Multi-Asset Market Making

Ioanid Roşu, Andriy Shkilko, Elvira Sojli, and Wing Wah Tham*

March 27, 2026

Abstract

We study how risk-averse market makers manage inventory by setting quotes in multiple assets. As the correlation between assets increases, an inventory shock in one asset causes the market maker to quote more in other assets but less in the affected asset. The model yields five predictions. Cross-hedging exists: after an inventory shock in one asset, the dealer quotes more aggressively in correlated assets. Cross-hedging increases with fundamental correlation, while own-hedging decreases (a substitution effect). A more risk-averse dealer cross-hedges less, because she mean-reverts each position to zero so quickly that portfolio rebalancing has no time to be valuable. Through an inventory-driven price-pressure channel, high dealer risk aversion generates a term structure of return correlations that attenuates at higher frequencies: the Epps effect. Finally, ignoring cross-hedging leads to systematic underestimation of market maker risk aversion; in the limit of perfect correlation, the bias reaches 50%. We test all five predictions using complete message-level data from the Toronto Stock Exchange.

KEYWORDS: Market making, inventory, multiple assets, cross-hedging, Epps effect.

*Roşu (corresponding author), rosu@hec.fr, HEC Paris; Shkilko, ashkilko@wlu.ca, Wilfried Laurier University; Sojli, e.sojli@unsw.edu.au, University of New South Wales; and Tham, w.tham@unsw.edu.au, University of New South Wales. We thank Bruno Biais, Davide Cianciaruso, Thierry Foucault, Denis Gromb, Terrence Hendershott, Stefano Lovo, and Daniel Schmidt for their suggestions. We are also grateful to finance seminar participants at HEC Paris and Peking University, as well as conference participants at the 10th Erasmus Liquidity Conference, the 10th Women in Market Microstructure Conference, and the 10th Annual Conference of Romanian Academic Economists from Abroad (ERMAS), for valuable comments.

> *"A market maker's success depends on it posting the best available prices and accurately responding to relevant market data in similar and correlated instruments."*
>
> — Virtu Financial, Inc., S-1 Registration Statement, U.S. Securities and Exchange Commission (2014)

Market makers are the liquidity supply side of financial markets, posting bid and ask quotes at which other participants trade. On the liquidity demand side, investors buy and sell assets to form and rebalance portfolios, motivated by differences in views, information, or endowment. While that demand side has been extensively analyzed through the lens of Modern Portfolio Theory, the supply side has received comparatively little attention, especially when the dealer is active in multiple assets simultaneously. Two questions then arise naturally. How does a market maker manage inventory risk when positions in different assets are correlated? And what effect does this multi-asset risk management have on prices?

In Modern Portfolio Theory, investors allocate capital across assets to trade off expected return against risk. Much less is known about the corresponding portfolio problem faced by the intermediaries who make those assets tradeable. A market maker typically provides liquidity in many securities at once, and shocks to her inventory in one name change the risk of her entire book. Virtu Financial, one of the largest electronic market makers in the world, puts it plainly in its public filings: a market maker's success depends on accurately responding to market data in *similar and correlated instruments* (Virtu Financial, Inc., 2014). The firm simultaneously quotes prices for over 25,000 securities across 235 venues in 37 countries, managing inventory across all of them as a single correlated portfolio. Yet the canonical inventory-based market-making literature largely analyses quote setting one traded asset at a time, with inventory control operating through quote shading in that single asset. Even when a dealer's background portfolio risk enters the objective, the theory rarely models jointly chosen quotes across multiple traded assets that allow active cross-hedging.[1] The single-asset framework de-

[1]The distinction is between a single traded asset with exogenous background portfolio risk, as in Ho and Stoll (1981) and related work, and a setting with multiple traded assets in which the dealer jointly sets quotes across all of them and can exploit correlations to hedge inventory actively.

veloped by Ho and Stoll (1981) and extended by Hendershott and Menkveld (2014) has been enormously productive. It is silent, however, on how a dealer allocates inventory adjustment across correlated assets inside her own book. This cross-asset dimension is precisely the margin that a portfolio perspective makes natural.

We propose a dynamic model of market making in N risky assets. Each trading round, a monopolist dealer sets ask and bid quotes in all assets at which investors trade. The market consists of one risk-free asset and several risky assets, whose fundamental value evolves according to a random walk with multivariate normal increments. Trading in the risky assets takes place in a market exchange, at discrete dates. After the dealer sets ask and bid quotes, the investors submit an aggregate market order that is the sum of an inelastic component, an elastic component linear in the fundamental value minus the quote, and a multivariate normal noise. The dealer maximizes her expected profit subject to a penalty that is quadratic in her inventory in the risky assets.

The equilibrium can be obtained in closed form, and has a simpler form when the discount coefficient approaches 1. As the correlation between two assets increases, an inventory shock in one asset (asset 1) causes the market maker to quote more in other assets (asset 2) but less in asset 1. Indeed, when the correlation between the two assets increases, the need of inventory hedging in asset 1 decreases due to a substitution effect. In the extreme case when the two assets become perfectly correlated, only the sum of inventories matters to the dealer and the assets become perfect substitutes. This implies that, compared to the non-perfectly correlated case, asset 1 is used less intensely for hedging, since asset 2 has become a better substitute.

If the market maker's risk aversion increases, her inventory in each asset reaches zero more often. Indeed, a more risk-averse market maker does not tolerate large inventories and sets quotes such that the inventory decreases to zero more quickly. This implies that the number of times an inventory reaches zero in an asset can be used as empirical proxy for the dealer's risk aversion. However, this has to be done with caution, as the failure to account for cross-asset inventory management leads to biased risk aversion estimates.

A more risk-averse market maker engages less in hedging inventory across assets. Intuitively, a very risk-averse dealer is more concerned about minimizing the inventory

penalty than maximizing the expected profit component. But in the limit, minimizing the inventory penalty means setting the quotes such that the inventory in each asset is expected to drop to approximately zero in each period, meaning that the liquidity shock in each asset is expected to be almost fully reversed in each period. As the liquidity shocks are assumed uncorrelated among the different assets, the pricing discounts are set almostly independently of each other, and therefore the pricing coefficient is approximately diagonal. Differently said, a very risk-averse dealer mean-reverts her inventory to zero very frequently, and hence hedging across assets has little value.

As a consequence of the previous result, when the market maker's risk aversion is high, the asset return correlation is low at higher frequencies. This is related to the effect discovered in Epps (1979), that stock returns are relatively uncorrelated at high frequencies (less than 10 minutes) but approach a constant correlation at low frequencies (more than 1 day). In our model, the inventory effect in prices induces a mean-reverting component at higher frequencies that disappears at lower frequencies, for which the random walk component of prices dominates. At low frequencies, the return correlation approaches the fundamental correlation, which is assumed constant over time.

Ignoring inventory cross-hedging by restricting analysis to a single asset leads to the underestimation of market maker risk aversion. Indeed, suppose the dealer's inventory in one asset has a half-life that allows reverse-engineering the dealer's risk aversion by considering the asset in isolation. This is equivalent to assuming that the other assets are uncorrelated and that there is no inventory cross-hedging. However, if the other assets are in fact correlated, it takes longer to reduce inventory to zero, as the other assets can be used for hedging and the need to reduce the inventory is effectively weaker. In the extreme case of perfect correlation with another asset, studying one asset in isolation would lead to an understimation of the dealer's risk aversion by 50%.

The key object is the $N \times N$ pricing matrix λ. The dealer's equilibrium midquote in each asset is the fundamental value minus a linear combination of her inventories in all assets:

$$p_t = v_t - \lambda x_t.$$

The diagonal entries λ_{ii} capture own-stock price pressure; the off-diagonal entries λ_{ij}

capture cross-stock price pressure. When the fundamental covariance Σ_v has positive off-diagonal entries, so does λ: a dealer who is long in asset 1 raises her midquote in asset 2 as well. This is inventory cross-hedging, and it is the central object of the paper. The model delivers five testable predictions.

Cross-hedging (Fact 1). After an inventory shock in stock 1, the dealer adjusts quotes in correlated stocks to attract offsetting flow. The off-diagonal entries of λ are positive whenever the fundamental correlation $a = \mathsf{corr}(\Delta v_1, \Delta v_2) > 0$.

Cross-hedging increases with correlation (Fact 2). As a increases, the off-diagonal entry λ_{12} increases. Correlated stocks become better substitutes for inventory reduction, so the dealer relies on them more.

Own-hedging decreases with correlation (Fact 3). As a increases, the diagonal entry λ_{11} decreases: a substitution effect. In the limit $a \to 1$, only the sum of inventories matters and the two assets are perfect substitutes.

Risk aversion reduces cross-hedging (Fact 4). As the risk aversion coefficient γ grows, λ approaches a diagonal matrix. A very risk-averse dealer mean-reverts each inventory to zero so rapidly that cross-hedging has no time to be useful: in the limit, the liquidity shock in each asset is fully reversed within a single period, pricing discounts are set independently across assets, and the off-diagonal entries of λ vanish.

The Epps effect (Fact 5). High dealer risk aversion implies diagonal λ, which implies that quotes in different assets respond independently to their respective inventories. At high frequencies, this inventory-driven price-pressure channel generates a mean-reverting component in prices that suppresses return correlations relative to their fundamental level. At low frequencies, the random-walk component of prices dominates and return correlations converge to the fundamental correlation a. This is a theoretical version of the Epps (1979) effect, that stock return correlations are lower at high frequencies than at low frequencies.

A further consequence of Fact 4 concerns the estimation of market maker risk aversion. Single-asset models, such as Ho and Stoll (1981) and Hendershott and Menkveld (2014), use the inventory half-life to back out γ. But with correlated assets and cross-hedging, the half-life of any individual asset is longer than it would be in isolation,

because the dealer can partially offset inventory using other assets and therefore faces less urgency to correct any individual position. Treating the other assets as if they were uncorrelated understates γ; the underestimation reaches 50% when two assets are perfectly correlated.

We test the model's core predictions using the complete message-level record of the Toronto Stock Exchange STAMP 4.0.3 feed for January 13, 2006. The sample contains 8.2 million time-stamped order and trade messages across 3,169 symbols. We focus on this date because the TSX feed provides the full order lifecycle with dealer-level identifiers, a combination that is rare in publicly available data.[2] We identify five market maker populations that differ in institutional obligations and discretion, ranging from Designated Market Makers (DMMs) with specialist responsibilities to electronic market makers (EMMs) whose liquidity provision is voluntary. Sub-second timestamps, full order-book reconstruction, and a user identifier on every message allow us to reconstruct each dealer's inventory trade by trade and to measure quoting responses in real time.

Our empirical strategy isolates inventory shocks and traces their propagation across the dealer's book. We define an inventory shock as a trade whose size exceeds two standard deviations above the dealer's pair-specific mean. Following a buy shock in one stock, we measure the change in the dealer's sell ratio, the fraction of new limit orders that are offers, in the other stocks in which the same dealer is an identified market maker. The results are sharp. Cross-hedging is pervasive: after a buy shock in one stock, dealers across all five populations shift toward sell-side quoting in other stocks where they make markets, with sell-ratio changes ranging from 0.019 to 0.184 depending on dealer type and stock universe. Three features of the response support an inventory-management interpretation rather than mechanically correlated customer flow. First, the cross-stock response builds gradually: it is statistically indistinguishable from zero at 60 seconds and reaches its full magnitude only at 300 seconds. Institutional portfolio execution can unfold over minutes, but the 300-second build is longer than typical basket-execution horizons and inconsistent with simultaneous correlated order flow, which reaches all quoted names within seconds. Second, the response is roughly

[2]The mechanisms we study operate on intra-day horizons. Our tests exploit cross-sectional variation across dealer-stock pairs within the day, rather than relying on time-series variation across days.

twice as large for involuntary shocks that create genuinely new inventory imbalances relative to shocks that are themselves inventory-management trades, supporting the causal direction. Third, the cross-stock adjustment arises through passive limit-order submission rather than aggressive trading, as the model implies. In addition, when asset pairs are negatively correlated, the model predicts a sign reversal in cross-stock quote adjustments, and we confirm this sign reversal in the data, a pattern that correlated customer flow stories do not generate cleanly.

