INNOCENZO GASPARINI INSTITUTE FOR ECONOMIC RESEARCH

Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Model Uncertainty, Thick Modelling and the Predictability of Stock Returns

Marco Aiolfi, Carlo Ambrogio Favero

Working Paper n. 221

July 2002

IGIER – Università Bocconi, Via Salasco 5, 20136 Milano –Italy
http://www.igier.uni-bocconi.it

Model Uncertainty, Thick Modelling and the Predictability of Stock Returns.*

Aiolfi M., C.A.Favero
Bocconi University, Bocconi University and CEPR

July 2002

Abstract

Recent financial research has provided evidence on the predictability of asset returns. In this paper we consider the results contained in Pesaran-Timmerman(1995), which provided evidence on predictability over the sample 1959-1992. We show that the extension of the sample to the nineties weakens considerably the statistical and economic significance of the predictability of stock returns based on earlier data.. We propose an extension of their framework, based on the explicit consideration of model uncertainty under rich parameterizations for the predictive models. We propose a novel methodology to deal with model uncertainty based on "thick" modeling, i.e. on considering a multiplicity of predictive models rather than a single predictive model. We show that portfolio allocations based on a thick modelling strategy sistematically overperforms thin modelling.

JEL Classification Numbers: G11, C53

1 Introduction

Recent financial research has provided ample evidence on the predictability of asset returns (see, for example, Keim and Stambaugh,(1986), Campbell and

*Addres for correspondance: Carlo A. Favero, IGIER-Università Bocconi, Via Salasco 5 20124 Milan Italy e-mail: *carlo.favero@uni-bocconi.it*. We are indebted to Alessandro Penati, Guido Tabellini and Francesco Corielli and seminar participants at 'Ente Einaudi for Monetary and Financial Studies' in Rome for comments and suggestions. We would like to thank Allan Timmerman and Hashem Pesaran for usefull comments and discussions.

Shiller(1988a, 1988b) Pesaran and Timmermann(1996), Lander at al.(1997) and, for a survey, Cochrane(2000)), Pesaran and Timmermann(1996,2000) have shown that, net of transactions costs, it could have been exploited by investors in the volatile markets of the 1970s.

Pesaran and Timmermann(1996) consider a time-varying parameterization for the forecasting model to find that the predictive power of various economic factors over stock returns changes through time and tends to vary with the volatility of returns. They apply a 'recursive modelling' approach, according to which at each point in time all the possible forecasting models are estimated and returns are predicted by relying on the best model, chosen on the basis of some given statistical criterion. The dynamic portfolio allocation, based on the signal generated by a time-varying model for asset returns, is shown to over-perform the buy-and-hold strategy over the period 1959-1992. The results obtained for the US are successfully replicated in a recent paper concentrating on the UK evidence, Pesaran and Timmermann (2000).

In this paper we propose a novel methodology which extend the proposal contained in the original paper to deal explicitly with model uncertainty. In the first section of the paper we discuss our proposal to deal with model uncertainty under rich parameterization for the predictive models. We then re-assess the original evidence on the statistical and economic significance of the predictability of stock returns by extending the data-set to the nineties and by evaluating comparatively thin and thick modelling.

2 Recursive modelling: thin or thick ?

Pesaran and Timmermann (1996) consider the problem of an investor allocating his portfolio between a safe asset denominated in dollar and US stocks. The decision on portfolio allocation is then completely determined by the forecast of excess returns on US stock. Their allocation strategy is such that portfolio is always totally allocated into one asset, which is the safe asset if predicted excess returns are negative, and shares if the predicted excess returns are positive. The authors forecast excess US stock returns by concentrating on an established benchmark set of regressors over which they conduct the search for a "satisfactory" predictive model. They focus on modelling the decision in real-time. To this end they implement the recursive modelling approach, according to which at each point in time, t, a search over a base set of observable k regressors is conducted to make one-period ahead forecast. In each period they estimate a set of regression spanned by all the possible permutations of the k regressors. This gives a total of 2^k different models for excess return. Models are estimated recursively, so that as the data-set is expanded by one observation in each period. Therefore a total of $2^k * 396$ models are estimated at each possible period from 1959:12 to 1992:11 to generate a portfolio allocation.

They estimate all the possible specifications of the following forecasting equation:

$$(x_{t+1} - r_{t+1}) = \boldsymbol{\beta}_i' \mathbf{X}_{t,i} + \varepsilon_{t+1,i} \tag{1}$$

where x_{t+1} are the monthly returns on US stocks and r_{t+1} are the monthly returns on the US dollar denominated safe asset (1-month T-bill), $\mathbf{X}_{t,i}$ is the set of regressors, observable at time t, included in the i-th specification $(i = 1, ...2^k)$ for the excess return. The relevant regressors are chosen from a benchmark set containing, the dividend yield YSP_t, the earning-price ratio PE_t, the 1-month T-bill rate $I1_t$ and its lag $I1_{t-1}$, the 12-month T-bill rate $I12_t$ and its lag $I12_{t-1}$, the year-on-year lagged rate of inflation π_{t-1}, the year-on-year lagged change in industrial output ΔIP_{t-1}, and the year-on-year lagged growth rate in the narrow money stock ΔM_{t-1}. A constant is always included and all variables based on macroeconomic indicators are measured by 12-month moving averages to decrease the impact of historical data revisions on the results.

At each sample point the investor computes OLS estimates of the unknown parameters for all possible models, chooses a forecast for excess returns given the predictions of 512 models and maps the forecast into a portfolio allocation by choosing shares if forecast is positive and the safe asset if the forecast is negative.

