# Efficient Market Hypothesis and Forecasting*

Allan Timmermann and Clive W.J. Granger

**Abstract**

The efficient market hypothesis gives rise to forecasting tests that mirror those adopted when testing the optimality of a forecast in the context of a given information set. However, there are also important differences arising from the fact that market efficiency tests rely on establishing profitable trading opportunities in 'real time'. Forecasters constantly search for predictable patterns and affect prices when they attempt to exploit trading opportunities. Stable forecasting patterns are therefore unlikely to persist for long periods of time and will self-destruct when discovered by a large number of investors. This gives rise to nonstationarities in the time series of financial returns and complicates both formal tests of market efficiency and the search for successful forecasting approaches.

Key words: efficient market hypothesis, forecast evaluation, model specification, learning.

*We thank two anonymous referees for constructive criticisms that were very helpful in writing this paper.

## I. Introduction

The efficient market hypothesis (EMH) is a backbreaker for forecasters. In its crudest form it effectively says that series we would very much like to forecast, the returns from speculative assets, are unforecastable. This is a venerable thesis, its earliest form appearing a century ago as the random walk theory (Bachelier (1900)). This theory was confirmed empirically in the 1960's (see Cootner (1964)) and many times since. Soon after the empirical evidence appeared, the EMH was proposed based on the overpowering logic that if returns were forecastable, many investors would use them to generate unlimited profits. The behavior of market participants induce returns that obey the EMH, otherwise there would exist a "money-machine" producing unlimited wealth, which cannot occur in a stable economy.

Intellectually, that might appear to be the end of the story. However, despite the force of the argument, it seems not to be completely convincing for many forecasters. Everyone with a new prediction method wants to try it out on returns from a speculative asset, such as stock market prices, rather than series that are known to be forecastable. Papers continue to appear attempting to forecast stock returns, usually with very little success.[1]

In this paper we discuss the efficient market hypothesis from the perspective of a modern forecasting approach. Despite its simplicity, the EMH is surprisingly difficult to test and considerable care has to be exercised in empirical tests. Forecasting experiments have to specify at least five factors, namely

(i) the set of forecasting models available at any given point in time, including estimation methods;

[1]It is often argued that there is a 'file drawer' bias in published studies due to the difficulty associated with publishing empirical studies that find insignificant effects. In studies of market efficiency, a reverse file drawer bias may be present. A researcher who genuinely believes he or she has identified a method for predicting the market has little incentive to publish the method in an academic journal and would presumably be tempted to sell it to an investment bank.

(ii) the search technology used to select the best (or a combination of best) forecasting model(s);

(iii) the available 'real time' information set, including public versus private information and ideally the cost of acquiring such information;

(iv) an economic model for the risk premium reflecting economic agents' trade-off between current and future payoffs;

(v) the size of transaction costs and the available trading technologies and any restrictions on holdings of the asset in question.

The EMH is special in that investors' current and future forecasts of payoffs affect their current and future trades which in turn affect returns. Investors' learning gives rise to the likely demise of stable forecasting models and this poses a unique challenge both to establishing successful forecasting procedures and to forecast evaluation.

Ignoring uncertainty about the best forecasting model (or set of forecasting models) the existence of a single successful prediction model is sufficient to demonstrate violation of the EMH. However, once model uncertainty is accounted for, this is no longer the case unless there is evidence of a search technology that would allow investors to identify successful models *ex ante*.

Acknowledging this point, we provide suggestions to the sort of forecasting procedure that could work even if the EMH is correct. It seems obvious that the old stand-by in the forecasting world, the standard constant-parameter model with a simple specification, such as the ARMA models discussed by Box and Jenkins, are not up to the task since it assumes stationarity. Consideration needs to be turned to quickly changing models that can detect and utilize any instances of temporary forecastability that might arise and quickly disappear as learning opportunities arise and close down.

The outline of the paper is as follows. Section 2 discusses the basic efficient market hypothesis in the context of classical definitions proposed in the literature.

Section 3 discusses the role of model specification uncertainty, while section 4 covers the effect of dynamic learning and 'feedback' effects on return predictability. Section 5 discusses the sort of forecasting approaches that may work even in an efficient market. Section 6 concludes.

## II. Definitions of Market Efficiency

Jensen (1978) defines market efficiency as follows[2]

> "*A market is efficient with respect to information set $\Omega_t$ if it is impossible to make economic profits by trading on the basis of information set $\Omega_t$.*"

A closely related definition of market efficiency is provided by Malkiel (1992):[3]

> "*A capital market is said to be efficient if it fully and correctly reflects all relevant information in determining security prices. Formally, the market is said to be efficient with respect to some information set, $\Omega_t$, if security prices would be unaffected by revealing that information to all participants. Moreover, efficiency with respect to an information set, $\Omega_t$, implies that it is impossible to make economic profits by trading on the basis of $\Omega_t$.*"

Three points are emphasized in these definitions, namely (i) the importance of the information set adopted in the test, $\Omega_t$; (ii) the ability to exploit this information in a trading strategy; and finally (iii) that the yardstick for testing if the EMH holds is measured in economic (i.e. risk-adjusted and net of transaction costs) profits. We discuss each of these elements in the following.

[2]Jensen uses $\theta_t$ as symbol for the information set. We have changed this to the symbol $\Omega_t$ which will be used throughout the paper.

[3]Malkiel uses $\phi$ as symbol for the information set and we have changed this to $\Omega_t$.

