COMPARING FINANCIAL PORTFOLIOS STYLE THROUGH QUANTILE REGRESSION[1]

Laura Attardi

Dipartimento di Progettazione Aeronautica, Università di Napoli Federico II
attardi@unina.it

Domenico Vistocco

Dipartimento di Scienze Economiche, Università di Cassino
vistocco@unicas.it

Abstract

The main goal of style analysis is to assess and compare the performance of different financial products. Style analysis models decompose the portfolio performance with respect to a set of known indexes (constituents). The model estimates the quotas of such indexes in the portfolio, with the aim to separate out their attribution to return. The estimated compositions are interpretable in terms of sensitivity of portfolio expected returns to constituents returns: the style of the products is determined by the way constituents exposure influences expected returns.

The classical model is based on a least squares constrained regression model (Sharpe 1992). Model constraints allow the coefficients to be exhaustive and non negative and to interpret the estimated coefficients in terms of composition quotas.

Quantile regression, as introduced by Koenker and Basset (Koenker & Basset 1978), may be viewed as an extension of classical least squares estimation of conditional mean models for conditional quantile functions. It then offers an useful look at the style analysis problem from a different point of view, allowing to obtain information on the entire returns distribution of analyzed products. Through these additional features, a more detailed comparison of the financial products is then obtainable, by combining classical model results and quantile regression results.

Keywords: Style analysis, Performance attribution, Constrained linear regression, Constrained linear quantile regression.

---

[1] *This paper is supported by MIUR grant 40% “Metodi statistici multivariati e di visualizzazione per l’analisi, la sintesi e la valutazione di indicatori di performance” (Resp.: Professor Maria Rosaria D’Esposito).*

1. INTRODUCTION

The aim of performance measurement is to assess and compare the performance of different portfolios (Feibel 2003). Basically this task is accomplished by exploiting past returns, as typically there is no other information available to external investors: the real composition of a financial portfolio is unknown but to the portfolio manager.

In order to assess the performance, a typical approach benchmarks a portfolio against a market index or a set of market indexes (Siegel 2003). Indexes are published by many different organizations, including financial data publishers, pension consultants, brokerage firms. Some of the provider publish complete families of indexes that represent the performance of the global reference market as well as the performances of the major asset classes.

Style analysis, as originally proposed by (Sharpe 1992), constructs a benchmark portfolio from a set of known indexes (for which returns are available) against which to compare the performance of an actively managed portfolio.

It has been widely applied for investment funds analysis (Brown & Goetzmann 1997), (Conversano & Vistocco 2004).

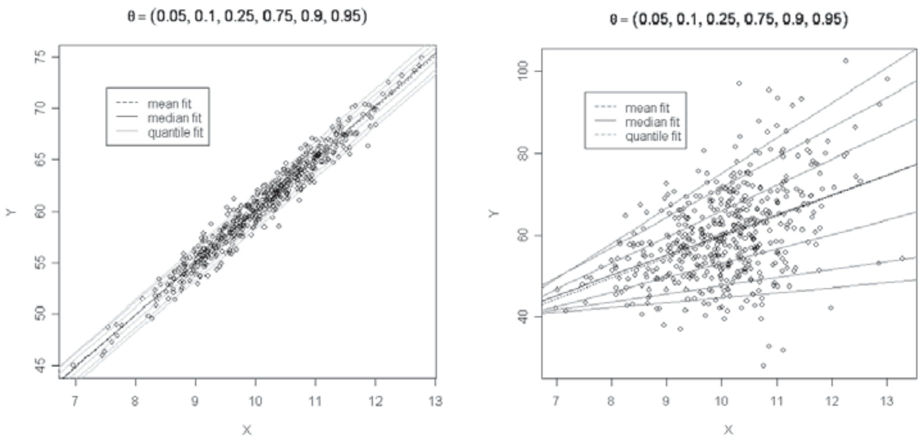
Aim of the paper is to investigate how to exploit the information provided by quantile regression models (for different values of conditional quantiles) in order to discriminate portfolios similar with respect to risk/return but with different internal composition. The quantile regression approach (Basset & Chen 2001) allows indeed to discriminate more active portfolios from more passive ones, stressing then the manager capability in outperforming passive benchmarks.

