Technical Analysis and Theory of Finance

Yingzi Zhu
Tsinghua University

and

Guofu Zhou
Washington University in St. Louis and CCFR*

JEL classification: G11; G12, G14; C11; C61

Keywords: Technical analysis; trading rules; asset allocation; predictability; learning

First draft: May 1, 2006
Current version: September 17, 2007

*We are grateful to Kerry Back, William Brock, Henry Cao (the CICF discussant), Bjørn Eraker, Ramo Gencay, Ravi Jagannathan, Raymond Kan, Hong Liu, Jun Liu, Lars Lochstoer (the EFA discussant), Ľuboš Pástor, Avanidhar Subrahmanyam, Tan Wang and seminar participants at Fudan University, Renmin University of China, Tulane University, Tsinghua University, Washington University, and participants at the 2007 China International Conference in Finance and the 2007 European Finance Association Annual Meetings for helpful comments. Zhou acknowledges China Center for Financial Research (CCFR) of Tsinghua University for visit in May 2006 where part of the paper was written. Zhu acknowledges financial support from Minoru Kobayashi China Economic Research Fund. We also thank Lynnea Brumbaugh-Walter for many helpful editorial comments.

Corresponding Author: Guofu Zhou, Olin Business School, Washington University, St. Louis, MO 63130. Phone: (314) 935-6384 and e-mail: zhou@wustl.edu

Technical Analysis and Theory of Finance

In this paper, we analyze the usefulness of technical analysis, specifically the widely used moving average trading rule, from an asset allocation perspective. We show that when stock returns are predictable, technical analysis adds value to commonly used allocation rules that invest fixed proportions of wealth in stocks. When there is uncertainty about predictability, the fixed allocation rules combined with technical analysis can outperform the prior-dependent optimal learning rule when the prior is not too informative. Moreover, the technical trading rules are robust to model specification, and they tend to substantially outperform the model-based optimal trading strategies when there is uncertainty about the model governing the stock price.

1 Introduction

Technical analysis uses past prices and perhaps other past statistics to make investment decisions. Proponents of technical analysis believe that these data contain important information about future movements of the stock market. In practice, all major brokerage firms publish technical commentary on the market and many of the advisory services are based on technical analysis. In his interviews with them, Schwager (1993, 1995) finds that many top traders and fund managers use it. Moreover, Covel (2005), citing examples of large and successful hedge funds, advocates the use of technical analysis exclusively without learning any fundamental information on the market.

Academics, on the other hand, have long been skeptical about the usefulness of technical analysis, despite its widespread acceptance and adoption by practitioners.[1] There are perhaps three reasons. The first reason is that there is no theoretical basis for it, which this paper attempts to provide. The second reason is that earlier theoretical studies often assume a random walk model for the stock price, which completely rules out any profitability from technical trading. The third reason is that earlier empirical findings, such as Cowles (1933) and Fama and Blume (1966), are mixed and inconclusive. Recently, however, Brock, Lakonishok, and LeBaron (1992), and Lo, Mamaysky, and Wang (2000) find strong evidence of profitability in technical trading based on more data and more elaborate strategies. Statistically, though, it is difficult to show the true effectiveness of technical trading rules because of a data-snooping bias (see, e.g., Lo and MacKinlay, 1990), which occurs when a set of data is used more than once for the purpose of inference and model selection. In its simplest form, rules that are invented and tested by using the same data set are likely to exaggerate their effectiveness. Accounting for the data-snooping bias, for example, Sullivan, Timmermann, and White (1999) show via bootstrap that Brock, Lakonishok, and LeBaron's results are much weakened. Using generic algorithms, Allen and Karjalainen (1999) find little profitability in technical trading. One could then argue that a bootstrap is subject to specification bias and that generic algorithms can be inadequate due to inefficient ways of learning. In any case, it appears that the statistical debate on the effectiveness of technical analysis is unlikely to get settled soon.

Our paper takes a new perspective. We consider the theoretical rationales for using technical

[1] Some academics take a strong view against technical analysis. For example, in his influential book, Malkiel (1981) says "technical analysis is anathema to the academic world."

analysis in a standard asset allocation problem.[2] An investor is to decide how to allocate his wealth optimally between a riskless asset and a risky one which we call stock. For tractability, we analyze the profitability of the simplest and seemingly the most popular technical trading rule – the moving average (MA) – which suggests that investors buy the stock when its current price is above its average price over a given period L.[3] The immediate question is what proportion of wealth the investor should allocate into the stock when the MA signals so. Previous studies use an all-or-nothing approach: the investor invests 100% of his wealth into the stock when the MA says 'buy', and nothing otherwise. This common and naive use of the MA is, in fact, not optimal from an asset allocation perspective because the optimal amount should be a function of the investor's risk aversion as well as the degree of predictability of the stock return. Intuitively, if the investor invests an optimal fixed portion of his money into the stock market, say 80%, when there is no MA signal, he should invest more than 80% when the MA signals a buy, and less so otherwise. The 100% allocation is therefore unlikely to be optimal. For a log-utility investor, we solve the problem of allocating the optimal amount of stock explicitly, which provides a clear picture of how the degree of predictability affects the allocation decision given the log-utility risk tolerance. We also solve the optimal investment problem both approximate analytically and via simulations in the more general power-utility case. The results show that the use of the MA can help increase the investor's utility substantially.

Moreover, given any investment strategy that allocates a fixed proportion of wealth to the stock, we show that the MA rule can be used in conjunction with to yield higher expected utility. In particular, it can improve the expected utility substantially for the popular fixed strategy that follows Markowitz's (1952) modern portfolio theory and Tobin's (1958) two-fund separation theorem. Since indexing, a strategy of investing in a well-diversified portfolio of stocks, comprises roughly one-third of the US stock market, and its trend is on the rise worldwide (see, e.g., Bhattacharya and Galpin (2006)), and since popular portfolio optimization strategies (see, e.g., Litterman, 2003, and Meucci, 2005) are also fixed strategies, any improvement over fixed strategies is of practical importance, which might be one of the reasons that technical analysis is widely used in practice.[4]

[2] In models of information asymmetry, Borwn and Jennings (1989), and Kim and Grundy (2002) show that technical analysis can have value.

[3] As time passes, the average price is always computed based on its current price and on those in the most recent L periods, and hence the average is called the moving average.

[4] Behaviorial reasons, such as limited attention and optimal learning with limited resources, may explain the use of simple technical rules in practice.

However, since the MA, as a simple filter of the available information on the stock price, disregards any information on predictable variables, trading strategies related to the MA must be in general dominated by the optimal dynamic strategy, which optimally uses all available information on both the stock price and on the predictable variables. An argument in favor of the MA could be that the optimal dynamic strategy is difficult for investors at large to implement due to both the difficulty of model identification and to the cost of collecting and to processing information. In particular, it is not easy to find predictable variables, nor are their observations at desired time frequencies readily available. This gives rise to the problem of predictability uncertainty in practice. In the presence of such uncertainty, Gennotte (1986), Barberis (2000) and Xia (2001), among others, show that the optimal dynamic strategy will depend on optimal learning about the unknown parameters of the model, and that, in turn, will depend on the investor's prior on the parameters. In the context of Xia's (2001) model, we find, interestingly, that, with the use of the MA rule, one can in fact outperform the optimal dynamic trading strategy when the priors are reasonable and yet not too informative. This seems due to the fact that the MA rule is less model dependent, and so it is more robust to the choice of underlying predictable variables.

Furthermore, the usefulness of the MA rule is more apparent when there is uncertainty about which model truly governs the stock price. In the real world, the true model is unknown to all investors. But for a wide class of plausible candidates of the true model, the optimal MA can be estimated easily, while the optimal trading strategy relies on a complete specification of the true model. When the wrong model is used to derive the optimal trading strategy, we show that the estimated optimal MA outperforms it substantially.

In typical applications, one usually chooses some ex-ante value as the lag length of the MA. The question of using the optimal lag has been done only by trial-and-error, and only for the pure MA strategy that takes an all-or-nothing allocation. Since this allocation itself is suboptimal, the associated optimal lag is suboptimal too. The asset allocation perspective provided here not only solves the optimal stock allocation problem for both the pure MA and its optimal combination with the fixed rules, but also determines the optimal lag of the MA. We find that the fixed rules in conjunction with the MA are fairly insensitive to the use of the optimal lags, while the optimal generalized MA is not.

The paper is organized as follows. Section 2 outlines the model and various investment strategies. Section 3 provides their explicit solutions in the log-utility case, while Section 4 compares them both approximate analytically and numerically in the power-utility case. Section 5 analyzes the strategies when there is parameter uncertainty, and Section 6 studies the case when there is model uncertainty. Section 7 explores the optimal choice of the MA lag length. Section 8 concludes.

2 The Model and Various Investment Strategies

For simplicity, we consider a two-asset economy in which a riskless bond pays a constant rate of interest r, and a risky stock can represent the aggregate equity market. Because of the ample evidence on the predictability of stock returns,[5] we, following Kim and Omberg (1996), and Huang and Liu (2006), among others, assume the following dynamics for the cum-dividend stock price S_t:

$$\frac{dS_t}{S_t} = (\mu_0 + \mu_1 X_t)dt + \sigma_s dB_t, \tag{1}$$
$$dX_t = (\theta_0 + \theta_1 X_t)dt + \sigma_x dZ_t, \tag{2}$$

where $\mu_0, \mu_1, \sigma_s, \theta_0, \theta_1$ and σ_x are parameters; X_t is a predictive variable; and B_t and Z_t are standard Brownian motions with correlation coefficient ρ. Note that θ_1 has to be negative to make X_t a mean-reverting process. The model is a special case of the general models of Merton (1992). In discrete-time, it is the well-known predictive regression model (e.g., Stambaugh (1999)).

Given an initial wealth W_0 and an investment horizon T, the standard allocation problem of an investor is to choose a portfolio strategy ξ_t to maximize his expected utility of wealth,

$$\max_{\xi_t} E[u(W_T)] \tag{3}$$

subject to the budget constraint

$$dW_t = rW_t\, dt + \xi_t(\mu_0 + \mu_1 X_t - r)dt + \xi_t \sigma_s dB_t. \tag{4}$$

The solution to this problem is the optimal trading strategy. In general, this strategy is a function of time and the associated state variables. We will refer to it in what follows as the optimal dynamic strategy, as it varies with time and states.

[5]There is a huge literature on predictability, examples of which are Fama and Schwert (1977), Campbell (1987), Ferson and Harvey (1991), Goyal and Welch (2003), and Ang and Bekaert (2007). Kandel and Stambaugh (1996), Barberis (2000), Pan and Liu (2004), and Huang and Liu (2007) are examples of studies on portfolio choice under predictability.

In this paper, we assume the power-utility

$$u(W_T) = \frac{W_T^{1-\gamma}}{1-\gamma}, \tag{5}$$

where γ is the investor's risk aversion parameter. In this case, the optimal dynamic strategy is known (see, e.g., Kim and Omberg, 1996, and Huang and Liu, 2007) and is given by

$$\xi_t^* = \frac{\mu_0 + \mu_1 X_t - r}{\gamma\sigma_s^2} + \frac{(1-\gamma)\rho\sigma_x}{\gamma\sigma_s}[\chi(t) + \zeta(t)X_t], \tag{6}$$

where $\chi(t)$ and $\zeta(t)$ satisfy the following ordinary differential equations:

$$\dot{\chi}(t) + a_1\zeta(t)\chi(t) + \frac{1}{2}a_2\chi(t) + a_4\zeta(t) + a_5 = 0, \tag{7}$$

$$\dot{\zeta}(t) + a_1\zeta^2(t) + a_2\zeta(t) + a_3 = 0, \tag{8}$$

with

$$a_1 = \frac{(1-\gamma)^2}{\gamma}\rho^2\sigma_x^2 + (1-\gamma)\sigma_x^2, \quad a_2 = 2\left(\frac{1-\gamma}{\gamma}\frac{\mu_1}{\sigma_s}\rho^2\sigma_x^2 + \theta_1\right),$$

$$a_3 = \frac{1}{\gamma}\left(\frac{\mu_1}{\sigma_s}\right)^2, \qquad a_4 = \frac{1-\gamma}{\gamma}\frac{\mu_0 - r}{\sigma_s}\rho\sigma_x + \theta_0, \quad a_5 = \frac{\mu_1(\mu_0 - r)}{\gamma\sigma_s^2},$$

and the terminal conditions $\chi(T) = \zeta(T) = 0$.

The assumption that stock returns are independently and identically distributed (iid) over time has played a major role in finance. It was the basis for much of the earlier market efficiency arguments, though was known later as only a sufficient condition. Nevertheless, some of the most popular investment strategies and theoretical models are based on this assumption. Under the iid assumption, the optimal strategy is

$$\xi_{\text{fix1}}^* = \frac{\mu_s - r}{\gamma\sigma_s^2}, \tag{9}$$

where μ_s is the long-term mean of the stock return. This strategy invests a fixed or constant portion of wealth, ξ_{fix1}^*, into the stock all the time. In discrete-time, this is the familiar suggestion of Markowitz's (1952) mean-variance framework and Tobin's (1958) two-fund separation theorem.[6] The strategy is one of the most important benchmark models used in practice today (see, e.g., Litterman (2003) and Meucci (2005)). Because of it, passive index investments have become increasingly popular (Rubinstein (2002)). Theoretically, the allocation rule ignores any time-varying investment opportunities and is clearly not optimal once the iid assumption is violated. A likely

[6] See Ingersoll (1987) or Back (2006) for an excellent textbook exposition.

practical motivation for its wide use is as follows. Although the stock returns are predictable, the predictability is small and uncertain. It could be costly for a small investor to collect news and reports about X_t whose benefits may outweigh the costs. As a result, the investor may simply follow a constant allocation rule even though there is a small degree of predictability.

The fixed rule ξ^*_{fix1} ignores any predictability completely. An interesting question is, then, whether one can obtain yet another fixed rule that accounts for the predictability. In other words, how should the investor invest his money when he knows the true predictable process but not the state variables? Mathematically, this amounts to solving the optimal allocation problem by restricting ξ_t to a constant. The solution is analytically obtained as (all proofs are given in the Appendix)

$$\xi^*_{\text{fix2}} = \frac{\mu_s - r}{\gamma\sigma_s^2 - (1-\gamma)(\mu_1^2 A + 2\mu_1\sigma_s B)}, \tag{10}$$

where

$$A = \frac{\sigma_x^2}{\theta_1^2}\left(1 + \frac{1 - e^{\theta_1 T}}{\theta_1 T}\right), \quad B = \frac{\rho\sigma_x}{\theta_1}\left(\frac{e^{\theta_1 T} - 1}{\theta_1 T} - 1\right).$$

Here we see that, for $\gamma = 1$, this optimal *constant* strategy is equal to ξ^*_{fix1}. In other words, for investors with log-utility, the optimal fixed strategy remains the same as before, even though the stock returns are predictable, a fact we can explain largely by the myopic behavior dictated by the log-utility. For $\gamma > 1$, however, there is an adjustment in the denominator of (10). In general, the adjustment can be either positive or negative.

