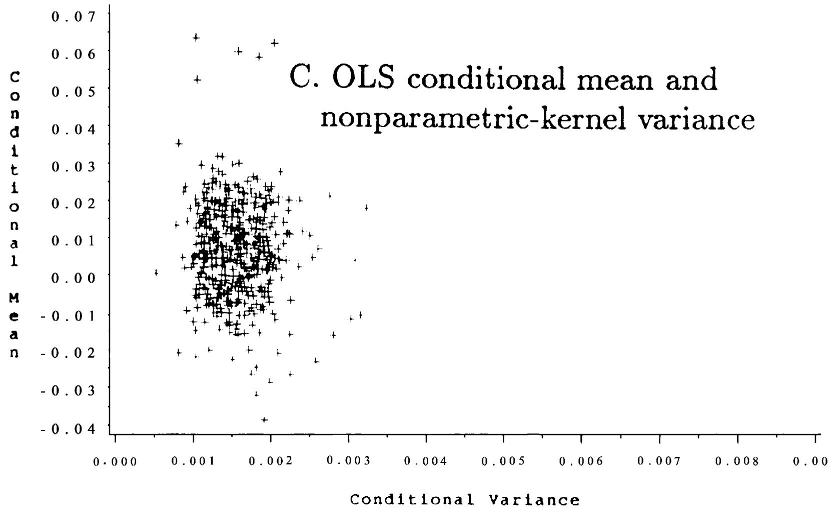


Fig. 12. The relation between the conditional mean and the conditional variance of NYSE value-weighted return, 1947–1988.



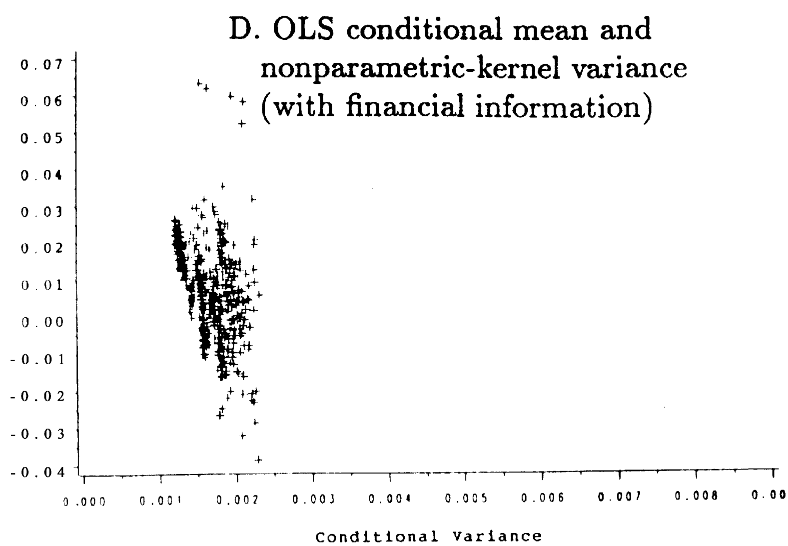


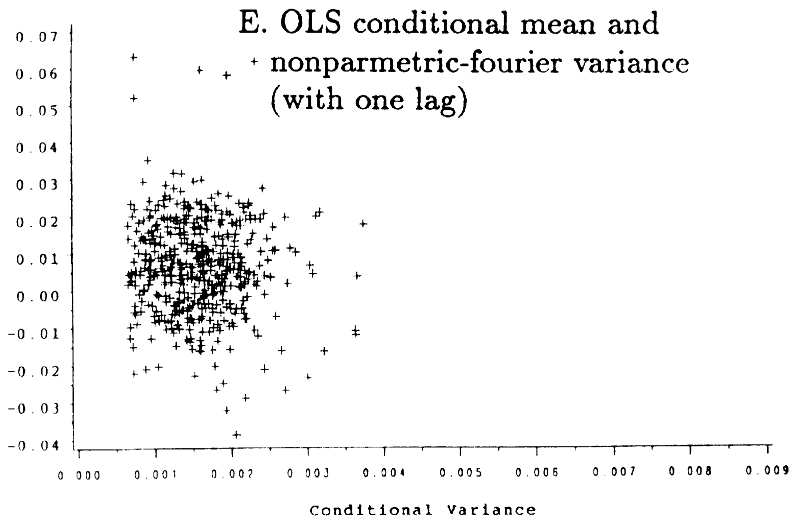
Fig. 12 (*continued*).

Consider two models

$$xvw_t = \beta_i \hat{\sigma}_{it}^2 + u_{it}, \tag{20}$$

$$xvw_t = \alpha_i + \beta_i \hat{\sigma}_{it}^2 + u_{it}. \tag{21}$$

The first model is designed to test whether the mean is proportional to the variance. The second model tests whether the mean is linear in the variance. The



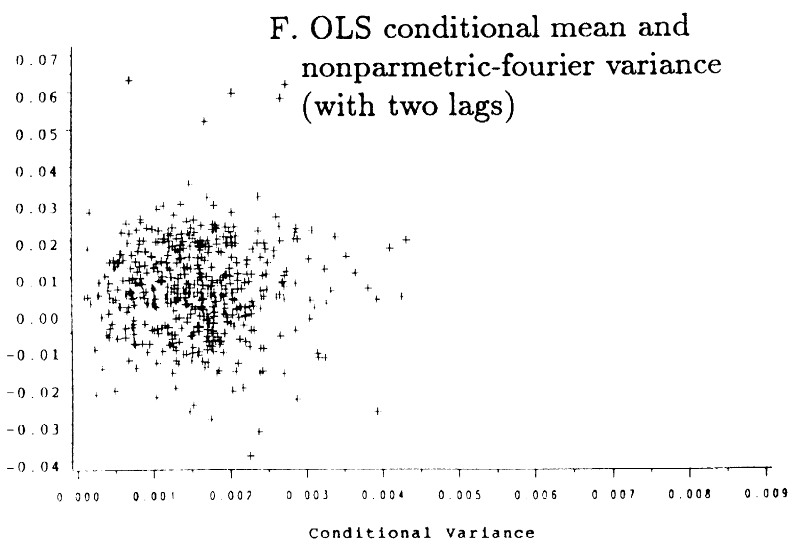
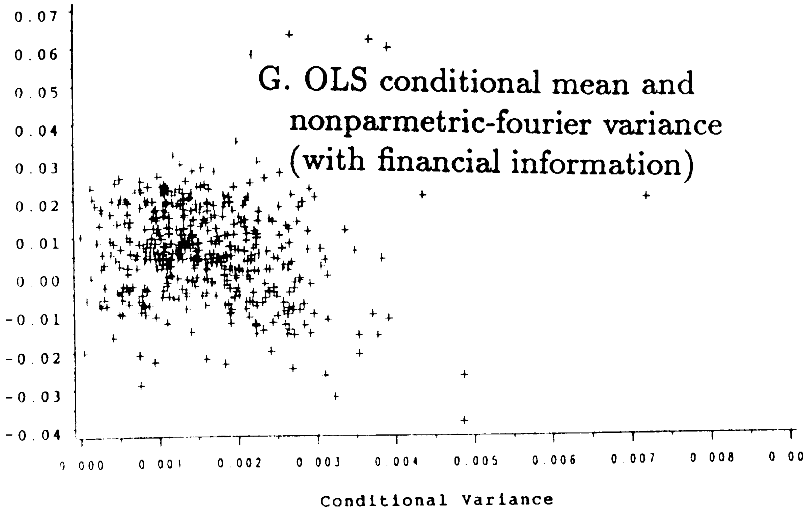