### *A. The Information Set*

Three forms of market efficiency are commonly entertained in the EMH literature based on the set of variables contained in the information set, $\Omega_t$, c.f. Roberts (1967) and Fama (1970). If $\Omega_t$ only comprises past and current asset prices (as well as possibly dividends and variables such as trading volume), the EMH in its weak form is being tested. Expanding $\Omega_t$ to include all publicly available information gives rise to the EMH in its semi-strong form. Finally, if all public and private information is included in $\Omega_t$, market efficiency in the strong form is being tested.

Most studies in the literature on predictability of stock market returns test the EMH in its weak or semi-strong form. For example, papers on the predictive performance of technical trading rules test weak form market efficiency since only past prices and maybe volume information are used as predictor variables. Studies that include an extended set of predictor variables such as default premia, term spreads and other business cycle indicators test semi-strong efficiency.

Restricting the information set in this way is designed to rule out private information that is harder to measure and perhaps also more expensive to acquire. For example, it is not usually asserted that a market is efficient with respect to inside information since this information is not widely accessible and hence cannot be expected to be fully incorporated in the current price. Strong form efficiency can be tested indirectly, e.g., by considering the performance of fund managers and testing if they manage to earn profits net of risk premia after accounting for the cost of acquiring private information.

Surveys of market efficiency such as Fama (1970, 1991) have focused on testing informational efficiency. Fama (1970) concludes that the empirical evidence is largely supportive of weak form and semi-strong form efficiency, while Fama (1992) reports stronger evidence of predictability in returns based both on lagged values of returns and publicly available information.

*B. Time-varying Risk Premia*

The EMH implies the absence of arbitrage opportunities. It does not rule out all forms of predictability in returns. Suppose for simplicity that transaction costs are either zero or small enough so that they can effectively be ignored. Also suppose that there is no uncertainty about the functional form and the parameter values of the best prediction model. Under no arbitrage, the current price of some financial asset, $P_t$, is then given as the conditional expectation of the asset's payoffs − comprising its future price, $P_{t+1}$ as well as any coupons or dividends, $D_{t+1}$− multiplied by a variable known as the 'stochastic discount factor' or 'pricing kernel', $Q_{t+1}$, that accounts for variations in economic risk premia:

$$P_t = E[Q_{t+1}(P_{t+1} + D_{t+1})|\Omega_t]. \tag{1}$$

Here $E[.|\Omega_t]$ is the mathematical expectation operator, or population expectation, conditional on the information set $\Omega_t$. Under a set of rather restrictive assumptions, the EMH therefore translates into a simple moment condition. This insight goes back to at least Harrison and Krebs (1979).

Most asset prices are trended. Tests for predictability typically eliminate such (global) trends by considering the excess rate of return, $R_{t+1}$, defined as the return, $(P_{t+1}+D_{t+1}-P_t)/P_t$, over and above the risk-free rate (e.g., the return on T-bills), $r_{ft}$. Dividing equation (1) through by $P_t$ and subtracting $r_{ft}$, we get

$$E[Q_{t+1}R_{t+1}|\Omega_t] = 0. \tag{2}$$

Since the process generating the risk-premium is model-dependent and is not observable, tests of the EMH can only be conducted jointly with auxiliary hypotheses about $Q_{t+1}$. This can most easily be seen by rearranging equation (2) to get

$$E[R_{t+1}|\Omega_t] = \frac{-Cov(R_{t+1}, Q_{t+1}|\Omega_t)}{E[Q_{t+1}|\Omega_t]}. \tag{3}$$

Predictability of returns thus need not violate the EMH. Forecasting models that work because they predict the conditional covariance of returns with the pricing kernel, $Q_{t+1}$, scaled by its conditional mean, are not ruled out.

It is worth pointing out, however, that although a variety of economic models, such as the Consumption Capital Asset Pricing Model, have been entertained most models of the risk premium generate insufficient variation in economic risk-premia to explain existing asset pricing puzzles.[4] This does not, of course, rule out that an economic model exists that could justify many patterns of predictability.

Even though many studies equate market efficiency with the random walk model for stock prices, this is clearly not generally true. More precisely, it can be shown that stock prices plus cumulated dividends discounted at the risk-free rate should follow a martingale process under the so-called 'risk-neutral' or equivalent martingale probability measure.[5] This measure incorporates information on risk-premia as well as the objective outcome probabilities. Let $B_{t+1}$ be the price at time $t+1$ of a risk-free zero coupon bond and define $P^*_{t+1} = P_{t+1}/B_{t+1}$ as the discounted value of the price at time $t+1$, while $\Sigma^*_{t+1} = \sum_{\tau=1}^{t+1}(D_\tau/B_\tau)$ is the cumulated sum of discounted dividends. It is easy to demonstrate from equation (1) that

$$E_{\pi^*}[P^*_{t+1} + \Sigma^*_{t+1}|\Omega_t] = P^*_t + \Sigma^*_t, \tag{4}$$

where $E_{\pi^*}$ indicates that the expectation is computed under the 'risk-neutral' probability measure formed as the product of the risk premium $Q_{t+1}$ and the 'objective' outcome probabilities. This will typically differ from the objective probability or empirical frequency.

Forecasting tests of the validity of the random walk model conducted on security prices are thus insufficient to demonstrate market inefficiency since they are only implied by the EMH under a set of special circumstances, i.e. when risk-premia do

[4]See, e.g. Hansen and Singleton (1983), Breeden, Gibbons and Litzenberger (1989).

[5]If the random walk hypothesis is interpreted in its strictest form to imply that price increments are identically and independently distributed, then the random walk model is rejected by the presence of conditional heteroskedasticity in returns irrespective of any risk premium effects. A weaker version of the random walk model adopted in some empirical studies only assumes that the price increments are uncorrelated. This need not be inconsistent with conditional heteroskedasticity.

not play an important role, in the absence of dividends and ignoring interest rate effects and transaction costs.