Among the technical trading rules, the rule that is based on the moving average of stock prices is the most popular one. Let $L > 0$ be the lag or the lookback period. A moving average (MA) of the stock price at any time t is defined as

$$A_t = \frac{1}{L}\int_{t-L}^{t} S_u \, du, \tag{11}$$

i.e., the average price over time period $[t - L, t]$. The simplest MA trading rule is the following stock allocation strategy,[7]

$$\eta_t = \eta(S_t, A_t) = \begin{cases} 1, & \text{if } \quad S_t > A_t; \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

This is well defined when $t > L$, and can be taken as zero or as another fixed constant when $t \leq L$.[8] This standard (pure) moving average rule is a market timing strategy that shifts investments

[7] In practice, the MA rule is computed based on ex-dividend prices which will be analyzed in Section 3.
[8] The Appendix discusses how we choose the initial value of an MA rule.

between cash and stock. Almost all existing studies on the MA strategy take a 100% position in the stock or nothing, i.e., the portfolio weight (on the stock) is η_t. This is clearly not optimal for two reasons. First, the MA rule should in general be a function of the risk-aversion parameter γ. Intuitively, γ reflects the investor's tolerance to stock risk, and it has to enter the allocation decision as is the case for the earlier optimal fixed strategies. Second, the degree of predictability must matter. The more predictable the stock, the more reliable the MA rule and hence the more allocation to the stock.

Other than the pure MA rule, we also consider the following generalized MA (GMA) rule,

$$\mathrm{GMA}(S_t, A_t, \gamma) = \xi_{\mathrm{fix}} + \xi_{\mathrm{mv}} \cdot \eta(S_t, A_t), \tag{13}$$

where ξ_{fix} and ξ_{mv} are constants. This trading strategy is a linear combination of a fixed strategy and a *pure* moving average strategy. It consists of all the previous strategies as special cases. For example, ξ^*_{fix1} is obtained by setting $\xi_{\mathrm{fix}} = \xi^*_{\mathrm{fix1}}$ and $\xi_{\mathrm{mv}} = 0$, and η_t is obtained by setting $\xi_{\mathrm{fix}} = 0$ and $\xi_{\mathrm{mv}} = 1$.[9]

There are three interesting questions associated with the GMA rule. First, what is the optimal choice of ξ_{fix} and ξ_{mv}, and how well does it perform as compared with other fixed strategies? Second, ξ_{fix} being equal to either ξ^*_{fix1} or ξ^*_{fix2}, whether the optimal choice of ξ_{mv} is zero or not indicates whether there is a gain in the expected utility when the fixed strategy is used in conjunction with the MA rule. Third, imposing $\xi_{\mathrm{fix}} = 0$, the optimal choice of ξ_{mv} indicates the optimal amount of investment based purely on the MA trading signal. If $\xi_{\mathrm{mv}} = 1$, the usual application of the MA with 100% stock allocation is optimal. However, as easily seen from our analysis later, the optimal value of ξ_{mv} is unlikely to be equal to one. These three questions will be answered first analytically for the log-utility, and then numerically for the power-utility.

Analytically, the distribution of the arithmetic moving average A_t is very complex and difficult to analyze. On the other hand, the geometric moving average,

$$G_t = \exp\left(\frac{1}{L}\int_{t-L}^{t} \log(S_u)\, du\right), \tag{14}$$

[9] It should be noted that if the initial state X_0 is drawn from the steady-state distribution, the investor will know X_0 when he chooses the constants ξ's in the GMA rule. Hence, the optimal ξ's will depend on X_0. However, our goal here is to find the optimal fixed ξ's that are independent of initial conditions. In other words, we solve in what follows the optimal allocation problem using the steady state distribution for X_0.

is tractable to allow explicit solutions. In addition, as shown in our later simulations, there are little performance differences in our main results with the use of either averages. Henceforth, we will focus our analysis on GMA(S_t, G_t, γ), i.e., the generalized MA strategy based on the geometric average.

3 Explicit Solutions: The Log-utility Case

In this section, we provide the explicit solutions of the optimal GMA strategies and compare them analytically with both the optimal fixed and the optimal dynamic allocations.

Under the stationarity condition for X_t, the wealth process corresponding to the GMA is

$$\frac{dW_t}{W_t} = [r + \text{GMA} \cdot (\mu_0 + \mu_1 X_t - r)]dt + \text{GMA} \cdot \sigma_s dB_t,$$

and hence, assuming $T > L$, we have

$$\begin{aligned}\log W_T = \log W_0 + rT &+ \int_0^L dt[\xi^*_{\text{fix1}}(\mu_0 + \mu_1 X_t - r - \frac{\sigma_s^2}{2}\xi^*_{\text{fix1}})] \\ &+ \int_L^T dt[\xi_{\text{fix}}(\mu_0 + \mu_1 X_t - r - \frac{\sigma_s^2}{2}\xi_{\text{fix}})] + \xi_{\text{mv}}\mu_1 \int_L^T dt\hat{X}_t\eta_t \\ &+ \int_L^T dt[\xi_{\text{mv}}(\mu_0 + \mu_1\bar{X} - r) - \frac{\sigma_s^2}{2}\xi_{\text{mv}}^2 - \sigma_s^2\xi_{\text{fix}}\xi_{\text{mv}}]\eta_t + \sigma_s \int_L^T (\xi_{\text{fix}} + \xi_{\text{mv}}\eta_t)dB_t, \quad (15)\end{aligned}$$

where $\hat{X}_t = X_t - \bar{X}$ with $\bar{X} = -\theta_0/\theta_1$. Then the expected log-utility is[10]

$$\begin{aligned}U_{\text{GMA}} = E\log W_T \;=\; & \log W_0 + rT + \frac{(\mu_0 + \mu_1\bar{X} - r)^2}{2\sigma_s^2}L \\ &+ \int_L^T dt\xi_{\text{fix}}[\mu_0 + \mu_1\bar{X} - r - \frac{\sigma_s^2}{2}\xi_{\text{fix}}] + \int_L^T dt\xi_{\text{mv}}\mu_1 E[\hat{X}_t\eta_t] \\ &+ \int_L^T dt[\xi_{\text{mv}}(\mu_0 + \mu_1\bar{X} - r) - \frac{\sigma_s^2}{2}\xi_{\text{mv}}^2 - \sigma_s^2\xi_{\text{fix}}\xi_{\text{mv}}]E[\eta_t]. \quad (16)\end{aligned}$$

To solve the optimization problem, let

$$b_1 \equiv E[\hat{X}_t\eta(S_t, G_t)], \qquad b_2 \equiv E[\eta(S_t, G_t)], \quad (17)$$

where b_1 is the covariance between X_t and the moving average strategy η_t and b_2 is the probability

[10]Consistent with footnote 9, the expectation operator E here is taken conditional on information set at $t = 0$ and with respect to the initial steady state distribution of X_0.

of $S_t > G_t$ at any given time. We show in the Appendix that

$$b_1 = E\hat{X}_t\eta(S_t, G_t) = \frac{C_{12}^Z}{\sqrt{C_{22}^Z}} N'(-\frac{m_2^Z}{\sqrt{C_{22}^Z}}), \quad (18)$$

$$b_2 = E\eta(S_t, G_t) = N(\frac{m_2^Z}{\sqrt{C_{22}^Z}}), \quad (19)$$

where

$$C_{12}^Z = (\frac{\mu_1\sigma_x^2}{2\theta_1^2} - \frac{\sigma_x\sigma_s\rho}{\theta_1})(1 - \frac{e^{\theta_1 L} - 1}{\theta_1 L}), \quad (20)$$

$$C_{22}^Z = (\sigma_s^2 + \frac{\mu_1^2\sigma_x^2}{\theta_1^2} - \frac{2\mu_1\sigma_x\sigma_s\rho}{\theta_1})\frac{L}{3} + (\frac{\mu_1^2\sigma_x^2}{2\theta_1^3} - \frac{\mu_1\sigma_x\sigma_s\rho}{\theta_1^2})\left[1 - \frac{2}{(\theta_1 L)^2}(1 - e^{\theta_1 L} + \theta_1 L e^{\theta_1 L})\right], \quad (21)$$

$$m_2^Z = (\mu_0 + \mu_1\bar{X} - \frac{\sigma_s^2}{2})\frac{L}{2}, \quad (22)$$

and $N(\cdot)$ and $N'(\cdot)$ are the distribution and density functions of the standard normal random variable, respectively. Since we assume X_t starts from its steady state distribution,[11] it turns out that b_1 and b_2 are independent of time t. Therefore, the expected log-utility of (16) becomes

$$\begin{aligned} U_{\text{GMA}} = E\log W_T \ = \ & \log W_0 + rT + \frac{(\mu_0 + \mu_1\bar{X} - r)^2}{2\sigma_s^2}L \\ & + \xi_{\text{fix}}[\mu_0 + \mu_1\bar{X} - r - \frac{\sigma_s^2}{2}\xi_{\text{fix}}](T - L) + \xi_{\text{mv}}\mu_1 b_1(T - L) \\ & + [\xi_{\text{mv}}(\mu_0 + \mu_1\bar{X} - r) - \frac{\sigma_s^2}{2}\xi_{\text{mv}}^2 - \sigma_s^2\xi_{\text{fix}}\xi_{\text{mv}}]b_2(T - L). \end{aligned} \quad (23)$$

With these preparations, we are ready to answer the three questions raised earlier. In doing so, we assume the investment horizon T is greater than or equal to the lag length L throughout. This assumption is clearly harmless.

3.1 Optimal GMA

On the question of finding an optimal fixed strategy that combines a fixed rule with the MA, the results are given by the following:

Proposition 1 In the class of strategies GMA(S_t, G_t, γ), the optimal choice of ξ_{fix} and ξ_{mv} under

[11] See, e.g., Karatzas and Shreve (1991, p. 358) for a discussion on the steady state. The details of the derivations are given in the Appendix of this paper.

the log-utility is

$$\xi^*_{\text{fix}} = \frac{\mu_s - r}{\sigma_s^2} - \frac{\mu_1 b_1}{(1-b_2)\sigma_s^2}, \tag{24}$$

$$\xi^*_{\text{mv}} = \frac{\mu_1 b_1}{b_2(1-b_2)\sigma_s^2}, \tag{25}$$

and the associated value function is

$$U^*_{\text{GMA1}} = U^*_{\text{fix1}} + \frac{\mu_1^2 b_1^2}{2b_2(1-b_2)\sigma_s^2}(T-L) \geq U^*_{\text{fix1}}, \tag{26}$$

where U^*_{fix1} is the value function associated with ξ^*_{fix1}.

Proposition 1 says that the improvement over ξ^*_{fix1} is always positive by combining a suitable fixed strategy with the moving average one unless $\mu_1 = 0$. In the case of $\mu_1 = 0$, the price return is unpredictable, and the fixed strategy ξ^*_{fix1} is optimal already. The point is that ξ^*_{fix1} is not optimal in general, and so the MA rule can help gain expected log-utility with the combination of another fixed strategy. Recall that, in the log-utility case, $\xi^*_{\text{fix2}} = \xi^*_{\text{fix1}}$. Hence, Proposition 1 applies to ξ^*_{fix2} as well, and ξ^*_{fix1} is the only fixed strategy to compare with.

It is interesting to observe that

$$\xi^*_{\text{fix}} + (b_2\xi^*_{\text{mv}}) = \xi^*_{\text{fix1}}. \tag{27}$$

If the predictable variable X_t is positively related to the stock market with $\mu_1 > 0$, the investor invests less than the standard fixed strategy by the amount of $b_2\xi^*_{\text{mv}}$ since $0 < b_2 < 1$ and $\xi^*_{\text{mv}} > 0$. Once the trend is up, as suggested by the moving average rule, the investor is more aggressive than the fixed strategy by investing an extra amount of $(1-b_2)\xi^*_{\text{mv}}$. This is consistent with the intuition that one should take advantage of the predictability of the stock market once it is detected by the MA rule.

If one strategy outperforms another over horizon T, it must continue to do so over a longer time. Hence, $U^*_{\text{GMA1}} - U^*_{\text{fix1}}$ must be an increasing function of T. What is striking here is that this relation is in fact linear in T in the log-utility case since b_1, b_2, μ_1 and σ_s are all horizon independent parameters.

Proposition 1 also makes possible an analytical comparison between GMA1 and the optimal dynamic strategy. Under the log-utility, the optimal dynamic rule (6) is the same as the myopic rule

$$\xi^*_{\text{opt}} = \frac{\mu_0 + \mu_1 X_t - r}{\sigma_s^2}.$$

By substituting this optimal rule into the wealth process, we obtain the optimal utility

$$U^*_{\text{opt}} = U^*_{\text{fix}} + \frac{1}{2}\frac{\mu_1^2 E\hat{X}_t^2}{\sigma_s^2}T. \tag{28}$$

Based on the value functions in both cases, we have

$$U^*_{\text{opt}} - U^*_{\text{GMA1}} \geq \frac{\mu_1^2}{2\sigma_s^2}\left[E\hat{X}_t^2 - \frac{b_1^2}{b_2(1-b_2)}\right](T-L). \tag{29}$$

Recalling that $b_1 = E\hat{X}_t\eta$ and $b_2 = E\eta$, we have $\text{var}(\eta) = E\eta^2 - (E\eta)^2 = b_2(1-b_2)$, and hence

$$\frac{b_1^2}{b_2(1-b_2)} = \frac{(E\hat{X}_t\eta)^2}{\text{var}(\eta)} = \frac{(\text{cov}(\hat{X}_t,\eta))^2}{\text{var}(\eta)} \leq \frac{E(\hat{X}_t^2)\text{var}(\eta)}{\text{var}(\eta)} = E\hat{X}_t^2.$$

Therefore, equation (29) is always positive, as it must be, since U^*_{opt} is the expected utility under the optimal dynamic strategy. It is seen that the smaller the σ_x^2, the smaller the difference. In other words, the less volatile the predictable variable, the closer the GMA1 to the optimal strategy. However, it should also be noted that, as σ_x^2 gets smaller, b_1 also gets closer to zero, i.e., the MA component becomes smaller too.

3.2 Combining A Fixed Rule with MA

Now, we answer the question of whether the moving average strategy can be used in conjunction with ξ^*_{fix1} to add value. To address this question, we need to solve the earlier optimization by imposing the constraint that $\xi_{\text{fix}} = \xi^*_{\text{fix1}}$. In this case, we have

Proposition 2 In the class of strategies GMA(S_t, G_t, γ) with ξ_{fix} being set at ξ^*_{fix1}, the optimal choice of ξ_{mv} under the log-utility is

$$\xi^*_{\text{mv}} = \frac{\mu_1 b_1}{b_2\sigma_s^2}, \tag{30}$$

and the associated value function is

$$U^*_{\text{GMA2}} = U^*_{\text{fix1}} + \frac{\mu_1^2 b_1^2}{2b_2\sigma_s^2}(T-L) \geq U^*_{\text{fix1}}, \tag{31}$$

where U^*_{fix1} is the value function associated with ξ^*_{fix1}.

As for U^*_{GMA1}, U^*_{GMA2} is at least as large as U^*_{fix1}. When there is predictability, it is clear that U^*_{GMA2} is strictly larger than U^*_{fix1}, implying that the MA rule helps to improve the expected utility, and does so strictly as long as the stock return is predictable.

An interesting observation is that ξ^*_{mv} in Proposition 2 differs from that in Proposition 1 by only a factor of $1 - b_2$ in the denominator. Because $0 < b_2 < 1$, ξ^*_{mv} is smaller now in absolute value. This is expected. Because ξ^*_{fix} is set at ξ^*_{fix1}, the risk exposure to the stock market is relatively high already as $\xi^*_{\text{fix1}} > \xi^*_{\text{fix}}$. Hence, when the MA rule detects an upward trend in the market, the investor acts more aggressively than ξ^*_{fix1}, but less aggressively than before. Finally, it is seen that

$$U^*_{\text{GMA2}} = U^*_{\text{GMA1}} - \frac{\mu_1^2 b_1^2}{2(1-b_2)\sigma_s^2}(T-L) \leq U^*_{\text{GMA1}} \leq U^*_{\text{opt}}. \tag{32}$$

While the second inequality, discussed earlier, is obvious, the first inequality should be true, too. The fixed component of GMA1 is optimally chosen, and hence its performance must be better than the GMA strategy with that component being set at ξ^*_{fix1}.