Pesaran and Timmermann select in each period only one forecast, i.e. that genrated by the best model selected on the basis of a specified selection criteria which weigths goodness of fit against parsimony of the specification(such as adjusted R^2, BIC, Akaike, Schwarz). We follow Granger (2000) and label this approach 'thin' modelling in that the forecast for excess returns and consequently the performance of the asset allocation are described over time by a thin line.

The advantage of this approach is that a process, potentially non-linear, is modeled by applying recursively a selection procedure among linear models. The specification procedure mimics a situation in which variables for predicting returns are chosen in each period from a pool of potentially relevant regressors. This choice fits well the behaviour often observed in financial markets of attributing different emphasis to the same variables in different periods.

Obviously, keeping track of the selected variables helps the reflection on the economic significance of the 'best' regression.

The main limit of thin modelling is that model, or specification, uncertainty is not considered. In each period the information coming from the discarded $2^k - 1$ models is ignored for the forecasting and portfolio allocation exercise. This choice seems to be particularly strong in the light of the results obtained by Bayesian line of research, which stresses the importance of the estimation risk for portfolio allocation (see for example, Barberis,2000, Kandel and Stanbaugh,1996). A natural way to interpret model uncertainty is to refrain from the assumption of the existence of a "true" model and attach instead probabilities to different possible models. This approach has been labelled 'Bayesian Model Averaging', see, for example, Hoeting J.et al.(1999), Raftery et al.(1997), and Avramov (2001). Bayesian methodology reveals the existence of in sample and out of sample predictability of stock returns, even when commonly adopted model selection criteria fail to demonstrate out of sample predictability.

The main difficulty with the application of Bayesian Model Averaging to problems like ours lies with the specification of prior distributions for parameters in all 2^k models of our interest. Recently, Doppelhofer et al. (2000) have proposed an approach labelled 'Bayesian Averaging of Classical Estimates'(BACE) which overcomes the need of specifying priors by combining the averaging of estimates across models, a Bayesian concept, with classical OLS estimation, interpretable in the Bayesian camp as coming from the assumption of diffuse, non-informative, priors.

In practice BACE averages parameters across all models by weigthing them proportionally to the logarithm of the likelihood function corrected for the degrees of freedom, using then a criterion similar to the Schwarz model selection criterion. It is important to note that the consideration of model uncertainty in our context generates potential for averaging at two different levels: averaging across the different predicted excess returns and averaging across the different portfolio choices driven by the excess returns.

The explicit consideration of estimation risks naturally generates 'thick' modelling, where both the prediction of models and the performance of the portfolio allocations over time are described by a thick line to take account of the multiplicity of models estimated. The thickness of the line is a direct reflection of the estimation risk.

Pesaran and Timmermann show that thin modelling allows to over-perform the buy and hold strategy. Re-evaluating their results from a thick modelling perspective raises immediately one question: "*why choose just a model to forecast excess returns?*"

3 A first look at the empirical evidence

We start be replicating[1] the exercise in Pesaran and Timmermann using the same dataset, keeping track of all the forecasts produced by taking into account the 2^k-1 combinations of regressors. We do so by looking at the within sample econometric performance, at the forecasting performance and at the performance of the portfolio allocation. The main feature of the series used in our exercises are described in Table 1.

Figure 1 allows to analyze the within sample econometric performance by reporting the adjusted R^2 for 2^k models estimated recursively. The difference in the selection criterion across different models is small. In fact, turning to predictive performance we find that it is possible to improve on the performance of the best model in terms of R^2 by using the information contained in the $2^k - 1$ models dominated (is many cases marginally) in terms of R^2. The result is most easily shown by using the metric of the sign test proposed by Pesaran-Timmermann(1996). The sign test is based on the proportion of times that the sign of a given variable y_t is correctly predicted in the sample by the sign of the predictor x_t . Under the null hypothesis that x_t has no power in predicting y_t the proportion of times that the sign is correctly predicted has a binomial distribution with known parameters, therefore a test of the null of predictive failure is constructed by comparing the observed proportion of sign correctly predicted with the proportion of sign correctly predicted under the null. Details on the derivation of the statistics and results are reported in Table 2. We report the tests obtained by basing the prediction on the signal of the best model (thin modelling) and on the signal of averages of a multiplicity of models, ranked on the basis of their R^2. The

[1] In fact, we replicate the allocation results in the case of no transaction costs. Transaction costs do not affect the portfolio choice in the original exercise, therefore they do not affect the mapping from forecasting to portfolio allocation, which is the main concern of our paper.

idea of averaging is in line with the observation in Clements and Hendry (2001) who demonstrate that when forecasting time series that are subject to deterministic shifts, the average of a group of forecasts from differently misspecified models can outperform them all. Results are shown in terms of increasing thickness, by reporting first sign test when averages of the top 1 per cent of the models are considered, to increase progressively the thickness of the modelling approach until we average across all 2^k models. The statistics show that the best performance in terms of the sign test is achieved when the average prediction of the best sixty per cent of the models in terms of R^2 is chosen as a predictor for excess returns. Curiously, averaging across all models deliver the same performance with thin modelling.

Lastly, we turn to the performance of the portfolio allocation based on the predictive regression. Figure 2 illustrates the cumulative wealth generated by the portfolio allocation based on the signal of all 2^k models, ranked in terms of their adjusted R^2. The value of the end-of-period wealth is not a decreasing function of the adjusted R^2. In fact, the highest value for the end-of-period wealth is achieved when portfolio is allocated according to the signals of the model ranked about eigthieth in terms of its R^2 and allocating portfolio in terms of the signal of one of the worst models in terms of R^2 generates an higher final wealth than that delivered by the allocation based on thin modelling.

The extension of the sample to the period 1993-2001 delivers a very different scenario: the adjusted R^2 of all models decreases substantially, the PT sign tests for predictive performance are not significant anymore, and the consequently the portfolio allocation performance generate lower wealth than the buy-and-hold strategy.

