REVERSING MOMENTUM: THE OPTIMAL DYNAMIC MOMENTUM STRATEGY

KAI LI[1] AND JUN LIU[2,3]

[1] Finance Discipline Group, UTS Business School
University of Technology Sydney, NSW 2007, Australia
[2] Rady School of Management
University of California San Diego, La Jolla, California 92093
[3] Shanghai Advanced Institute of Finance (SAIF)
Shanghai Jiao Tong University, Xuhui, Shanghai 200030, China
kai.li@uts.edu.au, junliu@ucsd.edu

Abstract. We study the optimal dynamic trading strategy between a riskless asset and a risky asset with momentum (momentum asset). The most salient characteristic of momentum is that positive price shocks predict positive future returns. This characteristic leads to big swings in returns over multiple periods. Investors with relative risk aversion greater than one dislike such big swings. We show that it is optimal for such investors to reverse momentum by holding less or even shorting the momentum asset. We find that the optimal portfolio weight also depends on the historical price path, in addition to momentum. Different historical price paths, even if they have the same momentum, lead to different optimal portfolio weights. In particular, with rebound path (a historical price path that decreases at the beginning and then rebounds later to have a positive momentum), it is optimal for investors to hold less or may short the momentum asset and hence suffer less or even benefit from momentum crashes.

Key words: Portfolio selection, momentum crashes, dynamic optimal momentum strategy.

JEL Classification: C32, G11

Date: March 11, 2016.

1. Introduction

This paper studies the optimal dynamic trading strategy between a riskless asset and a risky asset which has momentum (we term it momentum asset). Because future returns are predictable by past returns, momentum assets tend to have big swings in returns over multiple periods. Investors with relative risk aversion greater than one dislike these big swings. We show that it is optimal for such investors to reverse the momentum by holding less momentum asset over longer horizons. In fact, they may even short the asset with positive momentum, while the myopic strategy has long position. Effectively, the investors home-make their own asset with return reversal.

We find that the optimal portfolio weight also depends on historical price paths, not just on momentum which is determined by the beginning and end prices of the 'look-back period'. In general, there can be rebound paths that decrease at the beginning and then rebound later to have a positive momentum. Price paths can be also generally trending up to have a positive momentum. The optimal portfolio weights are different for different historical price paths (such as rebound and upward-trend price paths) even if the paths have the same momentum. In particular, the optimal strategy tends to have negative positions in the momentum asset with rebound paths, while has positive positions with upward-trend price paths. The myopic strategy widely used in academic literature and in practice only utilizes the momentum variable. Our paper shows that the optimal strategy also exploits the path dependence for a momentum asset, especially after sharp market rebounds.

Momentum is one of the most prominent empirical regularity in financial markets. Jegadeesh and Titman (1993) document cross-sectional momentum when 'look-back period' and 'holding period' are less than one year. Recently, Moskowitz, Ooi and Pedersen (2012) investigate time series momentum that characterizes strong positive predictability of a security's own past returns. The salient feature of momentum is the predictability of future return by the moving average of historical returns, which necessarily introduces time delays into price dynamics.

Stochastic processes with time delays are just recently studied in mathematics literature and are inherently non-Markovian. Stochastic control problems with time delays are quite involved because they give rise to infinite-dimensional non-Markovian systems, and the standard dynamic programming method cannot be used in this case. We solve the optimal dynamic portfolio strategy by applying the Cox and Huang (1989) approach. This approach makes the problem tractable. In particular, when horizon is shorter than the length of the look-back period, we can derive closed-form solutions.

Daniel and Moskowitz (2016) document momentum crashes after sharp market rebounds, which makes momentum strategy less appealing to risk-averse investors. They study optimal momentum portfolio by maximizing the Sharpe ratio to improve the performance of the standard momentum trading strategy. Their optimal portfolio weights are the mean-variance portfolio weights. Cujean and Hasler (2015) document a similar phenomenon in time series momentum. In these studies, the trading strategies are myopic. When the risky asset has positive momentum, the myopic portfolio weights are always positive because of the positive risk premium. Our dynamic optimal strategy suffers less or even benefits from the momentum crashes, because the dynamic portfolio weight is less than the myopic weight, and can be even negative for rebound price path. This is due to the fact that the momentum effect leads to long-lasting response of returns to a historical price shock, and makes the hedging demand strongly reacts to the historical price path, in addition to the momentum. Koijen, Rodríguez and Sbuelz (2009) study the optimal portfolio when the look-back period of momentum is infinite. In this case, the problem is much more tractable and the optimal weight becomes independent of historical price path.

The optimal portfolio weight also has other interesting features. For example, there are many bumps in the portfolio weight as a function of horizon, which is caused by the joint impact of momentum and the time-varying expected returns introduced by the path dependence. In contrast, the dynamic strategies with Markov state variables typically have monotonically smooth horizon dependence. In addition, the path dependence leads to big fluctuations in portfolio weights, which imply that market timing is important for momentum strategy.

The paper is organized as follows. We first provide an illustrative example in Section 2 to describe the intuition of reversing momentum. Section 3 discusses a formal model of momentum in continuous time. In Section 4, we analyze the return characteristics implied by the momentum model. The optimal portfolio selection problem is solved using the Cox-Huang approach in Section 5. Section 6 examines the properties of the optimal dynamic momentum strategy. More general momentum models are discussed in Section 7. Section 8 concludes. All the proofs are included in the appendices.

2. An Illustrative Example

This section discusses a two-period binomial-tree model of momentum to illustrate the intuition of reversing momentum in the paper. We study a financial market with two assets, a riskless asset with constant gross return R_f and a risky asset, whose gross return is characterized by a two-period binomial tree. The gross return over period 1 can be either U with probability P or D $(< U)$ with probability $1-P$. The return over period 2 is either $U+\Delta_S$ with probability P or $D+\Delta_S$ with probability $1-P$, where $S=U$ or D is the state at time 1. This setup keeps conditional volatility constant. When $\Delta_S \neq 0$, return becomes past-dependent. To model the momentum, we assume Δ_S is positive (negative) when $S=U$ (D). So the expected return over period 2 increases by Δ_U if period 1 realizes positive excess return, while decreases by Δ_D otherwise. The conditional volatility is a constant $P(1-P)(U-D)^2$ at each node of the tree.

The optimization problem for an investor with expected CRRA utility over terminal wealth at time 2 is given by

$$\max_{\phi_0,\phi_1} \mathbb{E}_0\left[\frac{\tilde{W}_2^{1-\gamma}}{1-\gamma}\right] = \max_{\phi_0,\phi_1} \mathbb{E}_0\left[\frac{\left(W_0\tilde{R}_1^p\tilde{R}_2^p\right)^{1-\gamma}}{1-\gamma}\right], \tag{2.1}$$

where ϕ is the portfolio weight invested in the momentum asset, $\tilde{W}$ is the wealth, $\tilde{R}^p$ is the portfolio return and $\gamma > 1$ is the constant relative risk aversion coefficient. Backward deduction implies that the above problem is equivalent to

$$\begin{aligned}
&\max_{\phi_0} \mathbb{E}_0\left[\frac{W_0^{1-\gamma}}{1-\gamma}(\tilde{R}_1^p)^{1-\gamma}\max_{\phi_1}\mathbb{E}_1\left[(\tilde{R}_2^p)^{1-\gamma}\right]\right] \\
=&\max_{\phi_0} \mathbb{E}_0\left[\frac{W_0^{1-\gamma}}{1-\gamma}(\tilde{R}_1^p)^{1-\gamma}\mathbb{E}_1\left[(\tilde{R}_2^*)^{1-\gamma}\right]\right],
\end{aligned} \tag{2.2}$$

where $\tilde{R}_2^*$ is the optimal portfolio return over period 2. By defining $\tilde{\varsigma} = \mathbb{E}_1\left[(\tilde{R}_2^*)^{1-\gamma}\right]$, (2.2) becomes

$$\max_{\phi_0} \mathbb{E}_0\left[\frac{W_0^{1-\gamma}}{1-\gamma}(\tilde{R}_1^p)^{1-\gamma}\tilde{\varsigma}\right]. \tag{2.3}$$

To rewrite (2.3) in terms of a standard portfolio problem, we need to define a new probability to eliminate $\tilde{\varsigma}$,

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \tilde{\varsigma}.$$

Then the original dynamic portfolio selection problem under the physical measure becomes a myopic problem under the new measure $\mathbb{P}^*$,

$$\max_{\phi_0} \mathbb{E}^*\left[\frac{W_0^{1-\gamma}}{1-\gamma}(\tilde{R}_1^p)^{1-\gamma}\right].$$

We choose Δ_S to guarantee both no arbitrage (i.e., $D+\Delta_S < R_f < U+\Delta_S$) and positive risk premium, so that any negative demand is not caused by negative risk premium. Appendix A.1 shows that the up state probability P^* under the new measure is smaller than P, and decreases as $\Delta_U - \Delta_D$ increases. This reduces the optimal stock position at time 0 comparing with the myopic strategy which only cares about the utility one period ahead. When $\Delta_U - \Delta_D$ is big enough, we have $\mathbb{E}^*[\tilde{R}_1] = P^*U + (1-P^*)D < R_f$, which is equivalent to $\phi_0 < 0$, a negative optimal demand at time 0.

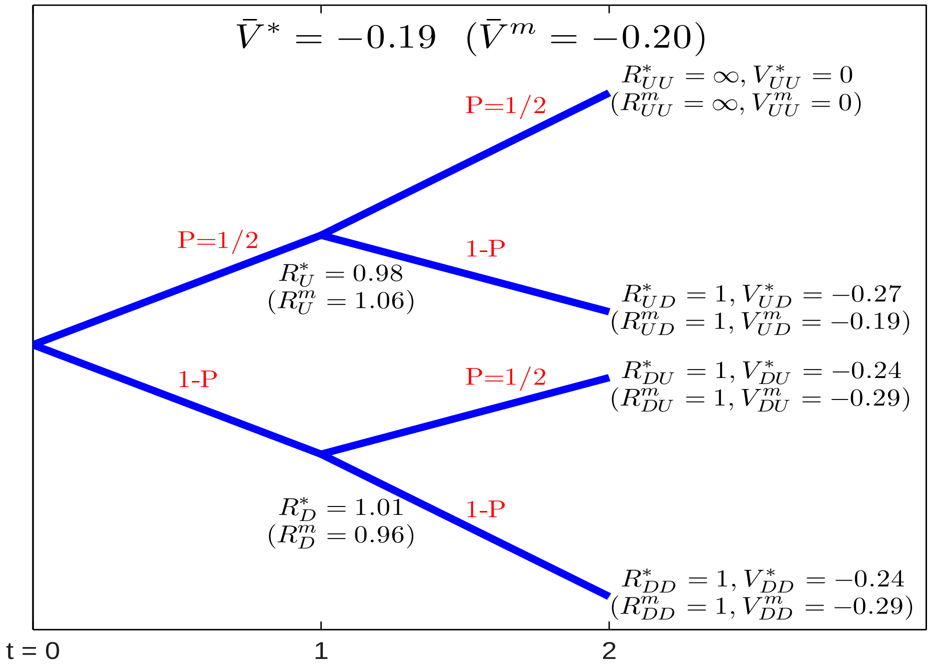


FIGURE 2.1. The portfolio returns for the optimal strategy (R^*) and the myopic strategy (R^m); and the terminal utilities for the optimal strategy (V^*) and the myopic strategy (V^m) at each market state. Here $R_f = 1$, $U = 1.5$, $D = 0.7$, $P = 0.5$, $\gamma = 5$, $W_0 = 1$, $\Delta_U = \overline{\Delta} = 0.3$ and $\Delta_D = \underline{\Delta} = -0.1$. At time 0, the optimal demand $\phi_0^* = -0.05 < 0$, while the myopic demand $\phi_0^m = 0.13 > 0$. The expected terminal utilities are $\bar{V}^* = -0.19$ and $\bar{V}^m = -0.20$ $(< \bar{V}^*)$ for the optimal and myopic strategies respectively.

Fig. 2.1 illustrates an example where the optimal position in momentum asset at time 0 is negative, while a myopic strategy always holds positive position whenever the risk premium is positive. For the optimal strategy, the short position at time 0 leads to smaller (greater) portfolio return at state U (D) over period 1 relative to the myopic strategy. The two strategies have the same returns at each state over period 2. Because $\tilde{\varsigma}$ reduces the 'probability' of up state, the expected terminal utility for the optimal strategy is, however, greater than that for the myopic strategy.

We complete this section with the following remark. When $\Delta_S = 0$, the risky asset has a standard i.i.d return process. Then $P^* = P$ and the optimal strategy always takes long position in the risky asset. Therefore, the short position in the risky asset illustrated in Fig. 2.1 is caused by the momentum $\Delta_S \neq 0$.

3. A Continuous-time Model of Momentum

In this section, we specify the price dynamics of the momentum asset. The uncertainty is represented by a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0})$, on which a one-dimensional Brownian motion B_t is defined. The price of the risky momentum asset at time t satisfies

$$\frac{dS_t}{S_t} = \big[\alpha m_t + (1-\alpha)\mu + r\big]dt + \sigma dB_t, \tag{3.1}$$

where m_t is the momentum variable, r is the short rate which is assumed to be constant, μ is a constant which can be shown later to be the average risk premium, and α measures the fraction of momentum in the expected returns. When $\alpha = 0$, the stock price (3.1) reduces to a standard geometric Brownian motion.

The time series momentum across different asset classes and markets documented in Moskowitz et al. (2012) shows that "*the past 12-month excess return of each instrument is a positive predictor of its future return.*"[1] Accordingly, the momentum variable m_t is defined as an equally-weighted moving average (MA) of historical excess returns over a past time interval $[t-\tau, t]$,

$$m_t = \frac{1}{\tau}\int_{t-\tau}^{t}\Big(\frac{dS_u}{S_u} - r du\Big), \tag{3.2}$$

where $\tau \geq 0$ is the 'look-back period' of the momentum. So m_t is determined by $\ln S_t - \ln S_{t-\tau}$. The equally-weighted MA (3.2) is mostly used in practice. For example, Neely, Rapach, Tu and Zhou (2014) show that this MA indicator displays statistically and economically significant predictive power to the equity risk premium. We focus on this momentum variable in our paper. We will also discuss other types of MA later in Section 7.

When $\tau = 0$, the momentum becomes the current rate of excess return, $m_t dt = dS_t/S_t - rdt$, and (3.1) reduces to the price with a standard geometric Brownian motion type. When $\tau = dt$, then the momentum variable becomes the last period excess return and hence stock return in (3.1) becomes a first order autoregressive (AR(1)) process. DeMiguel, Nogales and Uppal (2014) find that, by exploiting serial dependence, the arbitrage portfolios based on a first order autoregressive return model attains positive out-of-sample returns after adjusting for transaction costs. The greater the look-back period τ is, the less volatile the momentum variable is.

[1]For a large set of futures and forward contracts, Moskowitz et al. (2012) provide strong evidence for time series momentum based on the moving average of look-back excess returns. This effect based purely on a security's own past returns is related to, but different from, the cross-sectional momentum. Through return decomposition, Moskowitz et al. (2012) show that positive autocovariance between a security's excess return next month and it's lagged 1-year return is the main driving force for both time series momentum and cross-sectional momentum.

In all, our model not only includes the autoregressive models as special cases, but also can well capture the time series momentum effect documented in the literature. We will show later that a more general form of (3.2) can also include the mean-reverting Ornstein-Uhlenbeck process as its special case.