3.3 Optimal Pure MA

As discussed earlier, a standard or pure moving average rule is a market timing strategy that shifts money between cash and risky assets. Existing studies provide no guidance as to how much one should optimally invest in the stock even if one believes it is in an up-trend as signalled by the MA rule. Clearly, a 100% investment in the stock market is not optimal from a utility maximization point of view. Here we solve the optimal amount explicitly.

Proposition 3 In the class of strategies $\text{GMA}(S_t, G_t, \gamma)$ with restriction $\xi_{\text{fix}} = 0$, the optimal choice of ξ_{mv} under the log-utility is

$$\xi^*_{\text{mv}} = \frac{\mu_s - r}{\sigma_s^2} + \frac{\mu_1 b_1}{b_2 \sigma_s^2}, \tag{33}$$

and the associated value function, is

$$U^*_{\text{GMA3}} = U^*_{\text{fix1}} + \frac{(\mu_1 b_1 + (\mu_s - r)b_2)^2 - (\mu_s - r)^2 b_2}{2 b_2 \sigma_s^2}(T-L), \tag{34}$$

which can be either greater or smaller than U^*_{fix1}, the value function associated with ξ^*_{fix1}.

Consistent with our intuitive reasoning in the introduction, Proposition 3 says that, if an all-or-nothing investment strategy is taken based on the MA, the optimal stock allocation is unlikely to be 100%. Recognizing that 100% is not optimal, one may suggest a two-step approach for making use of the MA signal. In the first step, one determines the stock allocation, say ξ^*_{fix1}, based on a standard fixed allocation model, and then, in the second step, apply this in the market-timing decision: invest that amount into the stock if MA signals a 'buy', and nothing otherwise. Equation (33) says that this fixed amount differs from ξ^*_{mv} in general, and hence the decision is suboptimal too. The intuition is that one should invest more than that fixed amount if the trend is detected.

Proposition 3 also says that whether the pure MA strategy can outperform the fixed strategy depends on particular parameter values. It can be verified that, if the following relation about the risk premium is satisfied,[12]

$$\mu_s - r < \frac{\mu_1 b_1}{\sqrt{b_2} - b_2}, \tag{35}$$

the pure MA strategy does yield a higher expected utility than the fixed strategy ξ^*_{fix1}. However, with reasonable parameters calibrated from data, the above condition is not true. It implies that the optimal pure MA strategy usually performs worse than the simple fixed strategy. Indeed, our later simulations show that the pure MA strategy and its common analogues always perform the worst. Hence, if the MA rule is to be of any value to investors, it must be used wisely and in conjunction with the fixed strategies demonstrated by Propositions 1 and 2.

4 Comparison Under Power-Utility

In this section, we extend our earlier analysis to the power-utility case. First, we provide first-order accurate analytical solutions to the optimal GMA1, GMA2 and GMA3 strategies that provide insight on the role played by an investor's risk aversion parameter. Second, we derive second-order accurate analytical solutions to the strategies that are important for computing their performance under the power-utility. Finally, using data on S&P500 and three popular predictable variables from December 1926 to December 2004, we estimate the parameters of the model and examine how the different investment strategies differ in terms of their economic losses versus the optimal dynamic strategy.

[12] To appreciate the intuition behind the condition, we note that the denominator of the right hand side of the inequality is dominated by 0.25. Therefore, a sufficient condition for pure MA strategy to outperform a fixed rule is $\mu_1 b_1 > 4(\mu_s - r)$, which means that when mean reversion is stronger, the MA strategy is more likely to dominate the fixed rule. Similarly, if the equity premium is not too large, the MA strategy is more likely to dominate.

4.1 First-order Approximate Solutions

In the power-utility case, because of the complexity of the utility function, it is infeasible for us to derive an exact analytical solution for those trading strategies provided in the previous section. Nevertheless, we can obtain a first-order analytical approximation. The solutions are insightful as they reveal how the trading strategies are affected by γ, the investor's risk aversion.

By approximating $\int_0^T X_t dt$, $\int_0^T X_t \eta_t dt$ and $\int_0^T \eta_t dt$ with their mean values, we can write the expected utility under the GMA as

$$\begin{aligned} U_{\text{GMA}}(\gamma) \quad &\approx \quad \frac{(W_0 \exp(rT))^{1-\gamma}}{1-\gamma} \cdot \exp\left\{(1-\gamma)T\left[\xi_{\text{fix}}(\mu_0 + \mu_1 \bar{X} - r) - \frac{\gamma\sigma_s^2}{2}\xi_{\text{fix}}^2 + \xi_{\text{mv}}\mu_1 E[\hat{X}_t \eta_t]\right.\right. \\ &+ \quad \left.\left.[\xi_{\text{mv}}(\mu_0 + \mu_1 \bar{X} - r) - \frac{\gamma\sigma_s^2}{2}\xi_{\text{mv}}^2 - \gamma\sigma_s^2 \xi_{\text{fix}}\xi_{\text{mv}}]E\eta_t\right]\right\}. \end{aligned} \tag{36}$$

Optimizing this approximated utility function, we obtain

$$\text{GMA}(S_t, G_t, \gamma) = \frac{1}{\gamma}\text{GMA}(S_t, G_t, 1). \tag{37}$$

This says that the optimal generalized MA rules in the $\gamma \neq 1$ case is simply a scale of those in the log-utility case. Hence, much of the qualitative results obtained in the log-utility case carry over to the power-utility case, with accuracy up to the first-order approximation.

For example, the GMA1 strategy in the power-utility case is still of the earlier form, but with

$$\xi_{\text{fix}}^* = \frac{\mu_s - r}{\gamma\sigma_s^2} - \frac{\mu_1 b_1}{\gamma(1-b_2)\sigma_s^2}, \tag{38}$$

$$\xi_{\text{mv}}^* = \frac{\mu_1 b_1}{\gamma b_2(1-b_2)\sigma_s^2}. \tag{39}$$

This says that we simply scale down the stock investment by $1/\gamma$ when the investor is more risk-averse than the log-utility case. The same conclusion also holds for the GMA2 and GMA3 strategies. Interestingly, this scaling corresponds precisely to the way by which the usual fixed strategy is adjusted when the investor's preference changes from the log- to the power-utility. In particular, the optimal pure MA rule depends on γ. However, one should keep in mind that the simple inverse dependance on γ here is not exact, but only approximate with the first-order accuracy.

4.2 Second-order Approximate Solutions

While the previous approximate solutions make apparent the role of γ, they will not be accurate enough in simulations for measuring the true performance of the optimal GMA strategies, which are analytically unavailable. One may propose a numerical method, such as simulation, to compute the optimal GMA strategies, but this is feasible only for a given S_t, G_t and t. To evaluate the performance of these strategies, however, we need to compute the optimal GMA strategies at hundreds and thousands of draws of S_t and G_t and time t. Therefore, it is not possible to evaluate the performance of the optimal GMA strategies numerically without an efficient way to determine the strategies in the first place. To resolve this problem, we, in what follows, derive alternative analytical solutions to the strategies. These are more complex than the earlier ones, but are accurate to the second-order. As a compromise, they will be taken as the true strategies. Simulations will then be used to evaluate their performances.

Rather than ignoring the second-order terms of the random variables in (15), we approximate them by Gaussian processes that match both the first and second moments. Then, the power-utility,

$$U(\gamma) = \frac{1}{1-\gamma} E\left[W_T^{1-\gamma}\right] = \frac{1}{1-\gamma} E\left[\exp((1-\gamma)\log W_T)\right],$$

can be approximated by

$$\begin{aligned} U(\gamma) \quad = \quad & \frac{(W_0 \exp(rT))^{1-\gamma}}{1-\gamma} U_{\text{fix}}(\xi_{\text{fix}}) \exp\Big\{(1-\gamma)\xi_{\text{mv}} E[C_T + D_T + y(\xi_{\text{fix}}, \xi_{\text{mv}}) F_T] \\ & + \frac{1}{2}(1-\gamma)^2 \xi_{\text{mv}}^2 \text{var}[C_T + D_T + y(\xi_{\text{fix}}, \xi_{\text{mv}}) F_T] \\ & + (1-\gamma)^2 \xi_{\text{fix}} \xi_{\text{mv}} \text{cov}(A_T + B_T, C_T + D_T + y F_T)\Big\}, \end{aligned} \tag{40}$$

where $U_{\text{fix}}(\xi_{\text{fix}})$ is the value function associated with a given fixed strategy ξ_{fix},

$$y(\xi_{\text{fix}}, \xi_{\text{mv}}) = (\mu_0 + \mu_1 \bar{X} - r) - \frac{1}{2}\sigma_s^2 \xi_{\text{mv}} - \sigma_s^2 \xi_{\text{fix}},$$

and

$$C_T = \mu_1 \int_0^T \eta_t X_t dt, \qquad D_T = \sigma_s \int_0^T \eta_t dB_t, \qquad F_T = \int_0^T \eta_t dt,$$

$$A_T = \mu_1 \int_0^T X_t dt, \qquad B_T = \sigma_s \int_0^T dB_t.$$

Upon some further algebraic manipulation, we obtain the power-utility value function as

$$U(\gamma) = \frac{(W_0 \exp(rT))^{1-\gamma}}{1-\gamma} U_{\text{fix}}(\xi_{\text{fix}}) \exp\{(1-\gamma)\xi_{\text{mv}} \left[\phi_0 + \phi_1 \xi_{\text{mv}} + \phi_2 \xi_{\text{mv}}^2 + \phi_3 \xi_{\text{mv}}^3\right]\}, \tag{41}$$

where

$$
\begin{aligned}
\phi_0 &= EC_T + (\mu_0 + \mu_1 \bar{X} - r - \sigma_s^2 \xi_{\text{fix}}) EF_T \\
&\quad + (1-\gamma)\xi_{\text{fix}} \text{cov}(A_T + B_T, C_T + D_T + (\mu_0 + \mu_1 \bar{X} - r - \sigma_s^2 \xi_{\text{fix}}) F_T), \\
\phi_1 &= -\frac{1}{2}\sigma_s^2 EF_T + \frac{1}{2}(1-\gamma)\text{var}(C_T + D_T + (\mu_0 + \mu_1 \bar{X} - r - \sigma_s^2 \xi_{\text{fix}}) F_T) \\
&\quad + (1-\gamma)\xi_{\text{fix}} \text{cov}(A_T + B_T, -\frac{1}{2} F_T), \\
\phi_2 &= (1-\gamma)\text{cov}(C_T + D_T + (\mu_0 + \mu_1 \bar{X} - r - \sigma_s^2 \xi_{\text{fix}}) F_T, -\frac{1}{2}\sigma_s^2 F_T), \\
\phi_3 &= \frac{1}{2}(1-\gamma)\frac{\sigma_s^4}{4}\text{var}(F_T).
\end{aligned}
$$

Hence, for any given ξ_{fix}, we can solve the associated ξ^*_{mv}, which maximizes $U(\gamma)$ of (41), as

$$
\xi^*_{\text{mv}} = -\frac{\phi_2}{4\phi_3} - \left[\frac{q + \sqrt{q^2 + 4p^3/27}}{2}\right]^{\frac{1}{3}} + \frac{p}{3}\left[\frac{q + \sqrt{q^2 + 4p^3/27}}{2}\right]^{-\frac{1}{3}}, \tag{42}
$$

where

$$
p = \frac{\phi_1}{3\phi_3} - \frac{1}{3}\left(\frac{2\phi_2}{3\phi_3}\right)^{\frac{1}{3}}, \qquad q = \frac{\phi_0}{3\phi_3} - \frac{2}{27}\frac{\phi_0\phi_1\phi_2}{\phi_3^3} + \frac{2}{27}\left(\frac{2\phi_2}{3\phi_3}\right)^3. \tag{43}
$$

In particular, if $\xi_{\text{fix}} = \xi^*_{\text{fix1}}$ or ξ^*_{fix2} or 0, we obtain the corresponding ξ^*_{mv} from (42) that yields the approximate optimal GMA strategies. For easier reference, we will denote them as Fix1+MA, Fix2+MA, and Optimal MA, respectively. These three together with ξ^*_{fix1} and ξ^*_{fix2}, denoted as Fix1 and Fix2, consist of five strategies whose performances will be examined in detail in the next subsection.

4.3 Comparison in Calibrated Models

To get further insights into the practical importance of the differences in the five trading strategies, we in this subsection calibrate the model using monthly data on S&P500 and three popular predictable variables from December 1926 to December 2004, and then use simulation to obtain the certainty equivalent losses of the strategies as compared with the optimal dynamic one.

Table 1 reports the calibrated parameters (whose estimation details are provided in Appendix D). As expected, the stock volatility estimates are virtually the same as $\sigma_s = 0.1946$ across the three predictable models when the predictable variable is taken as the dividend yield, term-spread and payout ratio, respectively.[13] The same is true for the long-term mean of the stock return (not shown in the table). However, both the volatility of the predictable variable and its correlation with the stock return do vary across the models, making the comparison of the strategies more interesting.

Since utility values are difficult to interpret, we compute below the certainty-equivalent utility losses versus the optimal dynamic strategy. Normalizing the initial wealth at the level of one hundred dollars, $W_0 = 100$. Let $U^*_{\text{opt}}(W_0)$ be the expected utility based on the optimal dynamic strategy, and $U^*_f(W_0)$ be the expected utility based on any of the five suboptimal trading strategies: Fix1, Fix2, Fix1+MA, Fix2+MA, and Optimal MA. Since $U^*_{\text{opt}}(W_0) \geq U^*_f(W_0)$, there exists $\text{CE} \geq 0$ such that

$$U^*_{\text{opt}}(W_0 - \text{CE}) = U^*_f(W_0). \tag{44}$$

CE can be interpreted as the "perceived" certainty-equivalent loss at time zero to an investor who switches the optimal strategy to the suboptimal one. In other words, the investor would be willing to give up CE percent of his initial wealth to avoid investing in the suboptimal strategy. Similar measures are used by Kandel and Stambaugh (1996), Pástor and Stambaugh (2000), Fleming, Kirby, and Ostdiek (2001) and Tu and Zhou (2004), among others. In our computation of the certainty-equivalent losses below, we set the risk aversion $\gamma = 2$. In addition, we consider three ad hoc MA strategies, MA1, MA2 and MA3, whose stock allocations are 100%, Fix1 and Fix2, respectively when the MA indicates a 'buy' signal, and are nothing otherwise. The first of them is popular in practice. The other two are of interest because they provide information on how the optimal pure MA might be different from an MA strategy with seemingly well-chosen stock allocations.

Tables 2 and 3 report the CE losses in percentage points when $L = 50$ and 200 days, respectively.[14] The lag lengths are those used by Brock, Lakonishok, and LeBaron (1992), of which $L = 200$ is also the lag length of the popular moving average chart published by *Investor's Business Daily*, the major competitor of the *Wall Street Journal*. There are several interesting facts. First, the losses are substantial across all the strategies relative to the optimal dynamic one, and they vary substantially, too, across predictable models. When the predictable variable is taken as

[13] The dividend yield is not the dividend-price ratio. See, e.g., Goyal and Welch (2003) for a detailed description of these three predictable variables.

