OPTIMAL DYNAMIC MOMENTUM STRATEGIES

KAI LI$^1$ AND JUN LIU$^2$

$^1$ 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
kai.li@uts.edu.au, junliu@ucsd.edu

Abstract. Even though the momentum effect has been documented extensively, few papers have studied the optimal momentum trading strategy. Most studies use the myopic or mean-variance strategy which depends only on momentum and is optimal only for very short horizons. In this paper, we solve the optimal portfolio problem between a riskless asset and a risky asset with momentum. We show that the optimal portfolio weight depends on a weighted average of returns in the look-back period, with more recent returns receive higher weights, in addition to momentum. The optimal portfolio weight is higher than the myopic portfolio weight for hump-shaped paths, while is lower for rebound paths. In fact, the investor may even short the asset with a positive momentum if the price has a rebound path and effectively home-makes an asset with return reversal. Furthermore, the outperformance of the optimal portfolio relative to the myopic strategy is most significant during extreme market periods.

Key words: Momentum, optimal trading strategy.

JEL Classification: C32, G11

Date: December 16, 2016.

Acknowledgement: We would like to thank Michael Brennan, Georgy Chabakauri, Li Chen, Xue-Zhong He, Yuelei Li, Bernt Øksendal, Maureen O’Hara, and conference and seminar participants at the Inaugural Workshop on the Complex Information and Financial Market Dynamics (CIFMD 2016), Harbin Institute of Technology, SAIF and University of Technology Sydney for their helpful comments. An earlier version of this paper was prepared while Kai Li was visiting UCSD, whose hospitality he gratefully acknowledges. Financial support from the UTS CPD research grant is gratefully acknowledged. All remaining errors are ours. A previous version of this paper was circulated under the title “Reversing Momentum: The Optimal Dynamic Momentum Strategy”.

1
1. Introduction

Even though the momentum effect has been extensively documented, there exist few studies on the optimal trading strategy to explore it. Most studies on momentum use the myopic or mean-variance strategy. This is probably due to the fact that the price process for momentum is inherently path-dependent and therefore the standard dynamic programming approach of Merton (1971) cannot be used. In this paper, we model momentum in a continuous-time setting and study the optimal portfolio problem between a riskless asset and a risky asset with momentum (momentum asset) using Cox-Huang approach. The optimal portfolio weight is explicitly derived for investment horizon that is equal to or smaller than the length of the look-back period. Portfolio weights for longer horizons, although more involved, can be also studied using our setup.

The optimal portfolio weight of momentum asset has two components. The first is a mean-variance myopic portfolio weight and depends on momentum, which is the cumulative return of the look-back period. This is intuitive because the expected return of momentum asset depends linearly on momentum. The myopic portfolio weight is optimal for very short horizon and is used in most empirical studies. It is positive if momentum is positive.

In addition to momentum, the optimal portfolio weight depends also negatively (positively) on a weighted average of instantaneous returns in the look-back period, with more recent returns receive higher weights for investor with risk aversion coefficient greater (smaller) than one. For example, given a positive momentum, the weighted average of instantaneous returns can be positive (if more recent returns are higher) or negative (if more recent returns are lower), and hence the optimal portfolio weight can be lower or higher than the myopic portfolio weight for investors with risk aversion greater than one.

Furthermore, return paths can qualitatively change the portfolio strategy. Because more recent returns receive higher weights, the weighted average can be quite large for rebound paths, which can lead to a negative optimal portfolio weight. In this case, the momentum investor shorts the asset, even if the momentum is positive, and effectively home-makes an asset with return reversal.

Momentum strategies used in academic literature and in practice mostly are myopic or mean-variance. They depend only on the momentum variable and are independent of historical return paths. Our paper shows that the optimal strategy also depends on the path of returns. This is consistent with the recent empirical evidence that the profitability of the standard momentum strategies is affected by the shapes of the historical paths (e.g., Da, Gurun and Warachka, 2014, Cujean and Hasler, 2015 and Daniel and Moskowitz, 2016).
Although we study the optimal portfolio problem for time series momentum, the insight also applies to cross-sectional momentum if the cross-sectional momentum is mainly due to the predictability of future return by a moving average of historical returns. In this case, our results suggest that the optimal portfolio should take into account the paths of historical returns, not just momentum. For example, it is optimal to invest more in the momentum asset than a myopic strategy does with hump-shaped paths, but invest less with rebound paths.

Path dependence gives rise to some unique technical features. For example, as a function of horizon, the optimal portfolio weight for Markov prices are typically infinitely differentiable and is monotonic. On the other hand, the optimal portfolio weight for momentum assets is differentiable only once, and there are many intervals of increases and decreases (bumps).

Momentum has been extensively documented in the literature (e.g., Jegadeesh and Titman, 1993, Moskowitz et al., 2012 and Asness, Moskowitz and Pedersen, 2013, among many others). However, momentum strategy used in most studies is that of the mean-variance. Thus, when the risky asset has positive momentum, the portfolio weights are always positive. Our results suggest that they are not optimal for an investor with long investment horizons.

We show that the optimal portfolio significantly outperforms the myopic portfolio. By comparing the certainty equivalent, the optimal portfolio gains an extra return of 10% per year on average. The outperformance of the optimal momentum portfolio is more significant during the extreme periods with large up and down price movements than “normal time”. Moskowitz et al. (2012) also find that the standard time series momentum strategy delivers its highest profits during the most extreme market episodes.

Koijen, Rodríguez and Sbuelz (2009) study the optimal portfolio when the lookback period of momentum is infinite. This is a special case of our setting. In this case, the problem becomes Markovian and is much more tractable. The optimal weight is independent of historical paths.

The paper is organized as follows. Section 2 discusses a model of momentum in continuous time. The optimal portfolio selection problem is solved using the Cox-Huang approach in Section 3. Section 4 examines the properties of the optimal dynamic momentum strategy. More general momentum models are discussed in Section 5. We also provide an illustrative example in Section 6 to describe the intuition of reversing momentum. Section 7 concludes. Calculation details are included in the appendices.

Through return decomposition, Moskowitz, Ooi and Pedersen (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.
2. A Continuous-time Model of Momentum

In this section, we specify the price dynamics of momentum asset in continuous time and study the corresponding return characteristics. 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

\[
dS_t \bigg/ S_t = \left[ \alpha m_t + (1 - \alpha)\mu + r \right] dt + \sigma dB_t, \tag{2.1}
\]

where \(m_t\) is momentum, \(r\) is the short rate which is assumed to be constant, \(\mu\) is a constant and can be shown later to be the average risk premium, and the coefficient \(\alpha\) measures the fraction of momentum in the expected return.

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.” Accordingly, the momentum variable \(m_t\) is defined as an equally-weighted moving average (MA) of historical excess returns over a past time interval \([t - \tau, t]\),

\[
m_t = \frac{1}{\tau} \int_{t-\tau}^{t} \left( \frac{dS_u}{S_u} - r du \right), \tag{2.2}
\]

where \(\tau \geq 0\) is the “look-back period” of momentum.\(^2\) The momentum (2.2) can be also written as

\[
m_t = \frac{1}{\tau} \left( \ln S_t - \ln S_{t-\tau} \right) - r + \frac{\sigma^2}{2}, \tag{2.3}
\]

which states that momentum also depends on the cumulative return of the look-back period.

The equally-weighted MA (2.2) is mostly used in empirical studies and 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.\(^3\) We focus on this momentum variable in our paper. We will also discuss other types of MA later in Section 5. When \(\tau = 0\), \(m_t\) becomes the current instantaneous excess return, and the return process (2.1) becomes i.i.d. When \(\tau =

\(^{2}\)It follows from (2.1) and (2.2) that \(\mu dt = \mathbb{E} \left[ \frac{dS_t}{S_t} - r dt \right]\), which is the unconditional expectation of excess return.

\(^{3}\)Various MA indicators are widely used among practitioners (Lo and Hasanhodzic, 2010), and are found to exhibit significant forecasting powers to equity risk premium. Brock, Lakonishok and LeBaron (1992) document 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. Han, Zhou and Zhu (2016) provide a trend factor based on the MA of historical prices to exploit information over different look-back periods and show that it outperforms substantially the standard momentum factor.
The momentum variable becomes the last period excess return and hence stock return in (2.1) becomes a first order autoregressive (AR(1)) process. DeMiguel, Nogales and Uppal (2014) find that, by exploiting serial dependence, the arbitrage portfolios based on an AR(1) return model attains significantly positive out-of-sample returns. Therefore, our model not only includes the autoregressive models as special cases, but also well captures the momentum effect documented in the literature. We will show later that a general form of (2.2) can also include the Ornstein-Uhlenbeck process as its special case.

The MA of historical returns (2.2) utilizes past information and the resulting asset price model (2.1)-(2.2) is inherently non-Markovian. The following proposition shows that the model admits a unique positive solution $S > 0$.

**Lemma 2.1.** The system (2.1)-(2.2) has an almost surely continuously adapted unique solution $S$ for a given $\mathcal{F}_0$-measurable initial process $\varphi : \Omega \to \mathcal{C}([-\tau, 0], \mathbb{R})$. Furthermore, if $\varphi_t > 0$ for $t \in [-\tau, 0]$ a.s., then $S_t > 0$ for $t \geq 0$ a.s..

Two observations follow Lemma 2.1. First, although (2.3) shows that $m_t$ only depends on two prices at time $t$ and $t - \tau$ respectively, Lemma 2.1 states that we need the whole path of prices during $[t - \tau, t]$ to define the price process. This is because the historical price $S_u$ for $u \in (t - \tau, t)$ will be used to determine the future price at time $u + \tau$ ($> t$) according to (2.3). 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 $[t + \tau, t + 2\tau]$. Second, because $\mathcal{C}([-\tau, 0], \mathbb{R})$ is an infinite-dimensional space of initial conditions, system (2.1)-(2.2) also has infinite dimensions. The corresponding portfolio selection problem is conceptually much more difficult than in Markovian systems, which have finite dimensions.

**Lemma 2.2.** The return process and momentum process are stationary if and only if $-1 < \alpha < 1$.

By discretizing model (2.1)-(2.2), the return follows a restricted AR($N$) process with the same coefficient on lagged returns, where $N = \tau/\Delta t$. The stationary condition is determined by its characteristic equation, which is given by

$$\Phi(X) = X^N - \frac{\alpha}{N} (X^{N-1} + X^{N-2} + \cdots + 1) = 0.$$ 

Lemma 2.2 states that, when $-1 < \alpha < 1$, both return and momentum processes are stationary. The proof is included in Appendix C.2.

When $0 < \alpha < 1$, model (2.1)-(2.2) captures momentum effect. However, the stock return becomes i.i.d. when $\alpha = 0$, while has reversal effect when $-1 < \alpha < 0$ (Fama and French, 1988 and Poterba and Summers, 1988). In the remaining analysis, we focus on the case $0 < \alpha < 1$ to study momentum.
The momentum model (2.1)-(2.2) can be extended to a multiple risky asset model. Moskowitz et al. (2012) show that positive auto-covariance between a security’s excess return next month and its lagged 1-year return is the main driving force for both time series momentum and cross-sectional momentum. So our model characterizes a key feature of both time series momentum and cross-sectional momentum.

2.1. Return Characteristics. The return characteristics of the momentum asset are discussed in Appendix B. We show that the returns have many interesting features which cannot be observed in a standard Markovian price system. For example, different historical return paths (such as rebound and hump-shaped paths), even with the same momentum, lead to different expected returns and different Sharpe ratios. There is also a long-lasting response of returns to a price shock, which is inherent with momentum.

Especially, when $T \leq \tau$, we have

$$\ln S_T - \ln S_0 = \bar{\mu}_1 + e^{\frac{\alpha T}{\tau}} \left[ \int_{-\tau}^{T} \left[ 1 - e^{-\frac{\alpha(u+\tau)}{\tau}} \right] \frac{dS_u}{S_u} + \int_{-\tau+T}^{0} \left( 1 - e^{-\frac{\alpha u}{\tau}} \right) \frac{dS_u}{S_u} \right] + \sigma \int_{0}^{T} e^{\frac{\alpha(T-v)}{\tau}} dB_v,$$

(2.4)

where $\bar{\mu}_1 = \frac{\mu}{\alpha}(1 - \alpha)(r + \mu - \frac{\sigma^2}{2})(e^{\frac{\alpha T}{\tau}} - 1)$ is a constant. The first two terms are the expected return and the last term is the disturbance. (2.4) shows that future cumulative return $\ln S_T - \ln S_0$ of the momentum asset depends on a weighted average of instantaneous returns in the look-back period, with more recent instantaneous returns receive higher weights.