Cross-hedging strengthens with long-horizon return correlation, consistent with the model's off-diagonal comparative static. Using monthly return correlations from Compustat Global over January 2003 to December 2005 as a proxy for long-run co-movement, we find that a one-standard-deviation increase in a dealer's average portfolio correlation raises the cross-stock sell-ratio response by about 29 percent relative to its unconditional mean. The effect is more than twice as large for involuntary shocks, where the coefficient rises to 0.203 ($t = 3.21$), and it is strongest for Set C dealers, who combine formal specialist assignments with voluntary market making, where the coefficient is 0.225 ($t = 2.16$). The composition of adjustment also shifts with correlation. The share of total inventory response allocated to cross-stock channels increases monotonically from roughly 13 percent in the lowest-correlation quintile to 27 percent in the highest. Among DMMs, whose assignments are determined by exchange allocation rather than dealer choice, own-stock quote adjustment declines as average portfolio correlation rises, with a coefficient of -0.155 ($t = -1.96$), consistent with the level substitution predicted by the model. Among EMMs the opposite pattern appears, consistent with selection: the most active EMMs choose to make markets in correlated portfolios, so correlation and own-asset hedging intensity move together through a common underlying factor rather than through the substitution channel.

The portfolio dimension of inventory management matters for measurement and interpretation. In the single-asset literature, observing an inventory half-life in one asset and mapping it to risk aversion is equivalent to assuming that the dealer does not hedge through correlated positions. When the dealer is managing a correlated book, any one position is unwound more slowly because correlated assets absorb part of the adjustment burden, which weakens the incentive to eliminate the focal position directly.

In the knife-edge case of perfect correlation with another asset, studying one asset in isolation implies an underestimation of risk aversion by 50 percent; for empirically relevant correlations the bias is smaller but remains economically meaningful. More generally, single-asset measures of price impact, market quality, or inventory adjustment speed can be systematically biased when dealers manage correlated portfolios, and the direction and magnitude of the bias vary with the correlation structure in the way our model specifies.

Our paper contributes to several strands of the literature. The first is the theory of inventory-based market making. Ho and Stoll (1981) derive the optimal bid-ask spread for a dealer facing return uncertainty in a single asset, and Hendershott and Menkveld (2014) bring that framework to the data, identifying price pressure as the mechanism by which dealers restore inventory balance. Stoll (1978) and Amihud and Mendelson (1980) establish the foundational link between inventory costs and the bid-ask spread in single-asset settings. Our model extends this line of work to multiple correlated assets and shows that the single-asset pricing scalar generalises to an $N \times N$ pricing matrix whose off-diagonal elements are both theoretically disciplined and empirically detectable. Daures-Lescourret and Moinas (2023) model imperfect competition among two market makers posting prices for the same asset on two exchanges, focusing on how fragmentation affects liquidity; our focus is instead on how a single dealer optimally allocates inventory adjustment across correlated assets and what that implies for price formation, market quality, and the measurement of risk aversion.

The second strand is the empirical literature on market maker behavior. Menkveld (2013) documents the inventory dynamics of a large high-frequency market maker in a single asset, and Brogaard, Hendershott, and Riordan (2014) study high-frequency traders and price discovery across stocks while treating each stock independently. Copeland and Galai (1983) examine quote-setting under adverse selection in a single-asset framework. More broadly, the literature on dealer networks, including Hugonnier, Lester, and Weill (2020) and Di Maggio, Kermani, and Song (2017), examines how intermediaries interact with each other across markets but does not address how an individual dealer manages inventory across correlated positions within her own book. Our empirical contribution is to bring the cross-asset dimension into this literature directly. Rather than

studying each stock in isolation, we reconstruct the full multi-asset inventory position of each identified dealer and measure how shocks in one stock propagate into quoting behavior across the dealer's entire book. To our knowledge, we are the first to jointly derive a multi-asset closed-form inventory-tilting pricing matrix and test its cross-asset predictions using dealer-identified message-level order data across a broad equity universe with structurally distinct market maker populations.

Finally, our model provides a complementary theoretical mechanism for the Epps effect. Epps (1979) documents that stock return correlations decline systematically as the measurement frequency increases, a finding replicated across markets and time periods. The leading explanations attribute the attenuation to nonsynchronous trading (Scholes and Williams, 1977), bid-ask bounce (Roll, 1984), and differential information flow (de Jong and Nijman, 1997). Our model adds an inventory-based channel: a risk-averse dealer who mean-reverts inventory aggressively injects a transitory component into prices that attenuates measured co-movement at high frequencies but vanishes at lower frequencies where the random walk component of fundamental value dominates. This mechanism is complementary to rather than competing with the canonical explanations, and its distinctive prediction is that the degree of attenuation should vary with the aggregate risk aversion of the dealers active in a stock pair.

1 Model

1.1 Environment

Our model is close in spirit to the dealer models of Ho and Stoll (1981) and Hendershott and Menkveld (2014), except that we consider trading in multiple assets. Specifically, the market consists of one risk-free asset and N risky assets. Trading in the risky assets takes place in a market exchange, at discrete dates $t = 0, 1, 2, \ldots$, according to the mechanism described below. There are two types of market participants: (a) one monopolistic market maker called the dealer ("she") who sets ask and bid quotes at which others trade, and (b) traders, who submit market orders.

Assets. The risk-free asset is used as a numeraire and has a return of zero. Each

risky asset $n \in \{1, 2, \ldots, N\}$ has a fundamental value $v_{n,t}$ that follows a continuous multi-variate random walk process for which the increments have instant covariance matrix per unit of time equal to Σ_v. We define the fundamental value v_t of the N assets as the column vector with n'th entry equal to $v_{n,t}$. In Appendix B, we show that $v_{n,t}$ can be interpreted as the cash value that shareholders receive at liquidation, an event which may occur in each period with a constant probability.[3]

Dealer's Quotes and Objective. Suppose the dealer observes the fundamental value v_t before trading at t and sets the ask quote vector a_t and the bid quote vector b_t. Let x_t be the dealer's inventory in the risky assets before trading at t, which is also a column vector with N entries. If Q_t^b is the aggregate buy market order vector at t and Q_t^s is the aggregate sell market order vector at t, the dealer's inventory evolves according to

$$x_{t+1} \;=\; x_t - Q_t^b + Q_t^s. \tag{1}$$

For a given quoting strategy, the dealer's expected utility at $\tau = 1, 2, \ldots$ is:

$$\mathsf{E}_\tau \sum_{t=\tau}^{\infty} \beta^{t-\tau} \Big((v_t - b_t)' Q_t^s \;+\; (a_t - v_t)' Q_t^b \;-\; \gamma\, x_t' \Sigma_v x_t \Big), \tag{2}$$

where $\beta \in (0, 1)$ and $\gamma > 0$.[4] The dealer maximizes expected profit, but at each t faces a utility loss that is quadratic in the inventory. In Appendix B, we show that the expected utility of a risk-averse dealer is approximately of the form (2) when her risk aversion coefficient γ is sufficiently small.[5]

Trading. Upon observing the quotes, traders submit at t the following aggregate market order vectors:

$$\begin{aligned} Q_t^b \;&=\; \frac{k}{2}(v_t - a_t) + \ell + \varepsilon_t^b, \quad \text{with} \quad \varepsilon_t^b \overset{IID}{\sim} \mathcal{N}\left(0, \Sigma_L/2\right), \\ Q_t^s \;&=\; \frac{k}{2}(b_t - v_t) + \ell + \varepsilon_t^s, \quad \text{with} \quad \varepsilon_t^s \overset{IID}{\sim} \mathcal{N}\left(0, \Sigma_L/2\right), \end{aligned} \tag{3}$$

[3] Suppose there exists $\pi \in (0, 1)$ such that asset n liquidates in each period with probability π, in which case the shareholders receive $v_{n,t}$ per share. Then we show in Appendix B that the expected profit of a trader with quantities bought and sold at t equal to $-Q_t^b$ and $-Q_t^s$, respectively, has the form described in equation (2) with $\beta = 1 - \pi$.

[4] The equilibrium results are the same in the limit when $\beta = 1$ if the dealer's inventory penalty $\gamma x_t' \Sigma_v x_t$ in equation (2) is replaced by $\gamma x_{t+1}' \Sigma_v x_{t+1}$; see Proposition A.1 in the Appendix.

[5] This is done for one asset, but in principle the proof holds also for multiple assets.

where k and Σ_L are symmetric positive definite matrices, and ℓ is a column vector.[6] The parameter k is the investor elasticity matrix and ℓ is the inelasticity parameter vector.

Dealer's Information. The main results do not depend on any particular information structure, but to obtain explicit formulas we define the dealer's information set $\mathcal{I}_t$ before trading at t as the history of net order flow: $\mathcal{I}_t = \left\{Q^b_\tau - Q^s_\tau \mid \tau < t\right\}$.[7]

2 Equilibrium

2.1 Optimal Quotes

In this section, we solve for the dealer's optimal quotes and show that the results are obtained in closed form. We fix the dealer's risk aversion coefficient γ, the fundamental covariance Σ_v, the elasticity matrix k, the inelasticity parameter ℓ, and the noise covariance Σ_L. The matrices k and Σ_L are assumed symmetric and positive definite. Let I_N be the identity matrix in N dimensions.

Proposition 1 describes the optimal quotes set by the dealer. This result is obtained by applying standard methods in linear-quadratic dynamic programming, see Appendix A.[8] We fix γ, Σ_v, k, ℓ, and Σ_L. The matrices k and Σ_L are symmetric and positive definite. Let I_N denote the identity matrix.

Proposition 1. *If v_t is the fundamental value at t and x_t is the dealer's inventory before trading at t, then the dealer's optimal quotes at $t = 0, 1, \ldots$ are:*

$$a_t \;=\; v_t - \lambda x_t + h, \qquad b_t \;=\; v_t - \lambda x_t - h, \tag{4}$$

where h is an $N \times 1$ vector and λ is an $N \times N$ matrix satisfying:

$$h \;=\; k^{-1}\ell, \qquad \lambda \;=\; k^{-1}\Lambda, \tag{5}$$

[6] By extending the microfoundations from Roşu, Sojli, and Tham (2021), k can be shown to be proportional to $g^{-1}\Sigma_u^{-1}$, where g is the informed trader's risk aversion, $v + u$ is the asset's liquidation value, and v is the informed trader's signal. A similar result obtains from CARA investors with expected values equal to v, in which case k^{-1} is the conditional variance of $v + u$.

[7] Alternatively, the information set could contain both the buy and sell order flow separately, i.e., $\mathcal{I}_t = \left\{Q^b_\tau, Q^s_\tau | \tau \in \{0, 1, \ldots, t-1\}\right\}$, but the formulas would be more complicated.

[8] For a general treatment of such problems, see Ljungqvist and Sargent (2000).

where Λ is the unique symmetric positive definite matrix solving the quadratic equation:

$$\Lambda^2 + \Lambda\Big(S + \frac{1-\beta}{\beta} I_N\Big) \;=\; S, \quad \textit{with} \quad S \;=\; \gamma k \Sigma_v. \tag{6}$$

When β converges to 1, the parameter λ converges to:

$$\lambda \;=\; \frac{k^{-1}}{2}\Big((S^2 + 4S)^{1/2} - S\Big). \tag{7}$$

We consider only the limit $\beta = 1$ henceforth, in order to simplify formulas.

To get intuition for Proposition 1, suppose there is only one asset ($N = 1$) and the dealer is risk-neutral ($\gamma = 0$). Equation (7) shows that $\lambda = 0$, hence the dealer's inventory does not affect her strategy. Equation (4) implies that the dealer sets her quotes at equal distance around the value v_t: the ask quote is $a_t = v_t + h$ and the bid quote is $b_t = v_t - h$, where h is the constant half-spread. The equilibrium value $h = \ell/k$ in equation (5) reflects two opposite concerns for the dealer: too large a spread would decrease the profit that comes from the price-elastic investors (with elasticity k), while too small a spread would lower the profit that comes from the inelastic traders (who submit a fixed quantity ℓ at both quotes).

If the dealer is risk-averse, equation (7) shows that the dealer's inventory affects the quotes across assets in a linear way via the coefficient λ, i.e., the midquote price $p_t = (a_t + b_t)/2$ satisfies:

$$p_t \;=\; v_t - \lambda x_t. \tag{8}$$

To understand why, suppose before trading at t, the dealer has accommodated a positive net order flow and thus has a negative inventory ($x_t < 0$). To bring back her inventory to zero, she must encourage new sellers and discourage new buyers, which can be achieved by raising both quotes at t. Thus, the dealer raises the midquote above the value v_t, to a level equal to $p_t = v_t + \lambda(-x_t)$. For one asset, this is the price pressure mechanism identified in Hendershott and Menkveld (2014).

