Dynamic Portfolio Choice with Frictions∗

Nicolae Gărleanu and Lasse Heje Pedersen†

March, 2013

Abstract

We show that the optimal portfolio can be derived explicitly in a large class of models with transitory and persistent transaction costs, multiple assets, general correlation structure, and multiple return predictors with general dynamics. Our tractable continuous-time model is shown to be the limit of discrete-time models with endogenous transaction costs due to optimal dealer behavior. Depending on the dealers’ inventory dynamics, we show that transitory transaction costs survive, respectively vanish, in the limit, corresponding to an optimal portfolio with bounded, respectively quadratic, variation. Finally, we provide several economic applications and equilibrium implications.

∗We are grateful for helpful comments from Kerry Back, Darrell Duffie, Pierre Collin-Dufresne, Andrea Frazzini, Esben Hedegaard, David Lando, Hong Liu (discussant), Anthony Lynch, Ananth Madhavan (discussant), Andrei Shleifer, and Humbert Suarez, as well as from seminar participants at Stanford Graduate School of Business, University of California at Berkeley, Columbia University, NASDAQ OMX Economic Advisory Board Seminar, University of Tokyo, New York University, the University of Copenhagen, Rice University, University of Michigan Ross School, Yale University SOM, the Bank of Canada, and the Journal of Investment Management Conference.

†Gărleanu is at Haas School of Business, University of California, Berkeley, NBER, and CEPR; e-mail: garleanu@haas.berkeley.edu. Pedersen (corresponding author) is at New York University, Copenhagen Business School, the FRIC Center, AQR Capital Management, CEPR, and NBER, 44 West Fourth Street, NY 10012-1126; e-mail: lpederse@stern.nyu.edu, http://www.stern.nyu.edu/~lpederse/.
A fundamental question in financial economics is how to choose an optimal portfolio. Investors must consider the risks, correlations, expected returns, and transaction costs of all their available assets and their portfolio choice is a dynamic problem for several reasons: First, expected returns are driven by a number of economic factors that vary over time, leading to variation in the optimal portfolio.\(^1\) Second, transaction costs imply that an investor must consider the portfolio’s optimality both now and in the future. Third, investors continually face these trade-offs.

We provide a general and tractable framework to address these issues, deriving a simple expression for the optimal portfolio choice in light of all these dynamic considerations. Further, we show how the continuous-time solution is approached by discrete-time models in which transaction costs are modeled endogenously. We provide several additional applications of the framework and derive implications for equilibrium expected returns.

Our framework’s innovation is to consider a continuous-time model in which transaction costs are quadratic in the number of securities traded. The natural interpretation of a quadratic cost is that price impact is linear in the trade size, resulting in a quadratic cost. This assumption makes our framework highly tractable, allowing us to provide a closed-form optimal portfolio choice with multiple assets, multiple return-predicting factors, and general correlation structure. The tractability of our framework contrasts that of standard models in the literature. Indeed, models of proportional transaction costs are complex and rely on numerical solutions even in the case of a single asset with i.i.d. returns (i.e., no return predicting factors).\(^2\) In discrete time, quadratic costs have been shown to provide tractability and we rely in particular on Gârleanu and Pedersen (2013). (See also Heaton and Lucas (1996) and Grinold (2006).) However, it has been questioned whether market impact costs apply in continuous time or vanish in the limit. E.g., in the model of Cetin,

\(^1\)See, e.g, Campbell and Viceira (2002) and references therein.
\(^2\)There is an extensive literature on proportional transaction costs following Constantinides (1986). Davis and Norman (1990) provide a more formal analysis and Liu (2004) determines the optimal trading strategy for an investor with constant absolute risk aversion (CARA) and many independent securities with both fixed and proportional costs (without predictability). The assumptions of CARA and independence across securities imply that the optimal position for each security is independent of the positions in the other securities.
Jarrow, Protter, and Warachka (2006), transaction costs are irrelevant in continuous time. We provide a economic foundation for quadratic transaction costs that matter in continuous time. This is important for several reasons: First, if transaction costs did not matter in continuous time, then it would imply that the discrete-time models either rely heavily on the length of the time period or that transaction costs also have a small effect in these models, especially because investors’ trading frequencies are endogenous in the real world. Second, is it important to understand how models of different period length are connected and how parameters should be scaled as a function of the period length. Third, our continuous-time model is more tractable than its discrete-time counterpart and the continuous-time framework opens the door for further applications with all the usual benefits of continuous time.

To provide an economic foundation for a continuous-time model with transaction costs, we discretize the model and let transaction costs arise endogenously due to dealers’ inventory considerations a la Grossman and Miller (1988). We consider both persistent and transitory costs, corresponding to dealers who can lay off their inventory gradually or in one shot. We show that the discrete-time persistent market impact costs converge to a continuous-time model with the same persistent market impact parameter and a resiliency parameter which depends on the length of the time periods to the first order.

There are two ways to model how transitory costs depend on the trading frequency: (a) If dealers can always lay off their inventory in one time period, then shorter time periods imply that dealers need only hold inventories shorter time and, in this case, transitory costs vanish in the limit; (b) If, instead, the time it takes dealers to unload inventories does not go to zero even as trading frequencies increase, then transitory costs survive in the limit. In this case, the limit transaction costs are quadratic in the trading intensity, i.e., the number of shares traded per time unit.

We show that both trading costs and the optimal portfolio converge to their continuous-time counterparts as trading frequencies increase. In the case with vanishing transitory costs, inventory models with multiple correlated assets include Greenwood (2005) and Gârleanu, Pedersen, and Poteshman (2009).
the optimal continuous-time portfolio has quadratic variation. With transitory costs, however, our optimal continuous-time strategy is smooth and has a finite turnover. Our optimal strategy is qualitatively different from the strategy with proportional or fixed transaction costs, which exhibits long periods of no trading. Our strategy mimics a trader who is continuously posting limit orders close to the mid-quote, a strategy that is used in practice. The trading speed (the limit orders’ “fill rate” in our analogy) depends on how large transaction costs the trader is willing to accept (i.e., on where the limit orders are placed) as in our model. Our strategy has several advantages in the real world according to discussions with people who design trading systems: Trading continuously minimizes the order sizes at each point in time and exploits the liquidity that is available throughout the day/week/month, rather than submitting large infrequent orders when limited liquidity may be available. Consistently, the empirical literature generally finds transaction costs to be convex (e.g., Engle, Ferstenberg, and Russell (2008), Lillo, Farmer, and Mantegna (2003)), with some researchers estimating quadratic trading costs (e.g., Breen, Hodrick, and Korajczyk (2002) and Kyle and Obizhaeva (2011)).

The tractability of our framework makes it a potentially powerful “workhorse” for other applications involving transactions costs. As one such application, we embed the continuous-time model in an equilibrium setting. Rational investors facing transaction costs trade with several groups of noise traders who provide a time-varying excess supply or demand of assets. We show that, in order for the market to clear, the investors must be offered return premia depending on the properties of the noise-traders’ positions. In particular, the noise trader positions that mean revert more quickly generate larger alphas in equilibrium, as the rational investors must be compensated for incurring higher transaction costs per time unit. Long-lived supply fluctuations, on the other hand, give rise to smaller and more persistent alphas. This can help explain the short-term return reversals documented by Lehman (1990) and Lo and MacKinlay (1990), and their relation to transaction costs documented by Nagel (2011).

Finally, our work relates to several strands of literature in addition to the research cited above. One strand of literature studies equilibrium asset pricing with trading costs (Amihud
and Mendelson (1986), Vayanos (1998), Vayanos and Vila (1999), Lo, Mamaysky, and Wang (2004), Jang, Koo, Liu, and Loewenstein (2007), and Gârleanu (2009)) and time-varying trading costs (Acharya and Pedersen (2005), Lynch and Tan (2011)). Second, a strand of literature derives the optimal trade execution, treating the asset and quantity to trade as given exogenously (see, e.g., Perold (1988), Bertsimas and Lo (1998), Almgren and Chriss (2000), Obizhaeva and Wang (2006), and Engle and Ferstenberg (2007)). Finally, quadratic programming techniques are also used in macroeconomics and other fields, and, usually, the solution comes down to algebraic matrix Riccati equations (see, e.g., Ljungqvist and Sargent (2004) and references therein). We solve our model explicitly, including the Riccati equations.

The rest of the paper is organized as follows. Section 1 lays out our continuous-time framework and solves the model with transitory and persistent transaction costs. Section 2 provides a discrete-time foundation for the model, providing endogenous transaction costs and deriving the limit as the length of the time periods becomes small. Section 3 applies the framework to derive a number of economic implications. Section 5 concludes and all proofs are in appendix.

1 Continuous-Time Model

We start by introducing our tractable continuous-time framework and illustrating its solution. We first consider the case of purely transitory transaction costs, then introduce persistent transaction costs, and finally consider the case of purely persistent costs.

1.1 Purely Temporary Transaction Costs

An investor must choose an optimal portfolio among $S$ risky securities and a risk-free asset. The risky securities have prices $p$ with dynamics

$$dp_t = (r^f p_t + B f_t) dt + du_t, \tag{1}$$
Here, $f_t$ is a $K \times 1$ vector which contains the factors that predict excess returns, $B$ is an $S \times K$ matrix of factor loadings, and $u$ is an unpredictable “noise term,” i.e., a martingale (e.g., a Brownian motion) with instantaneous variance-covariance matrix $\text{var}_t(du_t) = \Sigma dt$. The return-predicting factors have dynamics given by

$$df_t = -\Phi f_t dt + d\varepsilon_t,$$

where $\Phi$ is a $K \times K$ matrix of mean-reversion coefficients, and the noise term $\varepsilon$ is a martingale with instantaneous variance-covariance matrix $\text{var}_t(d\varepsilon_t) = \Omega dt$. We require that the jump component of $\varepsilon$, $\Delta \varepsilon$, have mean zero, and we impose on $\Phi$ standard conditions sufficient to ensure that $f$ is stationary.