The MA of historical returns (3.2) uses latest past information and necessarily introduces into price dynamics time delays,[2] an inherent non-Markovian feature. The resulting asset price model (3.1)-(3.2) is path-dependent and is characterized by a non-Markovian system of stochastic delay differential equations (SDDEs), which is relatively new to the finance literature. Let $C([-\tau, 0], R)$ be the space of all continuous functions $\varphi : [-\tau, 0] \to R$. The following proposition shows that, for a given initial condition $S_t = \varphi_t$, $t \in [-\tau, 0]$, the system (3.1)-(3.2) admits a unique solution such that $S_t > 0$ almost surely for all $t \geq 0$ whenever $\varphi_t > 0$ for $t \in [-\tau, 0]$ almost surely.

Lemma 3.1. *The system* (3.1)-(3.2) *has an almost surely continuously adapted unique solution S 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.*

Two observations follow Lemma 3.1. First, although (3.2) implies that

$$m_t = \frac{1}{\tau}(\ln S_t - \ln S_{t-\tau}) - r + \frac{\sigma^2}{2}$$

only depends on two prices at time t and $t - \tau$ respectively, Lemma 3.1 states that, to define the price process, we need the whole path of prices during $[t - \tau, t]$. This is because the historical price S_u for $u \in (t - \tau, t)$ will be used to determine the future price at time $u + \tau$ $(> t)$. As time increases from t to $t + \tau$, all the historical prices during $[t - \tau, t]$ will be used successively. After this period, the prices over $[t, t + \tau]$ then become realized and will be used to form the prices over $[t + \tau, t + 2\tau]$... We will show later that the path-dependent feature is important for optimal dynamic momentum strategy. Second, notice $C([-\tau, 0], R)$ is an infinite-dimensional space of initial conditions. So Lemma 3.1 shows that the system (3.1)-(3.2) also has infinite dimensions. The corresponding portfolio selection problem is conceptually much more difficult than in the no-delay case, which has finite dimensions.

Although the continuous-time processes with time delays are relatively new in the theoretical finance literature, the MA rules have been widely used empirically. In addition to the papers cited above, various MA indicators are widely used among practitioners (Schwager, 1989 and Lo and Hasanhodzic, 2010), and have significant forecasting powers to equity risk premium. Brock, Lakonishok and LeBaron (1992)

[2]This is due to the fact that the lower limit of the integral, or the low boundary of the MA, is a function of time t.

find strong evidence of profitability of MA trading rules for Dow Jones Index. Zhu and Zhou (2009) demonstrate that, when stock returns are predictable or when parameter (or model) uncertainty exists, MA trading rules can well exploit the serial correlations of returns and hence significantly improve the portfolio performance.

4. Return Characteristics of the Momentum Model

In this section, we examine the return characteristics implied by the momentum model. Define $s_t = \ln S_t$. Notice that the solutions to (3.1) are given piecewisely as demonstrated in Appendix A.2. The expected values and variances of stock returns, and hence the Sharpe ratios, should also have different forms in different time intervals with length of τ. Proposition 4.1 confirms the conjectures.

Proposition 4.1. *For $\varpi \in [n\tau, (n+1)\tau]$, $n = 0, 1, 2, \cdots$, the cumulative returns of the stock over $[t, t+\varpi]$ are given by*

$$\begin{aligned} s_{t+\varpi} - s_t =& \frac{\tau}{\alpha}(1-\alpha)\Big(r+\mu-\frac{\sigma^2}{2}\Big)\Big[\sum_{i=0}^{n}\Big(\sum_{j=0}^{i}\frac{(-\frac{\alpha}{\tau})^j(\varpi-i\tau)^j}{j!}\Big)e^{\frac{\alpha}{\tau}(\varpi-i\tau)}-n-1\Big] \\ &+\Big[\sum_{i=0}^{n}\frac{(-\frac{\alpha}{\tau})^i(\varpi-i\tau)^i}{i!}e^{\frac{\alpha}{\tau}(\varpi-i\tau)}-1\Big]s_t \\ &-\frac{\alpha}{\tau}\int_{-\tau}^{0}\Big[\sum_{i=1}^{n}\frac{(-\frac{\alpha}{\tau})^{i-1}(\varpi-i\tau-u-t)^{i-1}}{(i-1)!}e^{\frac{\alpha}{\tau}(\varpi-i\tau-u-t)}\Big]s_{t+u}du \\ &-\frac{\alpha}{\tau}\int_{-\tau}^{\varpi-(n+1)\tau}\Big[\frac{(-\frac{\alpha}{\tau})^n[\varpi-(n+1)\tau-u-t]^n}{n!}e^{\frac{\alpha}{\tau}[\varpi-(n+1)\tau-u-t]}\Big]s_{t+u}du \\ &+\sigma\sum_{i=0}^{n}\int_{0}^{\varpi-i\tau}\frac{(-\frac{\alpha}{\tau})^i(\varpi-i\tau-u-t)^i}{i!}e^{\frac{\alpha}{\tau}(\varpi-i\tau-u-t)}dB_{t+u}, \end{aligned} \tag{4.1}$$

and the conditional mean and variance of the cumulative returns are given, respectively, by

$$\begin{aligned} \mathbb{E}_t\big[s_{t+\varpi} - s_t\big] =& \frac{\tau}{\alpha}(1-\alpha)\Big(r+\mu-\frac{\sigma^2}{2}\Big)\Big[\sum_{i=0}^{n}\Big(\sum_{j=0}^{i}\frac{(-\frac{\alpha}{\tau})^j(\varpi-i\tau)^j}{j!}\Big)e^{\frac{\alpha}{\tau}(\varpi-i\tau)}-n-1\Big] \\ &+\Big[\sum_{i=0}^{n}\frac{(-\frac{\alpha}{\tau})^i(\varpi-i\tau)^i}{i!}e^{\frac{\alpha}{\tau}(\varpi-i\tau)}-1\Big]s_t \\ &-\frac{\alpha}{\tau}\int_{-\tau}^{0}\Big[\sum_{i=1}^{n}\frac{(-\frac{\alpha}{\tau})^{i-1}(\varpi-i\tau-u-t)^{i-1}}{(i-1)!}e^{\frac{\alpha}{\tau}(\varpi-i\tau-u-t)}\Big]s_{t+u}du \\ &-\frac{\alpha}{\tau}\int_{-\tau}^{\varpi-(n+1)\tau}\Big[\frac{(-\frac{\alpha}{\tau})^n[\varpi-(n+1)\tau-u-t]^n}{n!}e^{\frac{\alpha}{\tau}[\varpi-(n+1)\tau-u-t]}\Big]s_{t+u}du, \end{aligned} \tag{4.2}$$

and

$$\begin{aligned}\mathbb{V}ar_t\big[s_{t+\varpi}-s_t\big] =&\sigma^2\bigg[\int_{-\tau}^{0} e^{-\frac{2\alpha}{\tau}u}du+\int_{-2\tau}^{-\tau}\bigg(\sum_{i=0}^{1}\frac{(-\frac{\alpha}{\tau})^i(-i\tau-u)^i}{i!}e^{\frac{\alpha}{\tau}(-i\tau-u)}\bigg)^2du\\ &+\cdots\\ &+\int_{-n\tau}^{-(n-1)\tau}\bigg(\sum_{i=0}^{n-1}\frac{(-\frac{\alpha}{\tau})^i(-i\tau-u)^i}{i!}e^{\frac{\alpha}{\tau}(-i\tau-u)}\bigg)^2du\\ &\int_{-\varpi}^{-n\tau}\bigg(\sum_{i=0}^{n}\frac{(-\frac{\alpha}{\tau})^i(-i\tau-u)^i}{i!}e^{\frac{\alpha}{\tau}(-i\tau-u)}\bigg)^2du\bigg].\end{aligned} \tag{4.3}$$

There are several observations from Proposition 4.1. First, the stock returns over $[t, t+\varpi]$ in (4.1) are given piecewisely. In (4.1), the first term is a deterministic function of horizon ϖ; the second, third and fourth terms are weighted sum of the historical prices s_u over $u \in [t-\tau, t]$. The last term is a weighted sum of the innovations dB_u for $u \in [t, t+\varpi)$ with the weights decreasing with u. Therefore, Proposition 4.1 states that the return process crucially depends on historical price realizations, not just on the beginning and end prices of the look-back period.

Second, the weights on all initial values s_u, $u \in [t-\tau, t]$ in the second, third and fourth terms of (4.1) sum up to zero. Therefore, the price level does not affect the returns of momentum asset. In particular, when the historical prices are chosen as the same constant value $\bar{s}$ (i.e., $s_u = \bar{s}$ for $u \in [t-\tau, t]$), then the expected return reduces to

$$\mathbb{E}_t\big[s_{t+\varpi}-s_t\big] = \frac{\tau}{\alpha}(1-\alpha)\bigg(r+\mu-\frac{\sigma^2}{2}\bigg)\bigg[\sum_{i=0}^{n}\bigg(\sum_{j=0}^{i}\frac{(-\frac{\alpha}{\tau})^j(\varpi-i\tau)^j}{j!}\bigg)e^{\frac{\alpha}{\tau}(\varpi-i\tau)}-n-1\bigg],$$

for $\varpi \in [n\tau, (n+1)\tau]$, $n = 0, 1, 2, \cdots$.

Third, when $\alpha = 0$, the returns in (3.1) becomes an i.i.d. process $dS_t/S_t = (\mu + r)dt + \sigma dB_t$. Proposition 4.1 implies

Corollary 4.2. *When* $\alpha = 0$*,*

$$\begin{aligned}&s_{t+\varpi}-s_t = \Big(\mu+r-\frac{\sigma^2}{2}\Big)\varpi+\sigma B_\varpi,\\ &\mathbb{E}_t\big[s_{t+\varpi}-s_t\big] = \Big(\mu+r-\frac{\sigma^2}{2}\Big)\varpi,\\ &\mathbb{V}ar_t\big[s_{t+\varpi}-s_t\big] = \sigma^2\varpi,\\ &\textit{Sharpe ratio} = \Big(\frac{\mu}{\sigma}-\frac{\sigma}{2}\Big)\sqrt{\varpi}.\end{aligned} \tag{4.4}$$

Fourth, we look at the case of $\varpi \le \tau$ in the following corollary and then numerically examine the case for $\varpi > \tau$.

Corollary 4.3. *For $\varpi \le \tau$,*

$$\begin{aligned} s_{t+\varpi} - s_t &= \frac{\tau}{\alpha}(1-\alpha)\left(r+\mu-\frac{\sigma^2}{2}\right)\left(e^{\frac{\alpha}{\tau}\varpi}-1\right) + \left(e^{\frac{\alpha}{\tau}\varpi}-1\right)s_t \\ &\quad - \frac{\alpha}{\tau}\int_{-\tau}^{\varpi-\tau} e^{\frac{\alpha}{\tau}(\varpi-\tau-u-t)} s_{u+t}du + \sigma\int_0^{\varpi} e^{\frac{\alpha}{\tau}(\varpi-u-t)}dB_{t+u}, \\ \mathbb{E}_t\left[s_{t+\varpi}-s_t\right] &= \frac{\tau}{\alpha}(1-\alpha)\left(r+\mu-\frac{\sigma^2}{2}\right)\left(e^{\frac{\alpha}{\tau}\varpi}-1\right) + \left(e^{\frac{\alpha}{\tau}\varpi}-1\right)s_t \\ &\quad - \frac{\alpha}{\tau}\int_{-\tau}^{\varpi-\tau} e^{\frac{\alpha}{\tau}(\varpi-\tau-u-t)} s_{u+t}du, \\ \mathbb{V}ar_t\left[s_{t+\varpi}-s_t\right] &= \frac{\sigma^2\tau}{2\alpha}\left(e^{\frac{2\alpha}{\tau}\varpi}-1\right). \end{aligned} \tag{4.5}$$

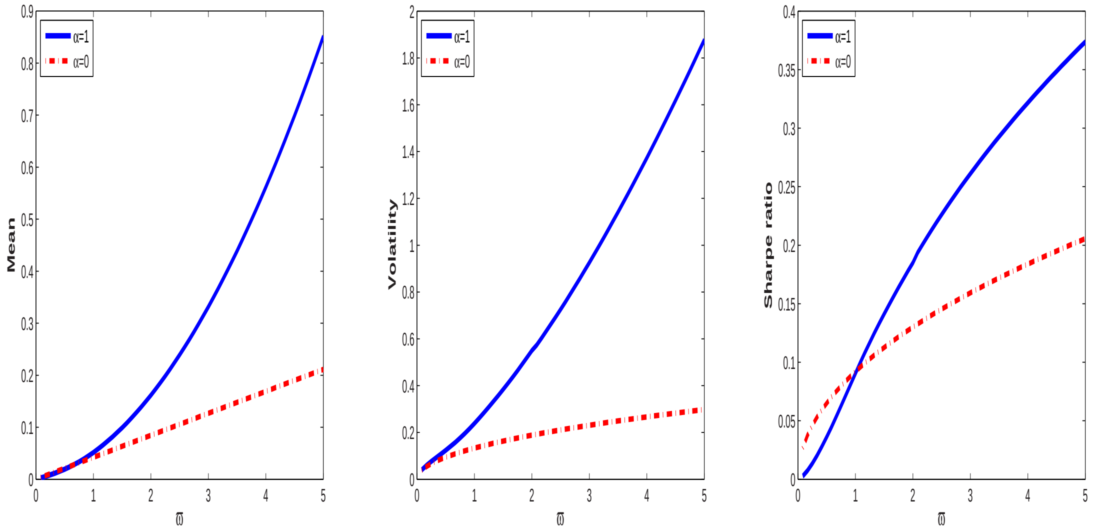
Especially, when the initial values are chosen as the same constant value $\bar{s}$ (i.e., $s_u = \bar{s}$ for $u \in [t-\tau, t]$), the expected cumulative return and the Sharpe ratio are given, respectively, by

$$\begin{aligned} &\mathbb{E}_t\left[s_{t+\varpi}-s_t\right] = \frac{\tau}{\alpha}(1-\alpha)\left(r+\mu-\frac{\sigma^2}{2}\right)\left(e^{\frac{\alpha}{\tau}\varpi}-1\right), \\ &\textit{Sharpe ratio} = (1-\alpha)\left(r+\mu-\frac{\sigma^2}{2}\right)\frac{\sqrt{2\tau}}{\sigma\sqrt{\alpha}}\sqrt{\frac{e^{\frac{\alpha}{\tau}\varpi}-1}{e^{\frac{\alpha}{\tau}\varpi}+1}} - \frac{\sqrt{2\alpha}r\varpi}{\sigma\sqrt{\tau(e^{\frac{\alpha}{\tau}\varpi}-1)}}. \end{aligned} \tag{4.6}$$

Interestingly, the Sharpe ratio in (4.6) depends on the riskless rate r. This is different from the standard geometric Brownian motion case in Corollary 4.2. By comparing the first and the second equalities in (A.13), we can see that the riskless rate in the Sharpe ratio is introduced by the momentum variable m_t, which is defined as a moving average of historical excess returns.

The expressions are more involved for the case $\varpi > \tau$ in Proposition 4.1. In order to examine the tradeoff between payoff and risk over longer time horizons, we numerically study how the expectations and variances of the cumulative returns and the Sharpe ratios evolve with respect to the horizon ϖ. We set the parameters according to the calibrations described in Section 6.1, and the historical prices are chosen as the same constant number $s_u = \bar{s}$ for $u \in [t-\tau, t]$. Fig. 4.1 illustrates the mean values and standard deviations of returns and the Sharpe ratios over a five-year horizon. Although the means and variances of the returns and the Sharpe ratios in Proposition 4.1 are given piecewisely, they are continuous in time as illustrated in Fig. 4.1.

