# ASSET ALLOCATION WITH TIME SERIES MOMENTUM AND REVERSAL

XUE-ZHONG HE*, KAI LI* AND YOUWEI LI**

*University of Technology Sydney, Business School
PO Box 123, Broadway, NSW 2007, Australia
**School of Management
Queen's University Belfast, Belfast, UK
tony.he1@uts.edu.au, kai.li@uts.edu.au, y.li@qub.ac.uk

Abstract. We develop a continuous-time asset price model to capture the well documented time series momentum and reversal. The optimal asset allocation strategy is derived theoretically and tested empirically. We show that, by combining with market fundamentals and timing opportunity with respect to market trend and volatility, the optimal strategy based on the time series momentum and reversal outperforms significantly, both in-sample and out-of-sample, the S&P500 and pure strategies based only on either time series momentum or mean reverting. The outperforming also holds for different time horizons and with short-sale constraints. Furthermore, the outperformance is immune to market states, investor sentiment and market volatility.

*Key words*: Momentum, reversal, optimal asset allocation, profitability.

*JEL Classification*: G12, G14, E32

---

*Date*: February 15, 2016.

*Acknowledgement*: We would like to thank Li Chen, Blake LeBaron, Michael Moore, Bernt Øksendal, Lei Shi, Jun Tu, Zhen Wu and Chao Zhou, and conference and seminar participants at the 2014 International Conference on CEF (Olso), 2015 MFA Annual Meeting (Chicago), Beijing Normal University, Central University of Finance and Economics, Chinese University of Hong Kong, National University of Singapore, Queen's University Belfast, Renmin University, Shanghai University of Finance and Economics, Tianjin University, University of Technology Sydney, Sun Yat-sen University, University of Urbino and York University for helpful comments and suggestions. Financial support from the Australian Research Council (ARC) under Discovery Grant (DP130103210) is gratefully acknowledged. The usual caveats apply.

*Corresponding author: Kai Li, University of Technology Sydney, UTS Business School, Finance Discipline Group, PO Box 123, Broadway, NSW 2007, Australia, kai.li@uts.edu.au.*

# Asset Allocation with Time Series Momentum and Reversal

Abstract

We develop a continuous-time asset price model to capture the well documented time series momentum and reversal. The optimal asset allocation strategy is derived theoretically and tested empirically. We show that, by combining with market fundamentals and timing opportunity with respect to market trend and volatility, the optimal strategy based on the time series momentum and reversal outperforms significantly, both in-sample and out-of-sample, the S&P500 and pure strategies based only on either time series momentum or mean reverting. The outperforming also holds for different time horizons and with short-sale constraints. Furthermore, the outperformance is immune to market states, investor sentiment and market volatility.

*Key words*: Momentum, reversal, optimal asset allocation, profitability.

*JEL Classification*: G12, G14, E32

## 1. Introduction

Equity return momentum in the short-run and reversal in the long-run are two of the most prominent financial market anomalies. Though market timing opportunities under mean reversion in equity return are well documented (see, for example, Campbell and Viceira (1999) and Wachter (2002)), time series momentum (TSM) that characterizes strong positive predictability of a security's own past returns has been explored recently in Moskowitz, Ooi and Pedersen (2012). If an investor incorporates both return momentum and reversal into a trading strategy optimally, the investor would expect to outperforming the strategies based only on return momentum or reversal, and even the market index.

This paper examines theoretically and empirically how to optimally explore time series momentum and reversal in financial markets. We first introduce a financial market asset price model to incorporate momentum and mean-reverting components. By solving a dynamic asset allocation problem, we derive the optimal investment strategy of combining momentum and mean reversion in closed form, which includes pure momentum and pure mean-reverting strategies as special cases. By estimating the model to monthly returns of the S&P 500 index, we show that the optimal strategy outperforms, measured by the utility of portfolio wealth and Sharpe ratio, not only the strategies based on the pure momentum and pure mean-reversion models but also the S&P 500 index.

This paper makes three contributions to the literature. Firstly, we find that the performance of TSM strategy can be significantly improved by combining with market fundamentals, while the performance of mean-reverting strategy can be significantly improved by combining with TSM. To demonstrate the outperformance of the optimal strategy over the pure TSM strategy, we derive a suboptimal portfolio purely based on the TSM effect. We find that this portfolio is not able to outperform the market portfolio and the optimal portfolio. Comparing the performance of the optimal strategy with the TSM strategy used in Moskowitz et al. (2012), we show that the optimal strategy outperforms the TSM and passive holding strategies. Essentially, in contrast to a TSM strategy based on trend only, the optimal strategy takes into account not only the trading signal based on momentum and fundamentals but also the size of position, which is associated with market volatility. Without considering the fundamentals, the pure momentum portfolio is highly leveraged, and hence suffers from higher risks. We also derive another suboptimal portfolio, the pure mean-reverting portfolio, by ignoring the TSM effect. We find that this portfolio is based conservatively on fundamental investments, leading to a stable growth rate of portfolio wealth, but is not able to explore the price trend, especially during extreme market periods, and hence underperforms the optimal portfolio. In addition to the above model-based results, we further investigate the model-free performance of the optimal strategy following Moskowitz et al. (2012). We find that our optimal strategy outperforms the TSM strategy with respect to Sharpe ratio and cumulative excess return.

Secondly, to the best of our knowledge, this paper is the first to theoretically examine the effect of the time horizon of the TSM on the performance of the optimal portfolio. Empirically, the time horizon in TSM is a fixed look-back period. It plays a crucial role in the performance of momentum strategies, which have been investigated extensively in the empirical literature.[1] However, due to the technical challenge, there are few theoretical results concerning the effect of the time horizon. The asset price model developed in this paper takes the time horizon of TSM into account directly. As the result, historical prices underlying the TSM component affect asset prices, leading to a non-Markov process characterized by stochastic delay differential equations (SDDEs). This is very different from the Markov asset price process documented in the literature (Merton 1969, 1971) in which it is difficult to model the time series momentum strategy explicitly. In the case of Markov processes, the stochastic control problem is most frequently solved using the dynamic programming method and HJB equation. However, solving the optimal control problem for SDDEs using the dynamic programming method becomes more challenging because it involves infinite-dimensional partial differential equations. One way to solve the problem is to apply a type of Pontryagin maximum principle, which has been developed recently by Chen and Wu (2010) and Øksendal et al. (2011) for the optimal control problem of SDDEs. By exploring these latest advances in the theory of the maximum principle for control problems of SDDEs, we derive the optimal strategies in closed form. This helps us to study thoroughly the impact of historical information on the profitability of different strategies based on different time horizons, in particular of TSM trading strategies based on moving averages over different time horizons. More interestingly, we show that the optimal strategy based on the estimated model performs the best when the TSM is based on the past 9 to 12 months.

Thirdly, we show that, in addition to price trend, position size is another very important factor for momentum trading. The optimal position size derived in this paper is determined by the level of trading signals and market volatility. In the empirical literature, momentum trading only considers the trading signals of price trend and takes a constant position to trade. We show that, if we only consider the sign of trading signals indicated by the optimal strategy and take a unit position to trade, the portfolio is not able to outperform the optimal portfolio for all time horizons. The robustness of the performance of the optimal strategy is also tested for

[1] See, for example, De Bondt and Thaler (1985) and Jegadeesh and Titman (1993).

different sample periods, out-of-sample predictions, short-sale constraints, market states, investor sentiment and market volatility.

This paper is closely related to the literature on reversal and momentum. Reversal is the empirical observation that assets performing well (poorly) over a long period tend subsequently to underperform (outperform). Momentum is the tendency of assets with good (bad) recent performance to continue outperforming (underperforming) in the short term. Reversal and momentum have been documented extensively for a wide variety of assets. On the one hand, Fama and French (1988) and Poterba and Summers (1988), among many others, document reversal for holding periods of more than one year, which induces negative autocorrelation in returns.[2] The value effect documented in Fama and French (1992) is closely related to reversal, whereby the ratio of an asset's price relative to book value is negatively related to subsequent performance. Mean reversion in equity returns has been shown to induce significant market timing opportunities (Campbell and Viceira 1999, Wachter 2002 and Koijen, Rodríguez and Sbuelz 2009). On the other hand, the literature mostly studies cross-sectional momentum.[3] More recently, Moskowitz et al. (2012) investigate TSM that characterizes strong positive predictability of a security's own past returns. For a large set of futures and forward contracts, Moskowitz et al. (2012) find that TSM based on excess returns over the past 12 months persists for between one and 12 months and then partially reverses over longer horizons. They provide strong evidence for TSM based on the moving average of look-back returns. This effect based purely on a security's own past returns is related to, but different from, the cross-sectional momentum phenomenon studied extensively in the literature. Through return decomposition, Moskowitz et al. (2012) argue that positive auto-covariance is the main driving force for TSM and cross-sectional momentum effects, while the contribution of serial cross-correlations and variation in mean returns is small. Intuitively, a strategy taking into account both the short-run momentum and long-run mean reversion in time series should be profitable and outperform pure momentum and pure mean-reversion strategies. In this paper, we provide a justification to this intuition theoretically and empirically.

The apparent persistent and sizeable profits of strategies based on momentum and reversal have attracted considerable attention, and many studies have tried to

[2]For instance, Jegadeesh (1991) finds that the next one-month returns can be negatively predicted by their lagged multiyear returns. Lewellen (2002) shows that the past one-year returns negatively predict future monthly returns for up to 18 months.

[3]Jegadeesh and Titman (1993) document cross-sectional momentum for individual U.S. stocks, predicting returns over horizons of 3–12 months using returns over the past 3–12 months. The evidence has been extended to stocks in other countries (Fama and French 1998), stocks within industries (Cohen and Lou 2012), across industries (Cohen and Frazzini 2008), and the global market with different asset classes (Asness, Moskowitz and Pedersen 2013).

explain the phenomena.[4] This paper is largely motivated by the empirical literature testing trading signals with combinations of momentum and reversal.[5] Asness, Moskowitz and Pedersen (2013) highlight that studying value and momentum jointly is more powerful than examining each in isolation.[6] Huang, Jiang, Tu and Zhou (2013) find that both mean reversion and momentum can coexist in the S&P 500 index over time. Extending the literature, this paper develops an asset price model by taking both mean reversion and time series momentum directly into account and demonstrates the explanatory power of the model through the outperformance of the optimal strategy.

This paper is also largely inspired by Koijen, Rodríguez and Sbuelz (2009), who propose a theoretical model in which stock returns exhibit momentum and mean-reversion effects. This paper is however different from Koijen et al. (2009) in two aspects. Firstly, in Koijen et al. (2009), the momentum is calculated from the entire set of historical returns with geometrically decaying weights, instead of a fixed look-back period. This effectively reduces the price dynamics to a Markovian system, and enables a thorough analysis of the performance of the hedging demand implied by the model. In this paper, we follow the empirical literature and model TSM by the standard moving average over a moving window with a fixed look-back period. Our model of momentum complements in a unique way to the theoretical study of Koijen et al. (2009) and many empirical studies that do not study systematically the role of momentum with different look-back period. We study explicitly the impacts of different look-back periods on the performance of momentum-related trading strategies. Secondly, instead of studying the economic gains of hedging due to momentum in Koijen et al. (2009), we focus more on the performance of

[4]Among which, the three-factor model of Fama and French (1996) can explain long-run reversal but not short-run momentum. Barberis, Shleifer and Vishny (1998) argue that these phenomena are the result of the systematic errors investors make when they use public information to form expectations of future cash flows. Models Daniel, Hirshleifer and Subrahmanyam (1998), with single representative agent, and Hong and Stein (1999), with different trader types, attribute the under reaction to overconfidence and overreaction to biased self-attribution. Barberis and Shleifer (2003) show that style investing can explain momentum and value effects. Sagi and Seasholes (2007) present an option model to identify observable firm-specific attributes that drive momentum. Vayanos and Woolley (2013) show that slow-moving capital can also generate momentum. He and Li (2015) find that momentum strategies can be self-fulfilling.

[5]For example, Balvers and Wu (2006) and Serban (2010) show empirically that a combination of momentum and mean-reversion strategies can outperform pure momentum and pure mean-reversion strategies for equity markets and foreign exchange markets respectively.

[6]They find that separate factors for value and momentum best explain the data for eight different markets and asset classes. Furthermore, they show that momentum loads positively and value loads negatively on liquidity risk; however, an equal-weighted combination of value and momentum is immune to liquidity risk and generates substantial abnormal returns.

the optimal strategy comparing with the market, TSM, and mean-reversion trading strategies.

The paper is organized as follows. We first present the model and derive the optimal asset allocation in Section 2. In Section 3, we estimate the model to the S&P 500 and conduct a performance analysis of the optimal portfolio and examine the impact of hedging demand. We then investigate the time horizon effect in Section 4. Section 5 concludes. All the proofs and robustness analysis are included in the appendices.

## 2. The Model and Optimal Asset Allocation

In this section, we introduce an asset price model and study the optimal investment decision problem. We consider a financial market with two tradable securities, a riskless asset $B$ satisfying

$$\frac{dB_t}{B_t} = rdt \tag{2.1}$$

with a constant riskless rate $r$, and a risky asset. Let $S_t$ be the price of the risky asset or the level of a market index at time $t$ where dividends are assumed to be reinvested. Empirical studies on return predictability, see for example Fama (1991), have shown that the most powerful predictive variables of future stock returns in the United States are past returns, dividend yield, earnings/price ratio, and term structure variables. Following this literature and Koijen et al. (2009), we model the expected return by a combination of a momentum term $m_t$ based on the past returns and a long-run mean-reversion term $\mu_t$ based on market fundamentals such as dividend yield. Consequently, we assume that the stock price $S_t$ follows

$$\frac{dS_t}{S_t} = \big[\phi m_t + (1-\phi)\mu_t\big]dt + \sigma_S' dZ_t, \tag{2.2}$$

where $\phi$ is a constant,[7] measuring the weight of the momentum component $m_t$, $\sigma_S$ is a two-dimensional volatility vector (and $\sigma_S'$ stands for the transpose of $\sigma_S$), and $Z_t$ is a two-dimensional vector of independent Brownian motions. The uncertainty is represented by a filtered probability space $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$ on which the two-dimensional Brownian motion $Z_t$ is defined. As usual, the mean-reversion process $\mu_t$ is defined by an Ornstein-Uhlenbeck process,

$$d\mu_t = \alpha(\bar{\mu} - \mu_t)dt + \sigma_\mu' dZ_t, \qquad \alpha > 0, \qquad \bar{\mu} > 0, \tag{2.3}$$

where $\bar{\mu}$ is the constant long-run expected return, $\alpha$ measures the rate at which $\mu_t$ converges to $\bar{\mu}$, and $\sigma_\mu'$ is a two-dimensional volatility vector. The momentum term

[7]The dominance of market fundamentals and TSM, measured by $\phi$, can be time-varying, depending on market condition. For simplicity we take $\phi$ as a constant parameter in this paper.

$m_t$ is defined by a standard moving average (MA) of past returns over $[t-\tau, t]$,

$$m_t = \frac{1}{\tau}\int_{t-\tau}^{t} \frac{dS_u}{S_u}, \tag{2.4}$$

where delay $\tau$ represents the time horizon. The way we model the momentum in this paper is motivated by the TSM strategy documented recently in Moskowitz et al. (2012), who demonstrate that the average return over a past period (say, 12 months) is a positive predictor of future returns, especially the return for the next month. The resulting asset price model (2.2)–(2.4) is characterized by a stochastic delay integro-differential system, which is non-Markovian and lacks analytical tractability. We show in Appendix A that the price process of (2.2)–(2.4) almost surely has a unique continuously adapted pathwise solution and the asset price stays positive for given positive initial values over $[-\tau, 0]$.

We now consider a typical long-term investor who maximizes the expected utility of terminal wealth at time $T(> t)$. Let $W_t$ be the wealth of the investor at time $t$ and $\pi_t$ be the fraction of the wealth invested in the stock. Then it follows from (2.2) that the change in wealth satisfies

$$\frac{dW_t}{W_t} = \big\{\pi_t[\phi m_t + (1-\phi)\mu_t - r] + r\big\}dt + \pi_t\sigma_S' dZ_t. \tag{2.5}$$

We assume that the preferences of the investor can be represented by a CRRA utility index with a constant coefficient of relative risk aversion equal to $\gamma$. The investment problem of the investor is then given by

$$J(W, m, \mu, t, T) = \sup_{(\pi_u)_{u\in[t,T]}} \mathbb{E}_t\Big[\frac{W_T^{1-\gamma} - 1}{1-\gamma}\Big], \tag{2.6}$$

where $J(W, m, \mu, t, T)$ is the value function corresponding to the optimal investment strategy. We apply the maximum principle for optimal control of stochastic delay differential equations and derive the optimal investment strategy in closed form. The result is presented in the following proposition and the proof can be found in Appendix B.

**Proposition 2.1.** *For an investor with an investment horizon $T-t$ and constant coefficient of relative risk aversion $\gamma$, the optimal wealth fraction invested in the risky asset is given by*

$$\pi_u^* = \frac{\phi m_u + (1-\phi)\mu_u - r}{\gamma\sigma_S'\sigma_S} + \frac{(z_u)_3\sigma_S}{p_u^3\sigma_S'\sigma_S}, \tag{2.7}$$

*where $z_u$ and $p_u$ are governed by a backward stochastic differential system* (B.6) *in Appendix B.2. Especially, when $\gamma = 1$, the preference is characterized by a log utility and the optimal allocation to stocks is given by*

$$\pi_t^* = \frac{\phi m_t + (1-\phi)\mu_t - r}{\sigma_S'\sigma_S}. \tag{2.8}$$

This proposition states that the optimal fraction (2.7) invested in the stock consists of two components. The first characterizes the myopic demand for the stock and the second is the intertemporal hedging demand (see, for instance, Merton 1971). When $\gamma = 1$, the optimal strategy (2.8) characterizes the myopic behavior of the investor with log utility. This result has a number of implications. Firstly, when the asset price follows a geometric Brownian motion process with mean-reversion drift $\mu_t$, namely $\phi = 0$, the optimal investment strategy (2.8) becomes

$$\pi_t^* = \frac{\mu_t - r}{\sigma_S' \sigma_S}. \tag{2.9}$$

This is the optimal investment strategy with mean-reverting returns obtained in the literature, say for example Campbell and Viceira (1999) and Wachter (2002). In particular, when $\mu_t = \bar{\mu}$ is a constant, the optimal portfolio (2.9) collapses to the optimal portfolio of Merton (1971).

Secondly, when the asset return depends only on the momentum, namely $\phi = 1$, the optimal portfolio (2.8) reduces to

$$\pi_t^* = \frac{m_t - r}{\sigma_S' \sigma_S}. \tag{2.10}$$

If we consider a trading strategy based on the trading signal indicated by the excess return $m_t - r$ only, with $\tau = 12$ months, the strategy of long/short when the trading signal is positive/negative is consistent with the TSM strategy used in Moskowitz et al. (2012). By constructing portfolios based on monthly excess returns over the past 12 months and holding for one month, Moskowitz et al. (2012) show that this strategy performs the best among all the momentum strategies with look-back and holding periods varying from one month to 48 months. Therefore, if we only take fixed long/short positions and construct simple buy-and-hold momentum strategies over a large range of look-back and holding periods, (2.10) shows that the TSM strategy of Moskowitz et al. (2012) can be optimal when mean reversion is not significant in financial markets. On the one hand, this provides a theoretical justification for the TSM strategy when market volatilities are constant and returns are not mean-reverting. On the other hand, note that the optimal portfolio (2.10) also depends on volatility. This explains the dependence of momentum profitability on market conditions and volatility found in empirical studies. In addition, the optimal portfolio (2.10) defines the optimal wealth fraction invested in the risky asset. Hence the TSM strategy of taking fixed positions based on the trading signal may not be optimal in general.

