Empirical Economics (2001) 26:293–305

EMPIRICAL
ECONOMICS

Portfolio style: Return-based attribution using quantile regression

Gilbert W. Bassett Jr., Hsiu-Lang Chen

Department of Finance, University of Illinois at Chicago, 601 South Morgan (mc 168), Chicago IL 60607-7121 (Email: gib@uic.edu; hsiulang@uic.edu)

Abstract. Return-based classification identifies a portfolio's style signature in the time series of its returns. Detection is based on a regression of portfolio returns on returns of factor mimicking indices. The method is easy to apply and does not require information about portfolio composition. Classification using least squares means that style is determined by the way factor exposure influences *expected* returns. We introduce regression quantiles as a complement to the standard analysis. The regression quantiles extract additional information from the time series of returns by identifying the way style affects returns at places other than the expected value. This allows discrimination among portfolios that would be otherwise judged equivalent based on conditional expectations. It also provides direct information about the impact of style on the tails of the conditional return distribution. Simple examples are presented to illustrate regression quantile classification.

Key words: Quantile Regression, Investment Style, Portfolio Attribution

1. Introduction

Professional investment managers follow a variety of approaches to deliver superior performance. Some rely solely on quantitative models. Others are quantitatively-oriented, but also consider information from sources such as firm managers, press reports and security analysts' recommendations. The variety of approaches complicates the problem of evaluating the performance of investment managers. The investment industry has responded to the pro-

We would like to thank the reviewers for their many helpful suggestions. An earlier version of the paper was presented at the International Conference on "Economic Applications of Quantile Regression" at the University of Konstanz, June 2000.

liferation of investment approaches by developing objective measures of a manager's investment style.

A money manager adopts an "investment style" by identifying securities with certain characteristics for potential inclusion in the portfolio. Fund managers with similar investment philosophies or styles will tend to perform more like each other than the overall market or managers with different styles.

The focus on style arises for several reasons. Accounting for style aids performance evaluation by giving a clearer picture about a manager's stock selection skill. The manager of a portfolio of small stocks may appear to have disappointing performance relative to a broad market index while performance may be outstanding relative to a small stock benchmark. A second motive arises from risk control. A pension plan sponsor might select a few active managers who are expected to achieve superior performance. If they all follow similar styles, however, they will tend to select similar stocks, yielding an overall plan portfolio that is highly undiversified.

Mutual fund managers are typically paid a fixed percentage of the fund's assets. Empirical evidence indicates that recent performance, relative to that of other funds or a benchmark index, has a positive effect on new inflows; see, for example, Ippolito (1992), Gruber (1996), and Sirri and Tufano (1998). In other words, mutual fund investors "chase returns" by channeling new investments towards better performing funds. It is not unlike a tournament in which managers compete for fund inflows and, ultimately, compensation. In such an investment tournament, managers have incentives to change their investment style, either ex ante, in order to enhance performance by under-representing the risk, or ex post, by switching styles in order to justify less than satisfactory returns. Therefore, there is a need for style classification of a manager's strategy. The traditional self-descriptions of objectives (such as growth, income, or balanced) are either too vague or too subjective to be useful. It is thus important to develop methods that can characterize mutual fund investment styles based on publicly available historical return information.

Return-based style classification is widely used in the investment management industry. The procedure regresses a fund's return on the returns to a variety of equity classes (or indexes). The estimated coefficients represent the mutual fund's style with respect to each of the indices. Such a classification thus identifies the fund's style signature in the time series of its returns. Models vary with respect to the choice of style indices. Carhart (1997) uses the Fama-French (1992, 1993) factors plus momentum as style indices. Connor and Korajczyk (1991) use style indices based on statistical factors extracted from stock returns using principal components. Sharpe (1992) proposes an asset allocation approach to yield a fund's effective asset mix. The Fama-French and Sharpe approaches are compared in Chan et al. (2000).

Previous methods have focused on the conditional *expectation* of a fund's return distribution. Fund managers however adopt criteria for stock selection (identifying underpriced securities, seeking growth potential, following past price trends, and so on) that affect the fund's entire return distribution. Hence, as an enhancement to the return-based approach we propose using quantile regression to identify patterns in the fund's entire return distribution. Examples are presented that illustrate how the quantile approach extracts additional information about portfolio style. The added detail permits a finer discrimination of management practices that can be used for style classification and performance evaluation.