To exploit the impact of momentum, we examine two values of the momentum fraction parameter: $\alpha = 1$ (the blue solid line) and $\alpha = 0$ (the red dash-dotted line). When $\alpha = 0$, the stock process (3.1) reduces to a standard geometric Brownian motion, and the mean and variance become linear functions of horizon ϖ. Fig. 4.1 shows that both the mean value and standard deviation of the returns of momentum



(a) Term structure of expected returns (b) Term structure of volatility (c) Term structure of Sharpe ratios

FIGURE 4.1. (a) The mean value $\bar{R}_\varpi = \mathbb{E}_t\big[s_{t+\varpi} - s_t\big]$ and (b) standard deviation $\sigma(R_\varpi) = \mathbb{S}\mathrm{td}_t\big[s_{t+\varpi} - s_t\big]$ of cumulative returns and (c) the Sharpe ratios $\mathrm{SR}_\varpi = (\bar{R}_\varpi - r)/\sigma(R_\varpi)$ conditional on a constant historical price path $s_u = \bar{s}$ for $u \in [t-\tau, t]$) as functions of horizon ϖ. Here $\tau = 1$, $r = 3\%$, $\mu = 2.38\%$, $\sigma = 13.3\%$ and the parameter of momentum fraction $\alpha = 0$ or 1.

asset are convex functions of horizon ϖ. The greater α is, the more convex they are and the greater their growth rates are. The Sharpe ratios for the momentum asset is smaller (greater) than in the standard geometric Brownian motion case for short (long) horizons ϖ.

Notice that the historical price path can also affect the mean value and Sharpe ratio, while cannot affect the variance. For example, if there is an increasing (decreasing) pattern in the historical path, then the growth rate of the mean value becomes bigger (smaller) because the weights on more recent past returns are relatively bigger in (4.1), implying that the initial trend has a positive impact on expected return and hence Sharpe ratio. This path effect on portfolio selection will be further explored in next two sections.

Corollary 4.4. *For $\varpi \in [n\tau, (n+1)\tau]$, $n = 0, 1, 2, \cdots$, the impulse-response function for the log price and return are given, respectively, by*

$$\mathcal{D}_t[s_{t+\varpi}] = \sigma \sum_{i=0}^{n} \frac{\big(-\frac{\alpha}{\tau}\big)^i (\varpi - i\tau)^i}{i!} e^{\frac{\alpha}{\tau}(\varpi - i\tau)}, \tag{4.7}$$

and

$$\mathcal{D}_t[ds_{t+\varpi}] = \sigma\Bigg[\sum_{i=1}^{n}\frac{\big(-\frac{\alpha}{\tau}\big)^{i}(\varpi-i\tau)^{i-1}}{(i-1)!}e^{\frac{\alpha}{\tau}(\varpi-i\tau)} - \sum_{i=0}^{n}\frac{\big(-\frac{\alpha}{\tau}\big)^{i+1}(\varpi-i\tau)^{i}}{i!}e^{\frac{\alpha}{\tau}(\varpi-i\tau)}\Bigg]dt. \tag{4.8}$$

Finally, to further exploit the path dependence property, we examine the response of returns to an innovation. The impulse-response function can be examined via the Malliavin derivatives (Detemple, Garcia and Rindisbacher, 2003). Corollary 4.4 states that the cumulative returns of momentum asset react to an initial shock piecewisely. For long horizons ϖ, the response decreases in horizon. However it is easy to check that the returns have no response to the shock when replacing m_t by an Ornstein-Uhlenbeck process, which is frequently used to model time-varying risk premium in the finance literature. This long-lasting response to a historical price shock is inherent with momentum, and we will see later that it generates a new type of hedging demand associated with the historical price path.

5. The Optimal Dynamic Momentum Strategy

We study the optimal dynamic trading strategy for an investor with expected utility over terminal wealth at time T and constant relative risk aversion $\gamma > 0$. The optimization problem of the investor is given by

$$\sup_{(\phi_t)_{t\in[0,T]}} \mathbb{E}_0\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right], \tag{5.1}$$

where ϕ_t is the fraction of wealth invested in the risky momentum asset.

Two approaches are most frequently used to solve the stochastic control problems: the dynamic programming method and the maximum principle (Yong and Zhou, 1999). Since the stochastic delay differential equations (SDDEs) are not Markovian, the dynamic programming method results in an infinite dimensional partial differential equation, which is difficult to be solved even numerically.[3] However, the maximum principle for the optimal control problem of SDDEs results in a full-coupled forward-backward stochastic delay differential system (Chen and Wu, 2010), and currently no algorithm exists for solving it numerically.

In this paper, we solve the optimization problem using the Cox and Huang (1989, 1991) approach, which is originated from finance literature and can be applied to the

[3]Larssen and Risebro (2001) show that the stochastic control problem for SDDEs can be reduced to a finite dimensional problem under the special conditions that time delays do not appear in both control variables and the value function, and parameters are also required to satisfy certain equalities (Theorem 5.1 in Larssen and Risebro, 2001). Their methods cannot be applied to the portfolio selection problems with past-dependent underlying stock processes, because in this case, the time delays affect the control variables.

non-Markov price models. The market is complete, and then there exists a unique state price density, which is given by

$$\pi_t = \exp\Big\{ -\int_0^t r du - \frac{1}{2}\int_0^t \theta_u^2 du - \int_0^t \theta_u dB_u \Big\}, \tag{5.2}$$

where

$$\theta_t = \frac{\alpha m_t + (1-\alpha)\mu}{\sigma}, \tag{5.3}$$

is the market price of risk. The process π_t can be interpreted as a system of Arrow-Debreu prices. Because θ_t is path-dependent, the price of a dollar at time t in each state is affected by the historical price path over $[t-\tau, t]$. The standard Cox-Huang approach leads to $W_T^* = (y\pi_T)^{-1/\gamma}$, where y is the Lagrange multiplier. Define

$$\xi_t = \exp\Big\{ -\frac{1}{2}\int_0^t \theta_u^2 du - \int_0^t \theta_u dB_u \Big\},$$

which is a martingale. Let $\varpi = T - t$ be the investment horizon and $\bar{\xi}_0 = \mathbb{E}_0\big[\xi_T^{(\gamma-1)/\gamma}\big]$. The following proposition provides the general results on the optimal dynamic momentum strategy and the value function.

Proposition 5.1. *For an investor with an investment horizon $\varpi = T - t$ and constant coefficient of relative risk aversion γ, the optimal wealth fraction invested in the risky asset is given by*

$$\phi_t^* = \frac{\alpha m_t + (1-\alpha)\mu}{\sigma^2} + \frac{\psi_t}{\sigma \pi_t W_t^*}, \tag{5.4}$$

where ψ_t is governed by

$$\pi_t W_t^* = W_0 + \int_0^t \psi_u dB_u. \tag{5.5}$$

and the remainder, $1-\phi_t^*$, is invested in the cash account. The corresponding optimal wealth process satisfies

$$W_t^* = W_0 \bar{\xi}_0^{-1} \pi_t^{-1} \mathbb{E}_t\Big[\xi_T^{(\gamma-1)/\gamma}\Big], \tag{5.6}$$

and the value function satisfies

$$V = \frac{1}{1-\gamma} W_0^{1-\gamma} \bar{\xi}_0^{\gamma} e^{(1-\gamma)rT}. \tag{5.7}$$

We show the details of the Cox-Huang approach in Appendix A.4. In order to derive the optimal portfolio weight, we need to compute

$$\mathbb{E}_t[\xi_T^{\frac{\gamma-1}{\gamma}}] = \xi_t^{\frac{\gamma-1}{\gamma}} \mathbb{E}_t\Big[\exp\Big\{ -\frac{\gamma-1}{\gamma}\Big(\int_t^T \theta_u dB_u + \frac{1}{2}\int_t^T \theta_u^2 du\Big)\Big\}\Big], \tag{5.8}$$

where the market price of risk θ_t is path-dependent. Unlike the Markovian system, we cannot apply the Feynman-Kac formula to the infinite dimensional SDDEs system in general. Due to the path dependence, the solution has to be given piecewisely.

In the following analysis, we mainly focus on the case when investment horizon is shorter than the length of the look-back period $0 \leq \varpi \leq \tau$. Subsection 5.1 provides closed-form solutions in this case. This investment problem with investment horizon shorter than look-back period are more important than with longer horizons for the following three reasons. First, for investment horizons longer than 1 year, returns observed in the data have reversals (Fama and French, 1988 and Poterba and Summers, 1988), which is not modelled in this paper. Second, in practice, momentum strategies are implemented only for holding periods shorter than 1 year (Jegadeesh and Titman, 1993, and Moskowitz et al., 2012). Third, the optimization problem for the case $\varpi > \tau$ is much more involved technically, however, we can solve it numerically.

5.1. **Closed-Form Solutions.**

Corollary 5.2. *When $0 \leq \varpi \leq \tau$, the optimal wealth fraction invested in the risky asset is given by*

$$\phi_t^* = \phi_t^m + \phi_t^{MH} + \phi_t^{PH}, \tag{5.9}$$

where

$$\begin{aligned} \phi_t^m &= \frac{\alpha m_t + \mu(1-\alpha)}{\gamma\sigma^2}, \\ \phi_t^{MH} &= \tau A_{1,\varpi} m_t, \\ \phi_t^{PH} &= A_{2,\varpi} + \Big(s_{t-\tau} + r\tau - \frac{\sigma^2\tau}{2}\Big) A_{1,\varpi}, \end{aligned} \tag{5.10}$$

and the corresponding optimal wealth process is given by

$$W_t^* = W_0 \bar{\xi}_0^{-1} e^{rt} \xi_t^{-1/\gamma} \exp\Big\{\frac{A_{1,\varpi}}{2} s_t^2 + A_{2,\varpi} s_t + A_{3,\varpi}\Big\}, \tag{5.11}$$

where

$$\begin{aligned} A_{1,\varpi} &= \frac{\alpha(\gamma-1)\big(1-e^{\frac{2\alpha}{\sqrt{\gamma}\tau}\varpi}\big)}{\gamma\sigma^2\tau\Big[(\sqrt{\gamma}-1)e^{\frac{2\alpha}{\sqrt{\gamma}\tau}\varpi} + (\sqrt{\gamma}+1)\Big]}, \\ A_{2,\varpi} &= \int_0^{\varpi} e^{\int_u^{\varpi}(\sigma^2 A_{1,v} + \frac{\alpha}{\gamma\tau})dv} \Bigg[\bigg(\Big(r - \frac{\sigma^2}{2}\Big)\Big(1-\frac{\alpha}{\gamma}\Big) + (1-\alpha)\frac{\mu}{\gamma} - \frac{\alpha}{\gamma\tau} s_{T-\tau-u}\bigg) A_{1,u} \\ &\quad + \frac{1-\gamma}{\gamma^2}\frac{\alpha}{\sigma^2\tau}\bigg((1-\alpha)\mu - \alpha\Big(r - \frac{\sigma^2}{2}\Big) - \frac{\alpha}{\tau} s_{T-\tau-u}\bigg)\Bigg] du, \\ A_{3,\varpi} &= \int_0^{\varpi} \Bigg[\frac{\sigma^2}{2} A_{2,u}^2 + \frac{\sigma^2}{2} A_{1,u} + \bigg(\Big(r - \frac{\sigma^2}{2}\Big)\Big(1-\frac{\alpha}{\gamma}\Big) + (1-\alpha)\frac{\mu}{\gamma} - \frac{\alpha}{\gamma\tau} s_{T-\tau-u}\bigg) A_{2,u} \\ &\quad + \frac{1-\gamma}{2\gamma^2\sigma^2}\bigg((1-\alpha)\mu - \alpha\Big(r - \frac{\sigma^2}{2}\Big) - \frac{\alpha}{\tau} s_{T-\tau-u}\bigg)^2\Bigg] du. \end{aligned} \tag{5.12}$$

When $\gamma > 1$, we have $\bar{A}_1 \leq A_{1,\varpi} \leq 0$, where $\bar{A}_1 = -\frac{\alpha(\sqrt{\gamma}+1)}{\gamma\sigma^2\tau} < 0$ is a stable steady state of $A_{1,\varpi}$. That is, $A_{1,\varpi}$ will converge to $\bar{A}_1$ as ϖ increases. When $\gamma < 1$, $A_{1,\varpi} \geq 0$ and goes to positive infinity as ϖ increases. Interestingly, the historical price realizations s_u for $u \in [t-\tau, t]$ can affect A_2 and A_3. This makes the optimal portfolio weight depend on the infinite-dimensional space of historical path. So investors need (and only need) the realized prices over the time period $[t-\tau, t]$ when making decision at time t. This is different from the portfolio selection problems under Markov prices, where investors make decisions only based on the current values of state variables.

The optimal weight (5.9) invested in the momentum asset consists of three components: ϕ_t^m, which is the myopic demand, ϕ_t^{MH} and ϕ_t^{PH}. We will call the last two components momentum hedging demand and path hedging demand respectively. The myopic demand ϕ_t^m follows a myopic momentum strategy studied in the empirical literature and depends only on the momentum variable at time t, which is the difference of the log prices at t and $t-\tau$. The second and the third components constitute the intertemporal hedging demand (Merton, 1971). Because $A_{1,\varpi}$ is a monotonic and deterministic function of time t, the momentum hedging demand ϕ_t^{MH} is linear in the state variable m_t and monotonic in investment horizon. Interestingly, although the historical path is not a state variable, it can still affect portfolio weight via the path hedging demand ϕ_t^{PH}. This is due to the fact that the historical price path leads to time-varying expected returns and further affects the time-varying coefficients in the intertemporal hedging demand. In particular, when $\varpi < \tau$, (5.12) states that A_2 only picks up the information of the historical path during $[t-\tau, T-\tau] \Big(\subset [t-\tau, t]\Big)$, and hence the path during $[T-\tau, t)$ cannot affect the portfolio weight. As investment horizon increases, more and more past information is taken into account by the path hedging demand. When $\varpi \geq \tau$, all the path $[t-\tau, t]$ will be considered.

When $\gamma < 1$, there is a finite critical horizon

$$\hat{\varpi} = \frac{\sqrt{\gamma}\tau}{2\alpha} \ln\Big(\frac{1+\sqrt{\gamma}}{1-\sqrt{\gamma}}\Big), \tag{5.13}$$

with which both the optimal portfolio weight and the expected utility approach infinity. It follows from (5.13) that small enough γ can make this infinite expected utility occur for the case $\varpi \leq \tau$.

When $\varpi > \tau$, because θ_t is path-dependent, it is difficult to solve the conditional expectation $\mathbb{E}_t\Big[\xi_T^{(\gamma-1)/\gamma}\Big]$ in closed form. We numerically solve it in next section.

6. Properties of the Optimal Dynamic Momentum Strategy

In this section, we first provide three lemmas on the properties of the portfolio weight, then calibrate the model and further examine the properties of the optimal momentum strategy numerically.

Because momentum is defined in terms of returns, the price level does not affect the future returns as shown in Proposition 4.1. It is natural to expect that the price level should also not affect the portfolio weights. The following lemma verifies this.

Lemma 6.1. *When the historical price path s_u is changed to $s_u + c$ for all $u \in [t-\tau, t]$, where c is a constant, ϕ_t^m, ϕ_t^{MH} and ϕ_t^{PH} do not change. So the three demand components depend on historical returns.*

It is easy to see that a change in the historical price path from s_u to $s_u + c$ for all $u \in [t-\tau, t]$, where c is a constant, cannot affect both myopic demand and momentum hedging demand because they depend only on momentum variable, which is determined by the log price difference $s_t - s_{t-\tau}$. Although ϕ_t^{PH} has different loadings on different historical prices, Lemma 6.1 states that the impacts on $A_{2,\varpi}$ and $A_{1,\varpi} s_{t-\tau}$ in ϕ_t^{PH} cancel out each other and hence the path hedging demand is also independent of a change in the historical price level. In other words, Lemma 6.1 implies that the total weights on historical prices sum up to zero in the myopic demand and the two types of hedging demands. However, all three demand components do depend on historical returns.