To our reading these results show that thick modelling has potential, but that refinements in the specification and the modelling selection strategy are called upon by the empirical evidence from the more recent data.

4 Our proposal for thick modelling

In the light of the evidence reported in the previous section we propose extensions of the original methodology both at the stage of model specification and of portfolio allocation.

We shall use thick modelling exclusively at the stage of portfolio allocation. The empirical evidence reported in the previous section shows clearly that the ranking of models in terms of their within sample performance does not match at all the ranking of models in terms of their ex-post forecasting power. This empirical evidence points clearly against BACE using within sample criteria to weight models. Consistently with this evidence, we opted for the selection method proposed by Granger (2000) of using a '... *procedure [which] emphasizes the purpose of the task at hand rather than just using a simple statistical pooling...'* Our task at hand is asset allocation.

4.1 Model specification

At the stage of model specification we consider two issues: the importance of balanced regressions and the optimal choice of the window of observations for estimation purposes.

A regression is balanced when the order of integration of the regressors matches that of the dependent variables. Excess returns are stationary, but not all variables candidate to explain that are stationary. To achieve a balanced regression in this case, cointegration among the included non-stationary variables is needed. As shown by Sims, Stock and Watson(1990) the appropriate stationary linear combinations of non-stationary variables will be naturally selected by the dynamic regression, when all non stationary variables potentially included in a cointegrating relations are included in the model. Therefore, when model selection criteria are applied, one must make sure that such criteria do not lead to exclude any component of the cointegrating vector from the regression. Following Pesaran and Timmermann (2001) we divide variables in focal, labelled A_t and secondary focal, labelled B_t. Focal variables are always included in all models, while the variables in B_t are subject to the selection process . We take these variables as those defining the long-run equilibria for the stock market Following the lead of traditional analysis[2] (Graham and Dodd Security Analysis, 4th edition, 1962,

---

[2]

> *"...Theoretical analysis suggests that both the dividend yield and the earnings yield on common stocks should be strongly affected by changes in the long-term interest rates. It is assumed that many investors are constantly making a choice between stock and bond purchases; as the yield on bonds advances, they would be expected to demand a correspondingly higher return on stocks, and conversely as bond yields decline..."*

The above statement suggests that either the dividend yield or the earnings yield on common stocks could be used p.510) and recent studies (Lander et al. (1997)) we have chosen to construct an equilibrium for the stock market by concentrating on a linear relation between the long term interest rates, R_t, and the logarithm of the earning price ratio, *ep*. Also recent empirical analysis (see Zhou, 1996) finds that stock market movements are closely related to shifts in the slope of the term structure. Such results might be explained by a correlation between the risk premia on long-term bonds and the risk premium on stocks. Therefore, we consider the term spread as a potentially important cointegrating relation. On the basis of this consideration we include in the set of focal variables the yield to maturity on 10-year government bonds (a variable which was not included in the original set of regressors by PT), the log of the earning price ratio and the interest rate on 12-month Treasury Bills, to ensure that the selected model is balanced and includes the two relevant cointegrating vectors. We do not impose any restrictions on the coefficients of the focal variables.

The second important issue at the stage of model selection is the choice of the window of observations for estimation.

In the absence of breaks in the DGP the usual method for estimation and forecasting is to use an expanding window. In this case, by augmenting an already selected sample period with new observations, more efficient estimates of the same fixed coefficients are obtained by using more information as it becomes available. However, if the parameters of the regression model are not believed to be constant over time, a rolling window of observations with a fixed size is frequently used. When a rolling window is used, the natural issue is the choice of its size. This problem has been already observed by Pesaran and Timmermann (1999) who provide an extensive analysis of model instability, structural breaks, and the choice of window observations. In line with their analysis we deal with the problem of window selection by starting from an expanding window, every time a new observation is available we run a backward CUSUM and CUSUM squared test to detect instability in the intercept and/or in the variance. We then keep expanding the window only when the null of no structural break is not rejected. Consider a sample of T observations and the following model:

$$y_{t,T} = \boldsymbol{\beta}^{i'} x^i_{t,T} + u_{t,T} \;\; i = 1, ..., 2^k$$

where $y_{t,T} = (y_t, y_t, y_{t+2}, ..., y_T)$ and $x^i_{t,T} = \left(x^i_t, x^i_{t+1}, x^i_{t+2}, ..., x^i_T\right)$ where $T - t + 1$ is the optimal window and T the last available observation, remember that we are interested in forecasting y_{T+1} given $x_{T+1,}\widehat{\boldsymbol{\beta}}^{i'}$. The problem of the optimal choice of t given model i, can be solved by running a CUSUM test with the order of the observations reversed in time starting from the m-th observation and going back to the first observation available. Once a structural break (either in the mean or in the variance) has been detected, we have found the optimal t. Clearly the optimal t can be the first observation in the sample (in this case we have an expanding window) or any number between 1 and m (flexible rolling window). This procedure allows us to optimally select the observation window[3] for each of the 2^k different models estimated at time t.