To get intuition for the case with multiple assets, we analyze in more detail the formula (7) for λ. If C is an $N \times N$ matrix, we denote by $|C|$ the infinity norm of C, i.e., the maximum absolute row sum of C. We use the big-O notation extended to matrices

in the standard way, with asymptotic behavior considered near the zero matrix.[9]

Corollary 1 provides approximations to the equation (7) in two extreme cases, when the risk aversion coefficient γ is either very high or very low.

Corollary 1. *In the context of Proposition 1, if* $\left|(\gamma k\Sigma_v)^{-1}\right| < 1/4$ *then:*

$$\lambda \;=\; k^{-1} + k^{-1}O\big((\gamma k\Sigma_v)^{-1}\big). \tag{9}$$

If $\left|\gamma k\Sigma_v\right| < 4$ *then:*[10]

$$\lambda \;=\; \gamma^{\frac{1}{2}}k^{-\frac{1}{2}}\Sigma_v^{\frac{1}{2}} + k^{-1}O\big(\gamma k\Sigma_v\big). \tag{10}$$

Equation (9) shows that when γ is large, $\lambda \approx k^{-1}$: the pricing matrix is approximately diagonal. A very risk-averse dealer is more concerned with minimizing the inventory penalty than with exploiting portfolio correlations. In the limit, she sets prices so that the inventory in each asset is expected to drop to zero in the next period. Since the liquidity shocks ε_t are independent across assets, the optimal pricing discounts are set independently, and λ becomes diagonal. A very risk-averse dealer mean-reverts each inventory so quickly that cross-hedging has no value.

2.2 Two Risky Assets

In this section, we focus on $N = 2$ with symmetric fundamental volatility and scalar elasticity to obtain simpler closed-form expressions.

Corollary 2. *Let* $N = 2$ *and suppose that:*

$$\Sigma_v \;=\; \sigma_v^2\begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}, \quad k = \begin{bmatrix} k_0 & 0 \\ 0 & k_0 \end{bmatrix}. \tag{11}$$

[9] If $X = X(c)$ is a scalar function with matrix values and $f(c)$ is a positive function (at least for small c), the notation $X = O(f(c))$ means that there exist $M, \eta > 0$ such that for all $c \in (0, \eta)$ the ratios $|X(c)|/f(c)|$ are bounded by M. If $X = X(C)$ is a matrix function with matrix values and $f(C)$ is positive definite for C positive definite and sufficiently small (in norm), the notation $X = O(f(C))$ means that the ratios $\big(f(C)\big)^{-1}X(C)$ are uniformly bounded in norm for C sufficiently small.

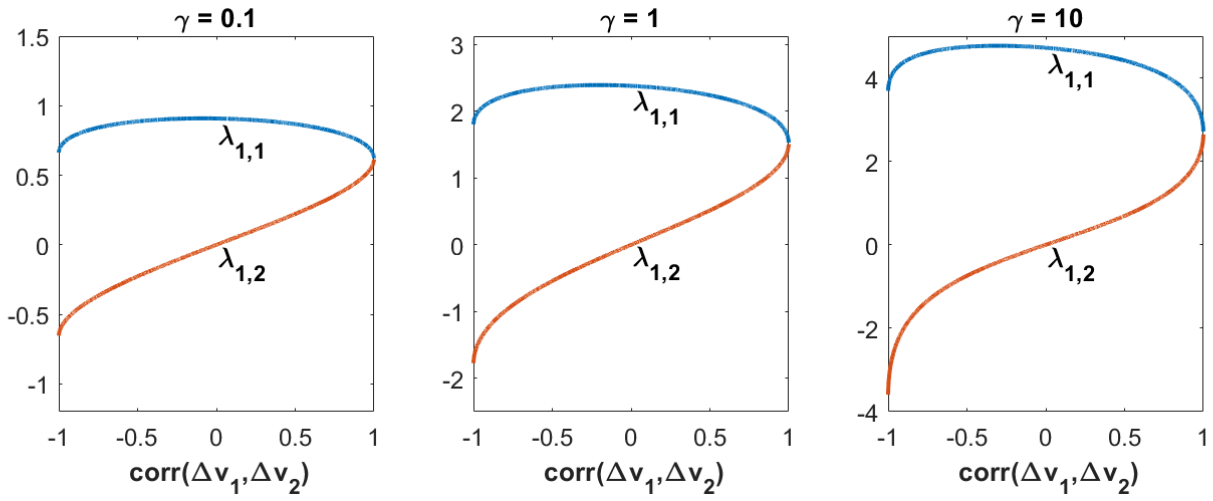
[10] In Appendix B, we analyze the microfoundations of the dealer's expected utility in the one-dimensional case ($N = 1$), and we show that the inequality $\gamma k\sigma_v^2 \ll 1$ is one of the conditions that are sufficient to obtain a quadratic penalty as in equation (2).

Then the dealer's pricing coefficient is:

$$\lambda = \frac{1}{k_0(m_1+1)(m_2+1)} \begin{bmatrix} m_1+m_2+2 & m_1-m_2 \\ m_1-m_2 & m_1+m_2+2 \end{bmatrix}, \quad \textit{with} \tag{12}$$
$$m_1 = \left(4\gamma^{-1}k_0^{-1}\sigma_v^{-2}(1-a)^{-1}+1\right)^{\frac{1}{2}}, \quad m_2 = \left(4\gamma^{-1}k_0^{-1}\sigma_v^{-2}(1+a)^{-1}+1\right)^{\frac{1}{2}}.$$

Figure 1 shows comparative statics of the pricing matrix λ with respect in the dealer's risk aversion coefficient (γ) and the fundamental correlation coefficient (a).

Figure 1: Dealer's Pricing Function and Risk Aversion. In the context of Corollary 2, this figure shows how the coefficients $\lambda_{1,1}$ and $\lambda_{1,2}$ of the dealer's pricing function λ vary with the increment correlation between the two components of the fundamental value: $a = \mathsf{corr}(\Delta v_1, \Delta v_2)$. The dealer's risk aversion coefficient is $\gamma = 0.1$ in the left graph, $\gamma = 1$ in the center graph, and $\gamma = 10$ in the right graph. In all graphs, the fundamental volatility is $\sigma_v = 1$ and the elasticity matrix is $k = k_0 I_2$, with $k_0 = 0.1$.



As Figure 1 shows, there is significant cross-hedging across the two assets. For instance, suppose the dealer's risk aversion coefficient is $\gamma = 1$, the fundamental volatility is $\sigma_v = 1$, and the elasticity matrix is $k = k_0 I_2$, with $k_0 = 0.1$.[11] Suppose the fundamental correlation between the two assets is $a = 0.6$. Corollary 2 implies that the pricing matrix is:

$$\lambda_{(a=0.6)} = \begin{bmatrix} 2.54 & 0.73 \\ 0.73 & 2.54 \end{bmatrix}. \tag{13}$$

Suppose the value of both assets is 10 and the dealer's starting inventory before trading

[11] We choose a relatively small value of k_0 to indicate that the order flow is not very elastic in the dealer's quotes.

at t is $x_t = [-1, +1]'$, i.e., the dealer is short one unit of asset 1 and long one unit of asset 2. The pricing equation (8) implies that the dealer must set the midquotes at $p_t = v_t - \lambda x_t$:

$$p_t = \begin{bmatrix} 10 \\ 10 \end{bmatrix} - \begin{bmatrix} 2.54 & 0.73 \\ 0.73 & 2.54 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 11.81 \\ 8.19 \end{bmatrix}. \tag{14}$$

If instead the dealer had zero units in asset 2, the midquotes would be set at:

$$p_t = \begin{bmatrix} 10 \\ 10 \end{bmatrix} - \begin{bmatrix} 2.54 & 0.73 \\ 0.73 & 2.54 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 12.54 \\ 10.73 \end{bmatrix}. \tag{15}$$

Thus, being long one unit of asset 2 helps reduce the pressure to increase inventory in asset 1, and as a result the dealer sets the midquote of asset 1 closer to the fundamental value of 10, at a rate equal to $0.73/2.54 \approx 29\%$ of the extra inventory in asset 2. The rate of cross-hedging increases with the correlation between assets 1 and 2, reaching 100% in the limit when the assets are perfectly correlated.

Another implication of the pricing equation (15) is that, even if the dealer is only short in asset 1, the dealer also increases her midquote in asset 2, by 0.73. The intuition for this fact is similar to that of standard portfolio theory, where in order to reduce the overall risk of a portfolio, it is more efficient to hedge using other assets as well.

If the fundamental correlation were to increase from $a = 0.6$ to $a = 0.8$, the dealer's pricing coefficient would change to:

$$\lambda_{(a=0.8)} = \begin{bmatrix} 2.38 & 1.06 \\ 1.06 & 2.38 \end{bmatrix}. \tag{16}$$

Comparing λ with the pricing matrix from equation (13), the cross-hedging coefficient $\lambda_{1,2}$ increases from 0.73 to 1.06. The intuition is straightforward: When the correlation between the two assets increases, asset 2 is used at a higher rate for hedging inventory in asset 1. At the same time, the own-hedging coefficient $\lambda_{1,1}$ decreases from 2.54 to 2.38. To understand this result, consider the more extreme case when the fundamental correlation becomes very close to $a = 1$, i.e., when the two assets are perfectly correlated.

In that case, the pricing matrix has all entries equal to each other (with a value of 1.79). Thus, only the sum of inventories matters to the dealer and the assets are almost perfect substitutes. This implies that, compared to the case when $a = 0.6$, asset 1 is used less intensely for hedging, since asset 2 has become a better substitute.

If the dealer's risk aversion γ is very large, Corollary 1 implies that λ is approximately diagonal.[12] (This can also be seen in Figure 1 when $\gamma = 10$.) Intuitively, a very risk-averse dealer is more concerned about minimizing the inventory penalty than maximizing the expected profit component. But in the limit, minimizing the inventory penalty means setting the prices such that the inventory in each asset is expected to drop to approximately zero in each period. This means that the dealer sets prices such that the liquidity shock in each asset is expected to be almost fully reversed in each period. As the liquidity shocks are assumed uncorrelated among the different assets, the pricing discounts are set almostly independently of each other, and therefore the pricing coefficient λ is approximately diagonal. Differently said, a very risk-averse dealer mean-reverts her inventory to zero very frequently, and hence hedging across assets has little value.

2.3 Correlation and the Dealer's Asset Selection

In this section, we analyze how the dealer chooses in which assets to provide quotes. For simplicity, we assume that the dealer already provides quotes in one asset and must choose one more asset, based on the fundamental correlation between them.

To solve this problem, we must calculate the value function of the dealer. The proof of Proposition 1 shows that in the general case with N risky assets, the dealer's value function is a quadratic function of the inventory x:

$$V(x) \;=\; \frac{1}{1-\beta}\Big(\ell' k^{-1}\ell - \beta\alpha \circ \Sigma_L\Big) \;-\; \frac{1}{\beta}x'\alpha x, \tag{17}$$

where $A \circ B = \sum_{i,j=1}^{N} A_{ij}B_{ij}$ denotes the Frobenius inner product (or double dot product) of two $N \times N$ matrices A and B, $\beta < 1$ is the discount rate, α is the $N \times N$ matrix

[12] In the case of $N = 2$ assets, this can be proved directly using Corollary 2: When γ approaches infinity, m_1 and m_2 are approximately equal to 1, and λ is approximately proportional to the identity matrix, i.e., $\lambda \approx k_0^{-1} I_2$.

that satisfies equation (**??**), and Σ_L is the covariance matrix of the liquidity demand $Q_t^b - Q_t^s$.

To simplify formulas, we define the dealer's normalized value function $W(x)$ as the limit of $(1-\beta)V(x)$ when β converges to 1. According to equation (17), the normalized value function is:

$$W(x) \;=\; \ell' k^{-1} \ell - \alpha \circ \Sigma_L. \tag{18}$$

Corollary 3 provides formulas for the simpler case when $N = 2$, as well as asymptotic limits in the extreme cases when the dealer is either almost risk-neutral or very risk-averse.