To simplify matters, define the risk-adjusted return as $R^*_{t+1} = R_{t+1}Q_{t+1}$. Informational efficiency then implies that for all vectors $\mathbf{z}_t \in \Omega_t$, the following equation must hold:

$$E[R^*_{t+1}|\mathbf{z}_t] = \mathbf{0}. \tag{5}$$

This means that all transformations $\psi(\mathbf{z}_t)$ of variables in the information set over which market efficiency is being tested should be orthogonal to $R^*_{t+1}$.

This test of the EMH is very similar to tests of optimality proposed in the forecasting literature. Suppose a forecaster is endowed with a loss function, $L(.)$, defined over the forecast error, $e_{t+1}$, given as the difference between the realization and prediction of some variable. The first order condition for forecast optimality is, c.f. Granger (1999),

$$E[L'(e_{t+1})|\Omega_t] = 0, \tag{6}$$

where the derivative is computed with respect to the forecast. Under quadratic loss the condition simplifies to

$$E[e_{t+1}\mathbf{z}_t] = \mathbf{0}, \tag{7}$$

for all $\mathbf{z}_t \in \Omega_t$. This implies unbiasedness and absence of serial correlation in the forecast error. If the shape of the loss function is unknown, of course a joint hypothesis testing problem arises when testing the informational efficiency of a forecast. This is similar to the problem associated with not observing the risk premium.

There is an important distinction, nevertheless. In tests of the EMH, the loss function is usually easily identified - namely economic profits net of transaction costs - but measurement of economic profits is made difficult by their dependence on the risk premium. One possibility is to carry out a sensitivity analysis with respect to different specifications of the pricing kernel, $Q_{t+1}$, and then adopt an approach such as that proposed by Diebold and Mariano (1995) to compare the economic profits generated by the forecasting model to those from, say, holding the market index.

Interestingly, the joint hypothesis problem in testing market efficiency in conjunction with a maintained specification for the economic risk premium was not considered particularly important in Fama (1970) since evidence of predictability was only found to be weak. Fama's early survey therefore assumed that the EMH implied that expected asset returns should be a 'fair game', suggesting that time-variations in the risk premium were not very important. This point of view very much changed in Fama (1991) which reports more extensive evidence of predictability and argues that an economic model for expected returns is required.[6]

### *C. Transaction costs and trading restrictions*

Transaction costs and trading restrictions change tests of market efficiency in some important ways. Most obviously, if transaction costs are very high, predictability is no longer ruled out by arbitrage, since it would be too expensive to take advantage of even a large, predictable component in returns.

An investor may predict that a particular stock is going to outperform the market by 2% next year, but if the transaction cost from buying the asset is 3%, then it may not be profitable to exploit this prediction. Predictability therefore has to be seen in relation to the transaction costs of the asset. Predictable patterns only invalidate the EMH once they are large enough to cover the size of transaction costs. Likewise, if short-selling is not possible, certain types of asymmetric predictability need not be inconsistent with the EMH simply because they cannot be exploited in a profitable trading strategy.

To formalize this idea, suppose that $\mathbf{c}_t$ is a vector of transaction cost parameters

[6] "In brief, the new work says that returns are predictable from past returns, dividend yields, and various term-structure variables. The new tests thus reject the old market efficiency-constant expected returns model that seemed to do well in the early work. This means, however, that the new results run head-on into the joint-hypothesis problem." (Fama (1991) page 1577).

incorporating factors such as bid-ask spreads and brokers' fees at time $t$. Also let $f_t$ embody the set of possible transactions at time $t$, which could include short sales constraints, limits on maximum holdings in individual stocks and so on. The EMH thus needs to be modified to a condition that, loosely speaking, takes the form

$$E[f_t(R^*_{t+1}, \mathbf{c}_t)|\Omega_t] = 0. \tag{8}$$

$f_t$ maps the position size into profits as a function of the future return and transaction cost parameters.

Transaction costs vary over time and have come down considerably for many assets. Tests of the EMH therefore require using real-time data on transaction costs and a condition such as (8) is difficult to test over long periods of time.

In practice transaction costs are also likely to rule out money machines since the market impact of increasingly large positions means that profit opportunities cannot simply be 'scaled up'. In this case one could also specify a limit on the size of the expected profits.

### *D. Market Efficiency and Intrinsic Asset Values*

The earlier definition of market efficiency is purely based on asset returns independently of how these are related to the underlying 'intrinsic' asset value. In fact, prices and values need not be closely related. For instance, the earlier definition of market efficiency does not rule out the presence of speculative bubbles unless a transversality condition is imposed on the asset price that solves equation (1).

An alternative definition of market efficiency built on the notion of value as distinct from price is due to Fisher Black (1986). In his presidential address to the American Finance Association, Black wrote:

> "*However, we might define an efficient market as one in which price is within a factor of 2 of value, i.e., the price is more than half of value and less than twice value. The factor of 2 is arbitrary, of course. Intuitively, though, it seems reasonable to me, in the light of sources of*

*uncertainty about value and the strength of the forces tending to cause price to return to value. By this definition, I think almost all markets are efficient almost all of the time. "Almost all" means at least 90%."* (page 533)

This definition of market efficiency focuses on the size of deviations of asset prices from true 'value'. Investors' information can be so 'noisy' at times that prices are far removed from fundamentals. However, the market is still efficient in this definition provided that such deviations do not last 'too long' or become 'too big' before they are corrected. A problem with this definition is, of course, that fundamentals are not observable. Testing the EMH according to this definition would therefore require measuring whether the difference between prices and values is growing over time. Ultimately this is unlikely to be a promising research strategy, however, since bubbles are not ruled out by the EMH on pure no-arbitrage grounds provided that they earn a risk premium that grows with their size and properly reflects the probability of their collapse. Furthermore, Evans (1991) finds that periodically collapsing bubbles are very difficult to detect by means of standard tests based on whether the growth in stock prices is more explosive than the growth in dividends. He therefore proposes parametrically modeling either the (unobserved) fundamentals or the speculative bubble and basing tests on the difference between observed prices and these components.