Fig. 12 (*continued*).

methodology involves estimating the parameters of these models and, given additional conditioning information, testing the overidentifying restrictions. In essence, the test checks to see whether the innovations, u_{it}, are correlated with the conditioning information.

Proportionality tests (Eq. (20)) using a linear model for the conditional variance are presented in Campbell (1987) and Harvey (1989) for US stock returns. These



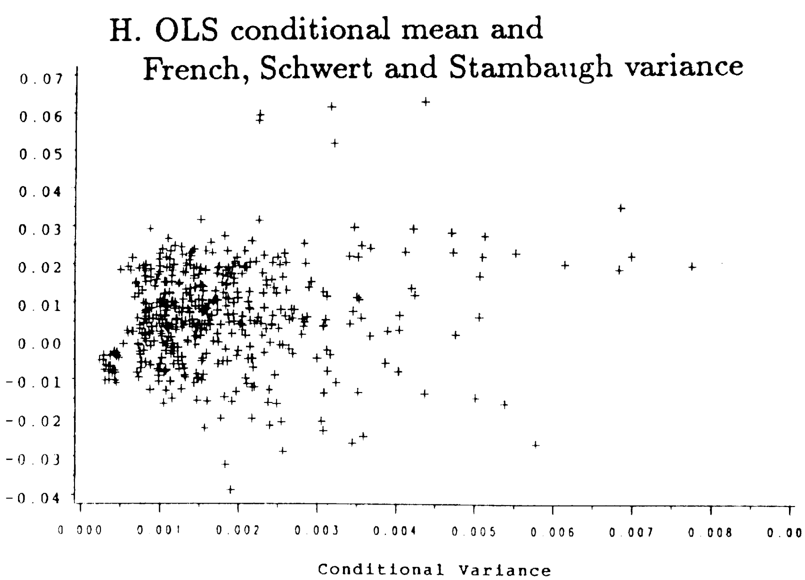


Fig. 12 (*continued*).

tests strongly reject proportionality. French et al. (1987) and Chou (1988) estimate Eq. (21) and find a positive linear relation between the mean and variance. Glosten et al. (1994), Pagan and Hong (in preparation), and Nelson (1991) find a negative relation. However, only the Glosten et al. (1994) study tests the overidentifying restrictions implied by Eq. (21).

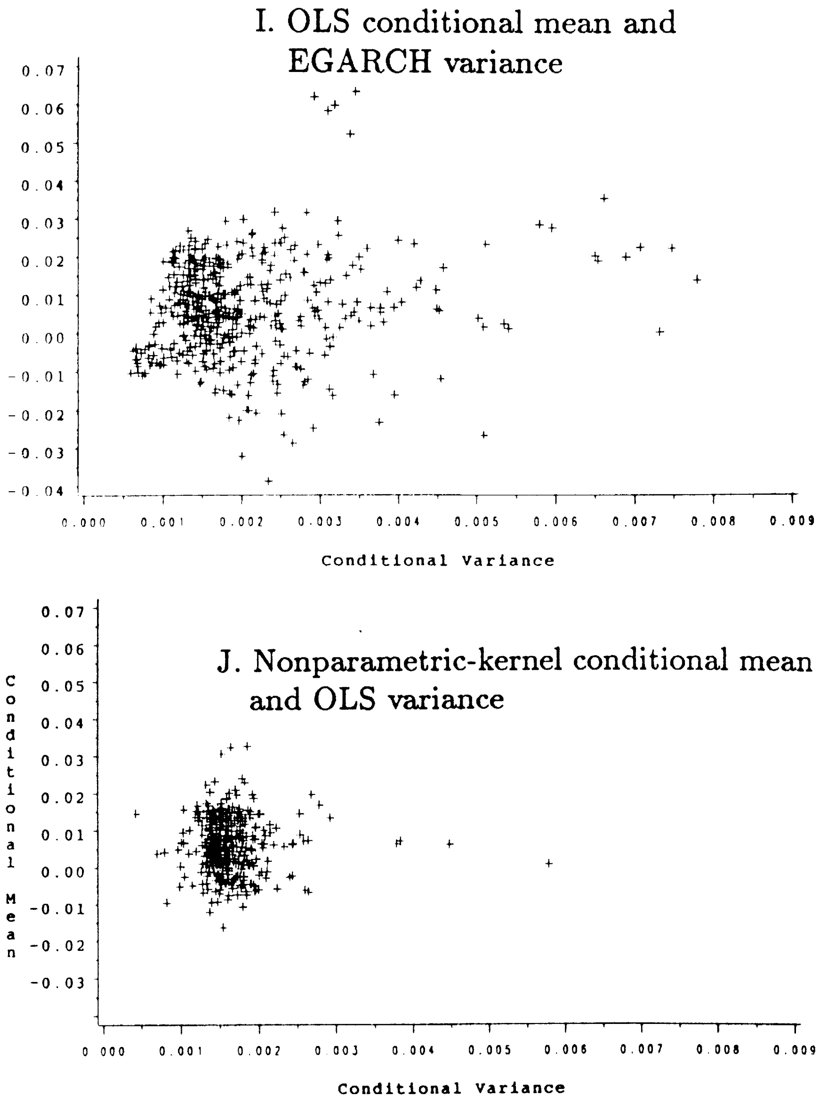


Fig. 12 (*continued*).