The paper is organized as follows. In section 2 we introduce briefly quantile regression. Then the style analysis problem is formulated in general terms in section 3.1; the classical least squares approach and the quantile regression approach are presented in section 3.2. The interpretation of the different information provided by least squares and quantile regression coefficients is illustrated through a set of three portfolios in section 3.3 before to go on some concluding remarks.

2. QUANTILE REGRESSION

Quantile regression as introduced in (Koenker & Basset 1978) may be viewed as an extension of classical least squares estimation of conditional mean models to the estimation of a set of conditional quantile functions.

The book of Koenker (Koenker 2005) collects the research results on quantile regression, encompassing models that are linear and nonlinear, parametric and nonparametric and focusing both on the model estimation and on the testing phase.



(a) (b)

Fig. 1: Least squares and quantile regression: *(a)* for an homogeneous model, and *(b)* for an heterogeneous model.

Here only a spot on quantile regression is offered, focusing on the different information provided in case of homogeneous and heterogeneous models.

Starting from two random variables, $X \sim N(10,1)$ and $e \sim N(0,1)$, a sample of $n = 500$ observation is extracted from the model $Y_1 = 10 + 5X + e$ (homogeneous error model) and from the model $Y_2 = 10 + 5X + 0.09X^2 + e$ (heterogeneous error model). Conditional mean fit, conditional median fit and conditional quantiles fits are represented for both the models in figure 1, (for $\theta = \{0.05, 0.1, 0.25, 0.75, 0.9, 0.95\}$, where θ denotes the particular conditional quantile).

For the homogeneous variance regression model – figure 1(a) – the only estimated effect is a change in central tendency of the distribution of Y conditional on the value of X (location model). Quantile regression slope estimates are then for a common parameter, and any deviation among the regression estimates is simply due to sampling variation: an estimate of the rate of change in the means from ordinary least squares regression is also an estimate of the same parameter as for the quantile regression.

When the predictor variable X exerts both a change in means and a change in variance on the distribution of Y (location-scale model), changes in the quantiles of Y cannot be the same for all the quantiles. Figure 1(b) shows that slope estimates differ across quantiles since the variance in Y changes as a function of X.

3. STYLE ANALYSIS

Style analysis is a return-based approach to measuring the performance of a financial portfolio. Nobel–prize Sharpe pioneered it in the early nineties (Sharpe 1992), (Sharpe 1998) by developing an asset class factor model to distinguish the performance of different mutual funds.

Portfolios are invested across a number of different sectors but typically there is no information available to external investors about the detailed choice of assets a particular portfolio holds. Since different sectors perform differently, it is difficult to separate out the contribution to return made by sector choice.

Basic aim of style analysis is to attribute portfolio returns to a set of known indexes for which returns are available. The indexes should represent the different asset classes. Obviously their returns should be available (Tierney 1991).

3.1 THE PROLEM

Style analysis exploits the only available information (portfolio returns and constituents returns) in order to estimate portfolio composition with respect to the used constituents.

Denoting by r_t^{port} and r_t^{const} the returns of the portfolio and of the n constituents, respectively, the equation underlying style analysis using n passive indexes is:

$$r_t^{port} = r_t^{const_1} w_t^{const_1} + \ldots + r_t^{const_n} w_t^{const_n} + e_t$$

where $t = 1, \ldots, T$ denotes the times for which returns are available, $w_t^{const_i}$ is the weight of the constituent i in the portfolio and e_t is the non factor component of the returns.

Rearranging the above equation as difference between the portfolio return and the return due to the indexes gives:

$$e_t = r_t^{port} - \left[r_t^{const_1} w_t^{const_1} + \ldots + r_t^{const_n} w_t^{const_n} \right]$$

where e_t is usually called tracking error.