## III. Uncertainty About Model Specification

An important weakness of the earlier definitions of market efficiency is that they do not account for investors' uncertainty about the 'best' model to use when forecasting future returns. In reality, investors face the difficult task of choosing a specific forecasting model or combining a subset of forecasting models from a huge, possibly infinite-dimensional, space of potential forecasting models.

This fundamental uncertainty about the best or even moderately successful prediction model is also the reason why predictability can exist in local 'pockets in time.' Investors with heterogenous beliefs and information simultaneously search for forecasting models that might work for some time. For instance, information criteria such as the $AIC$ or $BIC$ are commonly used to select among models that do not have to be nested. Alternatively, investors could have selected the forecasting model that historically would have generated the highest value of a particular financial performance measure.[7]

If agents do not know the true forecasting model, then the practice of using the mathematical expectations operator in the definition of market efficiency becomes rather less attractive. Instead it is more meaningful to define a market as being *efficient locally in time* with respect to information set $\Omega_t$ and the forecasting model $m_{it}(\mathbf{z}_t, \widehat{\boldsymbol{\theta}}_t)$ drawn from a set of available models, $M_t$ if

$$E[f_t(R^*_{t+1}, m_{it}(\mathbf{z}_t; \widehat{\boldsymbol{\theta}}_t), \mathbf{c}_t)] = 0. \tag{9}$$

Here $\widehat{\boldsymbol{\theta}}_t$ is a vector of parameters estimated using data up to time $t$ and $\mathbf{z}_t \in \Omega_t$. The $t$-subscript on $M_t$ indicates that only forecasting techniques that were available at time $t$ can be used in the modeling. We intend 'model' to be interpreted in the broadest sense to incorporate both the functional form, prediction variables, estimation method and choice of sample period (expanding window, rolling window, exponential discounting etc.).

In the absence of transaction costs, letting $\widehat{R}^*_{it+1}$ be the predicted value of $R^*_{t+1}$ generated by the $i$th forecasting model, we have

$$E[R^*_{t+1}\widehat{R}^*_{it+1}] = 0. \tag{10}$$

We can imagine that some models had predictive power before their discovery (e.g. neural networks may have worked well during, say, the 1960's). This would not constitute a violation of the EMH defined in equation (10) since such models

[7] If several search methods or model selection technologies are available, the question also arises which one to use. See Pesaran and Timmermann (1995) for a discussion of a possible 'hyper selection' method based on economic utility.

would not be elements in the relevant set, $M_t$. It would, however, violate equation (5) which is based on population expectations. Another way to state this is that the EMH in this definition does not rule out profits from new forecasting techniques. The latest techniques may have a 'honeymoon' period before their use becomes more widespread and they cease to generate profits. This would explain the economic incentive to develop these methods in the first instance.

If the market is not locally efficient at all points in time, it is possible that there exists a time interval, $\Upsilon = [t_{beg}, ..., t_{end}]$, $0 \leq t_{beg} \leq t_{end} \leq T$, formed as an interval on the collection of points, $t \in [0, 1, ..., T]$, and a forecasting model, $m_{it} \in M_t$, that could have been identified ex ante (i.e. at time $t_{beg} - 1$) using available model selection techniques such that

$$E[f_t(R^*_{t+1}, m_{it}(\mathbf{z}_t; \widehat{\boldsymbol{\theta}}_t), \mathbf{c}_t)] > 0, \quad \text{for all} \;\; t \in \Upsilon \tag{11}$$

or, under zero transaction costs

$$E[R^*_{t+1}\widehat{R}^*_{it+1}] > 0, \quad \text{for all} \;\; t \in \Upsilon \tag{12}$$

Conversely, a market is efficient over the period (horizon) $\tau$ if the length of these intervals, $t_{end} - t_{beg} + 1 \leq \tau$ for all forecasting models $m_{it} \in M_t$, i.e. (10) holds for less than $\tau$ periods. If transaction costs and market impact is taken into account, one can also specify quantitative limits on the expected profits.

An efficient market is thus a market in which predictability of asset returns, after adjusting for time-varying risk-premia and transaction costs, can still exist but only 'locally in time' in the sense that once predictable patterns are discovered by a wide group of investors, they will rapidly disappear through these investors' transactions.

At a more conceptual level one can even question whether it is reasonable to condition on the set of forecasting ($M_t$) and trading ($f_t, \mathbf{c}_t$) technologies that exist at a given point in time. For example, suppose that the application of ARCH-in-mean models to forecasting hourly returns would have been profitable if these techniques had been developed a few years prior to when they actually emerged in the study by Engle, Lilien and Robins (1987). If it would have been possible and inexpensive to develop such techniques early, the financial market was efficient with respect to its use of existing forecasting techniques but perhaps inefficient at developing new techniques. While such thoughts may seem rather speculative, they do show the issues arising once we acknowledge uncertainty about forecasting techniques and model specification.