The first panel of Table 8 presents the proportionality tests implied by Eq. (20). The coefficient of proportionality is positive because the average excess return and the average conditional variance are both positive. Consistent with the results of Campbell (1987) and Harvey (1989), a χ^2 test of the proportionality restriction is rejected at the 0.001 probability level for all nine variance estimators. Instead of

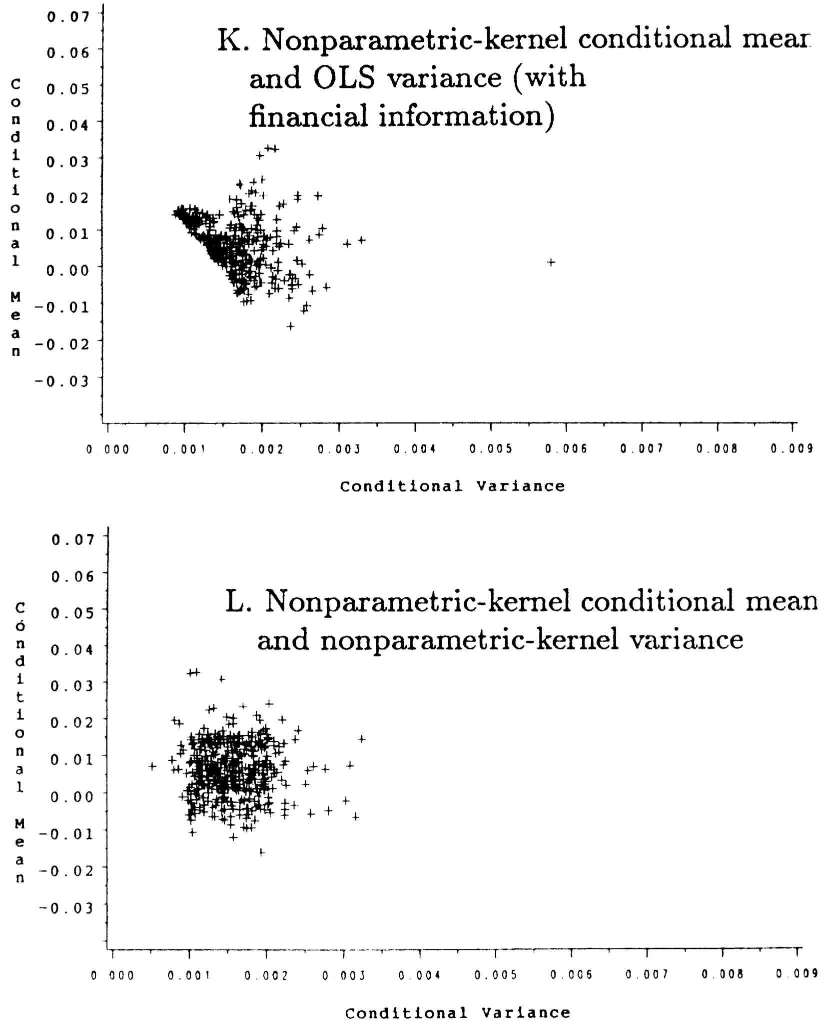
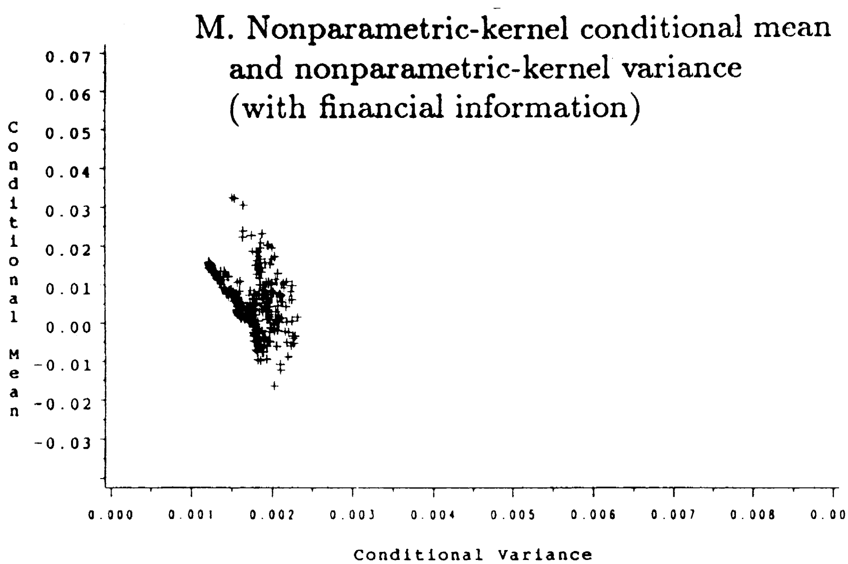


Fig. 12 (*continued*).

presenting the χ^2 statistic, I have presented the R^2 of the model residual on the information variables. The R^2 is significant in each case and closely resembles the R^2 in the conditional mean regression. This suggests that the ratio of mean to variance is time-varying.

The second panel provides tests of the linearity restriction. This restriction is rejected at the 0.001 probability level for each variance estimator. Similar to the results in the first panel, the R^2 of the disturbance on the financial innovation is of the same size as the R^2 of the conditional mean regression.



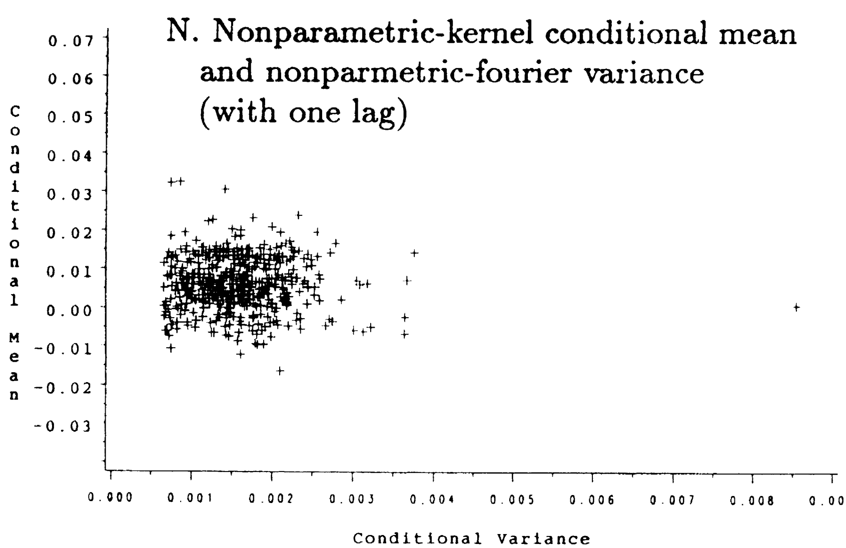
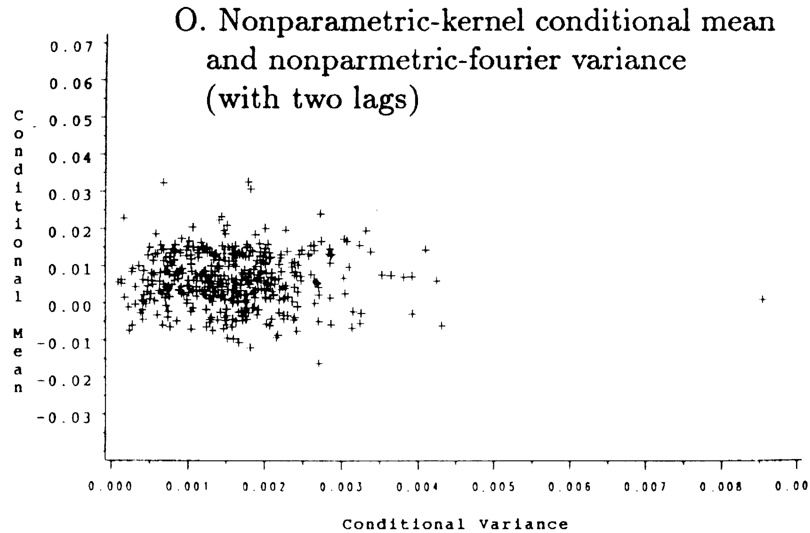