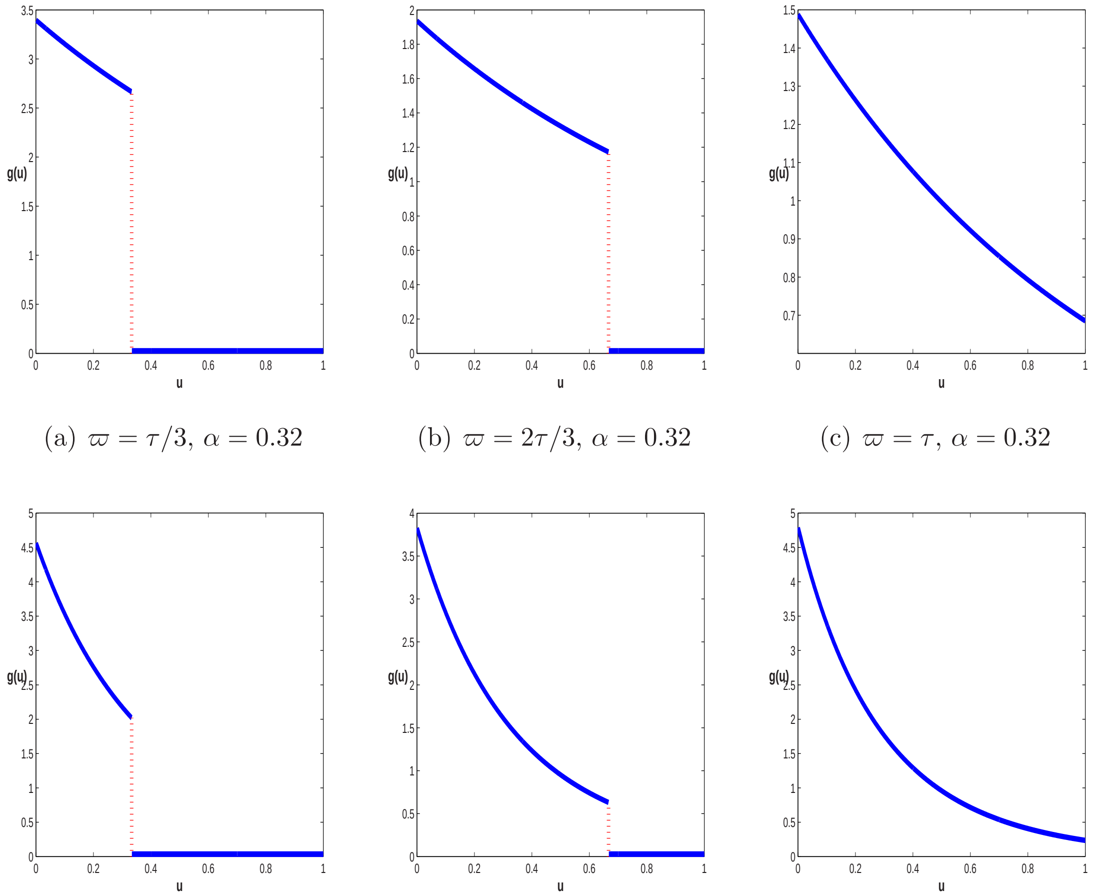
Lemma 6.2. *Assume $\alpha \in [0, 1]$. When the risky asset has a positive momentum ($m_t \geq 0$), the sum of myopic demand and momentum hedging demand is always positive ($\phi_t^m + \phi_t^{MH} > 0$) for investment horizon $\varpi \leq \tau$.*

Lemma 6.2 shows that in the bull market, the demands are always positive when ignoring the path effect no matter $\gamma > 1$ or < 1. However, we will see that the total demands can be negative (Fig. 6.3), that is, $\phi_t^m + \phi_t^{MH} < -\phi_t^{PH}$, even when $m_t \geq 0$.

The path hedging puts different loadings on different historical prices in $A_{2,\varpi}$. Let $g(u)$ be the loading of $-A_{2,\varpi}/A_{1,\varpi}$ on s_u over $u \in [t-\tau, t]$. Lemma 6.1 implies that $g(u) \geq 0$ and $\int g(u)du = 1$, and hence $g(u)$ can be regarded as a density function of the historical prices in $A_{2,\varpi}$. The following lemma characterizes the reaction of the path hedging demand to historical price realizations.

Lemma 6.3. *The density function $g(u)$, defined on $[t-\tau, t]$, is a decreasing function of u for $u \in [t-\tau, T-\tau]$ and becomes zero for $u \in [T-\tau, t]$.*

Lemma 6.3 shows that the path hedging reacts to $s_{t-\tau}$ the most and the reaction becomes smaller for the more recently historical prices during $[t-\tau, T-\tau)$ and no reaction to $[T-\tau, t)$. When $\gamma > 1$, $A_1 < 0$. Lemma 6.3 implies that the hedging demand reacts to the historical log price $s_{T-\tau-u}$ positively via $A_{2,\varpi}$ and hence an increase in a historical price increases the path hedging demand. However, when $\gamma < 1$, $A_1 > 0$ and the loading of path hedging on historical prices becomes negative.



(a) $\varpi = \tau/3$, $\alpha = 0.32$ (b) $\varpi = 2\tau/3$, $\alpha = 0.32$ (c) $\varpi = \tau$, $\alpha = 0.32$

(d) $\varpi = \tau/3$, $\alpha = 1$ (e) $\varpi = 2\tau/3$, $\alpha = 1$ (f) $\varpi = \tau$, $\alpha = 1$

FIGURE 6.1. The distribution function $g(u)$ over $u \in [t-\tau, t]$. Here $\tau = 1$ and $\gamma = 5$.

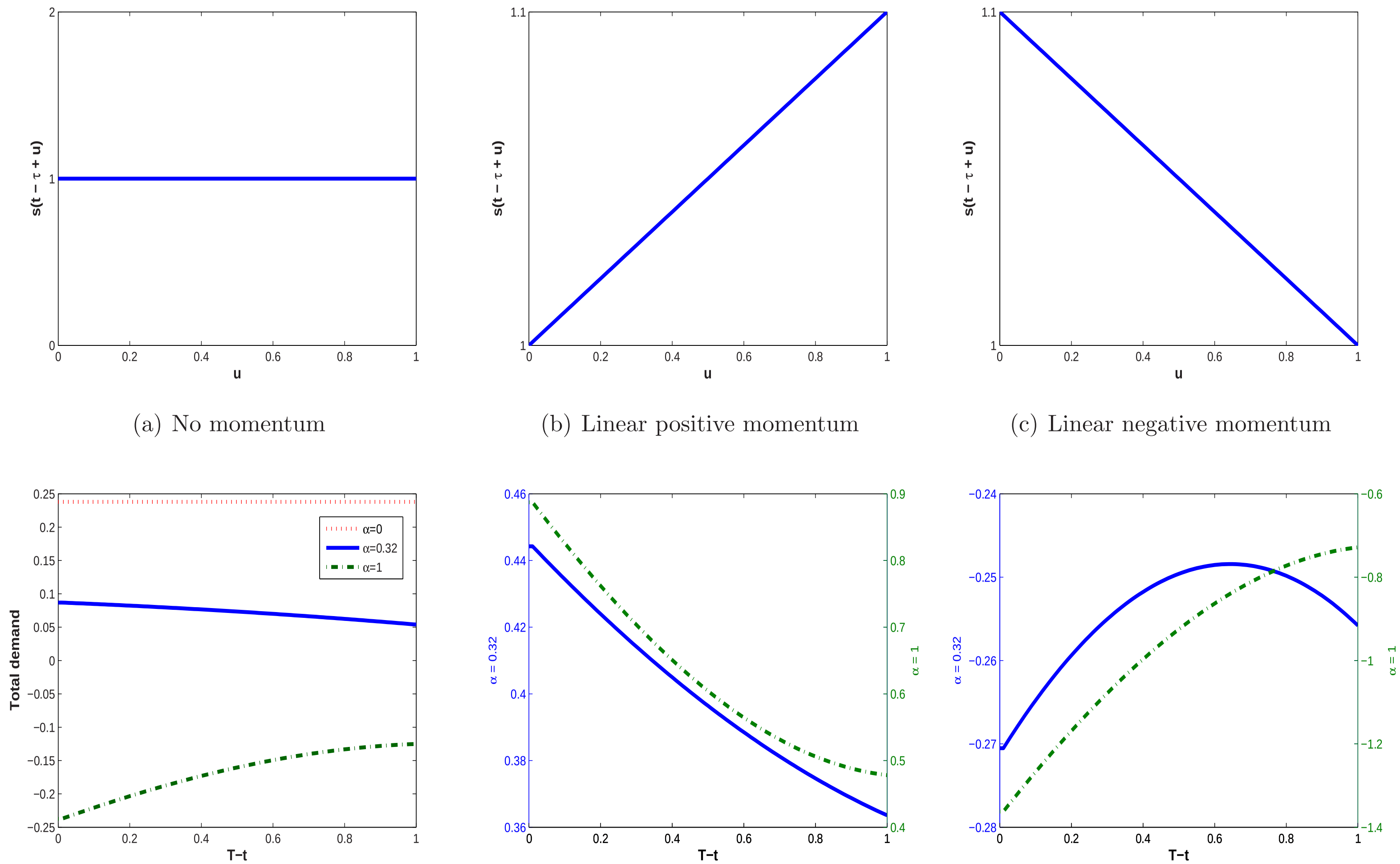
Fig. 6.1 illustrates the loadings of $-A_{2,\varpi}/A_{1,\varpi}$ on s_u over $u \in [t-\tau, t]$ for $\alpha = 0.32$ (the upper panel) and $\alpha = 1$ (the lower panel) with different investment horizons. It verifies Lemma 6.3. We find that the loadings decrease faster for greater α by comparing the cases $\alpha = 0.32$ and $\alpha = 1$.

6.1. **Calibration.** We calibrate the momentum model in this subsection and then numerically examine the portfolio weights in next subsection. To be consistent with the momentum literature, we discretize the continuous-time model (3.1)-(3.2) at a monthly frequency. This results in an autoregressive model of return, whose order depends on τ. We use monthly S&P 500 data over the period January 1871—December 2015 from the home page of Robert Shiller (*www.econ.yale.edu/~shiller/data.htm*).

The total return index is constructed by using the price index series and the dividend series. We set the instantaneous short rate $r = 3\%$ annually and $\mu = 2.38\%$ as the sample risk premium of the S&P 500 index. We estimate the autoregressive model using the maximum likelihood method and conduct the estimations separately for different look-back periods ranging from one month to 5 years. We find that the momentum fraction coefficient α is significantly positive for small look-back period, indicating a significant short-run momentum effect, but becomes negative and insignificant for large look-back period, indicating the long-run reversal. Empirically, Moskowitz et al. (2012) show that the time series momentum strategy based on a 12-month look-back period better predicts the next month's return than other look-back periods. When $\tau = 1$, we have the estimates $\alpha = 32\%$ and $\sigma = 13\%$ in annual terms, which are statistically significantly positive. The numerical results in this paper are based on this set of parameters, unless specified otherwise. Sometimes, we may choose $\alpha = 1$ to examine the impact of different momentum level.

6.2. **Portfolio Dynamics.** We first examine some simple cases with deterministic historical paths of log prices in Fig. 6.2 to provide basic understanding of the impact of historical price path. Fig. 6.2 (d) studies the impact of the non-price component in the optimal portfolio weight by choosing a constant level of historical path. When $\alpha = 0$, the stock returns become i.i.d. with constant risk premium of μ and hedging demands reduce to zero (the red dotted line). When $\alpha = 0.32$ (as estimated in Section 6.1), because the myopic demand corresponds to the value at $\varpi = 0$, the decreasing total demands imply that the hedging demands are negative and the level increases with investment horizon (the blue solid line). This is because momentum is conditionally positively correlated with returns, the hedging demands tend to be negative (positive) when $\gamma > 1$ ($\gamma < 1$). This observation is consistent with Koijen et al. (2009). When $\alpha = 1$, the historical excess returns in this case are negative, so the myopic demand is negative. The green dash-dotted line illustrates that the hedging demands become positive and increasing.

When there is a linear positive momentum, Fig. 6.2 (e) illustrates that the myopic demand follows a short-run momentum strategy with big positive stock holdings. The bigger α is, the stronger the momentum is and hence the more stocks are held by the investors in the bull market. The hedging demands are negative in the presence of positive momentum. For a linear decreasing trend, Fig. 6.2 (f) illustrates a hump-shaped demand function in this case. This implies that, in a bear market, investors can even hold positive hedging demands for small investment horizon, while negative hedging demands for longer horizons. Therefore, the optimal portfolio weights can be the same for certain two different horizons. When the historical path is nonlinear or stochastic, the portfolio weight becomes much more involved.



(a) No momentum

(b) Linear positive momentum

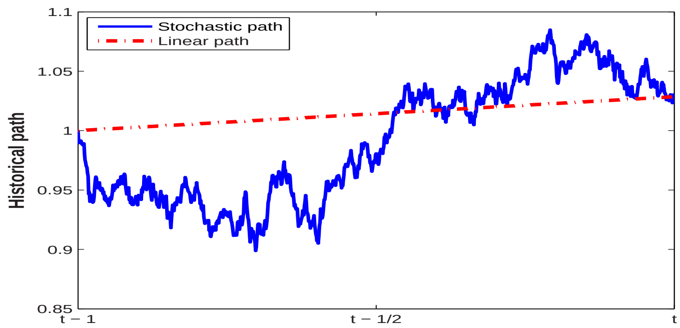
(c) Linear negative momentum

(d) Demands under no momentum

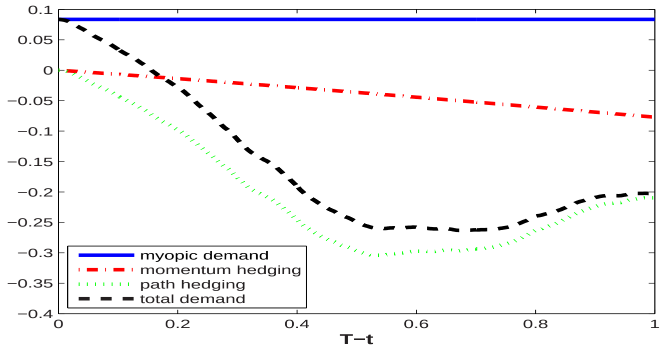
(e) Demands under positive momentum

(f) Demands under negative momentum

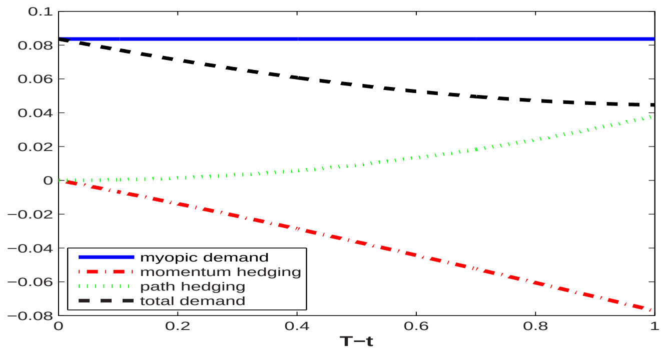
FIGURE 6.2. The historical path (upper panel) and the corresponding optimal portfolio dynamics (lower panel). The blue solid line and green dash-dotted line illustrate the portfolio weights against horizon for $\alpha = 0.32$ and $\alpha = 1$ respectively. Here $\tau = 1$ and $\gamma = 5$. The historical log price path is chosen as deterministic functions of time during $[t - \tau, t]$ with (a) no momentum; (b) linear positive momentum; and (c) linear negative momentum.