Thirdly, the optimal strategy (2.8) implies that a weighted average of momentum and mean-reverting strategies is optimal. Intuitively, it takes into account the short-run momentum and long-run reversal, both well-supported market phenomena. It also takes into account the timing opportunity with respect to market trend and volatility.

In summary, for the first time, we have provided a theoretical support for optimal strategies that combining of momentum and reversal documented in the empirical literature (see, for example, Balvers and Wu (2006) and Serban (2010)). In the rest of the paper, we first estimate the model to the S&P 500 and then evaluate and demonstrate empirically the performance of the optimal strategy comparing it to the market and other trading strategies recorded in the literature. In order to provide a better understanding of the performance, we start with the case $\gamma = 1$. The simple model and closed-form optimal strategy (2.8) facilitate model estimation and empirical analysis. We then numerically solve the optimal portfolio (2.7) and examine the values added by the hedging demand in Section 3.8.

## 3. Model Estimation and Performance Analysis

In this section we first estimate the model to the S&P 500. Based on these estimations, we then use utility of portfolio wealth and the Sharpe ratio to examine the performance of the optimal strategy (2.8), comparing to the performance of the market index and the optimal strategies based on pure momentum and pure mean-reversion models. To provide further evidence, we conduct out-of-sample tests on the performance of the optimal strategy and examine the effect of short sale constraints, market states, sentiment and volatility. In addition, we also compare the performance of the optimal strategy to that of the TSM strategy.

3.1. **Model Estimation.** In line with Campbell and Viceira (1999) and Koijen et al. (2009), the mean-reversion variable is affine in the (log) dividend yield,

$$\mu_t = \bar{\mu} + \nu(D_t - \mu_D) = \bar{\mu} + \nu X_t,$$

where $\nu$ is a constant, $D_t$ is the (log) dividend yield with $\mathbb{E}(D_t) = \mu_D$, and $X_t = D_t - \mu_D$ denotes the de-meaned dividend yield. Thus the asset price model (2.2)-(2.4) becomes

$$\begin{cases} \dfrac{dS_t}{S_t} = \big[\phi m_t + (1-\phi)(\bar{\mu} + \nu X_t)\big]dt + \sigma_S' dZ_t, \\ dX_t = -\alpha X_t dt + \sigma_X' dZ_t, \end{cases} \tag{3.1}$$

where $\sigma_X = \sigma_\mu/\nu$. The uncertainty in system (3.1) is driven by two independent Brownian motions. Without loss of generality, we follow Sangvinatsos and Wachter (2005) and assume the Cholesky decomposition on the volatility matrix $\Sigma$ of the dividend yield and return,

$$\Sigma = \begin{pmatrix} \sigma_S' \\ \sigma_X' \end{pmatrix} = \begin{pmatrix} \sigma_{S(1)} & 0 \\ \sigma_{X(1)} & \sigma_{X(2)} \end{pmatrix}.$$

Thus, the first element of $Z_t$ is the shock to the return and the second is the dividend yield shock that is orthogonal to the return shock.

To be consistent with the momentum and reversal literature, we discretize the continuous-time model (3.1) at a monthly frequency. This results in a bivariate Gaussian vector autoregressive (VAR) model on the simple return[8] and dividend yield $X_t$,

$$\begin{cases} R_{t+1} = \dfrac{\phi}{\tau}(R_t + R_{t-1} + \cdots + R_{t-\tau+1}) + (1-\phi)(\bar{\mu} + \nu X_t) + \sigma_S' \Delta Z_{t+1}, \\ X_{t+1} = (1-\alpha)X_t + \sigma_X' \Delta Z_{t+1}. \end{cases} \tag{3.2}$$

Note that both $R_t$ and $X_t$ are observable. We use monthly S&P 500 data over the period January 1871—December 2012 from the home page of Robert Shiller (www.econ.yale.edu/∼shiller/data.htm) and estimate model (3.2) using the maximum likelihood method. We set the instantaneous short rate $r = 4\%$ annually. As in Campbell and Shiller (1988*a*, 1988*b*), the dividend yield is defined as the log of the ratio between the last period dividend and the current index. The total return index is constructed by using the price index series and the dividend series.

The estimations are conducted separately for given time horizon $\tau$ varying from one to 60 months. Empirically, Moskowitz et al. (2012) show that the TSM strategy based on a 12-month horizon better predicts the next month's return than other time horizons. Therefore, in this section, we focus on the performance of the optimal strategy with a look-back period of $\tau = 12$ months and a one-month holding period. The effect of time horizon $\tau$ varying from one to 60 months is studied in the next section.

For comparison, we estimate the full model (FM) (3.2) with $0 < \phi < 1$, the pure momentum model (MM) with $\phi = 1$, and the pure mean-reversion model (MRM) with $\phi = 0$. For $\tau = 12$, Table 3.1 reports the estimated parameters, together with the 95% confidence bounds. For the pure momentum model ($\phi = 1$), there is only one parameter $\sigma_{S(1)}$ to be estimated. For the full model, as one of the key parameters, it shows that the momentum effect parameter $\phi \approx 0.2$, which is significantly different from zero. This implies that market index can be explained by about 20% of the momentum component and 80% of the mean-reverting component. Other parameter estimates in terms of the level and significance in Table 3.1 are consistent with those in Koijen et al. (2009).

We also conduct a log-likelihood ratio test to compare the full model ($0 < \phi < 1$) to the pure momentum model ($\phi = 1$) and pure mean-reversion model ($\phi = 0$). For the pure momentum model, the test statistic (13100) is much greater than 12.59, the critical value with six degrees of freedom at the 5% significance level. For the

[8]To be consistent with the momentum and reversal literature, we use simple return to construct $m_t$ and also discretize the stock price process into simple return rather than log return.

TABLE 3.1. Parameter estimations of the full model (FM), pure momentum model (MM) with $\tau = 12$, and pure mean-reversion model (MRM).

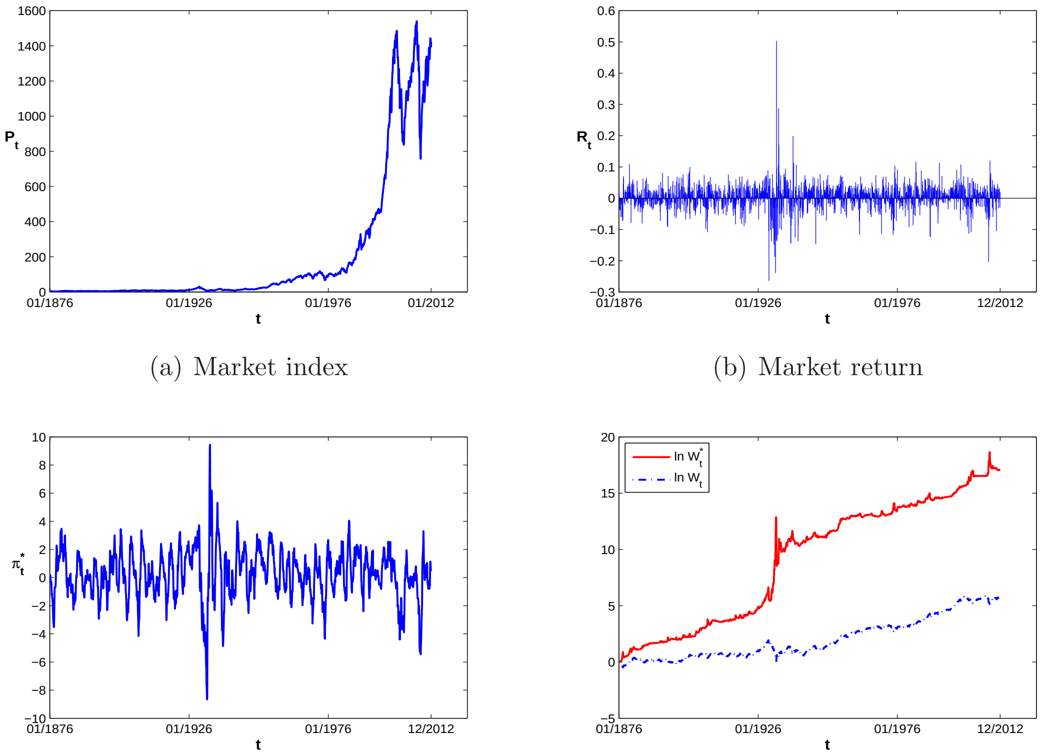
| Parameter | $\alpha$ | $\phi$ | $\bar{\mu}$ | $\nu$ |
|---|---|---|---|---|
| FM (%) | 0.46 | 19.85 | 0.36 | 0.20 |
| Bounds (%) | $(0.03, 0.95)$ | (8.70, 31.00) | (0.26, 0.46) | (-0.60, 1.00) |
| MM (%) | | | | |
| Bounds (%) | | | | |
| MRM (%) | 0.55 | | 0.37 | $2.67 * 10^{-5}$ |
| Bounds (%) | $(0.07, 1.03)$ | | (0.31, 0.43) | (-0.46, 0.46) |
| Parameters | $\sigma_{S(1)}$ | $\sigma_{X(1)}$ | $\sigma_{X(2)}$ | |
| FM (%) | 4.10 | -4.09 | 1.34 | |
| Bounds (%) | (3.95, 4.24) | (-4.24, -3.93) | (1.29, 1.39) | |
| MM (%) | 4.23 | | | |
| Bounds (%) | (4.09, 4.38) | | | |
| MRM (%) | 4.11 | -4.07 | 1.36 | |
| Bounds (%) | (3.97, 4.25) | (-4.22, -3.92) | (1.32, 1.40) | |

pure mean-reversion model, the test statistic (6200) is much greater than 3.841, the critical value with one degree of freedom at 5% significance level. Therefore the full model is significantly better than the pure momentum model and the pure mean-reversion model. This implies that the (full) model captures short-term momentum and long-term reversion in the market index and fits the data better than the pure momentum and pure mean-reverting models.

3.2. **Economic Value.** Based on the previous estimations, we examine the economic value of the optimal portfolio (2.8) based on log utility in terms of the utility of the portfolio wealth, comparing it to those of the market index and of the pure momentum and pure mean-reversion models. We evaluate the performance of a portfolio (or strategy) in terms of the Sharpe ratio in the next subsection.

We first compare the realized utility of the optimal portfolio wealth invested in the S&P 500 index based on the optimal strategy (2.8) with a look-back period $\tau = 12$ months and one-month holding period to the utility of a passive holding investment in the S&P 500 index with an initial wealth of \$1. As a benchmark, the log utility of an investment of \$1 to the index from[9] January 1876 to December 2012 is equal to 5.765. For $\tau = 12$, we calculate the moving average $m_t$ of past 12-month returns at any point of time based on the market index from January 1876 to December

[9]Considering the robustness analysis for $\tau$ varying from one to 60 months in the next section, all the portfolios start at the end of January 1876 (60 months after January 1871).




(a) Market index

(b) Market return

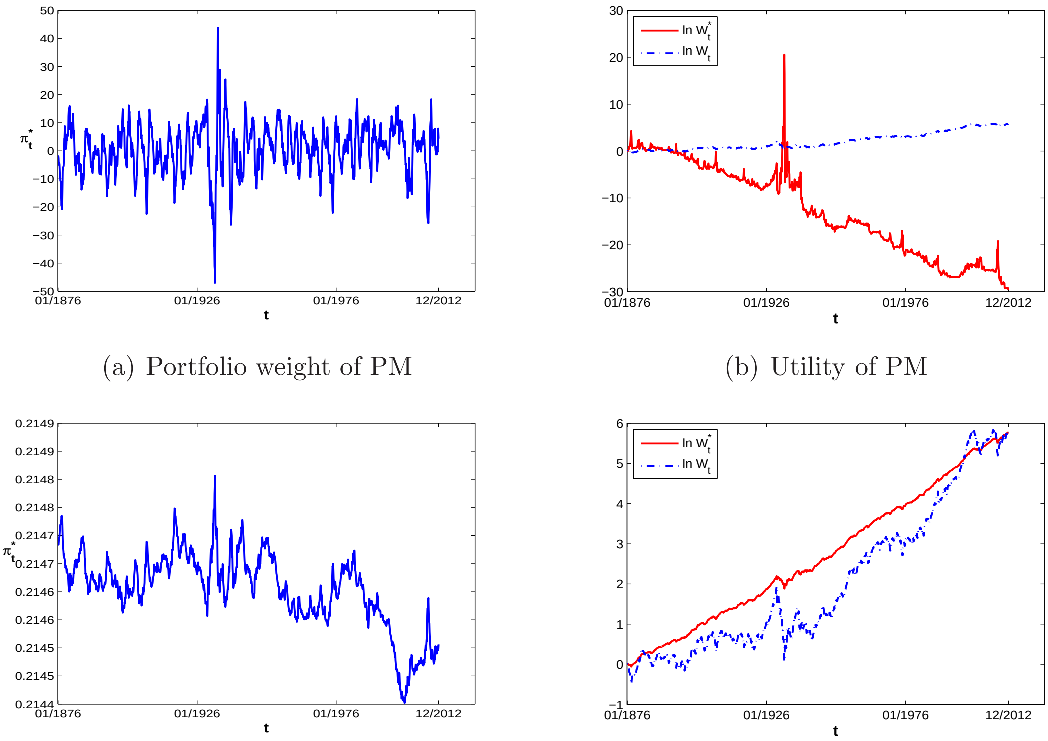
(c) Portfolio weight

(d) Utility

FIGURE 3.1. The time series of market index (a), the simple return of the S&P 500 (b); and the time series of the optimal portfolio (c) and the utility (d) of the optimal portfolio wealth from January 1876 until December 2012 for $\tau = 12$. In (d), the utilities of the optimal portfolio wealth and the market index are plotted in solid red and dash-dotted blue lines respectively.

2012. With an initial wealth of \$1 at January 1876 and the estimated parameters in Table 3.1, we calculate the monthly investment of the optimal portfolio wealth $W_t$ based on (2.8) and record the realized utilities of the optimal portfolio wealth from January 1876 to December 2012. Based on the calculation, we plot the index level and simple return of the S&P 500 index from January 1876 until December 2012 in Fig. 3.1 (a) and (b). Fig. 3.1 (c) reports the optimal wealth fractions $\pi_t$ of (2.8) and Fig. 3.1 (d) reports the evolution of the utilities of the optimal portfolio wealth over the same time period, showing that the optimal portfolios outperform the market index measured by the utility of wealth.

There are two interesting observations from Fig. 3.1. Firstly, the returns of optimal strategies and index are positively correlated (with a correlation of 0.335). Secondly, Fig. 3.1 (d) seems to indicate a big jump in the utilities of the optimal




(a) Portfolio weight of PM

(b) Utility of PM

(c) Portfolio weight of MRM

(d) Utility of MRM

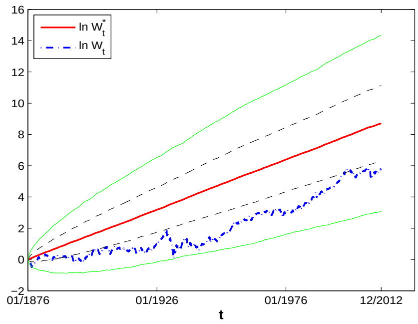
Figure 3.2. The time series of the optimal portfolio weight and the utility of the wealth for the pure momentum model with $\tau = 12$ (a) and (b) and the pure mean-reversion model (c) and (d) from January 1876 until December 2012. In (d), the utilities of the optimal portfolio wealth and the market index are plotted in solid red and dash-dotted blue lines respectively.

portfolio during the period of the Great Depression in the 1930s. This observation is consistent with Moskowitz et al. (2012), who find that the TSM strategy delivers its highest profits during the most extreme market episodes. However, the performance of the optimal portfolio is not completely driven by its performance during crisis periods.[10]

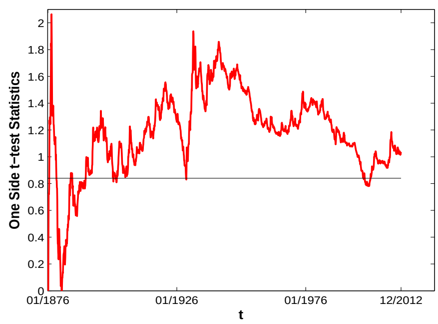
Next, we compare the economic value of the pure momentum and pure mean-reverting strategies to that of the market index. For the pure momentum model,

[10]To clarify this observation, we also examine performance using data from January 1940 to December 2012 to avoid the Great Depression periods. We re-estimate the model, conduct the same analysis. Our results show that the optimal strategies still outperform the market index over this time period. This indicates that the outperformance of the optimal strategy is not necessarily due to extreme market episodes, such as the Great Depression. Later in this section, we show that the outperformance is in fact immune to market conditions.

based on the estimated parameters in Table 3.1, Fig. 3.2 (a) and (b) illustrate the time series of the portfolio weights and the utilities of the optimal portfolio for the pure momentum model from January 1876 to December 2012. Compared to the full model illustrated in Fig. 3.1, the leverage of the pure momentum strategies is much higher, as indicated by the higher level of $\pi_t^*$. The optimal strategies for the pure momentum model suffer from high risk and perform worse than the market and hence the optimal strategies of the full model. Similarly, based on the estimates in Table 3.1, Fig. 3.2 (c) and (d) illustrate the time series of the portfolio weight and the utility of the wealth of the optimal portfolio for the pure mean-reversion model, showing that the performance of the strategy is about the same as the stock index but worse than the optimal strategies (2.8). Note that in this case there is not much variation in the portfolio weight and the optimal portfolio does not capture the timing opportunity of the market trend and market volatility. Therefore, both the pure momentum and pure mean-reversion strategies underperform the market and the optimal strategies of the full model.




(a) Average utility




(b) One-sided $t$ statistics

FIGURE 3.3. (a) Average utility ((the solid red line), the 95% confidence bounds (the solid green lines) and the 60% confidence bounds (the dotted blue lines) and (b) one-sided $t$-test statistics based on 1,000 simulations for $\tau = 12$.

To provide further evidence for the economic value of the optimal strategy, we conduct a Monte Carlo analysis. For $\tau = 12$ and the estimated parameters, we simulate model (3.1) and report the average portfolio utilities (the solid red line in the middle) based on 1,000 simulations in Fig. 3.3 (a), together with 95% confidence levels (the two solid green lines outside), comparing to the utility of the market index (the dotted blue line). It shows that firstly, the average utilities of the optimal portfolios are better than that of the S&P 500. Secondly, the utility for the S&P 500 falls into the 95% confidence bounds and hence the average performance of the optimal strategy is not statistically different from the market index at the 95%

confidence level. We also plot two black dashed bounds for the 60% confidence level. It shows that, at the 60% confidence level, the optimal portfolio significantly outperforms the market index. Fig. 3.3 (b) reports the one-sided $t$-test statistics to test $\ln W_t^* > \ln W_t^{SP500}$. The $t$-statistics are above 0.84 most of the time, which indicates a critical value at 80% confidence level. Therefore, with 80% confidence, the optimal portfolio significantly outperforms the market index. In summary, we have provided empirical evidence of the outperformance of the optimal strategy (2.8) compared to the market index, pure momentum and pure mean-reversion strategies.