[14] The results when $L = 100$ are similar and omitted for brevity.

the dividend yield, the losses (ignoring the ad hoc MA strategies, which will be dropped later for reasons below) vary from 7.8951% to 50.3555%. The range widens, from 18.0587% to 59.3592%, when the payout ratio is taken as the predictable variable. However, it narrows down to a low of 1.5504% and to a high of 42.9099% when the term-spread is taken as the predictable variable. The large losses suggest strongly that, in an asset allocation problem, it is very important to know both the true dynamics of stock returns and the associated optimal dynamic strategy. This may help explain why Wall Street firms spend enormous amounts of money collecting data and doing research. Kandel and Stambaugh (1996) show that the economic loss can be significant when one ignores predictability completely when there is in fact a small degree of predictability in the data. In a continuous-time version of their model, this is apparent when we examine the losses of Fix1 versus the optimal dynamic strategy. However, the optimal dynamic strategy is difficult to identify, while the fixed rules are more practical and easy to apply. Even if the optimal dynamic rule is available, the predictable variable(s) may not be available at all time frequencies while the stock price can be observed virtually continuously during trading hours for implementing any MA-based strategies.

Second, Fix2 performs better than Fix1, which is not surprising since Fix1 is optimal only under the iid assumption. The superior performance varies across predictable variables and achieves the best level when the term-spread is taken as the predictable variable. The performance difference is of significant economic importance even when $T = 10$. This suggests that ignoring predictability entirely can lead to substantial losses in expected utilities even within the class of fixed strategies.

Third, the MA rule adds value to both Fix1 and Fix2, and Fix2+MA is the best suboptimal strategy. For Fix1, the MA improves its performance substantially by cutting the losses by at least 1–2% as long as $T > 10$. However, the MA provides only small improvement over Fix2. This does not suggest necessarily that the practical value of the MA rule is small. In practice, it is extremely difficult to know precisely what process the stock follows and what variables exactly that drive the market. On the other hand, long-term stock return and volatility could be estimated with little error due to the long historical data. This means that Fix1 is a feasible strategy while Fix2 may not be, at least to a sizable number of investors. By the same token, the dynamic optimal rule is difficult to identify in practice, as we have commented earlier. Currently, there is one-third of the stocks are held by index funds and those invested in them are likely to invest their money with fixed allocations that resemble Fix1 more than Fix2. In addition, popular portfolio optimization strategies (see, e.g., Litterman, 2003, and Meucci, 2005) are more like Fix1 than Fix2. To the extent that this is true, the MA rule can have value. Theoretically, as explored in the next section, uncertainty about the degree of predictability can make the MA rule add value to the optimal dynamic rule, too, when the prior is not informative enough. Of course, there might be countless other reasons for the usefulness of the MA rule since so many successful practitioners put their money behind it in reality.

Fourth, the lag length makes only a small difference in the results except for the pure MA rule (and the ad hoc ones) which by definition depend on L more heavily. Since the fixed rules are independent of L, their values are the same across Tables 2 and 3. For both Fix1+MA and Fix2+MA, their values change only from 8.1765% and 7.8951% to 8.1253% and 7.8961%, respectively, in the dividend yield model with $T = 10$. When $T = 40$, the values are larger and so are the differences. But the larger differences are still less than 0.5%. In contrast, for the Optimal MA, the largest difference is as high as about 5%, occurring at $T = 40$.

Fifth, the optimal pure MA rules are much worse than other rules (except the ad hoc MA ones). For example, when the dividend yield is taken as the predictable variable and $L = 50$, it has a loss about twice as large as the fixed strategies rules when $T = 10$. The qualitative results change little as T increases. When the term-spread is taken as the predictable variable, the difference can be four times as large. The least difference, still over 5%, occurs when the payout rate is taken as the predictable variable. The results suggest strongly that one should not use MA alone, but only use it in conjunction with the fixed strategies.

Sixth, the ad hoc MA rules, MA1, MA2 and MA3, perform worse than the optimal pure MA. Theoretically, this is expected because the later is optimal. However, what is of interest here is that the under-performance can be of significant economic importance. Since these ad hoc rules perform poorly and do not add much information in comparison with other rules once we keep the optimal MA, we will eliminate them henceforth.

Finally, let us examine the impact of using either arithmetic moving averages or the ex-dividend stock prices in the computation of the GMA strategies. To see the influence of the first, Table 4 reports the same valuation as Table 3 except for that it replaces the previous geometric moving averages with the arithmetic ones. The results are little changed. For example, when $T = 40$ and when the dividend yield is taken as the predictable variable, Fix1+MA has a value of 27.3783%, which is virtually identical to the earlier value of 27.3408%. The largest difference occurs for the pure MA rule, which is still less than 0.5%. To see the effects of the dividends, Table 5 computes the losses of Table 3 by using the the ex-dividend prices instead, with an assumed annual dividend yield of 3%. Although the differences are larger now, they are confined only to the MA rule. They make no difference whatsoever for the GMA strategies. Overall, we find that our earlier conclusions are robust to using either arithmetic averages or ex-dividend stock prices in the implementation of the GMA rules.

5 Comparison Under Parameter Uncertainty

In previous sections, we assume that an economic agent making an optimal financial decision knows the true parameters of the model, as is commonly done in a theoretical set-up. However, the true parameters are rarely if ever known to the decision maker. In reality, model parameters have to be estimated, and different parameter estimates could provide entirely different results. This gives rise to the estimation risk associated with any trading strategy. In this section, we analyze the performance of the various investment strategies under such parameter uncertainty.

One remarkable feature of the pure moving average rule is that it is parameter- and model-free, and hence it is not subject to estimation risk once given an ex-ante allocation to the stock. Hence, it will not be surprising that the optimal GMA rule discussed below is quite robust to parameter uncertainty and does not require any prior estimate of the predictable parameter. In contrast, the performances of the optimal dynamic rules rely more heavily on how reliable the estimates of the true parameters are, which depends not only on the sample size, but also on the prior.

In a continuous-time model, it is well known that one can separate the estimation from the optimization problem (see, e.g., Gennotte (1986)), and parameter uncertainty affects the optimal portfolio choice through dynamic learning. Barberis (2000) and Xia (2001), among others, show that this dynamic learning effect not only changes the myopic portfolio holding, but also adds a new component for dynamic hedging arising from the parameter uncertainty. For tractability, we follow Xia's (2001) approach to model uncertainty about predictability to examine the usefulness of the GMA rule. In this case, the stock price dynamics can be re-parameterized as

$$\frac{dS_t}{S_t} = (\mu_0 + \mu_1 \bar{X} + \beta \hat{X}_t)dt + \sigma_s dB_t, \tag{45}$$

$$dX_t = (\theta_0 + \theta_1 X_t)dt + \sigma_x dZ_t, \tag{46}$$

where β is an unknown parameter to be inferred from the data. Uncertainty associated with β obviously measures an investor's uncertainty about predictability. All other parameters are assumed known. In particular, the long-term mean stock return, $\mu_0 + \mu_1 \bar{X}$, is known, where $\bar{X} = -\frac{\theta_0}{\theta_1}$ is the long-term mean of X_t. Assume β follows a diffusion process

$$d\beta = \lambda(\bar{\beta} - \beta)dt + \sigma_\beta dZ_t^\beta, \tag{47}$$

where the parameters of this process, i.e., the long term mean $\bar{\beta}$ and reversion speed λ, are known to investors. But the investor does not observe the innovation process Z_t^β directly, and has to infer the realization of β through observations on S_t and X_t. To complete the model, assume $E(dB_t dZ_t^\beta) = \rho_{\beta s} dt$, $E(dZ_t dZ_t^\beta) = \rho_{\beta x} dt$, $E(dB_t dZ_t) = \rho dt$. Xia (2001) provides a detailed analysis of this set-up and the associated properties.

Let $\mathcal{I}_t$ be the investor's filtration. Adapted to $\mathcal{I}_t$, the least square estimate of β is Gaussian, with mean and variance:

$$b_t = E[\beta_t | \mathcal{I}_t], \qquad \nu_t = E[(\beta_t - b_t)^2 | \mathcal{I}_t]. \tag{48}$$

Starting from a Gaussian prior for β with mean b_0 and variance ν_0, the Bayesian updating rule for the conditional mean and variance, b_t and ν_t, is (see, Xia, 2001)

$$db_t = \lambda(\bar{b} - b_t)dt + v_1 d\hat{B}_t + v_2 d\hat{Z}_t, \tag{49}$$

$$\frac{d\nu_t}{dt} = -2\lambda\nu_t + \sigma_\beta^2 - (v_1^2 + v_2^2 + 2v_1 v_2 \rho), \tag{50}$$

where

$$\begin{aligned}
\bar{b} &= \bar{\beta}, \\
v_1 &= \frac{\nu_t(X_t - \bar{X}) + \sigma_s \sigma_\beta(\rho_{\beta s} - \rho_{\beta x}\rho)}{\sigma_s(1-\rho^2)}, \\
v_2 &= \frac{-\nu_t(X_t - \bar{X})\rho_{xs} + \sigma_s \sigma_\beta(\rho_{\beta x} - \rho_{\beta s}\rho)}{\sigma_s(1-\rho^2)}, \\
d\hat{B}_t &= dB_t + \frac{(X_t - \bar{X})(\beta_t - b_t)}{\sigma_s} dt, \\
d\hat{Z}_t &= dZ_t.
\end{aligned}$$

To further simplify the problem, we assume log-utility. In this case, the optimal dynamic stock allocation can be solved analytically,

$$\xi^*_{\text{opt}} = \frac{\mu_s + b_t(X_t - \bar{X}) - r}{\sigma_s^2}. \tag{51}$$

Hence, the optimal log-utility level is

$$U^*_{\text{opt}} = E\log W_T = \int_0^T E\left[r + \xi^*_{\text{opt}}(\mu_0 + \mu_1\bar{X} + \beta(X_t - \bar{X}) - r) - \frac{1}{2}\xi^{*2}_{\text{opt}}\sigma_s^2\right] dt + \log W_0. \tag{52}$$

This value function can be computed easily via simulation.

In particular, the optimal fixed rule in the parameter uncertainty case, under the log-utility, can be explicitly obtained as

$$\xi^*_{\text{fix}} = \frac{\mu_s - r + C_T}{\sigma_s^2}, \tag{53}$$

where

$$C_T = \frac{1}{T}\int_0^T E\left[\beta\hat{X}_t\right] dt = \frac{\rho_{\beta x}\sigma_\beta\sigma_x}{(\theta_1 - \lambda)^2}\left[\frac{e^{(\theta_1-\lambda)T} - 1}{T} - 1\right].$$

Intuitively, C_T captures the covariance between the predictability parameter β and state variable X_t.

In our simulations below, we study the performance of the following three strategies in our parameter uncertainty setting:

1. The optimal dynamic learning rule ξ^*_{opt} as given by (51);
2. The optimal fixed strategy ξ^*_{fix} as given by (53);
3. The GMA rule, a combination of ξ^*_{fix} and the MA, with coefficients:

$$\xi_{\text{fix}} = \xi^*_{\text{fix}} - \frac{\bar{\beta}b_1}{b_2(1-b_2)\sigma_s^2}, \qquad \xi_{\text{mv}} = \frac{\bar{\beta}b_1}{b_2(1-b_2)\sigma_s^2}, \tag{54}$$

where b_1 and b_2 are defined similarly in (18) and (19) where the unknown μ_1 is replaced by the long term mean $\bar{\beta}$.

As in Xia (2001), we assume $\rho_{\beta x}$ to be zero. Then, neither the fixed rule nor the GMA rule depends on the unknown parameter β, and ξ^*_{fix} reduces to the optimal fixed rule $\xi^*_{\text{fix}2}$. In addition, for the mean-reverting process on β, we assume β_t starts from its calibrated long-term mean, $\beta_0 = 2.0715$, and set the reverting speed $\lambda = 0.115$ and the volatility $\sigma_\beta = 1.226$.

The results are provided in Table 6 with the dividend yield as the predictable variable, $L = 200$ days and $T = 10$ years. The first two columns are values for the prior mean and standard error, the third to the fifth columns are the expected utilities associated with the above three strategies, and the last two columns are the certainty-equivalent losses (in percentage points) of the fixed and GMA strategies relative to the optimal learning one. Because $\rho_{\beta x} = 0$, neither the fixed rule nor the GMA rule depends on the unknown parameter β, and hence their performances are independent of priors on β. Of course, the performance of the optimal updating rule depends on the prior. When the prior mean $b_0 = 0$, both the fixed and the GMA rule under-perform the optimal learning rule substantially, with losses from 10.67% to 12.40% and 10.07% to 11.80%, respectively. Among the priors, $\sqrt{\nu_0} = 2$ is clearly the best one, and hence it is not surprising to see that the associated loss is the largest. Interestingly, while it is unclear ex-ante whether or not $\sqrt{\nu_0} = 1$ is better than $\sqrt{\nu_0} = 3$, the former turns out to provide a higher expected utility for the optimal learning. The reason is that the model seems to penalize large beta values more than small ones relative to the true β_0. This is why that the losses become smaller when $\sqrt{\nu_0}$ further increases from 3. When the prior mean $b_0 = 4$, the results are similar qualitatively. However, when the prior $b_0 = 6$, which is not too informative about the true β_0, the optimal learning rule can now perform worse than either the fixed strategy or the GMA when $\sqrt{\nu_0} = 1$. When the prior mean moves further away at $b_0 = 7$, the losses increase substantially to over 10%. The optimal learning also depends on the investment horizon. As the horizon shortens, the optimal learning becomes worse, as shown by Table 7 with $T = 5$ years. This is expected because less time makes learning less effective. Overall, to the extent that uncertainty about predictability is high and the prior is not very informative, the widely used fixed strategy appears viable as it can outperform the optimal learning one. On the other hand, the MA rule can always add value to this fixed rule. Therefore, the MA rule or technical analysis seems capable of capturing information on the market that is useful to investors.

6 Comparison Under Model Uncertainty

In this section, we consider further the case in which the true model is not completely known to investors. Previously, the smart investors could obtain their optimal trading strategies based on their assumed true model, but now the true model is unknown both to these smart investors and to the technical traders. We examine how well the GMA strategy performs in this seemingly very realistic case because no one in the real world knows the exact model of stock prices. We find that the GMA strategy is quite robust to model specifications and outperforms the optimal trading strategies substantially when they are derived from the wrong models.

Recall that, in previous sections, we have solved the optimal GMA strategy in terms of the true parameters of the model, but this is not absolutely necessary. Indeed, we show now that the optimal GMA strategy can be estimated with much less dependence on the model. In other words, the strategy is robust to a wide class of model specifications. To see this, assume now that we have a very general stock price process

$$\frac{dS_t}{S_t} = R_t dt + \sigma dB_t, \tag{55}$$

where R_t is the instantaneous expected stock return that can be stochastic. For simplicity, σ is assumed, as before, as the constant volatility parameter. Then the log wealth process of the GMA strategy is

$$\log W_T = \log W_0 + rT + \int_0^T (\xi_{\text{fix}} + \xi_{\text{mv}}\eta_t)(R_t - r)dt + \int_0^T (\xi_{\text{fix}} + \xi_{\text{mv}}\eta_t)\sigma dB_t - \frac{1}{2}\int_0^T (\xi_{\text{fix}} + \xi_{\text{mv}}\eta_t)^2\sigma^2 dt.$$

Hence, the expected utility becomes

$$U = E\log W_T = \log W_0 + rT + \left(\xi_{\text{fix}} b_0 + \xi\text{mv} b_1 - \frac{1}{2}\xi_{\text{fix}}^2\sigma^2 - \xi_{\text{fix}}\xi_{\text{mv}}\sigma^2 b_2 - \frac{1}{2}\xi_{\text{mv}}^2\sigma^2 b_2\right)T, \tag{56}$$

where

$$\begin{aligned} b_0 =& \frac{1}{T}\int_0^T E[R_t - r]dt, \\ b_1 =& \frac{1}{T}\int_0^T E[\eta_t(R_t - r)]dt, \\ b_2 =& \frac{1}{T}\int_0^T E\eta_t dt. \end{aligned} \tag{57}$$

Optimizing the expected utility, we obtain

$$\hat{\xi}_{\text{fix}}^* = \frac{b_0}{\sigma^2} - b_2\hat{\xi}_{\text{mv}}^*, \qquad \hat{\xi}_{\text{mv}}^* = \frac{1}{\sigma^2(1 - b_2)}\left(\frac{b_1}{b_2} - b_0\right). \tag{58}$$

The parameters defined in (57) can be written in terms of moments,

$$b_0 = E[R_t] - r, \qquad b_1 = E[\eta_t R_t] - rb_2, \qquad b_2 = E[\eta_t]. \tag{59}$$

Thus, assuming stationarity as before, we can estimate them by their sample analogues. For example, to see how b_1 can be estimated, we write

$$R_t\Delta t = \frac{\Delta S_t}{S_t} - \sigma\Delta B_t.$$

With the law of iterative expectation, we have

$$b_1 = E[\eta_t E_t(R_t - r)] = E[\eta_t(\frac{\Delta S_t}{S_t \Delta t} - r)],$$

which can be estimated by using the corresponding sample average of the right hand side.