(a) A typical stochastic historical path and a deterministic path



(b) Demands under the stochastic path



(c) Demands under the linear path

FIGURE 6.3. (a) A typical stochastic historical price path (the blue solid line) and a linear path connecting the beginning and the end of the stochastic path (the red dash-dotted line). The three components of the optimal portfolios are plotted against investment horizon in (b) and (c) for the stochastic historical path and the corresponding deterministic path respectively. Here $\tau = 1$, $\alpha = 1$ and $\gamma = 5$.

Next we study a typical stochastic path generated from the model, and the portfolio weights are illustrated in Fig. 6.3. There are several observations. First, because momentum and stock price have the same shock, and they are conditionally positively correlated, the intertemporal hedging demand is negative. This leads to negative portfolio weight for large horizons even the historical price paths have positive momentum as illustrated in Panels (b) and (c).

Second, Panel (b) of Fig. 6.3 shows that the path hedging, which strongly reacts to the historical path, has a non-trivial effect and cannot be ignored. To study the impact of it, we compare two paths with the same momentum, while the first with rebounds path as illustrated by the blue solid line in Panel (a) and the second with a linear increasing path (the red dash-dotted line). Under the two market environments, investors hold the same myopic and momentum hedging demands for the two paths, but hold negative (positive) path hedging demand under the first (second) path. After sharp market rebounds, the investors tend to have negative total demands as shown in Fig. 6.3 (b). However, under this market condition, Cujean and Hasler (2015) show that the myopic time series momentum strategy may crash. Barroso and Santa-Clara (2015) and Daniel and Moskowitz (2016) document a similar phenomenon in the cross-section of stocks. So the optimal strategy can successfully benefit from the momentum crashes.

The reason for the negative demands is as follows. The momentum effect leads to long-lasting response to a historical price shock. Lemma 6.3 shows that historical price path has positive contribution to the path hedging demand when $\gamma > 1$, and the contribution becomes smaller for the more recent historical prices. So the path hedging strongly reacts to the historical path, especially to the beginning of it. When the beginning of the path has the opposite pattern to momentum, the path hedging demand, which dominates the total demand, tends to have an opposite sign to the myopic demand. This leads to negative total demand after sharp market rebounds.

Our paper is the first to show that the path hedging demand plays an important role in momentum trading. In general, it fits in with the myopic demand when the price path is generally trending up (or down), while opposes the myopic demand after a rebound (or hump-shaped) price path. Similarly, after a hump-shaped historical path, the optimal dynamic momentum strategy tends to have positive position in the momentum asset even in the presence of a negative momentum.

Third, Fig. 6.3 illustrates a 'smooth' portfolio dynamics even A_2 in (5.12) is stochastic. If we re-write ϕ_t with investment horizon ϖ at time t as ϕ_t^{ϖ}, then

$$\frac{d\phi^{\varpi}}{d\varpi} = s_t \dot{A}_{1,\varpi} + \dot{A}_{2,\varpi},$$

implying that ϕ_t^{ϖ} is not a diffusion process with respect to the investment horizon.

There are other interesting features. For example, there are many bumps in the portfolio weight as a function of horizon as illustrated in Fig. 6.3 (b). In fact, the path dependence introduces time variations in expected returns. They, coupled with momentum, lead to the bumps in horizon dependence. In contrast, the dynamic strategies with Markov state variables typically have monotonically smooth horizon dependence. In addition, the big fluctuation in the portfolio weights shown in Fig. 6.3 (b) implies market timing is important for momentum trading.

6.3. **Horizon longer than look-back period ($\varpi > \tau$).** When $\varpi > \tau$, we do not have closed-form solutions, while we can numerically solve the optimal portfolio weights based on the least squares Monte Carlo approach. The numerical method is described in Appendix B. Fig. 6.4 illustrates the optimal portfolio dynamics with investment horizon up to $\varpi = 5\tau$. We verify that the numerical solutions using the Monte Carlo estimations are the same as the closed-form solutions (5.9) for $\varpi \le \tau$. As horizon ϖ increases, Fig. 6.4 illustrates that the path impact becomes less important and the optimal portfolio weights approach a constant level. With the same set of parameters, we also find the same pattern over horizon $\varpi > \tau$ with different historical price paths (not reported here). This observation can provide guidance for analytically deriving optimal solutions over longer horizons. We leave this to future research.

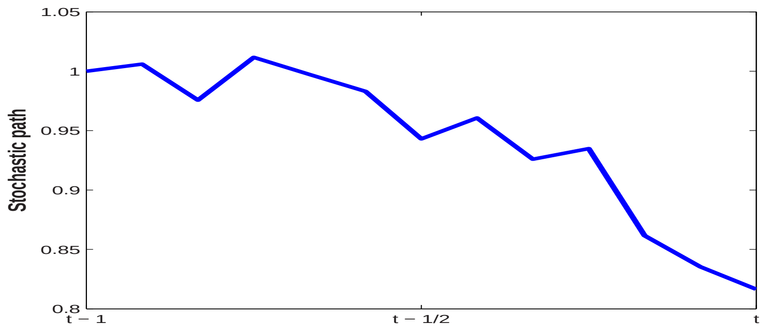
6.4. **The Impacts of Momentum and Look-back Period.** Now we examine the impact of momentum m_t. In order to eliminate the effect of the historical path, we choose a constant path $s_u = \bar{s}$ for $u \in [t-\tau, t)$. Fig. 6.5 plots the optimal portfolio weights with investment horizon of $\varpi = \tau$ against m_t. The optimal portfolio weights are positively linear in m_t. By comparing the three plots, less risk-averse investors will have greater momentum asset holdings.

Fig. 6.6 illustrates the impact of τ with a fixed 1-year investment horizon and the looking-back period τ varying from one to 50 years. Intuitively, momentum is a short-run property and long look-back period makes the momentum variable unable to capture trend. So the portfolio weight becomes less sensitive to momentum and becomes stable over longer look-back period.

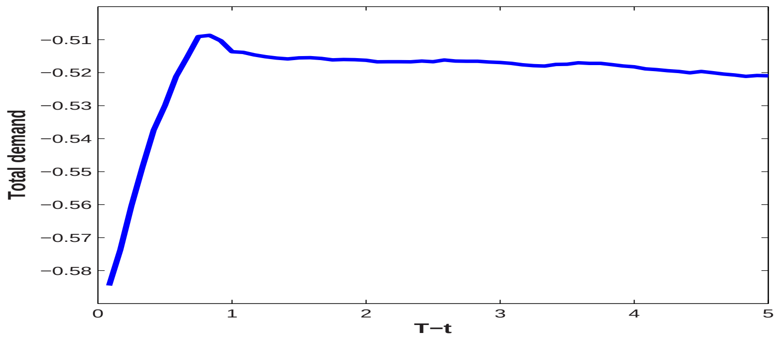
7. Model Extensions and More Discussions

7.1. **Exponentially Decaying Weighted Moving Average.** Alternative MA rules are also used in empirical studies. For example, the momentum variable m_t can be more generally defined as an exponentially decaying weighted MA of historical excess returns over a past time interval $[t-\tau, t]$

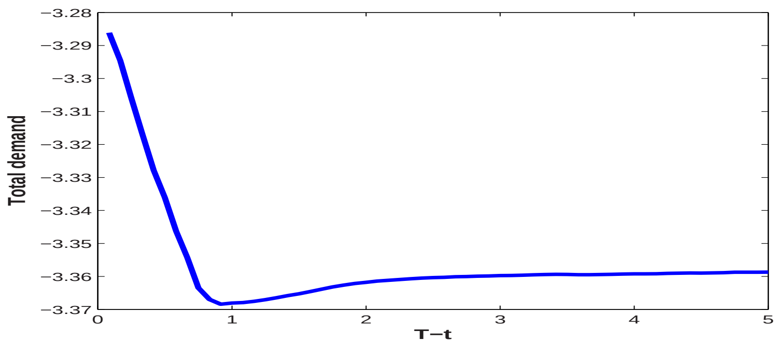
$$m_t = \frac{\lambda}{1-e^{-\lambda\tau}} \int_{t-\tau}^{t} e^{-\lambda(t-u)} \Big(\frac{dS_u}{S_u} - r du\Big), \tag{7.1}$$



(a) A typical stochastic historical path at a monthly frequency



(b) Portfolio weights when $\gamma = 5$



(c) Portfolio weights when $\gamma = 0.9$

FIGURE 6.4. A typical stochastic path (a) and the corresponding optimal portfolio dynamics with investment horizon $\varpi \in [0, 5]$ years for (b) $\gamma = 5$ and (c) $\gamma = 0.9$. The results are based on Monte Carlo simulations with $\tau = 1$. We choose the historical price path at a monthly frequency. (High frequencies are less practical for the Monte Carlo simulations.)

where $\tau \geq 0$ is the look-back period of the momentum, and λ is the decaying rate. Cox-Huang approach still applies. Especially, when $\lambda = 0$, the momentum variable becomes an equally-weighted MA of past excess returns in (3.2).

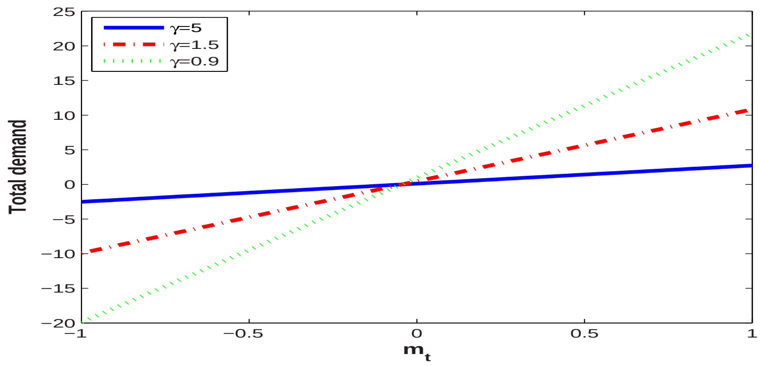


FIGURE 6.5. The optimal portfolio weights as a function of momentum. Here $\tau = 1$ and investment horizon is $\varpi = \tau$. The historical log prices are chosen as constant $s_u = \bar{s}$ during $u \in [t-\tau, t)$ to eliminate their impact.

Interestingly, when $\tau \to \infty$,

$$m_t = \lambda \int_{-\infty}^{t} e^{-\lambda(t-u)} \Big(\frac{dS_u}{S_u} - r du \Big), \tag{7.2}$$

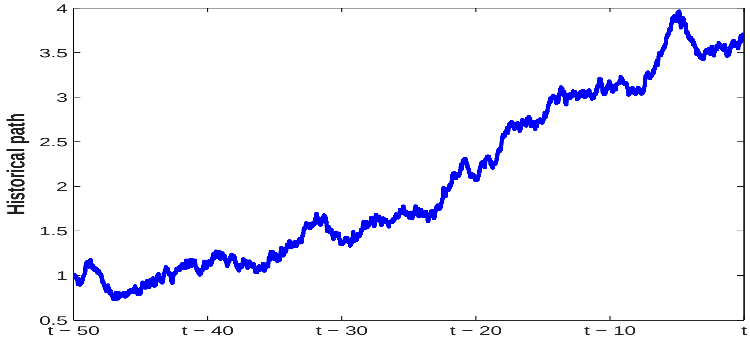
and hence

$$dm_t = \bar{\lambda}(\mu - m_t)dt + \sigma_m dB_t, \tag{7.3}$$

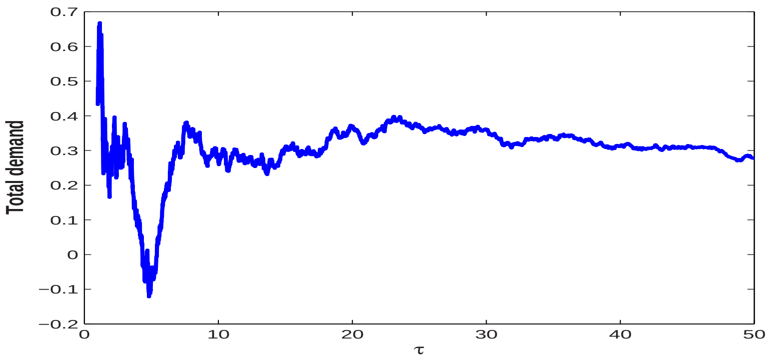
where $\bar{\lambda} = \lambda(1-\alpha)$ and $\sigma_m = \lambda\sigma$. In this case, m_t follows an Ornstein-Uhlenbeck process. The stock process with mean-reverting drift has been studied in Kim and Omberg (1996) and Liu (2007), among others. In this case, m_t is a Markov process, and the optimal portfolio weights are monotonically increasing with both horizon ϖ and state variable m_t when $\gamma > 1$, no matter what the historical path is. This is quite different from the optimal portfolio weights under momentum illustrated by Fig. 6.3, which exhibit big fluctuations with small bumps.

The exponentially decaying weighted MA with infinite look-back period (7.2) has been frequently used to model momentum in the theoretical finance literature because of its tractability. However, as we shown above, this variable cannot capture the short-run momentum, while becomes a long-run reversal indicator (Fama and French, 1988 and Poterba and Summers, 1988).

7.2. **Autoregressive Models.** Alternatively, if we instead consider the momentum variable as a historical excess return over a single time step, then the stock price is



(a) A fifty years typical stochastic historical path



(b) The portfolio weight as a function of the length of the look-back period

FIGURE 6.6. A typical historical path over 50 years $[t-50, t]$ (the upper panel) and the corresponding portfolio weight as a function of the length of the look-back period (the lower panel). Here $\gamma = 5$. We fix the investment horizon as $\varpi = 1$ year and let the looking-back period τ of momentum vary from 1 year to 50 years.

given by[4]

$$\frac{dS_t}{S_t} = \alpha \frac{dS_{t-\tau}}{S_{t-\tau}} + (1-\alpha)(\mu + r)dt + \sigma dB_t. \tag{7.4}$$

The coefficient α in this case measures the (τ/dt)-th order autocorrelation of the excess returns. However, (7.4) cannot well characterize the momentum effect. As emphasized in Moskowitz et al. (2012), "*The studies of autocorrelation examine, by definition, return predictability where the length of the "look-back period" is the same as the "holding period" over which returns are predicted. This restriction masks significant predictability that is uncovered once look-back periods are allowed to differ from predicted or holding periods. In particular, our result that the past*

[4]The stock price in this case is characterized by a stochastic neutral differential equation (Azbelev, Maksimov and Rakhmatullina, 2007).

12 months of returns strongly predicts returns over the next one month is missed by looking at one-year autocorrelations."

7.3. **Separating Current Price and Historical Price Path.** To provide further understanding of the roles of momentum hedging and path hedging, we rewrite system (3.1) as

$$ds_t = \Big[\frac{1}{\tau}(\alpha_1 s_t - \alpha_2 s_{t-\tau}) + (1-\alpha)\Big(r + \mu - \frac{\sigma^2}{2}\Big)\Big]dt + \sigma dB_t, \tag{7.5}$$

where $s_t = \ln S_t$ and α_1 and α_2 are parameters. When $\alpha_1 = \alpha_2 = \alpha$, (7.5) reduces to our momentum model (3.1)-(3.2). When $\alpha_2 = 0$, (7.5) becomes the Markov process studied in Kim and Omberg (1996).