2. Return based models: Expectation and quantiles

We restrict attention to domestic equity funds. While many style dimensions might be used, investment managers in practice tend to break the domestic equity investment universe into four classes: large-capitalization or small-capitalization growth stocks, and large-capitalization or small-capitalization value stocks.[1] Classification is based on a regression of a fund's return against the returns to the equity classes. In our implementation we use four equity classes: the Russell 1000 value index, the Russell 1000 growth index, the Russell 2000 value index, and the Russell 2000 growth index. The Russell indices measure returns to portfolios of the largest 1000 and the next largest 2000 stocks, where each size portfolio is further subdivided into growth and value categories depending on relative earnings and book value.

Listed below is the classification of the S&P500 portfolio for December 1997. The coefficients are least squares estimates based on S&P500 monthly returns for the prior 60 months (with the Russell index returns as explanatory variables).

S&P500 Attribution
December 1997

	Large Growth	Large Value	Small Growth	Small Value
Estimate	.57	.48	−.06	.01
Standard Error	.02	.03	.02	.03

The regression results based on returns confirms what is known about portfolio composition, namely that the S&P500 consists of large stocks with a 50-50 split between Growth and Value.[2]

The corresponding classification for the actively managed Fidelity Magellan Fund is shown below. These coefficients indicate a tilt toward Large-Value, but with significant portions in Large and Small Growth.

Magellan Attribution
December 1997

	Large Growth	Large Value	Small Growth	Small Value
Estimate	.14	.69	.21	−.03
Standard Error	.15	.20	.11	.20

The classification model implicit in these estimates is the conditional *expectation* model, $E(r_t|\boldsymbol{x_t}) = \boldsymbol{x_t \beta}$, or

$$E(r_t|\boldsymbol{x}_t) = \boldsymbol{x_t \beta} = \alpha + \beta_{LG} r1000g_t + \beta_{LV} r1000v_t + \beta_{SG} r2000g_t + \beta_{SV} r2000v_t$$

---

[1] For example, most pension plan sponsors use these categories when they screen investment managers. Recently another class, mid-capitalization stocks, has emerged.

[2] Sharpe (1992) considers unconstrained as well as estimates constrained to be nonnegative and sum to one. The latter provide an estimate of a long-only, passive allocation that best matches actual returns. For the equity-only portfolios that we will be considering, the two estimates are usually similar, and we consider unconstrained estimates. Constrained estimates however are readily accommodated by quantile regression. Indeed, it was precisely the difficulty of adding linear constraints to least squares as compared with least absolute deviations that resulted in one of the earliest applications of median regression; see Arrow et al. (1959).

where r_t is returns during month t, $\boldsymbol{x_t} = [1|\boldsymbol{z_t}]$ denotes the vector containing a constant for the intercept and a subvector z_t for the returns to the style indices, and $\boldsymbol{\beta}$ is a vector whose components measure the sensitivity of expected returns to factor exposures. The model says that, other things equal, a β coefficient represents the impact of a change in factor returns on the expected returns, for example,

$$\beta_{LG} = \partial E(r_t)/\partial r1000g.$$

As a consequence, style classification is determined by an estimate of how style exposure influences *expected* returns.

To complement the expectation-based classification, we consider the way style affects other parts of the return distribution. The model implicit in this classification is a linear conditional quantile model. The linear expectation model with iid errors is subsumed as a special case, but the conditional quantile model permits considerably greater flexibility. The linear conditional quantile model posits that the θ^{th} quantile of r_t is a linear function of the portfolio's style exposure, $Q(\theta|x_t) = \boldsymbol{x_t}\boldsymbol{\beta}(\boldsymbol{\theta})$, where as above $\boldsymbol{x_t}$ includes the Russell indices and an intercept.

In contrast to the expectation model, these α and β coefficients are indexed by θ. If the β's are constant, $\beta(\theta) = \beta$, $0 \leq \theta \leq 1$, then the model specializes to the conditional expectation model with homoscedastic errors, and $\alpha(\theta)$ reduces to the quantile function of the error term. When $\beta(\theta)$ varies with θ, the model specifies a form of heteroscedasticity in which the conditional return distribution depends on style. In particular, compared to the expectation model, the quantile model says,