The weights determined by optimization are called style weights. When combined with the indexes, they form the benchmark portfolio: the bracketed part of the equation can then be interpreted as the return of a weighted portfolio.

The weights $\mathbf{w}^{const}$ are estimated by minimizing the tracking error $e_t, \forall t$. They are constrained to sum up to 100% with the individual weights lying between 0% and 100%, as to allow an interpretation of the estimated coefficients in terms of constituents quotas in composing the portfolio.

The portfolio with optimized weights can be interpreted as a portfolio with the same style as the observed portfolio.

3.2 THE MODELS

Style analysis model regresses portfolio's returns on the returns of a variety of investment classes returns. Such a method thus identifies the portfolio's style in the time series of its returns and of constituents returns. Models can vary with respect to the choice of style indexes as well as with respect to the specific location of the response conditional distribution they are estimating.

The use of least squares (LS) model focuses on the conditional expectation of portfolio's returns distribution. The LS model can be formulated in a compact form as follows:

$$E(\mathbf{r}^{port} \mid \mathbf{r}^{const}) = \mathbf{r}^{const}\mathbf{w}^{const}$$

with constraints:

1. $\mathbf{w}^{const} \geq 0$
2. $\mathbf{1}^T\mathbf{w}^{const} = 1$

In the classical regression model, the $w_t^{const_i}$ coefficient represents the impact of a change in factor returns on the portfolio expected returns:

$$w^{const_i} = \frac{\partial E(r^{port})}{\partial r^{const_i}}$$

Using LS model, portfolio's style is determined estimating the style exposure influence on expected returns.

Extracting information at other places other than the expected value should provide useful insights as the style exposure could affect returns in different ways at different locations of the portfolio returns distribution.

Therefore quantile regression (QR) can be used as a complement to standard analysis, allowing discrimination among portfolios that would be otherwise judged equivalent using only conditional expectation.

The QR model for a given conditional quantile θ follows:

$$Q_\theta(\mathbf{r}^{port} \mid \mathbf{r}^{const}) = \mathbf{r}^{const}\mathbf{w}^{const}(\theta)$$

with constraints:

1. $\mathbf{w}^{const}(\theta) \geq 0$
2. $\mathbf{1}^T\mathbf{w}^{const}(\theta) = 1$

where: $(0 < \theta < 1)$

In a similar way as for the LS model, $w^{const_i}(\theta)$ coefficient of the QR model can be interpreted as the rate of change of the θ^{th} quantile of the portfolio returns distribution for an one unit change in returns of the i^{th} constituent returns:

$$w^{const_i}(\theta) = \frac{\partial Q_\theta(r^{port})}{\partial r^{const_i}}$$

The use of quantile regression offers then a more complete view of relationships among portfolio returns and constituents returns.

3.3 A CASE STUDY

In order to illustrate how to exploit the information provided by LS and QR models, an application on three bond portfolios follows. The portfolios have been obtained as a combination of G7 Merril Lynch indexes: they consist of indexes tracking the performance of the oustanding public local currency debt of G7 sovereign issuers. In particular, they refer to Canadian, French, German, Italian, Japanese, UK and US sovereign bonds issued in their respective domestic markets (Merril Lynch 2006).

Weekly data ($T = 209$) were used from 7 April 2001 to 23 February 2005. The three portfolios are similar with respect to risk/return behaviour – figure 2(a) – but have different levels of assets turnover. Portfolio p1 – figure 2(b) – is completely passive: the seven constituents have the same quota in the portfolio along the considered time interval. The composition of the portfolio p2 – figure 2(c) – is slightly different, allowing some deviations from the uniform composition. Portfolio p3 – figure 2(d) – is an active portfolio: the weights have been obtained using a classic optimization procedure (Elton & Gruber 1995).

Summary statistics for the three portfolios are contained in tables 1, 2 and 3. The first column of each table refers to portfolio returns while the others contains summary indicators for constituents weights.