4.2 Asset Allocation

We consider three different alternative ways of implementing thick modelling when allocating portfolios. Given the 2^k forecasts for excess returns in each period define α^S and $\left(1-\alpha^S\right)$ to be respectively the weight of stocks and safe asset(short term bills), let $\{y_i\}_{i=1}^{2^k}$ the full set of excess returns forecasts obtained in the previous step, and let $n = \boldsymbol{\omega}'2^k$, where $\boldsymbol{\omega} = [.01, .05, .1, .2, .3, .4, .5, .6, .7, .8, .9, 1]$ is the set of weights, in terms of the percentage of the model ordered according to their adjusted R^2, chosen to build up the appropriate trimmed means of the available forecasts. Then we use the following allocation criteria:

1. *Distribution-Thick-Modelling*: We look at the empirical distribution of the forecasts to apply the following criterion:

Criterion		Weigths
$\left[\frac{\sum_{i=1}^{n_{\omega_j}}(y_i>0)}{n_{\omega_i}}\right]$	> 0.5	$\alpha^S_{\omega_i} = 1, \alpha^B_{\omega_i} = 0$
$\left[\frac{\sum_{i=1}^{n_{\omega_j}}(y_i>0)}{n_{\omega_j}}\right]$	≤ 0.5	$\alpha^S_{\omega_i} = 0, \alpha^B_{\omega_i} = 1$

where $n_{\omega_j}(y_i > 0)$ is the number of models giving a positive prediction for excess returns within j-th class of the trimming grid (For example $n_{\omega_2}(y_i > 0)$ is the number of models in the best 10 per cent of the ranking in term of their adjusted R^2 predicting a positive excess return

[3]We impose that the shortest observation window automatically selected cannot be smaller than 2 or 3 times the dimension of the parameters' vector. So also the minimum observation window is a function of regressors included in each of 2^k different models.

2. *Meta-Thick-Modelling:* We use the same criterion as above, to derive a less aggressive portfolio allocation, in which corner solution are the exceptions rather than the rule:

Weigths
$\alpha_{\omega_i}^S = \left\{ \left[\frac{\sum_{i=1}^{n_{\omega_i}} (y_i > 0)}{n_{\omega_i}} \right] > 0.5 \right\}, \alpha_{\omega_i}^B = \left(1 - \alpha_{\omega_i}^S\right)$

3. *Kernel-Thick-Modelling*: we compute the weigthed average of predictions $\overline{y}$ (with weights based on the relative adjusted-R^2, through a kernel function that penalizes deviations from the best model in terms of R^2 and the bandwidth determined by the number of observations) and then we apply this rule:

Criterion	Weigths
$\overline{y} > 0$	$\alpha_{\omega_i}^S = 1, \alpha_{\omega_i}^B = 0$
$\overline{y} \leq 0$	$\alpha_{\omega_i}^S = 0, \alpha_{\omega_i}^B = 1$

4.3 Empirical Results

Our empirical results are reported in Table 3-4 and Figures 3-10.

Tables 3-4 illustrates the forecasting performance of models specified according different criteria and grouped according different trimming. Trimming are reported by rows and model estimation criteria are reported by columns. We have then five columns labeled respectively *Rec, Roll, Bal, Flex* and *Bal-Flex. Rec* replicates the original model estimation recursive and reports the results based on recursive estimation (expanding window of observations) with no focal variables . *Roll* reports the results based rolling estimation(with fixed window of 60 observations) with no focal variables. *Bal* reports the results based on recursive estimation(expanding window of observations), focal variables are : log of the price-earning ratio, yield-to maturity on long term bonds, yield on 12-month Treasury Bills. *Flex* reports the results based on rolling estimation(with optimally chosen window), no focal variables. *Bal-flex* reports the results based on rolling estimation(with optimally chosen window), focal variables are the log of the price-earning ratio, the yield-to maturity on long term bonds, and the yield on 12-month Treasury Bills. Table 3 reports the results for the sample 1954-1992 while Table 4 reports the results for the sample 1993-2001. The dominance of thick modelling over thin modelling is confirmed for the first sample across all the different columns of Table 3. The comparison of Table 3 with Table 4 confirms that the decrease in predictive power when the sample is extended. Interestingly, the Bal-Flex specification criterion, which was the worst performer in the 1954-1992 sample, dominates all in terms of percentage of correct signs in the nineties. This evidence suggest that combining the selection of focal variables and the selection of the optimal size for the estimation window provides the most robust performance in terms of sign test.

Figure 3-4 allow the evaluation of the performance of different portfolio allocation criteria, by comparing the end-of-period cumulative wealth associated to each of them with the cumulative wealth associated to a buy-and-hold strategy, always allocating the entire portfolio to shares[4]. We report the performance of different trimming criterion for all the model specification criteria for the sample 1960-1992 in Figure 3 and for the sample 1993-2001 in Figure 4. In each Figure the flat line is the end of period wealth of the buy and hold strategy associated to a beginning of period wealth of 100.

In the first part of the sample all econometric based allocation do better than the buy and hold strategy. Thick modelling does improve on thin modelling. In particular some form of distribution thick modelling dominates thin modelling independently of the model specification criteria. The dominance of thick modelling becomes stronger when Balanced and Flexible-Balanced specification criteria are chosen. Although more complicated selection criteria tend to give a weaker over-performance than the simple recursive specification.

In the second part of the sample over-performing the buy and hold strategy becomes much more difficult, however the dominance of thick modelling on thin modelling becomes stronger. More articulates model selection criteria now deliver better results than the simple recursive criterion. The best performance is achieved when the distribution-thick criterion is applied to the best 20 per cent of models in terms of their adjuster R^2.

5 Conclusions

In this paper, we have reassessed the results on the statistical and economic significance of the predictability of stock returns provided by Pesaran and Timmermann(1995) for the US data to propose a novel approach for portfolio allocation based on econometric modelling. We find that the results based

[4]Evaluation has been also conducted in terms of period returns and Sharpe-ratios, results are available upon request.

on the thin modelling approach originally obtained for the sample 1960-1992 are considerably weakened when the sample is extend to 2001.

We then show that the incorporation of model uncertainty substantially improves the performance of econometric based portfolio allocation

The portfolio allocation based on a strategy giving weights to a number of models rather than to just one model leads to systematic over-performance of portfolio allocations among 2 assets. However, even thick modelling does not guarantee a constant over-performance with respect to a typical market benchmark for our asset allocation problem. To this end we have observed that combining thick modelling with a model specification strategy that imposes balanced regressions and chooses optimally the estimation window reduces the volatility of the asset allocation performance and delivers a more consistent over performance with respect to the simple buy-and-hold strategy.