$$\beta_{LG}(\theta) = \partial Q(\theta|x_t)/\partial r1000g$$

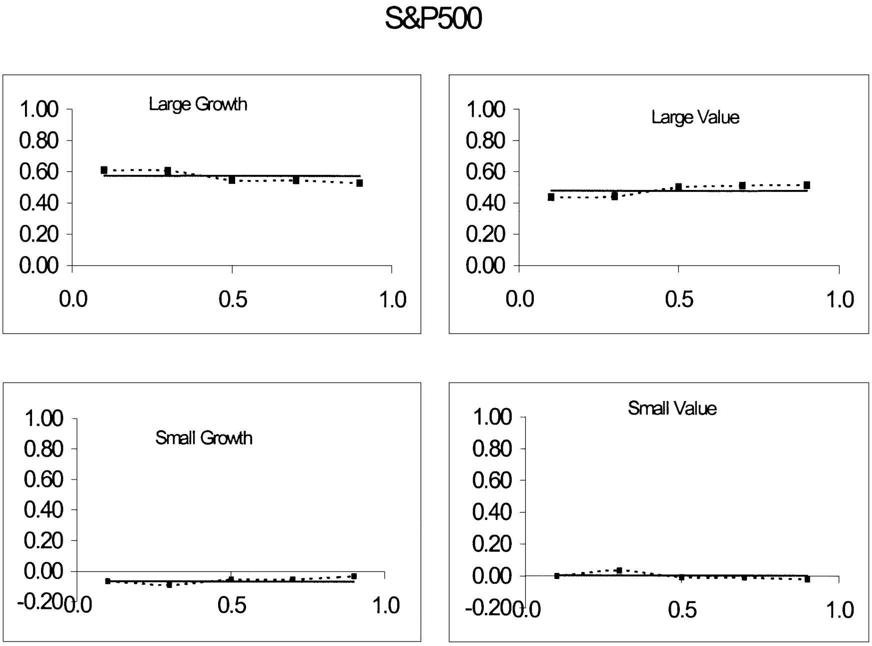
or, other things equal, a one unit change in returns of Large-Growth returns leads to a $\beta_{LG}(\theta)$ change in the θ^{th} quantile of the return distribution. These coefficients are allowed to vary by quantile and differ from the mean coefficient[3].

The quantile regression model is estimated using regression quantiles; see Koenker and Bassett (1978); for discussion of quantile models see Koenker (1982). The quantile regression estimates identify how style indices affect non-central parts of the return distribution. It is this feature of the quantile regression that we use to classify portfolios. Using both regression quantiles and least squares provides information regarding the impact of factor exposures at all parts of the return distribution.

3. Quantile style

The table and charts in Figure 1 show quantile estimates for the S&P500 for December 1997. The horizontal lines are the least squares estimates and the

[3] Attention will be restricted here to the simplest specification of the quantile model, but modifications and enhancements of the model can be readily implemented. For example, in the spirit of allowing "up and down betas", the quantiles can be allowed to differ depending on the sign of factor returns.



S&P500

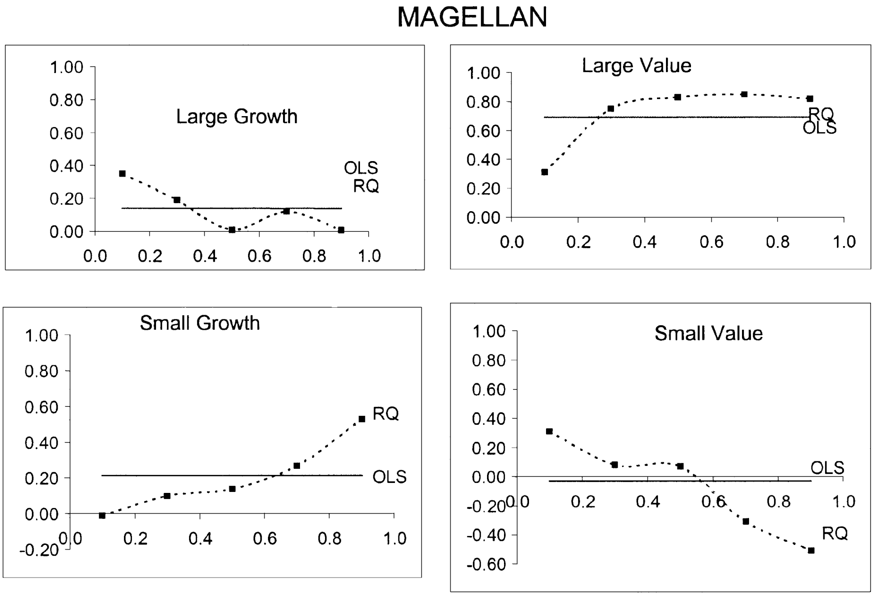
	OLS	$Q(.1)$	$Q(.3)$	$Q(.5)$	$Q(.7)$	$Q(.9)$
Large Growth	0.57	0.61	0.61	0.55	0.55	0.53
	0.02	0.04	0.04	0.03	0.04	0.05
Large Value	0.48	0.44	0.44	0.5	0.52	0.52
	0.03	0.06	0.05	0.04	0.05	0.08
Small Growth	−0.06	−0.06	−0.09	−0.05	−0.05	−0.03
	0.02	0.02	0.03	0.03	0.03	0.04
Small Value	0.01	0	0.04	−0.01	−0.01	−0.02
	0.03	0.05	0.05	0.05	0.06	0.08
Constant	0.01	−0.24	−0.14	0	0.13	0.31
	0.04	0.05	0.04	0.06	0.06	0.08