Fig. 12 (*continued*).

The sign of slope coefficient depends upon the variance model being used. The linear OLS and the nonparametric-kernel models that use no financial conditioning information produce positive, insignificant, coefficients. Both the French et al. (1987) estimator and the EGARCH model also deliver positive coefficients. These



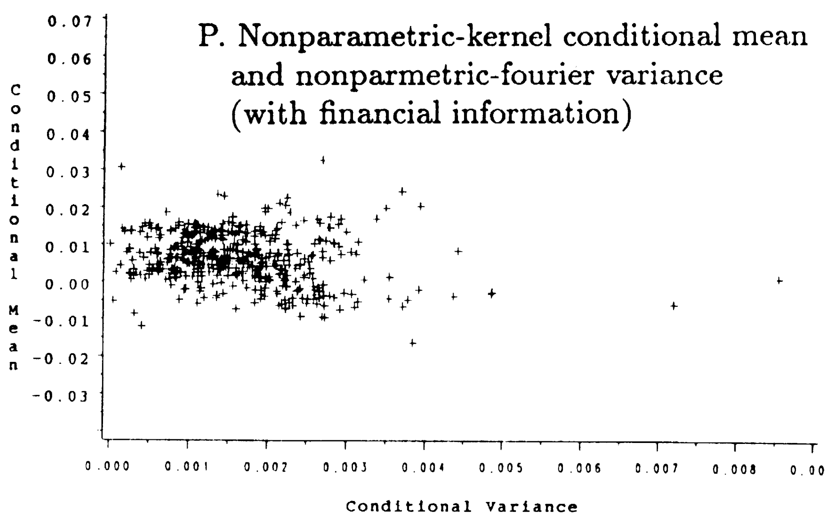
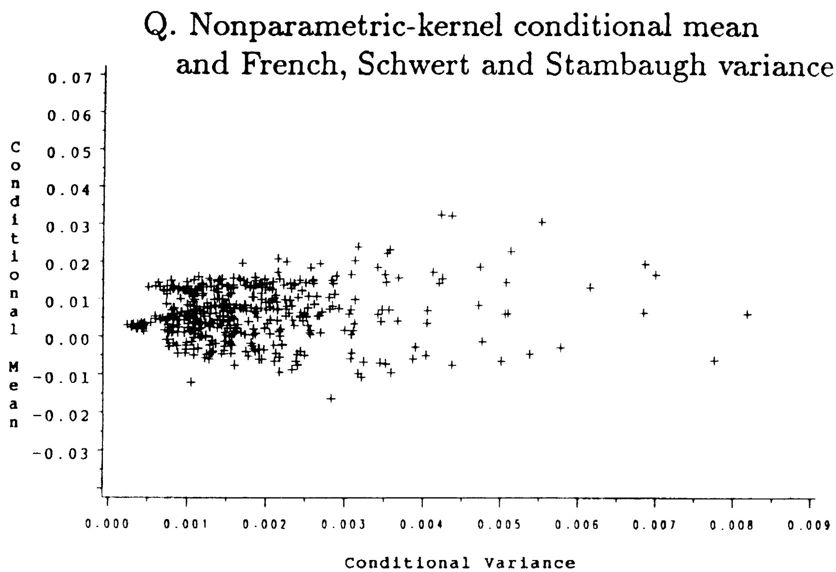


Fig. 12 (*continued*).

models also do not explicitly use the financial conditioning information. The linear OLS and the nonparametric-kernel models that explicitly take into account the conditioning information deliver significant negative coefficients. However, all the coefficient estimates should be interpreted with caution given that this linear specification is statistically rejected. In addition, even when the coefficients appear



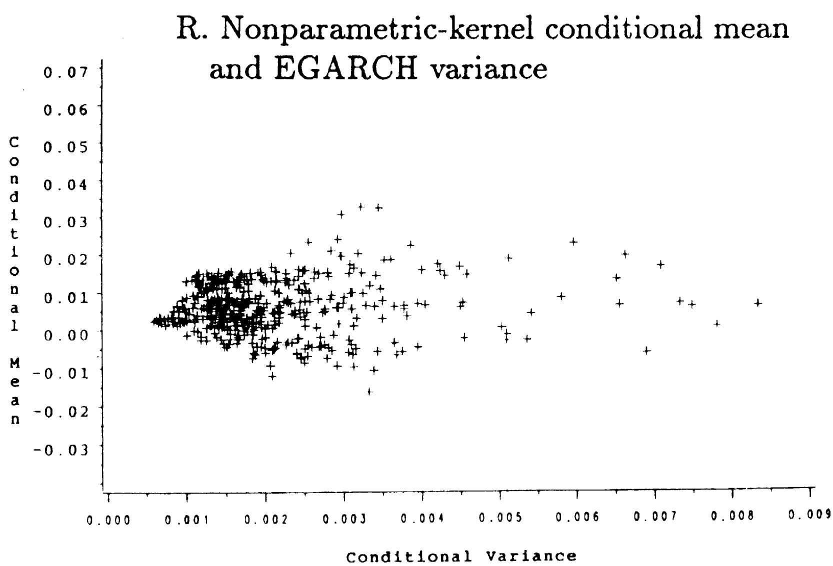
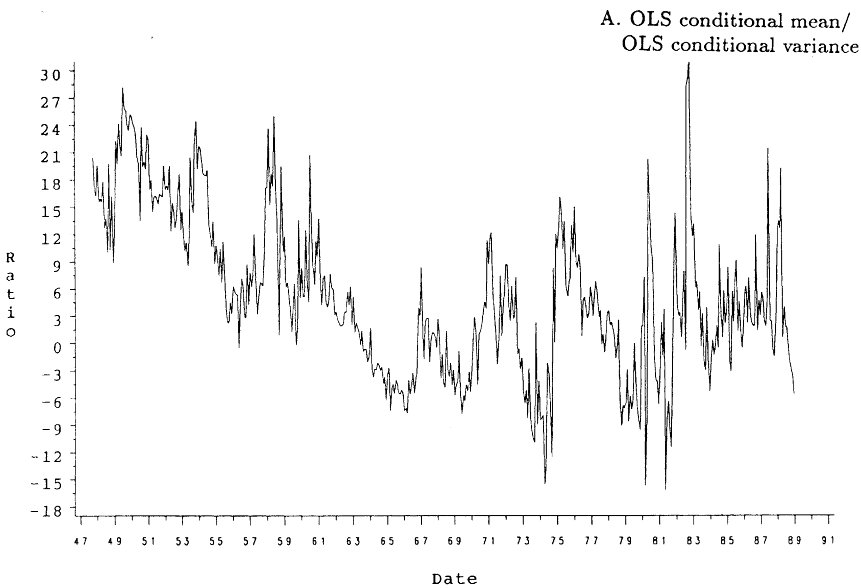