Corollary 3. *Let $N = 2$ and suppose that:*

$$\Sigma_v \;=\; \sigma_v^2 \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}, \quad k = \begin{bmatrix} k_0 & 0 \\ 0 & k_0 \end{bmatrix}, \quad \ell = \begin{bmatrix} \ell_1 \\ \ell_2 \end{bmatrix}, \quad \Sigma_L = L_0 \begin{bmatrix} 1 & b \\ b & 1 \end{bmatrix}. \tag{19}$$

Then the dealer's normalized value function $W(x)$ satisfies the equation:

$$\begin{aligned} W(x) \;&=\; k_0(\ell_1^2 + \ell_2^2) - \frac{\gamma \sigma_v^2 L_0}{2}\big(n_1(1-b) + n_2(1+b) + 2(1+ab)\big), \quad \textit{with} \\ n_1 &= \Big(4\gamma^{-1} k_0^{-1} \sigma_v^{-2}(1-a) + (1-a)^2\Big)^{\frac{1}{2}}, \\ n_2 &= \Big(4\gamma^{-1} k_0^{-1} \sigma_v^{-2}(1+a) + (1+a)^2\Big)^{\frac{1}{2}}. \end{aligned} \tag{20}$$

When γ approaches zero or infinity, respectively, the dealer's normalized value function satisfies the following asymptotic equations:

$$\begin{aligned} W(x) \;&=\; k_0(\ell_1^2 + \ell_2^2) - \gamma^{\frac{1}{2}} L_0 \sigma_v k_0^{-\frac{1}{2}} \Big((1-a)^{\frac{1}{2}}(1-b) + (1+a)^{\frac{1}{2}}(1+b)\Big) + O(\gamma^{\frac{3}{2}}), \\ W(x) \;&=\; k_0(\ell_1^2 + \ell_2^2) - 2\gamma L_0 \sigma_v^2 (1+ab) - \frac{2L_0}{k_0} + O\big(\gamma^{-1}\big). \end{aligned} \tag{21}$$

Consider the case in which the liquidity demand in the two assets is positively correlated, i.e., $b > 0$, and small. If the dealer is almost risk neutral, the term than dominates the value function is proportional to $-(1-a)^{1/2}(1-b) - (1+a)^{1/2}(1+b)$, which is a U-shaped function in a and attains local maximum at the extremes, $a = \pm 1$, with the value attained at $a = -1$ larger than the value attained at $a = 1$ (with equality when

$b = 0$). Thus, a dealer with small risk aversion prefers assets that are highly correlated, either positively or negatively, since there is more inventory cross-hedging in that case.

By contrast, if the dealer is very risk-averse, i.e., if γ is very high, the term that dominates the value function is proportional to $-(1+ab)$, which is decreasing in a. This is because a very risk-averse dealer does not do much cross-hedging to begin with, and instead is much more concerned about how correlated the inventories in the two assets are as a result of liquidity demand. This behavior is similar to the asset choice problem in mean-variance optimization, where investors prefer assets with correlation as low as possible.

In general, numerical results show that the dealer's value function is U-shaped, with the minimum at a positive fundamental correlation a_0. Thus, the solution of the dealer's choice problem depends on the various paratemers of the model and on what correlations are available to the dealer.

2.4 Inventory Half-Life and Risk Aversion

In this section, we analyze how the dealer's inventory evolves over time and how this evolution is related to the dealer's risk aversion. To get simpler solutions, we consider the particular case with $N = 2$ assets. As in in Corollary 2, we assume that the fundamental volatility is the same for both asets and that the elasticity matrix is proportional to the identity matrix.

We define the inventory half-life as the expected time it takes for the inventory in one asset to be reduced by half. Corollary 4 provide an approximate formula for the inventory half-life when the dealer's risk aversion coefficient is small.

Corollary 4. *Let $N = 2$ and suppose that:*

$$\Sigma_v \;=\; \sigma_v^2 \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}, \quad k = \begin{bmatrix} k_0 & 0 \\ 0 & k_0 \end{bmatrix}. \tag{22}$$

Let $m_1 = \left(4\gamma^{-1}k_0^{-1}\sigma_v^{-2}(1-a)^{-1}+1\right)^{1/2}$ and $m_2 = \left(4\gamma^{-1}k_0^{-1}\sigma_v^{-2}(1+a)^{-1}+1\right)^{1/2}$. Then,

for any $t \in \mathbb{N}$*:*

$$\left(I_2 - k_0\lambda\right)^t \;=\; \begin{bmatrix} \frac{1}{2}\left(\frac{m_2-1}{m_2+1}\right)^t + \frac{1}{2}\left(\frac{m_1-1}{m_1+1}\right)^t & \frac{1}{2}\left(\frac{m_2-1}{m_2+1}\right)^t - \frac{1}{2}\left(\frac{m_1-1}{m_1+1}\right)^t \\ \frac{1}{2}\left(\frac{m_2-1}{m_2+1}\right)^t - \frac{1}{2}\left(\frac{m_1-1}{m_1+1}\right)^t & \frac{1}{2}\left(\frac{m_2-1}{m_2+1}\right)^t + \frac{1}{2}\left(\frac{m_1-1}{m_1+1}\right)^t \end{bmatrix}, \tag{23}$$

and the inventory half-life T *starting from an inventory* $x_0 = \begin{bmatrix}1, 0\end{bmatrix}'$ *solves the equation:*

$$\left(\frac{m_1 - 1}{m_1 + 1}\right)^T + \left(\frac{m_2 - 1}{m_2 + 1}\right)^T \;=\; 1. \tag{24}$$

When γ *is small, the inventory half-life in asset* 1 *is approximately:*

$$T \;\approx\; \frac{m}{(1+a)^{\frac{1}{2}} + (1-a)^{\frac{1}{2}}}, \quad \textit{with} \quad m \;=\; \gamma^{-\frac{1}{2}} k_0^{-\frac{1}{2}} \sigma_v^{-1}. \tag{25}$$

When assets 1 and 2 are uncorrelated ($a = 0$), equation (25) implies that the inventory half-life $T_{a=0}$ is approximately $m/2$. When assets 1 and 2 are perfectly correlated ($a = 1$), the inventory half-life is $T_{a=1} = m/\sqrt{2} = T_{a=0} \times \sqrt{2}$, which is 41% larger than the inventory half-life in the uncorrelated case. Intuitively, with correlated assets it takes longer to reduce inventory to zero, as the other asset can be used for hedging and the need to reduce the inventory is effectively weaker.

Corollary 4 implies that market makers may appear significantly less risk-averse than they actually are. Indeed, suppose we observe a dealer's inventory half-life of T. Then, using equation (25) we can estimate the dealer's risk aversion coefficient under the assumption that $a = 0$ or $a = 1$:

$$\gamma_{a=0} \;=\; \frac{k_0^{-1}\sigma_v^{-2}}{4T^2}, \qquad \gamma_{a=1} \;=\; \frac{k_0^{-1}\sigma_v^{-2}}{2T^2}. \tag{26}$$

Thus, if we wrongly assumed that the assets were uncorrelated when in fact they were perfectly correlated, we would be wrong by a factor of $\gamma_{a=0}/\gamma_{a=1} = 1/2$, implying that we would underestimate the dealer's risk aversion by 50%.

2.5 Stationary Equilibria

We define a stationary equilibrium as an equilibrium for which the price covariance matrix is constant over time. One can show that there is a unique stationary equilibrium and that any equilibrium converges to the unique stationary equilibrium.

Proposition 2 describes the evolution of the dealer's inventory and price over time and provides formulas for various covariances in a stationary equilibrium.

Proposition 2. *In equilibrium, the dealer's inventory and price satisfy the following formulas:*

$$\begin{aligned} x_{t+1} &= (I_N - k\lambda)x_t - \varepsilon_t, \\ \Delta p_{t+1} &= \lambda\varepsilon_t + \Delta v_{t+1} + \lambda k\lambda x_t, \end{aligned} \tag{27}$$

where $\varepsilon_t = \varepsilon_t^b - \varepsilon_t^s$. *Suppose, moreover, that* k, Σ_v, *and* Σ_L *commute with each other. Then, in a stationary equilibrium, the inventory variance* $\Sigma_x = \mathsf{Var}(x_t)$, *price variance* $\Sigma_p = \mathsf{Var}(\Delta p_t)$, *and price autocovariance* $A_p = \mathsf{Cov}(\Delta p_{t+1}, \Delta p_{t+2})$ *are constant and satisfy:*

$$\begin{aligned} \Sigma_x &= k^{-1}\lambda^{-1}(2I_N - k\lambda)^{-1}\Sigma_L, \\ \Sigma_p &= 2\lambda^2(2I_N - k\lambda)^{-1}\Sigma_L + \Sigma_v, \\ A_p &= -k\lambda^3(2I_N - k\lambda)^{-1}\Sigma_L. \end{aligned} \tag{28}$$

If $\left|(\gamma k\Sigma_v)^{-1}\right| < 1/4$, *we have:*

$$\begin{aligned} \lambda &\approx k^{-1}, \\ \Sigma_x &\approx \Sigma_L, \\ \Sigma_p &\approx 2k^{-2}\Sigma_L + \Sigma_v, \\ A_p &\approx -k^{-2}\Sigma_L. \end{aligned} \tag{29}$$

If $\left|\gamma k \Sigma_v\right| < 4$*, we have:*

$$\begin{aligned}
\lambda &\approx \gamma^{\frac{1}{2}} k^{-\frac{1}{2}} \Sigma_v^{\frac{1}{2}}, \\
\Sigma_x &\approx \frac{1}{2} \left(\gamma k \Sigma_v\right)^{-\frac{1}{2}} \Sigma_L, \\
\Sigma_p &\approx \left(\gamma k^{-1} \Sigma_L + I_N\right) \Sigma_v, \\
A_p &\approx -\frac{1}{2} \gamma^{\frac{3}{2}} k^{-\frac{1}{2}} \Sigma_L \Sigma_v^{\frac{3}{2}}.
\end{aligned} \tag{30}$$

Corollary 5 focuses on the simpler case in which there are only $N = 2$ assets and provides closed-form results. As in Corollary 2, we assume that the fundamental volatility is the same for both asets and that the elasticity matrix is proportional to the identity matrix.

Corollary 5. *Let* $N = 2$ *and suppose that:*

$$\Sigma_v = \sigma_v^2 \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}, \quad k = \begin{bmatrix} k_0 & 0 \\ 0 & k_0 \end{bmatrix}, \quad \Sigma_L = \begin{bmatrix} L_0 & 0 \\ 0 & L_0 \end{bmatrix}. \tag{31}$$

Then, in a stationary equilibrium, the inventory variance $\Sigma_x = \mathsf{Var}(x_t)$*, price variance* $\Sigma_p = \mathsf{Var}(\Delta p_t)$*, price autocovariance* $A_p = \mathsf{Cov}(\Delta p_{t+1}, \Delta p_{t+2})$*, and price correlation*

$\rho_p = \mathsf{Corr}(\Delta p_{1,t}, \Delta p_{2,t})$ *are constant and satisfy:*

$$
\begin{aligned}
\Sigma_x &= \frac{L_0}{8m_1m_2}\begin{bmatrix} m_1^2m_2 + m_1m_2^2 + 4m_1m_2 + m_1 + m_2 & (m_1 - m_2)(1 - m_1m_2) \\ (m_1 - m_2)(1 - m_1m_2) & m_1^2m_2 + m_1m_2^2 + 4m_1m_2 + m_1 + m_2 \end{bmatrix}, \\
\Sigma_p &= \frac{2L_0\begin{bmatrix} m_1^2 + m_2^2 + m_1 + m_2 & (m_1 - m_2)(m_1 + m_2 + 1) \\ (m_1 - m_2)(m_1 + m_2 + 1) & m_1^2 + m_2^2 + m_1 + m_2 \end{bmatrix}}{k_0^2 m_1 m_2 (m_1 + 1)(m_2 + 1)} + \sigma_v^2 \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}, \\
A_p &= \frac{2L_0\begin{bmatrix} m_1^3 + m_2^3 + 2m_1^2 + 2m_2^2 + m_1 + m_2 & (m_1 - m_2)(m_1^2 + m_2^2 + 2m_1 + 2m_2 + m_1m_2 + 1 \\ (m_1 - m_2)(m_1^2 + m_2^2 + 2m_1 + 2m_2 + m_1m_2 + 1) & m_1^3 + m_2^3 + 2m_1^2 + 2m_2^2 + m_1 + m_2 \end{bmatrix}}{k_0^2 m_1 m_2 (m_1 + 1)^2 (m_2 + 1)^2} \\
\rho_p &= \frac{\frac{(m_1 - m_2)(m_1 + m_2 + 1)}{m_1 m_2 (m_1 + 1)(m_2 + 1)} + a\nu_0}{\frac{m_1^2 + m_2^2 + m_1 + m_2}{m_1 m_2 (m_1 + 1)(m_2 + 1)} + \nu_0}, \quad \textit{with} \quad \nu_0 = \frac{k_0^2 \sigma_v^2}{2L_0}, \\
m_1 &= \left(4\gamma^{-1}k_0^{-1}\sigma_v^{-2}(1 - a)^{-1} + 1\right)^{\frac{1}{2}}, \quad m_2 = \left(4\gamma^{-1}k_0^{-1}\sigma_v^{-2}(1 + a)^{-1} + 1\right)^{\frac{1}{2}}.
\end{aligned}
\tag{32}
$$

If $L_0 \gg k_0\sigma_v^2 \gg 1/\gamma$*, then* m_1 *is close to* m_2*,* ν_0 *is close to zero, and the price correlation* ρ_p *between assets 1 and 2 is close to zero.*

Thus, when the demand noise coefficient L_0 and the dealer's risk aversion γ are sufficiently high, the asset return correlation is low at higher frequencies. Indeed, a more risk-averse dealer does less inventory hedging across assets, and thus quotes are less correlated across assets. If by assumption the noise demand is also uncorrelated across assets, then the price correlation between assets is also low. In our model therefore we obtain a theoretical version of the Epps effect. Epps (1979) finds that stock returns are relatively uncorrelated at high frequencies (less than 10 minutes) but approach a constant correlation at low frequencies (more than 1 day). In our model, the inventory effect in prices induces a mean-reverting component at higher frequencies that disappears at lower frequencies, for which the random walk component of prices dominates. At low frequencies, the return correlation approaches the fundamental correlation, which is assumed constant over time.