Notes: Standard errors are listed below the estimates. The standard errors for the quantile regressions were computed using STATA and are based 100 bootstrap replications. All coefficients are for December 1997 and are based on monthly data for the previous 60 months.

Fig. 1. S&P500 Style: OLS and Quantile Regression Estimates

dashed lines correspond to the regression quantiles. For the S&P500 the standard and quantile classifications are nearly identical. This means that the impact of the various style indices is about the same at the expected value and other parts of the return distribution. All the estimates indicate that returns are affected by Large stocks with a 50-50 split between Growth and Value.

Figure 2 shows results for the Magellan fund. In this case quantile regression quantile and least squares estimates differ. At the expected value estimated by OLS, the portfolio has an important Large-Value tilt (.69) and



Magellan

	OLS	$Q(.1)$	$Q(.3)$	$Q(.5)$	$Q(.7)$	$Q(.9)$
Large Growth	0.14	0.35	0.19	0.01	0.12	0.01
	0.15	0.31	0.22	0.16	0.2	0.22
Large Value	0.69	0.31	0.75	0.83	0.85	0.82
	0.2	0.38	0.3	0.25	0.3	0.36
Small Growth	0.21	−0.01	0.1	0.14	0.27	0.53
	0.11	0.15	0.16	0.17	0.17	0.15
Small Value	−0.03	0.31	0.08	0.07	−0.31	−0.51
	0.2	0.31	0.27	0.29	0.32	0.35
Constant	−0.05	−1.9	−1.11	−0.3	0.89	2.31
	0.25	0.39	0.27	0.38	0.4	0.57

Notes: Standard errors are listed below the estimates. The standard errors for the quantile regressions were computed using STATA and are based 100 bootstrap replications. All coefficients are for December 1997 and are based on monthly data for the previous 60 months.

Fig. 2. Magellan Fund Style: OLS and Quantile Regression Estimates

otherwise is equally divided between Large-Growth (.14) and Small-Growth (.21). Compared to the mean regression estimates, the median regression estimates are even more strongly tilted toward Large-Value (.83) and there is little weight on Large-Growth.

4. Correct style?

With different values associated with the least squares and regression quantile estimates, it is tempting to ask which estimate best reflects the fund's true style? It depends on what we mean by style. Is style supposed to measure the impact on a portfolio's expected returns, its median returns, or is it best indicated by the way the tails of the return distribution are influenced by factor returns? Given the quantile model, each meaning of "style" can give a different answer.

A better response to the "true" style question is that it is a mistake to think a single measure should describe style. Given the quantile approach, a portfolio's style depends on how a factor influences the entire return *distribution*, and this cannot be described by a single number. Indeed the single number given by mean regression can obscure important results. The conditional expectation represents the average of what can be potentially very different and important quantile effects. This is evident in the above Small Value estimates for the Magellan fund where large positive and negative impacts in the tails cancel each other leaving barely the hint of an effect at the mean. The different answers given by the least squares and regression quantile estimates are not inconsistent because factors can have different impacts at different parts of the return distribution. All of the estimates are informative and useful. The least squares estimate attributes .69 to Large-Value, while $Q(.5)$ says the coefficient is .82. Insisting on "the" portfolio style presumes that the impact at the mean and median must the same. The focus on the entire return distribution means that least squares and regression quantiles together give a more detailed picture of the relation between returns and factor exposure.