Fig. 12 (*continued*).

statistically significant, the conditional variance estimator only explains a small portion of the variation in the conditional mean.

These models involve regressions of the actual market excess return on the estimated variance. The R^2 statistics reveal that the variance has a limited ability



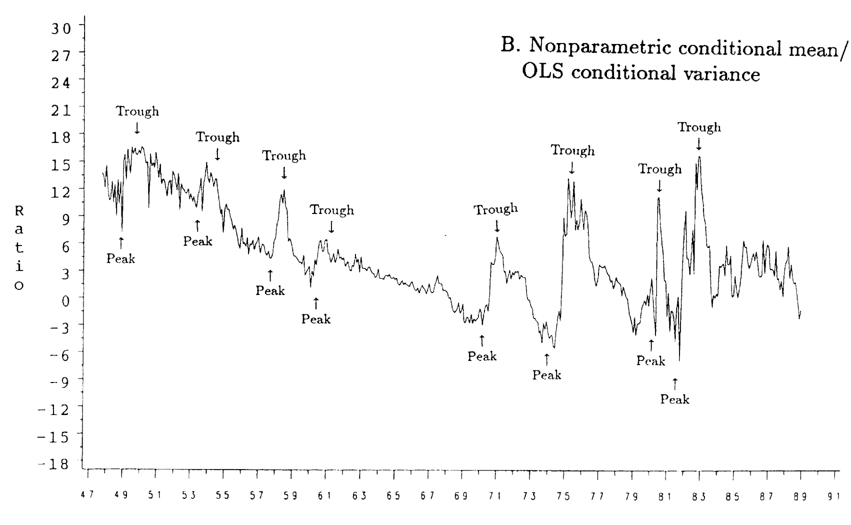


Fig. 13. Time-series variation in the ratio of the conditional mean to the conditional variance of the NYSE value-weighted return, 1947–1988.

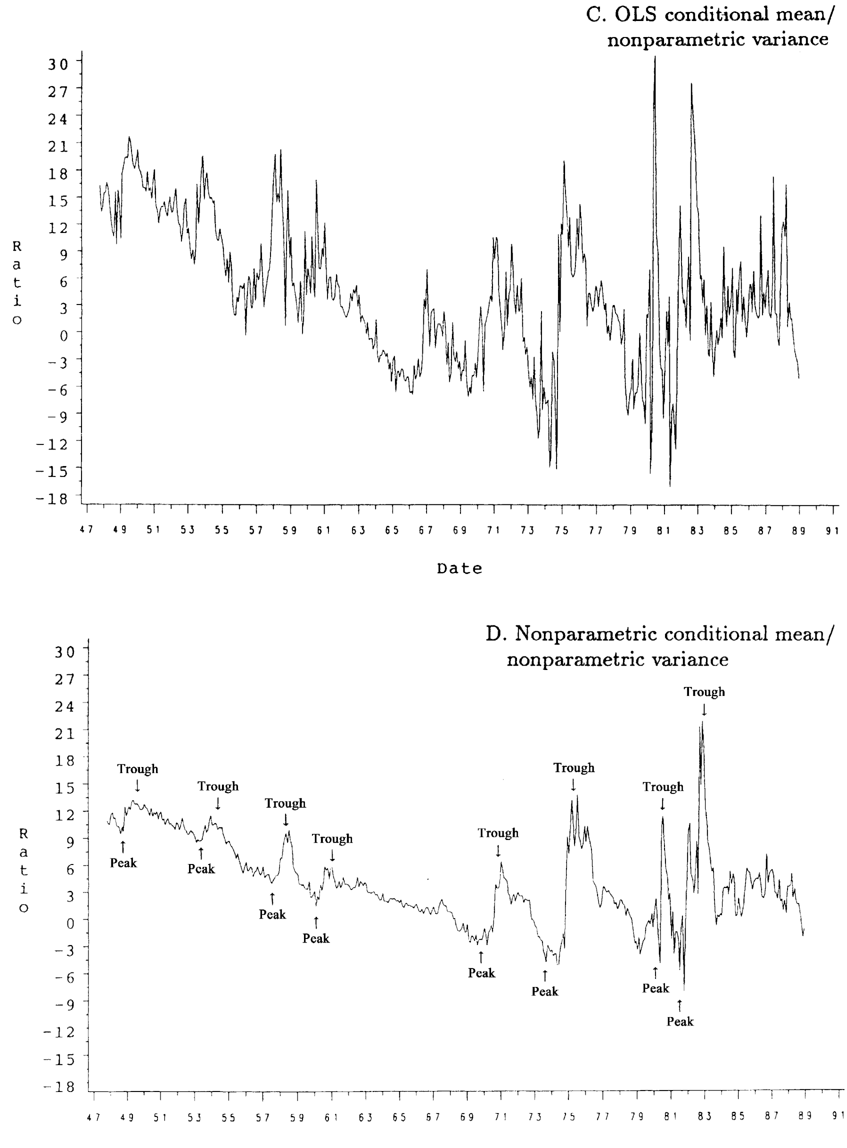
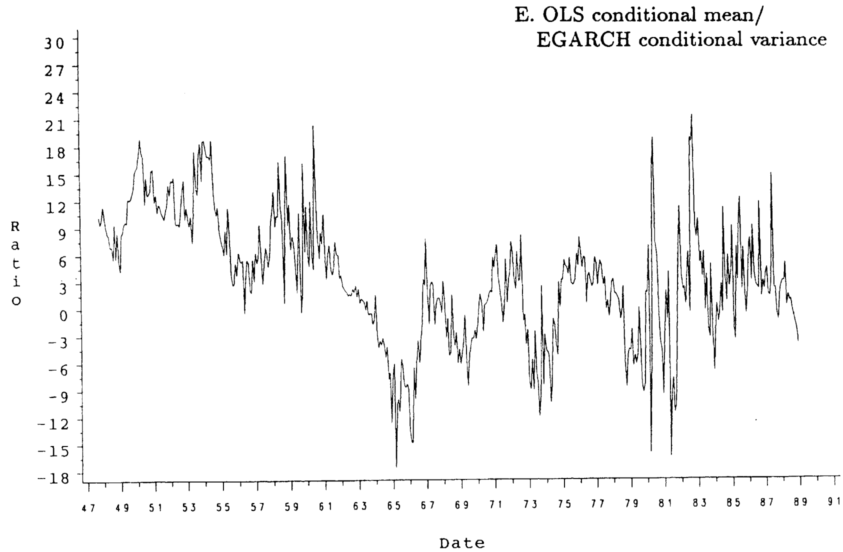