References

[1] Aiolfi, M., C. A. Favero, and G. Primiceri (2001) 'Recursive 'Thick' Modelling of Excess Return and Dynamic Portfolio Allocation', IGIER Working Paper N. 197.

[2] Avramov, D. (2002) 'Stock return predictability and model uncertainty', forthcoming in *Journal of Financial Economics*.

[3] Barberis N. (2000) 'Investing for the Long run when Returns are Predictable', *Journal of Finance*, 55, 1, 225-264

[4] Campbell J.Y. and R.Shiller (1987) 'Cointegration and tests of present value models" *Journal of Political Economy*,95, 1062-1088.

[5] Campbell J.Y. and R.Shiller (1988a) 'The dividend-price ratio and expectations of future dividends and discount factors', *Review of Financial Studies*, 1, 195-227.

[6] Campbell J.Y. and R.Shiller (1988b) 'Stock prices, earnings, and expected dividends', *Journal of Finance*, 43, 661-676.

[7] Campbell, J. Lo, A. and McKinlay (1997) *The Econometrics of Financial Markets* Princeton University Press

[8] Campbell J.Y. and L.Viceira (1999) 'Consumption and portfolio decisions when expected returns are time-varying', *Quarterly Journal of Economics*, 114, 433-495.

[9] Clements M.P., and D.F. Hendry (2001) 'Forecast Economic timee-Series', Cambridge University Press, Cambridge

[10] Cochrane J. (1999) 'Portfolio advice for a multifactor world', NBER wp 7170

[11] Doppelhofer G., Miller R.I., Sala-i-Martin (2000) 'Determinants of long-term growth:a Bayesian averaging of classical estimates(BACE) approach', NBER wp 7750

[12] Granger C.W.J. (2000) 'Thick modelling', UCSD, mimeo

[13] Granger C.W.J. and A.Timmermann (1999) 'Data mining with local model specification uncertainty: a discussion of Hoover and Perez', *The Econometrics Journal*, 2, 220-226

[14] Hoeting J., D.Madigan, A.Raftery and C.Volinsky (1999) 'Bayesian Model Averaging: a Tutorial' Technical Report 9814, Department of Statistics, Colorado State University

[15] Kandel S., and R.S. Stambaugh (1996) 'On the Predictability of Stock Returns: An Asset-Allocation Perspective', *Journal of Finance*, 51, 2, 385-424

[16] Keim D., and R.Stambaugh (1986) 'Predicting returns in the stock and bond markets' *Journal of Financial Economics*, 17, 357-390

[17] Lander J., Orphanides A. and M. Douvogiannis (1997) 'Earning forecasts and the predictability of stock returns: evidence from trading the S&P' Board of Governors of the Federal Reserve System, http:\\www.bog.frb.fed.org

[18] Lamont O.(1998) 'Earnings and expected returns' *Journal of Finance*, 53, 5, 1563-1587

[19] Pesaran M.H. and A.Timmermann (2001) 'A Recursive Modelling Approach to Predicting UK Stock Returns', *The Economic Journal*,

[20] Pesaran M.H. and A.Timmermann (1995) 'Predictability of Stock Returns: Robustness and Economic Significance', *Journal of Finance*, 50, 4, 1201-1228

[21] Pesaran M.H. and A. Timmermann (1992) 'A simple non-parametric test of predictive performance', *Journal of Business and Economics Statistics*, 10, 461-465

[22] Raftery A., Madigan D. and J.Hoeting(1997) 'Bayesian model averaging for linear regression models' *Journal of the American Statistical Association*, 92(437), 179-191

[23] Samuelson P.A.(1969) 'Lifetime portfolio selection by dynamic stochastic programming', *Review of Economics and Statistics, 51, 239-246.*

[24] Siegel, J. (1994) *Stocks for the Long Run*, Richard D.Irwin, Burr Ridge, III.

[25] Sullivan R., A.Timmermann and H.White (1999) 'Data-snooping, technical trading rules performance and the bootstrap', *Journal of Finance*, 54, 1647-1692

[26] Sims C., Stock J. and M.Watson(1990) 'Inference in linear time-series models with some unit roots', *Econometrica*, 58, 113-144

[27] Zhou, C. (1996) 'Stock market fluctuations and the term structure', Board of Governors of the Federal Reserve System, http:\\www.bog.frb.fed.org.

A Data Appendix

The extended dataset has been obtained merging PT95 original dataset (1954.1-1992.12) with new series retrived from DATASTREAM and FRED for the sample 1993.1-2001.9.

	Code	Description
$P_t^{stock,US}$	TOTMKUS(RI)	US-DS MARKET - TOT RETURN IND
dy_t^{US}	TOTMKUS(DY)	US -DS market- Dividend yield
pe_t^{US}	TOTMKUS(PE)	US-DS MARKET - PER
$r1_t^{US}$	ECUSD1M	US EURO-$ 1 MONTH (LDN:FT) - MIDDLE RATE
ppi_t^{US}	USOCPRODF	US PPI - MANUFACTURED GOODS NADJ
$r12_t^{US}$	ECUSD1Y	US EURO-$ 1 YEAR (LDN:FT) - MIDDLE RATE
ip_t^{US}	USINPRODG	US INDUSTRIAL PRODUCTION
$M0_t^{US}$	USM0....B	US MONETARY BASE CURA
$R10Y_t^{US}$	BMUS10Y(RY)	US YIELD-TO-MATURITY ON 10_YEAR GOV.BONDS