3.3. **The Sharpe Ratio.** We now use the Sharpe ratio to examine the performance of the optimal strategy. The Sharpe ratio is defined as the ratio of the mean excess return on a portfolio and the standard deviation of the portfolio return. When the Sharpe ratio of an active strategy exceeds the market Sharpe ratio, we say that the active portfolio outperforms or dominates the market portfolio (in an unconditional mean-variance sense). For empirical applications, the (ex-post) Sharpe ratio is usually estimated as the ratio of the sample mean of the excess return on the portfolio and the sample standard deviation of the portfolio return (Marquering and Verbeek 2004). The average monthly return on the total return index of the S&P 500 over the period January 1871–December 2012 is 0.42% with an estimated (unconditional) standard deviation of 4.11%. The Sharpe ratio of the market index is 2.1%. For the optimal strategy (2.8), the return of the optimal portfolio wealth at time $t$ is given by

$$R_t^* = (W_t^* - W_{t-1}^*)/W_{t-1}^* = \pi_{t-1}^* R_t + (1 - \pi_{t-1}^*) r. \tag{3.3}$$

Table 3.2 reports the Sharpe ratios of the passive holding market index portfolio and the optimal portfolios from January 1886 to December 2012 for $\tau = 12$ together with their 90% confidence intervals (see Jobson and Korkie 1981). It shows that, by taking the timing opportunity (with respect to the market trend and market volatility), the optimal portfolio outperforms the market. We also conduct a Monte Carlo analysis based on 1,000 simulations and obtain an average Sharpe ratio of 6.12% for the optimal portfolio. The result is consistent with the outperformance of the optimal portfolio measured by portfolio utility (with an average terminal utility of 8.71 for the optimal portfolio).

Table 3.2. The Sharpe ratios of the optimal portfolio and the market index with corresponding 90% confidence interval and the Sharpe ratio of the optimal portfolio based on Monte Carlo simulations.

| | Optimal portfolio | Market index | Monte Carlo |
|---|---|---|---|
| Sharpe ratio (%) | 5.85 | 2.11 | 6.12 |
| Bounds (%) | (1.86, 9.84) | (-1.88, 6.10) | (5.98, 6.27) |

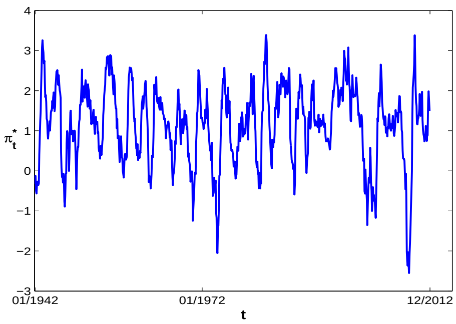
In summary, we have used two performance measures and provided empirical evidence of the outperformance of the optimal strategy (2.8) compared to the market index, pure momentum and pure mean-reversion strategies. The results provide empirical support for the analytical result on the optimal strategy derived in Section 2. In the following subsection, we conduct further empirical tests on these results. We first conduct out-of-sample tests on the predication power of the model and then examine the performance of the optimal strategy with short-sale constraints, market states, sentiment and volatility.

3.4. **Out-of-Sample Tests.** We implement a number of out-of-sample tests for the optimal strategies by splitting the whole data set into two sub-sample periods and using the first sample period to estimate the model. We then apply the estimated parameters to the second portion of the data to examine the out-of-sample performance of the optimal strategies.

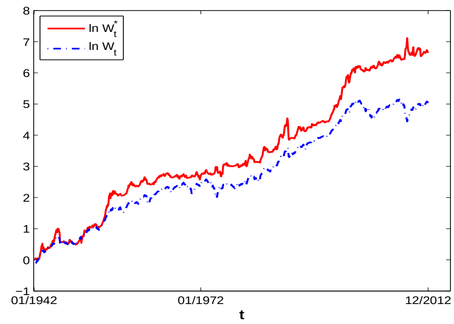
In the first test, we split the whole data set into two equal periods: January 1871 to December 1941 and January 1942 to December 2012. Notice the data in the two periods are quite different; the market index increases gradually in the first period but fluctuates widely in the second period as illustrated in Fig. 3.1 (a). With $\tau = 12$, Fig. 3.4 (a) and (b) illustrates the corresponding time series of the optimal portfolio and the utility of the optimal portfolio wealth from January 1942 to December 2012, showing that the utility of the optimal strategy grows gradually and outperforms the market index.

Many studies (see, for example, Jegadeesh and Titman 2011) show that momentum strategies perform poorly after the subprime crisis in 2008. In the second test, we use the subprime crisis to split the whole sample period into two periods and focus on the performance of the optimal strategies after the subprime crisis. The results are reported in Fig. 3.4 (c) and (d). It is clear that the optimal strategy still outperforms the market over the sub-sample period, in particular, during the financial crisis period around 2009 by taking large short positions in the optimal portfolios. We also use data from the last 10 years and 20 years as the out-of-sample test and find the results are robust.

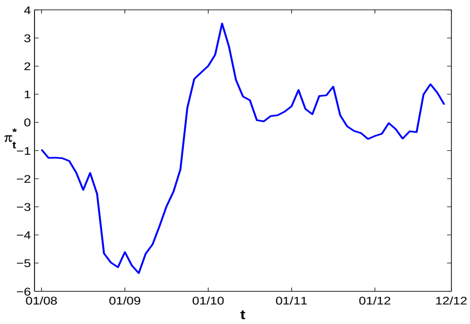
As the third test, we implement the rolling window estimation procedure to avoid look-ahead bias. For $\tau = 12$, we estimate parameters at each month by using the past 20 years' data and report the results in Fig. C.1 in Appendix C. We then report the time series of the index level (a), the simple return of the total return index of S&P 500 (b), the optimal portfolio (c), and the utility of the optimal portfolio wealth (d) in Fig. C.2 of Appendix C, showing a strong performance of the optimal portfolios over the market.



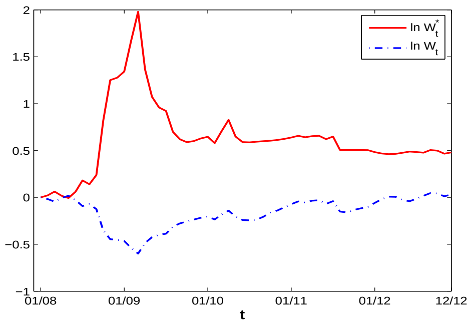

(a) The portfolio weight: 1/1942–12/2012




(b) The utility: 1/1942–12/2012




(c) The portfolio weight: 1/2008–12/2012




(d) The utility: 1/2008–12/2012

FIGURE 3.4. The time series of out-of-sample optimal portfolio weights and utility of the optimal portfolio wealth (the solid lines) from January 1942 until December 2012 in (a) and (b) and from January 2008 to December 2012 in (c) and (d) with $\tau = 12$ compared to the utility of the market index (the dotted line).

We also implement the out-of-sample tests for the pure momentum and pure mean-reversion models (not reported here) and find that they cannot outperform the market in most out-of-sample tests (last 10, 20 and 71 years), but do outperform the market for out-of-sample tests over the last five years. We also report the results of out-of-sample tests of the pure momentum in Fig. C.4 and the pure mean reversion in Fig. C.6 based on the 20-year rolling window estimates in Fig. C.3 and Fig. C.5 respectively, in Appendix C. We also implement the estimations for different window sizes of 25, 30 and 50 years (not reported here) and find that the estimated parameters are not very sensitive to the size of rolling window and the performance of strategies is similar to the case of 20-year rolling window estimation. Overall, the out-of-sample tests demonstrate the robustness of the outperformance of the optimal trading strategies compared to the market index, pure momentum and pure mean-reversion strategies.

3.5. **Short-sale Constraints.** Investors often face short-sale constraints. To evaluate optimal strategies under such constraints, we consider them when short selling and borrowing (at the risk-free rate) are not allowed. The portfolio weight $\pi$ in this case must lie between zero and 1. Since the value function is concave with respect to $\pi$, the optimal strategy becomes

$$\Pi_t^* = \begin{cases} 0, & \text{if} \quad \pi_t^* < 0, \\ \pi_t^*, & \text{if} \quad 0 \le \pi_t^* \le 1, \\ 1, & \text{if} \quad \pi_t^* > 1. \end{cases} \tag{3.4}$$

TABLE 3.3. The terminal utility of the portfolio wealth, Sharpe ratio, mean and stand deviation of the portfolio weights for the optimal portfolios with and without short-sale constraints, comparing with the market index portfolio.

| | Utility | Sharpe Ratio | Average weights | Std of weights |
|---|---|---|---|---|
| With constraints | 10.35 | 0.12 | 0.43 | 0.46 |
| Without constraints | 17.06 | 0.06 | 0.23 | 1.74 |
| Market index | 5.76 | 0.02 | | |

Table 3.3 reports the terminal utilities and the Sharpe ratio of the optimal portfolio with and without short-sale constraints, compared to the passive holding market index portfolio. The results show that the optimal portfolio with short-sale constraints outperforms the market, it even outperforms the optimal portfolio without short-sale constraints under the Sharpe ratio. It seems that the constraints improves portfolio performance. This observation is consistent with Marquering and Verbeek (2004 p. 419) who argue that "*While it may seem counterintuitive that strategies perform better after restrictions are imposed, it should be stressed that the unrestricted strategies are substantially more affected by estimation error.*" Indeed, we see from Table 3.3 that the estimated optimal portfolio weight without constraints has bigger standard error than that with constraints.

3.6. **Market States, Sentiment and Volatility.** The cross-sectional momentum literature has shown that momentum profitability can be affected by market states, investor sentiment and market volatility. For example, Cooper, Gutierrez and Hameed (2004) find that short-run (six months) momentum strategies make profits in an up market and lose in a down market, but the up-market momentum profits reverse in the long run (13–60 months). Hou, Peng and Xiong (2009) find momentum strategies with a short time horizon (one year) are not profitable in a down market, but are profitable in an up market. Similar profitability results are also reported in Chordia and Shivakumar (2002), specifically that common macroeconomic variables related to the business cycle can explain positive returns to momentum strategies during expansionary periods and negative returns during recessions.

To investigate the performance of optimal strategies under different market states,[11], we see from Table D.1 in Appendix D that the unconditional average excess return is 87 basis points per month. In up months, the average excess return is 81 basis points and it is statistically significant. In down months, the average excess return is 101 basis points; this value is economically significant although it is not statistically significant. The difference between down and up months is 20 basis points, which is not significantly different from zero, based on a two-sample $t$-test ($p$-value of 0.87).

Controlling for market risk, we use an up-month dummy[12] to capture incremental average return in up market months relative to down market months. We report the regression results in Table D.2 in Appendix D for the optimal strategy, the pure momentum strategy, pure mean-reversion strategy and the TSM strategy in Moskowitz et al. (2012) for $\tau = 12$ respectively. Except for the TSM, which earns significant positive returns in down markets, both down market returns $\alpha$ and the incremental returns in up market $\kappa$ are insignificant for all other strategies; these results are consistent with those in Table D.1. We also control for market risk in up and down months separately and obtain similar results in Table D.2.

Other way to see effects of market state on portfolio returns is to look at its predictive powers. Table D.3 reports predictive regression results of excess portfolio returns on the up-month dummy. We see that up market has no additional predictive power to portfolio returns over down market (insignificant $\kappa$), down market has significant predictive power to TSM returns. Down market has insignificant predictive power to the full model, pure momentum, and pure mean reversion, but among them, the effect is relatively strong in the full model, and weak in the pure mean reversion. We obtain similar results for the CAPM-adjusted return.

In terms of the effects of investor sentiment and market volatility on portfolio performance, Baker and Wurgler (2006, 2007) find that investor sentiment affects cross-sectional stock returns and the aggregate stock market. Wang and Xu (2012) find that market volatility has significant power to forecast momentum profitability. For TSMs, however, Moskowitz et al. (2012) find that there is no significant relationship of TSM profitability to either market volatility or investor sentiment. We

[11]We follow Cooper et al. (2004) and Hou et al. (2009) and define market state using the cumulative return of the stock index (including dividends) over the most recent 36 months. We label a month as an up (down) market month if the three-year return of the market is non-negative (negative). We compute the average return of the optimal strategy, compare the average returns between up and down market months, and report the results in Appendix D.

[12]The results are robust when we replace the up-month dummy with the lagged market return over the previous 36 months (not reported here).

find that both investor sentiment and market volatility have no predictive power on portfolio returns (see Table D.4 and D.5 in Appendix D).

Overall, we find that returns of the optimal strategies are not significantly different in up and down market states. We also find that both investor sentiment and market volatility have no predictive power for the returns of the optimal strategies. In fact, the optimal strategies have taken these factors into account and hence the returns of the optimal strategies have no significant relationship with these factors. Therefore, the optimal strategies are immune to market states, investor sentiment and market volatility.

3.7. **Comparison with TSM.** We now compare the performance of the optimal strategy to the TSM strategy of Moskowitz et al. (2012). The momentum strategies in the empirical studies are based on trading signals only. We first verify the profitability of the TSM strategies and then examine the excess return of buy-and-hold strategies when the position is determined by the sign of the optimal portfolio strategies (2.8) with different combinations of time horizons $\tau$ and holding periods $h$.

For a given look-back period $\tau$, we take long/short positions based on the sign of the optimal portfolio (2.8). Then for a given holding period $h$, we calculate the monthly excess return of the strategy $(\tau, h)$. Table E.1 in Appendix E reports the average monthly excess return (%) of the optimal strategies, skipping one month between the portfolio formation period and holding period to avoid the one-month reversal in stock returns, for different look-back periods (in the first column) and different holding periods (in the first row). The average return is calculated in the same way as in Moskowitz et al. (2012). We calculate the excess returns of the optimal strategies over the period from January 1881 (10 years after January 1871 with five years for calculating the trading signals and five years for holding periods) to December 2012.

For comparison, Table E.2 in Appendix E reports the average returns (%) for the pure momentum model.[13] Notice that Tables E.1 and E.2 indicate that strategy $(9, 1)$ performs the best. This is consistent with the finding in Moskowitz et al. (2012) that strategy $(9, 1)$ is the best strategy for equity markets although the 12-month horizon is the best for most asset classes.

Next we use the Sharpe ratio to examine the performance of the optimal strategy $\pi_t^*$ of (2.8) and compare it to the passive index strategy and two TSM strategies: one follows from Moskowitz et al. (2012) and the other is the TSM strategy based on the sign of the optimal strategies $sign(\pi_t^*)$ as the trading signal (instead of the average

[13]Notice the position is completely determined by the sign of the optimal strategies. Therefore, the position used in Table E.2 is the same as that of the TSM strategies in Moskowitz et al. (2012).

Table 3.4. The Sharpe ratio of the optimal portfolio, market index, TSM and MMR for $\tau = 12$ with corresponding 90% confidence interval.

| | Optimal portfolio | Market index | TSM | MMR |
|---|---|---|---|---|
| Sharpe ratio (%) | 5.85 | 2.11 | -0.03 | 4.16 |
| Bounds (%) | (1.86, 9.84) | (-1.88, 6.10) | (-4.01, 3.96) | (0.18, 8.15) |

excess return over a past period), which is called *momentum and mean-reversion (MMR) strategy* for convenience. For a time horizon of $\tau = 12$ months, we report the Sharpe ratios of the portfolios for the four strategies in Table 3.4 from January 1881 to December 2012. It shows that the TSM strategy underperforms the market while the MMR strategy outperforms it. The optimal strategy also significantly outperforms all the momentum, mean-reversion and TSM strategies. Note that the only difference between the optimal strategy and the MMR strategy is that the former considers the size of the portfolio position, which is inversely proportional to the variance, while the latter always takes one unit of long/short position. This implies that, in addition to trends, the size of the position is another very important factor for investment profitability.

Following Moskowitz et al. (2012), we examine the cumulative excess return. That is, the return at time $t$ is defined by

$$\hat{R}_{t+1} = sign(\pi_t^*)\frac{0.1424}{\hat{\sigma}_{S,t}}R_{t+1}, \tag{3.5}$$

where 0.1424 is the sample standard deviation of the total return index and the ex-ante annualized variance $\hat{\sigma}^2_{S,t}$ for the total return index is calculated as the exponentially weighted lagged squared month returns,

$$\hat{\sigma}^2_{S,t} = 12\sum_{i=0}^{\infty}(1-\delta)\delta^i(R_{t-1-i} - \bar{R}_t)^2, \tag{3.6}$$

here the constant 12 scales the variance to be annual, and $\bar{R}_t$ is the exponentially weighted average return based on the weights $(1-\delta)\delta^i$. The parameter $\delta$ is chosen so that the center of mass of the weights is $\sum_{i=1}^{\infty}(1-\delta)\delta^i = \delta/(1-\delta) =$ two months. To avoid look-ahead bias contaminating the results, we use the volatility estimates at time $t$ for time $t+1$ returns throughout the analysis.

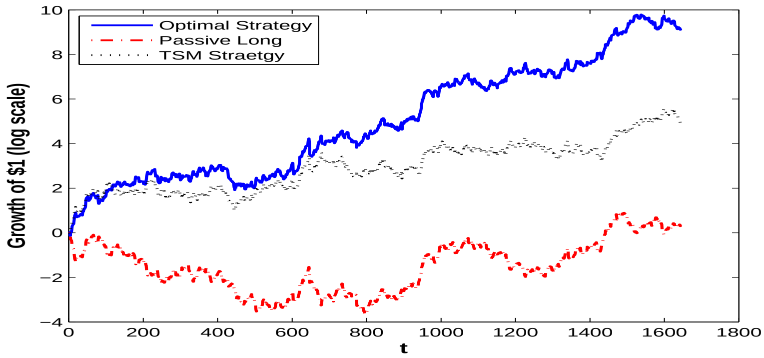



FIGURE 3.5. Log cumulative excess return of the optimal strategy and momentum strategy with $\tau = 12$ and passive long strategy from January 1876 to December 2012.

With a 12-month time horizon Fig. 3.5 illustrates the log cumulative excess return of the optimal strategy (2.8), the momentum strategy and the passive long strategy from January 1876 to December 2012. It shows that the optimal strategy has the highest growth rate and the passive long strategy has the lowest growth rate. The pattern of Fig. 3 in Moskowitz et al. (2012 p.239) is replicated in Fig. 3.5, showing that the TSM strategy outperforms the passive long strategy.[14] In summary, we have shown that the optimal strategy outperforms the TSM strategy of Moskowitz et al. (2012). By comparing the performance of two TSM strategies, we find that the TSM strategy based on momentum and reversal trading signal is more profitable than the pure TSM strategy of Moskowitz et al. (2012).[15]

3.8. **Discussions on Hedging Demand.** Taking the advantage of the closed-form solution, previous sections concentrate on the case of $\gamma = 1$. Ideally, we should examine the general case of $\gamma$ where the optimal portfolio weight is the sum of myopic and hedging demands for the stock. However, the optimal strategy (2.7) is determined by a coupled forward backward stochastic differential equations (FBSDEs), up till now, there is no efficient way to numerically solving FBSDEs with time delay (Ma and Yong, 1999 and Delong, 2013). Given current state of the art of FBSDEs

[14] In fact, the profits of the diversified time series momentum (TSMOM) portfolio in Moskowitz et al. (2012) are to some extent driven by the bonds when scaling for the volatility in equation (5) of their paper, and hence applying the TSM strategies to the stock index may have fewer significant profits than the diversified TSMOM portfolio.