Tab. 1: Summary statistics for portfolio p1.

	p1	CAN	FRA	GER	ITA	JAP	UK	USA
min	-0.0291	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429
Q1	-0.0063	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429
median	-0.0008	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429
mean	-0.0008	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429
Q3	0.0047	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429
max	0.0365	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429	0.1429
σ	0.0104	0	0	0	0	0	0	0

Tab. 2: Summary statistics for portfolio p2.

	p2	CAN	FRA	GER	ITA	JAP	UK	USA
min	-0.0286	0.132	0.1343	0.1263	0.1398	0.1373	0.1115	0.1418
Q1	-0.0061	0.1394	0.1429	0.1307	0.1462	0.1438	0.1161	0.1472
median	-0.0008	0.1442	0.1445	0.1335	0.148	0.1488	0.1201	0.1563
mean	-0.0007	0.1431	0.1462	0.135	0.1494	0.1491	0.1234	0.1538
Q3	0.0047	0.1479	0.1509	0.1397	0.1549	0.1547	0.1325	0.1579
max	0.0366	0.1556	0.1560	0.1504	0.1585	0.1647	0.1429	0.1657
σ	0.0103	0.0062	0.0055	0.0059	0.0049	0.0064	0.0087	0.0063

Tab. 3: Summary statistics for portfolio p3.

	p3	CAN	FRA	GER	ITA	JAP	UK	USA
min	-0.0258	0.0788	0.0834	0.0316	0.1111	0.0524	0.0159	0.1333
Q1	-0.0061	0.1267	0.1032	0.0412	0.1425	0.0882	0.0209	0.1957
median	-0.0007	0.1578	0.1376	0.0537	0.1652	0.1149	0.0287	0.3310
mean	-0.0004	0.1565	0.1329	0.0755	0.1717	0.1238	0.0478	0.2919
Q3	0.0046	0.1961	0.1563	0.1147	0.2079	0.1478	0.0813	0.3543
max	0.0374	0.2644	0.1829	0.2004	0.2444	0.3120	0.1429	0.4251
σ	0.0100	0.0483	0.0282	0.0446	0.0346	0.0507	0.0341	0.0865