### *A. Model specification search and forecast evaluation*

Most forecasters entertain several competing models before settling on their preferred model. When assessing a particular forecasting model's performance one must account for the effect of the specification search that preceded its discovery. It is necessary to ask whether the preferred forecasting model chosen from a larger set of models outperformed the benchmark model (which is taken as being efficient). Suppose there are $N_t$ models under consideration. Then we may be interested in testing whether, accounting for the effect of searching across $N_t$ models, the best model genuinely outperforms the benchmark:

$$H_0 : \max_{i=1,..,N_t} E[f_t(R^*_{t+1}, m_{it}(\mathbf{z}_t; \widehat{\boldsymbol{\theta}}_t), \mathbf{c}_t)] \leq 0.$$

See, e.g., Sullivan, Timmermann and White (2002) for a test that uses this setup. If the null is rejected, it suggests that the best forecasting model has genuine predictive power.

As data is represented in the form of a finite sample of asset returns, at any given point in time there will almost certainly be several possible forecasting models that appear to provide a good 'fit' to the data. Considering the full set of models used in the search and explicitly evaluating the performance of the best forecasting model as the outcome of the best draw from a larger universe handles this problem of data-snooping.

Many researchers have attempted to reject the EMH by presenting evidence that a particular model could forecast financial returns over a given sample period.

However, conditioning on a specific model, $m_{i_t^*} \in M_t$, and showing that it could have been used to forecast asset returns is insufficient to disprove the EMH.

Even if model $m_{i_t^*}$ was found *ex post* to have outperformed over a certain sample period, it is unclear that investors could possibly have selected this forecasting model *ex ante*. Both the set of forecasting models over which the search is conducted, $M_t$, and the set of search technologies $S_t$ thus have to be specified in a test of the EMH.

## IV. Learning and Non-stationarity of Returns

Because investors' beliefs affect the path of asset prices, even if the underlying payoffs such as dividends or coupons are stationary, the best forecasting model is unlikely to remain the same.[8] Individual forecasting models are likely to go through stages of success, declining value, and disappearance. The difficulty of selecting a successful forecasting model is of course compounded by the extremely noisy nature of most financial return series.

One place where such effects can be expected to show up is in the vast literature on financial market 'anomalies.' This literature has found evidence that the stocks of firms with high book to market values and low market capitalization pay higher returns than is compatible with standard models of risk premia. It has also found that returns at particular calendar frequencies tend to be abnormal (e.g., in January, or around week-ends, c.f. Lakonishok and Smidt (1988) and Thaler (1987)).

### *A. Self-Destruction of Predictability*

Once an anomaly has become 'public knowledge', we would expect it to disappear in future samples. This may happen simply because the anomaly was spurious in

[8]Conceptually it is difficult to distinguish between 'exogenous fundamentals' and investors' beliefs since the fundamental value that matters to current asset prices is the expected future payoff stream computed conditional on the current information that helps shape investors' beliefs.

the first instance, a result of excessive 'data mining'. In this case there is no reason why it should be repeated in future periods. However, even if the anomaly was originally genuine, we expect that its publication will attract sufficient new capital to exploit the predictable pattern and remove it in the process. This complicates any statistical tests of predictability.

To see why forecasting models 'self-destruct' in an efficient market, suppose that a particular forecasting model correctly predicts that small firms pay higher returns in recessions than is consistent with their risk premia. Investors who use forecasting models that identify this 'anomaly' should buy small firms' stocks when the economy is believed to be in a recession. This will bid up their price and lower the return on these stocks during recessions.

The market's learning may well take a long period until it is reasonably complete and hence patterns of (weak) predictability may exist for some time. Although an individual investor who has discovered a successful forecasting model is likely to act on the forecasting signals, bidding up the price of 'underpriced' assets and lowering the price of 'overpriced' assets, his individual actions are unlikely to lead to a full adjustment in the price. Over time, additional investors are likely to find similar models and allocate additional capital to exploit the patterns. Under these circumstances predictability is perhaps more likely to be present during volatile markets where it may be harder to detect and estimate a forecasting model that can benefit from predictability.

Out-of-sample forecastability is not necessarily indicative of whether predictability existed in-sample. In the absence of learning effects, returns in financial markets are often assumed to be stationary so that out-of-sample tests of predictability can be used to provide tests of genuine predictability in a way that controls for the possible contamination arising from data-mining. The outcome of such tests is difficult to interpret in financial markets where investors are constantly learning about the data generating process through new forecasting technologies. If it takes time for the markets to discover predictable patterns, such patterns may have been present up to a given point in time but disappear once they are made public. Their absence out-of-sample thus may not prove that predictability was not genuine in the first instance.

Recursive parameter estimation and model misspecification may lead to ex-post predictability in returns that nevertheless could not have been exploited in 'real time'. Suppose that stock returns display serial correlation in a particular sample. This would suggest that a simple ARMA model could have been used to generate profits. Such a conclusion is incorrect, however. Timmermann (1993) shows that investors' updating of their parameter estimates can lead to ex-post serial correlation in returns even if this did not exist in 'real time'. Suppose that the current asset price depends on the current model used by investors, $m_{i_t}$, as well as on the parameter estimates, $P_t = \eta(m_{i_t}(\widehat{\boldsymbol{\theta}}_t, \mathbf{z}_t))$. Similarly, next period's price will reflect the model chosen at this time as well as its parameter estimates: $P_{t+1} = \eta(m_{i_{t+1}}(\widehat{\boldsymbol{\theta}}_{t+1}, \mathbf{z}_{t+1}))$. The sample dependence in both $m_i$ and $\widehat{\boldsymbol{\theta}}$ can generate serial correlation in the time-series of excess returns, $R_1, R_2, ..., R_T$ since $R_{\tau+1} = (\eta(m_{i_{\tau+1}}(\widehat{\boldsymbol{\theta}}_{\tau+1}, \mathbf{z}_{\tau+1})) + D_{t+1} - \eta(m_{i_\tau}(\widehat{\boldsymbol{\theta}}_\tau, \mathbf{z}_\tau)))/\eta(m_{i_\tau}(\widehat{\boldsymbol{\theta}}_\tau, \mathbf{z}_\tau))$. Hence the expected rate of return, when computed under the model $m_{i_\tau}$, which we denote by $E_{m_{i_\tau}}[.]$, satisfies

$$E_{m_{i_\tau}}[R^*_{\tau+1}(m_{i_\tau}(\widehat{\boldsymbol{\theta}}_\tau, \mathbf{z}_\tau), m_{i_{\tau+1}}(\widehat{\boldsymbol{\theta}}_{\tau+1}, \mathbf{z}_{\tau+1}))] = 0.$$