By setting $\alpha_1 = 0$, we 'turn off' momentum, leaving with only path dependence,

$$ds_t = \Big[-\frac{\alpha_2}{\tau} s_{t-\tau} + (1-\alpha)\Big(r + \mu - \frac{\sigma^2}{2}\Big)\Big]dt + \sigma dB_t. \tag{7.6}$$

Then the optimal portfolio weight corresponding to (7.6) with investment horizon $\varpi \le \tau$ is given by

$$\phi_t^* = \frac{\alpha\Big(r - \sigma^2/2\Big) + (1-\alpha)\mu - \alpha_2 s_{t-\tau}/\tau}{\gamma\sigma^2}.$$

In this case, the hedging demand disappears. This is because the path is not a state variable and is uncorrelated with the innovation of stock price at time t. Therefore, the path hedging demand for the system (3.1)-(3.2) is caused by the joint impact of both s_t and $s_{t-\tau}$ in the expected returns.

8. Conclusion

We solve the optimal dynamic momentum strategy between a riskless asset and a momentum asset using the Cox-Huang approach.

We show that investors with relative risk aversion greater than one home-make their own reversal asset by holding less than the myopic investors or even shorting the asset with positive momentum over longer horizons. The optimal portfolio weight tends to be negative after sharp market rebounds, and hence can benefits from the momentum crashes.

Our model can be easily extended to the case with multiple risky assets or with time-varying momentum fractions. We leave this to the future research.

Appendix A. Proofs

A.1. **Details of the Illustrative Example.** To better provide the intuition, we first study a one-period portfolio selection problem. In this case, the expected utility is given by

$$\mathbb{E}\Big[\frac{\tilde{W}^{1-\gamma}}{1-\gamma}\Big] = \mathbb{E}\Big[\frac{W_0^{1-\gamma}}{1-\gamma}(\tilde{R}^p)^{1-\gamma}\Big], \tag{A.1}$$

where

$$\tilde{R}^p = R_f + \phi(\tilde{R} - R_f)$$

is the gross return of the portfolio, and ϕ is the fraction of wealth invested in the risky momentum asset. Define

$$A = \frac{R_f + \phi(U - R_f)}{R_f + \phi(D - R_f)} = \Big[\frac{P(U - R_f)}{(1-P)(R_f - D)}\Big]^{\frac{1}{\gamma}},$$

where the second equality is derived via the first order condition. No arbitrage condition implies $D < R_f < U$. Assume positive risk premium, then $A > 1$. The optimal portfolio weight can be given by

$$\phi = \frac{(A-1)R_f}{U - R_f + A(R_f - D)} > 0. \tag{A.2}$$

Now we go back to the two-period model. No arbitrage condition implies that $D + \Delta_U < R_f < U + \Delta_D$. Assume positive risk premium to guarantee any negative demand is not caused by the negative risk premium, which implies $P(U + \Delta_D) + (1-P)(D + \Delta_D) > R_f$. Then we have

$$\underline{\Delta} \le \Delta_D \le 0 \le \Delta_U \le \overline{\Delta},$$

where $\underline{\Delta} = P(R_f - U) + (1-P)(R_f - D)$ and $\overline{\Delta} = R_f - D$. It follows from (A.2) that the optimal portfolio weight given information at time 1 is given by

$$\begin{aligned} \phi_U &= \frac{(A_U - 1)R_f}{U + \Delta_U - R_f + A_U(R_f - D - \Delta_U)} > 0, \\ \phi_D &= \frac{(A_D - 1)R_f}{U + \Delta_D - R_f + A_D(R_f - D - \Delta_D)} > 0, \end{aligned} \tag{A.3}$$

where

$$A_U = \Big[\frac{P(U + \Delta_U - R_f)}{(1-P)(R_f - D - \Delta_U)}\Big]^{\frac{1}{\gamma}} > 1, \quad A_D = \Big[\frac{P(U + \Delta_D - R_f)}{(1-P)(R_f - D - \Delta_D)}\Big]^{\frac{1}{\gamma}} > 1.$$

At time 0, the optimization problem becomes

$$\begin{aligned}
&\max_{\phi_0,\phi_1} \mathbb{E}_0\Big[\frac{W_2^{1-\gamma}}{1-\gamma}\Big] \\
=&\max_{\phi_0,\phi_1} \mathbb{E}_0\Big[\frac{\big(W_0\tilde{R}_1^p\tilde{R}_2^p\big)^{1-\gamma}}{1-\gamma}\Big] \\
=&\max_{\phi_0} \mathbb{E}_0\Big[\frac{W_0^{1-\gamma}}{1-\gamma}(\tilde{R}_1^p)^{1-\gamma}\max_{\phi_1}\mathbb{E}_1\big[(\tilde{R}_2^p)^{1-\gamma}\big]\Big] \\
=&\max_{\phi_0} \mathbb{E}_0\Big[\frac{W_0^{1-\gamma}}{1-\gamma}(\tilde{R}_1^p)^{1-\gamma}\mathbb{E}_1\big[(\tilde{R}_2^*)^{1-\gamma}\big]\Big],
\end{aligned} \tag{A.4}$$

where $\tilde{R}_2^*$ is the return of the optimal portfolio over period 2. Define $\tilde{\varsigma} = \mathbb{E}_1[(\tilde{R}_2^*)^{1-\gamma}]$. By substituting ϕ_1 derived in (A.2) and replacing U and D by the corresponding gross return at different states, we have

$$\begin{aligned}
\varsigma_U &= \Big[\frac{R_f(U-D)}{U+\Delta_U-R_f+A_U(R_f-D-\Delta_U)}\Big]^{1-\gamma}\big[PA_U^{1-\gamma}+(1-P)\big] > 0, \\
\varsigma_D &= \Big[\frac{R_f(U-D)}{U+\Delta_D-R_f+A_D(R_f-D-\Delta_D)}\Big]^{1-\gamma}\big[PA_D^{1-\gamma}+(1-P)\big] > 0.
\end{aligned} \tag{A.5}$$

It is easy to verify that $\partial\varsigma/\partial\Delta < 0$, which implies that

$$\varsigma_U < \varsigma_D.$$

Then (A.4) becomes

$$\frac{W_0^{1-\gamma}}{1-\gamma}\big[P\varsigma_U+(1-P)\varsigma_D\big]\max_{\phi_0}\mathbb{E}_0\Big[(\tilde{R}_1^p)^{1-\gamma}\frac{\tilde{\varsigma}}{P\varsigma_U+(1-P)\varsigma_D}\Big]. \tag{A.6}$$

Therefore, the problem finally reduces to the standard one-period optimization problem in (A.1) except that the probabilities of up and down states are changed, respectively, to

$$P^* = \frac{\varsigma_U P}{\varsigma_U P+\varsigma_D(1-P)} < P, \quad 1-P^* = \frac{\varsigma_D(1-P)}{\varsigma_U P+\varsigma_D(1-P)} > 1-P.$$

So the new measure $\mathbb{P}^*$ under-weight the up state and over-weight the down state. The optimal portfolio weight at time 0 is then given by

$$\phi_0 = \frac{(A_0^*-1)R_f}{U-R_f+A_0^*(R_f-D)}, \tag{A.7}$$

where

$$A_0^* = \Big[\frac{P^*(U-R_f)}{(1-P^*)(R_f-D)}\Big]^{\frac{1}{\gamma}}.$$

The optimal portfolio returns at different states are given by

$$\begin{aligned}
\tilde{R}_U &= U - \frac{(U - R_f)(A_0^* D - U)}{U - R_f + A_0^*(R_f - D)}, \\
\tilde{R}_D &= D + \frac{(R_f - D)(U - A_0^* D)}{U - R_f + A_0^*(R_f - D)}, \\
\tilde{R}_{UU} &= U + \Delta_U + \frac{(R_f - U - \Delta_U)\big[U + \Delta_U - A_U(D + \Delta_U)\big]}{U + \Delta_U - R_f + A_U(R_f - D - \Delta_U)}, \\
\tilde{R}_{UD} &= D + \Delta_U + \frac{(R_f - D - \Delta_U)\big[U + \Delta_U - A_U(D + \Delta_U)\big]}{U + \Delta_U - R_f + A_U(R_f - D - \Delta_U)}, \\
\tilde{R}_{DU} &= U + \Delta_D + \frac{(R_f - U - \Delta_D)\big[U + \Delta_D - A_D(D + \Delta_D)\big]}{U + \Delta_D - R_f + A_D(R_f - D - \Delta_D)}, \\
\tilde{R}_{DD} &= D + \Delta_D + \frac{(R_f - D - \Delta_D)\big[U + \Delta_D - A_D(D + \Delta_D)\big]}{U + \Delta_D - R_f + A_D(R_f - D - \Delta_D)}.
\end{aligned} \tag{A.8}$$

To show why it can be optimal to have negative portfolio weight at time 0, we examine a special case when $\Delta_D = \underline{\Delta}$ and $\Delta_U = \overline{\Delta}$. In this case,

$$P^* = \frac{P}{1+P}, \quad A_0^* = \Big[\frac{P(U - R_f)}{R_f - D}\Big]^{\frac{1}{\gamma}},$$

and

$$\begin{aligned}
\tilde{R}_U &= R_f + \phi_0^o(U - R_f) < R_f, \\
\tilde{R}_D &= R_f + \phi_0^o(D - R_f) > R_f, \\
\tilde{R}_{UU} &= +\infty, \\
\tilde{R}_{UD} &= \tilde{R}_{DU}^o = \tilde{R}_{DD}^o = R_f.
\end{aligned} \tag{A.9}$$

The terminal utilities is given by

$$\begin{aligned}
\tilde{V}_{UU} &= -a\big(R_U R_{UU}\big)^{1-\gamma} = 0, \\
\tilde{V}_{UD} &= -a\big(R_U R_{UD}\big)^{1-\gamma} = -a\big[R_f + \phi_0(U - R_f)\big]^{1-\gamma} R_f^{1-\gamma}, \\
\tilde{V}_{DU} &= -a\big(R_D R_{DU}\big)^{1-\gamma} = -a\big[R_f + \phi_0(D - R_f)\big]^{1-\gamma} R_f^{1-\gamma}, \\
\tilde{V}_{DD} &= -a\big(R_D R_{DD}\big)^{1-\gamma} = -a\big[R_f + \phi_0(D - R_f)\big]^{1-\gamma} R_f^{1-\gamma},
\end{aligned} \tag{A.10}$$

for the myopic strategy when ϕ_0 is chosen as (A.2), and for the optimal strategy when ϕ_0 is chosen as (A.7).

We compare the optimal strategy with the myopic strategy which only cares about the expected utility one period ahead. The two strategies hold the same portfolio weights at time 1. So we only need to study the first period. The myopic portfolio weight is given by (A.2), which is based on the probabilities of market going up and down as P and $1 - P$ respectively. However, the optimal strategy is based on the probabilities of $P^* = P/(1+P)$ and $1 - P^* = 1/(1+P)$ respectively. For the optimal strategy, the probability of up market becomes 'smaller' ($P^* < P$) after adjusted for $\tilde{\varsigma}$ and hence it holds less risky asset than the myopic strategy. Especially, when

$$R_f - D > P(U - R_f), \tag{A.11}$$

$A_0^* < 1$. It follows from (A.7) that $\phi_0 < 0$, that is, the optimal strategy shorts the momentum asset in this case.

A.2. **Proof of Lemma 3.1.** The solution can be found by using forward induction steps of length τ. Let $t \in [0, \tau]$. Then the general system (3.1)-(7.1) becomes

$$\begin{cases} dS_t = S_t dN_t, & t \in [0, \tau], \\ S_t = \varphi_t \quad \text{for} & t \in [-\tau, 0] \quad \text{a.s.} \end{cases} \tag{A.12}$$

where

$$N_t = \int_0^t \left[\frac{\alpha\lambda}{1 - e^{-\lambda\tau}} \int_{s-\tau}^s e^{-\lambda(s-u)} \Big(\frac{d\varphi_u}{\varphi_u} - r du \Big) + (1-\alpha)\mu + r \right] ds + \sigma B_t$$

is a semimartingale. Then the system (A.12) has a unique solution

$$S_t = \varphi_0 \exp\Big\{ N_t - \frac{\sigma^2 t}{2} \Big\},$$

for $t \in [0, \tau]$. This implies that $S_t > 0$ for all $t \in [0, \tau]$ almost surely, when $\varphi_t > 0$ for $t \in [-\tau, 0]$ a.s. By a similar argument, it follows that $S_t > 0$ for all $t \in [\tau, 2\tau]$ a.s. Therefore $S_t > 0$ for all $t \geq 0$ a.s., by induction. Note that the above argument also implies existence and piecewise-uniqueness of the solution to the system (3.1)-(7.1).

A.3. **Proof of Proposition 4.1.** Let $s_t = \ln S_t$. Then we have

$$m_t = \frac{1}{\tau}(s_t - s_{t-\tau}) - \Big(r - \frac{\sigma^2}{2} \Big),$$

and hence

$$ds_t = \left[(1-\alpha)\Big(r + \mu - \frac{\sigma^2}{2} \Big) + \frac{\alpha}{\tau}(s_t - s_{t-\tau}) \right] dt + \sigma dB_t, \tag{A.13}$$

which implies

$$d\big(e^{-\frac{\alpha}{\tau}t} s_t\big) = e^{-\frac{\alpha}{\tau}t} \left[\left((1-\alpha)\Big(r + \mu - \frac{\sigma^2}{2} \Big) - \frac{\alpha}{\tau} s_{t-\tau} \right) dt + \sigma dB_t \right],$$

and

$$s_t = \frac{\tau}{\alpha}(1-\alpha)\Big(r+\mu-\frac{\sigma^2}{2} \Big) \big[e^{\frac{\alpha}{\tau}t} - 1 \big] + e^{\frac{\alpha}{\tau}t} P_0 - \frac{\alpha}{\tau} \int_{-\tau}^{t-\tau} e^{\frac{\alpha}{\tau}(t-\tau-v)} s_v dv + \sigma \int_0^t e^{\frac{\alpha}{\tau}(t-v)} dB_v. \tag{A.14}$$

We want to separate s_t into two parts, one determined by the initial values and another collecting all innovations. Notice the third term in (A.14) comprises the information of price s during $[-\tau, t-\tau]$. When $t \in [0,\tau]$, the third term is completely determined by the initial values and hence we have

$$s_t = \frac{\tau}{\alpha}(1-\alpha)\left(r+\mu-\frac{\sigma^2}{2}\right)\left[e^{\frac{\alpha}{\tau}t}-1\right]+e^{\frac{\alpha}{\tau}t}s_0-\frac{\alpha}{\tau}\int_{-\tau}^{t-\tau} e^{\frac{\alpha}{\tau}(t-\tau-v)}s_v dv+\sigma\int_0^t e^{\frac{\alpha}{\tau}(t-v)}dB_v. \tag{A.15}$$

When $t \in [\tau, 2\tau]$, (A.14) becomes

$$\begin{aligned} s_t =& \frac{\tau}{\alpha}(1-\alpha)\left(r+\mu-\frac{\sigma^2}{2}\right)\left[e^{\frac{\alpha}{\tau}t}-1\right]+e^{\frac{\alpha}{\tau}t}s_0 \\ &-\frac{\alpha}{\tau}\int_0^{t-\tau} e^{\frac{\alpha}{\tau}(t-\tau-v)}s_v dv-\frac{\alpha}{\tau}\int_{-\tau}^{0} e^{\frac{\alpha}{\tau}(t-\tau-v)}s_v dv+\sigma\int_0^t e^{\frac{\alpha}{\tau}(t-v)}dB_v. \end{aligned} \tag{A.16}$$