5. Magellan style over time

Regression quantiles also complement analysis of changing portfolio styles. Figures 3 and 4 depict least squares and selected regression quantile estimates for the Magellan fund going back to 1983. Figure 3 shows the relative size of the portfolio, while Figure 4 shows how style has changed in the Value-Growth dimension. Estimates are based on a rolling window of the previous 60 months of returns. The relative style regarding "size" is estimated by the difference between the large coefficients and the small coefficients; $(\beta_{LV} + \beta_{LG}) - (\beta_{SV} + \beta_{SG})$, while the relative Value/Grow style is measured by the difference in the Value and Growth coefficients; $(\beta_{LV} + \beta_{SV}) - (\beta_{LG} - \beta_{SG})$. Shown in the Figures are the OLS estimate, denoted $E()$, and the .3, .5, and .7 regression quantiles. The figure also shows the tenure of the Magellan Fund's managers.

Figure 3 shows that the relative size of the portfolio has increased since the early 1980s. Both least squares and the regression quantiles, which have tended generally to follow one another, indicate this move toward larger capitalization stocks. One exception to the comovement of the estimates was the period from 1988 to mid 1991 when $Q(.7)$ was larger than the other estimates. This foreshadowed a drift to a larger cap style throughout the 1990s. (The value for $Q(.7)$ means the upper tail of the return distribution was responding more strongly to the large cap indices than the other parts of the return distribution). The most recent period also shows large differences between the various estimates.

Figure 4 depicts movements of the fund in Value/Growth space. Until the start of 1992 all of the estimates indicated a neutral to slightly Growth tilt in