Table 1: Data description:excess returns

	1954:1-1992:12	1993:1:1-2001:9
Mean	0.005906	0.006766
Median	0.006802	0.010988
Max	0.162767	0.087631
Min	-0.220550	-0.152226
Std. Dev.	0.042428	0.043985
Skewness	-0.289853	-0.731198
Kurtosis	4.968427	3.806151
Jarque-Bera	82.10989	12.19959
Probability	0.000000	0.002243
Observations	468	105

Percentage of periods of inclusion of each regressor in the best model (selection criterion $R2$)

	1954:1-1992:12	1954:1-2001:8
ysp(-1)	69.7	72.2
ep(-1)	20.5	30.1
i1(-1)	99.2	99.4
i1(-2)	23.0	18.1
i12(-1)	44.7	56.2
i12(-2)	42.9	54.8
infl(-2)	55.8	65.0
d12ip(-2)	87.9	90.4
d12m(-2)	89.6	80.4

Table 2: Forecasting performance

% of correct signs of predicted excess returns based on thin and thick modelling and associated PT statistics						
Sample 1954.1-1992.12						
Best	Best 1%	Best 5%	Best 10%	Best 20%	Best 30%	Best 40%
0.596 3.3939	0.591 3.1693	0.591 3.2016	0.596 3.4097	0.5934 3.3470	0.6061 3.9218	0.6136 4.2904
Best 50%	Best 60%	Best 70%	Best 80%	Best 90%	Best 100%	
0.6162 4.4198	0.6136 4.3271	0.6162 4.4198	0.6111 4.2158	0.6010 3.8271	0.5960 3.6037	
Sample 1993.1-2001.10						
Best	Best 1%	Best 5%	Best 10%	Best 20%	Best 30%	Best 40%
0.457 −0.2201	0.447 −0.486	0.438 −0.5298	0.457 −0.2201	0.438 −0.7497	0.419 −1.2792	0.419 −1.2722
Best 50%	Best 60%	Best 70%	Best 80%	Best 90%	Best 100%	
0.419 −1.3787	0.409 −1.6375	0.428 −1.3361	0.438 −1.1868	0.476 −0.9425	0.476 −1.1871	

Each cell reports the percentage of correctly signed predictions and the associated PT- statistic. The PT-statistic is the Pesaran-Timmerman non-parametric test of predictive performance. Let $x_t = E(y_t, \Omega_{t-1})$ be the predictor of y_t found with respect to the information set, Ω_{t-1} , with n observations $(y_1, x_1), (y_2, x_2), \ldots, (y_n, x_n)$ available. The test proposed by Pesaran and Timmerman (1992) is based on the proportion of times that the direction of changes in y_t is correctly predicted by x_t. The test statistic is computed as

$$Sn = \frac{P - P^*}{\{V(P) - V(P^*)\}^{1/2}} \sim N(0,1) \qquad (2)$$

where:

$$\begin{aligned} P &= \bar{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i \\ P^* &= P_y P_x + (1 - P_y)(1 - P_x) \\ V(P^*) &= \frac{1}{n} P^* (1 - P^*) \\ V(P) &= n \begin{pmatrix} (2P_y - 1)^2 P_x (1 - P_x) + (2P_x - 1)^2 P_y (1 - P_y) + \\ + \frac{4}{n} P_y P_x (1 - P_y)(1 - P_x) \end{pmatrix} \end{aligned}$$

Z_i is an indicator variable which takes value of one when the sign of y_t is correctly predicted by x_t, and zero otherwise, P_y is the proportion of times y_t takes a positive value, P_x is the proportion of times x_t takes a positive value.

Table 3: Sample 1954.1-1992.12

Trimmed means	% of correct signs (PT-statistic)				
	Rec	*Roll*	*Bal*	*Flex*	*Bal-flex*
Best	0.596 3.3939	0.543 1.494	0.571 2.444	0.553 2.151	0.540 1.609
Best 1%	0.591 3.1693	0.545 1.607	0.573 2.615	0.558 2.330	0.530 1.257
Best 5%	0.591 3.2016	0.553 1.859	0.563 2.165	0.568 2.621	0.566 2.694
Best 10%	0.596 3.4097	0.558 2.084	0.568 2.351	0.563 2.331	0.571 2.825
Best 20%	0.5934 3.3470	0.550 1.767	0.563 2.205	0.586 3.257	0.551 2.015
Best 30%	0.6061 3.9218	0.556 2.014	0.573 2.780	0.576 2.850	0.556 2.217
Best 40%	0.6136 4.2904	0.568 2.555	0.571 2.690	0.591 3.460	0.551 2.015
Best 50%	0.6162 4.4198	0.575 2.899	0.568 2.599	0.583 3.187	0.548 1.902
Best 60%	0.6136 4.3271	0.573 2.716	0.563 2.418	0.576 2.871	0.540 1.609
Best 70%	0.6162 4.4198	0.601 3.751	0.568 2.599	0.586 3.257	0.540 1.633
Best 80%	0.6111 4.2158	0.588 3.193	0.551 1.968	0.591 3.481	0.535 1.360
Best 90%	0.6010 3.8271	0.596 3.493	0.548 1.949	0.596 3.623	0.551 1.900
Best 100%	0.5960 3.6037	0.588 2.915	0.553 2.199	0.601 3.770	0.556 2.103

Each cell reports the percentage of correctly signed predictions and the associated PT- statistic.

Rec: recursive estimation (expanding window of observations), no focal variables.

Rol: rolling estimation(with fixed window of 60 observations), no focal variables.

Bal: recursive estimation(expanding window of observations), focal variables: constant, log of the price-earning ratio, yield-to maturity on long term bonds, yield on 12-month Treasury Bills.

Flex: rolling estimation(with optimally chosen window), no focal variables.

Bal-flex: rolling estimation(with optimally chosen window), focal variables: constant, log of the price-earning ratio, yield-to maturity on long term bonds, yield on 12-month Treasury Bills.