The agent chooses his trading intensity $\tau_t \in \mathbb{R}^S$, which determines the rate of change of his position $x_t$:

$$dx_t = \tau_t dt.$$  \hspace{1cm} (3)

The transaction cost $TC$ per time unit of trading with intensity $\tau_t$ is

$$TC(\tau_t) = \frac{1}{2} \tau_t^\top \Lambda \tau_t.$$  \hspace{1cm} (4)

Here, $\Lambda$ is a symmetric positive-definite matrix measuring the level of trading costs.\footnote{We only consider smooth portfolio policies here because discrete jumps in positions or quadratic variation would be associated with infinite trading costs in this setting. E.g., if the agent trades $n$ shares over a time period of $\Delta t$, then the cost according to (4) is $\int_0^{\Delta t} TC(\tau_t) dt = \frac{1}{2} \Lambda \frac{n^2}{\Delta t}$, which approaches infinity as $\Delta t$ approaches 0.}  This quadratic transaction cost arises as the trade $\Delta x_t$ shares moves the price by $\frac{1}{2} \Lambda \Delta x_t$, and this results in a total trading cost of $\Delta x_t$ times the price move. This is a multi-dimensional version of Kyle’s lambda. Most of our results hold with this general transaction cost function, but some of the resulting expressions are simpler in the following special case.

\footnote{The assumption that $\Lambda$ is symmetric is without loss of generality. To see this, suppose that $TC(\Delta x_t) = \frac{1}{2} \Delta x_t^\top \Lambda \Delta x_t$, where $\Lambda$ is not symmetric. Then, letting $\Lambda$ be the symmetric part of $\Lambda$, i.e., $\Lambda = (\Lambda + \Lambda^\top)/2$, generates the same trading costs as $\Lambda$.}
**Assumption A.** Transaction costs are proportional to the amount of risk: $\Lambda = \lambda \Sigma$ for a scalar $\lambda > 0$.

This assumption is natural and, in fact, implied by the model of Găreleanu, Pedersen, and Poteshman (2009) as well as the micro-foundation that we provide in Section 2.2. To understand this, suppose that a dealer takes the other side of the trade $\Delta x_t$, holds this position for a period of time $dt$, and “lays it off” at the end of the period. Then the dealer’s risk is $\Delta x_t^\top \Sigma \Delta x_t dt^2$ and the trading cost is the dealer’s compensation for risk, depending on the dealer’s risk aversion reflected by $\lambda$. Section 2.2 further analyzes the conditions under which the compensation for risk is strictly positive.

The investor chooses his optimal trading strategy to maximize the present value of the future stream of expected excess returns, penalized for risk and trading costs:

$$
\max_{(\tau_s)_{s \geq t}} \mathbb{E}_t \int_t^\infty e^{-\rho(s-t)} \left( x_s^\top B f_s - \frac{\gamma}{2} x_s^\top \Sigma x_s - \frac{1}{2} \tau_s^\top \Lambda \tau_s \right) ds.
$$

This objective function means that the investor has mean-variance preferences over the change in his wealth $W_t$ each time period.

We conjecture and verify that the value function is quadratic:

$$
V(x, f) = -\frac{1}{2} x^\top A_{xx} x + x^\top A_{xf} f + \frac{1}{2} f^\top A_{ff} f + A_0.
$$

The following proposition states the solution to the model.

**Proposition 1**  
(i) There exists a unique optimal portfolio strategy.  
(ii) The optimal portfolio $x_t$ tracks a moving “aim portfolio” $\bar{M}^{aim} f_t$ with a tracking speed of $\bar{M}^{rate}$. That is, the optimal trading intensity $\tau_t = \frac{dx_t}{dt}$ is

$$
\tau_t = \bar{M}^{rate} (\bar{M}^{aim} f_t - x_t),
$$

7
where coefficient matrices are given by

$$\bar{M}^{rate} = \Lambda^{-1} A_{xx}$$  \hspace{1cm} (8)

$$\bar{M}^{aim} = A_{xx}^{-1} A_{xf}$$ \hspace{1cm} (9)

$$A_{xx} = -\frac{\rho}{2} \Lambda + \Lambda^\frac{1}{2} \left( \gamma \Lambda^{-\frac{1}{2}} \Sigma \Lambda^{-\frac{1}{2}} + \frac{\rho^2}{4} I \right)^\frac{1}{2} \Lambda^\frac{1}{2} $$ \hspace{1cm} (10)

$$\text{vec}(A_{xf}) = (\rho I + \Phi^\top \otimes I_K + I_S \otimes (A_{xx} \Lambda^{-1}))^{-1} \text{vec}(B).$$ \hspace{1cm} (11)

(iii) Under Assumption A, the solution simplifies: $A_{xx} = a \Sigma$ and

$$\bar{M}^{rate} = a/\lambda = \frac{1}{2} (\sqrt{\rho^2 + 4 \gamma/\lambda} - \rho)$$ \hspace{1cm} (12)

$$\bar{M}^{aim} = \gamma^{-1} \Sigma^{-1} B (I + a/\gamma \Phi)^{-1}.$$ \hspace{1cm} (13)

1.2 Temporary and Persistent Transaction Costs

We modify the set-up above by adding persistent transaction costs. Specifically, the agent transacts at price $\bar{p}_t = p_t + D_t$, where the distortion $D_t$ evolves according to

$$dD_t = -RD_t dt + Cdx_t = -RD_t dt + C\tau_t dt.$$ \hspace{1cm} (14)

The agent’s objective now becomes

$$\max_{(\tau_s)_{s\geq t}} E_t \int_t^\infty e^{-\rho(s-t)} \left( x_s^\top (B f_s - (r + R) D_s + C\tau_s) - \gamma x_s^\top \Sigma x_s - \frac{1}{2} \tau_s^\top \Lambda \tau_s \right) ds.$$ \hspace{1cm} (15)

We conjecture a quadratic value function, as before, in the state variable $(x_t, y_t)$, where we define $y = (f_t, D_t)$. Specifically, we write $V(x, y) = -\frac{1}{2} x^\top A_{xx} x + x^\top A_{xy} y + \frac{1}{2} y^\top A_{yy} y + A_0$.

**Proposition 2** The optimal trading intensity has the form

$$\tau_t = \bar{M}^{rate} (\bar{M}^{aim} y_t - x_t)$$ \hspace{1cm} (16)

for appropriate matrices $\bar{M}^{rate}$ and $\bar{M}^{aim}$ defined in the proof.
1.3 Purely Persistent Costs

The set-up is as above, but now we take Λ = 0. Under this assumption, it no longer follows that \( x_t \) has to be of the form \( dx_t = \tau_t dt \) for some \( \tau \). Indeed, with purely persistent price-impact costs, the optimal portfolio policy can have jumps and infinite quadratic variation (i.e., “wiggle” like a Brownian motion).

As before, \( D \) is the price distortion and it evolves as
\[
dD_t = -RD_t dt + Cdx_t. \tag{17}
\]

We define the objective of the trader to be
\[
E_t \int_t^\infty e^{-\rho(s-t)} \left( x_s^\top (\alpha_s - (r + R)D_s) - \frac{\gamma}{2} x_s^\top \Sigma x_s \right) ds \tag{18}
+ E_t \int_t^\infty e^{-\rho(s-t)} x_s^\top Cdx_s + \frac{1}{2} E_t \int_t^\infty e^{-\rho(s-t)} d[x, Cx]_s.
\]

The terms in the first row of (18) are as before. The second row captures the contemporaneous effect of trading on the price. Thus, the first term states that the mark-to-market price changes by \( Cdx \), applicable on the position \( x_{s-} \); the second term records the instantaneous mark-to-market gain on the just-purchased units \( dx_s \). While it might appear odd that the agent seems to benefit from buying and pushing the price higher, this benefit leads to a loss as the distortion \( D \) decays.

A helpful observation in this case is that making a large trade \( \Delta x \) over an infinitesimal time interval has an easily described impact on the value function. In fact, the ability to liquidate one’s position instantaneously, and then take a new position, at no cost relative to trading directly to the new position implies
\[
V(x, D, f) = V(0, D - Cx, f) - \frac{1}{2} x^\top Cx. \tag{19}
\]

We prove this intuitive conjecture by providing a verification argument for the optimal control and value function that we propose.

Before stating the result for this version of the model, we introduce the following natural
counterpart to Assumption A.

**Assumption B.** The persistent price impact is proportional to the level of risk, $C = c\Sigma$, where $c$ is a scalar. The resiliency $R$ is a scalar.

**Proposition 3** A quadratic value function exists of the form (A.38) in the appendix. The optimal portfolio is given by

$$x = J^{-1} \left[ \left( B - C^T R^T A_{Df} \right) f - \left( (r^f + R) + C^T R^T A_{DD} \right) \left( D - Cx_1 \right) \right] ,$$

where $A_{DD}$ and $A_{Df}$ are value-function coefficients given as solutions to (A.46) and (A.47), respectively.

## 2 Discrete-Time Foundation

In this section we consider the discrete-time counterpart to our continuous-time model and show that, under suitable conditions, the latter is the limit of the former as the length of time between trading dates goes to zero.