3 Empirical Evidence

We test Facts 1 through 5 using proprietary message-level data from the Toronto Stock Exchange. We focus on stocks with concentrated market-making, so that the monopolist dealer assumption underlying the model is most plausible.

3.1 Data and Market Maker Identification

The data consist of the complete STAMP Internal Feed from the TSX for January 13, 2006, containing 8,210,098 timestamped messages. We restrict to Regular Trading Hours (09:30–16:00 Eastern), during which we observe 337,303 trade executions totaling \$6.54 billion across 3,169 symbols. A `UserId` field on each message allows reconstruction of the complete intraday inventory path for every participant.

We define an active universe of 722 symbols meeting minimum liquidity thresholds ($\geq$50 daily trades, $\geq$\$50,000 daily volume, valid NBBO for $\geq$80% of the session, median spread $\leq$500 bps). Within this universe we identify two market maker populations. Designated Market Makers (DMMs) are broker-dealers with formal specialist obligations for assigned symbols, identified via `AccountType=ST` and `RTAutofill` flags; this yields 254 users across 1,832 (user, symbol) pairs. Endogenous Market Makers (EMMs) are voluntary two-sided quoters identified from quoting behavior, yielding 2,603 pairs (Set B) and a dual-role population (Set C) of 116 DMM users who also voluntarily quote non-assigned symbols.

For each (market maker, symbol) pair with at least five trades, a *shock* is a trade whose volume exceeds the pair-specific mean plus 2σ. A shock is *involuntary* if it creates a new position problem (a buy shock when pre-shock inventory is non-positive). For each shock we compute the sell ratio of the dealer's booked limit orders in 300-second pre- and post-shock windows: sell ratio $= \$_{\text{sell}}/(\$_{\text{buy}} + \$_{\text{sell}})$. The *sell ratio shift* $\Delta = \text{sell ratio}_{\text{post}} - \text{sell ratio}_{\text{pre}}$ is positive when the dealer tilts toward selling after a buy shock, which is the correct direction for inventory management.

Table 1 reports key characteristics of both populations. DMMs have a median trade-to-quote ratio of 2.20 (they execute more than they quote, consistent with absorbing flow against resting specialist quotes) and a median inventory half-life of 7.6 trades.

EMMs have a ratio of 0.63 and a half-life of 10.8 trades.

Table 1: Summary Statistics for Market Maker Populations

	DMM	**EMM (Set B)**
Panel A: Assigned stocks		
(User, symbol) pairs	1,832	2,603
Unique users	254	329
Booked orders per pair (median)	5	62
Trade-to-quote ratio (median)	2.20	0.63
Quote buy fraction (median)	0.50	0.50
Inventory half-life (trades, median)	7.6	10.8
Zero crossings per pair (median)	2	1
Mean |inventory| (shares, median)	1,313	3,056
Panel B: Full portfolio (qualified + non-assigned symbols)		
Total quoted symbols (median)	9.5	18.0
Two-sided quoted symbols (median)	6.0	10.0
User HHI (median)	4,472	5,471

Notes: Median values across (user, symbol) pairs during Regular Trading Hours, January 13, 2006. DMM pairs use `AccountType=ST` and `RTAutofill` flags. EMM Set B requires two-sided quoting, $\geq$1,000 shares booked, $\geq$5 trades, quote balance in $[0.25, 0.75]$, and quoting span $\geq$60 minutes. Inventory half-life is estimated from an AR(1) on the cumulative signed inventory path. User HHI is the Herfindahl index of booked volume across quoted symbols; HHI = 10,000 for a user active in one symbol.

3.2 Fact 1: Cross-Hedging Exists

After a buy shock in the shocked stock, we measure the sell ratio shift in four universes: (i) the shocked stock itself (own-stock hedging, λ_{11}); (ii) other stocks where the user is a qualified market maker (cross-stock hedging, λ_{12}); (iii) other stocks where the user posts two-sided quotes but does not meet the qualification threshold; and (iv) stocks with one-sided quoting only (a negative control).

Table 2 reports the dollar-weighted sell ratio shift at 300 seconds. Every population shows a significantly positive shift in the "Other qualified" column. The effect ranges from +0.019 ($t = 4.4$) for EMMs to +0.184 ($t = 7.2$) for DMMs. Three features establish that the cross-stock response reflects deliberate portfolio rebalancing rather than contemporaneous correlated order flow. First, the shift is indistinguishable from zero at 60 seconds and builds over 300 seconds; correlated flow would appear immediately. Second, involuntary shocks generate 25–50% larger responses than active unwinds, consistent with the shock being a genuine position problem rather than a choice. Third,

the cross-stock shift appears only in new passive limit orders, not in aggressive trades, which show momentum (more buying, not selling) after a buy shock. The operative mechanism is quote adjustment, precisely as the theory predicts.

Table 2: Fact 1: Cross-Stock Sell Ratio Shift After Buy Shocks (300s, Dollar-Weighted)

		Same stock		**Other qualified**	
Population	N	Δ	t	Δ	t
Panel A: DMMs					
All buy shocks	3,810	+.413	15.0	+.184	7.2
Involuntary shocks	1,950	+.498	18.1	+.221	6.9
Panel B: EMMs (Set B)					
All buy shocks	7,809	+.173	22.6	+.021	4.3
Involuntary shocks	2,977	+.253	17.2	+.021	2.0

Notes: Dollar-weighted sell ratio shift ($\Delta = \text{post}_{300s} - \text{pre}_{300s}$) after buy shocks (volume exceeding pair-specific mean + 2σ). "Same" = shocked stock. "Other qualified" = other stocks where the user meets the population's identification threshold. The DMM design uses a three-universe partition (same stock, other assigned, non-assigned); the EMM design uses a four-universe partition. Involuntary shocks are those where pre-shock inventory ≤ 0 (buy shock) or ≥ 0 (sell shock).

3.3 Fact 2: Cross-Hedging Increases with Correlation

For each inventory shock, we compute the average portfolio correlation: the mean pairwise monthly return correlation (estimated over 36 months of Compustat data, January 2003 to December 2005) between the shocked stock and all other stocks where the user is a qualified market maker. Table 3 reports OLS regressions of the cross-stock sell ratio shift on average portfolio correlation.

Table 3: Fact 2: Cross-Stock Shift Increases with Portfolio Correlation

	DMM		**EMM (Set B)**	
	β	t	β	t
Avg. portfolio correlation	+0.055	0.49	+0.098*	1.81
Involuntary shocks only	+0.141	1.62	+0.203***	3.21

Notes: OLS regression of the dollar-weighted cross-stock sell ratio shift (at 300s, buy shocks) on average monthly portfolio correlation. All specifications include MM-type fixed effects where applicable and cluster standard errors by user. $^{***}p < 0.01$, $^{*}p < 0.10$. The pooled-population coefficient is +0.084 ($t = 2.40$); for involuntary shocks it is +0.203 ($t = 3.21$, $N = 1{,}667$). Coverage is 53.7% of (shock, other-stock) pairs after matching to Compustat.

The coefficient on portfolio correlation is positive in both populations, and is signif-

icant for EMMs and in the involuntary-shock specification. A one-standard-deviation increase in average portfolio correlation (approximately 0.15) raises the cross-stock shift by about 29% of its unconditional mean. The stronger effect for involuntary shocks, where the shock is exogenous to the dealer's portfolio choice, provides the clearest identification.

3.4 Fact 3: Own-Hedging Substitutes Cross-Hedging

Table 4 reports OLS regressions of the own-stock sell ratio shift on average portfolio correlation, the same specification as Fact 2 but with the dependent variable changed to the same-stock shift. The theory predicts a negative coefficient.

Table 4: Fact 3: Own-Hedging Decreases with Portfolio Correlation (DMMs)

	β	t
Panel A: DMMs (assignment correlation more plausibly exogenous)		
All buy shocks	−0.155	−1.96
Involuntary shocks only	−0.192	−1.86
Panel B: EMMs (voluntary portfolio – self-selection confound)		
All buy shocks	+0.175	+2.39
Involuntary shocks only	+0.214	+2.17

Notes: OLS regression of the dollar-weighted own-stock sell ratio shift (at 300s, buy shocks) on average monthly portfolio correlation. Standard errors clustered by user. For DMMs, portfolio composition is determined by exchange assignment and is approximately exogenous to the within-day quoting response. For EMMs, the positive coefficient reflects self-selection: high-correlation EMMs are the most active market makers on all dimensions (more quotes per hour, shorter half-life, more two-sided symbols), so the correlation and the quoting intensity are jointly driven by an underlying activity factor rather than by the substitution mechanism in the theory.

DMMs show the predicted substitution effect ($\beta = -0.155$, $p = 0.05$): when the assigned portfolio is more correlated with the shocked stock, own-stock quote adjustment is reduced. EMMs show the opposite sign, because they self-select into correlated stocks: EMMs in the highest portfolio-correlation quintile post significantly more quotes per hour ($t = 4.3$), have shorter inventory half-lives ($t = -2.8$), and quote in more two-sided symbols ($t = 5.1$) than those in the lowest quintile. The substitution effect is present where the identification is clean (DMMs with exogenous assignment) and is obscured where portfolio composition is endogenous (EMMs).

3.5 Fact 4: Risk Aversion and Cross-Hedging

We proxy for dealer risk aversion using the inventory half-life (shorter half-life indicates faster mean-reversion, consistent with higher γ), zero crossings per pair, inverse user size (total trades), and a composite index that standardizes and averages all three. Table 5 reports the cross-stock sell ratio shift across terciles of each proxy.

Table 5: Fact 4: Cross-Stock Shift by Risk Aversion Proxy

	DMM		**EMM (Set B)**	
Risk aversion proxy	T3−T1	t	T3−T1	t
Half-life (short = high RA)	-0.059^{***}	-7.0	-0.041^{***}	-5.1
Zero crossings (more = high RA)	$+0.052^{***}$	$+5.4$	$+0.029^{***}$	$+3.8$
Inverse user size (small = high RA)	$+0.085^{***}$	$+9.2$	$+0.061^{***}$	$+7.1$
Composite RA index	$+0.089^{***}$	$+9.8$	$+0.063^{***}$	$+7.4$

Notes: T3−T1 is the difference in mean cross-stock shift (other qualified, 300s, buy shocks) between the top and bottom tercile of each risk aversion proxy. $^{***}p < 0.01$. The theory predicts $\text{T3} - \text{T1} < 0$ for all proxies. The half-life proxy has the correct sign: faster mean-reversion (higher RA) is associated with more cross-hedging. The zero-crossing and size proxies have the opposite sign and are confounded by a level effect (more active traders do more of everything) and by the dilution effect (sparse quoters mechanically show larger measured sell ratio shifts).

The half-life proxy yields the correct sign: dealers with shorter half-lives (who mean-revert faster, consistent with higher γ) show larger cross-stock shifts. The remaining proxies are wrong-signed, because they conflate risk aversion with trading activity: larger and more active market makers do more cross-hedging at every level of activity, and this level effect dominates. The half-life is the cleanest proxy because it measures mean-reversion speed in trade-units, which is not mechanically related to quote count or institution size. With the half-life proxy, Fact 4 is supported in sign, though not in the compositional dimension (the cross/same ratio rises rather than falls with RA, suggesting that active managers allocate more rather than less of their total response to cross-stock channels). Separating the risk preference channel from the activity-level channel requires instruments for γ beyond what a single trading day provides.