The future cumulative return can be also written as a weighted average of historical cumulative returns with beginning time varying from $-\tau$ to $-\tau + T$ and the same end time of 0,

$$\ln S_T - \ln S_0 = \bar{\mu}_2 + \frac{\alpha}{\tau} e^{\frac{\alpha(T-\tau)}{\tau}} \int_{-\tau}^{T} e^{-\frac{\alpha v}{\tau}} (\ln S_0 - \ln S_v) dv + \sigma \int_{0}^{T} e^{\frac{\alpha(T-v)}{\tau}} dB_v,$$

and $\bar{\mu}_2 = \bar{\mu}_1 + \frac{\sigma^2}{2\alpha} [e^{-\frac{\alpha T}{\tau}} (1 - \alpha) \tau - \alpha T - (1 - \alpha) \tau]$ is a constant. Therefore, future returns depend on the historical return path over $[-\tau, 0]$, and more recent historical returns play more important roles. This is different from a Markovian price system where future returns merely depends on the current value of state variables. In next section, we will show that the path-dependent feature is important for the optimal strategy.

3. The Optimal Dynamic Momentum Strategy

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

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

where \( \phi_t \) is the fraction of wealth invested in the risky momentum asset at time \( t \).

Because the price process for momentum is inherently path-dependent, the standard dynamic programming approach of Merton (1971) cannot be used to solve the optimal trading strategy. In this paper, we solve the optimization problem using the Cox and Huang (1989, 1991) approach, which can be applied to non-Markov prices.

Because the market is complete, the unique state price density is given by

\[
\pi_t = \exp \left\{ -\frac{1}{2} \int_0^t \theta_u^2 du - \int_0^t \theta_u dB_u \right\},
\]

where

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

is the market price of risk. Because \( \theta_t \) is path-dependent, the price of a dollar at time \( t \) in each state is affected by the historical return path over \([t - \tau, t]\). The standard Cox-Huang approach leads to \( W_T^* = \left( \lambda \pi_T \right)^{-1/\gamma} \), where \( \lambda \) is the Lagrange multiplier. Define

\[
\xi_t = \exp \left\{ -\frac{1}{2} \int_0^t \theta_u^2 du - \int_0^t \theta_u dB_u \right\},
\]

which is a martingale. The following lemma provides the general results of the Cox-Huang approach on the optimal momentum strategy and value function. We show the details in Appendix C.4.

**Lemma 3.1.** For an investor with an investment horizon \( T \) and constant coefficient of relative risk aversion \( \gamma \), the optimal wealth fraction invested in the risky momentum asset is given by

\[
\phi_0^* = \frac{\alpha m_0 + (1-\alpha)\mu}{\sigma^2} + \frac{\psi_0}{\sigma W_0},
\]

where \( \psi \) is governed by

\[
\pi_t W_t^* = W_0 + \int_0^t \psi_u dB_u.
\]

The remainder, \( 1 - \phi_0^* \), is invested in the cash account. The optimal wealth process satisfies

\[
W_t^* = W_0 \pi_t^{-1} \mathbb{E}_t \left[ \xi_T^{(\gamma-1)/\gamma} \right] / \mathbb{E}_0 \left[ \xi_T^{(\gamma-1)/\gamma} \right],
\]

and the value function satisfies

\[
V = \frac{1}{1-\gamma} W_0^{1-\gamma} e^{(1-\gamma)rT} \left( \mathbb{E}_0 \left[ \xi_T^{(\gamma-1)/\gamma} \right] \right)^\gamma.
\]
In order to derive the optimal portfolio weight, we need to compute $\mathbb{E}_{t}[\xi_{T}^{-}]$, in which the market price of risk $\theta$ is path-dependent. 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 look-back period $0 \leq T \leq \tau$. Subsection 3.1 provides closed-form solutions. This investment problem is more important than with longer horizons because momentum strategies are implemented only for holding periods shorter than 1 year (Jegadeesh and Titman, 1993, and Moskowitz et al., 2012). Portfolio weights for longer horizons, although more involved, can be also studied using our setup. We will numerically study the longer-horizon case in next section.

3.1. Closed-Form Solutions. The following proposition provides the results for the optimal dynamic momentum strategy when $0 \leq T \leq \tau$.

**Proposition 3.2.** When $0 \leq T \leq \tau$, the optimal wealth fraction invested in the risky momentum asset is given by

$$
\phi_{0}^{*} = \phi_{0}^{M} + \phi_{0}^{H},
$$

$$
\phi_{0}^{M} = \frac{\alpha m_{0} + (1 - \alpha)\mu}{\gamma \sigma^{2}} = \frac{\alpha}{\gamma \sigma^{2}} \int_{-\tau}^{0} \frac{1}{\tau} \left( \frac{dS_{v}}{S_{v}} - rdv \right) + \frac{(1 - \alpha)\mu}{\gamma \sigma^{2}}, \quad (3.8)
$$

$$
\phi_{0}^{H} = \frac{\alpha}{\gamma \sigma^{2}} \int_{-\tau}^{0} (1 - \gamma) \omega_{v} \left( \frac{dS_{v}}{S_{v}} - rdv \right) + C_{1},
$$

where

$$
\omega_{v} = \begin{cases} 
\int_{-\tau}^{v} \hat{\omega}_{u} du, & v \in [-\tau, -\tau + T], \\
\int_{-\tau}^{-\tau + T} \hat{\omega}_{u} du, & v \in [-\tau + T, 0], 
\end{cases} \quad (3.9)
$$

$\hat{\omega}_{u} > 0$ is given by (C.38) and $C_{1}$ is a constant given by (C.39) in Appendix C.5. The remainder, $1 - \phi_{0}^{*}$, is invested in the cash account.

For large $\gamma$, $\phi_{0}^{H}$ is (to the leading order of $1/\gamma$) given by

$$
\phi_{0}^{H} = -\frac{\alpha^{2}}{\gamma \sigma^{2}} \left\{ \int_{-\tau}^{-\tau + T} \frac{v + \tau}{\tau^{2}} \left( \frac{dS_{v}}{S_{v}} - rdv \right) + \int_{-\tau + T}^{0} \frac{T}{\tau^{2}} \left( \frac{dS_{v}}{S_{v}} - rdv \right) + \frac{1 - \alpha}{\alpha \tau} \mu T - \frac{rT^{2}}{2\tau^{2}} \right\},
$$

and the optimal portfolio weight is (to the leading order of $1/\gamma$) given by

$$
\phi_{0}^{*} = \frac{\alpha}{\gamma \sigma^{2}} \left\{ \int_{-\tau}^{-\tau + T} \frac{\tau - \alpha(v + \tau)}{\tau^{2}} \left( \frac{dS_{v}}{S_{v}} - rdv \right) + \int_{-\tau + T}^{0} \frac{\tau - \alpha T}{\tau^{2}} \left( \frac{dS_{v}}{S_{v}} - rdv \right) \right\} + \frac{(1 - \alpha)\mu(\tau - \alpha T)}{\gamma \sigma^{2} \tau} + \frac{\alpha^{2} r T^{2}}{2 \gamma \sigma^{2} \tau^{2}}.
$$

(3.10)
The optimal portfolio weight (3.8) on momentum asset consists of two components. The first component $\phi_{0}^{M}$ is a mean-variance myopic portfolio weight and depends on momentum, which is the equally weighted average of instantaneous excess returns of the look-back period. This is intuitive because the expected return of momentum asset depends linearly on momentum. The myopic portfolio weight is optimal for very short horizon and is used in most empirical studies. It is positive if momentum is positive.

The second component is the intertemporal hedging demand (Merton, 1971). Interestingly, it depends on a weighted average of instantaneous excess returns in the look-back period, with more recent returns receive higher weights. This gives rise to the dependence of portfolio weight on historical return paths. Different historical paths lead to different portfolio weights. So investors need the historical path over $[-\tau,0]$ when making decision at time $0$. It is different from the portfolio selection problems under Markov prices, where the optimal portfolio weight depends only on the current values of state variables.

Proposition 3.2 further implies that, for an investor with relative risk aversion coefficient greater than one ($\gamma > 1$), the intertemporal hedging demand depends negatively on the weighted average of historical returns. However, the hedging demand depends positively on the weighted average for investor with relative risk aversion coefficient less than one ($\gamma < 1$).

To provide more insights, we rewrite the portfolio weight in terms of price in the following corollary.

**Corollary 3.3.** The optimal portfolio weight (3.8) can be also written as

$$
\phi_{0}^{*} = \frac{\alpha}{\gamma \sigma^2} \left[ \ln S_{0} - \ln S_{-\tau} \right] + \left[ \frac{1}{\tau} \int_{-\tau}^{-\tau + T} (1 - \gamma) \hat{\omega}(\ln S_{0} - \ln S_{u})du \right] + C_{2}, (3.11)
$$

where $C_{2}$ is a constant given by (C.41).

For large $\gamma$, the optimal portfolio weight is (to the leading order of $1/\gamma$) given by

$$
\phi_{0}^{*} = \frac{\alpha}{\gamma \sigma^2} \left[ \ln S_{0} - \ln S_{-\tau} \right] - \left[ \frac{\alpha}{\tau^2} \int_{-\tau}^{-\tau + T} (\ln S_{0} - \ln S_{u})du \right] + C_{3}, (3.12)
$$

where $C_{3} = \frac{(\tau - \alpha T)(1 - \alpha)\mu - \alpha r}{\gamma \sigma^2}$, $\frac{\alpha^2 T(T - 2\tau)}{4\gamma^2}$.

Corollary 3.3 states that the optimal portfolio weight is a weighted average of historical cumulative returns with beginning time varying from $-\tau$ to $-\tau + T$ and the same end time of $0$. In (3.11) and (3.12), the first component in the square bracket is corresponding to the myopic demand, determined by the cumulative return over the look-back period of the momentum. The second component, caused by the hedging
demand, is a weighted average of historical cumulative returns over different lookback periods. For large $\gamma$, (3.12) shows that the second component depends on an equally weighted average of historical cumulative returns.

In all, Proposition 3.2 and Corollary 3.3 show that, in addition to momentum, the optimal portfolio weight also depends on a weighted average of instantaneous returns in the look-back period, or a weighted average of historical cumulative returns with beginning time varying from $-\tau$ to $-\tau + T$ and the same end time of 0. In next section, we will explore the path dependence of the portfolio. We will show that there are many novel features associated with the portfolio weight which cannot be observed in a standard Markovian setting.

4. Properties of the Optimal Dynamic Momentum Strategy

In this section, we examine the properties and the economic value of the optimal momentum strategy.

4.1. Path Dependence. Proposition 3.2 and Corollary 3.3 document the path dependence of the portfolio weight. The intertemporal hedging demand in Proposition 3.2 depends on a weighted average of historical returns. The following lemma characterizes its weights on different historical returns.

Corollary 4.1. The weight $\omega_v$ on historical instantaneous excess return $dS_v/S_v - rdv$ in the intertemporal hedging demand $\phi^H_0$ is an increasing function of $v$ for $v \in [-\tau, -\tau + T]$ and becomes constant for $v \in [-\tau + T, 0]$.

For large $\gamma$, the weight becomes

$$
\omega_v = \begin{cases} 
\frac{(\tau + v)}{\tau^2}, & v \in [-\tau, -\tau + T], \\
\frac{T}{\tau^2}, & v \in [-\tau + T, 0].
\end{cases}
$$

Unlike the myopic weight which depends on the cumulative returns (an equally weighted average of historical returns of the look-back period), Corollary 4.1 shows that the intertemporal hedging demand instead depends on an (unequally) weighted average of historical returns, where the weights are higher for more recent historical returns and depend on investment horizon $T$.

Especially, for large $\gamma$, Proposition 3.2 states that the weight $\omega_v$ on historical excess return $dS_v/S_v - rdv$ is linearly increasing with $v$ for $v \in [-\tau, -\tau + T]$ and is a constant for $v \in [-\tau + T, 0]$. For general $\gamma > 0$, the weight $\omega_v$ is approximately a linear function of $v$ for $v \in [-\tau, -\tau + T]$ and is a constant for $v \in [-\tau + T, 0]$ for typical parameters.\[4\]

\[4\]This implies that a hump-shaped path decreases the portfolio weight while a rebound path increases the portfolio weight, as will be shown later.
Fig. 4.1 highlights the weights of the hedging demand on historical instantaneous returns. When investment horizon is shorter than the length of the look-back period ($T < \tau$), the left panel shows that the hedging demand puts more weights on more recent historical returns over $[-\tau, -\tau + T]$ and the weights are the same over $[-\tau + T, 0]$. When investment horizon is equal to the length of the look-back period ($T = \tau$), the right panel shows that more recent historical returns receive higher weights over the entire look-back period $[-\tau, 0]$. Fig. 4.1 also shows that an increase in the momentum fraction coefficient $\alpha$ increases the weights.\(^5\)

\(^5\)In untabulated results, we find that the weights decrease with the volatility $\sigma$. 

\[ dS_v/S_v - rdv \text{ for } v \in [-\tau, 0] \text{ in the hedging demand } \phi_0^H \text{ in (3.8). Here } \sigma = 0.2, \mu = 2\%, r = 3\%, \tau = 1, \text{ and } \gamma = 5. \]
Figure 4.2. Panel (a) plots two typical stochastic historical (log) price paths with a rebound shape and a hump shape respectively generated from the model (2.1)-(2.2) and a linear historical price path. The three paths have the same momentum with the same beginning price and end price. The two components of the optimal portfolio and the total demand are plotted against investment horizon in Panels (b), (c) and (d) for the rebound path, the linear path and the hump-shaped path respectively. Here $\alpha = 0.3$, $\sigma = 0.2$, $\mu = 2\%$, $r = 3\%$, $\tau = 1$ and $\gamma = 5$. 
Moreover, Corollary 4.1 implies that the portfolio weight is path dependent. Fig. 4.2 highlights the portfolio weight dynamics for different shapes of historical paths. Panel (a) plots two typical stochastic paths generated from the momentum model: a rebound path (a path that has a downward trend early while an upward trend later) and a hump-shaped path (a path that has an upward trend early while a downward trend later). For comparison, we also examine a linear path connecting the beginning and the end of the stochastic paths. All three paths have the same beginning price and end price of the look-back period, and hence have the same positive momentum (the same positive cumulative returns). The two components of the optimal portfolio and the total demand for an investor with risk aversion greater than one are plotted against investment horizon in Panels (b), (c) and (d) for the rebound path, the linear path and the hump-shaped path respectively.