### 2.1 Discrete-Time Model and Solution

We start by presenting a discretely-sampled version of the continuous-time model. Securities are now traded at dates indexed by $t \in \{0, 1, 2, \ldots\}$, corresponding to calendar times $0, \Delta t, 2\Delta t, \ldots$, where $\Delta t$ is the length of the time periods. The securities’ price changes between times $t$ and $t+1$ in excess of the risk-free return, $p_{t+1} - (1 + r^f \Delta t)p_t$, are collected in an $S \times 1$ vector $r_{t+1}$. As before, excess returns can be predicted by the factors $f_t$:

$$r_{t+1} = B f_t \Delta t + u_{t+1} .$$

(21)
where $u_{t+1}$ is the unpredictable zero-mean noise term with variance $\text{var}(u_{t+1}) = \Sigma \Delta t$. Naturally, the returns and variance scale linearly in time, $\Delta t$. The return-predicting factor $f_t$ is known to the investor at time $t$ and it evolves according to

$$
\Delta f_{t+1} = -\Phi f_t \Delta t + \varepsilon_{t+1},
$$

where $\Delta f_{t+1} = f_{t+1} - f_t$ is the change in the factors, $\Phi$ is the matrix of mean-reversion coefficients, and $\varepsilon_{t+1}$ is the factor shock with variance $\text{var}(\varepsilon_{t+1}) = \Omega \Delta t$.

An investor in the economy faces transaction costs. The transaction cost ($TC$) associated with trading $\Delta x_t = x_t - x_{t-1}$ shares is given by

$$
TC(\Delta x_t) = \frac{1}{2} \Delta x_t^\top \Lambda(\Delta t) \Delta x_t,
$$

where $\Lambda(\Delta t)$ is the matrix of transitory market impact costs. The literature does not offer guidance for how $\Lambda(\Delta t)$ depends on $\Delta t$. (Of course, the dependence of the statistical-distribution parameters on the length of the time intervals follows directly from the fact that we consider a discretized version of the continuous-time model.) To address this issue, Section 2.2 provides this dependence of transaction costs on $\Delta t$ in a model of endogenous dealer behavior.

The investor’s objective is to choose the dynamic trading strategy $(x_0, x_1, \ldots)$ to maximize the present value of all future expected excess returns, penalized for risks and trading costs:

$$
\max_{x_0, x_1, \ldots} E_0 \left[ \sum_t (1 - \rho \Delta t)^{t+1} \left( x_t^\top r_{t+1} - \frac{\gamma}{2} x_t^\top \Sigma \Delta t x_t \right) - \frac{(1 - \rho \Delta t)^{t}}{2} \Delta x_t^\top \Lambda \Delta x_t \right],
$$

where the discount rate is $\rho \Delta t$ with $\rho \in (0, 1)$, and $\gamma$ is the risk-aversion coefficient (which naturally does not depend on $\Delta t$).

Gârleanu and Pedersen (2013) solve this discrete-time model using dynamic programming, but we re-derive the solution here for completeness. The value function $V(x_{t-1}, f_t)$ measures the value of entering period $t$ with a portfolio of $x_{t-1}$ securities and observing
return-predicting factors $f_t$. It solves the Bellman equation:

$$V(x_{t-1}, f_t) = \max_{x_t} \left\{ -\frac{1}{2} \Delta x_t^\top \Lambda \Delta x_t + (1 - \rho) \left( x_t^\top E_t[r_{t+1}] - \frac{\gamma}{2} x_t^\top \Sigma x_t + E_t[V(x_t, f_{t+1})] \right) \right\}. \quad (25)$$

The model has a unique solution and can be solved explicitly:

**Proposition 4 (Discrete-Time Solution with Transitory Costs)** The optimal portfolio $x_t$ tracks an “aim portfolio”, $M^{aim}(\Delta t)f_t$, with trading rate $M^{rate}(\Delta t)$:

$$\Delta x_t = M^{rate}(\Delta t) \left( M^{aim}(\Delta t)f_t - x_{t-1} \right), \quad (26)$$

where the coefficients are given by the value-function coefficients, made explicit in the appendix:

$$M^{rate}(\Delta t) = \Lambda^{-1} A_{xx} \quad (27)$$

$$M^{aim}(\Delta t) = A_{xx}^{-1} A_{xf}. \quad (28)$$

**Transitory and persistent transaction costs.**

To study the more general model with transitory and persistent transaction costs, we extend the model by letting the price be given by $\bar{p}_t = p_t + D_t$ and the investor incur the cost associated with the persistent price distortion $D_t$ in addition to the temporary trading cost $TC$ from before. Hence, the price $\bar{p}_t$ is the sum of the price $p_t$ without the persistent effect of the investor’s own trading (as before) and the new term $D_t$, which captures the accumulated price distortion due to the investor’s (previous) trades. Trading an amount $\Delta x_t$ pushes prices by $C\Delta x_t$ such that the price distortion becomes $D_t + C\Delta x_t$, where $C(\Delta t)$ is Kyle’s lambda for persistent price moves. Further, the price distortion mean reverts at a speed (or “resiliency”) $R(\Delta t)$. Section 2.2 shows how $C$ and $R$ depend on $\Delta t$. Given the persistent
price impact and resilience, the price distortion next period \((t + 1)\) is:

\[
D_{t+1} = (I - R) (D_t + C \Delta x_t).
\]  

(29)

The investor’s objective is as before, with a natural modification due to the price distortion:

\[
E_0 \left[ \sum_t (1 - \rho \Delta t)^{t+1} \left( x_t^\top \left[ B f_t - (R + r^f) (D_t + C \Delta x_t) \right] \Delta t - \frac{\gamma}{2} x_t^\top \Sigma x_t \Delta t \right) + (1 - \rho \Delta t)^t \left( -\frac{1}{2} \Delta x_t^\top \Lambda \Delta x_t + x_{t-1}^\top C \Delta x_t + \frac{1}{2} \Delta x_t^\top C \Delta x_t \right) \right].
\]  

(30)

Let us explain the various new terms in this objective function. The first term is the position \(x_t\) times the expected excess return of the price \(\bar{p}_t = p_t + D_t\) given inside the inner square brackets. As before, the expected excess return of \(p_t\) is \(B f_t\). The expected excess return due to the post-trade price distortion is

\[
D_{t+1} - (1 + r^f \Delta t)(D_t + C \Delta x_t) = -(R + r^f) (D_t + C \Delta x_t) \Delta t.
\]  

(31)

The second term is the penalty for taking risk as before. The three terms on the second line of (30) are discounted at \((1 - \rho)^t\) because these cash flows are incurred at time \(t\), not time \(t + 1\). The first of these is the temporary transaction cost as before. The second reflects the mark-to-market gain from the old position \(x_{t-1}\) from the price impact of the new trade, \(C \Delta x_t\). The last term reflects that the traded shares \(\Delta x_t\) are assumed to be executed at the average price distortion, \(D_t + \frac{1}{2} C \Delta x_t\). Hence, the traded shares \(\Delta x_t\) earn a mark-to-market gain of \(\frac{1}{2} \Delta x_t^\top C \Delta x_t\) as the price moves up an additional \(\frac{1}{2} C \Delta x_t\).

The value function is now quadratic in the extended state variable \((x_{t-1}, y_t) \equiv (x_{t-1}, f_t, D_t)\):

\[
V(x, y) = -\frac{1}{2} x^\top A_{xx} x + x^\top A_{xy} y + \frac{1}{2} y^\top A_{yy} y + A_0.
\]

As before, there exists a unique solution to the Bellman equation and the following proposition characterizes the optimal portfolio strategy.
Proposition 5 (General Discrete-Time Solution) The optimal portfolio $x_t$ is

$$
\Delta x_t = M^{rate} (M^{aim} y_t - x_{t-1}),
$$

which tracks an aim portfolio, $M^{aim} y_t$, that depends on the return-predicting factors and the price distortion, $y_t = (f_t, D_t)$. The coefficient matrices $M^{rate}(\Delta t)$ and $M^{aim}(\Delta t)$, which depend on the length $\Delta t$ of the time periods, are stated in the appendix.

2.2 Foundation for Transaction-Cost Specifications

We next consider the economic foundation for the quadratic transaction cost, the dependence on the trading frequency, and the limit as the trading frequency increases.

Transitory transaction costs.
To obtain a temporary price impact of trades endogenously, we consider an economy populated by three types of investors: (i) the trader whose optimization problem we study in the paper, referred throughout this section as “the trader,” (ii) “market makers,” who act as intermediaries, and (iii) “end users,” on whom market makers eventually unload their positions as described below.

The temporary price impact is due to the market makers’ inventories. We assume that there are a mass-one continuum of market makers indexed by the set $[0, h]$ and they arrive for the first time at the market at a time equal to their index. The market operates only at discrete times $\Delta t$ apart, and the market makers trade at the first trading opportunity. Once they trade — say, at time $t$ — market makers must spend $h$ units of time gaining access to end users. At time $t + h$, therefore, they unload their inventories at a price $p_{t+h}$ described below, and rejoin the market immediately thereafter. It follows that at each trading date in the market there is always a mass $\frac{\Delta t}{h}$ of competing market makers that clear the market.

The price $p$, the competitive price of end users, follows an exogenous process and corresponds to the fundamental price in the body of the paper. Market makers take this price as

---

6 We make the simplifying assumption that $\frac{h}{\Delta t}$ is an integer.
given and trade a quantity \( q \) to maximize a quadratic utility:

\[
\max_q \left\{ \hat{E}_t \left[ q(p_{t+h} - e^{r_h \hat{p}_t}) \right] - \frac{\gamma^M}{2} Var_t \left[ q(p_{t+h} - e^{r_h \hat{p}_t}) \right] \right\},
\]

(33)

where \( \hat{p}_t \) is the market price at time \( t \) and \( r \) is the (continuously-compounded) risk-free rate over the horizon. \( \hat{E} \) denotes expectations under the probability measure obtained from the market makers’ beliefs using their (normalized) marginal utilities corresponding to \( q = 0 \) as Radon-Nikodym derivative. Consequently,

\[
\hat{E}_t [p_{t+h}] = e^{r_h p_t},
\]

so that the maximization problem becomes

\[
\max_q \left\{ q(p_t - \hat{p}_t) - e^{-r_h} \frac{\gamma^M}{2} Var_t \left[ q p_{t+h} \right] \right\}.
\]

(34)

The price \( \hat{p} \) is set so as to satisfy the market-clearing condition

\[
0 = \Delta x_t + q \frac{\Delta t}{h},
\]

(35)

Since \( p \) is exogenous and Gaussian with variance \( V_h \) periods ahead that can be calculated easily,\(^7\) the maximization problem yields

\[
\hat{p}_t = p_t + e^{-r_h} \gamma^M V_h \frac{\Delta x_t}{\Delta t} h.
\]

(36)

Consequently, if the trader trades an amount \( \Delta x_t \), he trades at the unit price of \( p_t \) and pays an additional transaction cost of

\[
e^{-r_h} \gamma^M \Delta x_t^\top V_h \frac{\Delta x_t}{\Delta t} h,
\]

which has the quadratic form posited in the body of the paper.