3.6 Fact 5: The Epps Effect

We compute VWAP return correlations at four intraday frequencies (5, 15, 30, and 60 minutes) for all active-universe stock pairs from the STAMP data, and at daily and

monthly frequencies using 36 months of Compustat returns. The mean correlation rises monotonically with horizon: from 0.005 at 5-minute intervals to 0.046 at the monthly frequency, consistent with the Epps effect. The intraday estimates are noisy on a single trading day (standard error ≈ 0.11 for any pair with 77 five-minute returns), and the cross-frequency rank correlation between five-minute and daily pairwise correlations is $r = -0.007$, confirming that pair-level intraday correlations on one day are dominated by estimation noise. A definitive cross-sectional test of whether the Epps effect is increasing in aggregate dealer risk aversion requires multiple days of trade data. The daily/monthly Epps ratio is dominated by fundamental correlation ($t = 75.5$), and the market-making variables add little explanatory power for this aggregate effect.

4 Conclusion

We propose a dynamic model of market making in multiple assets and derive five testable predictions about the structure of cross-hedging. The pricing matrix λ summarizes how the dealer uses inventory in one asset to set quotes in correlated assets. Its off-diagonal entries are positive and increasing in fundamental correlation (Facts 1 and 2), while its diagonal entries decrease as better substitutes become available (Fact 3). High dealer risk aversion renders λ approximately diagonal, reducing cross-hedging but generating a term structure of return correlations that resembles the Epps effect (Facts 4 and 5). Ignoring this multi-asset structure when estimating risk aversion from single-asset inventory half-lives leads to downward bias of up to 50%.

Empirically, Facts 1 through 3 are supported by event-study evidence from the Toronto Stock Exchange. Cross-hedging is pervasive, builds over minutes (not seconds), operates through passive limit orders rather than aggressive trades, and is stronger for involuntary shocks. It increases with the fundamental correlation of the dealer's portfolio, and own-stock hedging decreases with correlation for designated market makers whose portfolio composition is more plausibly exogenous. The prediction that more risk-averse dealers cross-hedge less is supported in sign by the inventory half-life proxy, though the level effect of trading activity makes it difficult to isolate the pure risk preference channel. The Epps effect exists in the data but its cross-sectional link to dealer

risk aversion requires multiple days of intraday data to identify reliably.

REFERENCES

Amihud, Y. and H. Mendelson (1980). "Dealership market: Market-making with inventory." *Journal of Financial Economics*, 8, 31–53.

Brogaard, J., T. Hendershott, and R. Riordan (2014). "High-frequency trading and price discovery." *Review of Financial Studies*, 27, 2267–2306.

Copeland, T. E. and D. Galai (1983). "Information effects on the bid-ask spread." *Journal of Finance*, 38, 1457–1469.

Daures-Lescourret, L. and S. Moinas (2023). "Fragmentation and strategic market-making." *Journal of Financial and Quantitative Analysis*, 58(4), 1675–1700.

de Jong, F. and T. Nijman (1997). "High frequency analysis of lead-lag relationships between financial markets." *Journal of Empirical Finance*, 4, 259–277.

Di Maggio, M., A. Kermani, and Z. Song (2017). "The value of trading relationships in turbulent times." *Journal of Financial Economics*, 124, 468–493.

Epps, T. W. (1979). "Comovements in stock prices in the very short run." *Journal of the American Statistical Association*, 74(366a), 291–298.

Hendershott, T. and A. J. Menkveld (2014). "Price pressures." *Journal of Financial Economics*, 114(3), 405–423.

Ho, T. and H. R. Stoll (1981). "Optimal dealer pricing under transactions and return uncertainty." *Journal of Financial Economics*, 9, 47–73.

Hugonnier, J., B. Lester, and P.-O. Weill (2020). "Frictional intermediation in over-the-counter markets." *Review of Economic Studies*, 87, 1750–1796.

Ljungqvist, L. and T. J. Sargent (2000). *Recursive Macroeconomic Theory*. MIT Press, Cambridge, MA.

Menkveld, A. J. (2013). "High frequency trading and the new market makers." *Journal of Financial Markets*, 16, 712–740.

Roll, R. (1984). "A simple implicit measure of the effective bid-ask spread in an efficient market." *Journal of Finance*, 39, 1127–1139.

Roşu, I., E. Sojli, and W. W. Tham (2021). "Quoting activity and the cost of capital." *Journal of Financial and Quantitative Analysis*, 56(8), 2764–2799.

Scholes, M. and J. Williams (1977). "Estimating betas from nonsynchronous data." *Journal of Financial Economics*, 5, 309–327.

Stoll, H. R. (1978). "The supply of dealer services in securities markets." *Journal of Finance*, 33, 1133–1151.

Virtu Financial, Inc. (2014). "Form S-1 registration statement." U.S. Securities and Exchange Commission. Available at `https://ir.virtu.com/node/9491/html`.

Appendix A. Proofs

Proof of Proposition 1. Denote by $\mathcal{I}_t$ the dealer's information just before trading at t and by E_t the expectation operator conditional on $\mathcal{I}_t$. The dealer's strategy is a set of quotes (a_t, b_t) adapted to the filtration $\mathcal{I}_t$ for $t \in \mathbb{N}$, where a_t and b_t are $N \times 1$ vectors. Equivalently, the dealer chooses (h_t, δ_t), where h_t is the (quoted) half-spread and δ_t is the price discount:

$$h_t \;=\; \frac{a_t - b_t}{2}, \quad \text{and} \quad \delta_t \;=\; v_t - \frac{a_t + b_t}{2}. \tag{A1}$$

We show that $h_t = h$ and $\delta_t = \lambda x_t$, where h and λ are defined in equation (5).

Equation (3) implies that traders' buy and sell demands at t are given, respectively, by $Q_t^b = (k/2)(v_t - a_t) + \ell + \varepsilon_t^b$ and $Q_t^s = (k/2)(b_t - v_t) + \ell + \varepsilon_t^s$, with $\varepsilon_t^b, \varepsilon_t^s \sim \mathcal{N}(0, \Sigma_L/2)$. Define η_t to be the net demand noise at t:

$$\eta_t \;=\; \varepsilon_t^s - \varepsilon_t^b, \tag{A2}$$

which is uncorrelated with the past information and has a multivariate normal distribution $\mathcal{N}(0, \Sigma_L)$. If x_t is the dealer's inventory before trading at t, equation (1) implies that x_t evolves according to $x_{t+1} = x_t - Q_t^b + Q_t^s$, which translates into:

$$x_{t+1} \;=\; x_t - k\delta_t + \eta_t \quad \text{with} \quad \eta_t \overset{IID}{\sim} \mathcal{N}\big(0, \Sigma_L\big). \tag{A3}$$

Substituting Q_t^b and Q_t^s in the dealer's objective function from (2), we get:

$$\mathsf{E}\sum_{t=0}^{\infty} \beta^t\, \mathsf{E}_t\Big(-(a_t - v_t)'\frac{k}{2}(a_t - v_t) - (v_t - b_t)'\frac{k}{2}(v_t - b_t) + 2\ell' h_t - \gamma x_t' \Sigma_v x_t\Big). \tag{A4}$$

With the notation in equation (A1), we compute $\mathsf{E}_t(v_t - a_t)'k(v_t - a_t) = (\delta_t - h_t)'k(\delta_t - h_t)$ and $\mathsf{E}_t(v_t - b_t)'k(v_t - b_t) = (\delta_t + h_t)'k(\delta_t + h_t)$. Thus, the dealer solves:

$$\max_{(h_t, \delta_t)_{t\geq 0}} \; \mathsf{E}\sum_{t=0}^{\infty} \beta^t \Big(-\delta_t' k \delta_t - h_t' k h_t + 2\ell' h_t - \gamma x_t' \Sigma_v x_t\Big), \tag{A5}$$

where the inventory x_t evolves according to equation (A3). Using the Bellman principle of optimization, we reduce the dynamic optimization in (A5) to finding a value function

V that solves the static optimization problem:

$$V(x_t) \;=\; \max_{h_t,\delta_t} \Big(-\delta_t' k \delta_t - h_t' k h_t + 2\ell' h_t - \gamma x_t' \Sigma_v x_t + \beta\, \mathsf{E}_t\, V(x_{t+1}) \Big), \tag{A6}$$

We guess that $V : \mathbb{R}^N \to \mathbb{R}^N$ is a quadratic function of the form:

$$V(x) \;=\; W_0 - x' W_2 x, \tag{A7}$$

where W_0 is a scalar and W_2 is a symmetric $N \times N$ matrix. If A and B are two $N \times N$ matrices, denote their Frobenius inner product (or double dot product) by $A \circ B = \sum_{i,j=1}^{N} A_{ij} B_{ij}$. A direct calculation shows that $\mathsf{E}_t\big(\eta_t' W_2 \eta_t\big) = W_2 \circ \Sigma_L$. Substituting x_{t+1} from equation (A3), the problem becomes:

$$\begin{aligned} V(x_t) \;=\; \max_{h_t,\delta_t} \Big(& -\delta_t' k \delta_t - h_t' k h_t + 2\ell' h_t - \gamma x_t' \Sigma_v x_t + \beta W_0 \\ & - \beta x_t' W_2 x_t - \beta \delta_t' k W_2 k \delta_t + 2\beta \delta_t' k W_2 x_t - \beta W_2 \circ \Sigma_L \Big). \end{aligned} \tag{A8}$$

The first order condition with respect to h_t implies that the equilibrium half-spread satisfies $h_t = k^{-1}\ell$, which is the same as the constant h in the first part of equation (5).

The first order condition with respect to δ_t is $-2k\delta_t - 2\beta k W_2 k \delta_t + 2\beta k W_2 x_t \;=\; 0$, which implies that the equilibrium price discount is:

$$\delta_t \;=\; \lambda x_t, \quad \text{with} \quad \lambda \;=\; (k + \beta k W_2 k)^{-1} \beta k W_2. \tag{A9}$$

The second order condition with respect to δ_t is that the symmetric matrix $I_N + \beta W_2 k$ is positive definite, which we denote with an inequality sign:

$$I_N + \beta W_2 k \;>\; 0. \tag{A10}$$

Combining equations (A7) and (A8) and identifying the quadratic coefficients (of the form $x_t'(...)x_t$), it follows that W_2 satisfies:

$$W_2 \;=\; \gamma \Sigma_v + \beta W_2 - \beta^2 W_2 k (k + \beta k W_2 k)^{-1} k W_2, \tag{A11}$$

or, equivalently, after multiplying by k on both sides:

$$\gamma k\Sigma_v - \frac{1-\beta}{\beta} w_2 \;=\; w_2(I_N + w_2)^{-1} w_2, \quad \text{with} \quad w_2 \;=\; \beta k W_2. \tag{A12}$$

Denote by:

$$Z \;=\; (I_N + w_2)^{-1} w_2. \tag{A13}$$

Simple matrix manipulation shows that $w_2 = (I_N - Z)^{-1} Z$. Multiplying equation (A14) by $(I_N - Z)$ on both sides, we obtain:

$$(I_N - Z)S - \frac{1-\beta}{\beta} Z \;=\; Z^2, \quad \text{with} \quad S \;=\; \gamma k \Sigma_v. \tag{A14}$$

This is equivalent to $Z^2 + Z\big(S + \frac{1-\beta}{\beta} I_N\big) = S$, which is the same equation as (6). Following the existence and uniqueness of the square root of a symmetric positive definite matrix, it is straightforward to show that this equation has a unique solution in symmetric positive definite matrices. Thus, $Z = \Lambda$. Moreover, equations (A9) and (A13) imply that:

$$\lambda \;=\; (k + \beta k W_2 k)^{-1} \beta k W_2 \;=\; k^{-1}\Lambda. \tag{A15}$$

This proves that $\lambda = k^{-1}\Lambda$, which is the second part of equation (5).