Fig. 13 (*continued*).

to explain the conditional mean. Fig. 12 directly examines the relation between the conditional mean and the conditional variance using the different mean and estimators.



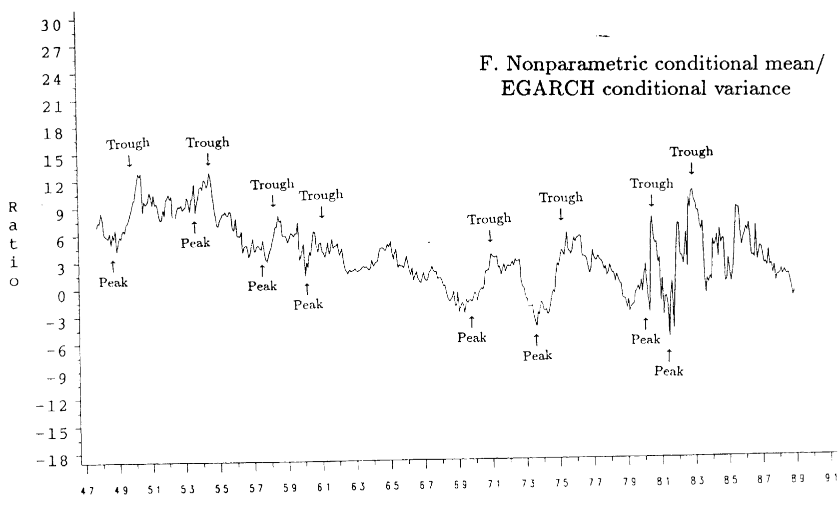


Fig. 13 (*continued*).

For most of the graphs, there is no obvious relation between the conditional mean and the conditional variance. For example, the French et al. and the EGARCH variance could be positively or negatively related to expected returns. However, the variance estimators that include the conditioning information show a distinctly negative relation between the conditional mean and the conditional variance. These results suggest that the specification of the conditional variance influences the inference about the relation between mean and variance.

But Fig. 13 does not reveal any information about the relation between mean and variance through time. Indeed, Table 8 suggests that there could be business-cycle patterns in the relation between the conditional mean and variance in that the residuals show the same level of correlation with the financial instrument a the conditional mean. These results suggest that a linear regression of returns on fitted volatility is not that well motivated and this is the reason that the overidentifying restrictions test suggests that the model is misspecified. Campbell (1987), Harvey (1989) and Kandel and Stambaugh (1990) have argued that the relation between conditional mean and variance changes through time. My results suggest that these results are robust to different definitions of conditional variances and means.

The time-series variation in the relation between the conditional mean and conditional variance can be viewed directly by plotting the ratio of the fitted values. Fig. 13 provides six different versions of this ratio—three conditional variance models and two conditional mean models. Given the earlier results on the conditional mean, it is no surprise that the ratio is much more volatile when the OLS conditional mean is used in the numerator.

Each ratio, OLS, nonparametric-kernel and EGARCH exhibits similar patterns. The nonparametric ratio has the most striking business-cycle patterns.[23] The ratio is low near business-cycle peaks and high near business-cycle troughs. In terms of a conditional CAPM, the expected compensation for market volatility depends on the stage of the business cycle. Investors require more expected return per unit of volatility at business-cycle troughs to invest in the market. At business cycle peaks, investors are willing to purchase equities with the knowledge that the expected return per unit of variance is small.

3. Conclusions

The goal of this paper was to explore the specification of conditional expectations. Sufficient conditions were developed whereby expectations would linear in the conditioning information with constant projection coefficients. Under more general conditions, conditional expectations can be obtained with nonparametric density estimation techniques. Given enough data, this method provides an empirical way of estimating the multivariate density. With this density, conditional expectations can be calculated.

Given that so many applications in finance rely on properly specifying conditional expectations, it seems reasonable to undertake a sensitivity analysis. The

[23] The business-cycle dating follows the NBER. These dates are established ex post. It would also be interesting to establish ex ante datings.

analysis in this paper considers both the expectation generating mechanism and the role of the conditioning information. The nonparametric technique is useful in both detecting deviations from linear conditional expectations and helping us understand the nature of potential nonlinearities.

There are numerous situations where nonparametric regression could be used. This paper focuses on four empirical applications. In the first, the returns of a market timer are simulated. The market timer allocates into market sensitive stocks in up markets and into market insensitive stock in down markets. This will create a nonlinearity in the relation between the portfolio return and the market return. The nonparametric density estimation technique has the ability to replicate the nonlinearity. This is impressive given that the technique has very little structure.

In the second application, the market model regression is examined for the recreation industry portfolio. The returns of this portfolio appear to be nonlinearly related to the market return. The results of the nonparametric regressions suggest that the nature of this nonlinearity is complicated. The nonparametric density estimation technique may be a useful diagnostic tool for detecting and potentially modeling complicated nonlinearities.

In the third example, nonparametric density estimation is used to create mimicking portfolios for hedging securities. The response function is used to create out-of-sample hedge positions for various Government National Mortgage Association pass-through securities. These instruments have embedded options which make them nonlinear. The nonparametric technique shows some ability to replicate the nonlinearity induced by the options.