[15]This paper studies the S&P 500 index over 140 years of data, while Moskowitz et al. (2012) focus on the futures and forward contracts that include equity indices, currencies, commodities, and sovereign bonds. Despite a large difference between the data investigated, we find similar patterns for the TSM in the stock index and replicate their results with respect to the stock index.

with time delay, we only be able to do some limited exploratory analysis on the hedging demand.

We follow the scheme developed in Bender and Denk (2007) which is based on Picard iterations. Due to the non-Markovian structure of time-delayed BSDEs, the conditional expectation in (B.6) in Appendix B has to be taken with respect to the whole information $\mathcal{F}_t$. Therefore, we estimate the expected values by approximating the Brownian motion by a symmetric random walk as in Ma, Protter, Martin and Torres (2002). Specifically, we first simulate the $2^T$ trajectories of the forward processes $S_t$ and $\mu_t$ for $t$ from 1 to $T$ based on the approximating binomial random walk. The parameters are chosen based on Table 3.1 and the initial values $S_t = \varphi_t$, $t \in [-\tau, 0]$ and $\mu_0 = \hat{\mu}$ are chosen as the corresponding initial values of S&P 500 and the dividend yield. A unique solution $(p, z)$ to the backward part (B.6) is obtained as the limit of the sequence of the processes $(p^{(n)}, z^{(n)})$ governed by

$$\begin{aligned} p_t^{(n)} =& \mathbb{E}\Big[\Phi_X(X_T) + \int_t^T \{(b_X^{\pi^{(n-1)}})^\top p_u^{(n-1)} + (\sigma_X^{\pi^{(n-1)}})^\top z_u^{(n-1)} + \\ & (b_{X_\tau}^{\pi^{(n-1)}}|_{u+\tau})^\top p_{u+\tau}^{(n-1)} + (\sigma_{X_\tau}^{\pi^{(n-1)}}|_{u+\tau})^\top z_{u+\tau}^{(n-1)}\} du \Big| \mathcal{F}_t\Big], \end{aligned}$$

with $(p^{(0)}, z^{(0)}) = (1, 0)$. For the $n$-th iteration, $\pi^{(n-1)}$ and hence $W^{(n-1)}$ can be obtained after knowing $(p^{(n-1)}, z^{(n-1)})$. This algorithm would be feasible for small terminal time $T$ but impractical for longer durations due to an enormous number of trajectories that has to be generated. We consider terminal time $T$ up to 12 months and choose the relative risk aversion $\gamma = 5$. To examine the values added by the hedging demand, we compare the optimal strategy and the myopic strategy. For both the optimal and suboptimal investment strategies, we determine the certainty equivalent return and report the annualized loss in certainty equivalent wealth by following the suboptimal strategic allocation. Specifically, the annualized utility costs are given by

$$C = \left[\frac{J_2 + 1/(1-\gamma)}{J_1 + 1/(1-\gamma)}\right]^{1/[T(1-\gamma)]} - 1, \tag{3.7}$$

where $J_1$ and $J_2$ are the value functions resulting from following the optimal and myopic strategies.

Table 3.5. The utility costs (in %) of behaving myopically for terminal times up to one year and time horizons up to six months. Here $\gamma = 5$.

| $T$ | 1 | 3 | 6 | 9 | 12 |
|---|---|---|---|---|---|
| $\tau = 1$ | -0.26 | -0.97 | -1.83 | -1.86 | -3.01 |
| $\tau = 3$ | -0.64 | -1.01 | -1.64 | -5.30 | -8.17 |
| $\tau = 6$ | -12.99 | -12.77 | -12.81 | -16.34 | -24.39 |

Table 3.5 reports the utility costs for terminal times up to one year and time horizons up to six months. Two observations follow Table 3.5. Firstly, intuitively, myopic strategy suffers a big loss for large investment horizons. Table 3.5 confirms this intuition and shows that the costs of myopic strategy increase as terminal time increases. Secondly, the time delay effect in stock returns also enlarges the costs of myopic strategy. We complete this section with the following remark. Notice the larger $\tau$ is, the more values of the backward processes are 0 across different market states for $t \in [T, T+\tau]$ and hence the less impact of time delay involved from the conditional expectations in the BSDEs. This effect can be observed from Table 3.5, especially the little difference when $T < \tau$ for $\tau = 6$.

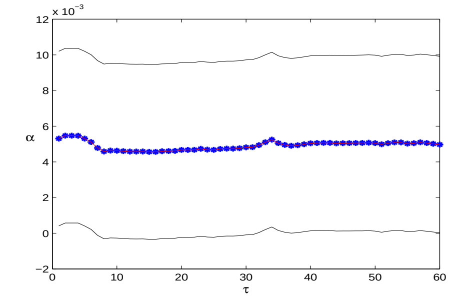
## 4. Time Horizon Effects

The impact of time horizon on investment profitability has been extensively investigated in the empirical literature, for example, De Bondt and Thaler (1985) and Jegadeesh and Titman (1993). Due to the closed-form optimal strategy (2.8), we are able to explicitly examine the dependence of the optimal results on different time horizons. Rather than focusing only on $\tau = 12$ months, in this section we examine the effect of a time horizon $\tau$ varying from one to 60 months on the outperformance of the optimal strategies.

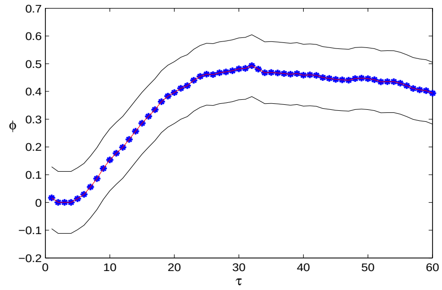
4.1. **Model Estimations and Comparison.** For a given time horizon $\tau$ we estimate the model 3.2. Fig. 4.1 reports the estimated parameters in monthly terms for $\tau$ ranging from one month to five years, together with the 95% confidence bounds. Fig. 4.1 (b) shows that the momentum effect parameter $\phi$ is significantly different from zero when time horizon $\tau$ is more than half a year, indicating a significant momentum effect for $\tau$ beyond six months.[16] Note that $\phi$ increases to about 50% when $\tau$ increases from six months to three years and then decreases gradually when $\tau$ increases further. This implies that market returns can be explained by both the momentum (based on different time horizons) and mean-reverting components. Other results in terms of the level and significance reported in Fig. 4.1 are consistent with Koijen et al. (2009).

Obviously, the estimations depend on the specification of the time horizon $\tau$. To explore the optimal value for $\tau$, we compare different information criteria, including Akaike (AIC), Bayesian (BIC) and Hannan–Quinn (HQ) information criteria for $\tau$ from one month to 60 months in Fig. F.1 of Appendix F. The results imply that the average returns over the past 18 months to two years can best predict future

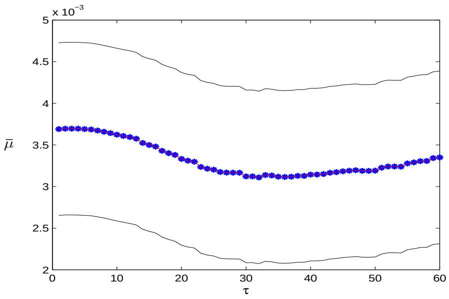
[16] For $\tau$ from one to five months, $\phi$ is indifferent from zero statistically and economically. Correspondingly, for small look-back periods of up to half a year, the model is equivalent to a pure mean-reversion model. This observation is helpful when explaining the results of the model for small look-back periods in the following discussion.



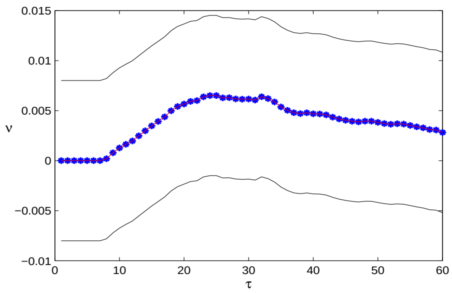
(a) Estimates of $\alpha$



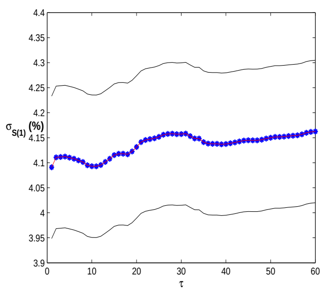
(b) Estimates of $\phi$



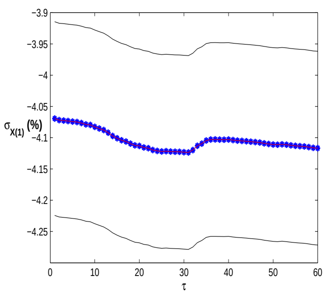
(c) Estimates of $\bar{\mu}$



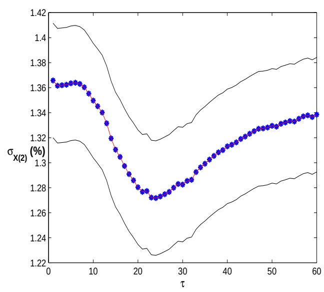
(d) Estimates of $\nu$



(e) Estimates of $\sigma_{S(1)}$



(f) Estimates of $\sigma_{X(1)}$

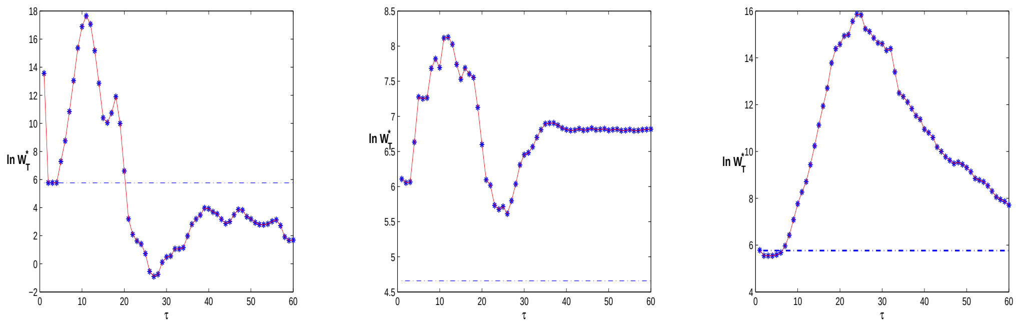


(g) Estimates of $\sigma_{X(2)}$

FIGURE 4.1. The estimates of (a) $\alpha$; (b) $\phi$; (c) $\bar{\mu}$; (d) $\nu$; (e) $\sigma_{S(1)}$; (f) $\sigma_{X(1)}$ and (g) $\sigma_{X(2)}$ as functions of $\tau$.

returns and the explanatory power for the market returns is reduced for longer time horizons. This is consistent with studies showing that short-term (one to two years), rather than long-term, momentum better explains market returns. Combining the results in Figs 4.1 (b) and F.1, we can conclude that market returns are better captured by short-term momentum and long-term reversion.

We also compare the performance of the optimal strategies with the pure momentum strategies ($\phi = 1$) for different $\tau$. Fig. F.2 (a) reports the estimates of $\sigma_{S(1)}$ and the 95% confidence bounds for $\tau \in [1, 60]$. It shows that as $\tau$ increases, the volatility of the index decreases dramatically for small time horizons and is then stabilized for large time horizons. It implies high volatility associated with momentum over short time horizons and low volatility over long time horizons. We also compare the information criteria for different $\tau$ (not reported here) and find that all the AIC, BIC and HQ reach their minima at $\tau = 11$. This implies that the average returns over the previous 11 months can predict future returns best for the pure momentum model. This is consistent with the finding of Moskowitz et al. (2012) that momentum returns over the previous 12 months better predict the next month's return than other time horizons. In addition, we conduct the log-likelihood ratio test to compare the full model to the pure momentum model ($\phi = 1$) and to the pure mean-reversion model for different $\tau$. We report the log-likelihood ratio test results in Fig. F.2 (b), which show that the full model is significantly better than the pure momentum model and pure mean-reversion model for all $\tau$.




(a) (b) (c)

FIGURE 4.2. The terminal utility of the optimal portfolio wealth (a) from January 1876 to December 2012, (b) from January 1945 to December 2012, and (c) the average terminal utility of the optimal portfolios based on 1000 simulations from January 1876 to December 2012, comparing with the terminal utility of the market index portfolio (the dash-dotted line).

4.2. **The effect of time horizon on performance.** For $\tau = 1, 2, \cdots, 60$, we estimate the full model over the full sample. Fig. 4.2 (a) reports the utility of terminal wealth, compared to the utility of the market portfolio at December 2012. It shows that the optimal strategies consistently outperform the market index for $\tau$ from five to 20 months.[17] The corresponding utilities are plotted in Fig. F.4.

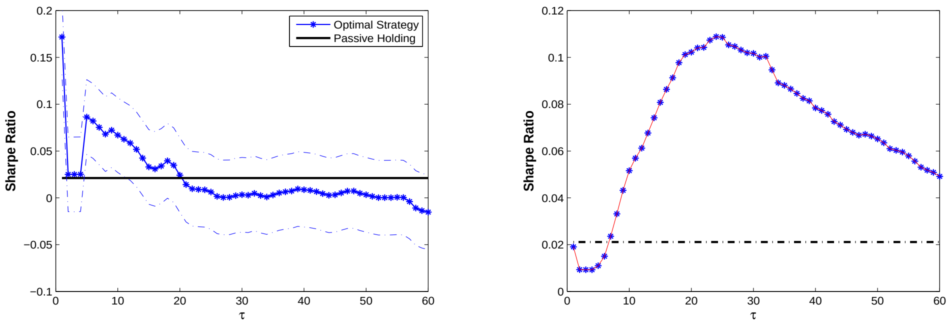
[17]When $\tau$ is less than half a year, Fig. 4.2 (a) shows that the optimal strategies do not perform significantly better than the market. As we indicate in footnote 16, the model with a small look-back period of up to half a year performs similarly to the pure mean-reversion strategy. Note the significant outperformance of the optimal strategy with a one-month horizon in Fig. 4.2 (a). This is due to the fact that the first order autocorrelation of the return of the S&P 500 is significantly

We have observed from Fig. 3.1 (d) for $\tau = 12$ that the Great Depression in the 1930s has greatly improved the utilities of the optimal portfolio . To clarify this observation, we also examine performance using the data from January 1940 to December 2012 to avoid the Great Depression period. We re-estimate the model, conduct the same analysis, and report the terminal utilities of the optimal portfolios in Fig. 4.2 (b) over this time period. It shows that the optimal strategies still outperform the market and the performance of the optimal strategies over the more recent time period becomes even better for all time horizons. Consistent with the results obtained in the previous section, the outperformance of the optimal strategy is not necessarily due to extreme market episodes, such as the Great Depression.

We also conduct further Monte Carlo analysis on the performance of the optimal portfolios based on the estimated parameters in Fig. 4.1 and 1,000 simulations and report the average terminal utilities in Fig. 4.2 (c). The result displays a different terminal performance from that in Fig. 4.2 (a). In fact, the terminal utility in Fig. 4.2 (a) is based on only one specific trajectory (the real market index), while Fig. 4.2 (c) provides the average performance based on 1,000 trajectories. We find that the optimal portfolios perform significantly better than the market index (the dash-dotted constant level) for all time horizons beyond half a year. In particular, the average terminal utility reaches its peak at $\tau = 24$, which is consistent with the result based on the information criteria in Fig. F.1, particularly the AIC. Therefore, according to the utility of portfolio wealth, the optimal strategies outperform the market index for most of the time horizons.

As the second performance measure, Fig. 4.3 (a) reports the Sharpe ratio of the passive holding market index portfolio from January 1881 to December 2012 and the Sharpe ratios of the optimal portfolios for $\tau$ from one month to 60 months together with their 90% confidence intervals (see Jobson and Korkie 1981). If we consider the optimal portfolio as a combination of the market portfolio and a risk-free asset, then the optimal portfolio should be located on the capital market line and hence should have the same Sharpe ratio as the market. However Fig. 4.3 (a) shows that, by taking the timing opportunity (with respect to the market trend and market volatility), the optimal portfolios (the dotted blue line) outperform the market (the solid black line) on average for time horizons from six to 20 months. The results are surprisingly consistent with that in Fig. 4.2 (a) under the utility measure. We also conduct a Monte Carlo analysis based on 1,000 simulations and report the average Sharpe ratios in Fig. 4.3 (b) for the optimal portfolios. It shows the outperformance of the optimal portfolios over the market index based on the Sharpe ratio for the look-back periods of more than six months. The results are consistent with that in

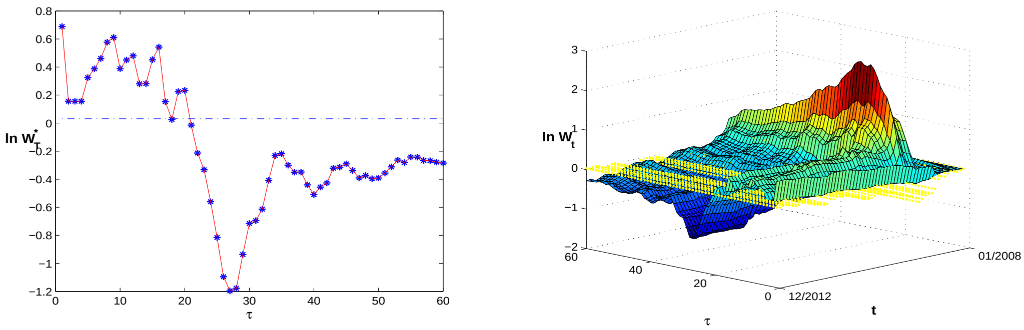
positive ($AC(1) = 0.2839$) while the autocorrelations with higher orders are insignificantly different from zero. This implies that the last period return could well predict the next period return.




(a) (b)

FIGURE 4.3. The Sharpe ratio of the optimal portfolio (the solid blue line) with corresponding 90% confidence intervals (a), the average Sharpe ratio based on 1,000 simulations (b) for $\tau \in [1, 60]$, compared to the passive holding portfolio of market index (the dotted black line) from January 1881 to December 2012.