Two cases suggest themselves naturally when considering the choice for the holding period

\(^7\)The resulting value is \( V_h = \Sigma h + B N_h \Omega N_h^\top B^\top \), where \( N_h = \int_0^h \int_u^h e^{-\Phi(t-u)} dt \, du = \Phi^{-1} h - \Phi^{-2} (I - e^{-\Phi h}) \) if \( \Phi \) is invertible. (Note that the first term, \( \Sigma h \), is of order \( h \), while the second of order \( h^2 \).)
$h$ as a function of $\Delta t$. In the first case, a decreasing $\Delta t$ is thought of as an improvement in the trading technology, attention, etc., of all market participants, and therefore $h$ decreases as $\Delta t$ does — in its simplest form, $h = \Delta t$, which yields a transaction cost of the order $\Delta t^2$. Generally, as long as $h \to 0$ as $\Delta t \to 0$, the transaction costs also vanishes.

The second case is that of a constant $h$: the dealers need a fixed amount of time to lay off a position regardless of the frequency with which our original traders access the market. It follows, in this case, that the price impact does not vanish as $\Delta t$ becomes small: in the continuous-time limit ($\Delta t \to 0$), the per-unit-of-time transaction cost is proportional to

$$\lim_{\Delta t \to 0} \frac{\Delta x_t^T}{\Delta t} V_h h^T \frac{\Delta x_t}{\Delta t} = \tau^T V_h h \tau,$$

as assumed in Section 1.1. One can therefore interpret $\Delta t$ in this case as the frequency with which the researcher observes the world, which does not impact (to the first order) equilibrium quantities — in particular, flow trades and costs. We summarize our results as follows:

**Proposition 6 (Time Dependence of Transitory Transaction Costs)**

(i) If dealers need a fixed amount of calendar time to lay off their inventory, then the transitory market-impact parameter $\Lambda(\Delta t)$ is of order $1/\Delta t$, $\Lambda(\Delta t) = \Lambda/\Delta t$.

(ii) If dealers can lay off their inventory during each time period, then the transitory market-impact parameter $\Lambda(\Delta t)$ is of order $\Delta t$, $\Lambda(\Delta t) = \Lambda \Delta t$.

**Persistent transaction costs.**

A similar model, but with a different specification of the market makers, can be used to justify a persistent price impact. Consider therefore the same model as in the previous section, but suppose now that market makers do not hold their inventories for a deterministic number $h$ of time units, but rather manage to deplete them, through trade with end users at price $p$, at a constant rate $\psi$. Thus, between two trading dates with the trader, a market maker’s
inventory evolves according to

\[ \Delta I_t = -\psi I_{t-1} \Delta t + q_t, \]  

(38)

where, in equilibrium,

\[ q_t = \Delta x_t. \]

The market makers continue to maximize a quadratic objective:

\[
\max_{\{q_s\}_{s \geq t}} \left\{ \hat{E}_t \sum_{s \geq t} e^{-r(s-t)} \left( \psi I_{s-1}^T p_s \Delta t - q_s^T \hat{p}_s - \frac{\gamma^M}{2} I_{s-1}^T V_{\Delta t} I_{s-1} \right) \right\},
\]

subject to (38) and expectations about \( q \) described below. Note that the market maker’s objective depends (positively) on the expected cash flows \( \psi I_{s-1}^T p_s \Delta t - q_s^T \hat{p}_s \) due to future trades with the end user and the trader and negatively on the risk of his inventory.

We assume that market makers cannot predict the trader’s order flow \( \Delta x \). More specifically, according to their probability distribution,

\[
\hat{E}_t [\Delta x_t | \mathcal{F}_s, s < t] = 0 \quad (40)
\]

\[
\hat{E}_t [(\Delta x_t)^2 | \mathcal{F}_s, s < t] = v. \quad (41)
\]

Moments of \( q_s \) and \( I_s \) follow immediately.

The first-order condition with respect to \( q_t \) is

\[
0 = \hat{E}_t \sum_{s \geq t} e^{-r(s-t)} \left( \psi p_s^T - \gamma^M I_{s-1}^T \frac{V_{\Delta t}}{\Delta t} \right) \frac{\partial I_{s-1}}{\partial q_t} \Delta t - \hat{p}_t^T. \quad (42)
\]

Using the fact that \( \frac{\partial I_t}{\partial q_t} = (1 - \psi \Delta t)^{s-t} \) for \( s \geq t \), the first-order condition yields

\[
\hat{p}_t = \hat{E}_t \sum_{s > t} e^{-r(s-t)} (1 - \psi \Delta t)^{s-t-1} \left( \psi p_s - \gamma^M \frac{V_{\Delta t}}{\Delta t} I_{s-1} \right) \Delta t. \quad (43)
\]
Using the facts that $\hat{E}_t[e^{-r(s-t)}p_s] = p_t$ and $\hat{E}_t[I_s] = (1 - \psi \Delta t)^{s-t}I_t$, we obtain

$$\hat{p}_t = p_t - \kappa_I I_t$$

for a constant matrix

$$\kappa_I = \sum_{n=0}^{\infty} e^{-r(n+1)\Delta t}(1 - \psi \Delta t)^{2n\Delta t} M V_{\Delta t} \Delta t. \quad (45)$$

The price $\hat{p}_t$ is only the price at the end of trading date $t$ — the price at which the last unit of the $q_t$ shares is traded. We assume that, during the trading date, orders of infinitesimal size come to market sequentially and the market makers’ expectation is that the remainder of date-$t$ order flow aggregates to zero — thus, the order flow is a martingale. It follows that the price paid for the $k$th percentile of the order flow $q_t$ is $p_t - \kappa_I (I_{t-1} + kq_t)$. This mechanism ensures that round-trip trades over very short intervals do not have transaction-cost implications.

This price specification is the same as in Section 2.1, with $\Lambda = 0$ and $D_t = -\kappa_I I_t$:

$$D_{t+1} = -\kappa_I I_{t+1}$$
$$= -\kappa_I (I_t - \psi I_t + \Delta x_t)$$
$$= D_t - \kappa_I \psi \kappa_I^{-1} D_t - \kappa_I \Delta x_t$$
$$= (I - R) (D_t + C \Delta x_t). \quad (46)$$

Proposition 7 (Time Dependence of Persistent Transaction Costs) The resiliency parameter $R$ is of order $\Delta t$, $R(\Delta t) = R\Delta t$. The persistent market impact $C$ does not depend on $\Delta t$.

Transitory and persistent transaction costs.

The two types of price impact can obtain simultaneously in this model so that we can have both kinds of transaction costs and consider their separate convergence to continuous time using Propositions 6–7. To see this, consider for instance an economy with the trader and two kinds of market makers. The trader transacts with the first group of market makers. After a
period of length $h$, these market makers clear their inventories with a second group of market makers, who specialize in locating end users and trading with them. This second group of market makers deplete their inventories only gradually (at a constant rate as above), giving rise to a persistent impact. The trader must compensate both groups of market makers for the risk taken, resulting in the two price-impact components.

2.3 Convergence as Length of Time Periods Vanishes

We now show that the continuous-time model and its solution are the limit of their discrete-time analogues. Our micro foundation for transaction costs highlights that there are two important cases that lead to different continuous-time limits as seen in Proposition 6.

**Proposition 8** (i) Suppose that dealers need a fixed amount of calendar time to lay off their inventory so that $\Lambda(\Delta t) = \Lambda / \Delta t$. Then the solution to the general discrete-time model with transitory and persistent transaction costs converges to the continuous-time solution with transitory and persistent costs in Proposition 2. The continuous-time matrix coefficients $M^{\text{rate}}$ and $M^{\text{speed}}$ are the limits of the discrete-time coefficients $M^{\text{rate}}$ and $M^{\text{speed}}$ as follows:

$$\lim_{\Delta t \to 0} \frac{M^{\text{rate}}(\Delta t)}{\Delta t} = \bar{M}^{\text{rate}}$$

$$\lim_{\Delta t \to 0} M^{\text{aim}}(\Delta t) = \bar{M}^{\text{aim}}.$$  \hspace{1cm} (47)

Given that $f$ is continuous almost everywhere on every path almost surely, (47) and (48) imply that, fixing $(x_t, f_t, D_t)$, \( \lim_{\Delta t \to 0} \frac{x_t + \Delta t - x_t}{\Delta t} \to \tau_t \) a.s.

(ii) Suppose that dealers can lay off their inventory each time period so that $\Lambda(\Delta t) = \Lambda \Delta t$. Then the solution to the general discrete-time model with transitory and persistent transaction costs converges to the solution to the continuous-time model with purely persistent costs in Proposition 3.

Equation (48) states that, given the current conditions, the aim portfolio in discrete time approaches the one in continuous time as the time period between trades becomes small. Equation (47) concerns the tracking speed: the per-unit-of-time fraction of the distance to
the aim covered by a trade is the same, for small $\Delta t$, as in continuous time. Consequently, the optimal solutions have the same dynamics, in the limit.

