Valuing Information Using Utility Functions: How Much Should We Pay for Linear Factor Models?

Soosung Hwang[1] Steve. E. Satchell

Cass Business School, London. Cambridge University, UK

July 2003

[1]Corresponding Author. Faculty of Finance, Cass Business School, 106 Bunhill Row, London, EC1Y 8TZ, UK. Tel: 44 (0)20 7040 0109. Fax: 44 (0)20 7040 8881. Email: s.hwang@city.ac.uk. The authors thank K. French for helpful comments and allowing us to use the Fama-French factor data.

Abstract

We investigates what is an appropriate level of investment management fees. We extend existing results and provide a several formula for the case of power utility and normal returns. Using the CRRA utility function with the range of the coefficient of the CRRA suggested by Mehra and Prescott (1985), we find that the value of information added by linear factor models of Fama and French (1992) exceeds observed management fees and only equals them for hitherto unmeasured magnitudes of risk aversion.

Keywords: CAPM, Fama-French Model, Value of Information.

JEL Codes: G20, G12.

1 Introduction

Information has become increasingly important in financial decision making. With the recent development of computing and network technique, vast amount of data, and various models, investors have more information than ever. Yet we do not know how valuable information is, and many of us take this recent development for granted. In our study we propose a valuation model for information in the expected utility theory framework.[1]

We extend Treich's (1997) model so that we could use it for the evaluation of a wider range of informational situations. In our study we focus on those prevalent in the world of asset management. Our model allows us to price performance of active fund managers both in terms of a down payment, and/or a profit share. Furthermore these payments can be related to the performance of the forecasting model provided by the active manager so that one can evaluate a model which provides an information coefficient of, say, 7%. We could answer if the level of fees charged by fund managers is appropriate. Christoffersen (2001) analyses fees for the period of 1990 to 1995 in the US market and reports that over half of money fund managers voluntarily waive their fees. However, it is not clear if the fees in the initial contracts are appropriate. Fund managers could overcharge fees and then waive them to increase expected fund flows.

We first provide the theoretical development. With the assumptions that asset returns and information are joint normally distributed, and for a constant relative risk aversion class of utility, closed form expressions for the value of information is derived. We show that whilst other decision making models such as the loss aversion (LA) utility of Kahneman and Tversky (1979) may be more appropriate to deal with these problems, they are much harder to implement. We can not provide analytical and numerical results due to the complexity of the utility function used in prospect theory. Then, we apply our analysis to evaluating the Fama-French model (1992) of risk premia based on styles, size and value. That is, we ask the question how much would a fund manager pay Professors Fama and French for the information that size and value can be used to select stocks. We find that within the range of the relative risk aversion coefficient (γ) suggested by some previous studies, i.e., $1 < \gamma < 2$, the value of information for the Fama-French model is too large, e.g., 1-2% a month. These results are consistent with Mehra and Prescott (1985) who show that, for constant relative risk aversion utility, investors would have to have the value of γ in excess of 30 to explain the historical equity premium. When we set $\gamma = 30$, we find that the value of information becomes smaller but still much higher than the management fee which investors pay to fund management companies in practice.

Our results show that the CRRA utility function needs a large value of the parameter to explain the management fees observed in practice. Using these high risk aversion models, we find that, even though eight years has passed since Fama and French (1992) published their paper, the linear factor model based on styles still provides additional information to the conventional CAPM. Also we find that the value of the additional information is more or less constant over time and has retained its value, notwithstanding substantial informational diffusion. We can infer an added value of active management of about 250 basis points a year based on the Fama-French model in excess of the capital asset pricing model. This can be contrasted with typical active fees that are currently about 45 basis points in the USA over the period 1990 to 1995. To compute a value of γ so that the price of information equals 48 basis points, we need γ to equal around 155. This can be interpreted in a number of ways. Firstly, power utility may be even more inadequate a tool in risk analysis than previously imagined. Secondly, and in contrast, fees for active management are cheap relative to the value that they add. Or finally, it is still possible to get excess returns with the Fama-French model.

In the next section, we first provide the theoretical development for a few important utility functions with the assumptions that asset returns and information are joint normally distributed. Then, in section 3, we apply our analysis to evaluating the Fama-French model (1992) of risk premia based on styles, size and value. Conclusions follow in section 4.

2 Standard Portfolio Model with Information

We shall structure the problem in terms of N assets with rates of return r_i, $i = 1, ..., N$, $\mathbf{r} = (r_1, r_2, ..., r_N)'$, and initial wealth w_0 plus one riskless asset, cash, with a rate of return, r_f. Then period 1 wealth, w_1, is

$$w_1 = w_0((1 + r_f) + \mathbf{x}'(\mathbf{r} - r_f\mathbf{e})) \tag{1}$$

where $\mathbf{x}$ is the vector of portfolio proportions, $\mathbf{e}$ is a $(N \times 1)$ vector of ones and $\mathbf{x}'\mathbf{e} = 1$.

Here we do not structure our analysis in terms of returns relative to a benchmark which is the analysis typically used in asset management. For such a case, risk is measured relative to a benchmark, and fees and performance related bonuses are often paid in terms of over/under performance relative to the benchmark as well. To deal with such a case, we would need to redefine the expected utility function. However, we prefer to present our results in a more conventional economic framework so that investors are motivated by final wealth, or total, not relative, returns.

The investor has an increasing concave Bernoulli utility function $u(w)$ and chooses $\mathbf{x}$ to maximise $E(u(w_1))$ subject to $\mathbf{x}'\mathbf{e} = 1$. Let

$$h = \max_{\mathbf{x}} \ E(u(w_1)). \tag{2}$$

Information enters the problem as a set of n conditioning variables, y, thus (2) is modified to

$$h(y) = \max_{\mathbf{x}} \ E(u(w_1)|y). \tag{3}$$

Pricing of information can be achieved in a variety of ways. In the relationships above, we have, as a possible pricing equation,

$$E_y(h(y)) = \max_{\mathbf{x}} \ E[u((w_0 + p_\theta)((1 + r_f) + \mathbf{x}'(\mathbf{r} - r_f\mathbf{e})))] \tag{4}$$

where p_0 would represent the minimum amount one would have to be compensated not to receive information Y, where Y are the random variables whose values are y. This has the interpretation of how much initial wealth would you be prepared to give up/receive to equate utility with information to utility without information. Treich's (1997) definition is similar to (4) except that

$$E_y(h(y)) = \max_{\mathbf{x}} \ E[u(w_1 + p_1)] \tag{5}$$

where w_1 is defined by (1) and p_1 is now the minimum amount of end-period wealth required to compensate the investor for not having the information. This definition assumes that fund managers will be compensated at the end of period.