Fig. 4.2 shows that the investor holds the same myopic demand for the three paths because all paths have the same momentum. However, the intertemporal hedging demands are different. The hedging demand is negative for the linear upward-trend path as illustrated in Panel (c), while the hedging demand for the rebound path in Panel (b) becomes smaller than that for the linear path. In facts, Panel (b) shows that the paths of past returns can qualitatively change the portfolio strategy. The total demand for the rebound path can be negative even if the path has positive momentum. Panel (d) shows that a hump-shaped path tends to oppose the effect.

The reason for the negative demands is as follow. Because more recent historical returns receive higher weights, the weighted average can be quite large in magnitude for a rebound path, which can lead to a negative optimal portfolio weight. In this case, the momentum investor shorts the asset with positive momentum and effectively home-makes an asset even with return reversal. Section 6 further illustrates the intuition of reversing momentum via a two-period model.

Because momentum is defined in terms of returns, price level does not affect the future returns (see Proposition B.1). Proposition 3.2 further implies that the two demand components are also independent of price levels. That is, when the historical price path $\ln S_u$ is changed to $\ln S_u + \text{c}$ for all $u \in [-\tau, 0]$, where $c$ is a constant, $\phi^M_0$ and $\phi^H_0$ do not change. So both demand components depend on historical returns.

4.2. Horizon Dependence. Interestingly, there are many intervals of increases and decreases (bumps) in the portfolio weight as a function of horizon as illustrated

---

6Because the price is a diffusion process, the instantaneous returns change signs for infinite number of times in any finite intervals. The rebound path generally has more negative returns early while more positive returns later.

7Similarly, after a hump-shaped historical path, the optimal dynamic momentum strategy tends to have positive position in the momentum asset for long investment horizons even in the presence of a negative momentum.
Figure 4.3. This figure illustrates the weights (4.1) of the hedging demand on historical instantaneous excess returns. For an investment horizon $T$, area I is the weights on historical returns during $[-\tau, 0]$. When $T$ increases, the weights increase by the amount illustrated by area II. The change in the weighted average can be positive or negative depending on historical return paths during $[-\tau + T, 0]$. This leads many intervals of increases and decreases in the portfolio weight as a function of horizon.

The following corollary provides further insight into the bumps.

**Corollary 4.2.** The portfolio weight $\phi^*_0$ is a non-diffusion process with respect to the investment horizon $T$:

$$\frac{\partial \phi^*_0}{\partial T} = (1 - \gamma) \frac{\alpha}{\gamma \sigma^2} \int_{-\tau}^{0} \varphi_v \left( \frac{dS_v}{S_v} - rdv \right) + C_4,$$

$$= \frac{\alpha}{\gamma \sigma^2} \int_{-\tau}^{-\tau + T} \frac{\partial \hat{\omega}(u, T)}{\partial T} (\ln S_0 - \ln S_u) du + \frac{\alpha}{\gamma \sigma^2} \hat{\omega}(-\tau + T, T)(\ln S_0 - \ln S_{-\tau + T}) + C_5,$$
where

\[ \varphi_v = \begin{cases} \int_{-\tau}^{v} \frac{\partial \omega(u,T)}{\partial T} du, & v \in [-\tau, -\tau + T], \\ \hat{\omega}(-\tau + T, T) + \int_{-\tau}^{-\tau + T} \frac{\partial \omega(u,T)}{\partial T} du, & v \in [-\tau + T, 0], \end{cases} \]

and \( C_4 \) and \( C_5 \) are given in Appendix C.8.

For large \( \gamma \), to the leading order of \( 1/\gamma \),

\[ \frac{\partial \phi_0^*}{\partial T} = -\frac{\alpha^2}{\gamma \sigma^2 \tau^2} \int_{-\tau + T}^{0} \left( \frac{dS_v}{S_v} - rdv \right) - \frac{\alpha}{\gamma \sigma^2 \tau} [(1 - \alpha)\mu - \alpha r] \]

\[ = -\frac{\alpha^2}{\gamma \sigma^2 \tau^2} (\ln S_0 - \ln S_{-\tau + T}) - \frac{\alpha^2}{\gamma \sigma^2 \tau^2} C_6, \]

where \( C_6 = \frac{\sigma^2 (\tau - T)}{2} + \frac{\tau}{\alpha} [(1 - \alpha)\mu - \alpha r] \).

**Figure 4.4.** This figure illustrates the horizon dependence of the portfolio weight. The blue solid line is a typical (log) price path during \([-\tau, 0]\) generated from the momentum model, and the green dash-dotted line illustrates the total demand for different investment horizon \( T \in [0, \tau] \). This figure verifies that the portfolio weight increases (decreases) with \( T \) when the cumulative return \( \ln S_0 - \ln S_{-\tau + T} + C_6 < 0 \) \((> 0)\) in (4.1). For example, \( \ln S_0 - \ln S_{-\tau + T} + C_6 \) is positive when \( T < 0.2 \), while becomes negative for \( 0.2 < T < 0.6 \). Accordingly, portfolio weight increases with horizon for \( T < 0.2 \) while decreases for \( 0.2 < T < 0.6 \). Here \( \alpha = 0.3, \sigma = 0.2, \mu = 2\%, r = 3\%, \tau = 1 \) and \( \gamma = 5 \).

The bumps can be also understood from (4.1). The cumulative return over \([-\tau + T, 0]\) can be positive or negative for different \( T \), leading to bumps in horizon
dependence. This is illustrated by Fig. 4.4. The blue solid line is a typical (log) price path during \([-\tau, 0]\) generated from the momentum model, and the green dash-dotted line illustrates the total demand for different investment horizon \(T \in [0, \tau]\). Eq. (4.1) states that the portfolio weight increases (decreases) with \(T\) when the cumulative return \(\ln S_0 - \ln S_{-\tau + T} + C_6 < 0\) (\(> 0\)). For example, \(\ln S_0 - \ln S_{-\tau + T} + C_6\) is positive when \(T < 0.2\), while becomes negative for \(0.2 < T < 0.6\). Accordingly, Fig. 4.4 shows that portfolio weight increases with horizon for \(T < 0.2\) while decreases for \(0.2 < T < 0.6\).

In fact, the expected return for different investment horizon depends on different “state variables” as shown in (2.4). Effectively, we have infinite number of state variables to constitute a sufficient statistic of the portfolio if we consider horizon dependence. The change of the set state variables leads to the bumps.

Moreover, even the portfolio weight as a function of investment horizon has bumps, it is still a differentiable function. Corollary 4.2 shows that the portfolio weight for momentum asset is ‘smooth’, although it is differentiable only once. On the other hand, the optimal portfolio weight for Markov prices are typically infinitely differentiable and is monotonic as a function of horizon.

In addition, the big fluctuation in the portfolio weights implies that market timing is important for momentum trading.

4.3. The Dependence on Other Variables. In untabulated analysis, we examine the dependence of the portfolio weight on momentum \(m_0\). We choose a linear (log) price path to eliminate the path effect. We find that the optimal portfolio weight is positively linear in momentum for a fixed investment horizon.

We also examine the impact of the look-back period \(\tau\) on portfolio weight for a fixed investment horizon. Because momentum is a short-run property, a large look-back period makes \(m_0\) unable to capture trend. We find that the portfolio weight becomes less sensitive to momentum in this case.

In addition, we find a non-monotonic dependence of the portfolio weight on the relative risk aversion coefficient \(\gamma\).

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

\[\sqrt{\gamma \tau / (2\alpha)} \ln [(1 + \sqrt{\gamma})/(1 - \sqrt{\gamma})],\]

with which both portfolio weight and expected utility approach infinity.\(^8\)

\(^8\)Although \(dS_t / S_t\) changes signs for infinite number of times in any finite interval, \(\partial \phi / \partial T\) changes signs less frequently depending on the recent price trend over \([-\tau + T, 0]\).

\(^9\)The infinite expected utility occurs for \(T \leq \tau\) when \(\gamma\) is small enough. This phenomenon is called ‘nirvana’ and has been observed when the expected return follows an Ornstein-Uhlenbeck process (Kim and Omberg, 1996). Our paper shows that the nirvana is also with momentum asset.
4.4. Economic Values of the Optimal Momentum Strategy. To assess the economic values of the optimal momentum strategy, we study the certainty equivalent wealths of the optimal strategy and the myopic strategy. Define the certainty equivalent wealth (CEW) of a strategy \( \phi \) for an investor as the amount of wealth that makes the investor indifferent between receiving CEW for sure at the terminal time \( T \) and having $1 today to invest up to the horizon using strategy \( \phi \). It follows from (C.19) that CEW of the optimal momentum strategy \( \phi^* \) satisfies

\[
\frac{\text{CEW}^{1-\gamma}_{\phi^*}}{1-\gamma} = \frac{1^{1-\gamma}}{1-\gamma} c_0^\gamma e^{(1-\gamma) r T},
\]

(4.2)
which is equivalent to

$$\text{CEW}_{φ}^∗ = \xi_0^{\gamma/(1-\gamma)} e^{rT} = e^{rT} \exp \left\{ \frac{\gamma}{1-\gamma} \left( \frac{A_{1,0}}{2} (\ln S_0)^2 + A_{2,0} \ln S_0 + A_{3,0} \right) \right\}. \quad (4.3)$$

The last equality follows from (C.27) and (C.28).

For comparison, we also study the CEW of the myopic strategy, which is given by

$$\text{CEW}_{φ}^m = e^{rT} \exp \left\{ \frac{1}{1-\gamma} \left( \frac{A_{1,0}^m}{2} (\ln S_0)^2 + A_{2,0}^m \ln S_0 + A_{3,0}^m \right) \right\}, \quad (4.4)$$

where $A^m$'s are given in Appendix C.10.

**Figure 4.6.** The maximum of the squared historical cumulative returns $X_{max}$ is plotted against the ratio of CEWs, where $X_{max} = \max_{t',t'' \in [-\tau, 0]} \left\{ (\ln S_{t'} - \ln S_{t''})^2 \right\}$ measures large up and down price movements. The results are based on 10000 simulated historical paths generated from model (2.1)-(2.2). The black solid line illustrates the linear regression $X_{max} = 0.98 + 0.11 \times \text{ratios of CEWs} + \epsilon$, where the slope is highly significantly positive (t-statistics of 78.77). Here $\alpha = 0.3$, $\sigma = 0.2$, $\mu = 2\%$, $r = 3\%$, $\tau = 1$ and $\gamma = 5$.

Fig. 4.5 illustrates the distributions of the certainty equivalent wealths (CEWs) of the optimal strategy and the myopic strategy, and the ratio of the CEWs. The results are based on 10000 simulated historical paths generated from model (2.1)-(2.2). Firstly, Panel (a) shows that both optimal and myopic strategies have CEWs greater than one, indicating that exploring momentum significantly increases the investor’s welfare. Secondly, Panel (b) shows that all the ratios of the CEWs are greater than one, indicating that the optimal strategy, which correctly hedges for
the momentum effect, further improves the investor’s welfare as compared to the myopic strategy. Thirdly, although we examine a one-year investment horizon, the optimal strategy can still significantly outperform the myopic strategy. The mean values of the CEW of the optimal strategy, the CEW of the myopic strategy and the ratio of CEWs are 2.19, 1.72 and 1.10 respectively, implying that the optimal strategy gains an extra return of 10% per year on average.