3 Extensions

3.1 General Predictor Dynamics

In this subsection we consider general dynamics for the predictor $f$. We show that the tractability of the model is not due to the linear dynamics of the base-case model, but rather solely to the linear-quadratic utility flow. In particular, we show that the nature of trading is the same as in the base case: When the cost has a transitory component, the agent trades at a constant rate — calculated exactly as before — towards an aim portfolio. In turn, the aim portfolio is a weighted average of all expected future fundamental “abnormal returns” $Bf$, as well as the distortion $D$ (if there is one). In the pure-persistent case, the optimal portfolio is a weighted average of expected future values $Bf$ and $D$.

To relate our results to the classical theory of investments, we employ the notation $\text{Markowitz}$ to refer to the instantaneous mean-various portfolio absent transaction costs. In agreement with the classical findings of Markowitz (1952),

$$\text{Markowitz}_t = (\gamma \Sigma)^{-1} Bf_t. \quad (49)$$

Note that this definition of the Markowitz portfolio concentrates exclusively on the expected returns due to fundamental fluctuations. The investor we consider experiences returns that reflect its price impact, too.

The result of this section is stated formally in the following proposition.

**Proposition 9** Suppose that $f$ is a general Markovian diffusion process.

(i) If the impact is purely transitory and Assumption A is satisfied, then the optimal portfolio
can be written as

\[ \text{aim}_0 = \int_0^\infty b e^{-bt} E[\text{Markowitz}_t | f_0] dt \]  

(50)

for a scalar \( b > 0 \). If Assumption A is not satisfied, then the result holds with \( b \) a matrix.

(ii) If the price impact has both a transitory and a permanent component, then the aim is of the form \( M_f^{\text{aim}}(f_0) + M_D^{\text{aim}}D_0 \) and the \( f \)-component of the aim equals

\[ M_f^{\text{aim}}(f_0) = \int_0^\infty b_1 e^{-b_2t} [0 1] \top (\gamma \Sigma) E[\text{Markowitz}_t | f_0] dt, \]

where \( b_1 \) and \( b_2 \) are appropriate matrices.

(iii) If the price impact is purely persistent, then the optimal position \( x_0 \) takes the form

\[ x_0 = M_f^x(f_0) + M_D^x D_0, \]

with

\[ M_f^x(f_0) = b_3 \text{Markowitz}_0 + b_1 \int_0^\infty e^{-b_2t} b_4 E[\text{Markowitz}_t | f_0] dt, \]

(52)

where \( b_i, i = 1, \ldots, 4 \), are appropriate matrices.

We make a couple of remarks concerning this result. First, the only difference from the base case is that the dependence of the optimal trade on the current value of the predictor, \( f_0 \), is no longer linear. Second, the proposition makes it clear that the dependence on \( f_0 \) stems from the (expected) future benefits of ownership, \( Bf_t \) (expressed here in terms of the Markowitz portfolio purely for expositional reasons). If the conditional expectation \( E[f_t | f_0] \) is not linear in \( f_0 \), then the trading strategy is not.

### 3.2 Time-Varying Parameters

Much of the tractability of the framework is preserved if one lets the risk aversion, transaction costs, or return variance vary over time. Specifically, the results derived above continue to hold, except that the value-function coefficients are functions of the time-varying parameter.\(^8\)

\(^8\)We continue to assume a Markovian structure.
We illustrate this statement in the simplest setting: the price impact is purely transitory, Assumption A holds, and $\text{var}_t(du_t) = v_t \bar{\Sigma}$ with

$$dv_t = \mu(v_t)dt + \sigma(v_t)dw_t. \quad (53)$$

Here, $w_t$ is a (one-dimensional) Wiener process, possibly correlated with $\varepsilon$ and $u$.

We conjecture the value function to be quadratic in $(x,f)$, but with coefficients that depend on $v$:

$$V(x,f,v) = -\frac{1}{2}x^\top A_{xx}(v)x + x^\top A_{xf}(v)f + \frac{1}{2}f^\top A_{ff}(v)f + A_0(v). \quad (54)$$

The HJB equation provides (second-order) differential equations for the coefficient functions — in particular, $A_{xx}$ and $A_{xf}$, which determine the trading strategy. In the special case $\mu(v) = \bar{\mu}v$, these ODEs can be solved explicitly: $A_{xx}$ is linear in $v$ and $A_{xf}$ is constant. A more empirically relevant case, however, is that of a mean-reverting volatility level $v$. The following proposition records some properties of the ensuing optimal trading strategy.

**Proposition 10 (Stochastic volatility)** Suppose that the price impact is purely transitory, Assumption A holds, and $\text{var}_t(du_t) = v_t \bar{\Sigma}$ and $\mu(v)$ in (53) is decreasing and crosses zero on the support of $v$. Then:

(i) There exists a cut-off value $\hat{v}$ such that, for $v_t \leq \hat{v}$, the trading intensity $\tau$ is higher than it would be if $v$ was constant and equal to $v_t$ for $s \geq t$; conversely for $v_t \geq \hat{v}$.

(ii) The loading $\bar{M}_\text{aim}$ of the aim on the predictor $f$ is lower than for constant $v$ if and only if $v_t \leq \hat{v}$.

The proposition shows how the volatility mean reversion impacts the trading strategy. In particular, if $v_t$ is low enough, so that it is expected to increase, then the trading intensity is higher than if $v$ were to stay constant: the higher utility cost due to increased future volatility, which is persistent, is mitigated by trading more currently, before the trading cost increases along with the volatility.

---

9One can ensure that $v$ is bounded by letting $\sigma(v)$ be zero outside some interval.
The impact on the aim portfolio can also be understood intuitively, and along the same lines. Specifically, when \( v \) is expected to rise, the utility cost due to risk is higher than if \( v \) is constant, and the investment is therefore more conservative.

We also note that, under Assumption A, letting \( \text{var}_t(du_t) \) be proportional to \( v_t \) and letting \( \Lambda_t \) and \( \gamma_t \) be proportional to \( v_t \) are isomorphic. The results also hold if \( \Lambda_t = \lambda \Sigma \) is constant (this is the same as letting only \( \gamma \) depend on \( v \)), as is intuitive.

4 Equilibrium Implications

In this section we study the restrictions placed on a security’s return properties by the market equilibrium. More specifically, we consider a situation in which an investor facing transaction costs absorbs a residual supply specified exogenously and analyze the relationship implied between the characteristics of the supply dynamics and the excess return.

For simplicity, we consider a model set in continuous time, as detailed in Section 1.1, featuring one security in which \( L \geq 1 \) groups of (exogenously given) noise traders hold positions \( z^l_t \) (net of the aggregate supply) given by

\[
\begin{align*}
 dz^l_t &= \kappa \left( f^l_t - z^l_t \right) dt \\
 df^l_t &= -\psi f^l_t dt + d\varepsilon^l_t.
\end{align*}
\]

It follows that the aggregate noise-trader holding, \( z_t = \sum z^l_t \), satisfies

\[
 dz_t = \kappa \left( \sum_{l=1}^{L} f^l_t - z_t \right) dt.
\]

We conjecture that the investor’s inference problem is as studied in Section 1.1, where \( f \) given by \( f \equiv (f^1, \ldots, f^L, z) \) is a linear return predictor and \( B \) is to be determined. We verify the conjecture and find \( B \) as part of Proposition 11 below.
Given the definition of $f$, the mean-reversion matrix $\Phi$ is given by

$$
\Phi = \begin{pmatrix}
\psi_1 & 0 & \cdots & 0 \\
0 & \psi_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
-\kappa & -\kappa & \cdots & \kappa
\end{pmatrix}.
$$  \tag{58}

Suppose that the only other investors in the economy are the investors considered in Section 1.1, facing transaction costs given by $\Lambda = \lambda \sigma^2$. In this simple context, an equilibrium is defined as a price process and market-clearing asset holdings that are optimal for all agents given the price process. Since the noise traders’ positions are optimal by assumption as specified by (55)–(56), the restriction imposed by equilibrium is that the dynamics of the price are such that, for all $t$,

$$
x_t = -z_t  \tag{59}
$$

$$
dx_t = -dz_t.  \tag{60}
$$

Using the results in Proposition 1, these equilibrium conditions lead to

$$
\frac{a}{\lambda} \sigma^{-2} B (a \Phi + \gamma I)^{-1} + \frac{a}{\lambda} e_{L+1} = -\kappa (1 - 2e_{L+1}),  \tag{61}
$$

where $e_{L+1} = (0, \cdots, 0, 1) \in \mathbb{R}^{L+1}$ and $1 = (1, \cdots, 1) \in \mathbb{R}^{L+1}$. It consequently follows that, if the investor is to hold $-z_t = -f_t^{L+1}$ at time $t$ for all $t$, then the factor loadings must be given by

$$
B = \sigma^2 \left[ -\frac{\lambda}{a} \kappa (1 - 2e_{L+1}) - e_{L+1} \right] (a \Phi + \gamma I).  \tag{62}
$$

For $l \leq L$, we calculate $B_l$ further as

$$
B_l = -\sigma^2 \kappa (\lambda \psi_l + \lambda \gamma a^{-1} + \lambda \kappa - a)
= -\lambda \sigma^2 \kappa (\psi_l + \rho + \kappa),  \tag{63}
$$
while

\[ B_{L+1} = \sigma^2(\rho\lambda\kappa + \lambda\kappa^2 - \gamma). \]  

(64)

Using this, it is straightforward to see the following key equilibrium implications:

**Proposition 11** The market is in equilibrium if and only if \( x_0 = -z_0 \) and the security’s expected excess return is given by

\[
\frac{1}{dt} E_t[dp_t - r_f p_t dt] = \sum_{l=1}^{L} \lambda \sigma^2 \kappa (\psi_l + \rho + \kappa) (-f^l_t) + \sigma^2 (\rho \lambda \kappa + \kappa^2 - \gamma) z_t. \]  

(65)

The coefficients \( \lambda \sigma^2 \kappa (\psi_k + \rho + \kappa) \) are positive and increase in the mean-reversion parameters \( \psi_k \) and \( \kappa \) and in the trading costs \( \lambda \sigma^2 \). In other words, noise trader selling \( (f^k_t < 0) \) increases expected excess returns, and especially so if its mean reversion is faster and if the trading cost is larger.