Now we are ready to define the estimated optimal GMA strategy as follows (which differs from the optimal GMA that solves from a given specification of the true model). At any time t, we use the available sample moments up to that time to estimate the parameters given by (59). Substituting the estimates into (58), we obtain the estimated optimal GMA strategy $\hat{\xi}^*_{\text{fix}} + \hat{\xi}^*_{\text{mv}}\eta_t$. Since the estimates $\hat{\xi}^*_{\text{fix}}$ and $\hat{\xi}^*_{\text{mv}}$ vary over time according to the moment estimates at time t, and since they do not depend on future information, the strategy is a feasible rolling strategy. One should note that no knowledge of the true model is needed other than the general form of equation (55).

To assess the model uncertainty effect, we assume that the true model is one of the three calibrated models, but this is unknown to the investors. There are three cases to consider, each of which corresponds to one of the three models as the true one, respectively. In the first case when the model with the dividend yield as the predictable variable is assumed the true data-generating process, Panel A of Table 8 reports the utility losses by using the estimated GMA and the optimal trading strategies based on the wrong models, the second and third one, respectively.[15] As before, the losses here are measured relative to the true optimal strategy. When $T = 5$, the largest loss of the estimate GMA is 5.3326%, far smaller than 17.2875%, the largest of the wrong optimal strategies.[16] It is also smaller than 6.5926%, the smallest of the latter. As investment horizon increases, the loss increases. The same conclusion also holds when the assumed true mode is the one with term-spread and payout ratio as the predictable variable, respectively, as indicated by the results in Panels B and C of the table.

Another interesting question is how well the estimated GMA compares with the estimated fixed strategy, i.e, $\hat{\xi}^*_{\text{fix}} = \hat{b}_0/\hat{\sigma}^2$ with $\hat{b}_0$ and $\hat{\sigma}^2$ as the moment estimators. The utility losses associated with $\hat{\xi}^*_{\text{fix}}$ are reported in the fourth column of Table 8. They are always larger than those associated with the GMA strategy, and are substantial so in many cases. This says that the estimated optimal

[15] For simplicity in making the points here, we have assumed that there is only one wrong model each time. In the case when there are many candidate models, one can imagine that the true model is still not any of them, and that the optimal aggregation of all the candidate models gives rise to a single wrong model.

[16] Although not reported, the estimated GMA differs only slightly from the optimal GMA. For example, in the first case, when $T = 5$ and $L = 50$, their difference is less than 0.5%.

GMA outperforms the estimated fixed strategy not only when the true model is known, as it is the case in Section 4, but also when the true model is unknown, as it is the case here.

Overall, our results show that, while the estimated GMA strategy has lower utility than the optimal one, it outperforms all the optimal strategies when they are derived from wrong models. Given that the true model is unknown and difficult to identify by investors in the real world, the robustness of the GMA strategy, or of the technical analysis in general, makes it a valuable tool in practice.

7 Optimal Lags

In previous sections, we have studied the various GMA strategies with some popular fixed lags. In this section, we ask how the lag can be optimized. We study this problem under the log-utility with the aid of the analytical solutions of Section 3. However, the optimal lag itself does not admit an explicit solution, but can be solved in closed form approximately that provides qualitative insights on the driving factors. Unlike Sections 5 and 6, we assume as usual in this section that the investor knows all the true parameters of the model to simplify the analysis.

To study the determinants of the optimal lag, we restrict parameter values to those of practical interest by assuming

$$\sigma_s^2 >> \frac{\mu_1^2\sigma_x^2}{\theta_1^2} - \frac{2\mu_1\sigma_x\sigma_s\rho}{\theta_1}. \tag{60}$$

This is because σ_x is much smaller relative to σ_s, and because the correlation ρ is close to zero. This relation holds for all the three calibrated models. Using the unit-free variable $x = \sqrt{|\theta_1|L}$, we can approximate equations (20), (21) and (22) by

$$\begin{aligned} C_{12}^Z &\approx C_1\Big(1 - \frac{1-e^{-x^2}}{x^2}\Big), \\ C_{22}^Z &\approx \frac{\sigma_s^2}{3}L = C_2x^2, \\ m_2^Z &= \frac{\mu_s - \sigma_s^2/2}{2}L = C_3x^2, \end{aligned}$$

where

$$C_1 = \frac{\mu_1\sigma_x^2}{2\theta_1^2} - \frac{\sigma_x\sigma_s\rho}{\theta_1}, \quad C_2 = \frac{\sigma_s^2}{3|\theta_1|}, \quad C_3 = \frac{\mu_s - \sigma_s^2/2}{2|\theta_1|}.$$

Therefore, equations (18) and (19) can be approximated as:

$$b_1 \approx C_4 \cdot \frac{1}{x}(1 - \frac{1 - e^{-x^2}}{x^2}) \cdot f(Ax) = C_4 h(x) f(Ax), \tag{61}$$

$$b_2 \approx N(Ax), \tag{62}$$

where

$$A = \frac{C_3}{\sqrt{C_2}} = \frac{\sqrt{3}}{2} \cdot \frac{\mu_s - \frac{\sigma_s^2}{2}}{\sigma_s \sqrt{|\theta_1|}}, \tag{63}$$

$$C_4 = \frac{C_1}{\sqrt{C_2}}, \tag{64}$$

$$h(x) = \frac{1}{x}(1 - \frac{1 - e^{-x^2}}{x^2}), \tag{65}$$

and $f(\cdot)$ is the standard normal density function. Then, we have

Proposition 4 In the class of strategies GMA(S_t, G_t, γ), if the investment horizon T is long enough, then the optimal lag L_{opt} under the log-utility is approximately given by

$$L_{\text{opt}} \approx \left[|\theta_1| \left(\frac{1 + A_i^2}{2} + \sqrt{(\frac{1 + A_i^2}{2})^2 - (\frac{5}{12} + \frac{A_i^2}{3})^2} \right) \right]^{-1}, \tag{66}$$

where $A_i = \frac{A}{\sqrt{2}}$ and A for the optimal GMA and Fix1+MA strategies, respectively.

Proposition 4 says that optimal lag is mainly a function of the unconditional mean return μ_s, stock volatility σ_s, and state variable mean reversion speed $|\theta_1|$ given that T is large. Since μ_s and σ_s are stable across different models, L_{opt} is mainly driven by the differences in θ_1. Figure 1 plots the CE losses of the above two strategies relative to the optimal dynamic one at various lag lengths when $T = 40$. Because of differences in θ_1, as predicted by Proposition 4, the optimal lag in the term-spread model is the smallest, and becomes the largest in the payout ratio model. There are in addition two interesting facts. First, the CE losses are much greater than those reported in Tables 2 and 3. This is expected because here $\gamma = 1$ while γ has a value of 2 in the earlier tables. The smaller the γ, the more the risk taking, and so the greater the impact of the various stock allocation strategies on the expected utility. Second, the performance across different lags do not vary much for Fix1+MA, implying that our earlier utility comparisons are insensitive to the use of the optimal lags. However, the optimal GMA rule is substantially more influenced by the use of the optimal lag than Fix1+MA. But this will not affect our earlier results because numerical studies on this rule are not provided due to the unavailability of its solution in the power-utility case.

Now consider the optimal lag for the pure MA strategy. Intuitively, given a lag length, the initial value of the moving average matters little when T is large. However, given T, the initial value matters significantly in choosing L. This is because L can be chosen as T. Indeed, since the pure MA under-performs ξ^*_{fix1} under the practical parameter values, it will be optimal to let $L = T$. In this case, the pure MA will be identical to Fix1 since the initial value is chosen as ξ^*_{fix1}. An alternative initial value for the pure MA is zero. In this case, it can be shown (see the Appendix) that

$$L_{\text{opt}} \approx \frac{2\log(|\theta_1|T)}{A|\theta_1|} \tag{67}$$

when $|\theta_1|T$ is large. This makes intuitive sense. The larger the speed of mean reversion, the shorter the lag length to capture the change of trends.

8 Conclusion

Although technical analysis is popular in investment practice, there are few studies on it. The empirical evidence is mixed, and we lack a theoretical understanding of why it might be useful. In this paper, we provide the theoretical justification for an investor to use the moving average (MA) rule, one of the widely used technical rules, in a standard asset allocation problem. The theoretical framework seems to offer quite a few insights about technical analysis. First, it answers the question of how much of his money a technical trader should allocate into the stock market if he receives a technical buy signal, while previous researchers determine it in ad hoc ways. Second, it shows how an investor might add value into his investment by using technical analysis, especially the MA, if he follows a fixed allocation rule that invests a fixed portion of wealth into the stock market (as dictated by the random walk theory of stock prices or by the popular mean-variance approach). Third, when model parameters are unknown and have to be estimated from data, the asset allocation framework illustrates that the combination of the fixed rule with the MA can even outperform the prior-dependent optimal learning rule when the prior is reasonable and yet not too informative. Finally, when the true model is unknown, as is the case in practice, we find that the optimal generalized MA is robust to model specification, and its estimate outperforms the optimal dynamic strategies substantially when they are derived from the wrong models.

For tractability, our exploratory study assumes a simple predictable process for a single risky asset and examines the simplest moving average rule. Studies that allow for both more general processes (such as those with jumps, factor structures, and multiple assets) and more elaborate rules are clearly called for. Broadly speaking, asset pricing anomalies such as the momentum effect can also be regarded as one of the profitable technical strategies that depend on historical information only. Questions that remain open are what underlying asset processes permit such anomalies and what the associated optimal investment strategies are. Moreover, related issues of interest are how past prices and trading volume reveal the strategies of the major market players, with their incomplete and complementary information, and how their interactions determine future asset prices. All of these are important and challenging topics for future research.

References

Allen, F., Karjalainen, R., 1999. Using genetic algorithms to find technical trading rules. Journal of Financial Economics 51, 245–271.

Ang, A., Bekaert, G., 2007. Stock return predictability: is it there? Review of Financial Studies 20, 651–707.

Back, K., 2006. Introduction to Asset Pricing and Portfolio Choice Theory. Manuscript. Texas A&M University, Texas.

Bhattacharya, U., Galpin, N., 2006. Is stock picking declining around the world? Unpublished working paper. Indiana University.

Barberis, N., 2000. Investing for the long run when returns are predictable. Journal of Finance 55, 225–264.

Brock, W., Lakonishok, J., LeBaron, B., 1992. Simple technical trading rules and the stochastic properties of stock returns. Journal of Finance 47, 1731–1764.

Brown, D., Jennings, R., 1989. On technical analysis. Review of Financial Studies 2, 527–551.

Campbell, J., 1987. Stock returns and the term structure. Journal of Financial Economics 18, 373–399.

Campbell, J., Lo, A., MacKinlay, C., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton, N.J.

Cowles, A., 1933. Can stock market forecasters forecast? Econometrica 1, 309–324.

Covel, M., 2005. Trend Following: How Great Traders Make Millions in Up or Down Markets. Prentice-Hall, New York.

Curtis, H., 1944. A derivation of Cardano's formula, American Mathematical Monthly 51, 35–35.

Fama, E., Blume, M., 1966. Filter rules and stock market trading. Journal of Business 39, 226–241.

Fama, E., Schwert, G., 1977. Asset returns and inflation. Journal of Financial Economics 5, 115–146.

Ferson, W., Harvey, C., 1991. The variation of economic risk premiums. Journal of Political Economy 99, 385–415.

Fleming, J., Kirby, C., Ostdiek, B., 2001. The economic value of volatility timing. Journal of Finance 56, 329–352.

Gennotte, G., 1986. Optimal portfolio choice under incomplete information. Journal of Finance 41, 733–746.

Goyal, A., Welch, I., 2003. Predicting the equity premium with dividend ratios. Management Science 49, 639–654.

Huang, L., Liu, H., 2007. Rational inattention and portfolio selection. Journal of Finance 62, 1999–2040.

Ingersoll, J., 1987. Theory of Financial Decision Making. Rowman & Littlefield, New York.

Kandel, S., Stambaugh, R., 1996. On the predictability of stock returns: An asset-allocation perspective. Journal of Finance 51, 385–424.

Karatzas, I., Shreve, S., 1991. Brownian Motion and Stochastic Calculus, Second Edition. Spriner-Verlag, New York.

Kim, T., Omberg, E., 1996. Dynamic nonmyopic portfolio behavior. Review of Financial Studies 9, 141–161.

Kim, Y., Grundy, B., 2002. Stock market volatility in a heterogeneous information economy. Journal of Financial and Quantitative Analysis 37, 1–27.

Litterman, B., 2003. Modern Investment Management. Wiley, New York.

Lo, A., MacKinlay, C., 1990. Data-snooping biases in tests of financial asset pricing models. Review of Financial Studies 3, 431–468.

Lo, A., Mamaysky, H., Wang, J., 2000. Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. Journal of Finance 55, 1705–1765.

Markowitz, H., 1952. Mean-variance analysis in portfolio choice and capital markets. Journal of Finance 7, 77–91.

Merton, R., 1992. Continuous-time Finance. Blackwell, Cambridge, MA.

Meucci, A., 2005. Risk and Asset Allocation Springer-Verlag, New York.

Pástor, Ľ., Stambaugh, R., 2000. Comparing asset pricing models: An investment perspective. Journal of Financial Economics 56, 335–381.

Rubinstein, M., 2002. Markowitz's "portfolio selection": A fifty-year retrospective. Journal of Finance 57, 1041–1045.

Schwager, J., 1993. Market Wizards: Interviews with Top Traders. Collins, New York.

Schwager, J., 1995. The New Market Wizards: Conversations with America's Top Traders, Wiley, New York.

Stambaugh, R., 1999. Predictive regressions. Journal of Financial Economics 54, 375–421.

Sullivan, R., Timmermann, A., White, H., 1999. Data-snooping, technical trading rule performance, and the bootstrap. Journal of Finance 53, 1647–1691.

Tobin, J., 1958. Liquidity preference as behavior towards risk. Review of Economic Studies 25, 65–86.

Tu, J., Zhou, G., 2004. Data-generating process uncertainty: what difference does it make in portfolio decisions? Journal of Financial Economics 72, 385–421.

Xia, Y., 2001. Learning about predictability: the effects of parameter uncertainty on dynamic asset allocation. Journal of Finance 55, 205–246.

Appendix

1 Proof of Equations (10), (18) and (19)

Let $y_t = \log S_t$. Then the model for the predictable variable and stock price process is:

$$\begin{cases} dX_t = (\theta_0 + \theta_1 X_t)dt + \sigma_x dZ_t, \\ dy_t = (\mu_0 + \mu_1 X_t - \sigma_s^2/2)dt + \sigma_s dB_t, \end{cases} \tag{A1}$$

where (Z_t, B_t) is a two-dimensional Brownian Motion with correlation coefficient ρ.