Finally, conditionally expected returns and variances of the excess return on the NYSE value-weighted portfolio are examined. An analysis of the conditional mean reveals that the highest R^2 is obtained with a linear-rather than nonparametric model. This indicates that the linear model may be a reasonable approximation for conditional expectations in asset pricing research. However, almost 30% of the conditionally expected excess returns from the linear specification are negative. This suggests that the data might have been overfit.

Nine different conditional variance estimators are examined. This allows us to assess the sensitivity of the variance estimator to the structure of the variance and the conditioning information. Empirical results suggest that the conditional mean should not be ignored when estimating conditional variances.

The relation between the conditional mean and the conditional variance generally depends on the conditioning information used in the variance estimation. Evidence is presented that suggests a negative relation between the conditional mean and the conditional variance if the variance estimator uses the same conditioning information as the mean. Parametric and nonparametric analysis of the data revealed a distinct business-cycle pattern in the ratio of the conditional mean to the conditional variance. Investors appear to require a large expected return per unit of volatility at business-cycle troughs and a low expected return per unit of volatility around business-cycle peaks.

Acknowledgements

I have benefited from the comments of Daniel Beneish, Marshall Blume, Doug Breeden, Wayne Ferson, Doug Foster, Mike Giarla, Mike Hemler, Ravi Jagannathan, Dan Nelson, Adrian Pagan, Tom Smith, Rob Stambaugh, S. Viswanathan, Aman Ullah and Guofu Zhou and from seminar participants at Duke University and the University of Chicago. Boaz Schwartz and Guofu Zhou provided excellent research assistance. I especially thank Peter Rossi, Mike Gibbons, and an anonymous referee for their detailed suggestions.

References

Asness, C., 1991. Negative Expected Returns and the Business Cycle. Working paper, Graduate School of Business, University of Chicago.
Backus, D.K., Gregory, A.W., 1989. Theoretical Relations Between Risk Premiums and Conditional Variances. Working paper, Federal Reserve Bank of Minneapolis, December.
Ball, R., Kothari, S.P., 1989. Nonstationary expected returns: implications for tests of market efficiency and serial correlation in returns. Journal of Financial Economics 25, 51–74.
Bierens, H.J., 1987. Kernel estimators of regression functions. In: Bewley, T.F. (Ed.), 1987. Advances in Econometrics 5th World Congress, vol. I, Cambridge Univ. Press, Cambridge, pp. 99–144.
Black, F., 1976. Studies of stock market volatility changes. Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section, pp. 177–181.
Black, F., Jensen, M., Scholes, M., 1972. The capital asset pricing model: some empirical results. In: Jensen, M. (Ed.), Studies in the Theory of Capital Markets. Praeger, New York.
Blake, I.F., Thomas, J.B., 1968. On class of processes arising in linear estimation theory. IEEE Transactions on Information Theory IT-14, 12–16.
Bollerstev, T., Engle, R.F., Wooldridge, J.M., 1988. A capital asset pricing model with time varying covariances. Journal of Political Economy 96, 116–131.
Braun, P., Nelson, D.B., Sunier, A.M., 1990. Good News, Bad News, Volatility and Betas. Journal of Finance 50, 1575–1603.
Breeden, D.T., 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. Journal of Financial Economics 7, 265–296.
Breeden, D.T., Giarla, M.J., 1989. Hedging interest rate risks with futures, swaps and options. In: Fabozzi, F.J. (Ed.), The Handbook of Mortgage Backed Securities.
Breeden, D.T., Gibbons, M.R., Litzenberger, R.H., 1989. Empirical tests of the consumption-oriented CAPM. Journal of Finance 44, 231–262.
Campbell, J.Y., 1987. Stock returns and the term structure. Journal of Financial Economics 18, 373–400.
Campbell, J.Y., 1993. Intertemporal Asset Pricing Without Consumption. American Economic Review 83, 487–512.
Cao, Q.C., Tsay, R.S., 1992. Nonlinear time series analysis of stock return volatility. Journal of Applied Econometrics 7, 165–185.
Chamberlain, G., 1983. A characterization of the distributions that imply mean-variance utility functions. Journal of Economic Theory 29, 185–201.
Chan, K.C., 1988. On the contrarian investment strategy. Journal of Business 61, 147–163.
Chan, K.C., Karolyi, G.A., Stulz, R.M., 1992. Global Financial Markets and the Risk Premium on US Equity. Journal of Financial Economics 32 (2), 137–168.

Chopra, N., Lakonishok, J., Ritter, J.R., 1992. "Measuring abnormal performance: Do stocks overreact?". Journal of Financial Economics 31 (2), 235–268.

Chou, R.Y., 1988. Volatility persistence and stock valuations: some empirical evidence using GARCH. Journal of Applied Econometrics 3, 279–294.

Christie, A.A., 1982. The stochastic behavior of common stock variances: value, leverage and interest rate effects. Journal of Financial Economics 10, 407–432.

Chu, K.C., 1973. Estimation and decision for linear systems with elliptically random processes. IEEE Transactions on Automatic Control AC-18, 499–505.

Davidian, M., Carroll, R.J., 1987. Variance function estimation. Journal of American Statistical Association 82, 1079–1091.

Debondt, W.F.M., Thaler, R.M., 1985. Does the stock market overreact? Journal of Finance 40, 793–805.

Devlin, S.J., Gnanadesikan, R.G., Kettenring, J.R., 1976. Some multivariate applications of elliptical distributions. In: Ideka, S. (Ed.), Essays in Probability and Statistics. Shinko Tsusho, Tokyo, pp. 365–393.

Diebold, F.X., Nason, J.M., 1989. Nonparametric Exchange Rate Prediction? Working paper, Board of Governors, Federal Reserve Bank, Washington.

Dybvig, P.H., Ross, S.A. et al., 1985. Differential information and performance measurement using security market line. Journal of Finance 40, 383–400.

Fama, E.F., 1984. The information in the term structure. Journal of Financial Economics 13, 509–528.

Fama, E.F., French, K.R., 1988. Dividend yields and expected stock returns. Journal of Financial Economics 22, 3–26.

Fama, E.F., French, K.R., 1989. Business conditions and expected returns on stocks and bonds. Journal of Financial Economics 25, 23–50.

Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: empirical tests. Journal of Political Economy 81, 607–636.

Fama, E.F., Schwert, G.W., 1977. Asset returns and inflation. Journal of Financial Economics 5, 115–146.

Ferson, W.E., Harvey, C.R., 1991. The variation of economic risk premiums. Journal of Political Economy 99, 385–415.

Foster, F.D., Viswanathan, S., 1993. The effect of public information and competition on trading volume and price volatility. Review of Financial Studies 6, 23–56.

French, K.R., Schwert, G.W., Stambaugh, R.F., 1987. Expected stock returns and volatility. Journal of Financial Economics 19, 3–30.