Naturally, noise-trader selling increases the expected excess return, while noise-trader buying lowers it, since the arbitrageurs need to be compensated to take the other side of the trade. Interestingly, the effect is larger when trading costs are larger and for noise-trader shocks with faster mean reversion because such shocks are associated with larger trading costs for the arbitrageurs.

### 5 Conclusion

The model’s tractability makes it potentially useful for future research on return predictability and transaction costs. As one such application of the model, we illustrate how transaction costs can lead to large alphas for short time periods in equilibrium. Indeed, we model an equilibrium with several “noise traders” who trade in and out of their positions with varying mean-reversion speeds and a rational arbitrageur — with trading costs and using the methodology that we derive — who takes the other side of these noise-trader positions to clear the market. We solve the equilibrium explicitly and show how noise trading leads to
return predictability and return reversals. Further, we show that noise-trader demand that
mean-reverts more quickly leads to larger return predictability because a fast mean reversion
is associated with high transaction costs for the arbitrageurs and, consequently, they must
be compensated in the form of larger return predictability.
References


A Proofs

Proof of Proposition 1. This proposition is a special case of Proposition 2, but it is sufficiently simpler to be worth sketching the calculation. The Hamilton-Jacoby-Bellman
(HJB) equation is
\[
\rho V = \sup_\tau \left\{ x^\top B f - \frac{\gamma}{2} x^\top \Sigma x - \frac{1}{2} \tau^\top \Lambda \tau + \frac{\partial V}{\partial x} \tau + \frac{\partial V}{\partial f} (-\Phi f) + \frac{1}{2} \text{tr} \left( \Omega \frac{\partial^2 V}{\partial f \partial f^\top} \right) \right\}. \tag{A.1}
\]

Maximizing this expression with respect to the trading intensity results in
\[
\tau = \Lambda^{-1} \frac{\partial V}{\partial x}. \tag{*}
\]

Under the natural quadratic-value function conjecture
\[
V(x, f) = -\frac{1}{2} x^\top A_{xx} x + x^\top A_{xf} f + \frac{1}{2} f^\top A_{ff} f + A_0, \tag{A.2}
\]

the optimal choice \( \tau \) equals
\[
\tau_t = -\Lambda^{-1} A_{xx} x_t + \Lambda^{-1} A_{xf} f_t. \tag{**}
\]

Once this is inserted in the HJB equation, it results in the following equations defining the value-function coefficients (using the symmetry of \( A_{xx} \)):
\[
-\rho A_{xx} = A_{xx} \Lambda^{-1} A_{xx} - \gamma \Sigma \tag{A.3}
\]
\[
\rho A_{xf} = -A_{xx} \Lambda^{-1} A_{xf} - A_{xf} \Phi + B \tag{A.4}
\]
\[
\rho A_{ff} = A_{xf}^\top \Lambda^{-1} A_{xf} - 2 A_{ff} \Phi. \tag{A.5}
\]

Pre- and post-multiplying (A.21) by \( \Lambda^{-\frac{1}{2}} \), we obtain
\[
-\rho Z = Z^2 + \frac{\rho^2}{4} I - U, \tag{A.6}
\]

that is,
\[
\left( Z + \frac{\rho}{2} f \right)^2 = U, \tag{A.7}
\]

30
where

\[
Z = \Lambda^{-\frac{1}{2}}A_{xx}\Lambda^{-\frac{1}{2}} \\
U = \gamma\Lambda^{-\frac{1}{2}}\Sigma\Lambda^{-\frac{1}{2}} + \frac{\rho^2}{4}I. 
\]  

(A.8)  

(A.9)

This leads to

\[
Z = -\frac{\rho}{2}I + U^{-\frac{1}{2}} \geq 0, 
\]  

(A.10)

implying that

\[
A_{xx} = -\frac{\rho}{2}\Lambda + \Lambda^{\frac{1}{2}}\left(\gamma\Lambda^{-\frac{1}{2}}\Sigma\Lambda^{-\frac{1}{2}} + \frac{\rho^2}{4}\right)^{\frac{1}{2}}\Lambda^{\frac{1}{2}}. 
\]  

(A.11)

The solution for \(A_{xf}\) follows from Equation (A.22), using the general rule that vec(\(XYZ\)) = (\(Z^\top \otimes X\)) vec(Y):

\[
\text{vec}(A_{xf}) = \left(\rho I + \Phi^\top \otimes I_K + I_S \otimes (A_{xx}\Lambda^{-1})\right)^{-1}\text{vec}(B). 
\]

If \(\Lambda = \lambda\Sigma\), then \(A_{xx} = a\Sigma\) with

\[
-\rho a = a^2\frac{1}{\lambda} - \gamma 
\]  

(A.12)

with solution

\[
a = -\frac{\rho}{2}\lambda + \sqrt{\gamma\lambda + \frac{\rho^2}{4}\lambda^2}. 
\]  

(A.13)

In this case, (A.22) yields

\[
A_{xf} = B\left(\rho I + \frac{a}{\lambda}I + \Phi\right)^{-1} = B\left(\frac{\gamma}{a}I + \Phi\right)^{-1}, 
\]

where the last equality uses (A.30).
Then, we have

\[ \tau_t = \frac{a}{\lambda} \left[ \Sigma^{-1} B (a \Phi + \gamma I)^{-1} f_t - x_t \right]. \] (A.14)

It is clear from (A.31) that \( \frac{a}{\lambda} \) decreases in \( \lambda \) and increases in \( \gamma \).

**Proof of Proposition 2.** Using the notation

\[ \Pi = \begin{bmatrix} \Phi & 0 \\ 0 & R \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} 0 \\ C \end{bmatrix}, \]
\[ \tilde{B} = \begin{bmatrix} B & -(R + r^f) \end{bmatrix}, \] (A.15)
\[ \tilde{\Omega} = \begin{bmatrix} \Omega & 0 \\ 0 & 0 \end{bmatrix}, \quad \tilde{\epsilon}_t = \begin{bmatrix} \epsilon_t \\ 0 \end{bmatrix}, \]

the HJB equation is

\[
\rho V = \max_{\tau} \left\{ x^\top (\tilde{B} y + C \tau) - \frac{\gamma}{2} x^\top \Sigma x - \frac{1}{2} \tau^\top \Lambda \tau + \frac{\partial V}{\partial x} \tau + \frac{\partial V}{\partial y} \left( -\Pi y + \tilde{C} \tau \right) + e \right\}
\]
\[
= \max_{\tau} \left\{ x^\top \tilde{B} y - \frac{\gamma}{2} x^\top \Sigma x - \frac{1}{2} \tau^\top \Lambda \tau + \tau^\top (Q_x x + Q_y y) - \frac{\partial V}{\partial y} \Pi y + e \right\},
\]

where

\[ e = \frac{1}{2} tr \left( \tilde{\Omega} \frac{\partial^2 V}{\partial y \partial y^\top} \right) \] (A.16)
\[ Q_x = -A_{xx} + \tilde{C}^\top A_{xy} + C^\top \]
\[ Q_y = A_{xy} + \tilde{C}^\top A_{yy}. \] (A.17)

It follows immediately that

\[ \tau_t = -\Lambda^{-1} Q_x [aim_t - x_t] \] (A.18)
\[ = \Lambda^{-1} \left( A_{xx} - \tilde{C}^\top A_{xy} - C^\top \right) [aim_t - x_t] \]
\[ \equiv \tilde{M}^{rate} [aim_t - x_t], \]
with

$$aim_t = -\left(Q_x^{-1}Q_y\right) y_t$$  \hspace{1cm} (A.19)  

$$= \left(A_{xx} - \tilde{C}^T A_{xy} - C^T\right)^{-1} \left(A_{xy} + \tilde{C}^T A_{yy}\right) y_t$$  

$$\equiv \tilde{M}aim_y t.$$  

The coefficient matrices solve the system

$$-\rho A_{xx} = -\gamma \Sigma + Q_x^T \Lambda^{-1} Q_x$$  

$$\rho A_{xy} = Q_x^T \Lambda^{-1} Q_y + \tilde{B} - A_{xy} \Pi$$  \hspace{1cm} (A.20)  

$$\rho A_{yy} = Q_y^T \Lambda^{-1} Q_y - A_{yy} \Pi - \Pi^T A_{yy}$$

$$= \left(A_{yy} + \tilde{C}^T A_{yy}\right) - A_{yy} \Pi - \Pi^T A_{yy},$$

We note that the equations above have to be solved simultaneously for $A_{xx}$, $A_{xy}$, and $A_{yy}$; there is no closed-form solution in general. The complication is due to the fact that current trading affects the persistent price component $D$ (that is, $C \neq 0$).

The special case $C = 0$ (and $D_0 = 0$) — i.e., purely transitory costs — does lead to closed-form solutions — which are even slightly simpler than in discrete time — as follows.