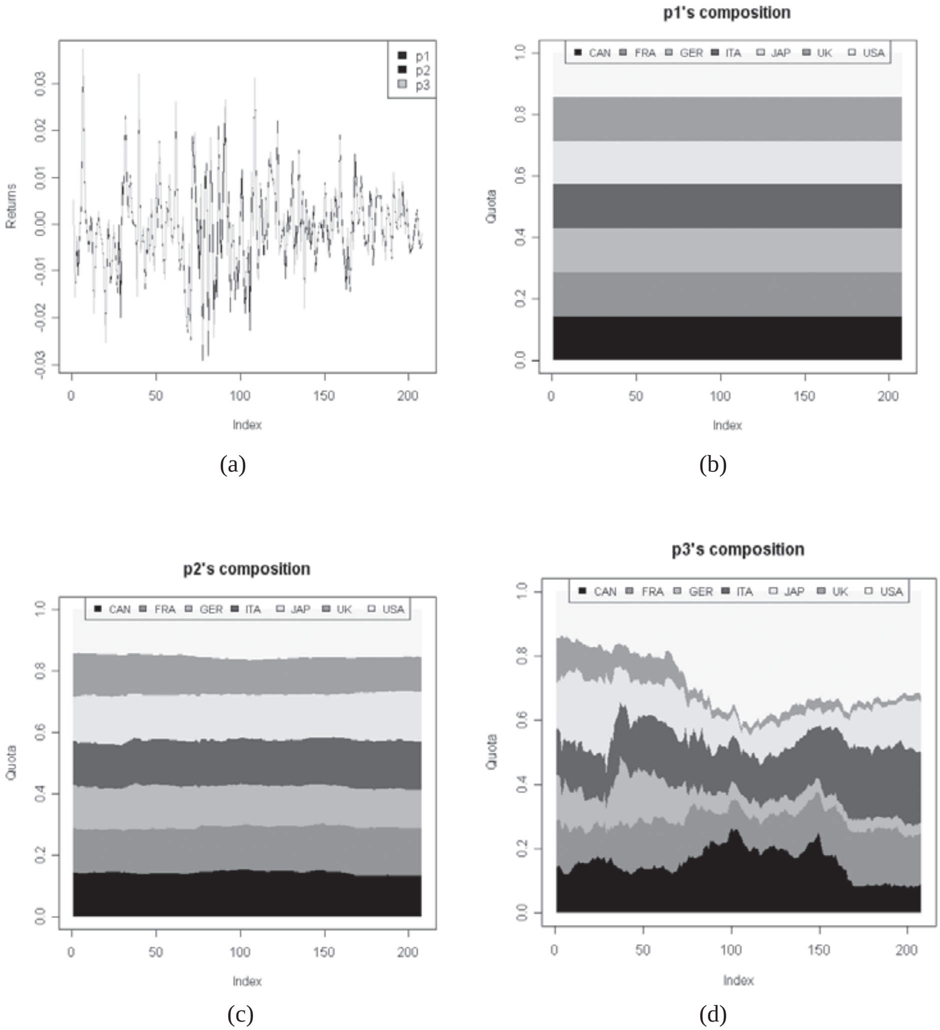


Fig. 2: The three used portfolios: *(a)* p1's, p2' s and p3' s returns, and *(b)* p1's real composition, *(c)* p2's real composition, *(d)* p3's real composition.

Tables 4, 5 and 6 collect results of LS models and of QR models for a set of selected quantiles ($\theta = \{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95, 0.99\}$): the seven portfolio constituents are on the rows while the columns refer to the different models, last column being the difference between maximum and minimum value for each row.

With respect to the first portfolio (table 4), the different models provide the

Tab. 4: LS and QR estimates for portfolio p1 (percentages).

p1	LS	Q(.05)	Q(.10)	Q(.25)	Q(.50)	Q(.75)	Q(.90)	Q(.95)	max - min
CAN	14.29	14.29	14.29	14.29	14.29	14.29	14.29	14.29	0.00
FRA	14.29	14.29	14.29	14.29	14.29	14.29	14.29	14.29	0.00
GER	14.29	14.29	14.29	14.29	14.29	14.29	14.29	14.29	0.00
ITA	14.29	14.29	14.29	14.29	14.29	14.29	14.29	14.29	0.00
JAP	14.29	14.29	14.29	14.29	14.29	14.29	14.29	14.29	0.00
UK	14.29	14.29	14.29	14.29	14.29	14.29	14.29	14.29	0.00
USA	14.29	14.29	14.29	14.29	14.29	14.29	14.29	14.29	0.00
	100	100	100	100	100	100	100	100	

Tab. 5: LS and QR estimates for portfolio p2 (percentages).

p2	LS	Q(.05)	Q(.10)	Q(.25)	Q(.50)	Q(.75)	Q(.90)	Q(.95)	max - min
CAN	14.31	14.36	14.28	14.25	14.19	14.16	14.09	14.08	0.28
FRA	14.32	14.39	14.43	14.25	14.27	14.28	14.33	14.34	0.17
GER	13.58	13.85	13.84	13.81	13.65	13.54	13.45	13.45	0.40
ITA	14.97	14.58	14.64	14.83	15.01	15.22	15.41	15.41	0.83
JAP	14.90	14.90	14.92	14.95	14.94	14.94	14.94	14.94	0.05
UK	12.30	12.65	12.60	12.52	12.29	12.11	11.87	11.86	0.78
USA	15.61	15.26	15.31	15.39	15.64	15.76	15.91	15.92	0.66
	100	100	100	100	100	100	100	100	

same values for each column, confirming the total passive nature of the portfolio.