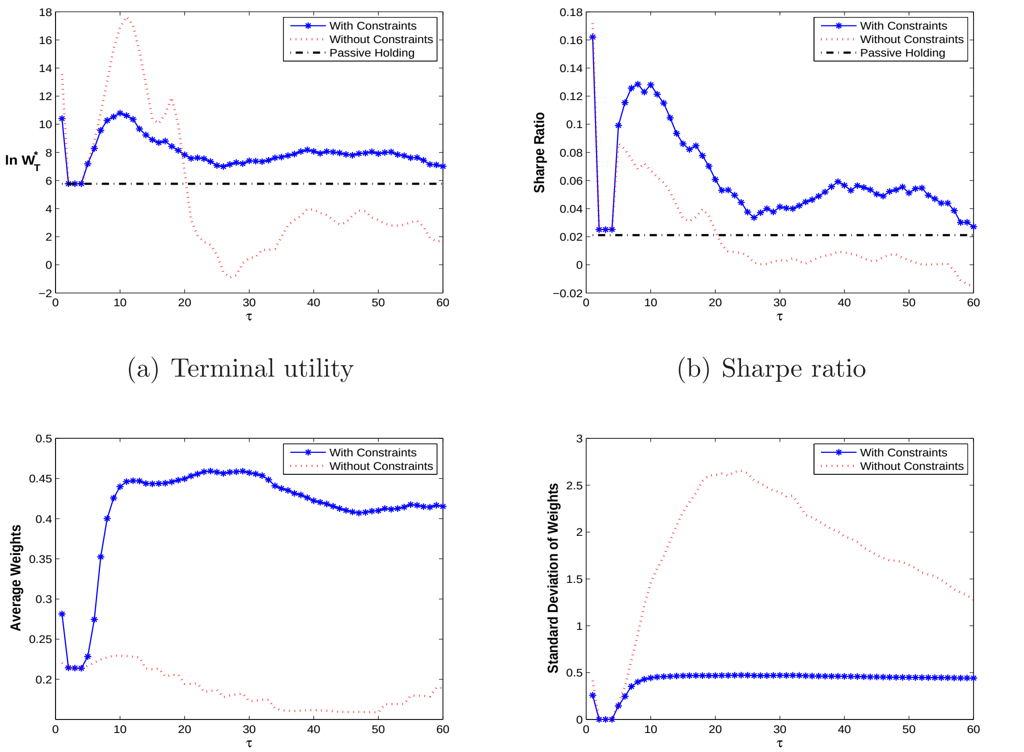
Fig. 4.2 (c) under the portfolio utility measure. In addition, we show in Fig. F.3 that the pure momentum strategies underperform the market in all time horizons from one month to 60 months. Therefore, we have demonstrated the consistent outperformance of the optimal portfolios over the market index and pure strategies under the two performance measures.




(a) Terminal utility (b) Series of utility

FIGURE 4.4. The terminal utility of the wealth for the optimal portfolio, compared to the passive holding market index portfolio (the dotted line), with out-of-sample data from January 2008 to December 2012 for $\tau \in [1, 60]$.

4.3. **The effect on the out-of-sample tests.** For $\tau$ from one month to 60 months, Fig. 4.4 reports the out-of-sample utility of the optimal portfolio wealth from January 2008 to December 2012. It clearly shows that the optimal strategies still outperform the market index for time horizons up to two years. We report additional out-of-sample tests in Appendix F.4 and rolling window estimates in Appendix F.5.



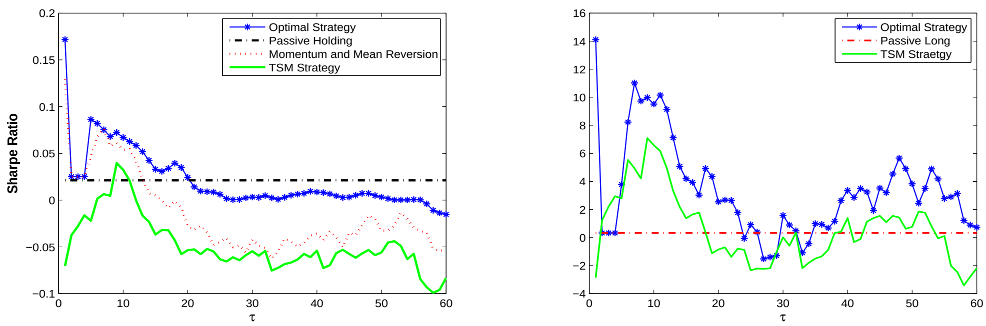

(a) Terminal utility (b) Sharpe ratio

(c) Mean of portfolio weights (d) Standard deviation of portfolio weights

FIGURE 4.5. The terminal utility of wealth (a) and the Sharpe ratio (b) for the optimal portfolio, the mean (c) and the standard deviation of the optimal portfolio weights, with and without short-sale constraints, compared with the market index portfolio.

4.4. **The effect on short-sale constraints.** For different time horizon, Fig. 4.5 (a) and (b) report the terminal utilities of the optimal portfolio wealth and the Sharpe ratio for the optimal portfolio with and without short-sales constraints, respectively, comparing with the passive holding market index portfolio. We also examine the mean and standard deviation of the optimal portfolio weights and report the results in Fig. 4.5 (c) and (d) with and without short-sale constraints. The results for $\tau = 12$ in the previous section also hold. That is, with the constraints, the optimal portfolio weights increase in the mean while volatility is low and stable. On the other hand, without constraints, the volatility of the optimal portfolio weights varies dramatically, which seems in line with the argument of Marquering and Verbeek (2004).

4.5. **Comparison with TSM with different time horizons.** As in the previous section for $\tau = 12$, we use the Sharpe ratio to examine the performance of the optimal strategy $\pi_t^*$ in (2.8) and compare with the passive index strategy and two TSM strategies for time horizons from 1 month to 60 months and one month holding period. We report the Sharpe ratios of the portfolios for the four strategies in Fig. 4.6 (a).[18] For comparison, we collect the Sharpe ratio for the optimal portfolio and the passive holding portfolio reported in Fig. 4.3 and report the Sharpe ratios of the TSM strategy using a solid green line and of momentum and mean-reversion strategy using a dotted red line together in Fig. 4.6 (a) from January 1881 to December 2012. We have three observations. First, the TSM strategy outperforms the market only for $\tau = 9, 10$ and the momentum and mean-reversion strategy outperform the market for short time horizons $\tau \leq 13$. Second, by taking the mean-reversion effect into account, the momentum and mean-reversion strategy performs better than the TSM strategy for all time horizons. Finally, the optimal strategy significantly outperforms both the momentum and mean-reversion strategy (for all time horizons beyond four months) and the TSM strategy (for all time horizons).




(a) Average Sharpe ratio (b) Terminal log cumulative excess return

FIGURE 4.6. (a) The average Sharpe ratio for the optimal portfolio, the momentum and mean-reversion portfolio and the TSM portfolio with $\tau \in [1, 60]$ and the passive holding portfolio from January 1881 until December 2012. (b) Terminal log cumulative excess return of the optimal strategies and TSM strategies with $\tau \in [1, 60]$ and passive long strategy from January 1876 to December 2012.

[18]The monthly Sharpe ratio for the pure mean-reversion strategy is 0.0250, slightly higher than that for the passive holding portfolio (0.0211).

Fig. 4.6 (b) shows the terminal values of the log cumulative excess returns of the optimal strategy (2.8) and the TSM strategy with $\tau \in [1, 60]$, together with the passive long strategy, from January 1876 to December 2012.[19] It shows that the optimal strategy outperforms the TSM strategy for all time horizons (beyond four months), while the TSM strategy outperforms the market for small time horizons (from about two to 18 months). The terminal values of the log cumulative excess return have similar patterns to the average Sharpe ratio reported in Fig. 4.6 (a), especially for small time horizons.

## 5. Conclusion

To characterize the time series momentum in financial markets, we propose a continuous-time model of asset price dynamics with the drift as a weighted average of mean reversion and moving average components. By applying the maximum principle for control problems of stochastic delay differential equations, we derive the optimal strategies in closed form. By estimating the model to the S&P 500, we show that the optimal strategy outperforms the TSM strategy and the market index. The outperformance holds for out-of-sample tests and with short-sale constraints. The outperformance is immune to the market states, investor sentiment and market volatility. The results show that the profitability pattern reflected by the average return of commonly used strategies in much of the empirical literature may not reflect the effect of portfolio wealth.

The model proposed in this paper is simple and stylized. The weights of the momentum and mean-reversion components are constant. When market conditions change, the weights can be time-varying. Hence it would be interesting to model their dependence on market conditions. This can be modelled, for example, as a Markov switching process or based on some rational learning process (Xia 2001). The portfolio performance is examined under log utility in this paper. It would be interesting to study the intertemporal effect under general power utility functions. We could also consider incorporating stochastic volatilities of the return process into the model. Finally, an extension of the model to a multi-asset setting to study cross-sectional optimal strategies would be helpful to understand cross-sectional momentum and reversal.

[19]Note that the passive long strategy introduced in Moskowitz et al. (2012) is different from the passive holding strategy studied in the previous sections. Passive long means holding one share of the index each period; however, passive holding in our paper means investing \$1 in the index in the first period and holding it until the last period.

## Appendix A. Properties of the Solutions to the System (2.2)–(2.4)

Let $C([-\tau,0],R)$ be the space of all continuous functions $\varphi:[-\tau,0]\to R$. For a given initial condition $S_t=\varphi_t$, $t\in[-\tau,0]$ and $\mu_0=\hat{\mu}$, the following proposition shows that the system (2.2)-(2.4) admits pathwise unique solutions such that $S_t>0$ almost surely for all $t\geq 0$ whenever $\varphi_t>0$ for $t\in[-\tau,0]$ almost surely.

**Proposition A.1.** *The system* (2.2)-(2.4) *has an almost surely continuously adapted pathwise unique solution* $(S,\mu)$ *for a given* $\mathcal{F}_0$*-measurable initial process* $\varphi:\Omega\to C([-\tau,0],R)$*. Furthermore, if* $\varphi_t>0$ *for* $t\in[-\tau,0]$ *almost surely, then* $S_t>0$ *for all* $t\geq 0$ *almost surely.*

*Proof.* Basically, the solution can be found by using forward induction steps of length $\tau$ as in Arriojas, Hu, Mohammed and Pap (2007). Let $t\in[0,\tau]$. Then the system (2.2)-(2.4) becomes

$$\begin{cases} dS_t = S_t dN_t, \qquad t\in[0,\tau], \\ d\mu_t = \alpha(\bar{\mu}-\mu_t)dt + \sigma'_\mu dZ_t, \qquad t\in[0,\tau], \\ \ S_t=\varphi_t \quad \text{for} \quad t\in[-\tau,0] \quad \text{almost surely} \quad \text{and} \quad \mu_0=\hat{\mu}. \end{cases} \tag{A.1}$$

where $N_t=\int_0^t\big[\frac{\phi}{\tau}\int_{s-\tau}^s\frac{d\varphi_u}{\varphi_u}+(1-\phi)\mu_s\big]ds+\int_0^t\sigma'_S dZ_s$ is a semimartingale. Denote by $\langle N_t,N_t\rangle=\int_0^t\sigma'_S\sigma_S ds$, $t\in[0,\tau]$, the quadratic variation. Then system (A.1) has a unique solution

$$\begin{cases} S_t=\varphi_0\exp\big\{N_t-\dfrac{1}{2}\langle N_t,N_t\rangle\big\}, \\ \mu_t=\bar{\mu}+(\hat{\mu}-\bar{\mu})\exp\{-\alpha t\}+\sigma'_\mu\exp\{-\alpha t\}\displaystyle\int_0^t\exp\{\alpha u\}dZ_u \end{cases}$$

for $t\in[0,\tau]$. This clearly implies that $S_t>0$ for all $t\in[0,\tau]$ almost surely, when $\varphi_t>0$ for $t\in[-\tau,0]$ almost surely. By a similar argument, it follows that $S_t>0$ for all $t\in[\tau,2\tau]$ almost surely. Therefore $S_t>0$ for all $t\geq 0$ almost surely, by induction. Note that the above argument also gives existence and pathwise-uniqueness of the solution to the system (2.2)-(2.4).

□

## Appendix B. Proof of Proposition 2.1

To solve the stochastic control problems, there are two approaches: the dynamic programming method (HJB equation) and the maximum principle. Since the SDDE is not Markovian, we cannot use the dynamic programming method. Recently, Chen and Wu (2010) introduced a maximum principle for the optimal control problem of SDDE. This method is further extended by Øksendal et al. (2011) to consider a one-dimensional system allowing both delays of moving average type and jumps. Because the optimal control problem of SDDE is relatively new to the field of economics and finance, we first briefly introduce the maximum principle of Chen and Wu (2010) and refer readers to their paper for details.

### B.1. The Maximum Principle for an Optimal Control Problem of SDDE.

Consider a past-dependent state $X_t$ of a control system

$$\begin{cases} dX_t = b(t, X_t, X_{t-\tau}, v_t, v_{t-\tau})dt + \sigma(t, X_t, X_{t-\tau}, v_t, v_{t-\tau})dZ_t, \qquad t \in [0, T], \\ X_t = \xi_t, \qquad v_t = \eta_t, \qquad t \in [-\tau, 0], \end{cases} \tag{B.1}$$

where $Z_t$ is a $d$-dimensional Brownian motion on $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$, and $b : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^n$ and $\sigma : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}^{n\times d}$ are given functions. In addition, $v_t$ is a $\mathcal{F}_t\,(t \geq 0)$-measurable stochastic control with values in $U$, where $U \subset \mathbb{R}^k$ is a nonempty convex set, $\tau > 0$ is a given finite time delay, $\xi \in C[-\tau, 0]$ is the initial path of $X$, and $\eta$, the initial path of $v(\cdot)$, is a given deterministic continuous function from $[-\tau, 0]$ into $U$ such that $\int_{-\tau}^{0} \eta_s^2 ds < +\infty$. The problem is to find the optimal control $u(\cdot) \in \mathcal{A}$, such that

$$J(u(\cdot)) = \sup\{J(v(\cdot)); v(\cdot) \in \mathcal{A}\}, \tag{B.2}$$

where $\mathcal{A}$ denotes the set of all admissible controls. The associated performance function $J$ is given by

$$J(v(\cdot)) = \mathbb{E}\Big[\int_0^T L(t, X_t, v_t, v_{t-\tau})dt + \Phi(X_T)\Big],$$

where $L : [0, T] \times \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^k \to \mathbb{R}$ and $\Phi : \mathbb{R}^n \to \mathbb{R}$ are given functions. Assume **(H1)**: the functions $b$, $\sigma$, $L$ and $\Phi$ are continuously differentiable with respect to $(X_t, X_{t-\tau}, v_t, v_{t-\tau})$ and their derivatives are bounded.

In order to derive the maximum principle, we introduce the following adjoint equation,

$$\begin{cases} -dp_t = \big\{(b^u_X)^\top p_t + (\sigma^u_X)^\top z_t + \mathbb{E}_t[(b^u_{X_\tau}|_{t+\tau})^\top p_{t+\tau} + (\sigma^u_{X_\tau}|_{t+\tau})^\top z_{t+\tau}] \\ \qquad + L_X(t, X_t, u_t, u_{t-\tau})\big\}dt - z_t dZ_t, \quad t \in [0, T], \\ p_T = \Phi_X(X_T), \qquad p_t = 0, \qquad t \in (T, T+\tau], \\ z_t = 0, \qquad t \in [T, T+\tau]. \end{cases} \tag{B.3}$$

We refer readers to Theorems 2.1 and 2.2 in Chen and Wu (2010) for the existence and uniqueness of the solutions of the systems (B.3) and (B.1) respectively.

Next, define a Hamiltonian function $H$ from $[0,T]\times\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^k\times\mathbb{R}^k\times L^2_{\mathcal{F}}(0,T+\tau;\mathbb{R}^n)\times L^2_{\mathcal{F}}(0,T+\tau;\mathbb{R}^{n\times d})$ to $\mathbb{R}$ as follows,

$$\begin{aligned}&H(t,X_t,X_{t-\tau},v_t,v_{t-\tau},p_t,z_t)\\ =&\langle b(t,X_t,X_{t-\tau},v_t,v_{t-\tau}),p_t\rangle+\langle\sigma(t,X_t,X_{t-\tau},v_t,v_{t-\tau}),z_t\rangle+L(t,X_t,v_t,v_{t-\tau}).\end{aligned}$$

Assume **(H2)**: the functions $H(t,\cdot,\cdot,\cdot,\cdot,p_t,z_t)$ and $\Phi(\cdot)$ are concave with respect to the corresponding variables respectively for $t\in[0,T]$ and given $p_t$ and $z_t$. Then we have the following proposition on the maximum principle of the stochastic control system with delay by summarizing Theorem 3.1, Remark 3.4 and Theorem 3.2 in Chen and Wu (2010).

**Proposition B.1.** (i) *Let $u(\cdot)$ be an optimal control of the optimal stochastic control problem with delay subject to* (B.1) *and* (B.2), *and $X(\cdot)$ be the corresponding optimal trajectory. Then we have*

$$\max_{v\in U}\langle H^u_v+\mathbb{E}_t[H^u_{v_\tau}|_{t+\tau}],v\rangle=\langle H^u_v+\mathbb{E}_t[H^u_{v_\tau}|_{t+\tau}],u_t\rangle,\qquad a.e.,\ a.s.;\tag{B.4}$$

(ii) *Suppose $u(\cdot)\in\mathcal{A}$ and let $X(\cdot)$ be the corresponding trajectory, $p_t$ and $z_t$ be the solution of the adjoint equation* (B.3)*. If **(H1)**, **(H2)** and* (B.4) *hold for $u(\cdot)$, then $u(\cdot)$ is an optimal control for the stochastic delayed optimal problem* (B.1) *and* (B.2).