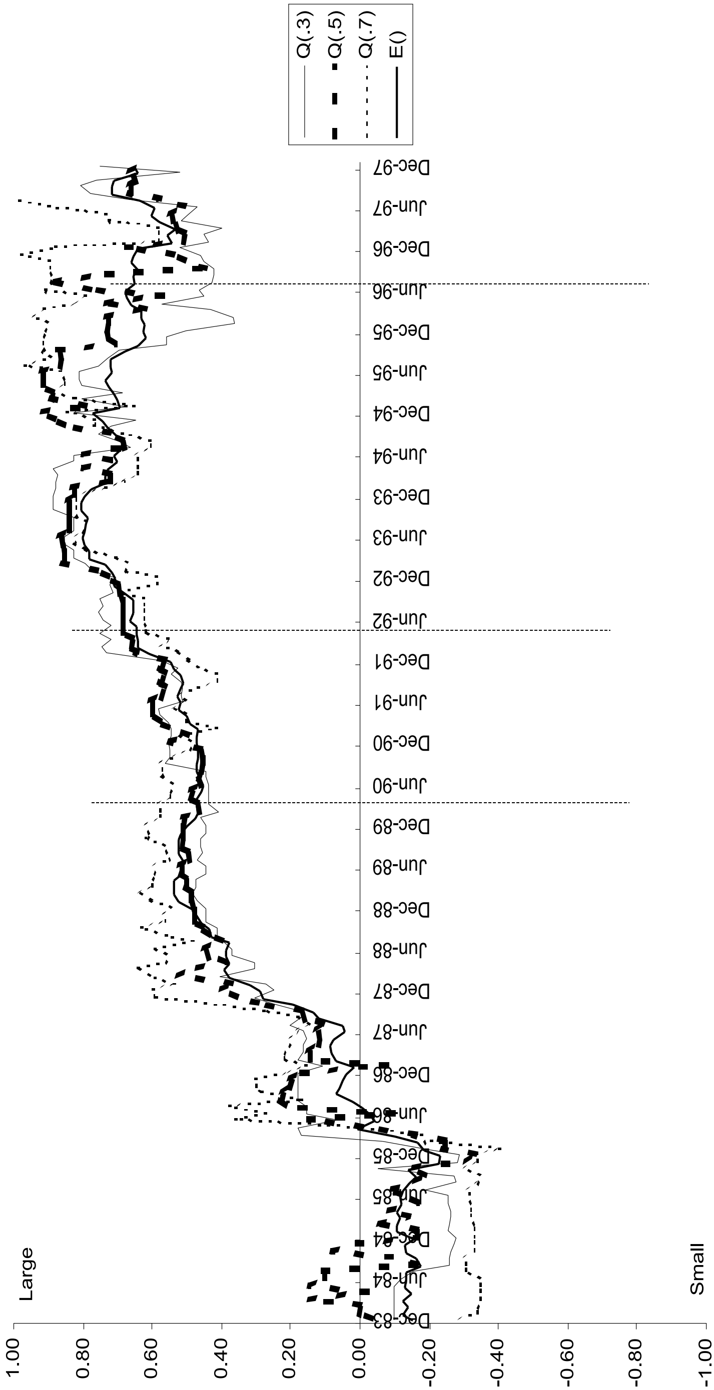


Fig. 3. Magellan Fund: OLS and Quantile Size Estimates
The dashed lines indicate fund managers: Peter Lynch, May '77–May'90: Morris Smith, June '90–June'92: Jeff Vinik July '92–June'96: Bob Stansky June '96–

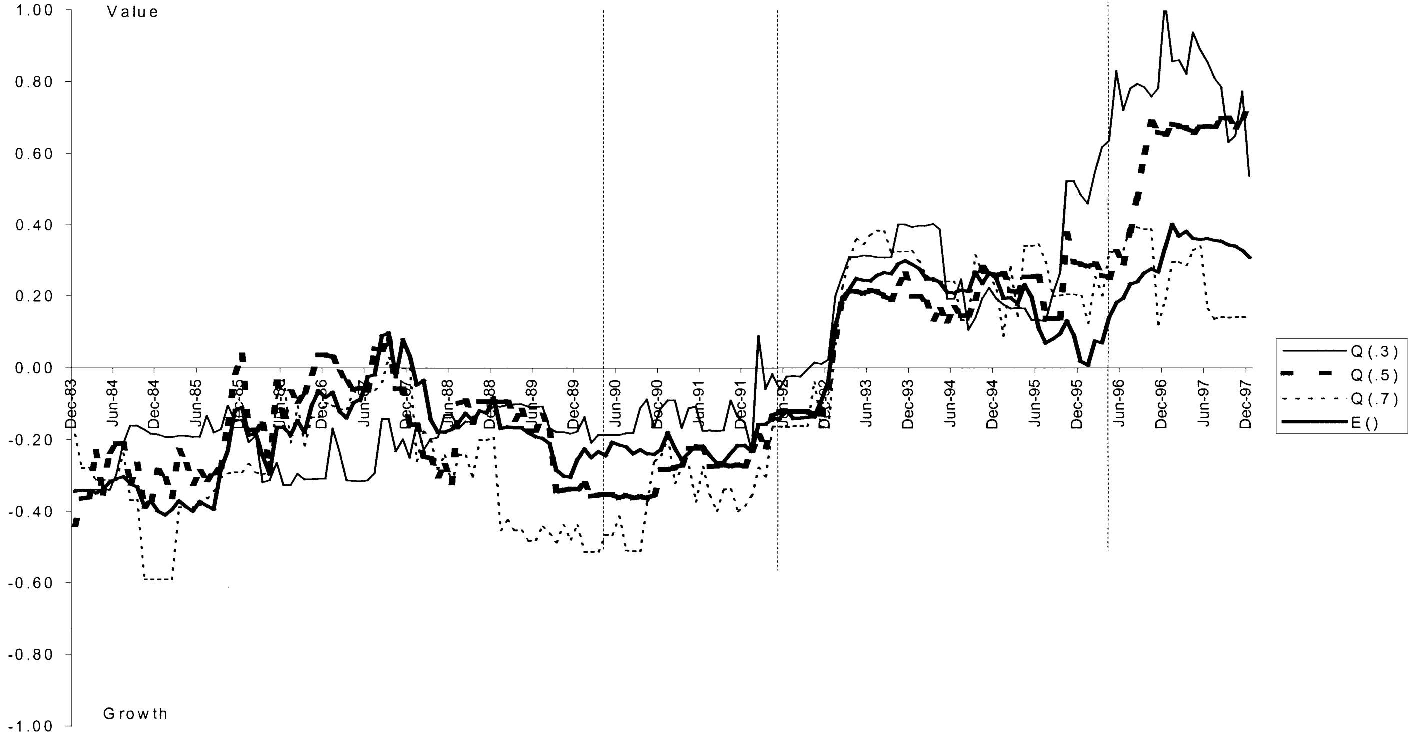


Fig. 4. Magellan Fund: OLS and Quantile Value/Growth Estimates
The dashed lines indicate fund managers: Peter Lynch, May ’77–May’90: Morris Smith, June ’90–June’92: Jeff Vinik July ’92–June’96: Bob Stansky June ’96–

the portfolio. Since 1992 when Jeff Vinik became manager the fund has shifted toward Value though the estimates differ regarding the extent of the shift. The gaps between the estimates at the end of the period correspond to the estimates reported previously in Figure 2 where the relative Value/Growth measured at the expectation was $E(\ ) = .31 = (.69 - .03) \times (.14 + .21)$, while at the median the exposure to value was much greater, $Q(.5) = .75 = (.83 + .07) - (.01 + .14)$. By the end of 1997, the median (compared to the mean) was much more sensitive to Value returns.