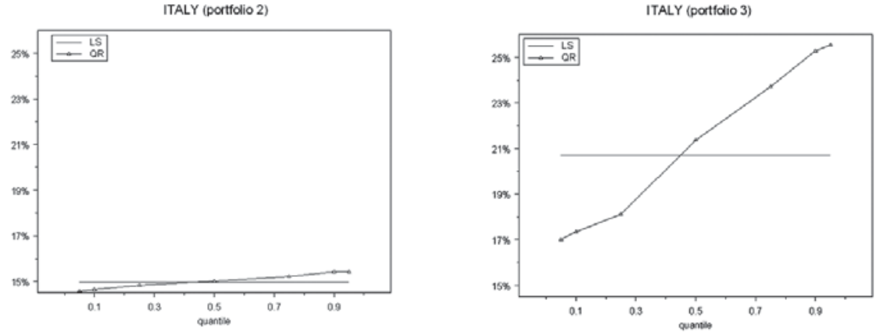
Moving from the second (table 5) to the third (table 6) portfolio, an increasing difference in the estimates provided by the different models is noticeable, providing information on the different level of turnover in the two portfolios along time.

The obtained results can be graphically inspected. The graphical representations allow to appreciate the different behaviour of the sectors in the same portfolio as well as to compare the same sectors in the different portfolios.

Figure 3 displays, with respect to the ITALY sector, the obtained estimates for portfolios p2 and p3. The different conditional quantiles are represented on the X axis while the coefficient values are on the Y axis. The horizontal line is placed at the value of the LS estimate.

Tab. 6: LS and QR estimates for portfolio p3 (percentages).

p3	LS	Q(.05)	Q(.10)	Q(.25)	Q(.50)	Q(.75)	Q(.90)	Q(.95)	max - min
CAN	14.50	16.09	15.63	14.84	13.44	12.34	11.35	11.04	5.05
FRA	12.02	13.66	13.62	13.20	13.57	12.14	12.27	12.26	1.64
GER	6.57	7.99	7.69	6.75	5.12	5.37	4.54	4.39	3.60
ITA	20.69	17.01	17.34	18.11	21.37	23.71	25.27	25.54	8.53
JAP	14.16	11.95	11.92	12.07	11.76	12.13	12.49	12.69	2.40
UK	1.89	4.60	4.73	4.45	2.40	1.63	0.30	0.00	4.73
USA	30.16	28.71	29.07	30.57	32.34	32.69	33.78	34.08	5.37
	100	100	100	100	100	100	100	100	



(a) (b)

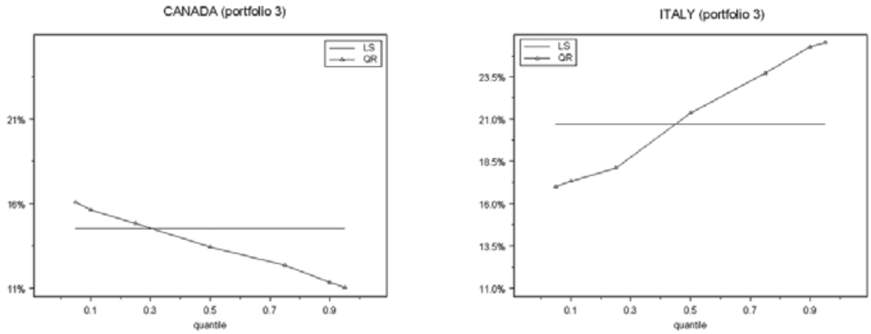
Fig. 3: LS and QR estimates: *(a)* portfolio p2 - ITALY sector, and *(b)* portfolio p3 -ITALY sector.

With respect to the second portfolio, Figure 3(a), the LS and QR information are nearly identical. This means that the impact of the various style indexes is about the same at the expected value and at the other parts of the returns distribution. Figure 3(b) show results for the more active portfolio p3. In this case quantile regression estimates and LS estimates differ: at the expected value estimated through LS, the quota of ITALY in the portfolio is about 21%, while lower estimates for lower quantiles and upper estimates for upper quantiles are evident for QR coefficients.

A different pattern is evident comparing the impact of the Italy and Canada constituents – figure 4 – in portfolio p3. For the Canada constituent the impact at different part of portfolio returns distribution is in some way opposite, as upper estimates correspond to lower quantiles while lower estimates are associated with upper quantiles.