Other versions are possible; we could have, for example,

$$h(y, p_0', p_1', \lambda) = \max_{\mathbf{x}} \ E[u((w_0 - p_0')((1 + r_f) + \mathbf{x}'(\mathbf{r} - r_f\mathbf{e})\lambda) - p_1')|y]. \tag{6}$$

Then equating

$$h = E(h(y, p_0', p_1', \lambda)), \tag{7}$$

i.e., equating equations (2) and (6), we define p_0' as the maximum amount of initial wealth the investor will pay to acquire information, p_1' is the maximum amount of final (fixed) wealth the investor will pay, and λ is the minimum proportion of active income retained so that $(1 - \lambda)$ is the maximum proportional active profit that the investor will give up.

In equation (6), term p_0', p_1', and λ have been jointly defined so that a contract could be specified in terms of a triplet $(p_0', p_1', 1 - \lambda)$ where of course any of the values could be set to zero. Some of the cases discussed have been analysed in the case of a CARA utility function by Nadiminti, Mukhopadhyay, and Kriebel (1996).

Although the definitions above have involved utility maximisation, i.e., maximising $u(w_1)$ with respect to $\mathbf{x}$, the definitions would still apply if $\mathbf{x}$ were partially fixed so that managers/investors need not be profit maximisers in all situations. Thus prices for information could be defined in terms of a current portfolio without information which is unmaximised but a maximised portfolio if information is received. In this study, we investigate the case that the value of information is paid at the beginning of the period.

Assuming compact returns (see Ingersoll, 1987), we can justify Taylor's series to second order in expanding (4), (5), and (7). Expanding (7) with $p_1' = 0$, $\lambda = 1$, we see that

$$\begin{aligned} h(y) &= \max\ E[u(w_0^*(1+r_f)) + \mathbf{x}'(\mathbf{r} - r_f\mathbf{e})w_0^* u'(w_0^*(1+r_f)) \\ &\quad + \frac{1}{2}(w_0^*)^2(\mathbf{x}'(\mathbf{r} - r_f\mathbf{e})(\mathbf{r} - r_f\mathbf{e})'\mathbf{x})u''(w_0^*(1+r_f))] \end{aligned} \tag{8}$$

where $w_0^* = w_0 - p_0'$. Denoting

$$\begin{aligned} u_0^* &= u(w_0^*(1+r_f)), \\ u_1^* &= u'(w_0^*(1+r_f)), \\ u_2^* &= u''(w_0^*(1+r_f)), \\ E(\mathbf{r} - r_f\mathbf{e}|y) &= \boldsymbol{\theta}, \\ E[(\mathbf{r} - r_f\mathbf{e})(\mathbf{r} - r_f\mathbf{e})'|y] &= \boldsymbol{\Sigma} + \boldsymbol{\theta}\boldsymbol{\theta}', \end{aligned}$$

we have

$$h(y) = \max \ u_0^* + w_0^* u_1^* \mathbf{x}'\boldsymbol{\theta} + \frac{1}{2}(w_0^*)^2 u_2^* \mathbf{x}'(\boldsymbol{\Sigma} + \boldsymbol{\theta}\boldsymbol{\theta}')\mathbf{x}. \tag{9}$$

How do we proceed in solving (7), i.e., solving for p_0' in

$$E(h(y)) = \max \ (u_0 + w_0 u_1 \mathbf{x}'\boldsymbol{\theta} + \frac{1}{2} w_0^2 u_2 \mathbf{x}'(\boldsymbol{\Sigma} + \boldsymbol{\theta}\boldsymbol{\theta}')\mathbf{x}), \tag{10}$$

where u_0, u_1, and u_2 are defined as above except that w_0^* is replaced by w_0? The answer clearly depends upon the distributional assumptions about $\mathbf{r} - r_f\mathbf{e}$ and y.

Assumption 1 We shall assume that $\mathbf{y}$ is an $(m \times 1)$ variable, and that $\mathbf{r} - r_f\mathbf{e}$ and $\mathbf{y}$ are joint normally distributed

$$\begin{pmatrix} \mathbf{r} - r_f\mathbf{e} \\ \mathbf{y} \end{pmatrix} \tilde{} N \left[\begin{pmatrix} \boldsymbol{\mu}_r \\ \boldsymbol{\mu}_y \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{rr} & \boldsymbol{\Sigma}_{ry} \\ \boldsymbol{\Sigma}_{yr} & \boldsymbol{\Sigma}_{yy} \end{pmatrix} \right]. \tag{11}$$

There are issues as to whether returns are in fact normally distributed. These have been discussed at great length in many places and we do not repeat the discussion here.

The implication of Assumption 1 is that

$$\begin{aligned} \boldsymbol{\theta} &= \boldsymbol{\mu}_r + \boldsymbol{\Sigma}_{ry}\boldsymbol{\Sigma}_{yy}^{-1}(\mathbf{y} - \boldsymbol{\mu}_y), \\ \boldsymbol{\Sigma} &= \boldsymbol{\Sigma}_{rr} - \boldsymbol{\Sigma}_{ry}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yr}. \end{aligned} \tag{12}$$

Note that

$$E(u(w_1)) = u_0^* + w_0^* u_1^* \mathbf{x}'\boldsymbol{\theta} + \frac{1}{2}(w_0^*)^2 u_2^* \mathbf{x}'(\boldsymbol{\Sigma} + \boldsymbol{\theta}\boldsymbol{\theta}')\mathbf{x},$$

and

$$E(u'(w_1)) = w_0^* u_1^* \boldsymbol{\theta} + (w_0^*)^2 u_2^* (\boldsymbol{\Sigma} + \boldsymbol{\theta}\boldsymbol{\theta}')\mathbf{x} = \mathbf{0}.$$