B.2. **Proof of Proposition 2.1.** We now apply Proposition B.1 to our stochastic control problem. Let $P_u:=\ln S_u$ and $V_u:=\frac{W_u^{1-\gamma}-1}{1-\gamma}$. Then the stochastic delayed optimal problem in Section 2 becomes to maximize $\mathbb{E}_u[\Phi(X_T)]:=\mathbb{E}_u[\frac{W_T^{1-\gamma}-1}{1-\gamma}]=\mathbb{E}_u[V_T]$, subject to

$$\begin{cases}dX_u=b(u,X_u,X_{u-\tau},\pi_u)du+\sigma(u,X_u,\pi_u)dZ_u, & u\in[t,T],\\ X_u=\xi_u,\qquad v_u=\eta_u, & u\in[t-\tau,t],\end{cases}\tag{B.5}$$

where

$$X_u=\begin{pmatrix}P_u\\ \mu_u\\ V_u\end{pmatrix},\qquad \sigma=\begin{pmatrix}\sigma_S'\\ \sigma_\mu'\\ \big((1-\gamma)V_u+1\big)\pi_u\sigma_S'\end{pmatrix},$$

$$b=\begin{pmatrix}\frac{\phi}{\tau}(P_u-P_{u-\tau})+(1-\phi)\mu_u-(1-\phi)\frac{\sigma_S'\sigma_S}{2}\\ \alpha(\bar\mu-\mu_u)\\ \big((1-\gamma)V_u+1\big)\Big\{-\frac{\gamma\pi_u^2\sigma_S'\sigma_S}{2}+\pi_u\Big[\frac{\phi}{\tau}(P_u-P_{u-\tau})+\frac{\sigma_S'\sigma_S}{2}\phi+(1-\phi)\mu_u-r\Big]+r\Big\}\end{pmatrix}.$$

Then we have the following adjoint equation

$$\begin{cases} -dp_u = \big\{(b_X^{\pi^*})^\top p_u + (\sigma_X^{\pi^*})^\top z_u + \mathbb{E}_u[(b_{X_\tau}^{\pi^*}|_{u+\tau})^\top p_{u+\tau} + (\sigma_{X_\tau}^{\pi^*}|_{u+\tau})^\top z_{u+\tau}] \\ \qquad + L_X\big\}du - z_u dZ_u, \quad u \in [t,T], \\ p_T = \Phi_X(X_T), \qquad p_u = 0, \qquad u \in (T, T+\tau], \\ z_u = 0, \qquad u \in [T, T+\tau], \end{cases} \tag{B.6}$$

where

$$p_u = (p_u^i)_{3\times 1}, \qquad z_u = (z_u^{ij})_{3\times 2},$$

$$(b_X^{\pi^*})^\top = \begin{pmatrix} \frac{\phi}{\tau} & 0 & [(1-\gamma)V_u + 1]\frac{\phi}{\tau}\pi_u^* \\ 1-\phi & -\alpha & [(1-\gamma)V_u+1](1-\phi)\pi_u^* \\ 0 & 0 & (1-\gamma)\Big\{-\frac{\gamma\pi_u^{*2}\sigma_S'\sigma_S}{2} + \pi_u^*\Big[\frac{\phi}{\tau}(P_u - P_{u-\tau}) + \frac{\sigma_S'\sigma_S}{2}\phi + (1-\phi)\mu_u - r\Big] + r\Big\} \end{pmatrix},$$

$$(b_{X_\tau}^{\pi^*}|_{u+\tau})^\top = \begin{pmatrix} -\frac{\phi}{\tau} & 0 & -[(1-\gamma)V_{u+\tau}+1]\frac{\phi}{\tau}\pi_{u+\tau}^* \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \Phi_X(X_T) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad L_X = 0,$$

$$(\sigma_X^{\pi^*})^\top = \begin{pmatrix} (\sigma_{1X}^{\pi^*})^\top \\ (\sigma_{2X}^{\pi^*})^\top \end{pmatrix}, \quad (\sigma_{iX}^{\pi^*})^\top = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & (1-\gamma)\pi_u^*\sigma_{S(i)} \end{pmatrix}, \quad i = 1,2, \quad (\sigma_{X_\tau}^{\pi^*}|_{u+\tau})^\top = \mathbf{0}_{2\times 3\times 3}.$$

The Hamiltonian function $H$ is given by

$$\begin{aligned} H = &\Big[\frac{\phi}{\tau}(P_u - P_{u-\tau}) + (1-\phi)\mu_u - (1-\phi)\frac{\sigma_S'\sigma_S}{2}\Big]p_u^1 + \alpha(\bar{\mu} - \mu_u)p_u^2 \\ &+ [(1-\gamma)V_{u+\tau}+1]\Big\{-\frac{\gamma\pi_u^2\sigma_S'\sigma_S}{2} + \pi_u\Big[\frac{\phi}{\tau}(P_u - P_{u-\tau}) + \frac{\sigma_S'\sigma_S}{2}\phi + (1-\phi)\mu_u - r\Big] + r\Big\}p_u^3 \\ &+ \sigma_S'\begin{pmatrix} z_u^{11} \\ z_u^{12} \end{pmatrix} + \sigma_\mu'\begin{pmatrix} z_u^{21} \\ z_u^{22} \end{pmatrix} + [(1-\gamma)V_u+1]\pi_u\sigma_S'\begin{pmatrix} z_u^{31} \\ z_u^{32} \end{pmatrix}, \end{aligned}$$

so that

$$H_\pi^{\pi^*} = [(1-\gamma)V_u+1]p_u^3\Big\{-\gamma\pi_u^*\sigma_S'\sigma_S + \frac{\phi}{\tau}(P_u - P_{u-\tau}) + \frac{\sigma_S'\sigma_S}{2}\phi + (1-\phi)\mu_u - r\Big\} + [(1-\gamma)V_u+1]\sigma_S'\begin{pmatrix} z_u^{31} \\ z_u^{32} \end{pmatrix}.$$

It can be verified that $\mathbb{E}_u[H_{\pi_\tau}^{\pi^*}|_{u+\tau}] = 0$. Therefore,

$$\langle H_\pi^{\pi^*} + \mathbb{E}_u[H_{\pi_\tau}^{\pi^*}|_{u+\tau}], \pi\rangle = \pi_u H_\pi^{\pi^*}.$$

Taking the derivative with respect to $\pi_u$ and letting it equal zero yields

$$\begin{aligned} \pi_u^* &= \frac{\frac{\phi}{\tau}(P_u - P_{u-\tau}) + \frac{\sigma_S'\sigma_S}{2}\phi + (1-\phi)\mu_u - r}{\gamma\sigma_S'\sigma_S} + \frac{\sigma_{S(1)}z_u^{31} + \sigma_{S(2)}z_u^{32}}{\gamma p_u^3\sigma_S'\sigma_S} \\ &= \frac{\phi m_u + (1-\phi)\mu_u - r}{\gamma\sigma_S'\sigma_S} + \frac{\sigma_{S(1)}z_u^{31} + \sigma_{S(2)}z_u^{32}}{\gamma p_u^3\sigma_S'\sigma_S}, \end{aligned} \tag{B.7}$$

where $z_u$ and $p_u$ are governed by the backward stochastic differential system (B.6). This gives the optimal investment strategy.

Especially, if $\gamma = 1$, the utility reduces to a log one. Then the parameter matrices in the adjoint equation (B.6) become

$$(b_X^{\pi^*})^\top = \begin{pmatrix} \frac{\phi}{\tau} & 0 & \frac{\phi}{\tau}\pi_u^* \\ 1-\phi & -\alpha & (1-\phi)\pi_u^* \\ 0 & 0 & 0 \end{pmatrix}, \quad (b_{X_\tau}^{\pi^*}|_{u+\tau})^\top = \begin{pmatrix} -\frac{\phi}{\tau} & 0 & -\frac{\phi}{\tau}\pi_{u+\tau}^* \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\Phi_X(X_T) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad L_X = 0, \quad (\sigma_X^{\pi^*})^\top = (\sigma_{X_\tau}^{\pi^*}|_{u+\tau})^\top = \mathbf{0}_{2\times 3\times 3}.$$

Since the parameters and terminal values for $dp_u^3$ are deterministic in this case, we can assert that $z_u^{31} = z_u^{32} = 0$ for $u \in [t, T]$, which leads to $p_u^3 = 1$ for $u \in [t, T]$. Then the Hamiltonian function $H$ is given by

$$\begin{aligned} H = & \Big[\frac{\phi}{\tau}(P_u - P_{u-\tau}) + (1-\phi)\mu_u - (1-\phi)\frac{\sigma_S'\sigma_S}{2}\Big]p_u^1 + \alpha(\bar{\mu} - \mu_u)p_u^2 \\ & + \Big\{ -\frac{\pi_u^2\sigma_S'\sigma_S}{2} + \pi_u\big[\frac{\phi}{\tau}(P_u - P_{u-\tau}) + \frac{\sigma_S'\sigma_S}{2}\phi + (1-\phi)\mu_u - r\big] + r\Big\}p_u^3 \\ & + \sigma_S'\begin{pmatrix} z_u^{11} \\ z_u^{12} \end{pmatrix} + \sigma_\mu'\begin{pmatrix} z_u^{21} \\ z_u^{22} \end{pmatrix}, \end{aligned}$$

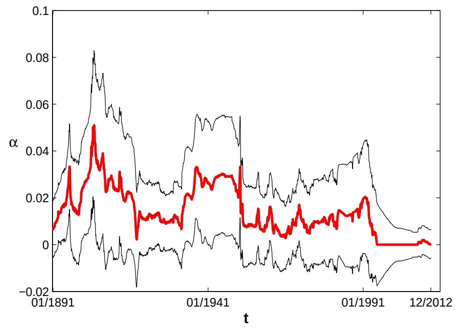
and the optimal strategy is given by

$$\pi_u^* = \frac{\phi m_u + (1-\phi)\mu_u - r}{\sigma_S'\sigma_S},$$

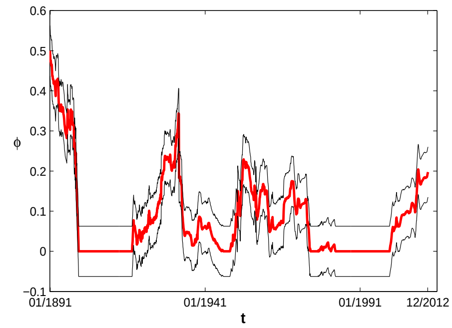
which is myopic.

## Appendix C. Additional out-of-sample tests and Rolling Window Estimations

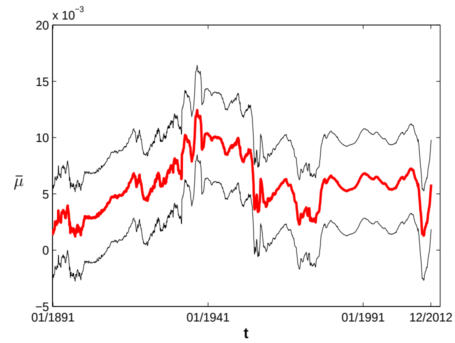
In this appendix, we provide some robustness analysis to out-of-sample tests and rolling window estimations.



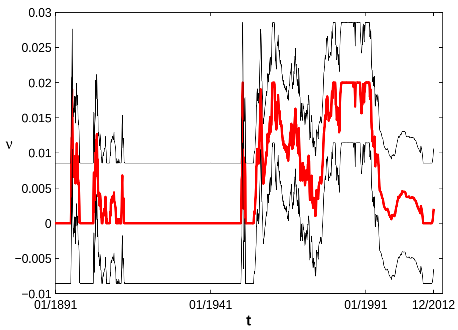

(a) Estimates of $\alpha$



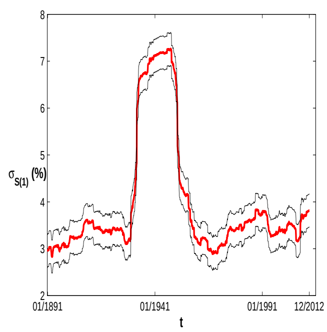

(b) Estimates of $\phi$



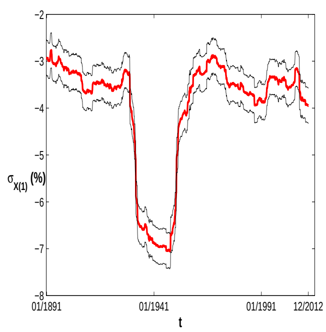

(c) Estimates of $\bar{\mu}$



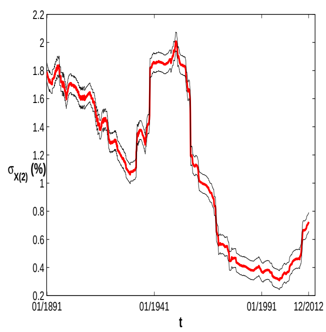

(d) Estimates of $\nu$




(e) Estimates of $\sigma_{S(1)}$




(f) Estimates of $\sigma_{X(1)}$

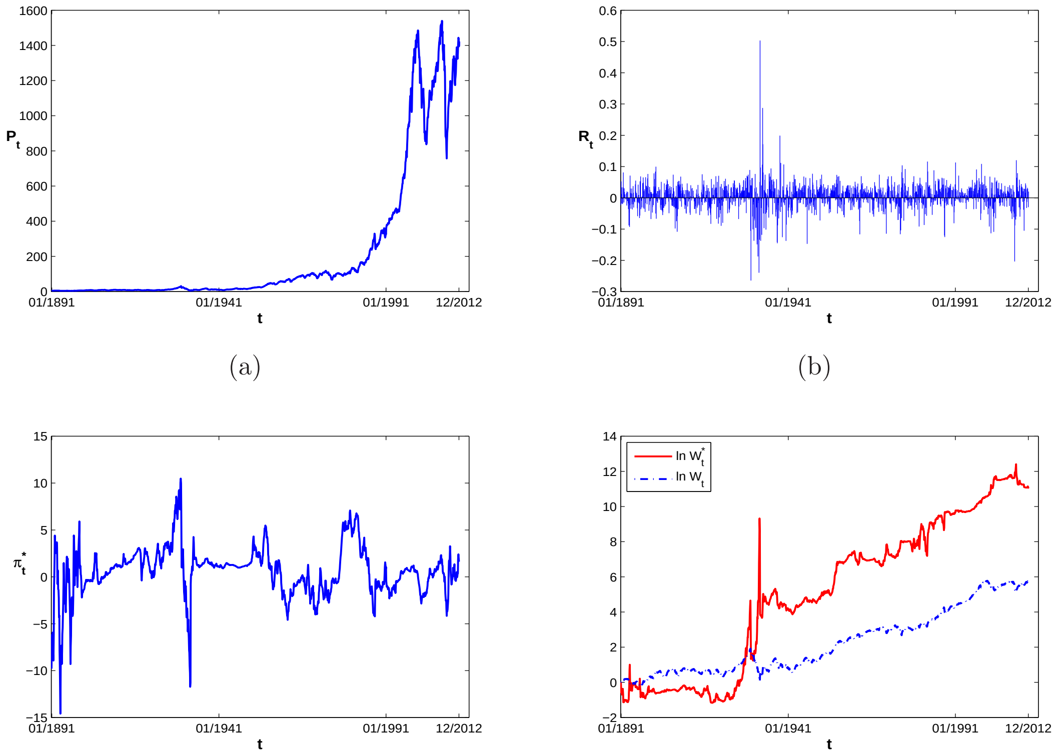



(g) Estimates of $\sigma_{X(2)}$

Figure C.1. The estimates of (a) $\alpha$; (b) $\phi$; (c) $\bar{\mu}$; (d) $\nu$; (e) $\sigma_{S(1)}$; (f) $\sigma_{X(1)}$ and (g) $\sigma_{X(2)}$ for $\tau = 12$ based on data from the past 20 years.

C.1. **Rolling Window Estimations.** For fix $\tau = 12$, we estimate parameters of (3.2) at each month by using the past 20 years' data to avoid look-ahead bias. Fig.

C.1 illustrates the estimated parameters. The big jump in estimated $\sigma_{S(1)}$ during 1930–1950 is consistent with the high volatility of market return illustrated in Fig. C.2 (b). Fig. C.1 also illustrates the interesting phenomenon: that the estimated $\phi$ is very close to zero for three periods of time, implying insignificant momentum but significant mean-reversion effect. By comparing Fig. C.1 (b) and (e), we observe that the insignificant $\phi$ is accompanied by high volatility $\sigma_{S(1)}$.




(a) (b) (c) (d)

FIGURE C.2. The time series of (a) the index level and (b) the simple return of the total return index of S&P 500; (c) the optimal portfolio and (d) the utility of wealth from December 1890 until December 2012 for $\tau = 12$ with 20-year rolling window estimated parameters.

Fig. C.2 illustrates the time series of (a) the index level and (b) the simple return of the total return index of S&P 500; (c) the optimal portfolio and (d) the utility of the optimal portfolio wealth from December 1890 to December 2012 for $\tau = 12$ with 20-year rolling window estimated parameters. The index return and $\pi_t^*$ are positively correlated with correlation 0.0620. In addition, we find that the profits are higher after the 1930s.

Fig. C.3 illustrates the estimates of $\sigma_{S(1)}$ for the pure momentum model ($\phi = 1$) based on data from the past 20 years; the big jump in volatility is due to the Great

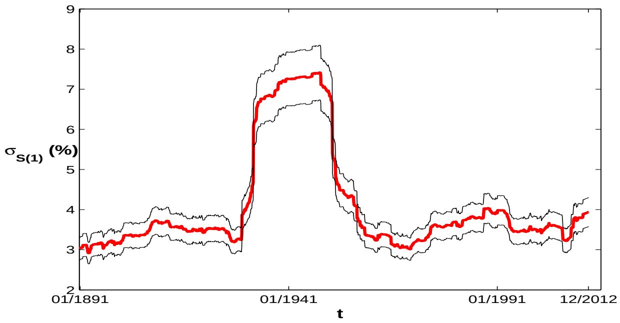
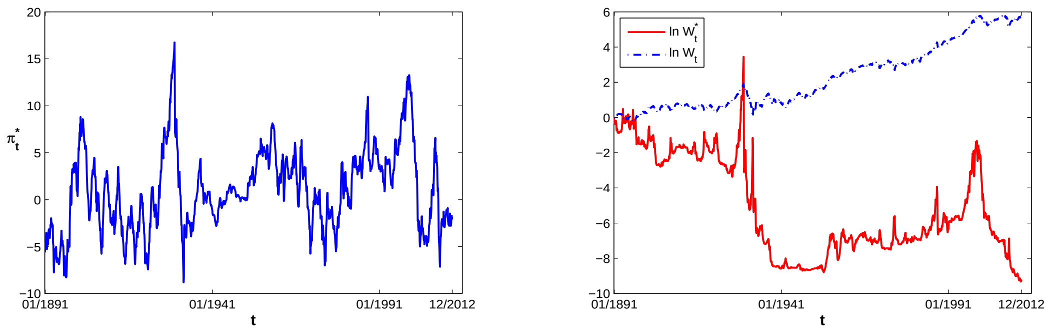



FIGURE C.3. Estimates of $\sigma_{S(1)}$ for the pure momentum model ($\phi = 1$) based on data from the past 20 years.



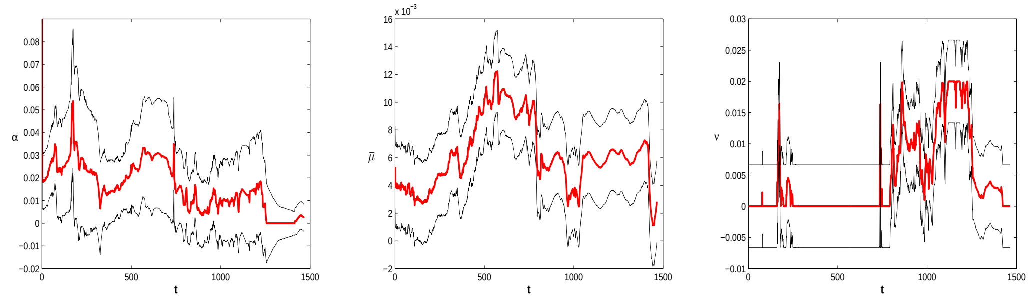

(a) (b)

FIGURE C.4. The time series of (a) the optimal portfolio and (b) the utility of wealth from December 1890 until December 2012 for $\tau = 12$ for the pure momentum model with 20-year rolling window estimated parameters.

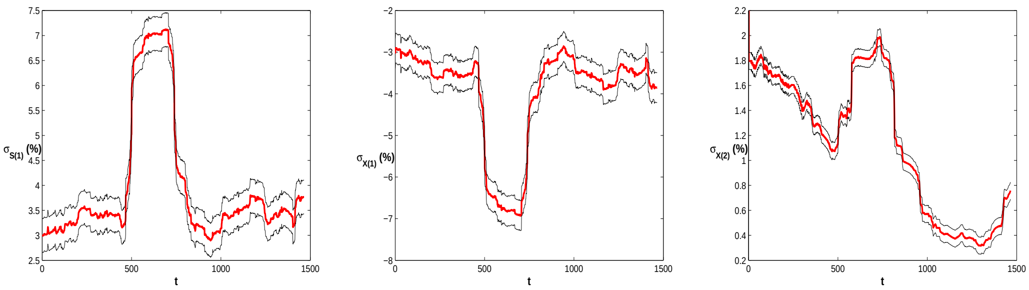
Depression in the 1930s. Fig. C.4 illustrates the time series of (a) the optimal portfolio and (b) the utility of wealth from December 1890 until December 2012 for $\tau = 12$ for the pure momentum model with the 20-year rolling window estimated $\sigma_{S(1)}$. By comparing Fig. C.3 and Fig. C.4 (b), the optimal strategy implied by the pure momentum model suffers huge losses during the high market volatility period. However, Fig C.2 illustrates that the optimal strategy implied by the full model makes big profits during the big market volatility period.