However, this condition does not imply that

$$E[R^*_{\tau+1}(m_{i_\tau}(\widehat{\boldsymbol{\theta}}_\tau, \mathbf{z}_\tau), m_{i_{\tau+1}}(\widehat{\boldsymbol{\theta}}_{\tau+1}, \mathbf{z}_{\tau+1}))|\Omega_T] = 0, \quad \tau \leq T.$$

The difference between using full-sample ($\widehat{\boldsymbol{\theta}}_T$) or recursive (or real-time) parameter estimates ($\widehat{\boldsymbol{\theta}}_t$) or model specifications $m_{i_t}, m_{i_{t+1}}$ can clearly be very significant.

As returns depend both on the selected model as well as the recursively updated parameter estimates, $R^*_{\tau+1}(m_{i_\tau}(\widehat{\boldsymbol{\theta}}_\tau, \mathbf{z}_\tau), m_{i_{\tau+1}}(\widehat{\boldsymbol{\theta}}_{\tau+1}, \mathbf{z}_{\tau+1}))$, the distribution of returns conditional on investors' forecasting variables is likely to be non-stationary. This has consequences both for forecast evaluation and for construction of forecasting models.

*B. Empirical Evidence*

There does not appear to be any surveys of market efficiencies that focus on the implications of the EMH for nonstationarities in returns that we have emphasized so far. For example, Fama (1991) focuses mainly on informational efficiency of security prices. However, some recent studies find evidence that appears to be consistent with the notion of self-destruction of predictable patterns in returns in that previously documented predictability disappeared at a time when a consensus was emerging that predictable patterns were present.

Dimson and Marsh (1999) find that the small-cap premium disappeared in the UK stock market after it became publicly known. Bossaert and Hillion (1999) investigate predictability of monthly stock returns in a variety of international stock markets and find that the apparent in-sample predictability breaks down out-of-sample some time around 1990. Aiolfi and Favero (2002) report that predictability in US stocks documented in earlier studies seems to have disappeared in the 1990s. Their results are consistent with the notion that the best model selected in 'real time' based on in-sample performance ceases to have predictive power out-of-sample.

Likewise, Sullivan, Timmermann and White (1999) find that the apparent historical ability of technical trading rules to generate excess returns has broken down after 1986. Intriguingly, Brock, Lakonishok and Lebaron (1992) found evidence of profitability from following technical trading rules using data up to 1986. Although their paper was only published in 1992, it was available well before that. The same factors that alerted these authors to the apparent success of technical trading rules may have brought traders to exploit these strategies and have led to the demise of these prediction rules.

## V. Which Forecasting Approaches may work?

We have argued that traditional time series forecasting methods relying on individual forecasting models or stable combinations of these are not likely to be useful. It is possible to return to fairly simple methods that are able to adapt or learn quickly and that can be applied in large numbers to a multitude of return series searching for possible 'hot spots' where forecastability is available. By the behavior of their forecast errors, they may be able to detect many such places, if they exist, where investment will be profitable. There will also be many false leads which will lead to unprofitable investments.

### *A. Wide Searches Across Models and Assets*

One possible procedure is to conduct a wide-spaced search for predictability across the many models in $M_t$. The methods envisaged in this search could include simple forms of nearest neighbor, genetic algorithms and neural networks. Since the set of models is high-dimensional, it is useful to consider two very different forecasting approaches,

*Random Selection Approach.* Randomly selected techniques are applied to randomly selected returns, each time interval. If there appears to be evidence of forecastability in recent periods from some evaluation procedure, a possibly profitable investment could occur.

Alternatively, one could consider

*Comprehensive Selection Approach.* All techniques are applied to all returns at all times. This method is both truly comprehensive but also more expensive computationally, but this cost is becoming less relevant each year.

The advantage of the latter approach is that some assets may be 'under-researched' and their prices not set efficiently. In a major stock exchange there will be about 20 to 30 thousand assets, including derivatives, producing returns, giving perhaps 200,000 asset returns world wide. The contents of $M_t$ and the evaluation process used will depend on the forecast organizer. Only application of the techniques will tell if they can actually produce worthwhile profits.

Advanced forecasting strategies have not yet been as extensively used in security selection from large cross-sections of assets as in the modeling of major indices such as the S&P500 stock market index. For sure, some pricing anomalies related to firm characteristics such as book-to-market value and firm size (market capitalization) have been proposed, but it is not clear that the list is limited to these attributes. Again a problem posed by a wider cross-sectional search is the effects of 'breaks' in the forecasting model that is likely to arise once detected patterns are published. Detection and exploration of predictable patterns in asset prices can only work if their self-destruction works sufficiently slowly to enable econometric methods to have sufficient power to identify them and produce reasonably precise out-of-sample forecasts. Returns data are very noisy and the predictive $R^2-$values tend to be low. This means that long data samples are required for identification and estimation of prediction models.