Thus we have

$$\mathbf{x} = -\frac{u_1^*}{w_0^* u_2^*}(\boldsymbol{\Sigma} + \boldsymbol{\theta}\boldsymbol{\theta}')^{-1}\boldsymbol{\theta},$$

and

$$\begin{aligned} h(y) &= \max E(u(w_1)) \\ &= u_0^* - \frac{1}{2}\frac{(u_1^*)^2}{u_2^*}\boldsymbol{\theta}'(\boldsymbol{\Sigma} + \boldsymbol{\theta}\boldsymbol{\theta}')^{-1}\boldsymbol{\theta} \\ &= u_0^* - \frac{1}{2}\frac{(u_1^*)^2}{u_2^*}\left(\boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta} - \frac{(\boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta})^2}{1 + \boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}\right) \\ &= u_0^* - \frac{1}{2}\frac{(u_1^*)^2}{u_2^*}\left(\frac{\boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}{1 + \boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}\right). \end{aligned} \tag{13}$$

So

$$E(h(y)) = u_0^* - \frac{1}{2}\frac{(u_1^*)^2}{u_2^*}E\left(\frac{\boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}{1 + \boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}\right). \tag{14}$$

Using the same method for the case of no information in (10), we have

$$u_0^* - \frac{1}{2}\frac{(u_1^*)^2}{u_2^*}E\left(\frac{\boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}{1 + \boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}\right) = u_0 - \frac{1}{2}\frac{{u_1}^2}{u_2}\frac{\boldsymbol{\mu}_r'\boldsymbol{\Sigma}_{rr}^{-1}\boldsymbol{\mu}_r}{1 + \boldsymbol{\mu}_r'\boldsymbol{\Sigma}_{rr}^{-1}\boldsymbol{\mu}_r} \tag{15}$$

and p_0' can be determined numerically for any given distribution satisfying Assumption 1 and any particular utility function.

2.1 Logarithmic Utility

This does raise the question as to whether (15) has any explicit solutions. It transpires that $u(w_1) = ln(w_1)$ has the property that $u_1^2/u_2 = -1$ for all value of w. We thus make Assumption 2.

Assumption 2 $u(w) = \ln(w)$.

Using Assumptions 1 and 2 and substituting into (15) gives

$$\ln[(w_0 - p_0')(1+r_f)] + \frac{1}{2}E\left(\frac{\boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}{1+\boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}\right) = \ln[w_0(1+r_f)] + \frac{1}{2}\frac{\boldsymbol{\mu}_r'\boldsymbol{\Sigma}_{rr}^{-1}\boldsymbol{\mu}_r}{1+\boldsymbol{\mu}_r'\boldsymbol{\Sigma}_{rr}^{-1}\boldsymbol{\mu}_r}.$$

Therefore

$$\begin{aligned} p_0' &= w_0\left(1 - \exp\left[\frac{1}{2}\frac{\boldsymbol{\mu}_r'\boldsymbol{\Sigma}_{rr}^{-1}\boldsymbol{\mu}_r}{1+\boldsymbol{\mu}_r'\boldsymbol{\Sigma}_{rr}^{-1}\boldsymbol{\mu}_r} - \frac{1}{2}E\left(\frac{\boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}{1+\boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}\right)\right]\right) \\ &= w_0\left(1 - \exp\left[\frac{1}{2}(R_0^2 - R_y^2)\right]\right), \end{aligned}$$

where

$$R_0^2 = \frac{\boldsymbol{\mu}_r'\boldsymbol{\Sigma}_{rr}^{-1}\boldsymbol{\mu}_r}{1+\boldsymbol{\mu}_r'\boldsymbol{\Sigma}_{rr}^{-1}\boldsymbol{\mu}_r}, \qquad (16)$$

and

$$R_y^2 = E\left(\frac{\boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}{1+\boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}\right), \qquad (17)$$

since they bear a close resemblance to conventional goodness of fit measures in linear regression. Furthermore, since $z/(1+z)$ is a concave function, we can apply Jensen's inequality so that $R_y^2 \geq \frac{E(\boldsymbol{\theta})'\boldsymbol{\Sigma}^{-1}E(\boldsymbol{\theta})}{1+E(\boldsymbol{\theta})'\boldsymbol{\Sigma}^{-1}E(\boldsymbol{\theta})} = \frac{\boldsymbol{\mu}_r'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_r}{1+\boldsymbol{\mu}_r'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_r}$, since $E(\boldsymbol{\theta}) = \boldsymbol{\mu}_r$.

Now $\boldsymbol{\Sigma} \leq \boldsymbol{\Sigma}_{rr}$ in the positive semi-definite sense, so that $\boldsymbol{\Sigma}^{-1} \geq \boldsymbol{\Sigma}_{rr}^{-1}$ in the positive semi-definite sense and $\frac{\boldsymbol{\mu}_r'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_r}{1+\boldsymbol{\mu}_r'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_r} \geq \frac{\boldsymbol{\mu}_r'\boldsymbol{\Sigma}_{rr}^{-1}\boldsymbol{\mu}_r}{1+\boldsymbol{\mu}_r'\boldsymbol{\Sigma}_{rr}^{-1}\boldsymbol{\mu}_r}$ since $z/(1+z)$ is an increasing function and $\boldsymbol{\Sigma}_{rr}^{-1}$ and $\boldsymbol{\Sigma}^{-1}$ are both positive definite. Thus $R_y^2 - R_0^2$ is a non-negative quantity that measures the value of the information in y.

Now the above arguments extend to the case where the investor is wishing to buy the information (y) and already has an information set (y_0). Thus the immediate generalisation of (8), (9) and (10) would be given by

$$h(y_0) = \max_x E(u(w)|y_0) \tag{18}$$

thus,

$$E(h(y_0)) = E(\max_x (E(u(w)|y, y_0))$$

where the appropriate pricing concept is used analogously to equations (4), (5), (6), and (7).