Moskowitz et al. (2012) find that the standard time series momentum strategy delivers its highest profits during the most extreme market episodes with large up and down price movements. We also examine the impact of market episodes on the economic value of the strategies. To measure the large up and down price movements, we define

$$X_{max} = \max_{t', t'' \in [-\tau, 0]} \left\{ (\ln S_{t'} - \ln S_{t''})^2 \right\},$$

(4.5)

which is the maximum of the squared historical cumulative returns over the look-back period. The greater $X_{max}$ is, the more likely the market is to experience large up and down moves.\(^{10}\)

In Fig. 4.6, the maximum of the squared historical cumulative returns $X_{max}$ is plotted against the ratio of CEWs. The results are based on 10000 simulated historical paths generated from model (2.1)-(2.2). It shows that the ratio of the CEWs increases with $X_{max}$, implying that the outperformance of optimal strategy tends to be more significant during the extreme market episodes with large up and down price movements. This is further confirmed by a linear regression of the ratio of the CEWs on $X_{max}$, where the slope is highly significantly positive with a t-statistics of 78.77.

In all, we show that the optimal momentum strategy significantly outperforms the myopic strategy, and the outperformance of the optimal portfolio becomes more significant during the extreme periods than during “normal time” by optimally hedging for the extreme events. This highlights the important roles of the hedging demand and the path effect especially during extreme market periods.

4.5. Horizon Longer Than the Length of the Look-back Period ($T \geq \tau$). When $T \geq \tau$, the optimization problem becomes much more involved technically, although the Cox-Huang approach still applies. In this subsection, we numerically solve the optimal portfolio weights based on the least squares Monte Carlo approach (Longstaff and Schwartz, 2001). The numerical method is described in Appendix D. We verify that the numerical solutions using Monte Carlo estimations are the same as the closed-form solutions (3.8) for $T \leq \tau$. Fig. 4.7 illustrates the optimal

---

\(^{10}\)Moskowitz et al. (2012) measures the extreme market episode by the squared market index return. To also characterize the path effect, we use the squared historical cumulative returns over the look-back period in (4.5). This measure is consistent with the CEWs in (4.3) and (4.4).
portfolios for investment horizon up to 5 years. The path impact becomes less important as the optimal portfolio weights become stable for large horizon. We still observe a non-monotonic horizon dependence of the portfolio weight for large investment horizons. In addition, the portfolio weight is still independent of price level as demonstrated in Appendix C.9.

5. Model Extensions and More Discussions

5.1. 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 \([t-\tau, t]\),

\[
m_t = \frac{\kappa}{1 - e^{-\kappa \tau}} \int_{t-\tau}^t e^{-\kappa(t-u)} \left( \frac{dS_u}{S_u} - rdu \right),
\]

where \( \tau \geq 0 \) is the look-back period of the momentum, and \( \kappa \) is the decaying rate. Cox-Huang approach still applies. Especially, when \( \kappa = 0 \), the momentum variable reduces to the equally-weighted MA of past excess returns (2.2) in Section 2.
Interestingly, when $\tau \to \infty$,

$$m_t = \kappa \int_{-\infty}^{t} e^{-\kappa(t-u)} \left( \frac{dS_u}{S_u} - ru \right) du,$$

(5.2)

and hence

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

(5.3)

where $\bar{\kappa} = \kappa(1-\alpha)$ and $\sigma_m = \kappa \sigma$. In this case, $m_t$ follows a mean-reverting Ornstein-Uhlenbeck process. The stock process with mean-reverting expected return 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 $T$ and state variable $m_t$ when $\gamma > 1$, no matter what the historical path is. This is very different from the optimal portfolio weights with momentum illustrated by Fig. 4.2.

The exponentially decaying weighted MA with infinite look-back period (5.2) has been frequently used to model momentum in the theoretical finance literature because of its tractability. However, as 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).

5.2. Autoregressive Models. Alternatively, if we instead define the momentum variable as a historical excess return over a single time step, then the stock price is given by

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

(5.4)

The coefficient $\alpha$ in this case measures the $(\tau/dt)$-th order autocorrelation of the excess returns. However, (5.4) 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 12 months of returns strongly predicts returns over the next one month is missed by looking at one-year autocorrelations."

5.3. Separating Current Price and Historical Price Path. To provide further understanding of hedging demand, we rewrite system (2.1) as

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

(5.5)
where \( s_t = \ln S_t \) and \( \alpha_1 \) and \( \alpha_2 \) are parameters. When \( \alpha_1 = \alpha_2 = \alpha \), (5.5) reduces to our momentum model (2.1)-(2.2) in Section 2. When \( \alpha_2 = 0 \), (5.5) becomes the Markov process studied in Kim and Omberg (1996).

By setting \( \alpha_1 = 0 \), we “turn off” momentum, leaving with only path dependence,

\[
d s_t = \left[ \frac{-\alpha_2}{\tau} s_{t-\tau} + (1 - \alpha) \left( r + \mu - \frac{\sigma^2}{2} \right) \right] dt + \sigma d B_t.
\]

When \( T - t \leq \tau \), the corresponding optimal portfolio weight is given by

\[
\phi_0^* = \frac{\alpha \left( r - \frac{\sigma^2}{2} \right) + (1 - \alpha) \mu - \alpha_2 s_{t-\tau}/\tau}{\gamma \sigma^2}.
\]

In this case, the hedging demand disappears because the path is uncorrelated with the innovation of stock price. This implies that the intertemporal hedging demand is caused by the joint impact of both \( s_t \) and \( s_{t-\tau} \) in the expected returns.

6. An Illustrative Two-period 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. The return dynamics of the risky asset are 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 momentum, we assume \( \Delta S \) is positive (negative) when \( S = U \) (\( D \)). So the expected return over period 2 increases by \( \Delta U \) if period 1 realizes positive excess return, while decreases by \( \Delta D \) otherwise. The conditional volatility is a constant \( \sqrt{P(1-P)(U-D)} \) 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

\[
\max_{\phi_0, \phi_1} E_0 \left[ \frac{\tilde{W}_2^{1-\gamma}}{1-\gamma} \right] = \max_{\phi_0, \phi_1} E_0 \left[ \frac{(W_0 \tilde{R}_1^p \tilde{R}_2^p)^{1-\gamma}}{1-\gamma} \right],
\]

where \( \phi \) 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

\[
\max_{\phi_0} E_0 \left[ \frac{W_0^{1-\gamma}}{1-\gamma} (\tilde{R}_1^p)^{1-\gamma} \max_{\phi_1} E_1 \left[ (\tilde{R}_2^p)^{1-\gamma} \right] \right] = \max_{\phi_0} E_0 \left[ \frac{W_0^{1-\gamma}}{1-\gamma} (\tilde{R}_1^p)^{1-\gamma} E_1 \left[ (\tilde{R}_2^p)^{1-\gamma} \right] \right],
\]

(6.2)
where $\tilde{R}_2^*$ is the optimal portfolio return over period 2. By defining $\xi = \mathbb{E}_1[(\tilde{R}_2^*)^{1-\gamma}]$, (6.2) becomes

$$\max_{\phi_0} \mathbb{E}_0 \left[ \frac{W_0^{1-\gamma}}{1-\gamma} (\tilde{R}_1^p)^{1-\gamma} \xi \right].$$  \hfill (6.3)

To rewrite (6.3) in terms of a standard portfolio problem, we need to define a new probability to eliminate $\xi$,

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

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 $\Delta S$ 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. The up state probability $P^*$ under the new measure is smaller than $P$ and decreases as $\Delta U - \Delta D$ increases, as shown in Appendix A. This reduces the optimal stock position at time 0 comparing with the myopic strategy which only cares about the utility one period ahead. When the difference $\Delta U - \Delta D$ 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.

In Fig. 6.1, the optimal portfolio weight of momentum asset at time 0 is negative. In contrast, the myopic strategy always holds positive position whenever 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 $\xi$ reduces the “probability” of up state, the expected utility for the optimal strategy is, however, greater than that for the myopic strategy.

We complete this section with the following two remarks. First, 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. 6.1 is caused by the momentum $\Delta S \neq 0$. Second, the two-period model can be easily extended to a multi-period model to examine the joint impact of different historical returns. However, we instead develop in next section a continuous-time model which enables us to conduct a unified analysis of the impact of historical return paths.

7. Conclusion

We solve the optimal dynamic momentum strategy between a riskless asset and a risky momentum asset using the Cox-Huang approach.
We show that the portfolio weight depends on paths of past returns in addition to momentum. For an investor with relative risk aversion coefficient greater than one, the optimal portfolio weight is higher than the myopic portfolio weight for hump-shaped paths, lower for upward-trending paths, and may even be negative for rebound paths. Thus the investor may short the momentum asset with a positive momentum if it has a rebound path in the look-back period and effectively home-makes an asset with return reversal.

Our model can be easily extended to a multiple risky assets model to study the cross-sectional momentum strategy. Because Moskowitz et al. (2012) shows that the positive auto-covariance 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, our results also shed lights on the optimal dynamic cross-sectional momentum strategy.

Many financial market stylized facts exhibit non-Markovian features, such as the long memory in volatility (Andersen, Bollerslev and Diebold, 2007), post earning announcement drift (Bernard and Thomas, 1989) and price delays (Hou and Moskowitz, 2005). The setup developed in our paper can be also used to study these phenomena. We leave this for future research.
Appendix A. 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

$$E\left[\frac{W^1}{1-\gamma}\right] = E\left[\frac{W^0}{1-\gamma}(\tilde{R}^p)^{1-\gamma}\right],$$

(A.1)

where

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

is the gross return of the portfolio, and \(\phi\) is the fraction of wealth invested in the risky momentum asset. Define

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

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

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

(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 that any negative demand is not caused by the negative risk premium, that is, \(P(U + \Delta_D) + (1 - P)(D + \Delta_D) > R_f\). Then we have

$$\Delta < \Delta_D \leq 0 \leq \Delta_U < \bar{\Delta},$$

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

$$\phi_U = \frac{(A_U - 1)R_f}{U + \Delta_U - R_f + A_U(R_f - D - \Delta_U)} > 0, \quad \phi_D = \frac{(A_D - 1)R_f}{U + \Delta_D - R_f + A_D(R_f - D - \Delta_D)} > 0,$$

(A.3)

where

$$A_U = \left[\frac{P(U + \Delta_U - R_f)}{(1 - P)(R_f - D - \Delta_U)}\right]^{1\gamma} > 1, \quad A_D = \left[\frac{P(U + \Delta_D - R_f)}{(1 - P)(R_f - D - \Delta_D)}\right]^{1\gamma} > 1.$$
At time 0, the optimization problem becomes
\[
\max_{\phi_0, \phi_1} E_0 \left[ \frac{W_2^{1-\gamma}}{1-\gamma} \right] = \max_{\phi_0} E_0 \left[ (\tilde{R}_1^p)^{1-\gamma} \right]
\]
\[
= \max_{\phi_0} E_0 \left[ (\tilde{R}_1^p)^{1-\gamma} \max_{\phi_1} E_1 \left[ (\tilde{R}_2^p)^{1-\gamma} \right] \right]
\]
(A.4)

where \( \tilde{R}_2^* \) is the return of the optimal portfolio over period 2. Define \( \varsigma = E_1[(\tilde{R}_2^p)^{1-\gamma}] \).

By substituting \( \phi_1 \) derived in (A.2) and replacing \( U \) and \( D \) by the corresponding gross return at different states, we have
\[
\varsigma_U = \left[ \frac{R_f(U - D)}{U + \Delta_U - R_f + A_U(R_f - D - \Delta_U)} \right]^{1-\gamma} \left[ P A_U^{1-\gamma} + (1 - P) \right] > 0,
\]
\[
\varsigma_D = \left[ \frac{R_f(U - D)}{U + \Delta_D - R_f + A_D(R_f - D - \Delta_D)} \right]^{1-\gamma} \left[ P A_D^{1-\gamma} + (1 - P) \right] > 0.
\]
(A.5)

It is easy to verify that \( \partial \varsigma / \partial \Delta < 0 \), which implies that
\[ \varsigma_U < \varsigma_D. \]

Then (A.4) becomes
\[
\frac{W_0^{1-\gamma}}{1-\gamma} \left[ P \varsigma_U + (1 - P) \varsigma_D \right] \max_{\phi_0} E_0 \left[ (\tilde{R}_1^p)^{1-\gamma} \frac{\varsigma}{P \varsigma_U + (1 - P) \varsigma_D} \right].
\]
(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
\[
\phi_0 = \frac{\left( A_0^* - 1 \right) R_f}{U - R_f + A_0^*(R_f - D)},
\]
(A.7)

where
\[ A_0^* = \left( \frac{P^* (U - R_f)}{(1 - P^*) (R_f - D)} \right)^{\frac{1}{\gamma}}. \]
The optimal portfolio returns at different states are given by

\[
\begin{align*}
\hat{R}_U &= U + \frac{(U - R_f)(A_0^* D - U)}{U - R_f + A_0^*(R_f - D)}, \\
\hat{R}_D &= D + \frac{(R_f - D)(U - A_0^* D)}{U - R_f + A_0^*(R_f - D)}, \\
\hat{R}_{UU} &= U + \Delta_U + \frac{(R_f - U - \Delta_U)[U + \Delta_U - A_U(D + \Delta_U)]}{U + \Delta_U - R_f + A_U(R_f - D - \Delta_U)}, \\
\hat{R}_{UD} &= D + \Delta_U + \frac{(R_f - D - \Delta_U)[U + \Delta_U - A_U(D + \Delta_U)]}{U + \Delta_D - R_f + A_D(R_f - D - \Delta_U)}, \\
\hat{R}_{DU} &= U + \Delta_D + \frac{(R_f - U - \Delta_D)[U + \Delta_D - A_D(D + \Delta_D)]}{U + \Delta_D - R_f + A_D(R_f - D - \Delta_D)}, \\
\hat{R}_{DD} &= D + \Delta_D + \frac{(R_f - D - \Delta_D)[U + \Delta_D - A_D(D + \Delta_D)]}{U + \Delta_D - R_f + A_D(R_f - D - \Delta_D)}.
\end{align*}
\]