6. Performance evaluation using quantiles

One of the applications of style classification is in assessing a manager's skill (or luck), while controlling for the portfolio's overall style. Forecasted returns based on style indices provide a benchmark for determining manager performance relative to a passively managed style portfolio. When estimated by a single parameter, performance is measured relative to only expected returns, absent the possibility that style influences the entire conditional return distribution. This contrasts with the quantile approach where performance evaluation requires a more comprehensive account of the way style influences returns.

Figures 5 and 6 compare actual performance of the Magellan fund for each month of 1997 holding style fixed at its estimated value at the start of the year. The first set of estimates in Figure 5 is based on least squares. The forecast of the return distribution for each month in 1997 is estimated by, $N(x_i\hat{\beta}, \hat{\sigma})$, a normal random variable with mean $x_i\hat{\beta}$, and $\hat{\sigma}$ the standard error of the mean regression. The forecast uses actual monthly returns of the indices, x_i, during 1997 and estimated style coefficients $\hat{\beta}$ estimated as of December 1996. The estimate, $\hat{\sigma} = 1.55$, is the standard error of the residual that determines the dispersion of the forecasted distribution. The points depicted in the figure are (i) the .9 and .1 quantiles of the normal distribution, (ii) the mean, $x_i\hat{\beta}$, and (iii) the actual return. Assuming style remained constant, the differences between actual and forecasted values are due to non-style influences such as skill or luck.

Comparison of the actual returns to the forecasted distribution of returns provides an indication of performance, holding style constant. Notice that the forecasted return distribution varies only by location with the dispersion remaining constant. This is indicted in the Figure by the constant difference between the .9 and .1 quantiles for each month.

Figure 6 presents the same picture based on quantile regression. The monthly forecasts are based on, $\hat{Q}(\theta) = x_i\hat{\beta}(\theta)$ where x_i is monthly Russell returns, and $\hat{\beta}(\theta)$ are the regression quantiles estimated as of the start of the year.

With the analysis based on quantile regression the entire forecasted return distribution changes each month according to realizations of the style indices. This is seen by the varying spread between the .9 and .1 quantiles. In February, for example, the forecast interval is very narrow and actual returns fall below $Q(.1)$ (thus signaling a poor month of active management), while in May the interval is very wide and returns fall near the median as well as OLS forecast.