In particular, if we make assumptions 1 and 2, and define $\Psi = (y, y_0)$ to be the augmented information set, then an argument identical to the one above shows that we can calculate p_0' as

$$p_0' = w_0(1 - \exp[\frac{1}{2}(R^2_{y_0} - R^2_\Psi)]) \tag{19}$$

The definition of $R^2_{y_0}$ and R^2_Ψ are as follows. Let

$$\begin{aligned} \boldsymbol{\theta}_{y_0} &= \boldsymbol{\mu}_r + \boldsymbol{\Sigma}_{ry_0}\boldsymbol{\Sigma}^{-1}_{y_0y_0}(y_0 - \boldsymbol{\mu}_{y_0}) \\ \boldsymbol{\Sigma}_1 &= \boldsymbol{\Sigma}_{rr} - \boldsymbol{\Sigma}_{ry_0}\boldsymbol{\Sigma}^{-1}_{y_0y_0}\boldsymbol{\Sigma}_{y_0r} \end{aligned}$$

and

$$\begin{aligned} \boldsymbol{\theta}_{\Psi} &= \boldsymbol{\mu}_r + \boldsymbol{\Sigma}_{r\Psi}\boldsymbol{\Sigma}^{-1}_{\Psi\Psi}(\Psi - \boldsymbol{\mu}_{\Psi}) \\ \boldsymbol{\Sigma}_2 &= \boldsymbol{\Sigma}_{rr} - \boldsymbol{\Sigma}_{r\Psi}\boldsymbol{\Sigma}^{-1}_{\Psi\Psi}\boldsymbol{\Sigma}_{\Psi r} \end{aligned}$$

where $\boldsymbol{\Sigma}_{r\Psi} = (\boldsymbol{\Sigma}_{ry_0}, \boldsymbol{\Sigma}_{ry})$, $\boldsymbol{\Sigma}_{\Psi\Psi} = \begin{pmatrix} \boldsymbol{\Sigma}_{y_0y_0} & \boldsymbol{\Sigma}_{y_0y} \\ \boldsymbol{\Sigma}_{yy_0} & \boldsymbol{\Sigma}_{yy} \end{pmatrix}$, etc, and

$$\begin{aligned} R^2_{y_0} &= E\left(\boldsymbol{\theta}'_{y_0}\boldsymbol{\Sigma}^{-1}_1\boldsymbol{\theta}_{y_0}/(1 + \boldsymbol{\theta}'_{y_0}\boldsymbol{\Sigma}^{-1}_1\boldsymbol{\theta}_{y_0})\right) \qquad (20) \\ R^2_{\Psi} &= E\left(\boldsymbol{\theta}'_{\Psi}\boldsymbol{\Sigma}^{-1}_2\boldsymbol{\theta}_{\Psi}/(1 + \boldsymbol{\theta}'_{\Psi}\boldsymbol{\Sigma}^{-1}_2\boldsymbol{\theta}_{\Psi})\right). \end{aligned}$$

By extensions of the previous argument, $R^2_\Psi - R^2_{y_\Theta}$ is non-negative and so we see that the extra information is valued in terms of extra R^2.

2.2 Power Utility and Loss Aversion Utility

We now consider a more general CRRA utility function with the power utility function. Note that when the coefficient of the power utility function is one, the utility function is equivalent to the logarithmic utility function. The power utility function we use in this study is as follows;

Assumption 3

$$u(w) = \frac{w^{1-\gamma}}{1-\gamma}, \gamma > 0, \gamma \neq 1.$$

In this case,

$$\begin{aligned} u_0^* &= \frac{1}{1-\gamma}(w^*)^{1-\gamma}, \\ u_1^* &= (w^*)^{-\gamma}, \\ u_2^* &= -\gamma(w^*)^{-(1+\gamma)}, \end{aligned}$$

and

$$\frac{(u_1^*)^2}{u_2^*} = -\frac{(w^*)^{1-\gamma}}{\gamma},$$

where $w^* = w_0^*(1+r_f)$. Substituting the above equations into (15) gives

$$\begin{aligned} & ((w_0 - p_0')(1+r_f))^{1-\gamma}\left[\frac{1}{1-\gamma} + \frac{1}{2\gamma}E\left(\frac{\boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}{1+\boldsymbol{\theta}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\theta}}\right)\right] \\ = \;& (w_0(1+r_f))^{1-\gamma}\left[\frac{1}{1-\gamma} + \frac{1}{2\gamma}\left(\frac{\boldsymbol{\mu}_r'\boldsymbol{\Sigma}_{rr}^{-1}\boldsymbol{\mu}_r}{1+\boldsymbol{\mu}_r'\boldsymbol{\Sigma}_{rr}^{-1}\boldsymbol{\mu}_r}\right)\right] \end{aligned}$$

and we have

$$
\begin{aligned}
p_0' &= w_0 \left[1 - \left[\frac{\frac{1}{1-\gamma} + \frac{1}{2\gamma} \frac{\boldsymbol{\mu}_r' \boldsymbol{\Sigma}_{rr}^{-1} \boldsymbol{\mu}_r}{1+\boldsymbol{\mu}_r' \boldsymbol{\Sigma}_{rr}^{-1} \boldsymbol{\mu}_r}}{\frac{1}{1-\gamma} + \frac{1}{2\gamma} E\left(\frac{\boldsymbol{\theta}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\theta}}{1+\boldsymbol{\theta}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\theta}} \right)} \right]^{\frac{1}{1-\gamma}} \right] \qquad (21) \\
&= w_0 \left[1 - \left[\frac{\frac{1}{1-\gamma} + \frac{1}{2\gamma} R_{y\Theta}^2}{\frac{1}{1-\gamma} + \frac{1}{2\gamma} R_{\Psi}^2} \right]^{\frac{1}{1-\gamma}} \right]
\end{aligned}
$$

if we consider the situation as in (19).

Many previous studies, either theoretically or empirically, suggest that the admissable range of the coefficient of the CRRA, γ, is between one and two (see Arrow, 1971; Tobin and Dolde, 1971; Friend and Blume, 1975; Kydland and Prescott, 1982; Kehoe, 1984). However, with such a range of γ, a large equity premium in the US cannot be explained. Mehra and Prescott (1985) suggest that the puzzle can only be solved when γ is of the order of 30. In addition, the portfolios of US investors do not seem to be explained by standard portfolio choice models such as CRRA. This can be summarised as excessive equity positions of the US investors unless investors are unreasonably risk averse (see Campbell and Viceira, 1998).

The puzzle has been tackled from various perspectives by other studies such as Rietz (1988), and Benartzi and Thaler (1995). Recently, Ang, Bekeart and Liu (2000) suggest that the disappointment aversion preferences of Gul (1991) may explain the portfolio choice of US investors. This has similarities to the loss aversion class of utility introduced by Kahneman and Tversky (1979, 1992). Although prospect theory utility seems to be promising way of pricing information, the closed form solution for (7) is very difficult to obtain and approximations of the sort given in this paper do not bring about significant simplifications. So, whilst such theories probably lead to more realistic prices, they do not lend themselves to the methodologies advanced in this paper. Our approach therefore will be to carry out computations with power and log-utility and report results for a range of values of γ.