(A.8)

To show why it can be optimal to have negative portfolio weight at time 0, we examine a special case when \(\Delta_D = \bar{\Delta}\) and \(\Delta_U = \bar{\Delta}.\)\(^{11}\) In this case,

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

and

\[
\hat{R}_U = R_f + \phi_0^*(U - R_f) < R_f, \\
\hat{R}_D = R_f + \phi_0^*(D - R_f) > R_f, \\
\hat{R}_{UU} = +\infty, \\
\hat{R}_{UD} = \hat{R}_{DU} = \hat{R}_{DD} = R_f.
\]

(A.9)

The terminal utilities are given by

\[
\begin{align*}
\hat{V}_{UU} &= a(R_U R_{UU})^{1-\gamma} = 0, \\
\hat{V}_{UD} &= a(R_U R_{UD})^{1-\gamma} = a[R_f + \phi_0(U - R_f)]^{1-\gamma} R_f^{1-\gamma}, \\
\hat{V}_{DU} &= a(R_D R_{DU})^{1-\gamma} = a[R_f + \phi_0(D - R_f)]^{1-\gamma} R_f^{1-\gamma}, \\
\hat{V}_{DD} &= a(R_D R_{DD})^{1-\gamma} = a[R_f + \phi_0(D - R_f)]^{1-\gamma} R_f^{1-\gamma},
\end{align*}
\]

(A.10)

for the myopic strategy when \(\phi_0\) is chosen as (A.2), and for the optimal strategy when \(\phi_0\) is chosen as (A.7), where \(a = W_0^{1-\gamma}/(1 - \gamma) < 0.\)

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\(^{11}\)In this special case, there exists arbitrage opportunity because \(D + \Delta_U = R_f.\) However, it is obvious that the optimal strategy will still short the momentum asset when \(\Delta_D\) and \(\Delta_U\) are chosen close enough to \(\bar{\Delta}\) and \(\bar{\Delta}\) respectively, as illustrated in Fig. 6.1.
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 $\zeta$ and hence it holds less risky asset than the myopic strategy. Especially, when

$$R_f - D > P(U - R_f),$$

(A.11)

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

**Appendix B. Return Characteristics of the Momentum Model**

In this section, we examine the return characteristics implied by the momentum model (2.1)-(2.2). Define $s_t = \ln S_t$. Notice that the solutions to (2.1)-(2.2) are derived piecewisely as demonstrated in Appendix C.1. The expected return, volatility and Sharpe ratio should also have different forms in different time intervals with length of $\tau$. Proposition B.1 confirms the conjectures.

**Proposition B.1.** For $T - t \in [n\tau, (n + 1)\tau]$, $n = 0, 1, 2, \ldots$, the cumulative return of the stock over $[t, T]$ is given by

$$s_T - s_t = \left(1 - \alpha\right) \left(\bar{r} + \mu - \frac{\sigma^2}{2}\right) \left[\sum_{i=0}^{n} \left(\sum_{j=0}^{i} \frac{(-\frac{\alpha}{\tau})^j(T - t - i\tau)^j}{j!}\right) e^{\frac{\alpha}{\tau}(T-t-i\tau)} - 1\right]$$

$$+ \left[\sum_{i=0}^{n} \frac{(-\frac{\alpha}{\tau})^i(T - t - i\tau)^i}{i!} e^{\frac{\alpha}{\tau}(T-t-i\tau)} - 1\right] s_t$$

$$- \frac{\alpha}{\tau} \int_{-\tau}^{0} \left[\sum_{i=1}^{n} \frac{(-\frac{\alpha}{\tau})^{i+1}(T - t - i\tau - u)^{i+1}}{(i-1)!} e^{\frac{\alpha}{\tau}(T-t-i\tau-u)}\right] s_{t+u} du$$

$$- \frac{\alpha}{\tau} \int_{-\tau}^{T-t-(n+1)\tau} \left[\frac{(-\frac{\alpha}{\tau})^n[T - t - (n + 1)\tau - u]^n}{n!} e^{\frac{\alpha}{\tau}[T-t-(n+1)\tau-u]}\right] s_{t+u} du$$

$$+ \sigma \sum_{i=0}^{n} \int_{0}^{T-t-i\tau} \frac{(-\frac{\alpha}{\tau})^i(T - t - i\tau - u)^i}{i!} e^{\frac{\alpha}{\tau}(T-t-i\tau-u)} dB_{t+u}.\quad (B.1)$$

There are three observations from Proposition B.1. First, the stock returns (B.1) over $[t, T]$ are weighted sum of the historical prices $s_u$ over $u \in [t - \tau, t]$. Therefore, the return process crucially depends on historical instantaneous returns, not just on the beginning and end prices of the look-back period.

More importantly, the weights on different historical prices in (B.1) are different. This implies that different historical price paths (such as rebound and upward-trend price paths) even with the same momentum lead to different expected returns and
different Sharpe ratios. However, volatility is independent of the historical price path. This path effect has important implications for portfolio selection.

In addition, the weights on all historical prices in (B.1) sum up to zero. So the price level does not affect the returns of the momentum asset.

![Graphs showing term structure of expected returns, volatility, and Sharpe ratios](image)

**Figure B.1.** (a) The mean value $\bar{R}_T = \mathbb{E} \left[ \ln S_T - \ln S_0 \right]$ and (b) standard deviation $\sigma(R_T) = \text{Std} \left[ \ln S_T - \ln S_0 \right]$ of cumulative returns and (c) the Sharpe ratio $(\bar{R}_T - r)/\sigma(R_T)$ conditional on a constant historical price path $\ln S_u = \bar{s}$ for $u \in [-\tau, 0]$) as functions of horizon $T$. The conditional mean and variance are given in Appendix C.3. Here $\tau = 1$, $r = 3\%$, $\mu = 2\%$, $\sigma = 20\%$, and the parameter of momentum fraction $\alpha = 0$ (the red dash-dotted line) or 0.9 (the blue solid line).

Second, the last term in (B.1) is a weighted sum of the innovations $dB$ over $u \in [t, T)$. The weights are positive and decreasing with $u$. Moreover, for $T - t \in [n\tau, (n + 1)\tau]$, $n = 0, 1, 2, \ldots$, the response of returns to an innovation is given by

$$D_t[ds_T] = \sigma \left[ \sum_{i=1}^{n} \left( -\frac{\alpha}{\tau} \right)^i (T - t - i\tau)^{i-1} (i-1)! e^{\frac{\tau}{T} \sum_{i=0}^{n} \left( -\frac{\alpha}{\tau} \right)^{i+1} (T - t - i\tau)^{i} i!} e^{\frac{\tau}{T} (T - t - i\tau)} \right] dt.$$  

It is positive and vanishes for long horizons. This long-lasting response of returns to a price shock is inherent with momentum.

---

12 The conditional mean value of the cumulative return $\mathbb{E}_t[s_T - s_t]$ is just the first four terms on the right-hand side of (B.1), and the variance $\text{Var}_t[s_T - s_t]$ is given by (C.11).

13 The impulse-response function can be examined via the Malliavin derivatives (Detemple, Garcia and Rindisbacher, 2003).
Finally, the momentum increases return volatility, while its impact on expected return depends on the historical path which has infinite dimensions. We examine a simple case in which the historical prices are chosen as the same constant number \( s_u = \bar{s} \) for \( u \in [t - \tau, t] \) to provide a first glance at the impact of momentum on the payoff-risk tradeoff over different time horizons. In this case, the second, third and fourth terms of (B.1) become zero. For comparison, we also examine the i.i.d. return process (\( \alpha = 0 \)). Fig. B.1 illustrates the mean values and standard deviations of returns and the Sharpe ratios over a five-year horizon. When \( \alpha = 0 \), the stock price (2.1) reduces to a standard geometric Brownian motion, and the mean and variance are linear in horizon \( T \). However, both mean and standard deviation of the returns of the momentum asset (\( \alpha = 0.9 \)) are convex functions of horizon. The greater \( \alpha \) is, the greater their curvature and growth rate are. The Sharpe ratio also grows faster with respect to horizon as the momentum fraction \( \alpha \) increases.

Although the means and variances are given piecewisely in Proposition B.1, they are continuous in time as illustrated in Fig. B.1.

**Appendix C. Proofs**

C.1. **Proof of Lemma 2.1.** The solution can be found by using forward induction steps of length \( \tau \). Let \( t \in [0, \tau] \). For a given \( \mathcal{F}_0 \)-measurable initial process \( \varphi : \Omega \rightarrow C([-\tau, 0], R) \), where \( C([-\tau, 0], R) \) is the space of all continuous functions defined on \([-\tau, 0]\), the general system (2.1)-(5.1) becomes

\[
\begin{aligned}
\left\{
\begin{array}{ll}
\frac{dS_t}{S_t} = r_t dN_t, & t \in [0, \tau], \\
S_t = \varphi_t & t \in [-\tau, 0] \quad \text{a.s.}
\end{array}
\right.
\end{aligned}
\]  

(C.1)

where

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

is a semimartingale. Then the system (C.1) has a unique solution

\[
S_t = \varphi_0 \exp \left\{ N_t - \frac{\sigma^2 t}{2} \right\},
\]

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 (2.1)-(5.1).

C.2. **Proof of Lemma 2.2.** The discretization of model (2.1)-(2.2) is given by

\[
r_t = \alpha \frac{r_{t-1} + \cdots + r_{t-N}}{N} + (1 - \alpha)(\mu + r) \Delta t + \sigma \epsilon_t,
\]

(C.2)

where \( N = \tau/\Delta t \) is a positive integer and \( \epsilon_t \sim \mathcal{N}(0, \Delta t) \). So the return process of the momentum asset follows a restricted AR(\( N \)) process with the same coefficient.
on lagged returns. The stationary condition is determined by the characteristic equation for (C.2), which is given by

$$ \Phi(X) = X^N - \frac{\alpha}{N}(X^{N-1} + X^{N-2} + \cdots + 1) = 0. $$

(C.3)

It is easy to check that the process of momentum $m_t$ has the same characteristic equation (C.3), and hence has the same stationary condition as return's.

The return process (C.2) is stationary if and only if all the roots of (C.3) lie inside the unit circle, which is, according to Jury's test, equivalent to

(C.1) $\Phi(1) > 0$;  
(C.2) $(-1)^N\Phi(-1) > 0$;

and

(C.3) the $(N-1) \times (N-1)$ matrices

$$ A_{N-1}^\pm = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ -\alpha/N & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ -\alpha/N & \cdots & -\alpha/N & 1 \end{pmatrix} \pm \begin{pmatrix} 0 & 0 & \cdots & \cdots & 0 & -\alpha/N \\ 0 & 0 & \cdots & -\alpha/N & -\alpha/N \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & -\alpha/N & \cdots & -\alpha/N & -\alpha/N \\ -\alpha/N & -\alpha/N & \cdots & \cdots & -\alpha/N \end{pmatrix} $$

are positive innerwise.

Condition (C.1) is equivalent to $\alpha < 1$. Because

$$ (-1)^N\Phi(-1) = \begin{cases} 1 & \text{if } N \text{ is even;} \\ 1 + \frac{\alpha}{N} & \text{if } N \text{ is odd,} \end{cases} $$

condition (C.2) is equivalent to $\alpha > -N$.

It can be verify that

$$ A_M^+ = \begin{cases} (1 + \frac{\alpha}{M})^{\frac{M}{2}} (1 - \alpha) & \text{if } M \text{ is even;} \\ (1 + \frac{\alpha}{M}) \frac{M-1}{2} (1 - \alpha) & \text{if } M \text{ is odd,} \end{cases} $$

and

$$ A_M^- = \begin{cases} (1 + \frac{\alpha}{M})^{\frac{M}{2}} & \text{if } M \text{ is even;} \\ (1 + \frac{\alpha}{M}) \frac{M+1}{2} & \text{if } M \text{ is odd.} \end{cases} $$

Condition (C.3) implies that $A_M^+ > 0$ and $A_M^- > 0$ for $M = 1, \cdots, N-1$, which is equivalent to $-1 < \alpha < 1$.