While undoubtedly such systematic data mining already is practiced both by individual forecasters and certainly across different forecasting groups, it is also highly likely to lead to spurious in-sample predictability that will not continue to prevail out-of-sample. If sufficiently many forecasting models and trading rules are considered on a finite data sample, by pure chance some apparently successful strategies will show up even if, truly, they do not work. Hence for this a-theoretical strategy to work, it is necessary to control for the data-mining that it involves. Sullivan, Timmermann and White (1999) provide an application to technical trading rules that deals with this issue.

The literature on forecasting with nonlinear models suggests that these models only produce accurate predictions that improve upon simple linear alternatives in limited parts of the sample space. One could certainly imagine a 'regime switching' strategy that used different types of linear or nonlinear models in different blocks in time. The real question is whether financial data is too noisy and has too little persistence to successfully identify such regimes.

*B. Selection of Data Window*

Nonstationarities in returns introduced by large macroeconomic shocks, the markets' learning and institutional shifts means that the question of how to set the data window used to estimate the parameters of the forecasting model is likely to receive increased attention in the future. It is common practice to use rolling windows of five years of data in finance, but it is not clear why this strategy would be optimal in the presence of the types of nonstationarities found in finance. Another possibility is to use a geometrically declining discount factor on historical data which weights older data less than more recent data.

Pesaran and Timmermann (2002) propose instead a breakpoint monitoring method that attempts to detect breaks in the forecasting model and determines the data window on the basis of the outcome of such real-time break tests. They propose a reversed ordered Cusum (ROC) method that reverses the order of the data and tests for the most recent break point in a given prediction model. Data after the most recent break is then used to estimate the forecasting model, although it may also be optimal to use pre-break data in some circumstances. Their empirical findings suggest that the market timing information in the predictions is improved by accounting for breaks.

*C. Thick Modeling*

There is no lack of techniques that appear to have been successful in forecasting returns on some occasions as seen from the contents of the recent two-volume collection of papers by Mills (2002). For example, Lo and Mackinlay (1997) discuss portfolios of stock returns and bonds chosen to maximize predictability and claim to have found many examples. Occasional successes with nonlinear models, such as interest rates, threshold autoregressive models, and nearest-neighbor techniques have also been reported, but none of these have been consistently successful. To this list could be added time-varying parameter regressions and autoregressions. There is clearly no shortage of candidate models that can be considered in a forecasting search engine, either individually or in combination. If a particular asset becomes temporarily forecastable, as is suggested here, then we would expect several of these techniques to indicate it, contemporaneously, but some will be more efficient than others.

Rather than "thin modeling" where decisions are based on just a single "best" model, we recommend the use of "thick modeling" where a decision is based on a combination of outputs of models with statistically similar outputs, c.f. Granger (2000). The best technology for doing this is still being developed and critical values for the eventual output have to be constructed by bootstrap simulations. This aspect is acceptable in macroeconomic but possibly not yet in high-speed finance.

In a very promising application of this approach adopted to monthly returns on US stocks, Aiolfi and Favero (2002) find that accounting for model uncertainty using this type of approach leads to substantial improvement in the asset allocation performance based on recursive predictions.

*D. Modeling the Predictive Density of Returns*

A final area that we view as very promising in future work is the modeling of the predictive density of returns. Derivative contracts such as options and futures are now traded on many assets. For such assets it is natural to conduce efficiency tests jointly on the returns on both the derivatives and the underlying asset. Predictability no longer covers only the first moment of returns on the underlying asset since the payoff on options is a nonlinear function of the asset return.

Consider a call option with a strike price, $X$, trading at a market price of $Call_t$. At the option's expiration $(t+1)$ the option's payoff will be $Max(0, P_{t+1} - X)$. Suppose that the investor knows the pricing kernel, $Q_{t+1}$, and has a model for the joint probability distribution of $Q_{t+1}$ and $P_{t+1}$, $F(Q_{t+1}, P_{t+1}|\Omega_t)$. The theoretical value of the option should be

$$\int_Q \int_X^\infty Q_{t+1} Max(0, P_{t+1} - X) dF(Q_{t+1}, P_{t+1}|\Omega_t),$$

If this value is sufficiently different from the current market value of the option to cover transaction costs, then it is possible that a profitable trading strategy can be designed.

Notice that the kinked option payoff means that it is generally insufficient to price options solely on the basis of knowledge of the expected payoff on the underlying stock. Tracing out the above expression for different values of $X$ effectively requires knowledge of the full conditional probability density of $R_{t+1}$ and the EMH thus requires modeling not just the first conditional moment but the full conditional probability distribution of $R_{t+1}$ given information at time $t$.

The EMH therefore does not imply that all changes in this density are unpredictable. It does, however, require that certain functions of the probability distribution are not predictable.

As a concrete example, there is now substantial evidence that volatility of asset returns varies over time in a way that can be partially predicted. For this reason there has been considerable interest in improved volatility forecasting models in the context of option pricing, see e.g. Engle, Hong, Kane and Noh (1993). Does this violate market efficiency? Clearly the answer is no unless a trading strategy could be designed that would use this information in the options markets to identify under- and over-valued options. If options markets are efficient, option prices should incorporate the best volatility forecasts at all points in time.

To our knowledge no similar results exist yet for the full predictive density of asset returns. However it is likely that the methods now being developed for predicting the conditional skew, kurtosis and higher order moments of asset returns will also find some use in tests of market efficiency.