Table 4: Sample 1993.1-2001.10

Trimmed means	% of correct signs (PT-statistic)				
	Rec	Roll	Bal	Flex	Bal-flex
Best	0.457 −0.2201	0.514 0.334	0.485 0.470	0.467 0.545	0.476 1.139
Best 1%	0.447 −0.4860	0.552 1.264	0.485 0.470	0.429 −0.107	0.505 1.440
Best 5%	0.438 −0.5298	0.533 0.742	0.495 0.872	0.438 0.062	0.467 0.545
Best 10%	0.457 −0.2201	0.552 1.038	0.504 1.021	0.419 −0.516	0.467 0.414
Best 20%	0.438 −0.7497	0.485 −0.334	0.457 0.132	0.457 −0.220	0.457 −0.105
Best 30%	0.419 −1.2722	0.504 −0.038	0.485 0.470	0.467 −0.182	0.457 −0.557
Best 40%	0.419 −1.2722	0.504 −0.038	0.485 0.350	0.448 −0.927	0.448 −1.151
Best 50%	0.419 −1.3787	0.457 −0.890	0.495 0.499	0.467 −0.856	0.505 −0.500
Best 60%	0.409 −1.6375	0.485 −0.447	0.476 −0.031	0.476 −0.942	0.524 −0.442
Best 70%	0.428 −1.3361	0.466 −0.971	0.447 −1.038	0.476 −1.315	0.524 −0.569
Best 80%	0.438 −1.1868	0.485 −1.040	0.447 −1.038	0.448 −2.162	0.552 −0.098
Best 90%	0.476 −0.9425	0.485 −1.587	0.466 −0.742	0.495 −1.595	0.562 0.065
Best 100%	0.476 −1.1871	0.495 −1.760	0.476 −0.709	0.495 −1.716	0.590 0.687

Each cell reports the percentage of correctly signed predictions and the associated PT- statistic.

Rec: recursive estimation (expanding window of observations), no focal variables.

Rol: rolling estimation(with fixed window of 60 observations), no focal variables.

Bal: recursive estimation(expanding window of observations), focal variables: constant, log of the price-earning ratio, yield-to maturity on long term bonds, yield on 12-month Treasury Bills.

Flex: rolling estimation(with optimally chosen window), no focal variables.

Bal-flex: rolling estimation(with optimally chosen window), focal variables: constant, log of the price-earning ratio, yield-to maturity on long term bonds, yield on 12-month Treasury Bills.

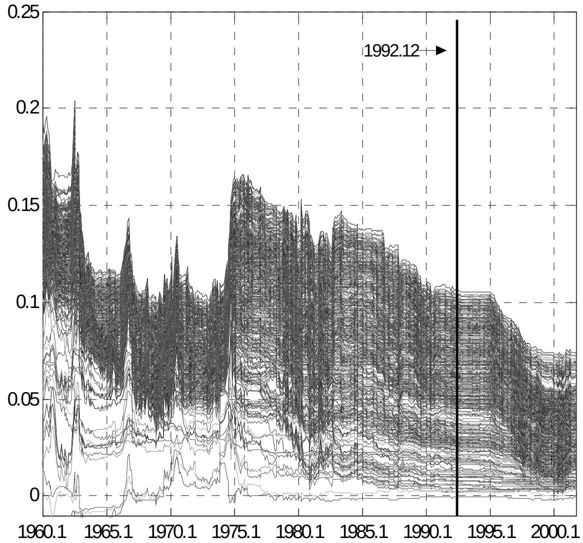

Figure 1: Recursive adjusted R2: first sample 1954.1-1959.12, last sample 1954.1-2001.8.

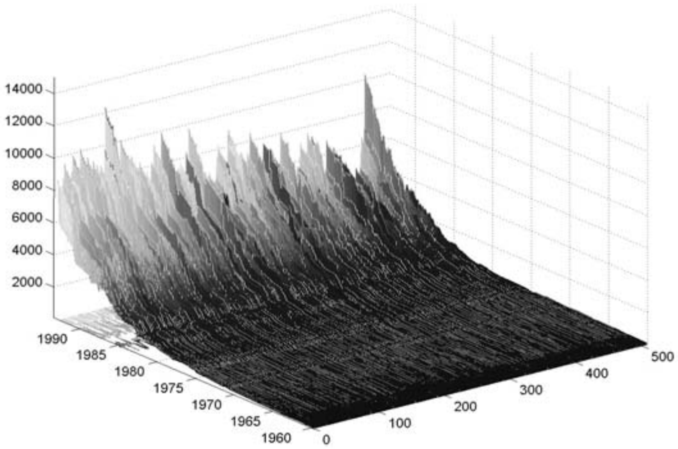

Figure 2: Cumulative wealth obtained from 511 different portfolios.

Figure 3: The performance of different portfolio allocations over the sample 1960-1992