In all, Conditions (C.1)-(C.3) are equivalent to $-1 < \alpha < 1$, which is the necessary and sufficient condition for both return and momentum being stationary.
C.3. **Proof of Proposition B.1.** Let \( s_t = \ln S_t \). Then we have

\[
m_t = \frac{1}{\tau} (s_t - s_{t-\tau}) - \left( r - \frac{\sigma^2}{2} \right),
\]

and hence

\[
ds_t = \left[ (1 - \alpha) \left( r + \mu - \frac{\sigma^2}{2} \right) + \frac{\alpha}{\tau} (s_t - s_{t-\tau}) \right] dt + \sigma dB_t, \tag{C.4}
\]

which implies that

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

and

\[
s_t = \frac{\tau}{\alpha} \left( 1 - \alpha \right) \left( r + \mu - \frac{\sigma^2}{2} \right) \left[ e^{\frac{\alpha}{\tau}t} - 1 \right] + e^{\frac{\alpha}{\tau}t} s_0 - \frac{\alpha}{\tau} \int_{-\tau}^{t-\tau} e^{\frac{\alpha}{\tau}(t-\tau-v)} s_v dv + \sigma \int_0^t e^{\frac{\alpha}{\tau}(t-v)} dB_v. \tag{C.5}
\]

We want to separate \( s_t \) into two parts, one determined by the initial values and another collecting all innovations. Notice that the third term in (C.6) 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

\[
s_t = \frac{\tau}{\alpha} \left( 1 - \alpha \right) \left( r + \mu - \frac{\sigma^2}{2} \right) \left[ e^{\frac{\alpha}{\tau}t} - 1 \right] + e^{\frac{\alpha}{\tau}t} s_0 \tag{C.6}
\]

When \( t \in [\tau, 2\tau] \), (C.6) becomes

\[
\begin{align*}
  s_t &= \frac{\tau}{\alpha} \left( 1 - \alpha \right) \left( r + \mu - \frac{\sigma^2}{2} \right) \left[ e^{\frac{\alpha}{\tau}t} - 1 \right] + e^{\frac{\alpha}{\tau}t} s_0 \\
  &\quad - \frac{\alpha}{\tau} \int_{-\tau}^{t-\tau} e^{\frac{\alpha}{\tau}(t-\tau-v)} s_v dv + \frac{\sigma}{\tau} \int_0^t e^{\frac{\alpha}{\tau}(t-v)} dB_v.
\end{align*} \tag{C.7}
\]

Notice that when \( v \in [0, t-\tau] \subseteq [0, \tau] \), \( s_v \) is given by (C.7). By replacing \( s_v \) in the third term of (C.8) by (C.7), we have

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

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 (C.6) for
\[ t \in [n\tau, (n+1)\tau] \text{ as} \]

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

By substituting \( s_v, v \in [i\tau, (i+1)\tau], \ i = 0, 1, \cdots, n - 1 \) into the last term of (C.9), 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 that

\[
s_t = \frac{\tau}{\alpha} (1 - \alpha) \left( r + \mu - \frac{\sigma^2}{2} \right) \left[ \sum_{i=0}^n \left( \sum_{j=0}^i \left( -\frac{\alpha}{\tau} \right)^j (t - i\tau)^j \right) \right] e^{\frac{\alpha t}{\tau}} - n - 1 \\
+ \sum_{i=0}^n \left( -\frac{\alpha}{\tau} \right)^i (t - i\tau)^i e^{\frac{\alpha t}{\tau}} s_0 - \frac{\alpha}{\tau} \int_0^0 \left[ \sum_{i=1}^n \left( -\frac{\alpha}{\tau} \right)^{i-1} (t - i\tau - v)^{i-1} \right] (i-1)! e^{\frac{\alpha (t-\tau-v)}{\tau}} s_v dv \\
- \frac{\alpha}{\tau} \int_{-\tau}^{t-(n+1)\tau} \left[ \left( -\frac{\alpha}{\tau} \right)^n (t - (n+1)\tau - v)^n \right] e^{\frac{\alpha (t-(n+1)\tau-v)}{\tau}} s_v dv \\
+ \sigma \sum_{i=0}^n \int_0^{t-\tau} \left( -\frac{\alpha}{\tau} \right)^i (t - i\tau - v)^i e^{\frac{\alpha (t-\tau-v)}{\tau}} dB_v, \quad t \in [n\tau, (n+1)\tau].
\]

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

\[
\text{Var}_0[\ln(S_t/S_0)] = \text{Var}_0 \left[ \sigma \sum_{i=0}^n \int_0^{t-\tau} \left( -\frac{\alpha}{\tau} \right)^i (t - i\tau - v)^i e^{\frac{\alpha (t-\tau-v)}{\tau}} dB_v \right] \\
= \sigma^2 \text{Var}_0 \left[ \int_0^{t-n\tau} \left( \sum_{i=0}^n \left( -\frac{\alpha}{\tau} \right)^i (t - i\tau - v)^i \right) e^{\frac{\alpha (t-\tau-v)}{\tau}} dB_v \right. \\
+ \int_{t-n\tau}^{t-(n-1)\tau} \left( \sum_{i=0}^{n-1} \left( -\frac{\alpha}{\tau} \right)^i (t - i\tau - v)^i \right) e^{\frac{\alpha (t-\tau-v)}{\tau}} dB_v} + \cdots \\
+ \int_{t-2\tau}^{t-\tau} \left( \sum_{i=0}^{1} \left( -\frac{\alpha}{\tau} \right)^i (t - i\tau - v)^i \right) e^{\frac{\alpha (t-\tau-v)}{\tau}} dB_v \left. + \int_{t-\tau}^{t} e^{2\frac{\alpha v}{\tau}} dB_v \right] \\
= \sigma^2 \left[ \int_0^{t-n\tau} \left( \sum_{i=0}^n \left( -\frac{\alpha}{\tau} \right)^i (t - i\tau - v)^i \right) e^{\frac{\alpha (t-\tau-v)}{\tau}} dv \right. \\
+ \int_{t-n\tau}^{t-(n-1)\tau} \left( \sum_{i=0}^{n-1} \left( -\frac{\alpha}{\tau} \right)^i (t - i\tau - v)^i \right) e^{\frac{\alpha (t-\tau-v)}{\tau}} dv + \cdots \\
+ \int_{t-2\tau}^{t-\tau} \left( \sum_{i=0}^{1} \left( -\frac{\alpha}{\tau} \right)^i (t - i\tau - v)^i \right) e^{\frac{\alpha (t-\tau-v)}{\tau}} dv} + \int_{t-\tau}^{t} e^{2\frac{\alpha v}{\tau}} dv \right].
\]
By changing of variable $u = v - t$, the variance is given by

$$\text{Var}_0[ s_t - s_0] = \sigma^2 \left[ \int_{-\tau}^{0} e^{-\frac{2\alpha}{T} u} du + \int_{-2\tau}^{-\tau} \left( \sum_{i=0}^{\frac{n-1}{2}} \frac{(-\frac{\alpha}{T})^i}{i!} e^\frac{\alpha}{T}(-i\tau-u) \right)^2 du + \cdots 
+ \int_{-n\tau}^{-(n-1)\tau} \left( \sum_{i=0}^{\frac{n-1}{2}} \frac{(-\frac{\alpha}{T})^i}{i!} e^\frac{\alpha}{T}(-i\tau-u) \right)^2 du + \int_{-\tau}^{0} \left( \sum_{i=0}^{\frac{n-1}{2}} \frac{(-\frac{\alpha}{T})^i}{i!} e^\frac{\alpha}{T}(-i\tau-u) \right)^2 du \right].$$

(C.11)

C.4. Proof of Lemma 3.1. It follows from (2.1) that the market price of risk is given by

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

(C.12)

which satisfies the Novikov’s condition

$$\mathbb{E} \left[ \exp \left\{ \frac{1}{2} \int_0^T \theta_t^2 dt \right\} \right] < \infty.$$  

(C.13)

So the state price density is given by

$$\pi_t = \exp \left\{ -\int_0^t rdu - \frac{1}{2} \int_0^t \theta_u^2 du - \int_0^t \theta_u dB_u \right\}.$$  

(C.14)

Define

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

(C.15)

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 $\phi_t$ 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

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

(C.16)

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

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

Then the problem reduces to the unconstrained maximization of

$$\mathbb{E}_0 \left[ \frac{W_T^{1-\gamma}}{1-\gamma} \right] + \lambda \left( W_0 - \mathbb{E}_0[\pi_T W_T] \right),$$  

where $\lambda$ is the Lagrange multiplier. Proofs of this well-known result can be found in Harrison and Kreps (1978), Cox and Huang (1989) and Karatzas and Shreve (1998). The first order condition leads to the following optimal terminal wealth

$$W_T = (\lambda \pi_T)^{-\gamma}.$$  

(C.17)
Define $\bar{\xi} = \mathbb{E}_0[\xi_T^{(\gamma-1)/\gamma}]$. Then

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

and hence the Lagrange multiplier is given by

$$\lambda = \bar{\xi}_0 W_0^{-\gamma} e^{(1-\gamma)rT}. \quad \text{(C.18)}$$

It follows from (C.17) and (C.18) that the value function satisfies

$$V = \mathbb{E}_0 \left[ \frac{W_T^{1-\gamma}}{1-\gamma} \right] = \frac{1}{1-\gamma} W_0^{-\gamma} \bar{\xi}_0 e^{(1-\gamma)rT}. \quad \text{(C.19)}$$

The optimal wealth process is then given by

$$W_t = \pi_t^{-1} \mathbb{E}_t(\pi_TW_T) = W_0 \tilde{\xi}_0 e^{(1-\gamma)\theta_t + \psi_t} \xi_t^{(\gamma-1)/\gamma}. \quad \text{(C.20)}$$

It follows from (C.16) that

$$d(e^{-rt}W_t) = \sigma e^{-rt} \phi_t W_t dB_t^Q. \quad \text{(C.21)}$$

In addition, Ito’s formula implies that

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

where $\psi_t$ is governed by

$$\pi_t W_t = W_0 + \int_0^t \psi_u dB_u. \quad \text{(C.22)}$$

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

$$\phi_t = \frac{\theta_t}{\sigma} + \frac{\psi_t}{\sigma \pi_t W_t} = \frac{\alpha m_t + (1-\alpha)\mu}{\sigma^2} + \sigma^{-1} W_0^{-1} \bar{\xi}_0 \psi_t \left( \mathbb{E}_t \left[ \xi_t^{(\gamma-1)/\gamma} \right] \right)^{-1}. \quad \text{(C.23)}$$

C.5. **Proof of Proposition 3.2.** We rewrite (2.1) as

$$ds_t = \left[ (1-\alpha) \left( r + \mu - \frac{\sigma^2}{2} \right) + \frac{\alpha}{\gamma} (s_t - s_{t-\gamma}) \right] dt + \sigma dB_t, \quad \text{(C.24)}$$

where $s_t = \ln S_t$, and

$$\theta_t = \frac{1}{\sigma} \left[ (1-\alpha) \mu - \alpha \left( r - \frac{\sigma^2}{2} \right) \right] + \frac{\alpha}{\gamma \sigma} (s_t - s_{t-\gamma}). \quad \text{(C.25)}$$

We define a new measure

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \exp \left\{ - \int_t^T \frac{\gamma-1}{\gamma} \theta_u dB_u - \int_t^T \frac{(\gamma-1)^2}{2\gamma^2} \theta_u^2 du \right\}, \quad \text{(C.26)}$$

and under the new measure,

$$ds_t = \left[ \left( r - \frac{\sigma^2}{2} \right)(1-\alpha) + (1-\alpha) \frac{\mu}{\gamma} + \frac{\alpha}{\gamma \sigma} (s_t - s_{t-\gamma}) \right] dt + \sigma dB_t^*, \quad \text{(C.27)}$$
and
\[ \mathbb{E}_0[\xi_{T}^{\gamma-1}] = \mathbb{E}_0^* \left[ \exp \frac{1 - \gamma}{2 \gamma^2 \sigma^2} \int_0^T \left[ (1 - \alpha) \mu - \alpha \left( r - \frac{\sigma^2}{2} \right) + \frac{\alpha}{\tau} (s_u - s_{u-\tau}) \right]^2 du \right] \]. 