Specializing (A.18) and (A.19) to this case, we obtain

$$\tau_t = -\Lambda^{-1} A_{xx} x_t + \Lambda^{-1} A_{xf} f_t,$$

while (A.20) becomes

$$-\rho A_{xx} = A_{xx} \Lambda^{-1} A_{xx} - \gamma \Sigma$$  \hspace{1cm} (A.21)  

$$\rho A_{xf} = -A_{xx} \Lambda^{-1} A_{xf} - A_{xf} \Phi + B$$  \hspace{1cm} (A.22)  

$$\rho A_{ff} = A_{xf}^T \Lambda^{-1} A_{xf} - 2A_{ff} \Phi.$$  \hspace{1cm} (A.23)
Pre- and post-multiplying (A.21) by $\Lambda^{-\frac{1}{2}}$, we obtain

$$-\rho Z = Z^2 + \frac{\rho^2}{4} I - Y,$$

that is,

$$\left(Z + \frac{\rho}{2} I\right)^2 = Y,$$  \hspace{1cm} (A.24)

where

$$Z = \Lambda^{-\frac{1}{2}} A_{xx} \Lambda^{-\frac{1}{2}}$$ \hspace{1cm} (A.26)

$$Y = \gamma \Lambda^{-\frac{1}{2}} \Sigma \Lambda^{-\frac{1}{2}} + \frac{\rho^2}{4} I.$$ \hspace{1cm} (A.27)

This leads to

$$Z = -\frac{\rho}{2} I + Y^{\frac{1}{2}} \geq 0,$$  \hspace{1cm} (A.28)

implying that

$$A_{xx} = -\frac{\rho}{2} \Lambda + \Lambda^{\frac{1}{2}} \left(\gamma \Lambda^{-\frac{1}{2}} \Sigma \Lambda^{-\frac{1}{2}} + \frac{\rho^2}{4}\right)^{\frac{1}{2}} \Lambda^{\frac{1}{2}}.$$  \hspace{1cm} (A.29)

The solution for $A_{xf}$ follows from Equation (A.22), using the general rule that $\text{vec}(XYZ) = (Z^\top \otimes X) \text{vec}(Y)$:

$$\text{vec}(A_{xf}) = \left(\rho I + \Phi^\top \otimes I_S + I_K \otimes (A_{xx} \Lambda^{-1})\right)^{-1} \text{vec}(B).$$

Just as in discrete time, the form of the solution further simplifies under Assumption A. Given $\Lambda = \lambda \Sigma$, $A_{xx} = a \Sigma$ with

$$-\rho a = a^2 \frac{1}{\lambda} - \gamma,$$  \hspace{1cm} (A.30)
with positive solution

\[
a = -\frac{\rho}{2} \lambda + \sqrt{\gamma \lambda + \frac{\rho^2}{4} \lambda^2}.
\]  \hspace{1cm} (A.31)

In this case, (A.22) yields

\[
A_{xf} = B \left( \frac{\rho I + a I}{\lambda} + \Phi \right)^{-1}
\]

\[
= B \left( \frac{\gamma a I + \Phi}{\lambda} \right)^{-1},
\]

where the last equality uses (A.30).

Then, we have

\[
\tau_t = \frac{a}{\lambda} \left[ \Sigma^{-1} B (a \Phi + \gamma I)^{-1} f_t - x_t \right].
\]  \hspace{1cm} (A.32)

It is clear from (A.31) that \( \frac{a}{\lambda} \) decreases in \( \lambda \) and increases in \( \gamma \). \( \blacksquare \)

**Proof of Proposition 3.** Let’s start with the complete problem:

\[
V(x_t, D_t, f_t) = E_t \int_t^\infty e^{-\rho(s-t)} \left( x_s^\top (Bf_s - (r + R)D_s) - \frac{\gamma}{2} x_s^\top \Sigma x_s \right) ds
\]

\[
+ E_t \int_t^\infty e^{-\rho(s-t)} x_s^\top C d x_s + \frac{1}{2} E_t \int_t^\infty e^{-\rho(s-t)} d [x_s, C x_s].
\]  \hspace{1cm} (A.33)

We let \( x \) have the representation

\[
dx_t = \mu_t dt + \sigma_t d Z_t^x + \Delta x_t.
\]  \hspace{1cm} (A.34)

As is customary with such problems, we write and solve the HJB equation, then use the fact that it is satisfied to provide a so-called verification argument for the proposed optimal control and value function. We also use the conjectured form (19) and introduce the notation
\hat{V}(D, f) = V(0, D, f), \text{ so that }
\begin{equation}
V(x, D, f) = \hat{V}(D - Cx, f) - \frac{1}{2} x^\top Cx.
\end{equation}

Note that
\begin{equation}
d(D_s - Cx_s) = -RD_s ds,
\end{equation}
so that \(D^0 \equiv D - C\hat{x}\) is a continuous and finite-variation process.

The HJB equation is
\begin{equation}
0 = \sup_{\Delta x, \mu, \sigma} \left\{ x^\top (Bf - (r^f + R)D) - \frac{\gamma}{2} x^\top \Sigma x + x^\top C \frac{1}{dt} E_t [dx_t] + \frac{1}{2} \frac{1}{dt} E_t d[x_t, Cx_t] - \rho \hat{V} + \frac{1}{2} \rho x^\top Cx + \hat{V}_D(-RD) + \hat{V}_f(-\Phi f) \\
-x^\top C \frac{1}{dt} E_t [dx_t] - \frac{1}{2} \frac{1}{dt} E_t d[x_t, Cx_t] + \frac{1}{dt} E_t d[f, V_ff] \\
+ \frac{1}{\Delta t} E_t \left[ x^\top C \Delta x + \frac{1}{2} \Delta x^\top C \Delta x + \hat{V}(D^0, f_- + \Delta f) - \hat{V}(D^0, f_-) \right] \right\}.
\end{equation}
Here, we suppressed the notational dependence on time and also wrote \(x_-\) for \(x_{t-}\) and similarly \(D_-\) and \(f_-\).

We conjecture a quadratic form for the value function \(\hat{V}\):
\begin{equation}
\hat{V}(D, f) = \frac{1}{2} D^\top A_{DD} D + D^\top A_{Df} f + f^\top A_{ff} f + A_0,
\end{equation}
which, given \(E_t [\Delta f_t] = 0\), leads to the simplification
\begin{equation}
0 = \sup_{x} \left\{ -\rho \hat{V} + \frac{\rho}{2} x^\top Cx + x^\top Bf - x^\top (r^f + R)(D^0 + Cx) - \frac{\gamma}{2} x^\top \Sigma x \\
- \hat{V}_D R(D^0 + Cx) - \hat{V}_f \Phi f + tr(A_{ff} \Omega) \right\}.
\end{equation}
We remark on the fact that \(A.39\) has the standard continuous-time form. The first two terms in \(A.39\) equal the value function decay rate \(-\rho V(\hat{x}, \hat{D}, f)\), while the remaining terms
represent the flow benefit from taking position \( x \) for the next infinitesimal time period: the expected excess return, the distortion decay summed with the opportunity cost of funds (the risk-free rate), from which the position \( x \) will suffer over \( dt \), the risk cost, and the change over time in \( V \) induced by the decay of \( D \) and of \( f \), as well as the convexity and jump adjustments for \( f \). Note that, in order for the problem to be well defined, it is necessary that \( \rho < 2(r^f + R) + \gamma \Sigma \) — otherwise, the agent gains too much from pushing the prices up currently relative to the perceived cost of the risk and the decay in the distortion.

In order to write down the solution, let

\[
J = \frac{1}{2} (J_0 + J_0^\top) \quad \text{(A.40)}
\]

\[
J_0 = \gamma \Sigma + (2R + 2r^f - \rho)C \quad \text{(A.41)}
\]

\[
j = (B - C^\top R^\top A_{Df}) f - (C^\top R^\top A_{DD} + r^f + R) D^0. \quad \text{(A.42)}
\]

It follows that

\[
x = J^{-1} \left( B f - (r^f + R) D^0 - C^\top R^\top \hat{V}_D \right) \quad \text{(A.43)}
\]

\[
= J^{-1} j \quad \text{(A.44)}
\]

and the HJB equation becomes

\[
0 = \frac{1}{2} j^\top J^{-1} j - \hat{V}_D R D^0 - \hat{V}_f \Phi f - \rho \hat{V} + \frac{1}{2} \text{tr} \left( \hat{V}_f \Omega \right). \quad \text{(A.45)}
\]

The constant matrices \( A_{DD} \) and \( A_{Df} \) are computed in the usual way:

\[
\rho A_{DD} = (A_{DD} R C + r^f + R^\top) J^{-1} \left( C^\top R^\top A_{DD} + r^f + R \right) - A_{DD} R - R^\top A_{DD} \quad \text{(A.46)}
\]

\[
\rho A_{Df} = (A_{DD} R C + r^f + R^\top) J^{-1} \left( -B + C^\top R^\top A_{Df} \right) - R^\top A_{Df} - A_{Df} \Phi. \quad \text{(A.47)}
\]

To prove that the proposed solution does, indeed, solve the trader’s problem, we follow a verification argument. Let \( \hat{x} \) be an arbitrary trading strategy (satisfying technical conditions
XXX) and \( V \) quadratic, defined by (19) and the matrices \( A \). Since it holds generally that
\[
e^{-\rho t}V(\hat{x}_t, \hat{D}_t, f_t) = e^{-\rho t} \hat{V}(\hat{D}_t - C\hat{x}_t, f_t) - \frac{1}{2} e^{-\rho t} \hat{x}_t^\top C\hat{x}_t
\]
\[
= e^{-\rho T} \hat{V}(\hat{D}_T - C\hat{x}_T, f_T) - \frac{1}{2} e^{-\rho T} \hat{x}_T^\top C\hat{x}_T
\]
\[
- \int_t^T d \left( e^{-\rho s} \hat{V}(\hat{D}_s - C\hat{x}_s, f_s) - \frac{1}{2} e^{-\rho s} \hat{x}_s^\top C\hat{x}_s \right),
\]
it is sufficient to show that \( \lim_{T \to \infty} e^{-\rho T} E_t \left[ \hat{V}(\hat{D}_T - C\hat{x}_T, f_T) - \frac{1}{2} \hat{x}_T^\top C\hat{x}_T \right] = 0 \) and
\[
E_t \left[ - \int_t^T d \left( e^{-\rho s} \hat{V}(\hat{D}_s - C\hat{x}_s, f_s) - \frac{1}{2} e^{-\rho s} \hat{x}_s^\top C\hat{x}_s \right) \right]
\geq E_t \int_t^T e^{-\rho s} \left( \hat{x}_s^\top \left( Bf_s - (r + R)\hat{D}_s \right) - \frac{\gamma}{2} \hat{x}_s^\top \Sigma \hat{x}_s \right) ds
\]
\[
+ E_t \int_t^T e^{-\rho s} \hat{x}_s^\top C\hat{x}_s + \frac{1}{2} E_t \int_t^T e^{-\rho s} d[\hat{x}_s, C\hat{x}_s],
\]
and that the inequality holds with equality at the conjectured optimum control.