Notice when $v \in [0, t-\tau] \subseteq [0,\tau]$, s_v is given by (A.15). By replacing s_v in the third term of (A.16) by (A.15), we have

$$\begin{aligned} s_t =& \frac{\tau}{\alpha}(1-\alpha)\left(r+\mu-\frac{\sigma^2}{2}\right)\left(e^{\frac{\alpha}{\tau}t}+[1-\frac{\alpha}{\tau}(t-\tau)]e^{\frac{\alpha}{\tau}(t-\tau)}-2\right)+\left[e^{\frac{\alpha}{\tau}t}-\frac{\alpha}{\tau}(t-\tau)e^{\frac{\alpha}{\tau}(t-\tau)}\right]s_0 \\ &-\frac{\alpha}{\tau}\int_{-\tau}^{0} e^{\frac{\alpha}{\tau}(t-\tau-v)}s_v dv+\frac{\alpha^2}{\tau^2}\int_{-\tau}^{t-2\tau}(t-2\tau-v)e^{\frac{\alpha}{\tau}(t-2\tau-v)}s_v dv \\ &+\sigma\int_0^t e^{\frac{\alpha}{\tau}(t-v)}dB_v-\frac{\sigma\alpha}{\tau}\int_0^{t-\tau}(t-\tau-v)e^{\frac{\alpha}{\tau}(t-\tau-v)}dB_v. \end{aligned} \tag{A.17}$$

After s_t's are expressed as the sum of a term with initial values and a term with Brownian motions for $t \in [i\tau, (i+1)\tau]$, $i = 0, 1, \cdots, n-1$, we can re-write (A.14) for $t \in [n\tau, (n+1)\tau]$ as

$$\begin{aligned} s_t =& \frac{\tau}{\alpha}(1-\alpha)\left(r+\mu-\frac{\sigma^2}{2}\right)\left[e^{\frac{\alpha}{\tau}t}-1\right]+e^{\frac{\alpha}{\tau}t}s_0+\sigma\int_0^t e^{\frac{\alpha}{\tau}(t-v)}dB_v \\ &-\frac{\alpha}{\tau}\left(\int_{-\tau}^{0}+\int_0^{\tau}+\cdots+\int_{(n-1)\tau}^{t-\tau}\right)e^{\frac{\alpha}{\tau}(t-\tau-v)}s_v dv. \end{aligned} \tag{A.18}$$

By substituting s_v, $v \in [i\tau, (i+1)\tau]$, $i = 0, 1, \cdots, n-1$ into the last term of (A.18), we can separate s_t for $t \in [n\tau, (n+1)\tau]$ into an initial values component and a

Brownian motions component. Therefore, mathematical induction implies

$$
\begin{aligned}
s_t = & \frac{\tau}{\alpha}(1-\alpha)\left(r+\mu-\frac{\sigma^2}{2}\right)\left[\sum_{i=0}^{n}\left(\sum_{j=0}^{i}\frac{(-\frac{\alpha}{\tau})^j(t-i\tau)^j}{j!}\right)e^{\frac{\alpha}{\tau}(t-i\tau)}-n-1\right] \\
& + \sum_{i=0}^{n}\frac{(-\frac{\alpha}{\tau})^i(t-i\tau)^i}{i!}e^{\frac{\alpha}{\tau}(t-i\tau)}s_0 \\
& - \frac{\alpha}{\tau}\int_{-\tau}^{0}\left[\sum_{i=1}^{n}\frac{(-\frac{\alpha}{\tau})^{i-1}(t-i\tau-v)^{i-1}}{(i-1)!}e^{\frac{\alpha}{\tau}(t-i\tau-v)}\right]s_v dv \\
& - \frac{\alpha}{\tau}\int_{-\tau}^{t-(n+1)\tau}\left[\frac{(-\frac{\alpha}{\tau})^n[t-(n+1)\tau-v]^n}{n!}e^{\frac{\alpha}{\tau}[t-(n+1)\tau-v]}\right]s_v dv \\
& + \sigma\sum_{i=0}^{n}\int_{0}^{t-i\tau}\frac{(-\frac{\alpha}{\tau})^i(t-i\tau-v)^i}{i!}e^{\frac{\alpha}{\tau}(t-i\tau-v)}dB_v, \qquad t\in[n\tau,(n+1)\tau].
\end{aligned}
$$

The mean value of $\ln(S_t/S_0)=s_t-s_0$ is just the first four terms minus s_0. The variance is given by

$$
\begin{aligned}
\mathbb{V}\mathrm{ar}_0\big[\ln(S_t/S_0)\big] = & \mathbb{V}\mathrm{ar}_0\left[\sigma\sum_{i=0}^{n}\int_{0}^{t-i\tau}\frac{(-\frac{\alpha}{\tau})^i(t-i\tau-v)^i}{i!}e^{\frac{\alpha}{\tau}(t-i\tau-v)}dB_v\right] \\
= & \sigma^2\mathbb{V}\mathrm{ar}_0\Bigg[\int_{0}^{t-n\tau}\left(\sum_{i=0}^{n}\frac{(-\frac{\alpha}{\tau})^i(t-i\tau-v)^i}{i!}e^{\frac{\alpha}{\tau}(t-i\tau-v)}\right)dB_v \\
& + \int_{t-n\tau}^{t-(n-1)\tau}\left(\sum_{i=0}^{n-1}\frac{(-\frac{\alpha}{\tau})^i(t-i\tau-v)^i}{i!}e^{\frac{\alpha}{\tau}(t-i\tau-v)}\right)dB_v \\
& + \cdots \\
& + \int_{t-2\tau}^{t-\tau}\left(\sum_{i=0}^{1}\frac{(-\frac{\alpha}{\tau})^i(t-i\tau-v)^i}{i!}e^{\frac{\alpha}{\tau}(t-i\tau-v)}\right)dB_v \\
& + \int_{t-\tau}^{t}e^{2\frac{\alpha}{\tau}(t-v)}dB_v\Bigg] \qquad\qquad (A.19) \\
= & \sigma^2\Bigg[\int_{0}^{t-n\tau}\left(\sum_{i=0}^{n}\frac{(-\frac{\alpha}{\tau})^i(t-i\tau-v)^i}{i!}e^{\frac{\alpha}{\tau}(t-i\tau-v)}\right)^2 dv \\
& + \int_{t-n\tau}^{t-(n-1)\tau}\left(\sum_{i=0}^{n-1}\frac{(-\frac{\alpha}{\tau})^i(t-i\tau-v)^i}{i!}e^{\frac{\alpha}{\tau}(t-i\tau-v)}\right)^2 dv \\
& + \cdots \\
& + \int_{t-2\tau}^{t-\tau}\left(\sum_{i=0}^{1}\frac{(-\frac{\alpha}{\tau})^i(t-i\tau-v)^i}{i!}e^{\frac{\alpha}{\tau}(t-i\tau-v)}\right)^2 dv \\
& + \int_{t-\tau}^{t}e^{2\frac{\alpha}{\tau}(t-v)}dv\Bigg].
\end{aligned}
$$

By changing of variable $u = v - t$, the variance is given by

$$\begin{aligned}\mathbb{V}\mathrm{ar}_0\big[s_t - s_0\big] =&\sigma^2\bigg[\int_{-\tau}^{0} e^{-\frac{2\alpha}{\tau}u}du + \int_{-2\tau}^{-\tau}\bigg(\sum_{i=0}^{1}\frac{(-\frac{\alpha}{\tau})^i(-i\tau-u)^i}{i!}e^{\frac{\alpha}{\tau}(-i\tau-u)}\bigg)^2du\\ &+\cdots\\ &+\int_{-n\tau}^{-(n-1)\tau}\bigg(\sum_{i=0}^{n-1}\frac{(-\frac{\alpha}{\tau})^i(-i\tau-u)^i}{i!}e^{\frac{\alpha}{\tau}(-i\tau-u)}\bigg)^2du\\ &\int_{-t}^{-n\tau}\bigg(\sum_{i=0}^{n}\frac{(-\frac{\alpha}{\tau})^i(-i\tau-u)^i}{i!}e^{\frac{\alpha}{\tau}(-i\tau-u)}\bigg)^2du\bigg].\end{aligned}$$

A.4. **Proof of Proposition 5.1.** It follows from (3.1) that the market price of risk is given by

$$\theta_t = \frac{\alpha m_t + (1-\alpha)\mu}{\sigma},$$

which satisfies the Novikov's condition

$$\mathbb{E}\Big[\exp\Big\{\frac{1}{2}\int_0^T \theta_t^2 dt\Big\}\Big] < \infty. \tag{A.20}$$

So the state price density is given by

$$\pi_t = \exp\Big\{-\int_0^t r du - \frac{1}{2}\int_0^t \theta_u^2 du - \int_0^t \theta_u dB_u\Big\}. \tag{A.21}$$

Define

$$\xi_t = \exp\Big\{-\frac{1}{2}\int_0^t \theta_u^2 du - \int_0^t \theta_u dB_u\Big\},$$

which is a martingale under the objective probability measure $\mathbb{P}$.

The wealth process follows

$$dW_t = W_t(r + \sigma\theta_t\phi_t)dt + \sigma W_t\phi_t dB_t,$$

where ϕ_t is the fraction of wealth invested in the risky asset. Define the martingale measure $\mathbb{Q}$ by $\frac{d\mathbb{Q}}{d\mathbb{P}} = \xi_T$. Under the martingale measure, the wealth process W_t follows

$$e^{-rt}W_t = W_0 + \sigma\int_0^t e^{-ru}W_u\phi_u dB_u^Q, \quad 0 \le t \le T, \tag{A.22}$$

where $B_t^Q = B_t + \int_0^t \theta_u du$ is a Brownian motion under $\mathbb{Q}$. The budget constraint can be given by

$$\mathbb{E}_0[\pi_T W_T] \le W_0.$$

Then the problem reduces to the unconstrained maximization of

$$\mathbb{E}_0\bigg[\frac{W_T^{1-\gamma}}{1-\gamma}\bigg] + y\big(W_0 - \mathbb{E}_0[\pi_T W_T]\big),$$

where y is the Lagrange multiplier. Proofs of this well-known result can be found in Harrison and Kreps (1979), Cox and Huang (1989) and Karatzas and Shreve (1998). The first order condition leads to the following optimal terminal wealth

$$W_T = (y\pi_T)^{-1/\gamma}. \tag{A.23}$$

Define $\bar{\xi}_0 = \mathbb{E}_0\big[\xi_T^{(\gamma-1)/\gamma}\big]$. Then

$$W_0 = \mathbb{E}_0[\pi_T W_T] = \mathbb{E}_0[\pi_T^{(\gamma-1)/\gamma}]y^{-1/\gamma} = \bar{\xi}_0 e^{(1-\gamma)rT/\gamma} y^{-1/\gamma},$$

and hence the Lagrange multiplier is given by

$$y = \bar{\xi}_0^{\gamma} W_0^{-\gamma} e^{(1-\gamma)rT}. \tag{A.24}$$

It follows from (A.23) and (A.24) that the value function satisfies

$$V = \mathbb{E}_0\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right] = \frac{1}{1-\gamma} W_0^{1-\gamma} \bar{\xi}_0^{\gamma} e^{(1-\gamma)rT}. \tag{A.25}$$

The optimal wealth process is then given by

$$W_t = \pi_t^{-1}\mathbb{E}_t[\pi_T W_T] = W_0 \tilde{\xi}_0^{-1} e^{rt} \xi_t^{-1} \mathbb{E}_t\left[\xi_T^{(\gamma-1)/\gamma}\right]. \tag{A.26}$$

It follows from (A.22) that

$$d(e^{-rt}W_t) = \sigma e^{-rt}\phi_t W_t dB_t^Q.$$

In addition, Ito's formula implies

$$d(e^{-rt}W_t) = d(\pi_t \xi_t^{-1} W_t) = \xi_t^{-1}(\pi_t\theta_t W_t + \psi_t) dB_t^Q,$$

where ψ_t is governed by

$$\pi_t W_t = W_0 + \int_0^t \psi_u dB_u. \tag{A.27}$$

By matching the volatility, the optimal portfolio weight is given by

$$\phi_t = \frac{\theta_t}{\sigma} + \frac{\psi_t}{\sigma\pi_t W_t} = \frac{\alpha m_t + (1-\alpha)\mu}{\sigma^2} + \sigma^{-1} W_0^{-1} \bar{\xi}_0 \psi_t \Big(\mathbb{E}_t\Big[\xi_T^{(\gamma-1)/\gamma}\Big]\Big)^{-1}.$$

A.5. **Proof of Corollary 5.2.** We rewrite (3.1) as

$$ds_t = \left[(1-\alpha)\Big(r + \mu - \frac{\sigma^2}{2}\Big) + \frac{\alpha}{\tau}(s_t - s_{t-\tau})\right]dt + \sigma dB_t, \tag{A.28}$$

where $s_t = \ln S_t$, and

$$\theta_t = \frac{1}{\sigma}\left[(1-\alpha)\mu - \alpha\Big(r - \frac{\sigma^2}{2}\Big)\right] + \frac{\alpha}{\tau\sigma}(s_t - s_{t-\tau}). \tag{A.29}$$

We define a new measure

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \exp\left\{ -\int_t^T \frac{\gamma-1}{\gamma}\theta_u dB_u - \int_t^T \frac{(\gamma-1)^2}{2\gamma^2}\theta_u^2 du \right\}, \tag{A.30}$$

and under the new measure,

$$
\begin{aligned}
ds_t &= \left[\left(r - \frac{\sigma^2}{2}\right)\left(1 - \frac{\alpha}{\gamma}\right) + (1-\alpha)\frac{\mu}{\gamma} + \frac{\alpha}{\gamma\tau}(s_t - s_{t-\tau})\right]dt + \sigma dB_t^*, \\
\mathbb{E}_t[\xi_T^{\frac{\gamma-1}{\gamma}}] &= \xi_t^{\frac{\gamma-1}{\gamma}}\mathbb{E}_t^*\left[\exp\left\{\frac{1-\gamma}{2\gamma^2\sigma^2}\int_t^T\left[(1-\alpha)\mu - \alpha\left(r - \frac{\sigma^2}{2}\right) + \frac{\alpha}{\tau}(s_u - s_{u-\tau})\right]^2 du\right\}\right].
\end{aligned}
\tag{A.31}
$$

Let $X_u = s_u$ and $X_u^\tau = s_{u-\tau}$ for $u \in [t, T]$. We rewrite process s_t as

$$
dX_t = \left[\left(r - \frac{\sigma^2}{2}\right)\left(1 - \frac{\alpha}{\gamma}\right) + (1-\alpha)\frac{\mu}{\gamma} + \frac{\alpha}{\gamma\tau}(X_t - X_t^\tau)\right]dt + \sigma dB_t^*. \tag{A.32}
$$

When $0 \le T - t \le \tau$, X_u^τ is the realized log price and is known at t and hence X_t in (A.32) can be considered as a Markov process. Denote

$$
f(X,t) = \mathbb{E}_t^*\left[\exp\left\{\frac{1-\gamma}{2\gamma^2}\int_t^T\frac{1}{\sigma^2}\left[(1-\alpha)\mu - \alpha\left(r - \frac{\sigma^2}{2}\right) + \frac{\alpha}{\tau}(X_u - X_u^\tau)\right]^2 du\right\}\right]. \tag{A.33}
$$