Gallant, A.R., 1981. On the bias in flexible functional forms and an essentially unbiased form: the Fourier flexible form. Journal of Econometrics 15, 211–224.

Gallant, A.R., Tauchen, G.E., 1989. Seminonparametric estimation of conditionally constrained heterogeneous processes. Econometrica 57, 1091–1120.

Gallant, A.R., Hsieh, D.A., Tauchen, G.E., 1989. On fitting recalcitrant series: the pound–dollar exchange rate, 1974–83. In: Barnett, W.A. (Ed.), Nonparametric and Seminonparametric Models in Econometrics and Statistics: Proceedings of the Fifth International Symposium in Economic Theory and Econometrics. Cambridge Univ. Press, Cambridge.

Gallant, A.R., Hansen, L.P., Tauchen, G.E., 1990. Using conditional moments of asset payoffs to infer the volatility of intertemporal marginal rates of substitution. Journal of Econometrics 45, 141–185.

Gallant, A.R., Rossi, P., Tauchen, G.E., 1992. Stock prices and volume. Review of Financial Studies 5 (2), 199–242.

Gallant, A.R., Rossi, P., Tauchen, G.E., 1993. Nonlinear dynamic structures. Econometrica 61, 871–907.

Gibbons, M.R., 1982. Multivariate tests of financial models: a new approach. Journal of Financial Economics 10, 3–27.

Gibbons, M.R., Ferson, W.E., 1985. Tests of asset pricing models with changing expectations and an unobservable market portfolio. Journal of Financial Economics 14, 217–236.

Gibbons, M.R., Ross, S.A., Shanken, J., 1989. A test of the efficiency of given portfolio. Econometrica 57, 1121–1152.

Glosten, L.R., Jagannathan, R., Runkle, D., 1994. On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance.

Hansen, L.P., Richard, S.F., 1987. The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. Econometrica 55, 587–613.

Harvey, C.R., 1989. Time-varying conditional covariances in tests of asset pricing models. Journal of Financial Economics 24, 289–317.

Harvey, C.R., 1991. The world price of covariance risk. Journal of Finance 46, 111–157.

Harvey, C.R., Whaley, R.E., 1991. S&P 100 index option volatility. Journal of Finance 46, 1551–1561.

Harvey, C.R., Whaley, R.E., 1992. Market Volatility Prediction and the Efficiency of S&P 100 Index Option Market. Journal of Finance 31, 43–73.

Harvey, C.R., Zhou, G., 1990. Bayesian inference in asset pricing tests. Journal of Financial Economics 26, 221–254.

Hasbrouck, J., 1986. A note on linear heteroskedasticity models. Economic Letters, 349–351.

Hodrick, R.J., 1992. Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement. Review of Financial Studies 5, 357–386.

Hollifield, B., 1991. Linear Asset Pricing with Predictable Time-Varying Betas and Risk Premia: A General Equilibrium Approach. Working paper, Graduate School of Industrial Administration, Carnegie-Mellon University.

Ingersoll, J., 1987. Theory of Financial Decision Making. Rowan and Littlefield, Totawa, NJ.

Jensen, M.C., 1969. Risk, the pricing of capital assets, and the evaluation of investment portfolios. Journal of Business 42, 167–247.

Kandel, S., Stambaugh, R.F., 1990. Expectations and volatility of consumption and asset returns. Review of Financial Studies 3, 207–232.

Keim, D.B., Stambaugh, R.F., 1986. Predicting returns in the bond and stock market. Journal of Financial Economics 17, 357–390.

Kelker, D., 1970. Distribution theory of spherical distributions and location-scale parameter generalization. Sankhyā, Series A, 419–430.

Li, Y., 1990. Expected Stock Returns, Risk Premiums, and Volatilities of Economic Factors. Journal of Empirical Finance 5 (2), 69–97.

MacKinlay, A.C., 1987. On multivariate tests of the capital asset pricing model. Journal of Financial Economics 18, 341–372.

Merton, R.C., 1980. On estimating the expected return on the market: an exploratory investigation. Journal of Financial Economics 8, 323–362.

Nadaraya, E.A., 1964. On estimating regression. Theory of Probability and Its Applications 9, 142–143.

Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59, 347–370.

Officer, R.R., 1973. The variability of the market factor of the NYSE. Journal of Business 46, 434–453.

Pagan, A., Hong, Y., 1989. Nonparametric estimation and the risk premium. In: Barnett, W.A., Powell, J.A., Tauchen, G.E. (Eds.), Nonparametric and Seminon-Parametric Models in Econometrics and Statistics: The 5th International Symposium in Economic Theory and Econometrics. Cambridge Univ. Press, Cambridge, pp. 51–75, ch. 2.

Pagan, A., Schwert, G.W., 1990. Alternative models for conditional stock volatility. Journal of Econometrics 45, 267–290.

Pagan, A., Ullah, A., 1988. The econometric analysis of models with risk terms. Journal of Applied Econometrics 3, 87–105.

Robinson, P.M., 1983. Nonparametric estimators for time series. Journal of Time Series Analysis 4, 185–207.

Robinson, P.M., 1988. Semiparametric econometrics: a survey. Journal of Applied Econometrics 3, 35–51.
Schwert, G.W., 1989a. Why does stock market volatility change over time? Journal of Finance 44, 1115–1154.
Schwert, G.W., 1989b. Business cycles, financial crises and stock volatility. Carnegie-Rochester Conference Series on Public Policy 31, 83–126.
Shanken, J., 1985. Multivariate tests of the zero-beta CAPM. Journal of Financial Economics 14, 327–348.
Shanken, J., 1990. Intertemporal asset pricing: an empirical investigation. Journal of Econometrics 45, 99–120.
Silverman, B.W., 1986. Density Estimation for Statistics and Data Analysis. Chapman & Hall, London.
Stambaugh, R.F., 1982. On the exclusion of assets from tests of the two parameter model: a sensitivity analysis. Journal of Financial Economics 10, 237–268.
Stanton, R., 1990. Rational Prepayment and the Valuation of Mortgage-Backed Securities. Review of Financial Studies 8, 677–708.
Stone, M., 1974. Cross-validatory choice and assessment of statistical predictions. Journal of the Royal Statistical Society B 36, 111–133.
Ullah, A., 1988. Nonparametric estimation of econometric functionals. Canadian Journal of Economics 21, 625–658.
Vershik, A.M., 1964. Some characteristics properties of Gaussian stochastic processes. Theory of Probability and Its Applications 9, 353–356.
Watson, G.S., 1964. Smooth regression analysis. Sankhyā, Series A 26, 359–372.