Ito’s lemma implies
\[
d \left( \hat{V}(\hat{D}_s - C\hat{x}_s, f_s) - \frac{1}{2} \hat{x}_s^\top C\hat{x}_s \right)
\]
\[
= \hat{V}_D(d\hat{D}_s - C\hat{x}_s) + \hat{V}_f(df_s - \Delta f_s) - \hat{x}_s^\top C\hat{x}_s - \frac{1}{2} d[\hat{x}_s, C\hat{x}_s] + \frac{1}{2} d[f_s, \hat{V}_f f_s] - \frac{1}{2} \Delta f_s^\top \hat{V}_f \Delta f_s + \hat{V}(\hat{D}_s - C\hat{x}_s, f_s) - \hat{V}(\hat{D}_s - C\hat{x}_s, f_s_+).\]

Taking conditional expectations of (A.49), one gets
\[
-\hat{V}_D R\hat{D}_s - \hat{x}_s^\top C E_s \frac{1}{ds} [d\hat{x}_s] - \frac{1}{2} \frac{1}{ds} d[\hat{x}_s, C\hat{x}_s] + \hat{V}_f(-\Phi) f_s + \text{tr}(A_{f f}\Omega),
\]
the negative of which we wish to be larger than
\[
\rho \hat{x}_s^\top C\hat{x}_s - \rho \hat{V} + \hat{x}_s^\top \left( Bf_s - (r + R)\hat{D}_s \right) - \frac{\gamma}{2} \hat{x}_s^\top \Sigma \hat{x}_s + \hat{x}_s^\top C \frac{1}{ds} E_s [d\hat{x}_s] + \frac{1}{2} \frac{1}{ds} E_s [d\hat{x}_s, C\hat{x}_s].
\]
This outcome is ensured by the HJB equation (A.39), which is satisfied by \( \hat{V} \) and \( x \), for any value of \( D^0 \). Furthermore, \( x \) satisfies the constraint with equality. ■

**Proof of Proposition 4.** To be added. ■

**Proof of Proposition 5.** To be added. ■

**Proof of Proposition 6.** To be added. ■

**Proof of Proposition 7.** To be added. ■

**Proof of Proposition 8.** To be added. ■

**Proof of Proposition 9.** (i) The conjectured value function, HJB equation, and optimal trading intensity are given, respectively, by

\[
V(x, f) = -\frac{1}{2} x^\top A_{xx} x + x^\top A_{xf}(f) + A_{ff}(f) \tag{A.50}
\]

\[
\rho V = -\frac{\gamma}{2} x^\top \Sigma x + x^\top B f - \frac{1}{2} \tau^\top \Lambda \tau + \tau^\top V_x^\top + x^\top D A_{xf}(f) + D A_{ff}(f) \tag{A.51}
\]

\[
\tau = \Lambda^{-1} A_{xx} \left( A_{xx}^{-1} A_{xf}(f) - x \right) \tag{A.52}
\]

Let \( \alpha_s \equiv E_0[\text{aim}_s] \), \( \dot{\alpha}_s = \frac{d}{ds} \alpha_s \), and \( M_s \equiv E_t[\text{Markowitz}_s] \). The expected aim dynamics, using (A.51), are

\[
\dot{\alpha}_s = A_{xx}^{-1} E_t [DA_{xf}(f_s)]
\]

\[
= A_{xx}^{-1} E_t [\rho A_{xf} + A_{xx} \Lambda^{-1} A_x - B f_s)]
\]

\[
= \rho \alpha_s + \Lambda^{-1} A_{xx} \alpha_s - \gamma A_{xx}^{-1} \Sigma M_s
\]

\[
= \gamma A_{xx}^{-1} \Sigma \alpha_s - \gamma A_{xx}^{-1} \Sigma M_s. \tag{A.53}
\]
The last equality made use of (A.21), which continues to hold, and implies

\[ A_{xx} (\rho + \Lambda^{-1} A_{xx}) = \gamma \Sigma. \]

We therefore obtain

\[ \text{aim}_0 = \alpha_0 = \int_0^\infty \gamma A_{xx}^{-1} \Sigma e^{-\gamma A_{xx}^{-1} \Sigma} M_s ds. \]

Under Assumption A, \( A_{xx}^{-1} \Sigma \) is a scalar.

(ii) In this case, the conjectured value function is

\[ V = -\frac{1}{2} x^T A_{xx} x + x^T A_{xf}(f) + A_{ff}(f) + \frac{1}{2} D^T A_{DD} D + D^T A_{Df}(f) + x^T A_{xD} D \quad (A.54) \]

\[ \rho V = -\frac{\gamma}{2} x^T \Sigma x + x^T (B f + C \tau) - \frac{1}{2} \tau^T \Lambda \tau + \tau^T V_x^T + (C \tau - RD)^T V_D^T + \]

\[ x^T D A_{xf}(f) + DA_{ff}(f) \]

\[ \tau = \Lambda^{-1} \left( A_{xx} + C^T + C^T A_{Dx} \right) \times \]

\[ \left( (A_{xx} + C^T + C^T A_{Dx})^{-1} \left( A_{xf}(f) + C^T A_{Df}(f) + (A_{xD} + C^T A_{DD}) D \right) - x \right). \]

Plugging (A.56) in (A.55) and then proceeding as in part (i), we obtain

\[ \dot{\alpha}_s = b_2 \alpha_s - b_4 \left[ \begin{array}{c} 0 \\ \gamma \Sigma \end{array} \right] M_s, \]

where we redefine \( \alpha_s \) as the expected component of the aim depending on \( f \),

\[ \alpha_s = E_0 \left[ (A_{xx} + C^T + C^T A_{Dx})^{-1} \left( A_{xf}(f) + C^T A_{Df}(f) \right) \right], \]
and

\[
\tilde{M}_1 = \begin{bmatrix}
- (A_{xx} - A_{xD}C - C) \\
A_{xD}^\top - A_{DD}C
\end{bmatrix} \Lambda^{-1} \begin{bmatrix} I & C^\top \end{bmatrix}
\]

\[
b_2 = \rho - \tilde{M}_1 + \begin{bmatrix} 0 & 0 \\
0 & R^\top
\end{bmatrix}
\]

\[
b_4 = (A_{xx} - C^\top A_{xD}^\top - C^\top)^{-1} \begin{bmatrix} I & C^\top \end{bmatrix}.
\]

The result follows immediately.

(iii) We follow the same strategy. The conjectured value function, etc., are

\[
\hat{V} = \frac{1}{2} D^0 \Lambda D^0 + D^0 A_D f(f) + A_{ff}(f)
\]

\[
\rho \hat{V} = \sup_x \left\{ \frac{\rho}{2} x^\top C x + x^\top B f - x^\top (r^f + R)(D^0 + C x) - \frac{\gamma}{2} x^\top \Sigma x \\
- \hat{V}_D R(D^0 + C x) - D^0 A_D f(f) + D A_{ff}(f) \right\}
\]

\[
x = J^{-1} \left( B f - C^\top R^\top A_D f(f) - (C^\top R^\top A_{DD} + r^f + R) D^0 \right).
\]

This time let

\[
\alpha_s = E_0 [A_D f(f)]
\]

and derive, using (A.58),

\[
\dot{\alpha}_s = -\rho \alpha_s - (A_{DD} R C + r + R^\top) J^{-1} \left( B E [f_t | f_0] - C^\top R^\top \alpha_s \right) - R^\top \alpha_s
\]

\[
\equiv b_2 \alpha_s - b_4 M^f_s
\]

for constant matrices \(b_i\). Solving for \(\alpha_0\) and using (A.59) yields the result. ■
Proof of Proposition 10. Specifically, the HJB equation reads

\[
0 = \sup_{\tau_s} \left\{ x_s^\top B_f s - \frac{\gamma}{2} x_s^\top \Sigma x_s v_s - \frac{\lambda}{2} \tau_s^\top \Sigma \tau_s v_s - \rho V + V_x \tau_s - V_f \Phi f + \frac{1}{2} \text{tr} \left( V_{ff} \Omega \right) \\
+ V_v \mu + \frac{1}{2} V_{vv} \sigma^2 + V_v f \frac{d}{ds} [f, v]_s \right\},
\]

with

\[
V_v = -\frac{1}{2} x^\top A'_{xx} x + x^\top A'_{xf} f + \frac{1}{2} f^\top A'_{ff} f + A'_0 \tag{A.61}
\]

\[
V_{vv} = -\frac{1}{2} x^\top A''_{xx} x + x^\top A''_{xf} f + \frac{1}{2} f^\top A''_{ff} f + A''_0. \tag{A.62}
\]

\[
\]

Proof of Proposition 11. Suppose that \( E_t[dp_t - r^t p_t dt] = B_f t dt \) with \( B \) given by (62) and apply Proposition 1 to conclude that, if \( x_t = -f^{K+1}_t \), then \( dx_t = -df^{K+1}_t \). The comparative-static results are immediate. \( \blacksquare \)
References