For future use, we identify the constant terms (independent of x_t) in equation (A8), and we obtain:

$$W_0 \;=\; \frac{1}{1-\beta}\Big(\ell' k^{-1} \ell - \beta W_2 \circ \Sigma_L\Big). \tag{A16}$$

When β converges to 1, equation (6) becomes $\Lambda^2 + \Lambda S = S$. This equation has a unique symmetric positive definite solution:

$$\Lambda \;=\; \frac{1}{2}\Big((S^2 + 4S)^{1/2} - S\Big). \tag{A17}$$

This proves equation (7). □

Proposition A.1. *Suppose the dealer's objective in equation (2) is modified by replacing the inventory penalty*

$$\gamma x_t' \Sigma_v x_t$$

with

$$\gamma x'_{t+1}\Sigma_v x_{t+1}.$$

Then, for any fixed $\beta \in (0,1)$, the dealer's dynamic optimization problem is equivalent to the original problem in Proposition 1, with the same primitives except that the risk-aversion coefficient γ is replaced by γ/β. Consequently, the equilibrium quotes remain of the form

$$a_t = v_t - \lambda x_t + h, \qquad b_t = v_t - \lambda x_t - h,$$

with

$$h = k^{-1}\ell, \qquad \lambda = k^{-1}\Lambda,$$

where Λ is the unique symmetric positive definite solution of

$$\Lambda^2 + \Lambda\Big(\frac{\gamma}{\beta}k\Sigma_v + \frac{1-\beta}{\beta}I_N\Big) = \frac{\gamma}{\beta}k\Sigma_v.$$

In particular, when $\beta \to 1$, the limiting equilibrium is identical to the one in Proposition 1, namely

$$\lambda = \frac{k^{-1}}{2}\Big(\big(S^2+4S\big)^{1/2} - S\Big), \qquad S = \gamma k\Sigma_v.$$

Proof. Let Π_t denote the dealer's period-t expected profit term,

$$\Pi_t \;=\; (v_t - b_t)'Q_t^s + (a_t - v_t)'Q_t^b.$$

Consider the modified objective

$$\mathsf{E}_\tau \sum_{t=\tau}^{\infty} \beta^{\,t-\tau}\Big(\Pi_t - \gamma x'_{t+1}\Sigma_v x_{t+1}\Big). \tag{A18}$$

Fix the state x_τ. Re-indexing the penalty term yields

$$\sum_{t=\tau}^{\infty} \beta^{\,t-\tau} x'_{t+1}\Sigma_v x_{t+1} = \sum_{s=\tau+1}^{\infty} \beta^{\,s-1-\tau} x'_s\Sigma_v x_s = \frac{1}{\beta}\sum_{s=\tau+1}^{\infty} \beta^{\,s-\tau} x'_s\Sigma_v x_s.$$

Hence

$$\sum_{t=\tau}^{\infty} \beta^{\,t-\tau} x'_{t+1}\Sigma_v x_{t+1} = \frac{1}{\beta}\sum_{s=\tau}^{\infty} \beta^{\,s-\tau} x'_s\Sigma_v x_s - \frac{1}{\beta}x'_\tau\Sigma_v x_\tau.$$

Substituting into (A18), we obtain

$$\mathsf{E}_\tau \sum_{t=\tau}^{\infty} \beta^{t-\tau}\Big(\Pi_t - \gamma x_{t+1}' \Sigma_v x_{t+1}\Big) = \mathsf{E}_\tau \sum_{t=\tau}^{\infty} \beta^{t-\tau}\Big(\Pi_t - \frac{\gamma}{\beta} x_t' \Sigma_v x_t\Big) + \frac{\gamma}{\beta} x_\tau' \Sigma_v x_\tau.$$

The last term depends only on the current state x_τ and not on the dealer's choice variables at dates $t \geq \tau$. Therefore, for every fixed $\beta \in (0,1)$, maximizing (A18) is equivalent to maximizing the original objective (2) with γ replaced by γ/β.

Proposition 1 then applies directly, with

$$S = \frac{\gamma}{\beta} k \Sigma_v.$$

Thus, the optimal quotes are again linear,

$$a_t = v_t - \lambda x_t + h, \qquad b_t = v_t - \lambda x_t - h,$$

with

$$h = k^{-1}\ell, \qquad \lambda = k^{-1}\Lambda,$$

where Λ solves

$$\Lambda^2 + \Lambda\Big(\frac{\gamma}{\beta} k \Sigma_v + \frac{1-\beta}{\beta} I_N\Big) = \frac{\gamma}{\beta} k \Sigma_v.$$

Finally, when $\beta \to 1$, we have $(\gamma/\beta) k \Sigma_v \to \gamma k \Sigma_v$, so the limiting matrix equation becomes

$$\Lambda^2 + \Lambda S = S, \qquad S = \gamma k \Sigma_v,$$

which is exactly the limit equation in Proposition 1. Therefore the limiting equilibrium when $\beta \to 1$ is the same under the two specifications of the inventory penalty. □

Proof of Corollary 1. First, assume that $\left|S^{-1}\right| < 1/4$. The Taylor series for $(1+x)^{1/2}$ is $1 + x/2 - x^2/8 + \cdots$ and converges when $|x| < 1$. A similar analysis for matrices shows that the Taylor series for $(1 + 4S^{-1})^{1/2}$ is $1 + 2S^{-1} - 2S^{-2} + \cdots$ and converges when $|S^{-1}| < 1/4$. Equation (7) implies that the matrix $\Lambda = k\lambda$ satisfies:

$$\Lambda \;=\; \frac{S}{2}\Big(\big(I_N + 4S^{-1}\big)^{1/2} - I_N\Big) \;=\; \frac{S}{2}\Big(2S^{-1} - 2S^{-2} + S^{-1}O\big(S^{-2}\big)\Big). \tag{A19}$$

Thus, $\Lambda = I_N - S^{-1} + O\big(S^{-2}\big) = I_N + O\big(S^{-1}\big)$, which proves equation (9).

Now, assume that $|S| < 4$. Rewriting equation (7), we obtain that:

$$\begin{aligned} 2\Lambda \;&=\; \Big(4S\Big(1+\frac{S}{4}\Big)\Big)^{1/2} - S \;=\; 2S^{1/2}\Big(1+\frac{S}{8}+O\big(S^2\big)\Big) - S \\ &=\; 2S^{1/2}\Big(1 - S^{1/2} + \frac{S}{8} + O\big(S^2\big)\Big). \end{aligned} \tag{A20}$$

Thus, $\Lambda = S^{1/2} + O(S)$, which proves equation (10). □

Proof of Corollary 2. Let

$$P \;=\; \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Then

$$P'\Sigma_v P = \sigma_v^2 \,\mathrm{diag}(1+a, 1-a), \qquad P'kP = k_0 I_2.$$

Hence, with $S = \gamma k \Sigma_v$, we obtain

$$P'SP = \gamma k_0 \sigma_v^2 \,\mathrm{diag}(1+a, 1-a).$$

By equation (7), the matrix $\Lambda = k\lambda$ is given by

$$\Lambda = \frac{1}{2}\Big((S^2+4S)^{1/2} - S\Big).$$

Since S is diagonalized by P, so is Λ. If we denote by

$$s_+ = \gamma k_0 \sigma_v^2 (1+a), \qquad s_- = \gamma k_0 \sigma_v^2 (1-a),$$

its two eigenvalues, then the corresponding eigenvalues of Λ are

$$\mu_+ = \frac{1}{2}\big(\sqrt{s_+^2 + 4s_+} - s_+\big), \qquad \mu_- = \frac{1}{2}\big(\sqrt{s_-^2 + 4s_-} - s_-\big).$$

Using the identity

$$\frac{1}{2}\big(\sqrt{s^2+4s} - s\big) = \frac{2}{1+\sqrt{1+4/s}},$$

we obtain

$$\mu_+ = \frac{2}{m_2+1}, \qquad \mu_- = \frac{2}{m_1+1},$$

where

$$m_1 = \left(1 + \frac{4}{\gamma k_0 \sigma_v^2 (1-a)}\right)^{1/2}, \qquad m_2 = \left(1 + \frac{4}{\gamma k_0 \sigma_v^2 (1+a)}\right)^{1/2}.$$

Therefore,

$$P'\Lambda P = \operatorname{diag}\left(\frac{2}{m_2+1}, \frac{2}{m_1+1}\right),$$

and rotating back gives

$$\Lambda = P \operatorname{diag}\left(\frac{2}{m_2+1}, \frac{2}{m_1+1}\right) P' = \frac{1}{(m_1+1)(m_2+1)} \begin{bmatrix} m_1+m_2+2 & m_1-m_2 \\ m_1-m_2 & m_1+m_2+2 \end{bmatrix}.$$

Since $\lambda = k^{-1}\Lambda = k_0^{-1}\Lambda$, equation (12) follows. □

Proof of Corollary 3. Equation (A7) implies that for every $x \in \mathbb{R}^2$,

$$\lim_{\beta \to 1} (1-\beta) V(x) \;=\; \ell' k^{-1} \ell - \alpha \circ \Sigma_L, \tag{A21}$$

where α satisfies $(k^{-1}+\alpha)^{-1} = \gamma \alpha^{-1} \Sigma_v \alpha^{-1}$. Let

$$P \;=\; \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Then $P'\Sigma_v P = \sigma_v^2 \operatorname{diag}(1+a, 1-a)$ and $P'\Sigma_L P = L_0 \operatorname{diag}(1+b, 1-b)$. As $k = k_0 I_2$, the matrix

$$A \;=\; \gamma^{-1} \Sigma_v^{-1} k^{-1}$$

is diagonalized by the same matrix P and satisfies

$$P'AP = \operatorname{diag}\left(\frac{1}{\gamma k_0 \sigma_v^2 (1+a)}, \frac{1}{\gamma k_0 \sigma_v^2 (1-a)}\right).$$

If $X \;=\; \gamma^{-1} \Sigma_v^{-1} \alpha$, then

$$X^2 - X - A = 0,$$

so that $X = \big(I_2 + (I_2 + 4A)^{1/2}\big)/2$. Hence X is diagonalized by P, with eigenvalues

$$x_+ = \frac{1}{2}\left(1 + \sqrt{1 + \frac{4}{\gamma k_0 \sigma_v^2 (1+a)}}\right), \qquad x_- = \frac{1}{2}\left(1 + \sqrt{1 + \frac{4}{\gamma k_0 \sigma_v^2 (1-a)}}\right).$$

Using the notation of the corollary,

$$n_2 = (1+a)\sqrt{1 + \frac{4}{\gamma k_0 \sigma_v^2 (1+a)}}, \qquad n_1 = (1-a)\sqrt{1 + \frac{4}{\gamma k_0 \sigma_v^2 (1-a)}},$$

so that

$$\gamma\sigma_v^2(1+a)x_+ = \frac{\gamma\sigma_v^2}{2}\big((1+a) + n_2\big), \qquad \gamma\sigma_v^2(1-a)x_- = \frac{\gamma\sigma_v^2}{2}\big((1-a) + n_1\big).$$

Therefore,

$$P'\alpha P = \frac{\gamma\sigma_v^2}{2}\,\mathrm{diag}\big((1+a) + n_2, (1-a) + n_1\big).$$

Since the Frobenius product is invariant under orthogonal changes of basis,

$$\begin{aligned}
\alpha \circ \Sigma_L &= (P'\alpha P) \circ (P'\Sigma_L P) \\
&= \frac{\gamma\sigma_v^2 L_0}{2}\Big(((1+a) + n_2)(1+b) + ((1-a) + n_1)(1-b)\Big) \\
&= \frac{\gamma\sigma_v^2 L_0}{2}\Big(n_1(1-b) + n_2(1+b) + 2(1+ab)\Big).
\end{aligned}$$

Substituting this identity into equation (A21) proves equation (20).

For the asymptotics as $\gamma \to 0$, we have

$$n_1 = 2\gamma^{-\frac{1}{2}} k_0^{-\frac{1}{2}} \sigma_v^{-1} (1-a)^{\frac{1}{2}} + O(\gamma^{\frac{1}{2}}), \qquad n_2 = 2\gamma^{-\frac{1}{2}} k_0^{-\frac{1}{2}} \sigma_v^{-1} (1+a)^{\frac{1}{2}} + O(\gamma^{\frac{1}{2}}),$$

which substituted into equation (20) yields the first line of equation (21). For the asymptotics as $\gamma \to \infty$, we use

$$n_1 = (1-a) + \frac{2}{\gamma k_0 \sigma_v^2} + O(\gamma^{-2}), \qquad n_2 = (1+a) + \frac{2}{\gamma k_0 \sigma_v^2} + O(\gamma^{-2}),$$

which substituted into equation (20) yields the second line of equation (21). □

Proof of Corollary 4. Let

$$P \;=\; \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

By the proof of Corollary 2, the matrix $k_0\lambda$ is diagonalized by P and satisfies

$$P'(k_0\lambda)P = \operatorname{diag}\left(\frac{2}{m_2+1}, \frac{2}{m_1+1}\right).$$

Therefore,

$$P'(I_2 - k_0\lambda)P = \operatorname{diag}\left(\frac{m_2-1}{m_2+1}, \frac{m_1-1}{m_1+1}\right).$$

Taking the t-th power and rotating back, we obtain

$$\begin{aligned}(I_2 - k_0\lambda)^t &= P \operatorname{diag}\left(\left(\frac{m_2-1}{m_2+1}\right)^t, \left(\frac{m_1-1}{m_1+1}\right)^t\right) P' \\ &= \begin{bmatrix} \frac{1}{2}\left(\frac{m_2-1}{m_2+1}\right)^t + \frac{1}{2}\left(\frac{m_1-1}{m_1+1}\right)^t & \frac{1}{2}\left(\frac{m_2-1}{m_2+1}\right)^t - \frac{1}{2}\left(\frac{m_1-1}{m_1+1}\right)^t \\ \frac{1}{2}\left(\frac{m_2-1}{m_2+1}\right)^t - \frac{1}{2}\left(\frac{m_1-1}{m_1+1}\right)^t & \frac{1}{2}\left(\frac{m_2-1}{m_2+1}\right)^t + \frac{1}{2}\left(\frac{m_1-1}{m_1+1}\right)^t \end{bmatrix},\end{aligned}$$

which proves the first displayed formula.