3 Empirical Tests

One of the important benefits in our derivation of information value in the previous section is that we can now calculate the value of information with an assumption of a particular utility function. In this section, we calculate the value of information with the logarithmic utility function in (19) as well as the power utility function in (21) for different values of γ.

We calculate the value of information using CAPM and linear factor models, especially for the Fama and French model (1992) of risk premia on value and size. Thus, we use three factors as explanatory variables; excess market returns and two factor mimicking portfolio (FMP) returns for value and size. Factor portfolios may be portfolios of equities corresponding to attractive equity characteristics such as size or value. The factor portfolios may be actual indices such as the S&P500, FTSE100 (for size) or artificially constructed portfolios long in high value equities and short in low value equities (for value).

The data we use are the log-returns of the S&P500 index and 450 individual equities belonging to the S&P500 index, 3 month US treasury bill, and FMP returns for value and size. A total number of 180 monthly returns from April 1984 to March 1999 is used. Excess individual equity returns and market returns are calculated by taking the 3 month US treasury bill from individual equity returns and market returns, respectively. Due to the changes of the components in the S&P500 index and unavailable equity returns in early stages of our sample period, the number of available equities at the beginning of the sample period is 350.

3.1 Factor Mimicking Portfolios (FMPs)

A factor mimicking portfolio (FMP) is a portfolio of assets whose returns are designed to be highly correlated with the (unobservable) factor values. Portfolios constructed from eigenvectors in principal component analysis are examples of factor mimicking portfolios.

In this study, we use two sets of value and size factors; our own factors as described below, and the Fama-French factors. We first explain how our FMPs are constructed in detail. The explanation on the Fama-French factors can be found in Fama and French (1993). Our factor mimicking portfolios are constructed with the same method as in Hwang and Satchell (1999) and Hall, Hwang, and Satchell (2002). To calculate the FMPs, we use a total number of 2046 US equities in the MSCI universe from April 1984 to March 1999. For each factor f, the universe is ranked by an attribute of f. For instance, f might be the size factor and the universe of stocks would be ranked by their size attribute data. Then an equally weighted portfolio is formed that is long the top n-tile ranked by the f attribute, and short the bottom n-tile, ranked by the f attribute. The resulting hedge portfolio is the factor mimicking portfolio of factor f. The order of the n-tile should typically be small. Thus, the use of quartiles ($n = 4$) or thirds ($n = 3$) is probably more appropriate than the use of, for example, deciles ($n = 10$), because of the greater diversification produced. In this particular model thirds have been used.

The two attributes used in this study are value (VL) and size (SZ). The FMPs calculated for the two styles are the factor mimicking portfolios for value (FVL) and size (FSZ). Note that the returns of the FMPs are calculated each month. Thus, the number of equities used for the calculation of FMPs changes over time (but the maximum number is 2046). The values of the two attributes for each equity are defined as :

$$\begin{aligned} SZ_t &= \log(SharePrice(t) \times ShareNumber(t)), \\ VL_t &= \frac{DP_t + EP_t + SP_t + BP_t + CP_t}{5}, \end{aligned}$$

where

$$\begin{aligned} DP_t &= \frac{Dividend(t)}{Share\ Price(t)}, \\ EP_t &= \frac{Earnings\ Per\ Share(t)}{Share\ Price(t)}, \\ SP_t &= \frac{Net\ Sales\ Per\ Share(t)}{Share\ Price(t)}, \\ BP_t &= \frac{Book\ Value\ Per\ Share(t)}{Share\ Price(t)}, \end{aligned}$$

$$CP_t = \frac{Cash\ Flow\ Per\ Share(t)}{Share\ Price(t)}.$$

The above procedure can be seen as a natural way to make the factor mimicking portfolios approximately uncorrelated. If the true factors are uncorrelated, then sorting by the attributes and constructing long and short positions relative to factor 1, say, should produce a portfolio with little to no systematic exposure to factor 2. Actually, the attributes themselves may be strongly correlated and, as a consequence, fully uncorrelated portfolios cannot be expected. However, it is considered that this procedure is preferable to using factor analysis or principal components, since then one usually loses any understanding of what the factors signify.

Some of the statistical properties of excess market returns and FMP returns are reported in panel A of table 1. None of the factor returns have mean returns significantly different from zero. Note that correlations between the excess market returns and the two FMPs are significant and all three explanatory variables are not normally distributed. Thus our assumption of multivariate normality may be too strong. However, our choice and calculation of explanatory variables are not different from those used in practice, and thus our calculations should have some parallels with the economic realities of fund management.

As explained above, we also use two Fama-French Benchmark Factors; the performance of small stocks relative to big stocks (SMB, Small Minus Big) and the performance of value stocks relative to growth stocks (HML, High Minus Low). Note that Fama-French SMB and HML are equivalent to our factor mimicking portfolios, FSZ and FVL, respectively. The more detailed explanation on these factor returns can be found in Fama and French (1993).

However, since the construction methods and the universe between our factors and Fama-French factors are not exactly same, some of the properties are expected to be different. The correlation coefficients between SMB and FSZ as well as between HML and FVL are -0.773 and 0.395, respectively. The two size related factors are highly correlated, but the two value related factors are not highly correlated. This is because Fama and French use one attribute, book-to-market value, whilst our FVL is obtained with five attributes as in the above.

We report the properties of these factor return in panel B of table 1. The mean values of FSZ and FVL are very large compared with those of SMB and HML. Other statistics such as skewness and excess kurtosis do not seem to be similar. As explained above, these results come from the differences in the construction of the factors. Despite these differences, we note that none of the statistics in panel A are statistically significantly different from those of panel B of table 1. However, the correlation matrixes show that we can find similar relationship between the factors and the market portfolio in both cases; FVL is significantly negatively correlated with the market portfolio returns and the FSZ, and HML is also significantly negatively correlated with the market portfolio returns and the SMB.

The correlation matrixes in panels A and B suggest that the value of information obtained with these two different sets of factors may be similar. In the next subsection, we try to answer this question by investigating how much we have to pay for the CAPM and the Fama-French model with these factors.