## VI. Conclusion

The classic EMH does not discuss how the information variables in $\Omega_t$ are used to produce actual forecasts. It is not difficult to change the definition of market efficiency to cover this aspect. For example, Jensen's (1978) definition could be extended as follows:

*A market is efficient with respect to the information set, $\Omega_t$, search technologies, $S_t$, and forecasting models, $M_t$, if it is impossible to make economic profits by trading on the basis of signals produced from a forecasting model in $M_t$ defined over predictor variables in the information set $\Omega_t$ and selected using a search technology in $S_t$.*

If the behavior of investors produces efficient markets by their continuous profit seeking, the reverse is that the EMH does not rule out predicting many other variables that, although of general interest, are not the basis for a profit making strategy. A simple example, discussed by Theil (1966) looked at the number of companies whose share advanced in a day minus the number that declined on the Amsterdam stock exchange. He found this quantity not white noise, and later studies on the New York Stock Exchange reached similar conclusions. In each case one must ask how a particular variable fits the criterion of being useful in designing profitable trading strategies. As an example consider abnormal trading volume. This is used as an input variable in many technical trading strategies and also affects the liquidity and costs of trading strategies. This may well explain why it is considered very difficult to predict.

Ultimately, there are likely to be short-lived gains to the first users of new financial prediction methods. Once these methods become more widely used, their information may get incorporated into prices and they will cease to be successful. This race for innovation coupled with the market's adoption of new methods is likely to give rise to many new generations of financial forecasting methods.

## References


[1] Aiolfi, M. and C.A. Favero (2002). Model Uncertainty, Thick Modeling and the Predictability of Stock Returns. Mimeo, Bocconi University.

[2] Bachelier, L. (1964). Theory of Speculation. In: P. Cootner (Ed.), *The Random Character of Stock Market Prices*. Cambridge: Massachusetts Institute of Technology Press.

[3] Black, F. (1986). Noise. *Journal of Finance,* 41, 529-543.

[4] Bossaert, P., and Hillion, P. 1999. Implementing Statistical Criteria to Select Return Forecasting Models: What do we Learn? *Review of Financial Studies*, 12, 405-428.

[5] Breeden, D., M. Gibbons and R. Litzenberger, 1989, Empirical Test of the Consumption-Oriented CAPM. *Journal of Finance* 44, 231-262.

[6] Brock, W., J. Lakonishok and B. LeBaron, 1992. Simple Technical Trading Rules and the Stochastic Properties of Stock Returns. *Journal of Finance* 47, 1731-1764.

[7] Cootner, P. (ed.) (1964). *The Random Character of Stock Market Prices.* Cambridge: Massachusetts Institute of Technology Press.

[8] Diebold, F.X. and R. Mariano, 1995. Comparing Predictive Accuracy. *Journal of Business and Economic Statistics* 13, 253-65.

[9] Dimson, E. and P. Marsh, 1999. Murphy's law and market anomalies. *Journal of Portfolio Management*, 25(2), 53-69.

[10] Engle, R., C. Hong, A. Kane and J. Noh. 1993. Arbitrage Valuation of Variance Forecasts Using Simulated Options. *Advances in Futures and Options Research* 6, 393-415.

[11] Engle, R., D. Lilien and R. Robins, 1987, Estimating Time-Varying Risk Premia in the Term Structure: The ARCH-M Model. *Econometrica* 55, 391-407.

[12] Evans, G.W. (1991). Pitfalls in Testing for Explosive Bubbles in Asset Prices. *American Economic Review*, 81, 922-930.

[13] Fama, E.F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 25, 383-417.

[14] Fama, E.F. (1991). Efficient Capital Markets II. *Journal of Finance*, 46, 1575-1617.

[15] Granger, C.W.J. (1999). Outline of Forecast Theory using Generalized Cost Functions. *Spanish Economic Review,* 1, 161-173.

[16] Granger, C.W.J. (2000). Thick Modelling. UCSD Mimeo.

[17] Hansen, L. and K. Singleton, 1983, Stochastic Consumption, Risk Aversion and the Temporal Behavior of Asset Returns. *Journal of Political Economy* 91, 249-68.

[18] Harrison, M. and D. Krebs (1979). Martingales and Arbitrage in Multiperiod Securities Markets, *Journal of Economic Theory,* 20, 381-408.

[19] Jensen, M. (1978). Some Anomalous Evidence Regarding Market Efficiency, *Journal of Financial Economics,* 6, 95-101.

[20] Lakonishok, J. and S. Smidt. 1988. Are Seasonal Anomalies Real? A Ninety-Year Perspective. *Review of Financial Studies* 1, 403-425.

[21] Lo, A.W. and A. C. MacKinlay. (1997). Maximizing Predictability in the Stock and Bond Markets. *Macroeconomic Dynamics* 1, 102-234.

[22] Malkiel, B. (1992). Efficient Market Hypothesis. In: Newman et al. (Eds.), *New Palgrave Dictionary of Money and Finance*, London: Macmillan.

[23] Mills, T. (ed.) (2002). Forecasting Financial Markets. Edward Elgar Publications. Cheltenham UK.

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

[25] Pesaran, M.H. and A. Timmermann (2002) Market Timing and Return Prediction under Model Instability. Forthcoming in *Journal of Empirical Finance.*

[26] Roberts, H. (1967). Statistical versus clinical prediction of the stock market. Unpublished manuscript, Center for Research in Security Prices, University of Chicago.

[27] Sullivan, R., Timmermann, A., White, H., 1999. Data-Snooping, Technical Trading Rule Performance, and the Bootstrap. *Journal of Finance,* 54, 1647-1692.

[28] Sullivan, R., A. Timmermann, & H. White (2002). Forecast Evaluation with Shared Data Sets, *International Journal of Forecasting* (forthcoming).

[29] Thaler, R.H., 1987. Anomalies: Weekend, Holiday, Turn of the Month, and Intraday Effects. *Journal of Economic Perspectives* 1, 169-178.

[30] Theil, H. (1966). *Applied Economic Forecasting*, Amsterdam: North-Holland.

[31] Timmermann, A. (1993). How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices, *Quarterly Journal of Economics,* 108, 1135-1145.