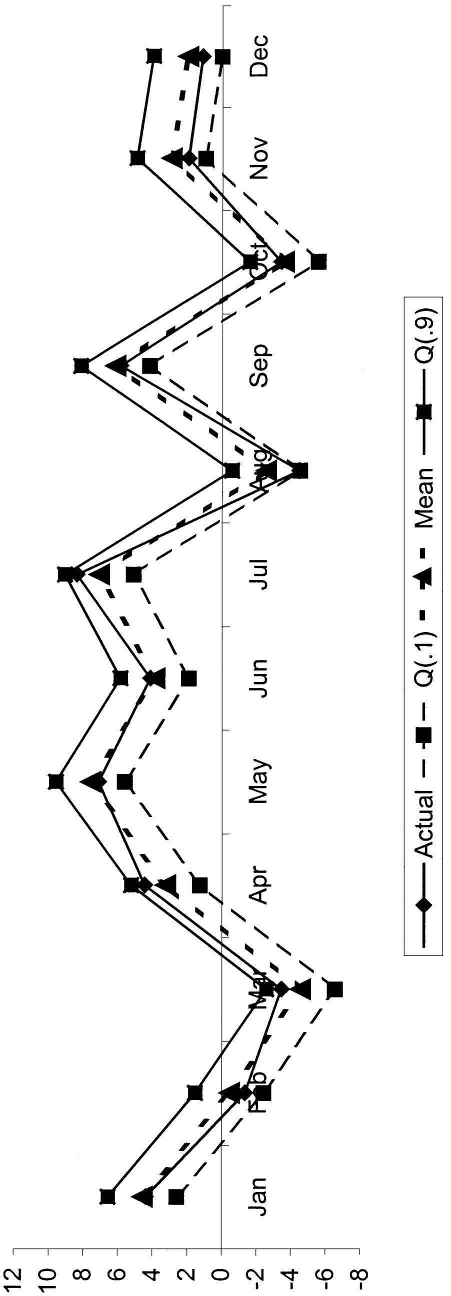


Fig. 5. Forecasted Quantile Returns based on OLS for Magellan in 1997

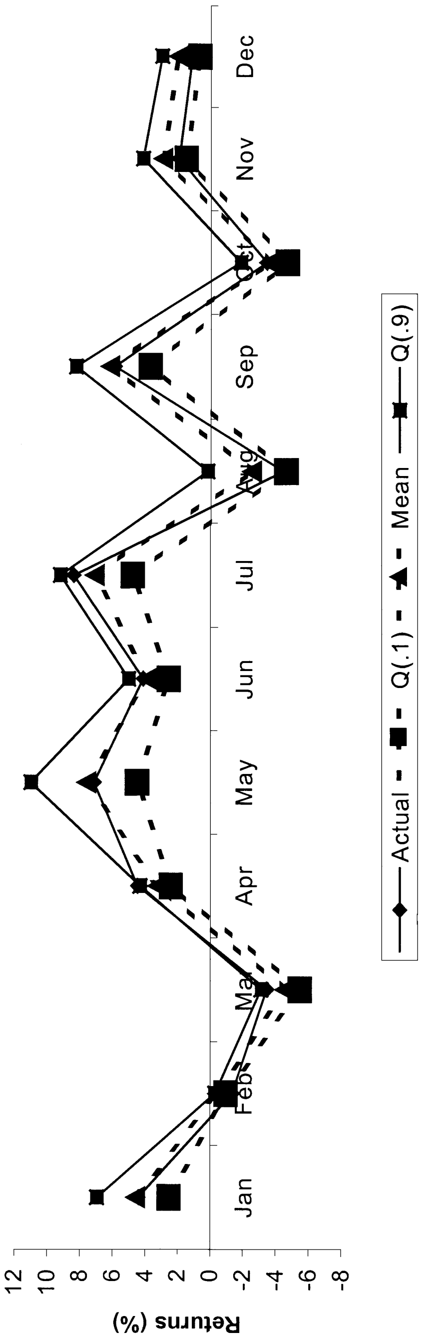


Fig. 6. Forecasted Quantile Returns based on Quantile Regression for Magellan in 1997

7. Conclusions

Quantile regression is proposed as an addition to the style classification toolkit. The quantile approach complements least squares by identifying the impact of style on the conditional return distribution at places other than the expected value. The method does not require additional data and is easy to implement. The illustrations presented in the paper show how the conditional return distribution can respond to factors in different ways at alternative parts of the return distribution.

Among topics for additional research is consideration of alternative specifications of the quantile model; for example, allowing quantiles to depend on market direction. Practical applications of quantile regression to the large universe of existing funds will indicate the extent to which it enhances returns-based classification of portfolio style.

References

Arrow K, Hoffenberg M (with Markowitz H, Shephard R) (1959) A time series analysis of inter-industry demands. North Holland

Carhart MM (1997) On persistence in mutual fund performance. Journal of Finance 52:57–82

Chan L, Chen H-L, Lakonishok J (2000) On mutual fund investment styles. Forthcoming Review of Financial Studies

Fama EF, French KR (1992) The cross section of expected returns. Journal of Finance 47:427–465

Fama EF, French KR (1993) Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33:3–56

Gruber MJ (1996) Another puzzle: The growth in actively managed mutual funds. Journal of Finance 51:783–810

Ippolito RA (1992) Consumer reaction to measures of poor quality: Evidence from the mutual fund industry. Journal of Law and Economics 35:45–70

Koenker R (1982) Robust methods in econometrics. Econometric Reviews 1:213–215

Koenker R, Bassett G (1978) Regression quantiles. Econometrica 46:33–50

Sharpe WF (1992) Asset allocation: Management style and performance measurement. Journal of Portfolio Management (Winter), 7–19

Sirri ER, Tufano P (1998) Costly search and mutual fund flows. Journal of Finance 53:1589–1622