To rule out any explosive behavior, we assume $\theta_1 < 0$ throughout, which is consistent with empirical applications. Furthermore, we assume that X_t is a stationary process for $t \geq 0$. Integrating the stochastic differential equation (A1) for X_t, we have

$$X_t = X_0 e^{\theta_1 t} - \frac{\theta_0}{\theta_1}(1 - e^{\theta_1 t}) + \sigma_x \int_0^t e^{\theta_1 (t-s)} dZ_s. \tag{A2}$$

It follows that X_t is normally distributed with mean and covariance

$$EX_t = EX_0 e^{\theta_1 t} - \frac{\theta_0}{\theta_1}(1 - e^{\theta_1 t}), \tag{A3}$$

$$\mathrm{cov}(X_t, X_s) = [V(0) - \frac{\sigma_x^2}{2\theta_1}(e^{-2\theta_1 t \wedge s)} - 1)]e^{\theta_1 (t+s)}, \tag{A4}$$

respectively, where EX_0 and $V(0)$ are the mean and variance of X_0. Then, the steady state mean and variance of X_t can be obtained by taking $t \to +\infty$ in (A3) and (A4), i.e.,

$$\bar{X} = -\frac{\theta_0}{\theta_1}, \quad \bar{V}_x = -\frac{\sigma_x^2}{2\theta_1}.$$

The necessary and sufficient condition for X_t to be stationary for $t \geq 0$ is that X_0 start from the steady state, i.e., X_0 is normally distributed with mean $\bar{X}$ and variance $V(0) = \bar{V}_x$. Under the stationarity condition, the first two moments (A3) and (A4) that characterize the distribution of X_t can thus be simplified as:

$$EX_t = \bar{X} = -\frac{\theta_0}{\theta_1}, \qquad \mathrm{cov}(X_t, X_s) = -\frac{\sigma_x^2}{2\theta_1} e^{\theta_1 |t-s|}. \tag{A5}$$

With initial conditions $X|_{t=0} = X_0$, $y|_{t=0} = y_0$, we integrate stochastic differential equations (A1) to obtain

$$\begin{cases} X_t = X_0 e^{\theta_1 t} - \frac{\theta_0}{\theta_1}(1 - e^{\theta_1 t}) + \sigma_x \int_0^t e^{\theta_1 (t-s)} dZ_s, \\ y_t = y_0 + \int_0^t (\mu_0 + \mu_1 X_s - \sigma_s^2/2)ds + \sigma_s B_t. \end{cases} \tag{A6}$$

Let $M_t = \log G_t$, where G_t is the geometric moving average at time t, then

$$M_t = \frac{1}{L}\int_{t-L}^{t} y_s ds.$$

To derive (10), we note, under constant holding ξ_{fix2}, the wealth process is

$$\log W_T = \log W_0 + rT + \xi_{\text{fix2}}(\mu_0 - r - \xi_{\text{fix2}}\sigma_s^2/2)T + \xi_{\text{fix2}}\mu_1 \int_0^T X_t dt + \xi_{\text{fix2}}\sigma_s B_T. \tag{A7}$$

Then, optimizing over ξ_{fix2} the power-utility

$$\begin{aligned}\frac{1}{1-\gamma}E\left[\exp((1-\gamma)\log W_T)\right] &= \frac{1}{1-\gamma}\exp\left[(1-\gamma)(\log W_0 + rT + \xi_{\text{fix2}}(\mu_0 - r - \xi_{\text{fix2}}\sigma_s^2/2)T)\right] \\ &\quad \cdot E\exp\left[(\xi_{\text{fix2}}\mu_1\int_0^T X_t dt + \xi_{\text{fix2}}\sigma_s B_T)(1-\gamma)\right],\end{aligned} \tag{A8}$$

we obtain the solution

$$\xi^*_{\text{fix2}} = \frac{(\mu_0 - r) + \mu_1 E[\frac{1}{T}\int_0^T X_t dt]}{\gamma\sigma_s^2 - (1-\gamma)(\mu_1^2 A + 2\mu_1\sigma_s B)}, \tag{A9}$$

where

$$A = \frac{1}{T}\text{var}[\int_0^T X_t dt], \qquad B = \frac{1}{T}\text{cov}[\int_0^T X_t dt, B_T].$$

With (A6) and (A5), A and B can be simplified as

$$\begin{aligned} A &= \int_0^T dt \int_0^T ds < X_t X_s >= -\frac{\sigma_x^2}{2\theta_1}\int_0^T dt\int_0^T ds e^{\theta_1|t-s|} \\ &= \frac{\sigma_x^2}{\theta_1^2}\left(T + \frac{1-e^{\theta_1 T}}{\theta_1}\right),\end{aligned} \tag{A10}$$

and

$$B = \int_0^T < X_t, B_T > dt = \frac{\rho\sigma_x}{\theta_1}\left(\frac{e^{\theta_1 T}-1}{\theta_1} - T\right), \tag{A11}$$

where $< \cdot, \cdot >$ denotes the covariance operator conditional on information at time 0 throughout the Appendix for brevity, and we have made use of the following fact that for $t \le T$

$$\begin{aligned} < X_t, B_T > &= \sigma_x \int_0^t e^{\theta_1(t-s)} < dZ_s, B_T > \\ &= \sigma_x \int_0^t \rho e^{\theta_1(t-s)} ds = \frac{\rho\sigma_x}{\theta_1}(e^{\theta_1 t} - 1).\end{aligned}$$

Now, to derive (18) and (19), taking expectation on (A6) and making use of (A5), we obtain

$$\begin{aligned} Ey_t &= y_0 + (\mu_0 + \mu_1\bar{X} - \sigma_s^2/2)t, \\ EM_t &= y_0 + (\mu_0 + \mu_1\bar{X} - \sigma_s^2/2)(t - \frac{L}{2})\end{aligned}$$

when $t > L$. These results allow us to compute the following second moments for $t > L$:

$$
\begin{aligned}
< X_t, X_{t-L} > &= -\frac{\sigma_x^2}{2\theta_1} e^{\theta_1 L}, \\
< y_t, X_{t-L} > &= \int_0^t \mu_1 < X_s, X_{t-L} > ds + \sigma_x \sigma_s \int_0^{t-L} e^{\theta_1 (t-L-s)} < dW_s, B_t > \\
&= \int_0^{t-L} \mu_1 < X_s, X_{t-L} > ds + \int_{t-L}^t \mu_1 < X_s, X_{t-L} > ds + \sigma_x \sigma_s \rho \int_0^{t-L} e^{\theta_1 (t-L-s)} ds \\
&= \frac{\mu_1 \sigma_x^2}{2\theta_1^2} (2 - e^{\theta_1 (t-L)} - e^{\theta_1 L}) - \frac{\sigma_x \sigma_s \rho}{\theta_1} (1 - e^{\theta_1 (t-L)}), \qquad \text{(A12)} \\
< X_t, y_{t-L} > &= \int_0^{t-L} \mu_1 < X_s, X_t > ds + \sigma_x \sigma_s \int_0^t e^{\theta_1 (t-s)} < dW_s, B_{t-L} > \\
&= (\frac{\mu_1 \sigma_x^2}{2\theta_1^2} - \frac{\sigma_x \sigma_s \rho}{\theta_1}) (e^{\theta_1 L} - e^{\theta_1 t}), \qquad \text{(A13)} \\
< y_t, y_t > &= \sigma_s^2 t + \int_0^t \int_0^t \mu_1^2 < X_s, X_u > ds du + 2\sigma_s \int_0^t \mu_1 < X_s, B_t > ds \\
&= (\sigma_s^2 + \frac{(\mu_1 \sigma_x)^2}{\theta_1^2} - \frac{2\mu_1 \sigma_x \sigma_s \rho}{\theta_1}) t + (\frac{(\mu_1 \sigma_x)^2}{\theta_1^3} - \frac{2\mu_1 \sigma_s \sigma_x \rho}{\theta_1^2}) (1 - e^{\theta_1 t}),
\end{aligned}
$$

where we have used the fact $< X_s, B_t >= \sigma_x \int_0^s e^{\theta_1 (s-u)} \rho du$, for $s \leq t$, an equality

$$
\int_0^t \int_0^t < X_s, X_u > ds du = \frac{\sigma_x^2}{\theta_1^2} t + \frac{\sigma_x^2}{\theta_1^3} (1 - e^{\theta_1 t}),
$$

and another equality

$$
\begin{aligned}
< y_t, y_{t-L} > &= < y_{t-L}, y_{t-L} > + \int_{t-L}^t \mu_1 < X_s, y_{t-L} > ds \qquad \text{(A14)} \\
&= (\sigma_s^2 + \frac{(\mu_1 \sigma_x)^2}{\theta_1^2} - \frac{2\mu_1 \sigma_s \rho \sigma_x}{\theta_1}) (t - L) + (\frac{(\mu_1 \sigma_x)^2}{2\theta_1^3} - \frac{\mu_1 \sigma_s \rho \sigma_x}{\theta_1^2}) (1 - e^{\theta_1 (t-L)} + e^{\theta_1 L} - e^{\theta_1 t}).
\end{aligned}
$$

Next, we compute the following second moments involving M_t using (A12) and (A14):

$$
\begin{aligned}
< X_t, M_t > &= \frac{1}{L} \int_{t-L}^t < y_s, X_t > ds \\
&= \frac{1}{L} (-\frac{\mu_1 \sigma_x^2}{2\theta_1^3} + \frac{\sigma_x \sigma_s \rho}{\theta_1^2}) (1 - e^{\theta_1 L}) - (\frac{\mu_1 \sigma_x^2}{2\theta_1^2} - \frac{\sigma_x \sigma_s \rho}{\theta_1}) e^{\theta_1 t}, \\
< y_t, M_t > &= \frac{1}{L} \int_{t-L}^t < y_t, y_s > ds \\
&= (\sigma_s^2 + \frac{(\mu_1 \sigma_x)^2}{\theta_1^2} - \frac{2\mu_1 \sigma_s \sigma_x \rho}{\theta_1}) (T - \frac{L}{2}) + (\frac{(\mu_1 \sigma_x)^2}{2\theta_1^3} - \frac{\mu_1 \sigma_x \rho \sigma_s}{\theta_1^2}) (1 - e^{\theta_1 T}) \\
&\quad - (\frac{(\mu_1 \sigma_x)^2}{2\theta_1^3} - \frac{\mu_1 \sigma_x \rho \sigma_s}{\theta_1^2}) \frac{1}{\theta_1 L} (1 - e^{\theta_1 L} - e^{\theta_1 (T-L)} + e^{\theta_1 T}).
\end{aligned}
$$

Finally, in order to compute $< M_t, M_t >$, we note first

$$
\begin{aligned}
M_t &= \frac{1}{L}\int_{t-L}^{t} y_s ds = \frac{1}{L}\int_0^L [y_{t-L} + (y_{t-L+s} - y_{t-L})]ds \\
&= y_{t-L} + \frac{1}{L}\int_0^L \hat{y}_{t-L+s} ds,
\end{aligned}
$$

where $\hat{y}_{t-L+s} = y_{t-L+s} - y_{t-L}$. Then, we can write $< M_t M_t >$ as

$$
\begin{aligned}
< M_t, M_t > &= < (y_{t-L} + \frac{1}{L}\int_0^L \hat{y}_{t-L+s} ds), (y_{t-L} + \frac{1}{L}\int_0^L \hat{y}_{t-L+s} ds) > \\
&= < \hat{M}_L, \hat{M}_L > + \frac{2}{L}\int_0^L < y_{t-L}, y_{t-L+s} > ds - < y_{t-L}, y_{t-L} >,
\end{aligned}
$$

where $\hat{M}_t = \frac{1}{t}\int_0^t y_s ds$. Using (A14), we obtain

$$
\begin{aligned}
< \hat{M}_t, \hat{M}_t > &= \frac{1}{t^2}\int_0^t \int_0^t < y_s, y_u > dsdu \\
&= \frac{t}{3}(\sigma_s^2 + \frac{(\mu_1\sigma_x)^2}{\theta_1^2} - \frac{2\mu_1\sigma_x\sigma_s\rho}{\theta_1}) \\
&\quad + (\frac{(\mu_1\sigma_x)^2}{2\theta_1^3} - \frac{\mu_1\sigma_x\rho\sigma_s}{\theta_1^2}) \left[1 - \frac{2e^{\theta_1 t}}{\theta_1 t} - \frac{2}{(\theta_1 t)^2}(1 - e^{\theta_1 t})\right].
\end{aligned}
$$

For the term $\int_0^L < y_{t-L}, y_{t-L+s} > ds$, equation (A14) can be used for its computation. Hence, we get the last term for determining the covariance matrix of the trio (X_t, y_t, M_t) as

$$
\begin{aligned}
< M_t, M_t > = &(\sigma_s^2 + \frac{(\mu_1\sigma_x)^2}{\theta_1^2} - \frac{2\mu_1\sigma_x\sigma_s\rho}{\theta_1})(t - \frac{2L}{3}) \\
&+ \left[\frac{(\mu_1\sigma_x)^2}{2\theta_1^3} - \frac{\mu_1\sigma_x\rho\sigma_s}{\theta_1^2}\right]\left[1 - \frac{1}{(\theta_1 L)^2}(1 - e^{\theta_1 L} + \theta_1 L e^{\theta_1 L}) - \frac{2}{\theta_1 L}(1 - e^{\theta_1 L})(1 - e^{\theta_1 (t-L)})\right].
\end{aligned}
$$

Summarizing above, we have

Lemma 1 For $t > L$, the trio (X_t, y_t, M_t) are jointly normally distributed with mean $n = (n_1, n_2, n_3)$ given by

$$
\begin{aligned}
n_1 &= -\frac{\theta_0}{\theta_1}, \\
n_2 &= y_0 + (\mu_0 - \frac{\mu_1\theta_0}{\theta_1} - \sigma_s^2/2)t, \\
n_3 &= y_0 + (\mu_0 - \frac{\mu_1\theta_0}{\theta_1} - \sigma_s^2/2)(t - \frac{L}{2}),
\end{aligned}
$$

and covariance matrix $D = (D_{ij})$ given by

$$\begin{aligned}
D_{11} &= -\frac{\sigma_x^2}{2\theta_1}, \\
D_{22} &= (\sigma_s^2 + \frac{(\mu_1\sigma_x)^2}{\theta_1^2} - \frac{2\mu_1\sigma_x\sigma_s\rho}{\theta_1})t + (\frac{\sigma_x^2}{\theta_1^3} - \frac{2\mu_1\sigma_x\sigma_s\rho}{\theta_1^2})(1-e^{\theta_1 t}), \\
D_{33} &= (\sigma_s^2 + \frac{(\mu_1\sigma_x)^2}{\theta_1^2} - \frac{2\mu_1\sigma_x\sigma_s\rho}{\theta_1})(t - \frac{2L}{3}) \\
&\quad + (\frac{(\mu_1\sigma_x)^2}{2\theta_1^3} - \frac{\mu_1\sigma_s\rho\sigma_x}{\theta_1^2})\left[1 - \frac{2}{(\theta_1 L)^2}(1 - e^{\theta_1 L} + \theta_1 L e^{\theta_1 L}) - \frac{2}{\theta_1 L}(1-e^{\theta_1 L})(1-e^{\theta_1(t-L)})\right], \\
D_{12} &= (\frac{\mu_1\sigma_x^2}{2\theta_1^2} - \frac{\sigma_x\sigma_s\rho}{\theta_1})(1-e^{\theta_1 t}), \\
D_{13} &= \frac{1}{L}(-\frac{\mu_1\sigma_x^2}{2\theta_1^3} + \frac{\sigma_x\sigma_s\rho}{\theta_1^2})(1-e^{\theta_1 L}) - (\frac{\mu_1\sigma_x^2}{2\theta_1^2} - \frac{\sigma_x\sigma_s\rho}{\theta_1})e^{\theta_1 t}, \\
D_{23} &= (\sigma_s^2 + \frac{(\mu_1\sigma_x)^2}{\theta_1^2} - \frac{2\mu_1\sigma_s\sigma_x\rho}{\theta_1})(t - \frac{L}{2}) + (\frac{(\mu_1\sigma_x)^2}{2\theta_1^3} - \frac{\mu_1\sigma_s\rho\sigma_x}{\theta_1^2})(1-e^{\theta_1 t}) \\
&\quad - (\frac{(\mu_1\sigma_x)^2}{2\theta_1^3} - \frac{\mu_1\sigma_s\rho\sigma_x}{\theta_1^2})\frac{1}{\theta_1 L}(1 - e^{\theta_1 L} - e^{\theta_1(t-L)} + e^{\theta_1 t}).
\end{aligned}$$