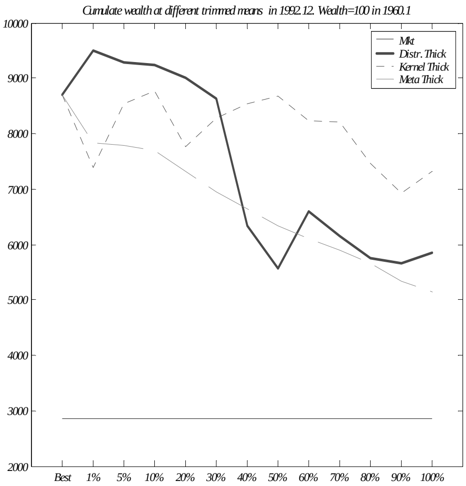

Figure 3a: Recursive

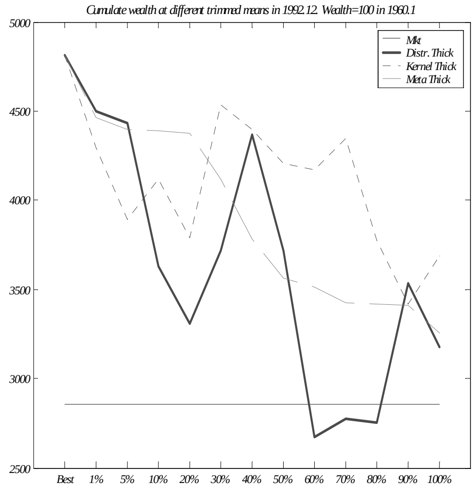

Figure 3b: Recursive Balanced

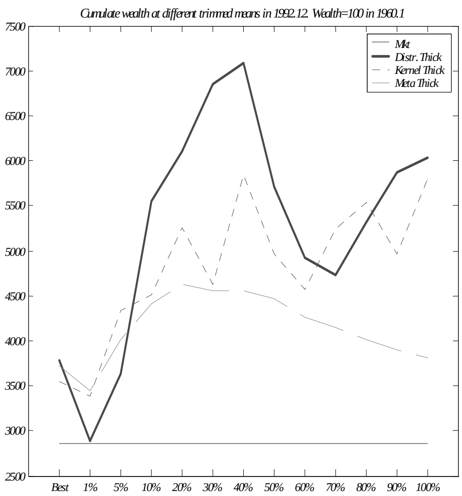

Figure 3c: Flexible Rolling

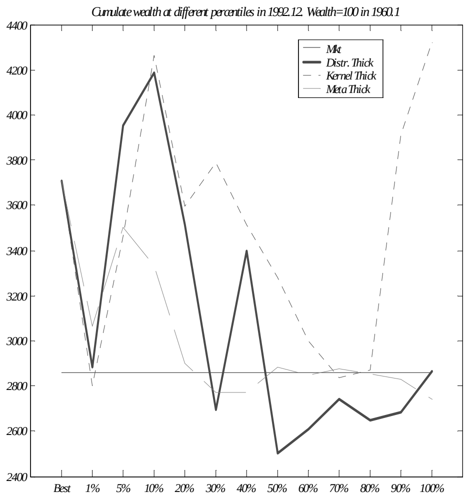

Figure 3d: Flexible-Balanced Rolling

Figure 4: The performance of different portfolio allocations over the sample 1993-2001

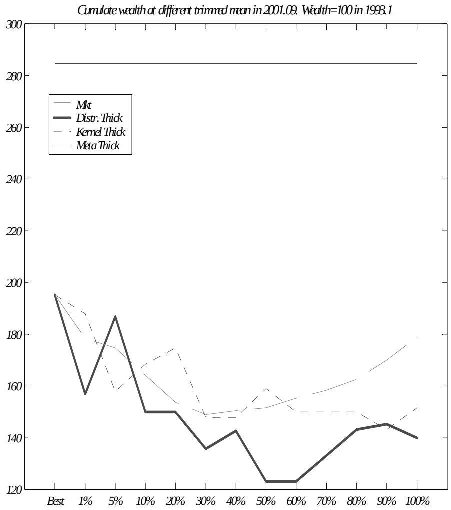

Figure 4a: Recursive

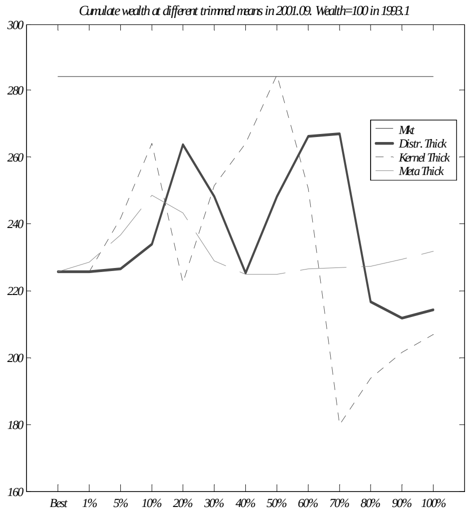

Figure 4b: Recursive Balanced

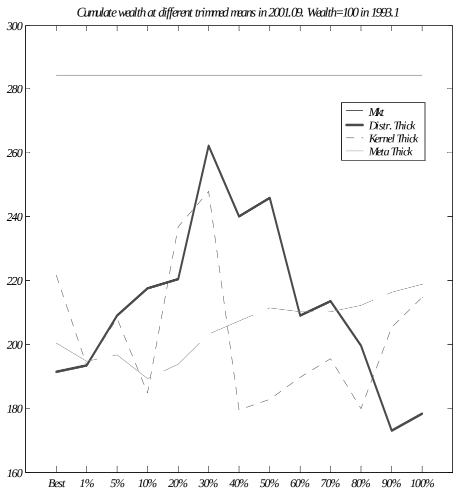

Figure 4c: Flexible rolling

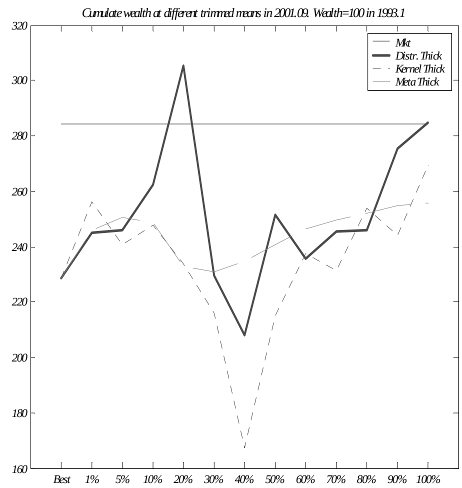

Figure 4d: Flexible-Balanced Rolling