Fig. C.5 illustrates the estimated parameters for the pure mean-reversion model based on data from the past 20-years.

Fig. C.6 illustrates the time series of the optimal portfolio and the utility of wealth from December 1890 until December 2012 for the pure mean-reversion model with 20-year rolling window estimated parameters. After eliminating the look-ahead bias, the pure mean-reversion strategy cannot outperform the stock index any longer.

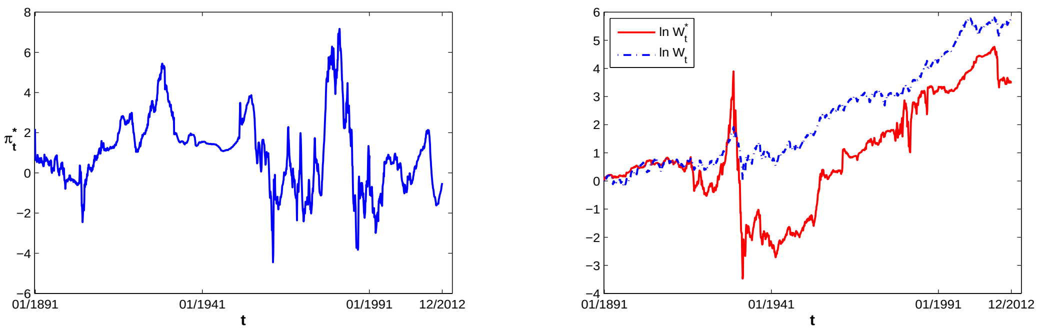


(a) Estimates of $\alpha$ (b) Estimates of $\bar{\mu}$ (c) Estimates of $\nu$



(d) Estimates of $\sigma_{S(1)}$ (e) Estimates of $\sigma_{X(1)}$ (f) Estimates of $\sigma_{X(2)}$

FIGURE C.5. The estimates of (a) $\alpha$; (b) $\phi$; (c) $\bar{\mu}$; (d) $\nu$; (e) $\sigma_{S(1)}$; (f) $\sigma_{X(1)}$ and (g) $\sigma_{X(2)}$ for the pure mean-reversion model based on data from the past 20 years.




(a) (b)

FIGURE C.6. The time series of (a) the optimal portfolio and (b) the utility of wealth from December 1890 until December 2012 for the pure mean-reversion model with 20-year rolling window estimated parameters.

## Appendix D. Regressions on the Market States, Sentiment and Volatility

D.1. **Market States.** First, we follow Cooper et al. (2004) and Hou et al. (2009) and define market state using the cumulative return of the stock index (including dividends) over the most recent 36 months.[20] We label a month as an up (down) market month if the market's three-year return is non-negative (negative). There are 1,165 up months and 478 down months from February 1876[21] to December 2012.

Table D.1. The average excess return of the optimal strategy for $\tau = 12$.

| | Observations (N) | Average excess return |
|---|---|---|
| Unconditional return | 1,643 | 0.0087 |
| | | (2.37) |
| Up market | 1,165 | 0.0081 |
| | | (4.09) |
| Down market | 478 | 0.0101 |
| | | (0.87) |

We compute the average return of the optimal strategy and compare the average returns between up and down market months. Table D.1 presents the average unconditional excess returns and the average excess returns for up and down market months. The unconditional average excess return is 87 basis points per month. In up market months, the average excess return is 81 basis points and it is statistically significant. In down market months, the average excess return is 101 basis points; this value is economically but not statistically significant. The difference between down and up months is 20 basis points, which is not significantly different from zero based on a two-sample $t$-test ($p$-value of 0.87).[22]

We use the following regression model to test for the difference in returns:

$$R_t^* - r = \alpha + \kappa I_t(UP) + \beta(R_t - r) + \epsilon_t, \tag{D.1}$$

where $R_t^* = (W_t^* - W_{t-1}^*)/W_{t-1}^*$ in (3.3) is the month $t$ return of the optimal strategy, $R_t - r$ is the excess return of the stock index, and $I_t(UP)$ is a dummy variable that takes the value of 1 if month $t$ is in an up month, and zero otherwise. The regression intercept $\alpha$ measures the average return of the optimal strategy in down market months, and the coefficient $\kappa$ captures the incremental average return in up market

[20] The results are similar if we use the alternative 6-, 12- or 24- month market state definitions, even though they are more sensitive to sudden changes in market sentiment.

[21] We exclude January 1876 in which there is no return to the optimal strategies.

[22] The $p$-values for the pure momentum strategy, pure mean-reversion strategy and TSM are 0.87, 0.87 and 0.67 respectively.

months relative to down months. We also replace the market state dummy in (D.1) with the lagged market return over the previous 36 months (not reported here), and the results are robust.

TABLE D.2. The coefficients for the regression (D.1)-(D.2).

| | FM | MM | MRM | TSM | FM | MM | MRM | TSM |
|---|---|---|---|---|---|---|---|---|
| $\alpha$ | 0.0094 | 0.0476 | -0.0000 | 0.0060 | 0.0086 | 0.0423 | 0.0002 | 0.0058 |
| | (1.46) | (1.34) | (-0.01) | (3.23) | (1.44) | (1.32) | (0.18) | (3.29) |
| $\kappa$ | 0.0005 | 0.0041 | -0.0005 | -0.0014 | -0.0008 | -0.0034 | -0.0002 | -0.0017 |
| | (0.06) | (0.10) | (-0.32) | (-0.63) | (-0.11) | (-0.09) | (-0.14) | (-0.81) |
| $\beta$ | -1.0523 | -6.7491 | 0.3587 | -0.1548 | | | | |
| | (-12.48) | (-14.60) | (22.97) | (-6.39) | | | | |
| $\beta_1$ | | | | | 0.1994 | 0.7189 | 0.0708 | 0.1341 |
| | | | | | (1.90) | (1.27) | (3.84) | (4.30) |
| $\beta_2$ | | | | | -2.5326 | -15.5802 | 0.6991 | -0.4964 |
| | | | | | (-22.16) | (-25.31) | (34.88) | (-14.63) |

The first four columns of Table D.2 reports the regression coefficients of (D.1) for the full model, the pure momentum model, pure mean-reversion model and the TSM strategy in Moskowitz et al. (2012) for $\tau = 12$ respectively. We see that for all strategies, the differences in returns between down and up market are not significant. Also, the returns in down market are not significant, except for the TSM which earns significant positive returns in down market. The results are consistent with those in Table D.1.

To further control for market risk in up and down market months, we now run the following regression:

$$R_t^* - r = \alpha + \kappa I_t(UP) + \beta_1(R_t - r)I_t(UP) + \beta_2(R_t - r)I_t(DOWN) + \epsilon_t. \quad \text{(D.2)}$$

The regression coefficients are reported in the last four columns of Table D.2. Again, we obtain similar results to (D.1).

TABLE D.3. The coefficients for the regression (D.3)-(D.4).

| | Excess return | | | | CAPM-adj return | | | |
|---|---|---|---|---|---|---|---|---|
| | FM | MM | MRM | TSM | FM | MM | MRM | TSM |
| $\alpha$ | 0.0083 | 0.0409 | 0.0002 | 0.0057 | 0.0094 | 0.0476 | 0.0000 | 0.0060 |
| | (1.22) | (1.08) | (0.14) | (3.03) | (1.46) | (1.34) | (-0.01) | (3.23) |
| $\kappa$ | 0.0006 | 0.0037 | -0.0002 | -0.0012 | 0.0005 | 0.0041 | -0.0005 | -0.0014 |
| | (0.07) | (0.08) | (-0.14) | (-0.53) | (0.06) | (0.10) | (-0.32) | (-0.63) |

When we regress excess return on dummy variable of previous month's state $I_{t-1}(UP)$:

$$R_t^* - r = \alpha + \kappa I_{t-1}(UP) + \epsilon_t, \tag{D.3}$$

we find insignificant $\kappa$s in Table D.3, which indicate the insignificant incremental predictive power of up market state for returns of all strategies. However, we observe significant estimate of $\alpha$ for TSM, and larger $t$-statistics of $\alpha$ for FM and MM (although not significant at conventional level) than that of MRM. This implies that down market state predicts TSM returns, and it also has stronger predictive power for the optimal strategy and pure momentum strategy compared to that of mean reversion. We obtain the same result for the CAPM-adjusted returns:

$$\begin{aligned} &R_t^* - r = \alpha^{CAPM} + \beta^{CAPM}(R_t - r) + \varepsilon_t, \\ &R_t^* - r - \beta^{CAPM}(R_t - r) = \alpha + \kappa I_{t-1}(UP) + \epsilon_t. \end{aligned} \tag{D.4}$$

In summary, whereas cross-sectional momentum usually generates higher returns in up months in Hou et al. (2009), we do not find significant differences in returns between up and down months for the strategies from our model and the TSM. The TSM has significant positive returns in down market months, which also has significant predictive power to next month's TSM returns.

Table D.4. The coefficients for the regression (D.5).

| | FM | MM | MRM | TSM |
|---|---|---|---|---|
| $a$ | 0.0059 | 0.0267 | 0.0005 | 0.0040 |
| | (1.77) | (1.74) | (1.49) | (2.57) |
| $b$ | 0.0040 | 0.0134 | -0.0003 | 0.0023 |
| | (1.20) | (0.87) | (-1.01) | (1.48) |

D.2. **Investor Sentiment.** In this subsection, we examine if investor sentiment predicts returns of the optimal strategies:

$$R_t^* - r = a + bT_{t-1} + \epsilon, \tag{D.5}$$

where $T_t$ is the sentiment index constructed by Baker and Wurgler (2006).[23] We see from Table D.4 that none of the estimates of $b$ is significant, which suggests that investor sentiment has no predictive power for returns of optimal strategies and of the TSM. We also examine monthly changes of the level of sentiment by replacing $T_t$ with its monthly changes and their orthogonalized indexes. The results are similar.

Table D.5. The coefficients for the regression (D.6)-(D.7).

| | FM | MM | MRM | TSM | FM | MM | MRM | TSM |
|---|---|---|---|---|---|---|---|---|
| $\alpha$ | 0.0037 | 0.0421 | -0.0014 | 0.0053 | -0.0020 | -0.0151 | 0.0012 | 0.0025 |
| | (1.06) | (1.11) | (-1.00) | (2.78) | (-0.27) | (-0.36) | (0.80) | (1.19) |
| $\kappa$ | 0.0138 | 0.0141 | 0.0137 | -0.0232 | | | | |
| | (0.25) | (0.05) | (1.21) | (-1.48) | | | | |
| $\kappa_1$ | | | | | 0.1043 | 0.5763 | -0.0127 | 0.0016 |
| | | | | | (1.34) | (1.33) | (-0.80) | (-0.07) |
| $\kappa_2$ | | | | | 0.1026 | 0.5564 | -0.0098 | 0.0084 |
| | | | | | (1.80) | (1.75) | (-0.84) | (0.53) |

D.3. **Market Volatility.** Finally, we examine the predictability of market volatility for portfolio returns:

$$R_t^* - r = \alpha + \kappa \hat{\sigma}_{S,t-1} + \epsilon_t, \tag{D.6}$$

where the ex-ante annualized volatility $\hat{\sigma}_{S,t}$ is given by (3.6). We see from Table D.5 that the estimated $\kappa$s are not significant, which implies that volatility has no predictive power for returns of optimal strategies and of the TSM. We obtain similar results even if we separate volatility into up and down market months as Wang and Xu (2012):

$$R_t^* - r = \alpha + \kappa_1 \hat{\sigma}_{S,t-1}^{+} + \kappa_2 \hat{\sigma}_{S,t-1}^{-} + \epsilon_t, \tag{D.7}$$

where $\hat{\sigma}_{S,t}^{+}$ ($\hat{\sigma}_{S,t}^{-}$) is equal to $\hat{\sigma}_{S,t}$ if the market state is up (down) and otherwise equal to zero.

[23]The data on the Baker-Wurger sentiment index from 07/1965 to 12/2010 is obtained from the Jeffrey Wurglers website (http://people.stern.nyu.edu/jwurgler/).

## Appendix E. Comparison with Moskowitz, Ooi and Pedersen (2012)

Table E.1. The average excess return (%) of the optimal strategies for different look-back period $\tau$ (different row) and different holding period $h$ (different column).

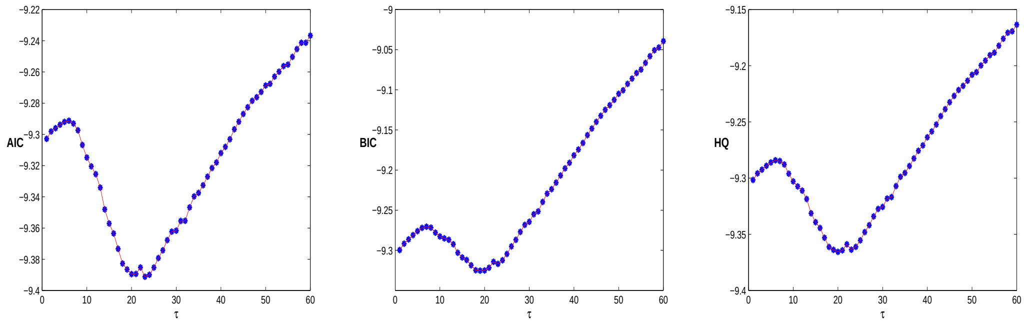
| $(\tau \setminus h)$ | 1 | 3 | 6 | 9 | 12 | 24 | 36 | 48 | 60 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.1337 | 0.1387 | 0.1874* | 0.1573* | 0.0998 | 0.0222 | 0.0328 | 0.0479 | 0.0362 |
| | (1.28) | (1.84) | (3.29) | (2.83) | (1.84) | (0.42) | (0.63) | (0.90) | (0.66) |
| 3 | 0.0972 | 0.0972 | 0.0972 | 0.0972 | 0.0972 | 0.0972 | 0.0972 | 0.0972 | 0.0972 |
| | (0.93) | (0.93) | (0.93) | (0.93) | (0.93) | (0.93) | (0.93) | (0.93) | (0.93) |
| 6 | 0.2022 | 0.2173* | 0.2315* | 0.1462 | 0.0700 | -0.0414 | 0.0199 | 0.0304 | 0.0014 |
| | (1.93) | (2.28) | (2.60) | (1.75) | (0.88) | (-0.58) | (0.32) | (0.53) | (0.02) |
| 9 | 0.3413* | 0.3067* | 0.2106* | 0.1242 | 0.0333 | -0.0777 | -0.0095 | 0.0000 | -0.0450 |
| | (3.27) | (3.12) | (2.28) | (1.45) | (0.41) | (-1.16) | (-0.17) | (0.00) | (-1.11) |
| 12 | 0.1941 | 0.1369 | 0.0756 | -0.0041 | -0.0647 | -0.0931 | -0.0234 | -0.0137 | -0.0587 |
| | (1.85) | (1.40) | (0.80) | (-0.04) | (-0.76) | (-1.30) | (-0.41) | (-0.30) | (-1.46) |
| 24 | -0.0029 | -0.0513 | -0.0776 | -0.0591 | -0.0557 | -0.0271 | 0.0261 | -0.0020 | -0.0082 |
| | (-0.03) | (-0.51) | (-0.79) | (-0.62) | (-0.61) | (-0.34) | (0.40) | (-0.03) | (-0.15) |
| 36 | 0.0369 | 0.0602 | 0.0517 | 0.0419 | 0.0416 | 0.0657 | 0.0351 | 0.0273 | 0.0406 |
| | (0.35) | (0.59) | (0.52) | (0.43) | (0.44) | (0.81) | (0.49) | (0.42) | (0.64) |
| 48 | 0.1819 | 0.1307 | 0.1035 | 0.0895 | 0.0407 | -0.0172 | 0.0179 | 0.0500 | 0.0595 |
| | (1.74) | (1.30) | (1.06) | (0.93) | (0.43) | (-0.21) | (0.24) | (0.70) | (0.86) |
| 60 | -0.0049 | -0.0263 | -0.0800 | -0.1160 | -0.1289 | -0.0396 | 0.0424 | 0.0518 | 0.0680 |
| | (-0.05) | (-0.26) | (-0.81) | (-1.20) | (-1.41) | (-0.49) | (0.55) | (0.69) | (0.92) |

Table E.2. The average excess return (%) of the optimal strategies for different look-back period $\tau$ (different row) and different holding period $h$ (different column) for the pure momentum model.

| $(\tau \setminus h)$ | 1 | 3 | 6 | 9 | 12 | 24 | 36 | 48 | 60 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | -0.0144 | 0.0652 | 0.0714 | 0.0689 | 0.0568 | -0.0040 | 0.0006 | 0.0010 | -0.0133 |
| | (-0.14) | (0.89) | (1.34) | (1.52) | (1.37) | (-0.12) | (0.02) | (0.04) | (-0.57) |
| 3 | 0.1683 | 0.1915* | 0.1460 | 0.1536* | 0.0764 | -0.0360 | -0.0290 | -0.0143 | -0.0395 |
| | (1.61) | (2.16) | (1.91) | (2.20) | (1.17) | (-0.69) | (-0.72) | (-0.45) | (-1.38) |
| 6 | 0.2906* | 0.2633* | 0.2635* | 0.1884* | 0.1031 | -0.0484 | -0.0130 | 0.0157 | -0.0281 |
| | (2.78) | (2.79) | (3.01) | (2.29) | (1.34) | (-0.75) | (-0.26) | (0.40) | (-0.77) |
| 9 | 0.4075* | 0.3779* | 0.2422* | 0.1538 | 0.0545 | -0.0735 | -0.0217 | -0.0047 | -0.0460 |
| | (3.91) | (3.78) | (2.62) | (1.76) | (0.66) | (-1.05) | (-0.38) | (-0.10) | (-1.12) |
| 12 | 0.2453* | 0.1660 | 0.0904 | 0.0122 | -0.0748 | -0.1195 | -0.0602 | -0.0454 | -0.0798 |
| | (2.35) | (1.67) | (0.94) | (0.13) | (-0.86) | (-1.63) | (-1.02) | (-0.95) | (-1.88) |
| 24 | 0.0092 | -0.0242 | -0.0800 | -0.0962 | -0.0955 | -0.0682 | -0.0081 | -0.0140 | -0.0211 |
| | (0.09) | (-0.24) | (-0.81) | (-1.03) | (-1.06) | (-0.88) | (-0.13) | (-0.24) | (-0.39) |
| 36 | -0.0005 | 0.0194 | 0.0219 | 0.0212 | 0.0113 | 0.0030 | 0.0127 | 0.0241 | 0.0206 |
| | (-0.01) | (0.19) | (0.23) | (0.22) | (0.12) | (0.04) | (0.18) | (0.37) | (0.33) |
| 48 | 0.0779 | 0.0733 | 0.0231 | 0.0019 | -0.0392 | -0.0676 | -0.0004 | 0.0435 | 0.0382 |
| | (0.74) | (0.73) | (0.24) | (0.02) | (-0.42) | (-0.83) | (-0.01) | (0.61) | (0.55) |
| 60 | -0.0568 | -0.0852 | -0.1403 | -0.1706 | -0.1986 | -0.1091 | -0.0043 | 0.0157 | 0.0239 |
| | (-0.54) | (-0.84) | (-1.41) | (-1.77) | (-2.15) | (-1.36) | (-0.06) | (0.22) | (0.34) |

## Appendix F. The Effects of Time Horizons

F.1. **Information Criteria.** We present different information criteria, including Akaike (AIC), Bayesian (BIC) and Hannan–Quinn (HQ) information criteria for $\tau$ from one month to 60 months in Fig. F.1. We see that the AIC, BIC and HQ reach their minima at $\tau = 23, 19$ and $20$ respectively. Also, we observe a common increasing pattern of the criteria level for longer $\tau$.