In order to measure the reliability of obtained estimates, classical inferential procedures should be interpreted with caution, due to the imposition of inequality linear constraints (Andrews 1999), (Judge & Takayama 1966). Some general results are available for the normal linear regression model (Geweke 1986) while a bayesian procedure is formulated in (Davis 1978). In the framework of Sharpe style weights, an approximate solution is proposed in (Lobosco & Di Bartolomeo 1997) while a Bayesian approach is engaged in (Christodoulakis 2005).

Figure 5 depicts the LS and QR inferences for the Canada coefficient of portfolio p3: the solid line with filled dots represents the point estimates for 99 distinct



(a) (b)

Fig. 4: LS and QR estimates: *(a)* portfolio p3 - CANADA sector, and *(b)* portfolio p3 - ITALY sector.

conditional quantiles ($\theta = \{0.01, 0.02, \ldots, 0.98, 0.99\}$) with the shaded area depicting a 95% pointwise confidence band. Confidence intervals are computed by the rank inversion method described in (Koenker 2005).

Superimposed on the plot is a line representing the LS estimate, with two dotted lines representing a 95% confidence interval for this coefficient obtained with the approximation proposed in (Lobosco & Di Bartolomeo 1997).

4. SOME CONCLUDING REMARKS

Quantile regression models allow to obtain information on the impact of exposure choices on the entire conditional returns distribution of observed portfolios. The estimated coefficients can then be used to discriminate portfolios according to their assets turnover.

Further investigation to make the best use of quantile regression potential for style analysis should concerns the simulation of a more numerous set of portfolios. Moreover the method should be tested on different kind of portfolios.

Further developments could regard the estimation of conditional returns distribution of the observed portfolios starting from style analysis results. In such a way inferences on portfolios returns could be conducted in a semiparametric framework, without the need to assume the usual conditions that are used in common practice.

Moreover a valuable topic should be the extension of inferential least squares results in presence of linear inequality constraints to the quantile regression version of the same problem.

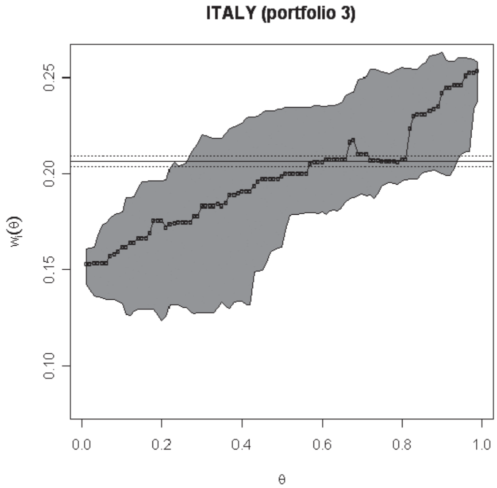


Fig. 5: LS and QR estimates with 95% confidence interval: portfolio p3 - ITALY sector.

REFERENCES

ANDREWS, D.W.K. (1999). Estimation when the Parameter are on the Boundary. *Econometrica,* 67:1341-1383.

BASSET, G.W., CHEN, H.L. (2001). Portfolio Style: Return-Based Attribution Using Quantile Regression. In Fitzenberger B., Koenker R. and Machado J.A.F. (eds), *Economic Applications of Quantile Regression (Studies in Empirical Economics),* Physica-Verlag, 293-305.

BROWN, S., GOETZMANN, W. (1997). Mutual Funds Style Analysis. *The Journal of Financial Economics,* 43:373-399.

CHRISTODOULAKIS, G.A. (2005). Sharpe Style Analysis in the MSCI Sector Portfolios: a Monte Carlo Integration Approach. In Knight J. and Satchell S. (eds), *Linear Factor Models in Finance),* Elsevier, 83-94.

CONVERSANO, C., VISTOCCO, D. (2004). Model Based Visualization of Portfolio Style Analysis. In Antock J. (eds), *Proceedings of the International Conference "COMPSTAT 2004"),* Physica-Verlag, 815-822.