Feynman-Kac formula implies

$$
\begin{aligned}
&\frac{\partial f}{\partial t} + \left[\left(r - \frac{\sigma^2}{2}\right)\left(1 - \frac{\alpha}{\gamma}\right) + (1-\alpha)\frac{\mu}{\gamma} + \frac{\alpha}{\gamma\tau}(X - X^\tau)\right]\frac{\partial f}{\partial X} + \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial X^2} \\
&+ \frac{1-\gamma}{2\gamma^2\sigma^2}\left[(1-\alpha)\mu - \alpha\left(r - \frac{\sigma^2}{2}\right) + \frac{\alpha}{\tau}(X - X^\tau)\right]^2 f = 0.
\end{aligned}
\tag{A.34}
$$

By guessing and substituting

$$
f(X_t, \varpi) = \exp\left\{\frac{A_{1,\varpi}}{2}X_t^2 + A_{2,\varpi}X_t + A_{3,\varpi}\right\}, \tag{A.35}
$$

into (A.34), where $\varpi = T - t$ is the investment horizon, and replacing $X_{T-\varpi}^\tau$ by $s_{T-\tau-\varpi}$, we have

$$
\begin{aligned}
\dot{A}_{1,\varpi} &= \sigma^2 A_{1,\varpi}^2 + \frac{2\alpha}{\gamma\tau}A_{1,\varpi} + \frac{1-\gamma}{\gamma^2}\frac{\alpha^2}{\tau^2\sigma^2}, \\
\dot{A}_{2,\varpi} &= \left(\sigma^2 A_{1,\varpi} + \frac{\alpha}{\gamma\tau}\right)A_{2,\varpi} + \left[(1-\alpha)\frac{\mu}{\gamma} + \left(r - \frac{\sigma^2}{2}\right)\left(1 - \frac{\alpha}{\gamma}\right) - \frac{\alpha}{\gamma\tau}s_{T-\tau-\varpi}\right]A_{1,\varpi} \\
&\quad + \frac{1-\gamma}{\gamma^2}\frac{\alpha}{\sigma^2\tau}\left[(1-\alpha)\mu + \alpha\left(\frac{\sigma^2}{2} - r\right) - \frac{\alpha}{\tau}s_{T-\tau-\varpi}\right], \\
\dot{A}_{3,\varpi} &= \frac{\sigma^2}{2}A_{2,\varpi}^2 + \frac{\sigma^2}{2}A_{1,\varpi} + \left[(1-\alpha)\frac{\mu}{\gamma} + \left(r - \frac{\sigma^2}{2}\right)\left(1 - \frac{\alpha}{\gamma}\right) - \frac{\alpha}{\gamma\tau}s_{T-\tau-\varpi}\right]A_{2,\varpi} \\
&\quad + \frac{1-\gamma}{2\gamma^2\sigma^2}\left[(1-\alpha)\mu + \alpha\left(\frac{\sigma^2}{2} - r\right) - \frac{\alpha}{\tau}s_{T-\tau-\varpi}\right]^2,
\end{aligned}
\tag{A.36}
$$

with initial condition $A_{1,0} = A_{2,0} = A_{3,0} = 0$. The realized price $s_{T-\tau-\varpi}$ in A_2 and A_3 is continuous but non-differentiable. So (A.36) is a non-autonomous ordinary differential equation system. This is different from the dynamic strategies with

Markov state variables, where Feynman-Kac formula results in an autonomous partial differential equation, which can be further reduced to an autonomous system of Riccati equations. The solution to (A.36) is given by (5.12) by applying the method of variation of constants.

By substituting (A.35) into (A.26), we have

$$\begin{aligned}\frac{dW_t}{W_t} &= \left[\frac{\theta_t^2}{\gamma} + \sigma\theta_t(A_{1,T-t}X_t + A_{2,T-t}) + r\right]dt + \left[\frac{\theta_t}{\gamma} + \sigma(A_{1,T-t}X_t + A_{2,T-t})\right]dB_t,\\ \psi_t &= W_t\pi_t\Big[\frac{1-\gamma}{\gamma}\theta_t + \sigma(A_{1,T-t}X_t + A_{2,T-t})\Big],\end{aligned} \tag{A.37}$$

and hence the optimal portfolio weight is given by

$$\begin{aligned}\phi_t &= \frac{\theta_t}{\gamma\sigma} + A_{1,T-t}s_t + A_{2,T-t}\\ &= \frac{\theta_t}{\gamma\sigma} + \tau A_{1,\varpi}m_t + \Big[A_{2,\varpi} + \Big(s_{t-\tau} + r\tau - \frac{\sigma^2\tau}{2}\Big)A_{1,\varpi}\Big].\end{aligned} \tag{A.38}$$

The last equality follows from $m_t = \frac{1}{\tau}(s_t - s_{t-\tau}) - r + \frac{\sigma^2}{2}$, which can be directly derived from (3.2).

A.6. **Proof of Lemma 6.1.** At time t, the optimal portfolio (5.9) only consists of the historical prices over $[t-\tau, t]$, so we examine that the historical path changes from s_u to $s_u + c$ for $u \in [t-\tau, t]$, where c is a constant. The price trend $s_t - s_{t-\tau}$ in the myopic demand is still the same and hence the myopic demand does not change. For the hedging demand, (5.12) implies that $A_{2,\varpi}$ changes by

$$\begin{aligned}&A_{2,\varpi}(s_u + c) - A_{2,\varpi}(s_u)\\ =&- ce^{\int_0^\varpi(\sigma^2 A_{1,v} + \frac{\alpha}{\gamma\tau})dv}\int_0^\varpi\left(\frac{\alpha}{\gamma\tau}A_{1,v} + \frac{1-\gamma}{\gamma^2}\frac{\alpha^2}{\sigma^2\tau^2}\right)e^{-\int_0^v(\sigma^2 A_{1,u} + \frac{\alpha}{\gamma\tau})du}dv,\end{aligned}$$

and the second component of the optimal portfolio changes by $cA_{1,\varpi}$. We only need to show that the sum of the two changes equals zero.

In fact, the first equation in (A.36) implies that

$$\begin{aligned}&\dot{A}_{1,\varpi}e^{-\int_0^\varpi(\sigma^2 A_{1,v} + \frac{\alpha}{\gamma\tau})dv} - \left(\sigma^2 A_{1,\varpi} + \frac{\alpha}{\gamma\tau}\right)A_{1,\varpi}e^{-\int_0^\varpi(\sigma^2 A_{1,v} + \frac{\alpha}{\gamma\tau})dv}\\ =&\left(\frac{\alpha}{\gamma\tau}A_{1,\varpi} + \frac{1-\gamma}{\gamma^2}\frac{\alpha^2}{\sigma^2\tau^2}\right)e^{-\int_0^t(\sigma^2 A_{1,v} + \frac{\alpha}{\gamma\tau})dv}.\end{aligned}$$

By taking the integral from 0 to ϖ, we have

$$A_{1,\varpi}e^{-\int_0^\varpi(\sigma^2 A_{1,v} + \frac{\alpha}{\gamma\tau})dv} = \int_0^\varpi\left(\frac{\alpha}{\gamma\tau}A_{1,v} + \frac{1-\gamma}{\gamma^2}\frac{\alpha^2}{\sigma^2\tau^2}\right)e^{-\int_0^v(\sigma^2 A_{1,u} + \frac{\alpha}{\gamma\tau})du}dv,$$

so the change in the hedging demand

$$cA_{1,\varpi} + A_{2,\varpi}(s_u + c) - A_{2,\varpi}(s_u) = 0.$$

Therefore, both myopic demand and hedging demand do not react to a change in the level of historical prices. But the weights on historical price s_u in A_2 are different for different u, so the optimal portfolio weight is affected by the historical patterns in terms of returns.

A.7. **Proof of Lemma 6.2.** Assume $m_t \geq 0$ and $\alpha \in [0,1]$. The sum of myopic demand and momentum hedging demand is given by

$$\phi_t^m + \phi_t^{MH} = \Big(\frac{\alpha}{\gamma\sigma^2} + \tau A_{1,\varpi}\Big)m_t + \frac{\mu}{\gamma\sigma^2}(1-\alpha) > \Big(\frac{\alpha}{\gamma\sigma^2} + \tau A_{1,\varpi}\Big)m_t.$$

In order to demonstrate that $\phi_t^m + \phi_t^{MH} > 0$ for $\varpi \leq \tau$, it is sufficient to show that $\frac{\alpha}{\gamma\sigma^2} + \tau A_{1,\varpi} > 0$.

Notice that $A_{1,\varpi}$ in (A.36) has two steady states

$$\bar{A}_1^{\pm} = \frac{\alpha}{\gamma\sigma^2\tau}(-1 \pm \sqrt{\gamma}),$$

and the corresponding eigenvalues are given by

$$\chi^{\pm} = \pm\frac{2\alpha}{\sqrt{\gamma}\tau}.$$

Therefore, the steady state $\bar{A}_1^-$ is locally asymptotically stable while $\bar{A}_1^+$ is unstable. When $\gamma < 1$, both steady states are negative. Because the greater steady state $\bar{A}_1^+$ is unstable and the initial value A_1 is $A_{i,0} = 0$, $A_{1,\varpi}$ will go to positive infinity as ϖ increases. It is well known that all solutions of a one-dimensional ordinary differential equation are monotonic functions of time, so $A_{1,\varpi} \geq 0$ for all $\varpi \in [0,\tau]$, implying that $\frac{\alpha}{\gamma\sigma^2} + \tau A_{1,\varpi} > 0$.

When $\gamma > 1$, the unstable steady state $\bar{A}_1^+$ is positive while the stable steady state $\bar{A}_1^-$ is negative. So $A_{1,\varpi}$ with 0 initial value will monotonically decrease to $\bar{A}_1^-$ as ϖ increases. It is easy to verify that

$$\frac{\alpha}{\gamma\sigma^2} + \tau A_{1,\varpi} > \frac{\alpha}{\gamma\sigma^2} + \tau A_{1,\tau} = \frac{\alpha\big[1+\sqrt{\gamma}+(1-\sqrt{\gamma})e^{2\alpha/\sqrt{\gamma}}\big]}{\sqrt{\gamma}\sigma^2\big[(\sqrt{\gamma}-1)e^{2\alpha/\sqrt{\gamma}}+1+\sqrt{\gamma}\big]} > 0,$$

whenever $\varpi \leq \tau$ and $\gamma > 1$.

A.8. **Proof of Lemma 6.3.** It follows from (5.12) that $g(u)$ is given by

$$g(u) = \begin{cases} \exp\Big\{\int_{T-\tau-u}^{T-t}(\sigma^2 A_{1,v} + \frac{\alpha}{\gamma\tau})dv\Big\}\Big[\frac{\alpha A_{1,T-\tau-u}}{\gamma\tau A_{1,\varpi}} - \frac{(\gamma-1)\alpha^2}{\gamma^2\sigma^2\tau^2 A_{1,\varpi}}\Big], & u \in [t-\tau, T-\tau], \\ 0, & u \in [T-\tau, t], \end{cases} \tag{A.39}$$

where $A_{1,\varpi}$ is negative when $\gamma > 1$ and positive when $\gamma < 1$. Then the derivative of the density with respect to u for $u \in [t-\tau, T-\tau]$ is given by

$$\frac{\partial g}{\partial u} = -\frac{\alpha^2}{\gamma\tau^2}\frac{A_{1,T-\tau-u}}{A_{1,\varpi}}\exp\Big\{\int_{T-\tau-u}^{\varpi}(\sigma^2 A_{1,v} + \frac{\alpha}{\gamma\tau})dv\Big\} \leq 0, \tag{A.40}$$

implying the decreasing loadings on historical prices.

Appendix B. Monte Carlo Simulation Method for $\varpi > \tau$

The conditional expectations are calculated using the least squares Monte Carlo approach (Longstaff and Schwartz, 2001). More specifically, we simulate 10,000 time series of prices over $[t, T]$ for a given historical path during $[t-\tau, t]$ generated from model (3.1)-(3.2). The conditional expectation $E_t\big[\xi_T^{(\gamma-1)/\gamma}\big]$ in Proposition 5.1 is just the average of $\xi_T^{(\gamma-1)/\gamma}$. The conditional expectation $E_{t+dt}\big[\xi_T^{(\gamma-1)/\gamma}\big]$ is derived by regressing the realizations of $\xi_T^{(\gamma-1)/\gamma}$ on a constant and the corresponding shocks $d\hat{B}_t$ at time $t+dt$ by following Longstaff and Schwartz (2001). We find that adding more regressors, such as $(d\hat{B}_t)^2$, $(d\hat{B}_t)^3$, or prices $\hat{s}_{t+dt}$, $\hat{s}^2_{t+dt}$ or $\hat{s}^3_{t+dt}$, cannot affect the results. Then ψ_t in (5.5) can be derived by regressing

$$d(\pi_t W_t) = W_0 \bar{\xi}_0^{-1}\left(\mathbb{E}_{t+dt}\left[\xi_T^{(\gamma-1)/\gamma}\right] - \mathbb{E}_t\left[\xi_T^{(\gamma-1)/\gamma}\right]\right)$$

on $d\hat{B}_t$. The total demands follow (5.4).

References

Azbelev, N., Maksimov, V. and Rakhmatullina, L. (2007), *Introduction to the Theory of Functional Differential Equations: Methods and Applications*, Hindawi Publishing Corporation.

Barroso, P. and Santa-Clara, P. (2015), 'Momentum has its moments', *Journal of Financial Economics* **116**, 111–120.

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

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

Cox, J. and Huang, C. (1989), 'Optimal consumption and portfolio policies when asset process follow a diffusion process', *Journal of Economic Theory* **49**, 33–83.

Cox, J. and Huang, C. (1991), 'A variational problem arising in financial economics', *Journal of Mathematical Economics* **20**, 465–487.

Cujean, J. and Hasler, M. (2015), Why does return predictability concentrate in bad times?, working paper, University of Toronto.

Daniel, K. and Moskowitz, T. (2016), 'Momentum crashes', *Journal of Financial Economics* **forthcoming**.

DeMiguel, V., Nogales, F. and Uppal, R. (2014), 'Stock return serial dependence and out-of-sample portfolio performance', *Review of Financial Studies* **27**, 1031–1073.

Detemple, J., Garcia, R. and Rindisbacher, M. (2003), 'A Monte Carlo method for optimal portfolios', *Journal of Finance* **58**, 401–446.

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

Harrison, J. and Kreps, D. (1979), 'Martingales and multiperiod securities markets', *Journal of Economic Theory* **20**, 381–408.

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

Karatzas, I. and Shreve, S. (1998), *Methods of Mathematical Finance*, Springer-Verlag, New York.

Kim, T. and Omberg, E. (1996), 'Dynamic nonmyopic portfolio behavior', *Review of Financial Studies* **9**, 141–161.

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

Larssen, B. and Risebro, N. (2001), When are HJB-equations for control problems with stochastic delay equations finite dimensional?, preprint, University of Oslo.

Liu, J. (2007), 'Portfolio selection in stochastic environments', *Review of Financial Studies* **20**, 1–39.

Lo, A. and Hasanhodzic, J. (2010), *The Evolution of Technical Analysis: Financial Prediction from Babylonian Tablets to Bloomberg Terminals*, John Wiley & Sons, Hoboken, NJ, USA.

Longstaff, F. and Schwartz, E. (2001), 'Valuing American options by simulation: A simple least-squares approach', *Review of Financial Studies* **14**, 113–147.

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.

Neely, C., Rapach, D., Tu, J. and Zhou, G. (2014), 'Forecasting the equity risk premium: The role of technical indicators', *Management Science* **60**, 1772–1791.

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

Schwager, J. (1989), *Market Wizards*, John Wiley & Sons, Hoboken, NJ, USA.

Yong, J. and Zhou, X. (1999), *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York.

Zhu, Y. and Zhou, G. (2009), 'Technical analysis: An asset allocation perspective on the use of moving averages', *Journal of Financial Economics* **92**, 519–544.