(C.27)

Lemma C.1. When \( 0 \leq T \leq \tau \),
\[ \mathbb{E}_0^* \left[ \exp \frac{1 - \gamma}{2 \gamma^2 \sigma^2} \int_0^T \left[ (1 - \alpha) \mu - \alpha \left( r - \frac{\sigma^2}{2} \right) + \frac{\alpha}{\tau} (s_u - s_{u-\tau}) \right]^2 du \right] = \exp \left\{ \frac{A_{1,t} s_0^2}{2} + A_{2,t} s_0 + A_{3,t} \right\}, \]

where
\[ A_{1,t} = \frac{\alpha (\gamma - 1) (1 - e^{2\alpha(T-1)})}{\gamma \sigma^2 \tau \left( \sqrt{\gamma} - 1 \right) e^{2\alpha(T-1)} + (\sqrt{\gamma} + 1)}, \]
\[ A_{2,t} = \int_t^T e^{\int_u^t (\alpha^2 A_{1,u} + \frac{\sigma^2}{2}) du} \left[ \left( \left( r - \frac{\sigma^2}{2} \right) \left( 1 - \frac{\alpha}{\gamma} \right) + (1 - \alpha) \mu - \alpha \gamma \gamma \right) A_{1,u} + \frac{1 - \gamma}{\gamma^2} \left( (1 - \alpha) \mu - \alpha \left( r - \frac{\sigma^2}{2} \right) - \frac{\alpha}{\tau} s_{u-\tau} \right) \right] du, \]
\[ A_{3,t} = \int_t^T \left[ \frac{\sigma^2}{2} A_{2,u} + \frac{\sigma^2}{2} A_{1,u} + \left( \left( r - \frac{\sigma^2}{2} \right) \left( 1 - \frac{\alpha}{\gamma} \right) + (1 - \alpha) \mu - \alpha \gamma \gamma \right) A_{2,u} + \frac{1 - \gamma}{2 \gamma^2 \sigma^2} \left( (1 - \alpha) \mu - \alpha \left( r - \frac{\sigma^2}{2} \right) - \frac{\alpha}{\tau} s_{u-\tau} \right) \right]^2 du. \]

Proof. When \( t \leq T \), \( s_{t-\tau} \) is the realized log price and is known at time 0. Then \( s_t \) in (C.26) can be considered as a Markov process. Denote
\[ f(s, t) = \mathbb{E}_t^* \left[ \exp \frac{1 - \gamma}{2 \gamma^2} \int_t^T \frac{1}{\sigma^2} \left[ (1 - \alpha) \mu - \alpha \left( r - \frac{\sigma^2}{2} \right) + \frac{\alpha}{\tau} (s_u - s_{u-\tau}) \right]^2 du \right]. \]

(C.29)

Feynman-Kac formula implies that
\[ \frac{\partial f}{\partial t} + \left[ \left( r - \frac{\sigma^2}{2} \right) \left( 1 - \frac{\alpha}{\gamma} \right) + (1 - \alpha) \mu \frac{\gamma}{\gamma} + \frac{\alpha}{\gamma} (s - s_{t-\tau}) \right] \frac{\partial f}{\partial s} + \frac{\sigma^2}{2} \frac{\partial^2 f}{\partial s^2} \]
\[ + \frac{1 - \gamma}{2 \gamma^2 \sigma^2} \left( (1 - \alpha) \mu - \alpha \left( r - \frac{\sigma^2}{2} \right) + \frac{\alpha}{\tau} (s - s_{t-\tau}) \right)^2 f = 0. \]  

(C.30)

By conjecturing and substituting
\[ f(s_t, t) = \exp \left\{ \frac{A_{1,t} s_t^2}{2} + A_{2,t} s_t + A_{3,t} \right\}, \]

(C.31)
into (C.30), we have
\[
\dot{A}_{1,t} = -\sigma^2 (A_{1,t})^2 - \frac{2\alpha}{\gamma} A_{1,t} - \frac{1 - \gamma}{\gamma^2} \frac{\alpha^2}{r^2 \sigma^2},
\]
\[
\dot{A}_{2,t} = -\left(\sigma^2 A_{1,t} + \frac{\alpha}{\gamma} \right) A_{2,t} - \left(1 - \alpha\right) \frac{\mu}{\gamma} + \left(\frac{r - \sigma^2}{2}\right) \left(1 - \frac{\alpha}{\gamma}\right) - \frac{\alpha}{\gamma} s_{t-\tau} \right] A_{1,t}
\]
\[
- \frac{1 - \gamma}{\gamma^2} \frac{\alpha}{\sigma^2} \left(1 - \alpha\right) \mu + \alpha \left(\frac{\sigma^2}{2} - r\right) - \frac{\alpha}{T} s_{t-\tau},
\]
\[
\dot{A}_{3,t} = -\frac{\sigma^2}{2} A_{2,t} + \frac{\sigma^2}{2} A_{1,t} - \left(1 - \alpha\right) \frac{\mu}{\gamma} + \left(\frac{r - \sigma^2}{2}\right) \left(1 - \frac{\alpha}{\gamma}\right) - \frac{\alpha}{\gamma} s_{t-\tau} \right] A_{2,t}
\]
\[
- \frac{1 - \gamma}{2\gamma^2 \sigma^2} \left(1 - \alpha\right) \mu + \alpha \left(\frac{\sigma^2}{2} - r\right) - \frac{\alpha}{T} s_{t-\tau},
\]
(3.2)
with terminal conditions \(A_{1,T} = A_{2,T} = A_{3,T} = 0\).

In (3.2), \(A_{1,t}, A_{2,t}\) and \(A_{3,t}\) are governed by ordinary differential equations (ODEs). The realized price \(s_{t-\tau}\) in \(A_2\) and \(A_3\) is continuous but non-differentiable. So (3.2) is a non-autonomous ordinary differential 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 (3.2) is given by
\[
A_{1,t} = \frac{\alpha (\gamma - 1) \left(1 - e^{\frac{2\alpha(T-t)}{\sqrt{\gamma} + 1}}\right)}{\gamma \sigma^2 t \left(\sqrt{\gamma} - 1\right) e^{\frac{2\alpha(T-t)}{\sqrt{\gamma} + 1}} + \left(\sqrt{\gamma} + 1\right)},
\]
\[
A_{2,t} = \int_t^T e^{t-u} \left(\frac{\sigma^2 A_{1,u} + \alpha}{\gamma} \right) du \left[\left(\frac{r - \sigma^2}{2}\right) \left(1 - \frac{\alpha}{\gamma}\right) + \left(1 - \alpha\right) \frac{\mu}{\gamma} - \frac{\alpha}{\gamma} \ln S_{u-\tau}\right] A_{1,u}
\]
\[
+ \frac{1 - \gamma}{\gamma^2} \frac{\alpha}{\sigma^2} \left(1 - \alpha\right) \mu + \alpha \left(\frac{r - \sigma^2}{2}\right) - \frac{\alpha}{\gamma} \ln S_{u-\tau}\right] \right] du,
\]
\[
A_{3,t} = \int_t^T \left[\frac{\sigma^2}{2} A_{2,u} + \frac{\sigma^2}{2} A_{1,u} + \left(\frac{r - \sigma^2}{2}\right) \left(1 - \frac{\alpha}{\gamma}\right) + \left(1 - \alpha\right) \frac{\mu}{\gamma} - \frac{\alpha}{\gamma} \ln S_{u-\tau}\right] A_{2,u}
\]
\[
+ \frac{1 - \gamma}{2\gamma^2 \sigma^2} \left(1 - \alpha\right) \mu + \alpha \left(\frac{r - \sigma^2}{2}\right) - \frac{\alpha}{\gamma} \ln S_{u-\tau}\right) \right] du.
\]
(3.3)

Now we are ready to prove Proposition 3.2.

By substituting (3.1) into (3.20), we have
\[
\frac{dW_t}{W_t} = \left[\frac{\theta_t^2}{\gamma} + \sigma \theta_t (A_{1,t} s_t + A_{2,t}) + r\right] dt + \left[\frac{\theta_t}{\gamma} + \sigma (A_{1,t} s_t + A_{2,t})\right] dB_t,
\]
(3.4)
\[
\psi_t = W_t \pi_t \left[\frac{1 - \gamma}{\gamma} \theta_t + \sigma (A_{1,t} s_t + A_{2,t})\right],
\]
and hence the optimal portfolio weight is given by

\[ \phi_0 = \phi_0^M + \phi_0^H, \]

\[ \phi_0^M = \frac{\theta_0}{\gamma \sigma}, \quad \text{and} \quad \phi_0^H = A_{1,0}s_0 + A_{2,0}, \tag{C.35} \]

and the optimal wealth process is given by

\[ W_t^* = W_0 \xi_t^{-1} e^{rt} \xi_t^{-1/\gamma} \exp \left\{ \frac{A_{1,t}}{2} (\ln S_t)^2 + A_{4,t} \ln S_t + A_{3,t} \right\}. \tag{C.36} \]

It follows from (C.33) that

\[
\phi_0^H = -\frac{1}{\gamma^2} \int_0^{T-\tau} e^{\xi_{u+\tau}^+ \left( \sigma^2 A_{1,\alpha} + \frac{\alpha}{\tau} \right) \xi_u^+} \left\{ \frac{\gamma - 1}{\gamma^2 \sigma^2} \left( \left( 1 - \alpha \right) \mu + \left( \gamma - \alpha \right) \left( r - \frac{\sigma^2}{2} \right) \right) \frac{\tau}{\alpha} \left( -\frac{\alpha}{\gamma^2} A_{1,u+\tau} \right) + \frac{\gamma - 1}{\gamma^2 \sigma^2} \alpha \right\} \xi_u^+ \, du,
\]

\[
= -\frac{1}{\gamma^2} \int_0^{T-\tau} e^{\xi_{u+\tau}^+ \left( \sigma^2 A_{1,\alpha} + \frac{\alpha}{\tau} \right) \xi_u^+} \left\{ \frac{\gamma - 1}{\gamma^2 \sigma^2} \left( \left( 1 - \alpha \right) \mu + \left( \gamma - \alpha \right) \left( r - \frac{\sigma^2}{2} \right) \right) \frac{\tau}{\alpha} \left( -\frac{\alpha}{\gamma^2} A_{1,u+\tau} \right) + \frac{\gamma - 1}{\gamma^2 \sigma^2} \alpha \right\} \xi_u^+ \, du,
\]

\[
= -\int_0^{T-\tau} \left( \int_{-\tau}^{\xi_{u+\tau}^+} e^{\xi_{u+\tau}^+ \left( \sigma^2 A_{1,\alpha} + \frac{\alpha}{\tau} \right) \xi_u^+} \left\{ \frac{\gamma - 1}{\gamma^2 \sigma^2} \left( \left( 1 - \alpha \right) \mu + \left( \gamma - \alpha \right) \left( r - \frac{\sigma^2}{2} \right) \right) \frac{\tau}{\alpha} \left( -\frac{\alpha}{\gamma^2} A_{1,u+\tau} \right) + \frac{\gamma - 1}{\gamma^2 \sigma^2} \alpha \right\} \xi_u^+ \, du \right) \xi_u^+ \, ds_v
\]

\[
- \int_{-\tau}^{T-\tau} \left( \int_{-\tau}^{\xi_{u+\tau}^+} e^{\xi_{u+\tau}^+ \left( \sigma^2 A_{1,\alpha} + \frac{\alpha}{\tau} \right) \xi_u^+} \left\{ \frac{\gamma - 1}{\gamma^2 \sigma^2} \left( \left( 1 - \alpha \right) \mu + \left( \gamma - \alpha \right) \left( r - \frac{\sigma^2}{2} \right) \right) \frac{\tau}{\alpha} \left( -\frac{\alpha}{\gamma^2} A_{1,u+\tau} \right) + \frac{\gamma - 1}{\gamma^2 \sigma^2} \alpha \right\} \xi_u^+ \, du \right) \xi_u^+ \, ds_v
\]

\[
= \int_{-\tau}^{0} \kappa_v \left( \frac{dS_v}{S_v} - rdv \right) + A_4,
\]

where

\[
\kappa_v = \begin{cases} \int_{-\tau}^{v} \hat{\kappa}_u \, du, & -\tau \leq v \leq T - \tau, \\ \int_{-\tau}^{T-\tau} \hat{\kappa}_u \, du, & T - \tau \leq v \leq 0, \end{cases}
\]

\[
\hat{\kappa}_u = e^{\xi_{u+\tau}^+ \left( \sigma^2 A_{1,\alpha} + \frac{\alpha}{\tau} \right) \xi_u^+} \left( \frac{\alpha}{\gamma^2} A_{1,u+\tau} - \frac{\gamma - 1}{\gamma^2 \sigma^2} \alpha \right).
\]
and
\[
A_4 = \int_{-\tau}^{T-\tau} e^{u\tau} (\sigma^2 A_{1,u+\gamma} + \frac{\alpha}{\gamma}) du \left\{ (1-\alpha)\mu + \gamma (r - \frac{\sigma^2}{2}) - \alpha \right\} \tau \left( \frac{\alpha}{\gamma} A_{1,u+\tau} - \gamma - \frac{1}{\gamma^2} \frac{\alpha^2}{\sigma^2 \tau^2} + (r - \frac{\sigma^2}{2}) \frac{(\gamma - 1)\alpha}{\gamma \sigma^2 \tau} \right\} du + r \int_{-\tau}^{0} \kappa_v dv.
\]