DAVIS, W.W. (1978). Bayesian Analysis of the Linear Model Subject to Linear Inequality Constraints. *Journal of the American Statistical Association,* (78):573-579.

ELTON, E.J., GRUBER, M.J. (1995). *Modern Portfolio Theory and Investment Analysis.* John Wiley & Sons, Chichester.

FEIBEL, B.J. (2003). *Investment Performance Measurement.* Wiley Finance, John Wiley & Sons, Inc., Hoboken, New Yersey.

GEWEKE, J. (1986). Exact Inference in the Inequality Constrained Normal Linear Regression Model. *Journal of Applied Econometrics,* vol. 1, n. 2, (April):127-141.

JUDGE, G.G., TAKAYAMA, T. (1966). Inequality Restrictions in Regression Analysis. *Journal of the American Statistical Association,* (61):166-181.

KOENKER, R., BASSET, G.W. (1978). Regression Quantiles. *Econometrica,* 46:33-50.

KOENKER, R. (2006). quantreg: Quantile Regression. R package version 4.01, http://www.r-project.org.

KOENKER, R. (2005). *Quantile Regression.* Econometric Society Monographs.

LOBOSCO, A., DI BARTOLOMEO, D. (1997). Approximating the Confidence Intervals for Sharpe Style Weights. *Financial Analysts Journal,* July/Aug.:80-85.

MERRIL LYNCH (2006). *MLIndex System.* http://www.mlindex.ml.com.

R DEVELOPMENT CORE TEAM (2006). R: *A Language Environment for Statistical Computing.* http://www.r-project.org, R Foundation for Statistical Computing, Vienna, Austria.

SHARPE,W. (1992). Asset Allocation: Management Styles and Performance Measurement. *The Journal of Portfolio Management.*

SHARPE, W. (1998). Determining a Fund's Effective Asset Mix. *Investment Management Review,* vol. 2, n. 6, (December):59-69.

SIEGEL, L.B. (2003). *Benchmark and Investment Management.* CFA Institute.

TIERNEY, D.E., WINSTON, K. (1991). Using Generic Benchmarks to Present Management Styles. *The Journal of Portfolio Management*, 17:33-36.

UN APPROCCIO BASATO SULLA *REGRESSIONE QUANTILE* PER L'ANALISI DELLO STILE DI GESTIONE DI PORTAFOGLI FINANZIARI

Riassunto

L'analisi dello stile di gestione mira a valutare la performance di portafogli finanziari caratterizzati da strategie di investimento simili. Il metodo si basa sulla costruzione di un portafoglio di riferimento, o portafoglio benchmark, per ciascuno dei prodotti finanziari oggetto di analisi. Tali portafogli benchmark si caratterizzano per una struttura di rischio/rendimento simile a quella dei portafogli reali, rispetto ai quali si differenziano perché è possibile stimarne la composizione interna. Fine ultimo del metodo è quindi scomporre la performance globale dei portafogli per imputarla ai differenti settori di investimento. Il metodo di analisi è stato proposto in letteratura da Sharpe nel 1992 e si basa su un modello di regressione lineare vincolata in cui i vincoli assicurano che i coefficienti di regressione stimati possano essere letti come dati composizionali. L'utilizzo dell'approccio classico alla regressione comporta l'interpretazione dei risultati in termini di in- fluenza della performance di ciascuno dei vari componenti sulla performance attesa del portafoglio. La tecnica della regressione quantile, proposta originariamente da Koenker e Basset nel 1978, offre pertanto una naturale estensione del modello classico attraverso l'analisi dell'effetto che le performance delle varie componenti del portafoglio hanno sui differenti quantili della distribuzione della performance del portafoglio. Attraverso tali informazioni è quindi possibile discriminare portafogli caratterizzati da uno stesso comportamento in termini di rendimento atteso ma con un diverso grado di movimentazione interna rispetto ai differenti settori.