3.2 Results

For the calculation of the value of information, we need to invert a variance-covariance matrix in equation (20). We group the equities randomly such that in each group 50 equities are assigned. For each group, we use the first 60 observations from April 1984 to March 1989 to obtain the value of information at March 1989, and then use the next 60 observations from May 1984 to April 1989 to obtain the value of information at April 1989, and so on. From these rolling windows we obtain 121 information values from March 1989 to March 1999 for each group. We then calculate the value of information at each time by obtaining an average value over the different groups.

In most cases, however, our interests lie in the forecasting power of the CAPM or the linear factor models. Whilst it might be a useful exercise to look at current explanatory variables to explain current returns for the added value in more efficient risk control management, this is probably secondary to forecasting in terms of the relevance of the application. Thus, we use one month lagged excess market returns for the CAPM, and one month lagged excess market returns and one month lagged factor mimicking portfolio (FMP) returns for value and size for the Fama and French linear factor model. The information in the original CAPM and the Fama and

French model is simultaneous with equity returns (current information), whilst the information in forecasting equity returns in this study is past (past information). Of course, one could argue that this information is more useful to diversify stocks than to forecast stocks. Our mythical fund manager who is buying this information is using it to forecast. This setting is more realistic since many models in practice are based on these approaches.

Before we discuss the results, it is worthwhile considering what US/UK management fees were in the period 1989 to 2000. Christoffersen (2001) analyses fee data from two sources: Lipper Analytical Services and the IBC Donoghue Quarterly Report on Mutual Fund Performance. These sources cover the period 1990 to 1995 in the US. She reports fees for institutional funds involving contracted advisory and non-advisory fees that average about 50 basis points per annum. The corresponding amount for retail funds is about 65 basis points.

Table 2 reports the values of information during sample period, which is obtained by averaging the 121 values of information for various values of γ. For the range of $0.5 < \gamma < 2$, the value of information lies between 5.55% and 2.5% a month for the CAPM and between 9.79% and 4.59% a month for the Fama-French model with our factors, whilst it lies between 9.73% and 4.56% a month for the Fama-French model with the Fama-French factors. Therefore, for the same ranges of γ, the value of additional information lies between 4.53% and 2.15% a month with our factors and between 4.47% and 2.12% a month with the Fama-French factors. These numbers seem to be too large. Finally, the values of information for the values of γ up to 35 show, unsurprisingly, that the fee paid decreases with γ. Taking $\gamma = 30$ as our proxy for the real utility function, we see that active management adds 240 basis points a year to the CAPM. To achieve 4 basis points a month we need γ equivalent to 155!

Many authors have argued that large values of the relative risk aversion coefficient are incompatible with economic theory. This incompatibility stems from models where the representative agent has power utility and the equity risk premium can be calculated as γcov(returns to equity, change in consumption), where γ is the coefficient of RRA. It is claimed that the continuous time consumption CAPM will not allow large values of γ as it will lead to negative interest rates (see Weil, 1989; Cochrane, 1997). Other macroeconomic anomalies are also mentioned in Cochrane (pages 16-17, 1997). We argue that our utility function is that of a typical (but not representative) fund manager; thus large values of γ can co-exist with any representative agent utility, except under rather idealised situations where power utility aggregation may apply.

The empirical results reported in table 1 show that there are few differences between the two different sets of factors; our factors and Fama-French factors. This was expected since correlation matrices of the two sets of factors are nearly the same and higher moments are not considered. Thus, the explanation below only uses the results with our factors.

We plot the 121 information values from March 1989 to March 1999 with $\gamma = 1.5$

and $\gamma = 30$. Figures 1 and 2 suggest, not surprisingly, that the value of information changes over time. From the figures we can pick up some period when the CAPM and the Fama-French model are highly valued; e.g., from the beginning of 1991 to early 1992, middle 1993 to the end of 1994, and late 1998 to the end of sample period. The changing value of information is consistent with empirical results on the varying importance of style factor (see Bauman, Conover, and Miller, 1999; Hwang and Satchell, 1999). It is worth noting that the values of information are highly autocorrelated; the estimates of the autocorrelation coefficient with lag one are around 0.84.

The figures also suggest that the value of information does not seem to show any long run trend. Furthermore, the value of the additional information shows no significant difference over time. This means that the Fama-French model can be still valued after their initial publication of their research in 1992 and that the value has not been diluted by the diffusion of information. Thus one explanation, that as a theory becomes widely known, its incremental value decreases as many investors incorporate it into their trading strategies, does not appear to be upheld by our analysis. There may be a linkage between the persistence of the value of information and the information variables being sources of risk rather than anomalies.

Whilst we might stand accused of rediscovering the equity premium puzzle by yet another circuitous route, our use of power utility with a γ of 30 or higher can be seen as a proxy for a more complex utility function. Our proxy has been calibrated to get the risk premia approximately right and hence hopefully, the value of information right. Based on this interpretation, our informational values could be deemed to be sensible.

We note several limitations of our exercise, which may reduce the high value of information for the CAPM and the Fama-French model. Firstly, we do not exclude short-selling and this might influence results. In our study, the sum of the absolute values of portfolio weights was about eight on average. Thus our "optimal" portfolios were characterised by large numbers of large long and short positions. If our fund manager was a long-short investor with longs at 130% and shorts at 30%, we might express the absolute magnitude of portfolio weights to be 1.6. Thus our idealised portfolios are traded about five times more than we might expect for an active manager and are five times more geared than we might expected for a long-short manager.

Secondly, we ignore transaction costs. Since we assume that our representative agent is an institutional investor who is trading S&P500 stocks through portfolio trades, her transaction costs are likely to be of the order of 25 basis points per trade. Since trading is monthly and turn-over is of the order of 80 percent per annum, annual transaction costs may only be of the order of 20 basis points. However, we found that our turn-over, was on average, about 40% a month; this is obviously very high and our value of information would be much smaller if we traded at this level and paid even modest costs.

4 Conclusions

Our paper has provided a procedure for the evaluation of information that generalises existing procedures. Furthermore, in certain cases, we can derive closed form expressions for the price of information. This approach is applied to the valuation of the Fama-French model. Thus we may answer the question 'do active managers overcharge for providing stock selection skills based on linear factor models?' Our results suggest that an active manager who uses past factor information to forecast returns adds about 240 basis points over and above market information relative to a fee of between 30 to 70 basis points. Thus performance is considerably in excess of current fee structures and suggests that current fees are cheap, even if we allow for 50 basis points for annual transaction costs in trading.