Let
\[
C_{0,u} = \frac{\alpha}{\gamma(1-\gamma)} A_{1,u} + \frac{\alpha^2}{\gamma^2 \sigma^2 \tau^2}, \quad \text{and} \quad \omega_v = \frac{\gamma \sigma^2 \kappa_v}{(1-\gamma)\alpha},
\]
then the weights (C.37) can be written as
\[
\omega_v = \begin{cases} 
\int_{-\tau}^{v} \omega_u du, & v \in [-\tau, -\tau + T], \\
\int_{-\tau+T}^{-\tau} \omega_u du, & v \in [-\tau + T, 0],
\end{cases}
\]
\[
\hat{\omega}_u = \frac{\gamma \sigma^2}{\alpha} C_{1,u+\tau} \exp \left\{ \int_{0}^{u+\tau} \left[ \frac{\gamma(1-\gamma)\sigma^2 \tau}{\alpha} C_{0,u} + \frac{\alpha}{\tau} \right] du \right\} > 0,
\]
and
\[
C_{0,u} = \frac{\alpha^2 \left( e^{2\alpha(T-u)/\gamma} + 1 \right)}{\gamma^{3/2} \sigma^2 \tau^2 \left[ (\sqrt{\gamma} - 1) e^{2\alpha(T-u)/\gamma} + (\sqrt{\gamma} + 1) \right]},
\]
\[
C_1 = \int_{0}^{T} \exp \left\{ \int_{0}^{u} \left[ \frac{\gamma(1-\gamma)\sigma^2 \tau}{\alpha} C_{0,u} + \frac{\alpha}{\tau} \right] du \right\} \left\{ \left[ (1-\alpha)\mu - \alpha r \right]/\gamma \right. \\
+ r - \frac{\sigma^2}{2} \frac{\gamma(1-\gamma)\tau}{\alpha} C_{0,u} + \frac{\alpha}{\tau} \left( 1 - \gamma \right) r \frac{\alpha}{\gamma \sigma^2} \int_{-\tau}^{0} \omega_v dv 
\right\} \int_{-\tau}^{u} \omega_u du.
\]

When \( \gamma \to \infty \), to the leading order of \( 1/\gamma \), we have \( \hat{\omega} = \alpha \gamma^{-1} \tau^{-2} \). The weight \( \omega_v \) on the historical instantaneous return \( dS_v/S_v \) becomes
\[
\omega_v = \begin{cases} 
\frac{\alpha}{\gamma \tau} (v + \tau), & v \in [-\tau, -\tau + T], \\
\frac{\alpha}{\gamma \tau} T, & v \in [-\tau + T, 0],
\end{cases}
\]
and the hedging demand \( \phi_0^H \) reduces to (3.10).

C.6. Proof of Corollary 3.3. Using (3.8),
\[
\phi_0 = \phi_0^M + (1-\gamma) \frac{\alpha}{\gamma \sigma^2} \int_{-\tau}^{0} \omega_v \frac{dS_v}{S_v} + C_1
\]
\[
= \phi_0^M + (1-\gamma) \frac{\alpha}{\gamma \sigma^2} \int_{-\tau}^{-\tau+T} \int_{u}^{0} \frac{dS_v}{S_v} \omega_u du + C_1
\]
\[
= \frac{\alpha}{\gamma \sigma^2} \tau (\ln S_0 - \ln S_{-\tau}) + (1-\gamma) \frac{\alpha}{\gamma \sigma^2} \left( \int_{-\tau}^{-\tau+T} \hat{\omega}_u du \ln S_0 - \int_{-\tau}^{-\tau+T} \hat{\omega}_u \ln S_u du \right) + C_2,
\]
where $C_2$ is a constant given by
\[
C_2 = C_1 - (1 - \gamma) \frac{\alpha}{\gamma \sigma^2} \int_{-\tau}^{-\tau+T} \frac{\sigma^2 u \hat{\omega}_u}{2} du + \frac{(1 - \alpha)\mu - \alpha r}{\gamma \sigma^2} + \frac{\alpha}{2\gamma}.
\] (C.41)

C.7. **Proof of Corollary 4.1.** It is easy to check that $\hat{\omega}_v > 0$. Therefore, the weight $\omega_v$ of historical return $dS_v/S_v$ in the hedging demand $\phi^H_0$ is an increasing function of $v$ for $v \in [-\tau, -\tau + T]$ and becomes constant for $v \in [-T, 0]$.

C.8. **Proof of Corollary 4.2.**
\[
\frac{\partial \phi_0}{\partial T} = (1 - \gamma) \frac{\alpha}{\gamma \sigma^2} \int_{-\tau}^{0} \omega(v, T) \left( \frac{dS_v}{S_v} - rdv \right) du + C_4,
\]
where
\[
\int_{-\tau}^{0} \omega(v, T) \left( \frac{dS_v}{S_v} - rdv \right) = \int_{-\tau}^{-\tau+T} \left[ \int_{-\tau}^{0} \hat{\omega}(u, T) \left( \frac{dS_v}{S_v} - rdv \right) du, \right.
\]
and $C_4 = \frac{\partial C_4(T)}{\partial T}$. So
\[
\frac{\partial \left[ \int_{-\tau}^{0} \omega(v, T) \left( \frac{dS_v}{S_v} - rdv \right) \right]}{\partial T} = \int_{-\tau}^{0} \varphi_v \left( \frac{dS_v}{S_v} - rdv \right),
\]
where
\[
\varphi_v = \begin{cases} \int_{-\tau}^{v} \frac{\partial \omega(u, T)}{\partial T} du, & v \in [-\tau, -\tau + T], \\ \hat{\omega}(-\tau + T, T) + \int_{-\tau}^{-\tau+T} \frac{\partial \omega(u, T)}{\partial T} du, & v \in [-T, 0]. \end{cases}
\]

In addition, $\frac{\partial \phi_0}{\partial T}$ can be also written in terms of cumulative returns:
\[
\frac{\partial \phi_0}{\partial T} = \int_{-\tau}^{-\tau+T} \frac{\alpha}{\gamma \sigma^2} \hat{\omega}(u, T) \left( \ln S_0 - \ln S_u \right) du + \frac{\alpha}{\gamma \sigma^2} \hat{\omega}(-\tau + T, T) \left( \ln S_0 - \ln S_{-\tau+T} \right) + C_5,
\]
where $C_5 = \frac{\alpha(\tau-T)\hat{\omega}(\tau, T)}{2\gamma} - \frac{\alpha}{2\gamma} \int_{-\tau}^{-\tau+T} \frac{\partial \omega(u, T)}{\partial T} du$.

C.9. **Price Level Independence of the Optimal Portfolio.**

**Corollary C.2.** When the historical price path $s_u$ is changed to $s_u + c$ for all $u \in [-\tau, 0]$, where $c$ is a constant, $\phi^M_0$ and $\phi^H_0$ do not change. So both demand components depend on historical returns.

**Proof.** At time 0, the optimal portfolio (3.8) only consists of the historical prices over $[-\tau, 0]$, so we examine that the historical path changes from $s_u$ to $s_u + c$ for $u \in [-\tau, 0]$, where $c$ is a constant. The price trend $s_0 - s_{-\tau}$ in the myopic demand is still the same and hence (C.4) implies that the myopic demand does not change.

It follows from (C.12), (C.15) and (C.29) that a constant change in the historical path does not affect $\mathbb{E}_t \left[ S_t^{(\gamma-1)/\gamma} \right]$. Then (C.20) implies that the constant path change affects $W_t$ only via $W_0$. So the change does not affect the total demand $\phi_0$ according to (C.21).
Therefore, both myopic demand and intertemporal hedging demand do not react to a change in the level of historical prices. But the weights on historical price $s_u$ in the hedging demand are different for different $u$, so the optimal portfolio weight is affected by the historical patterns in terms of returns. □

C.10. Certainty Equivalent Wealth of Myopic Momentum Strategy. The wealth process of an investor using the myopic momentum strategy follows

$\frac{dW_t}{W_t} = (r + \sigma \theta^m_t)dt + \sigma \phi^m_t dB_t$, \hspace{1cm} (C.42)

The corresponding value function is given by

$V^m_t = E_t \left[ \frac{W^{1-\gamma}_T}{1-\gamma} \right]$

$= e^{(1-\gamma) \tau \gamma} W^{1-\gamma}_t E_t \left[ \exp \left\{ \int_t^T \left( 2(2-1)(1-\gamma) \frac{\theta^2}{2\gamma^2} du + \frac{1-\gamma}{\gamma} \theta dB_u \right) \right\} \right]$

$= e^{(1-\gamma) \tau \gamma} W^{1-\gamma}_t E_t \left[ \exp \left\{ \frac{1-\gamma}{2\gamma} \int_t^T \left( (1-\alpha) \mu - \alpha \left( r - \frac{\sigma^2}{2} \right) + \alpha \tau (s_u - s_{u-t}) \right)^2 du \right\} \right]$, \hspace{1cm} (C.43)

where the last equality follows from the change of measure (C.25).

When $T - \tau \leq u \leq T$, Feynman-Kac formula implies that

$\frac{\partial V^m_t}{\partial t} + \left[ \left( r - \frac{\sigma^2}{2} \right) \left( 1 - \frac{\alpha}{\gamma} \right) + (1-\alpha) \frac{\mu}{\gamma} + \frac{\alpha}{\gamma \tau} (s - s_{t-\tau}) \right] \frac{\partial V^m_t}{\partial s} + \frac{\sigma^2}{2} \frac{\partial^2 V^m_t}{\partial s^2}$

$+ \frac{1-\gamma}{2\gamma \sigma^2} \left( (1-\alpha) \mu - \alpha \left( r - \frac{\sigma^2}{2} \right) + \alpha \tau (s - s_{t-\tau}) \right)^2 V^m_t = 0$. \hspace{1cm} (C.44)

The solution to (C.44) is given by

$V^m_t(s_t, t) = \exp \left\{ \frac{A^m_t s_t^2}{2} + A^m_t s_t + A^m_{3,t} \right\}$, \hspace{1cm} (C.45)
where $A_{1,t}^m, A_{2,t}^m$ and $A_{3,t}^m$ are governed by the ordinary differential equations (ODEs)

\[\dot{A}_{1,t}^m = -\sigma^2(A_{1,t}^m)^2 - \frac{2\alpha}{\gamma} A_{1,t}^m - \frac{1 - \gamma}{\gamma} \frac{\alpha^2}{\tau^2 \sigma^2},\]

\[\dot{A}_{2,t}^m = -\left(\sigma^2 A_{1,t}^m + \frac{\alpha}{\gamma T}\right) A_{2,t}^m - \left[(1 - \alpha)\frac{\mu}{\gamma} + \left(r - \frac{\sigma^2}{2}\right) (1 - \alpha) - \frac{\alpha}{\gamma T} s_{t-\tau}\right] A_{1,t}^m\]

\[-\frac{1 - \gamma}{\gamma} \frac{\alpha}{\sigma^2 T} \left[(1 - \alpha)\mu + \frac{\sigma^2}{2} - r\right] - \frac{\alpha}{\tau} s_{t-\tau}\right],\]

\[\dot{A}_{3,t}^m = -\frac{\sigma^2}{2} (A_{2,t}^m)^2 - \frac{\sigma^2}{2} A_{1,t}^m - \left[(1 - \alpha)\frac{\mu}{\gamma} + \left(r - \frac{\sigma^2}{2}\right) (1 - \alpha) - \frac{\alpha}{\gamma T} s_{t-\tau}\right] A_{2,t}^m\]

\[-\frac{1 - \gamma}{2\gamma \sigma^2} \left[(1 - \alpha)\mu + \frac{\sigma^2}{2} - r\right] - \frac{\alpha}{\tau} s_{t-\tau}\right]^2,\]

with terminal conditions $A_{1,T}^m = A_{2,T}^m = A_{3,T}^m = 0$.

The certainty equivalent wealth satisfies

\[
\text{CEW}_{1-\gamma}^\gamma = \frac{\text{CEW}_{1-\gamma}^\gamma}{1 - \gamma} = \frac{1}{1 - \gamma},
\]

which is equivalent to

\[
\text{CEW}_{1-\gamma}^\gamma = e^{rT}(V_0)_{1-\gamma}^1 = e^{rT} \exp\left\{\left(\frac{A_{1,0}^m s_0^2}{2} + A_{2,0}^m s_0 + A_{3,0}^m\right)/(1 - \gamma)\right\}.
\]

**APPENDIX D. MONTE CARLO SIMULATION METHOD FOR $T > \tau$**

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 (2.1)-(2.2). The conditional expectation $E_t[\xi_T^{(\gamma-1)/\gamma}]$ in Proposition 3.1 is just the average of $\xi_T^{(\gamma-1)/\gamma}$. The conditional expectation $E_{t+dt}[\xi_T^{(\gamma-1)/\gamma}]$ 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 $\hat{s}_{t+dt}$, $\hat{s}_{t+dt}^2$ or $\hat{s}_{t+dt}^3$, cannot affect the results. Then $\psi_t$ in (3.5) can be derived by regressing

\[d(\pi_t W_t) = W_0 \xi_0^{-1} \left(E_{t+dt}[\xi_T^{(\gamma-1)/\gamma}] - E_t[\xi_T^{(\gamma-1)/\gamma}]\right)\]

on $d\hat{B}_t$. The total demands follow (3.4).
REFERENCES