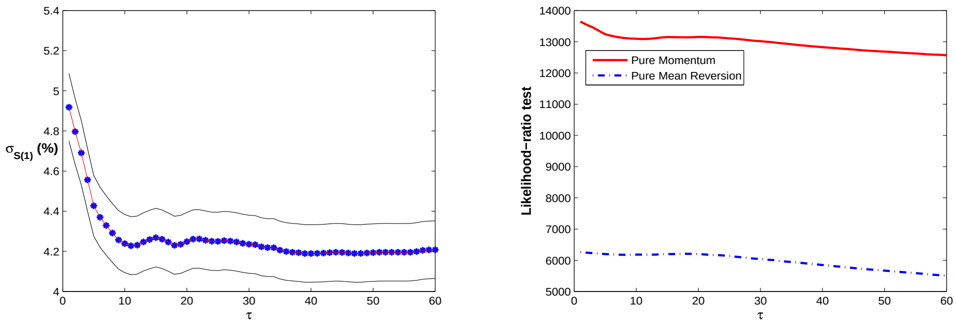



(a) AIC (b) BIC (c) HQ

Figure F.1. (a) Akaike information criteria, (b) Bayesian information criteria, and (c) Hannan–Quinn information criteria for $\tau \in [1, 60]$.

F.2. **The Estimation and Log-likelihood Ratio Test for the Pure Momentum Strategy.** We present the estimation of parameter $\sigma_{S(1)}$ for the pure momentum model in Fig. F.2(a) for $\tau \in [1, 60]$. We also present the results of the log-likelihood ratio test to compare the full model to the estimated pure momentum model ($\phi = 1$) and pure mean-reversion model with respect to different $\tau$ in Fig. F.2 (b), where the solid red line shows the test statistic when compared to the pure momentum model. The statistic is much greater than 12.59, the critical value with six degrees of freedom at the 5% significance level. The dash-dotted blue line illustrates the test statistic when comparing to the pure mean-reversion model. The test statistic is much greater than 3.841, the critical value with one degree of freedom at 5% significance level.

F.3. **The Performance for Different Time Horizons.** To compare with the performance of the pure momentum strategies to the market index, based on estimated parameters in Fig. F.2, we report the terminal utilities of the portfolios of the pure momentum model at December 2012 in Fig. F.3.




(a) The estimates of $\sigma_{S(1)}$ (b) The log-likelihood ratio test

FIGURE F.2. The estimates of $\sigma_{S(1)}$ for the pure momentum model (a) and the log-likelihood ratio test (b) for $\tau \in [1, 60]$.

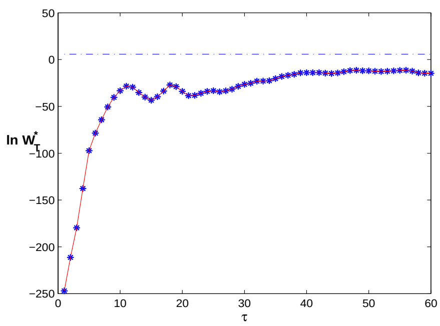



FIGURE F.3. The utility of terminal wealth of the pure momentum model for $\tau \in [1, 60]$.

For $\tau = 1, 2, \cdots, 60$, Fig. F.4 illustrates the evolution of the utility of the optimal portfolio wealth (the dark and more volatile surface) and of the passive holding index portfolio (the yellow and smooth surface) from January 1876 until December 2012. It indicates that the optimal strategies outperform the market index for $\tau$ from five months to 20 months consistently.

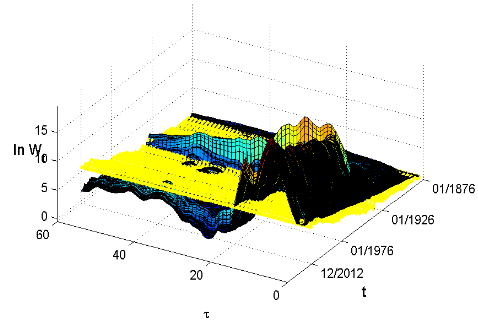



FIGURE F.4. The utility of the optimal portfolio wealth from January 1876 until December 2012 comparing with the passive holding portfolio for $\tau \in [1, 60]$.

F.4. **The Out-of-sample Test for Different Time Horizons.** To see the effect of the time horizon on the results of out-of-sample tests, we split the whole data set into two equal periods: January 1871–December 1941 and January 1942–December 2012. For given $\tau$, we estimate the model for the first sub-sample period and do the out-of-sample test over the second sub-sample period. We report the utility of terminal wealth for $\tau \in [1, 60]$ using sample data of the last 71 years in Fig. F.5. Clearly the optimal strategies still outperform the market for $\tau \in [1, 14]$.

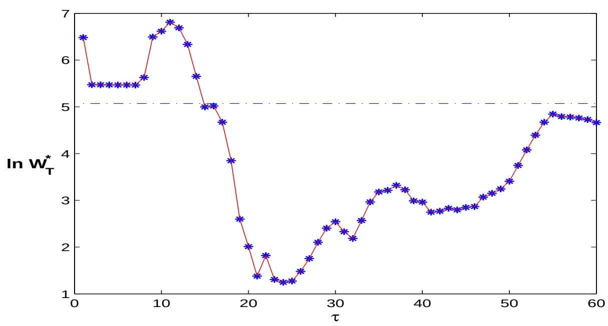



FIGURE F.5. The utility of terminal wealth for $\tau \in [1, 60]$ based on the out-of-sample period January 1942 to December 2012.

F.5. **Rolling Window Estimates for Different Time Horizons.** We also implement rolling window estimations for different time horizons. Fig. F.6 illustrates the correlations of the estimated $\sigma_{S(1)}$ with (a) the estimated $\phi$ and the return of the optimal strategies for (b) the full model, (c) the pure momentum model and (d) the TSM return for $\tau \in [1, 60]$. Interestingly, higher volatility is accompanied by a less significant momentum effect with small time horizons ($\tau \leq 13$). But $\phi$ and $\sigma_{S(1)}$ are positive correlated when the time horizon becomes large. One possible reason is that a long time horizon makes the trading signal less sensitive to changes in price

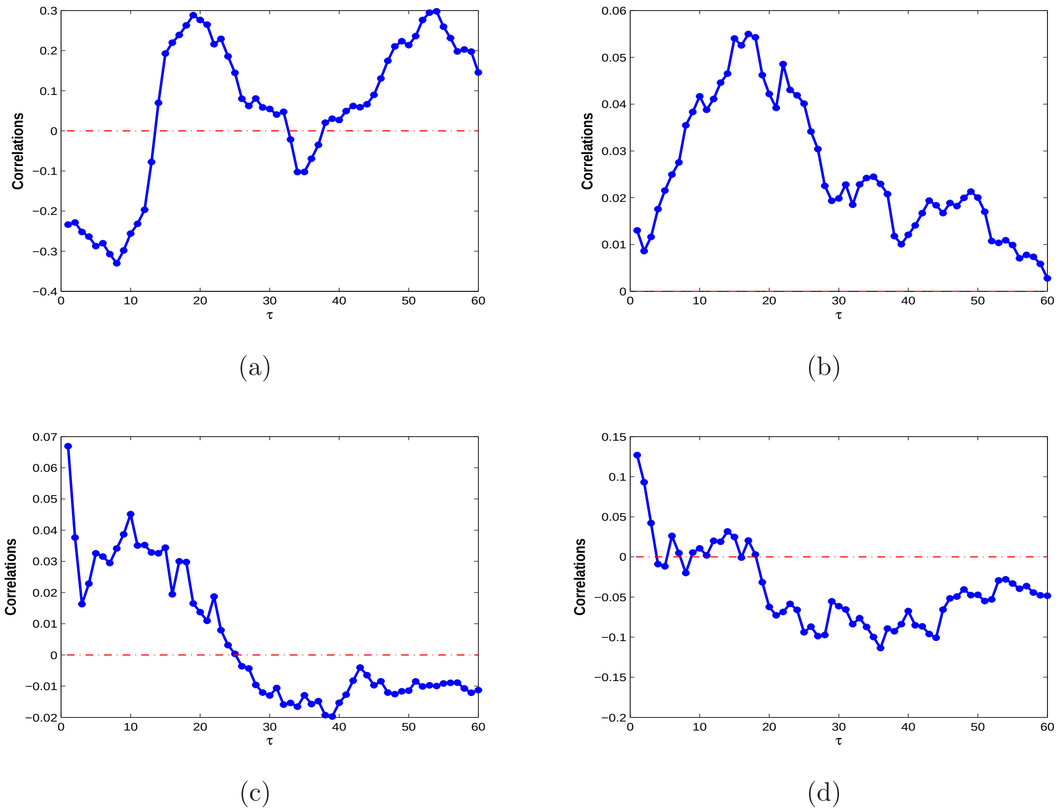



FIGURE F.6. The correlations of the estimated $\sigma_{S(1)}$ with (a) the estimated $\phi$ and the return of the optimal strategies for (b) the full model, (c) the pure momentum model and (d) the TSM return.

and hence the trading signal is significant only when the market price changes dramatically in a high volatility period. Fig. F.6 (c) and (d) show that the profitability of the optimal strategies for the pure momentum model and the TSM strategies are sensitive to the estimated market volatility. The return is positively (negatively) related to market volatility for short (long) time horizons. But Fig. F.6 (b) shows that the optimal strategies for the full model perform well even in a highly volatile market.

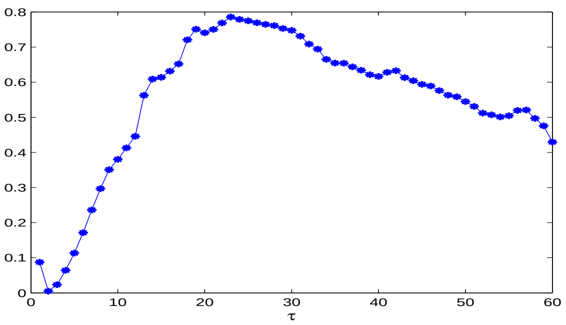



FIGURE F.7. The fraction of $\phi$ significantly different from zero for $\tau \in [1, 60]$.

We also study other time horizons. We find that the estimates of $\sigma_{S(1)}$, $\sigma_{X(1)}$ and $\sigma_{X(2)}$ are insensitive to $\tau$ but the estimates of $\phi$ are sensitive to $\tau$. Fig. F.7 illustrates the corresponding fraction of $\phi$, which is significantly different from zero for $\tau \in [1, 60]$. It shows that the momentums with 20–30 month horizons occur most frequently during the period from December 1890 until December 2012.

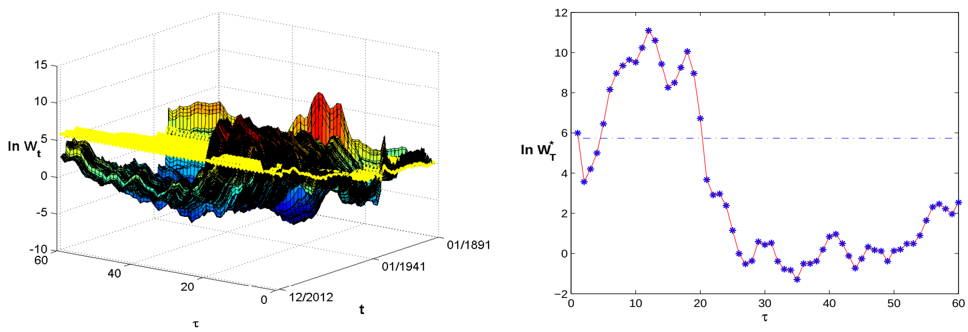



FIGURE F.8. The utility of wealth from December 1890 to December 2012 for the optimal portfolio with $\tau \in [1, 60]$ and the passive holding portfolio with 20-year rolling window estimated parameters.

Fig. F.8 (a) illustrates the utility of wealth from December 1890 until December 2012 for the optimal portfolio with $\tau \in [1, 60]$ and the passive holding portfolio. Especially, the utility of terminal wealth illustrated in Fig. F.8 (b) shows that the optimal strategies work well for short horizons $\tau \leq 20$ and the terminal utility reaches its peak at $\tau = 12$.

## References

Arriojas, M., Hu, Y., Mohammed, S. and Pap, G. (2007), 'A delayed Black and Scholes formula', *Stochastic Analysis and Applications* **25**, 471–492.

Asness, C., Moskowitz, T. and Pedersen, L. (2013), 'Value and momentum everywhere', *Journal of Finance* **68**, 929–985.

Baker, M. and Wurgler, J. (2006), 'Investor sentiment and the cross-section of stock returns', *Journal of Finance* **57**, 1645–1679.

Baker, M. and Wurgler, J. (2007), 'Investor sentiment in the stock makret', *Journal of Economic Perspectives* **21**, 129–151.

Balvers, R. and Wu, Y. (2006), 'Momentum and mean reversion across national equity markets', *Journal of Empirical Finance* **13**, 24–48.

Barberis, N. and Shleifer, A. (2003), 'Style investing', *Journal of Financial Economics* **68**, 161–199.

Barberis, N., Shleifer, A. and Vishny, R. (1998), 'A model of investor sentiment', *Journal of Financial Economics* **49**, 307–343.

Bender, C. and Denk, R. (2007), 'A forward scheme for backward SDEs', *Stochastic Processes and their Applications* **117**, 1793–1812.

Campbell, J. and Shiller, R. (1988*a*), 'The dividend-price ratio and expectations of future dividends and discount factors', *Review of Financial Studies* **1**, 195–228.

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

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

Chen, L. and Wu, Z. (2010), 'Maximum principle for the stochastic optimal control problem with delay and application', *Automatica* **46**, 1074–1080.

Chordia, T. and Shivakumar, L. (2002), 'Momentum, business cycle, and time-varying expected returns', *Journal of Finance* **57**, 985–1019.

Cohen, L. and Frazzini, A. (2008), 'Economic links and predictable returns', *Journal of Finance* **63**, 1977–2011.

Cohen, L. and Lou, D. (2012), 'Complicated firms', *Journal of Financial Economics* **104**, 383–400.

Cooper, M., Gutierrez, R. and Hameed, A. (2004), 'Market states and momentum', *Journal of Finance* **59**, 1345–1365.

Daniel, K., Hirshleifer, D. and Subrahmanyam, A. (1998), 'Investor psychology and investor security market under-and overreactions', *Journal of Finance* **53**, 1839–1886.

De Bondt, W. and Thaler, R. (1985), 'Does the stock market overreact?', *Journal of Finance* **40**, 793–808.

Delong, Ł. (2013), *Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications*, Springer, London.

Fama, E. (1991), 'Efficient capital markets: II', *Journal of Finance* **46**, 1575–1615.

Fama, E. and French, K. (1988), 'Permanent and temporary components of stock prices', *Journal of Political Economy* **96**, 246–273.

Fama, E. and French, K. (1992), 'The cross section of expected stock returns', *Journal of Finance* **47**, 427–465.

Fama, E. and French, K. (1996), 'Multifactor explanations of asset pricing anomalies', *Journal of Finance* **51**, 55–84.

Fama, E. and French, K. (1998), 'Value versus growth: The international evidence', *Journal of Finance* **53**, 1975–1999.

He, X. and Li, K. (2015), 'Profitability of time series momentum', *Journal of Banking and Finance* **53**, 140–157.

Hong, H. and Stein, J. (1999), 'A unified theory of underreaction, momentum trading, and overreaction in asset markets', *Journal of Finance* **54**, 2143–2184.

Hou, K., Peng, L. and Xiong, W. (2009), A tale of two anomalies: The implications of investor attention for price and earnings momentum, working paper, Ohio State University.

Huang, D., Jiang, F., Tu, J. and Zhou, G. (2013), Mean reversion, momentum and return predictability, working paper, Washington University in St. Louis.

Jegadeesh, N. (1991), 'Seasonality in stock price mean reversion: Evidence from the U.S. and U.K.', *Journal of Finance* **46**, 1427–1444.

Jegadeesh, N. and Titman, S. (1993), 'Returns to buying winners and selling losers: Implications for stock market efficiency', *Journal of Finance* **48**, 65–91.

Jegadeesh, N. and Titman, S. (2011), 'Momentum', *Annual Review of Financial Economics* **3**, 493–509.

Jobson, J. and Korkie, B. (1981), 'Performance hypothesis testing with the Sharpe and Treynor measures', *Journal of Finance* **36**, 889–908.

Koijen, R., Rodríguez, J. and Sbuelz, A. (2009), 'Momentum and mean reversion in strategic asset allocation', *Management Science* **55**, 1199–1213.

Lewellen, J. (2002), 'Mometnum and autocorrelation in stock returns', *Review of Financial Studies* **15**, 533–563.

Ma, J., Protter, P., Martin, J. and Torres, S. (2002), 'Numerical method for backward stochastic differential equations', *Annals of Applied Probability* **12**, 302–316.

Ma, J. and Yong, J. (1999), *Forward-Backward Stochastic Differential Equations and Their Applications*. in Lecture Notes in Mathematics: 1702, Springer, Berlin.

Marquering, W. and Verbeek, M. (2004), 'The economic value of predicting stock index returns and volatility', *Journal of Financial and Quantitative Analysis* **39**, 407–429.

Merton, R. (1969), 'Lifetime portfolio selection under uncertainty: The continuous-time case', *Review of Economics and Statistics* **51**, 247–257.

Merton, R. (1971), 'Optimum consumption and portfolio rules in a continuous time model', *Journal of Economic Theory* **3**, 373–413.

Moskowitz, T., Ooi, Y. H. and Pedersen, L. H. (2012), 'Time series momentum', *Journal of Financial Economics* **104**, 228–250.

Øksendal, B., Sulem, A. and Zhang, T. (2011), 'Optimal control of stochastic delay equations and time-advanced backward stochastic differential equations', *Advances in Applied Probability* **43**, 572–596.

Poterba, J. and Summers, L. (1988), 'Mean reversion in stock prices: Evidence and implications', *Journal of Financial Economics* **22**, 27–59.

Sagi, J. and Seasholes, M. (2007), 'Firm-specific attributes and the cross-section of momentum', *Journal of Financial Economics* **84**, 389–434.

Sangvinatsos, A. and Wachter, J. A. (2005), 'Does the failure of the expectations hypothesis matter for long-term investors?', *Journal of Finance* **60**, 179–230.

Serban, A. (2010), 'Combining mean reversion and momentum trading strategies in foreign exchange markets', *Journal of Banking and Finance* **34**, 2720–2727.

Vayanos, D. and Woolley, P. (2013), 'An institutional theory of momentum and reversal', *Review of Financial Studies* **26**, 1087–1145.

Wachter, J. (2002), 'Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets', *Journal of Financial and Quantitative Analysis* **37**, 63–91.

Wang, K. and Xu, J. (2012), Market volatility and momentum, working paper, University of Toronto.

Xia, Y. (2001), 'Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation', *Journal of Finance* **55**, 205–246.