We found that a degree of relative risk aversion that equates fees in practise with the fees derived from theory requires that γ, the coefficient of relative risk aversion, be set to 155! From here we conclude that either power utility is even worse than we previously thought or that active management is currently rather cheap at about 50 basis points a year.

However, because of the limitations in this study, the value of information in the CAPM and the Fama-French model seems to be much less than calculated in this study and thus the coefficient of relative risk aversion will be smaller than the numbers we found. Despite the limitations, we still expect that there may be a non-trivial information value for the Fama-French model, since the value does not show any upward or downward trend over last several years.

References

Ang, A., Bekeart, G., Liu, J. (2000), 'Why Stocks May Disappoint', National Bureau of Economic Research Working Paper 7783.

Arrow, K. J. (1971), *Essays in the Theory of Risk-bearing*, North-Holland, Amsterdam.

Balduzzi, P., Lynch, A. (1999), 'Transaction costs and predictability: Some utility cost calculations', *Journal of Financial Economics* 52(1), 47-78.

Barberis, N. (2000), 'Investing for the Long Run when Returns are Predictable', *Journal of Finance* 55, 225-264.

Bauman, W. S., Conover, C. M., Miller, R. E. (1999), 'Investor Overreaction in International Stock Markets', *Journal of Portfolio Management* 25(4), 102-111.

Benartzi, S., Thaler, R. H. (1995), 'Myopic Loss Aversion and the Equity Premium Puzzle', *Quarterly Journal of Economics* 110(1), 73-92.

Brandt, M. W. (1999), 'Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach', *Journal of Finance* 54, 1609-1646.

Brandt, M. W., Ait-Sahalia, Yacine. (2001), 'Variable Selection for Portfolio Choice', *Journal of Finance* 56, 1297-1351.

Campbell, J., Viceira, L. (1999), 'Consumption and Portfolio Decisions When Expected Returns are Time-varying', *Quarterly Journal of Economics* 114, 433-495.

Christoffersen, S. E. K. (2001), 'Why Do Money Fund Managers Voluntarily Waive Their Fees?', *Journal of Finance* 56, 1117-1140.

Cochrane, J. H. (1997), 'Where is the Market Going? Uncertain Facts and Novel Theories', *Economic Perspectives* XXI: 6 (November/December 1997), Federal Reserve Bank of Chicago.

Fama, E., French, K. R. (1992), 'The Cross-Section of Expected Stock Returns', *Journal of Finance* 47, 4274-65.

Fama, E., French, K. R. (1993), 'Common Risk factors in the Returns on Stocks and Bonds', *Journal of Financial Economics* 33, 3-56.

Friend, I., Blume, M. E. (1975), 'The Demand for Risky Assets', *American Economic Review* 65, 900-922.

Gul, F. (1991), 'A Theory of Disappointment Aversion', *Econometrica* 59(3), 667-686.

Hall, A., Hwang, S., Satchell, S. E. (2002), 'Using Bayesian Variable Selection Methods to Choose Style Factors in Global Stock Return Models', *Journal of Banking and Finance* 26, 2301-2325.

Hwang, S., Satchell, S. E. (1999), 'The Disappearance of Style in the US Equity Market', Financial Econometrics Research Centre Working Paper WP99-6, Cass Business School.

Ingersoll, J. E. (1987), *Theory of Financial Decision Making*, Rowman & Littlefield Publishers, Inc.

Kahneman, D., Tversky, A. (1979), 'Prospect Theory: An Analysis of Decision Under Risk', *Econometrica* 47, 263-291.

Kahneman, D., Tversky, A. (1992), 'Advances in Prospect Theory: Cumulative Representation of Uncertainty', *Journal of Risk and Uncertainty* 5, 297-323.

Kehoe, P. J. (1983), 'Dynamics of the Current Account: Theoretical and Empirical Analysis', Working paper, Harvard University.

Kydland, F. E., Prescott, E. C. (1982), 'Time to Build and Aggregate Fluctuations', *Econometrica* 50, 1345-1370.

Mehra, R., Prescott, E. C. (1985), 'The Equity Premium: A Puzzle', *Journal of Monetary Economics* 15, 145-161.

Nadiminti, R., Mukhopadhyay, T., Kriebel, C. H. (1996), 'Risk Aversion and the Value of Information', *Decision Support Systems* 16(3), 241-254.

Rietz, T. A. (1988), 'The Equity Risk Premium: A Solution', *Journal of Monetary Economics* 22(1), 117-31.

Tobin, J., Dolde, W. (1971), Wealth, Liquidity and Consumption, *in Consumer Spending and Monetary Policy: The Linkage* (Federal reserve Bank of Boston, Boston, MA), 99-146.

Treich, N. (1997), 'Risk Tolerance and Value of Information in the Standard Portfolio Model', *Economics Letters* 55(3), 361-63.

Weil, P. (1989), 'The Equity Premium Puzzle and the Risk-free Rate Puzzle', *Journal of Monetary Economics* 24, 401-421.

Xia, Y. (2001), 'Learning about Predictability: the Effect of Parameter Uncertainty on Dynamic Asset Allocation', *Journal of Finance* 56, 205-246.

Notes

[1]We focus on pricing of information rather than the forecastibility of returns or the welfare gains from asset allocation. See Balduzzi and Lynch (1999), Barberis (2000), Brandt (1999, 2001), Campbell and Viceira (1999), Xia (2001) for the implications for portfolio choice and welfare. The implication of predictability is also ignored in this paper, since we restrict attention to investors with only one-month horizon investor. Other work of interest has been carried out by Barberis (2000) and Xia (2001) and others who take into consideration that the predictability relations derived from the data are only estimates. Thus the investor will have uncertainty about whether this relation is in the data generation model or not, and will learn about it over time.

Table 1 Properties of Excess Market Returns and Factor Mimicking Portfolios

A. Factor Mimicking Portfolio of Hwang and Satchell (2000)

	Excess Market Returns	Factor Mimicking Portfolio (Value)	Factor Mimicking Portfolio (Size)
Mean	0.935	1.571	1.474
Standard Deviation	4.399	2.765	2.200
Skewness	-1.460	-1.038	-0.777
Excess Kurtosis	6.738	1.247	2.035
Correlation Matrix			
Excess Market Returns	1.000		
(Value)	-0.242	1.000	
Factor Mimicking Portfolio (Size)	0.113	-0.350	1.000

Notes: The above statistics for the excess market returns and two factor mimicking portfolio returns are calcualted with 180 monthly returns from April 1984 to March 1999. The excess market returns are obtained with S&P500 index total returns minus 3 month US Treasury bill rate. The two factor mimicking portfolio returns are calculated as described in Hall, Hwang and Satchell (2000) with 2046 US equities.