With Lemma 1, the proof of (18) and (19) follows from

Lemma 2 Let $\hat{X}_t = X_t - \bar{X}$ and $Z_t = y_t - M_t$. Then $(\hat{X}_t, Z_t)$ is normally distributed with mean $m^Z = (n_1, n_2 - n_3)$, and covariance $C^Z = (C_{ij}^Z)$ given by

$$C_{11}^Z = D_{11}, \quad C_{22}^Z = D_{22} + D_{33} - 2D_{23}, \quad C_{12}^Z = D_{12} - D_{13}.$$

Moreover,

$$\begin{aligned}
E[1_{Z_t \geq 0}] &= N(\frac{m_2^Z}{\sqrt{C_{22}^Z}}), \\
E[X_t 1_{Z_t \geq 0}] &= m_1^Z N(\frac{m_2^Z}{\sqrt{C_{22}^Z}}) + \frac{C_{12}^Z}{\sqrt{C_{22}^Z}} N'(-\frac{m_2^Z}{\sqrt{C_{22}^Z}}).
\end{aligned} \tag{A15}$$

Proof: It is sufficient to prove only equation (A15), which is generally true for any jointly normal random variable (x, z), with mean (m_x, m_z), standard deviation (σ_x, σ_z), and correlation ρ, i.e.,

$$E[x 1_{z \leq 0}] = m_x N(\frac{m_z}{\sigma_z}) + \rho\sigma_x N'(-\frac{m_z}{\sigma_z}). \tag{A16}$$

Indeed, after standardization,

$$\hat{x} = \frac{x - m_x}{\sigma_x}, \quad \hat{z} = \frac{z - m_z}{\sigma_z},$$

we can write

$$\hat{x} = \rho\hat{z} + \sqrt{1-\rho^2}\hat{e},$$

where $\hat{e}$ is the standard normal variable that is independent of $\hat{z}$. Generally, for $m_z \geq 0$, which is satisfied by our application, where $E[Z_t] = E[y_t] - E[M_t] > 0$. Therefore, we have

$$\begin{aligned} E[x1_{z\leq 0}] &= E[(\sigma_x\hat{x} + m_x)1_{\hat{z}\leq -\frac{m_z}{\sigma_z}}] \\ &= m_x E1_{\hat{z}\leq -\frac{m_z}{\sigma_z}} + \rho\sigma_x E[\hat{z}1_{\hat{z}\leq -\frac{m_z}{\sigma_z}}] \\ &= m_x N(-\frac{m_z}{\sigma_z}) - \rho\sigma_x N'(-\frac{m_z}{\sigma_z}). \end{aligned}$$

Therefore,

$$E[x1_{z\geq 0}] = E[x] - E[x1_{z\leq 0}] = m_x N(\frac{m_z}{\sigma_z}) + \rho\sigma_x N'(-\frac{m_z}{\sigma_z})$$

which proves (A16).

2 Proof of Propositions 1, 2 and 3

Notice first that all three GMA strategies involve MA which is only well defined for $t > L$. When $t \leq L$, we define them here as the optimal fixed strategy ξ^*_{fix2} which is the same as ξ^*_{fix1} under the log-utility. Thus, the complete GMA rule is

$$\text{GMA}(S_t, G_t, \gamma = 1) = \begin{cases} \xi_{\text{fix}} + \xi_{\text{mv}} \cdot \eta(S_t, G_t), & \text{for} \quad t > L; \\ \xi^*_{\text{fix1}}, & \text{for} \quad t \leq L. \end{cases} \tag{A17}$$

This makes comparison across the strategies fair since they all start from ξ_{fix1}. For example, if the pure MA had started from zero, it would surely under-perform the other two over $[0, L]$ assuming a positive risk premium. Analytically, the same starting point makes the expressions simpler. But this implies that we assign an initial value of ξ^*_{fix1} for the pure MA strategy, while the initial values of the MA components in the first two strategies are different (to make the combined GMA start from the same point). Clearly, for a fixed L, the initial value has little impact if any when T is large. This is also consistent with the numerical results in Sections 3 and 4. However, when study optimal lags, the initial value does matter because the optimal lag of pure MA strategy can be close to T (see Section 7).

With any of the MA strategies, the key is to maximize the expected log-utility, which follows from Appendix A.1 and (23), as a function of $\xi_{\rm fix}$ and $\xi_{\rm mv}$,

$$\begin{aligned} U_{\rm GMA}(\xi_{\rm fix},\xi_{\rm mv}) &= \log W_0 + rT + \frac{(\mu_0+\mu_1\bar{X}-r)^2}{2\sigma_s^2}L \\ &\quad +\xi_{\rm fix}[\mu_0+\mu_1\bar{X}-r-\frac{\sigma_s^2}{2}\xi_{\rm fix}](T-L)+\xi_{\rm mv}\mu_1 b_1(T-L) \\ &\quad +[\xi_{\rm mv}(\mu_0+\mu_1\bar{X}-r)-\frac{\sigma_s^2}{2}\xi_{\rm mv}^2-\sigma_s^2\xi_{\rm fix}\xi_{\rm mv}]b_2(T-L). \end{aligned} \tag{A18}$$

where b_1 and b_2 are defined in (18) and (19).

To prove Proposition 1, we need to maximize $U_{\rm GMA}(\xi_{\rm fix},\xi_{\rm mv})$ with respect to both $\xi_{\rm fix}$ and $\xi_{\rm mv}$. The first order conditions are

$$\begin{aligned} \frac{\partial U_{\rm GMA}(\xi_{\rm fix},\xi_{\rm mv})}{\partial \xi_{\rm fix}}|_{\xi_{\rm fix}=\xi_{\rm fix}^*,\xi_{\rm mv}=\xi_{\rm mv}^*} = 0, \\ \frac{\partial U_{\rm GMA}(\xi_{\rm fix},\xi_{\rm mv})}{\partial \xi_{\rm mv}}|_{\xi_{\rm fix}=\xi_{\rm fix}^*,\xi_{\rm mv}=\xi_{\rm mv}^*} = 0, \end{aligned} \tag{A19}$$

which implies

$$\begin{aligned} &\mu_0+\mu_1\bar{X}-r-\sigma_s^2\xi_{\rm fix}-\sigma_s^2\xi_{\rm mv}b_2=0, \\ &b_1+(\mu_0+\mu_1\bar{X}-r)b_2-\sigma_s^2(\xi_{\rm fix}+\xi_{\rm mv})b_2=0. \end{aligned}$$

With some algebra, we obtain the optimal solution:

$$\begin{aligned} \xi_{\rm fix}^* &= \frac{\mu_0+\mu_1\bar{X}-r}{\sigma_s^2}-\frac{\mu_1 b_1}{(1-b_2)\sigma_s^2}, \\ \xi_{\rm mv}^* &= \frac{\mu_1 b_1}{b_2(1-b_2)\sigma_s^2}. \end{aligned}$$

Since the value function for log-utility associated with $\xi_{\rm fix1}^*$ is

$$U_{\rm fix1}^* = \log W_0 + rT + \frac{(\mu_0+\mu_1\bar{X}-r)^2}{2\sigma_s^2}T,$$

we obtain equation (26) by substituting this into $U_{\rm GMA}(\xi_{\rm fix},\xi_{\rm mv})$ evaluated at the optimal solution $(\xi_{\rm fix}^*,\xi_{\rm mv}^*)$.

To prove Proposition 2, we simply let $\xi_{\rm fix}=\xi_{\rm fix1}^*$, and optimize $U_{\rm GMA}(\xi_{\rm fix1}^*,\xi_{\rm mv})$ over $\xi_{\rm mv}$ alone. Similar algebra yields the solution. The proof of Proposition 3 follows analogously.

3 Proof of Equation (42)

To maximize $U(\gamma)$ of (41) over ξ_{mv}, it is equivalent to maximize

$$\max_{\xi_{\mathrm{mv}}} f(\xi_{\mathrm{mv}}) = \xi_{\mathrm{mv}}(\phi_0 + \phi_1\xi_{\mathrm{mv}} + \phi_2\xi_{\mathrm{mv}}^2 + \phi_3\xi_{\mathrm{mv}}^3).$$

The first-order condition is

$$f'(\xi_{\mathrm{mv}}) = \phi_0 + 2\phi_1\xi_{\mathrm{mv}} + 3\phi_2\xi_{\mathrm{mv}}^2 + 4\phi_3\xi_{\mathrm{mv}}^3 = 0, \tag{A20}$$

which in turn can be transformed to

$$y^3 + py + q = 0, \tag{A21}$$

where

$$y = \xi_{\mathrm{mv}} + \frac{\phi_2}{4\phi_3}$$

with p and q given in (43). Numerical computations show that, for a wide range of parameters of interest, we have

$$q^2 + \frac{4p^3}{27} > 0. \tag{A22}$$

The solution to cubic equation (A21) is known as Cardano solution (e.g., Curtis (1944)), which is given by

$$y^* = -\left[\frac{q + \sqrt{q^2 + 4p^3/27}}{2}\right]^{\frac{1}{3}} + \frac{p}{3}\left[\frac{q + \sqrt{q^2 + 4p^3/27}}{2}\right]^{-\frac{1}{3}}.$$

Under condition (A22), this is the unique real root. Hence

$$\xi_{\mathrm{mv}}^* = -\frac{\phi_2}{4\phi_3} + y^*$$

which is the same as equation (42). Furthermore, it can be verified that $\phi_1 < 0$, and so this solution to (A20) is indeed a maximum.

4 Computing the ML Estimators

Following Huang and Liu (2007), the continuously compounded return $R_{t+1} = \log(S_{t+1}/S_t)$ and X_{t+1} are jointly Gaussian, and the log-likelihood function, conditional on X_0, can be written as

$$\begin{aligned}
\pounds(\Theta) &= \sum_{t=1}^{T} \log f\left(R_t, X_t | X_{t-1}; \Theta\right) \\
&= -\frac{T}{2}\left(2\log 2\pi + \log\sigma_1^2 + \log\sigma_2^2 + \log\left(1-\rho_{12}^2\right)\right) \\
&\quad -\frac{1}{2\left(1-\rho_{12}^2\right)}\sum_{t=1}^{T}\left\{\frac{\left(R_t - a_{11} - a_{12}X_{t-1}\right)^2}{\sigma_1^2} + \frac{\left(X_t - b_{11} - b_{12}X_{t-1}\right)^2}{\sigma_2^2}\right. \\
&\quad \left. -\frac{2\rho_{12}\left(R_t - a_{11} - a_{12}X_{t-1}\right)\left(X_t - b_{11} - b_{12}X_{t-1}\right)}{\sigma_1\sigma_2}\right\},
\end{aligned}$$

where $\Theta \equiv (a_{11}, a_{12}, b_{11}, b_{12}, \sigma_1, \sigma_2, \rho_{12})$ with

$$a_{11} = (\mu_0 - \frac{1}{2}\sigma_s^2 - \frac{\mu_1\theta_0}{\theta_1})\Delta t + \frac{\mu_1\theta_0}{\theta_1^2}\left(e^{\theta_1\Delta t} - 1\right), a_{12} = \frac{\mu_1}{\theta_1}\left(e^{\theta_1\Delta t} - 1\right), b_{11} = \frac{\theta_0}{\theta_1}\left(e^{\theta_1\Delta t} - 1\right), b_{12} = e^{\theta_1\Delta t},$$

$$\sigma_1^2 = (\sigma_s^2 + \frac{\mu_1^2}{\theta_1^2}\sigma_x^2 - \frac{2\mu_1}{\theta_1}\rho\sigma_s\sigma_x)\Delta t + \frac{1}{2\theta_1}(e^{2\theta_1\Delta t} - 1)\frac{\mu_1^2}{\theta_1^2}\sigma_x^2 + \frac{2\mu_1}{\theta_1^2}(e^{\theta_1\Delta t} - 1)(\rho\sigma_s\sigma_x - \frac{\mu_1}{\theta_1}\sigma_x^2),$$

$$\sigma_2^2 = \frac{\sigma_x^2}{2\theta_1}(e^{2\theta_1\Delta t} - 1),$$

$$\rho_{12}\sigma_1\sigma_2 = \frac{\mu_1}{2\theta_1^2}(e^{\theta_1\Delta t} - 1)^2\sigma_x^2 + \frac{\rho\sigma_s\sigma_x}{\theta_1}(e^{\theta_1\Delta_t} - 1).$$

Let Y be a $T \times 2$ matrix formed by observation on R_t and X_t, and Z be formed by a T-vector of ones and the T values of X_{t-1}. Define

$$B = \begin{pmatrix} a_{11} & b_{11} \\ a_{12} & b_{12} \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}. \tag{A23}$$

Then, the estimator of B is $\hat{B} = (X'X)^{-1}X'Y$, and that of Σ is $\hat{\Sigma} = (Y - X\hat{B})'(Y - X\hat{B})/T$. The estimator for the original parameters, such as μ_0, can be backed out from these estimates.

5 Proof for the Optimal Lags

Now we need to optimize equation (26), (31) and (34) over L. To prove Proposition 4, consider $U^*_{\text{GMA1}} - U^*_{\text{fix1}}$ and $U^*_{\text{GMA2}} - U^*_{\text{fix1}}$. Ignoring some constants, the target functions are

$$U_1 = \frac{b_1^2}{b_2(1-b_2)}(1-\frac{L}{T}) = V_1(1-\frac{L}{T}), \quad \text{and} \quad U_2 = \frac{b_1^2}{b_2}(1-\frac{L}{T}) = V_2(1-\frac{L}{T}), \tag{A24}$$

where V_1 and V_2 are defined accordingly. Since V_1 and V_2 are T independent, so are their maximum over L. As T is large, $1-\frac{L}{T}$ can be ignored, and hence we need only to maximize V_1 and V_2.

The first-order condition for maximizing V_2 is

$$V_2' = \frac{2b_1b_1'b_2 - b_1^2b_2'}{b_2^2} = 0. \tag{A25}$$

Substituting those approximate expressions (61) and (62) for b_1 and b_2, we have

$$2h'(x)f(Ax) - 2Axh(x)f(Ax) - \frac{Ah(x)f^2(Ax)}{N(Ax)} = 0. \tag{A26}$$

This is a transcendant equation that is difficult to solve without further simplifications. It can be shown that the third term is dominated by the first one when $x < 1$, and by the second one when $x > 1$. Ignoring the third term, we need only to optimize

$$b_1 = h(x) \cdot f(Ax). \tag{A27}$$

The Taylor expansion for $h(x)$ is

$$h(x) = \frac{x}{2} - \frac{x^3}{6} + \frac{x^5}{24} + \cdots, \tag{A28}$$

which implies that (A27) can be approximated by

$$(\frac{x}{2} - \frac{x^3}{6} + \frac{x^5}{24})\exp(-\frac{A^2x^2}{2}). \tag{A29}$$

Taking derivative with respect to x and letting it be equal to zero, we obtain, after ignoring higher-order terms,

$$(\frac{5}{24} + \frac{A^2}{6})x^4 - \frac{1+A^2}{2}x^2 + \frac{1}{2} = 0.$$

The smaller root of the above quadratic equation, which corresponds to the maximum, is the solution for the second case of Proposition 4.