B. Factor Mimicking Portfolio of Fama and French

	Excess Market Returns	HML (the performance of value stocks relative to growth stocks)	SMB (the performance of small stocks relative to big stocks)
Mean	0.935	0.091	-0.317
Standard Deviation	4.399	2.514	2.526
Skewness	-1.460	0.106	0.130
Excess Kurtosis	6.738	0.279	1.104
Correlation Matrix			
Excess Market Returns	1.000		
HML (the performance of value stocks relative to growth stocks)	-0.477	1.000	
SMB (the performance of small stocks relative to big stocks)	0.077	-0.278	1.000

Notes: The above statistics for the excess market returns and two factor mimicking portfolio returns are calcualted with 180 monthly returns from April 1984 to March 1999. The excess market returns are obtained with S&P500 index total returns minus 3 month US Treasury bill rate. The two factor mimicking portfolio returns are the Fama-French factors downloaded from http://web.mit.edu/kfrench/www.

Table 2 Value of Information CAPM and Linear Factor Models with Lagged Factor Returns

Utility Functions	CAPM (A)	Excess Market Return, FMP(Value), FMP(Size) (B)	Excess Market Return, HML(Value), SMB(Size) (C)	Additional Value of Information with Hwang and Satchell Factors (B)-(A)	Additional Value of Information with Fama-French Factors (C)-(A)
Power Utility with γ=0.5	5.55%	9.79%	9.73%	4.53%	4.47%
Logarithmic Utility	3.94%	7.10%	7.06%	3.31%	3.26%
Power Utility with γ=1.5	3.06%	5.57%	5.54%	2.61%	2.57%
Power Utility with γ=2	2.50%	4.59%	4.56%	2.15%	2.12%
Power Utility with γ=5	1.19%	2.23%	2.21%	1.05%	1.03%
Power Utility with γ=10	0.64%	1.20%	1.19%	0.57%	0.56%
Power Utility with γ=15	0.43%	0.82%	0.81%	0.39%	0.38%
Power Utility with γ=20	0.33%	0.62%	0.62%	0.29%	0.29%
Power Utility with γ=25	0.27%	0.50%	0.50%	0.24%	0.23%
Power Utility with γ=30	0.22%	0.42%	0.42%	0.20%	0.20%
Power Utility with γ=35	0.19%	0.36%	0.36%	0.17%	0.17%

Notes: The above results are obtained with monthly excess returns of randomly selected 450 equities belonging to the S&P500 index from April 1984 to March 1999. Due to the changes of the constituents in the S&P500 index and unavailable equity returns in early stages of our sample period, the number of available equity returns in early stages of our sample period is 350. The excess monthly market returns were calculated with S&P500 index total returns and the 3 month US Treasury bill. To avoid difficulties in calculation, we group the equities randomly such that in each group 50 equities are assigned. The value of information was calculated with the assumption of the logarithmic and power utility functions. By 'Lagged Factor Returns', we mean one period (month) before excess market return and FMP returns as explanaroty factors.

Figure 1 Information Values of CAPM and Fama-French Linear Factor Model with Lagged Factor Returns (γ=1.5)

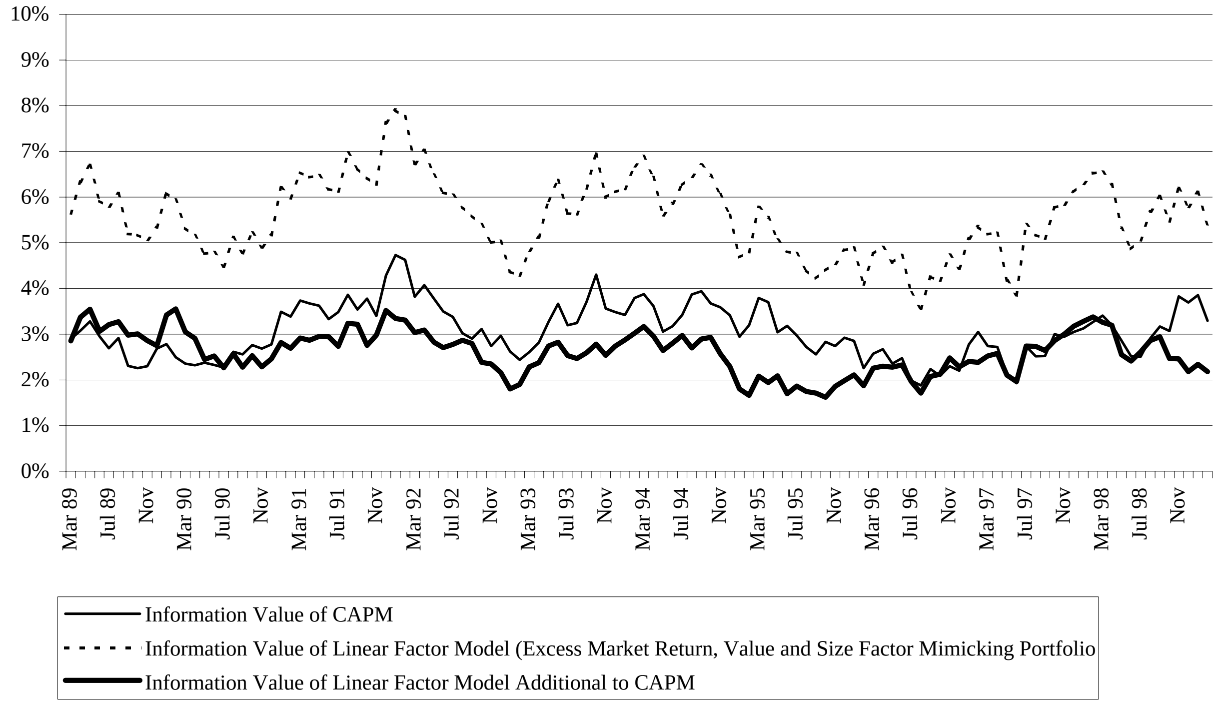

Figure 2 Information Values of CAPM and Fama-French Linear Factor Model with Lagged Factor Returns (γ=30)