To provide solution for the first case, we now maximize V_1. Its denominator can be approximated by $N(Ax) \cdot N(-Ax)$, and hence

$$V_1 \approx \frac{h^2(x) f^2(Ax)}{N(Ax)N(-Ax)} = \frac{1}{C \cdot N(Ax)} \left[h(x)\sqrt{f(Ax)} \right]^2 .$$

where we have used the approximation $N(-Ax) \approx C \cdot f(Ax)$ for $Ax > 0$ and large. Similar to the earlier case, we can ignore $N(Ax)$, and hence the target function becomes to $h(x) \cdot \sqrt{f(Ax)}$. This has the same form as (A27) with $\frac{A}{\sqrt{2}}$ plays the role of earlier A. Therefore, the solution follows.

Finally, to derive (67), we need to maximize $U_3 = U^*_{\text{GMA3}} - \frac{(\mu_s - r)^2}{2\sigma_s^2} L$. Similarly, this can be replaced by a target function

$$\begin{aligned} V_3 &= [\mu_1 C_4 h(x) f(Ax) + C_5 N(Ax)] \cdot (1 - \frac{x^2}{|\theta_1|T}) \\ &= [\mu_1 C_4 \cdot \frac{1}{x}(1 - \frac{1 - e^{-x^2}}{x^2}) f(Ax) + C_5 N(Ax)] \cdot (1 - \frac{x^2}{|\theta_1|T}) \\ &\approx C_5 N(Ax) \cdot (1 - \frac{x^2}{|\theta_1|T}), \end{aligned} \qquad \text{(A30)}$$

where the last approximation is due to the dominance of the second term in the bracket. The first-order condition is

$$f(Ax) \cdot (1 - \frac{x^2}{|\theta_1|T}) - \frac{2}{|\theta_1|T} x N(Ax) = 0. \qquad \text{(A31)}$$

Since there is only one solution, we can verify that

$$|\theta_1|T >> 1, \quad Ax >> 1, \quad \frac{x^2}{|\theta_1|T} \to 0, \qquad \text{(A32)}$$

and hence we can reduce the first-order condition to $Af(Ax) \approx \frac{2}{|\theta_1|T} x$. This implies (67).

Table 1: Calibrated Model Parameters

The table reports parameter estimates for the following cum-dividend price process,

$$\frac{dS_t}{S_t} = (\mu_0 + \mu_1 X_t)dt + \sigma_s dB_t,$$
$$dX_t = (\theta_0 + \theta_1 X_t)dt + \sigma_x dZ_t,$$

where $\mu_0, \mu_1, \sigma_s, \theta_0, \theta_1$ and σ_x are parameters, X_t is a predictive variable, and B_t and Z_t are standard Brownian motions with correlation coefficient ρ. The estimation is based on monthly returns on S&P500 from December 1926 to December 2004, and on X_t which is the dividend yield, term-spread and payout ratio, respectively, in the time period.

	Dividend yield	Term-spread	Payout ratio
μ_0	0.0310	0.0969	0.2824
μ_1	2.0715	1.2063	-0.2917
σ_s	0.1946	0.1947	0.1942
θ_0	0.0100	0.0087	0.0140
θ_1	-0.2532	-0.5270	-0.0296
σ_x	0.0122	0.0132	0.0497
ρ	-0.0730	0.0014	-0.0031

Table 2: Utility Losses Versus Optimal Strategy ($L = 50$)

The table reports the utility losses, measured as percentage points of initial wealth, that one is willing to give up to switch from a given strategy to the optimal dynamic one. The moving average (MA) lag length L is set equal to 50 days. Fix1 is the standard fixed allocation rule and Fix2 is such a rule accounting for predictability. Fix1+MA and Fix2+MA are those combined with the MA. Optimal MA is the strategy that uses the MA optimally without any combination of fixed rules, and MA1, MA2 and MA3 are ad hoc MA only strategies whose stock allocations are 100%, Fix1 and Fix2, respectively when the MA indicates a 'buy' signal, and are nothing otherwise.

	Dividend yield	Term-spread	Payout ratio
	T=10		
Fix1	8.8445	3.8948	20.8564
Fix2	7.9044	1.5676	18.0614
Fix1+MA	8.1765	2.6154	18.6388
Fix2+MA	7.8951	1.5504	18.0613
Optimal MA	16.3033	13.0875	27.6918
MA1	17.7622	14.0367	28.0962
MA2	17.2341	14.4233	30.8149
MA3	16.8626	13.6139	28.4319
	T=20		
Fix1	16.6797	7.6093	31.2747
Fix2	15.1708	3.1122	30.6817
Fix1+MA	15.3441	4.6476	30.6190
Fix2+MA	15.1608	3.0586	30.6814
Optimal MA	29.3260	23.9314	41.3094
MA1	31.5715	25.6522	43.1087
MA2	30.6943	26.2543	43.9738
MA3	30.0402	24.8735	42.9918
	T=40		
Fix1	30.3693	14.6129	50.6936
Fix2	28.0266	6.1289	49.4951
Fix1+MA	27.9065	7.9488	50.2724
Fix2+MA	27.9847	5.8799	49.4951
Optimal MA	50.3555	42.9099	59.3592
MA1	53.6836	45.2836	63.6320
MA2	51.9720	45.5609	60.0394
MA3	51.1360	43.7584	61.3035

Table 3: Utility Losses Versus Optimal Strategy ($L = 200$)

The table reports the utility losses, measured as percentage points of initial wealth, that one is willing to give up to switch from a given strategy to the optimal dynamic one. The moving average (MA) lag length L is set equal to 200 days. Fix1 is the standard fixed allocation rule and Fix2 is such a rule accounting for predictability. Fix1+MA and Fix2+MA are those combined with the MA. Optimal MA is the strategy that uses the MA optimally without any combination of fixed rules, and MA1, MA2 and MA3 are ad hoc MA only strategies whose stock allocations are 100%, Fix1 and Fix2, respectively when the MA indicates a 'buy' signal, and are nothing otherwise.

	Dividend yield	Term-spread	Payout ratio
	T=10		
Fix1	8.8445	3.8948	20.8564
Fix2	7.9044	1.5676	18.0614
Fix1+MA	8.1253	2.4974	18.1453
Fix2+MA	7.8961	1.5472	18.0587
Optimal MA	15.1814	11.5260	24.8460
MA1	17.2853	14.0991	25.8806
MA2	16.3831	13.6423	28.1928
MA3	16.1825	13.3845	26.0471
	T=20		
Fix1	16.6797	7.6093	31.2747
Fix2	15.1708	3.1122	30.6817
Fix1+MA	14.9916	4.4603	30.5210
Fix2+MA	15.1677	3.0395	30.6817
Optimal MA	26.5693	21.4349	38.7195
MA1	30.6418	24.8091	41.2517
MA2	29.2722	24.6615	41.5063
MA3	28.7557	23.6940	40.5912
	T=40		
Fix1	30.3693	14.6129	50.6936
Fix2	28.0266	6.1289	49.4951
Fix1+MA	27.3408	6.9752	50.6936
Fix2+MA	28.0207	5.8872	49.4951
Optimal MA	45.3152	36.6855	55.4043
MA1	49.7462	40.3596	59.8256
MA2	48.7554	41.7133	56.5179
MA3	47.5196	39.0324	57.9567

Table 4: Utility Losses Versus Optimal Strategy for Arithmetic Average

The table reports the utility losses, measured as percentage points of initial wealth, that one is willing to give up to switch from a given strategy to the optimal dynamic one. The moving average (MA) lag length L is set equal to 200 days. Unlike Table 3, the MA is now based on the arithmetic average instead of the geometric average. Fix1 is the standard fixed allocation rule and Fix2 is such a rule accounting for predictability. Fix1+MA and Fix2+MA are those combined with the MA. Optimal MA is the strategy that uses the MA optimally without any combination of fixed rules.

	Dividend yield	Term-spread	Payout ratio
	T=10		
Fix1	8.8445	3.8948	20.8564
Fix2	7.9044	1.5676	18.0614
Fix1+MA	8.0907	2.5136	18.1525
Fix2+MA	7.9008	1.5466	18.0701
Optimal MA	15.0735	11.6682	24.9513
	T=20		
Fix1	16.6797	7.6093	31.2747
Fix2	15.1708	3.1122	30.6817
Fix1+MA	15.0358	4.3909	30.5307
Fix2+MA	15.1663	3.0526	30.7185
Optimal MA	26.8732	21.3547	38.9168
	T=40		
Fix1	30.3693	14.6129	50.6936
Fix2	28.0266	6.1289	49.4951
Fix1+MA	27.3783	6.6970	50.6600
Fix2+MA	28.0170	5.9626	49.6410
Optimal MA	45.7154	36.2868	55.8563

Table 5: Utility Losses Versus Optimal Strategy with Ex-dividend Price

The table reports the utility losses, measured as percentage points of initial wealth, that one is willing to give up to switch from a given strategy to the optimal dynamic one. The moving average (MA) lag length L is set equal to 200 days. Unlike Table 3, the MA is now based on the ex-dividend price instead of the cum-dividend price. Fix1 is the standard fixed allocation rule and Fix2 is such a rule accounting for predictability. Fix1+MA and Fix2+MA are those combined with the MA. Optimal MA is the strategy that uses the MA optimally without any combination of fixed rules.

	Dividend yield	Term-spread	Payout ratio
	T=10		
Fix1	8.8445	3.8948	20.8564
Fix2	7.9044	1.5676	18.0614
Fix1+MA	8.1528	2.7371	18.1492
Fix2+MA	7.8978	1.5212	18.0638
Optimal MA	16.0852	13.2148	25.7408
	T=20		
Fix1	16.6797	7.6093	31.2747
Fix2	15.1708	3.1122	30.6817
Fix1+MA	15.1637	4.5174	30.5592
Fix2+MA	15.1619	3.0577	30.7276
Optimal MA	28.6068	23.0312	40.0777
	T=40		
Fix1	30.3693	14.6129	50.6936
Fix2	28.0266	6.1289	49.4951
Fix1+MA	27.3789	6.7349	50.8201
Fix2+MA	28.0283	6.0231	49.6831
Optimal MA	47.1715	38.6881	56.7177

Table 6: Comparison Under Parameter Uncertainty (T=10)

The table reports both the utilities of the optimal learning, the standard fixed and the combination of the fixed with MA (GMA) strategies, and the associated certainty-equivalent losses, measured as percentage points of initial wealth, relative to the optimal learning strategy. The MA length is 200 days and investment horizon is $T = 10$ years. The predictability parameter β is captured by a mean-reverting process starting from its long-term level $\bar{\beta}_0 = 2.0715$. The standard normal prior on β_0 has a prior mean b_0 and standard deviation $\sqrt{\nu_0}$.

b_0	$\sqrt{\nu_0}$	U_{optl}	U_{fix}	U_{GMA}	CE_{fix}	CE_{GMA}
0	1	1.1371	1.0144	1.0204	12.27	11.67
0	2	1.1384	1.0144	1.0204	12.40	11.80
0	3	1.1340	1.0144	1.0204	11.96	11.36
0	4	1.1211	1.0144	1.0204	10.67	10.07
4	1	1.1470	1.0144	1.0204	13.26	12.66
4	2	1.1490	1.0144	1.0204	13.46	12.86
4	3	1.1451	1.0144	1.0204	13.07	12.47
4	4	1.1313	1.0144	1.0204	11.69	11.09
6	1	0.9989	1.0144	1.0204	-1.55	-2.15
6	2	1.0153	1.0144	1.0204	0.09	-0.51
6	3	1.0295	1.0144	1.0204	1.51	0.91
6	4	1.0349	1.0144	1.0204	2.05	1.45
7	1	0.8880	1.0144	1.0204	-12.64	-13.24
7	2	0.9151	1.0144	1.0204	-9.93	-10.53
7	3	0.9424	1.0144	1.0204	-7.20	-7.80
7	4	0.9606	1.0144	1.0204	-5.38	-5.98

Table 7: Comparison Under Parameter Uncertainty (T=5)

The table reports both the utilities of the optimal learning, the standard fixed and the combination of the fixed with MA (GMA) strategies, and the associated certainty-equivalent losses, measured as percentage points of initial wealth, relative to the optimal learning strategy. The MA length is 200 days and investment horizon is $T = 5$ years. The predictability parameter β is captured by a mean-reverting process starting from its long-term level $\bar{\beta}_0 = 2.0715$. The standard normal prior on β_0 has a prior mean b_0 and standard deviation $\sqrt{\nu_0}$.

b_0	$\sqrt{\nu_0}$	U_{optl}	U_{fix}	U_{GMA}	CE_{fix}	CE_{GMA}
0	1	0.5026	0.4567	0.4603	4.59	4.23
0	2	0.5035	0.4567	0.4603	4.68	4.32
0	3	0.5005	0.4567	0.4603	4.38	4.02
0	4	0.4914	0.4567	0.4603	3.47	3.11
4	1	0.5144	0.4567	0.4603	5.77	5.41
4	2	0.5147	0.4567	0.4603	5.80	5.44
4	3	0.5107	0.4567	0.4603	5.40	5.04
4	4	0.5002	0.4567	0.4603	4.35	3.99
6	1	0.4037	0.4567	0.4603	-5.30	-5.66
6	2	0.4143	0.4567	0.4603	-4.24	-4.60
6	3	0.4226	0.4567	0.4603	-3.41	-3.77
6	4	0.4241	0.4567	0.4603	-3.26	-3.62
7	1	0.3190	0.4567	0.4603	-13.77	-14.13
7	2	0.3375	0.4567	0.4603	-11.92	-12.28
7	3	0.3552	0.4567	0.4603	-10.15	-10.51
7	4	0.3658	0.4567	0.4603	-9.09	-9.45

Table 8: Comparison Under Model Uncertainty

The table reports the utility losses of the estimated optimal GMA, the optimal strategies derived from the wrong models and the estimated fixed strategy, measured as percentage points of initial wealth relative to the optimal strategy of knowing the true model. The moving average (MA) lag length L and investment horizon T are set equal to 50, 100, 200 days and 5, 10 and 20 years, respectively.

	Estimated Optimal GMA			Estimated	Uncertain Model	
	L=50	L=100	L=200	Fixed Strategy	Wrong Model 1	Wrong Model 2
Panel A: Dividend yield						
T=5	5.2284	5.3326	5.3326	5.6161	6.5926	17.2875
T=10	13.5583	13.4199	13.3593	13.9894	15.4393	38.9453
T=20	28.2943	27.9709	27.8483	28.3660	31.0094	70.7737
Panel B: Term-spread						
T=5	1.3607	1.4297	1.4198	1.8330	6.5926	9.8685
T=10	3.8922	3.7576	3.6517	5.4083	15.4393	23.4327
T=20	8.7347	8.5703	8.3598	11.2636	31.0232	50.3464
Panel C: Payout ratio						
T=5	3.3718	3.4588	3.6420	4.1897	17.2875	9.8685
T=10	12.3133	12.3922	12.7245	16.4312	38.9453	23.4327
T=20	34.9361	35.3770	35.4674	40.1124	70.7737	50.3365

Figure 1: Effect of Lag Length

The figure plots the certainty-equivalent losses versus the moving average lag length measured in days in the three predictable models.