If $x_0 = [1, 0]'$, then the expected inventory after t periods is $(I_2 - k_0\lambda)^t x_0$, so the expected inventory in asset 1 is reduced by half exactly when

$$\frac{1}{2}\left(\frac{m_2-1}{m_2+1}\right)^T + \frac{1}{2}\left(\frac{m_1-1}{m_1+1}\right)^T = \frac{1}{2},$$

which is equivalent to

$$\left(\frac{m_1-1}{m_1+1}\right)^T + \left(\frac{m_2-1}{m_2+1}\right)^T = 1.$$

This proves the equation that characterizes the inventory half-life.

Finally, suppose that γ is small and let $m = \gamma^{-1/2} k_0^{-1/2} \sigma_v^{-1}$. Then

$$m_1 = 2m(1-a)^{-1/2} + O(m^{-1}), \qquad m_2 = 2m(1+a)^{-1/2} + O(m^{-1}).$$

Moreover,

$$\log\left(\frac{m_i-1}{m_i+1}\right) = -\frac{2}{m_i} + O(m_i^{-3}), \qquad i = 1, 2.$$

Hence the half-life equation becomes

$$\exp\left(-\frac{2T}{m_1} + O\left(\frac{T}{m_1^3}\right)\right) + \exp\left(-\frac{2T}{m_2} + O\left(\frac{T}{m_2^3}\right)\right) = 1.$$

Keeping the first-order terms in $1/m_1$ and $1/m_2$, we obtain

$$1 - \frac{2T}{m_1} + 1 - \frac{2T}{m_2} \simeq 1,$$

so that

$$T \simeq \frac{1}{2\left(m_1^{-1} + m_2^{-1}\right)} = \frac{m}{(1+a)^{1/2} + (1-a)^{1/2}}.$$

This proves equation (25). □

Proof of Proposition 2. Proposition 1 implies that $p_t = v_t - \lambda x_t$, hence $v_t - p_t = \lambda x_t$. Also, $Q_t^b - Q_t^s = k(v_t - p_t) + \varepsilon_t = k\lambda x_t + \varepsilon_t$. As the dealer's inventory changes by the amount of the net order flow, $x_{t+1} = x_t - (Q_t^b - Q_t^s) = x_t - k\lambda x_t - \varepsilon_t$, which proves the first part of equation (27).

Denote by $Y = k\lambda$. At $t+1$, the equation $p_t = v_t - \lambda x_t$ becomes $p_{t+1} = v_{t+1} - \lambda x_{t+1}$, hence $\Delta p_{t+1} = \Delta v_{t+1} - \lambda(x_{t+1} - x_t)$. As $x_{t+1} - x_t = -Y x_t - \varepsilon_t$, we obtain $\Delta p_{t+1} = \Delta v_{t+1} + \lambda\varepsilon_t + \lambda Y x_t$, which proves the second part of equation (27).

In a stationary equilibrium, the inventory variance remains constant, i.e., $\mathsf{Var}(x_{t+1}) = \mathsf{Var}(x_t) = \Sigma_x$. By assumption, k and Σ_v commute with each other, hence equation (**??**) implies that k and λ also commute with each other (they are also both symmetric). The equation $x_{t+1} = (I_N - Y)x_t - \varepsilon_t$ implies that $\Sigma_x = (I_N - Y)\Sigma_x(I_N - Y) + \Sigma_L$. Hence, Σ_x satisfies the linear matrix equation $Y\Sigma_x + \Sigma_x Y - Y\Sigma_x Y = \Sigma_L$. If k and Σ_v also commute with Σ_L, Σ_x satisfies the equation $Y(2I_N - Y)\Sigma_x = \Sigma_L$, hence $\Sigma_x = Y^{-1}(2I_N - Y)^{-1}\Sigma_L$, which proves the first part of equation (28).

The price increment at $t+1$ is $\Delta p_{t+1} = \Delta v_{t+1} + \lambda\varepsilon_t + \lambda Y x_t$, hence $\Sigma_p = \Sigma_v + \lambda\Sigma_L\lambda +$

$\lambda Y \Sigma_x Y \lambda$. Using the formula $Y\Sigma_x Y = Y\Sigma_x + \Sigma_x Y - \Sigma_L$, we compute:

$$\Sigma_p \;=\; \Sigma_v + \lambda(Y\Sigma_x + \Sigma_x Y)\lambda, \quad \text{with} \quad Y \;=\; k\lambda. \tag{A22}$$

If Σ_L commutes with k and Σ_v, equation (A22) becomes $\Sigma_p = \Sigma_v + 2k\lambda^3\Sigma_x = \Sigma_v + 2\lambda^2(2I_N - k\lambda)^{-1}\Sigma_L$, which proves the second part of equation (28).

The price increment at $t+2$ is $\Delta p_{t+2} = \Delta v_{t+2} + \lambda\varepsilon_{t+1} + \lambda Y x_{t+1} = \Delta v_{t+2} + \lambda\varepsilon_{t+1} + \lambda Y(I_N - Y)x_t - \lambda Y\varepsilon_t$. Thus, the autocovariance $A_p = \mathsf{E}\big((\Delta p_{t+2})\,(\Delta p_{t+1})'\big) = \lambda Y(I_N - Y)\Sigma_x Y\lambda - \lambda Y\Sigma_L\lambda = \lambda Y\big((I_N - Y)\Sigma_x Y - \Sigma_L\big)\lambda$. Using the equation $\Sigma_x Y - Y\Sigma_x Y = \Sigma_L - Y\Sigma_x$, we compute $A_p = -\lambda Y^2\Sigma_x\lambda$. If Σ_L commutes with k and Σ_v, we compute $A_p = -k^2\lambda^4\Sigma_x = -k\lambda^3(2I_N - k\lambda)^{-1}\Sigma_L$, which proves the third part of equation (28).

□

Proof of Corollary 5. Let

$$P \;=\; \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

By Corollary 2, the matrix λ is diagonalized by P and satisfies

$$P'\lambda P = \operatorname{diag}\left(\frac{2}{k_0(m_2+1)}, \frac{2}{k_0(m_1+1)}\right).$$

Since $\Sigma_L = L_0 I_2$ and $\Sigma_v = \sigma_v^2\begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}$, we also have

$$P'\Sigma_L P = L_0 I_2, \qquad P'\Sigma_v P = \operatorname{diag}\big(\sigma_v^2(1+a), \sigma_v^2(1-a)\big).$$

We now apply Proposition 2 in this diagonal basis.

First,

$$P'\Sigma_x P = k_0^{-1}(P'\lambda P)^{-1}(2I_2 - k_0 P'\lambda P)^{-1}L_0 I_2 = \operatorname{diag}\left(\frac{L_0(m_2+1)^2}{4m_2}, \frac{L_0(m_1+1)^2}{4m_1}\right).$$

Rotating back gives

$$\Sigma_x = P\,\mathrm{diag}\left(\frac{L_0(m_2+1)^2}{4m_2}, \frac{L_0(m_1+1)^2}{4m_1}\right)P'$$
$$= \frac{L_0}{8m_1m_2}\begin{bmatrix} m_1^2m_2+m_1m_2^2+4m_1m_2+m_1+m_2 & (m_1-m_2)(1-m_1m_2) \\ (m_1-m_2)(1-m_1m_2) & m_1^2m_2+m_1m_2^2+4m_1m_2+m_1+m_2 \end{bmatrix}.$$

This proves the formula for Σ_x.

Second, Proposition 2 implies

$$P'(\Sigma_p-\Sigma_v)P = 2(P'\lambda P)^2(2I_2-k_0P'\lambda P)^{-1}L_0I_2 = \mathrm{diag}\left(\frac{4L_0}{k_0^2m_2(m_2+1)}, \frac{4L_0}{k_0^2m_1(m_1+1)}\right).$$

Hence

$$\Sigma_p = P\,\mathrm{diag}\left(\frac{4L_0}{k_0^2m_2(m_2+1)}, \frac{4L_0}{k_0^2m_1(m_1+1)}\right)P' + \sigma_v^2\begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}$$
$$= \frac{2L_0}{k_0^2m_1m_2(m_1+1)(m_2+1)}\begin{bmatrix} m_1^2+m_2^2+m_1+m_2 & (m_1-m_2)(m_1+m_2+1) \\ (m_1-m_2)(m_1+m_2+1) & m_1^2+m_2^2+m_1+m_2 \end{bmatrix}$$
$$+ \sigma_v^2\begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix}.$$

This proves the formula for Σ_p.

Third,

$$P'A_pP = -k_0(P'\lambda P)^3(2I_2-k_0P'\lambda P)^{-1}L_0I_2 = -\,\mathrm{diag}\left(\frac{4L_0}{k_0^2m_2(m_2+1)^2}, \frac{4L_0}{k_0^2m_1(m_1+1)^2}\right).$$

Therefore,

$$A_p = -P\,\mathrm{diag}\left(\frac{4L_0}{k_0^2m_2(m_2+1)^2}, \frac{4L_0}{k_0^2m_1(m_1+1)^2}\right)P'$$
$$= -\frac{2L_0}{k_0^2m_1m_2(m_1+1)^2(m_2+1)^2}\begin{bmatrix} m_1^3+m_2^3+2m_1^2+2m_2^2+m_1+m_2 & (m_1-m_2)(m_1^2+ \\ (m_1-m_2)(m_1^2+m_2^2+2m_1+2m_2+m_1m_2+1) & m_1^3+m_2^3 \end{bmatrix}$$

This proves the formula for A_p.

Finally, the matrix Σ_p has the form $\begin{bmatrix} u & v \\ v & u \end{bmatrix}$, hence $\rho_p = v/u$. Using the formula for Σ_p above, we get

$$u = \frac{2L_0}{k_0^2} \frac{m_1^2 + m_2^2 + m_1 + m_2}{m_1 m_2 (m_1+1)(m_2+1)} + \sigma_v^2,$$

$$v = \frac{2L_0}{k_0^2} \frac{(m_1 - m_2)(m_1 + m_2 + 1)}{m_1 m_2 (m_1+1)(m_2+1)} + a\sigma_v^2.$$

Dividing numerator and denominator by $2L_0/k_0^2$ and using $\nu_0 = k_0^2\sigma_v^2/(2L_0)$ proves the stated formula for ρ_p. □

Proof of Proposition ??. Proposition 1 implies that $p_t = (a_t + b_t)/2 = w_t - \lambda x_t$, hence $Q_t^b - Q_t^s = k(v_t - p_t) + \varepsilon_t = k(v_t - w_t) + k\lambda x_t + \varepsilon_t = k\lambda x_t + ke_t + \varepsilon_t$. Note that $x_{t+1} = x_t - (Q_t^b - Q_t^s)$, hence $x_t \in \mathcal{I}_t$. Thus, observing the net order flow at t is equivalent to receiving a signal $s_t = ke_t + \varepsilon_t$, where both e_t and ε_t are orthogonal to $\mathcal{I}_t$. We compute:

$$\begin{aligned} w_{t+1} \;&=\; \mathsf{E}(v_t|\mathcal{I}_t) \;=\; w_t + \mathsf{E}(e_t \mid ke_t + \varepsilon_t) \;=\; w_t + c_t(ke_t + \varepsilon_t), \quad \text{with} \\ c_t \;&=\; \mathsf{Cov}(e_t, ke_t + \varepsilon_t)\,\mathsf{Var}(ke_t + \varepsilon_t)^{-1} \;=\; \Sigma_{e,t}k\big(k\Sigma_{e,t}k + \Sigma_L\big)^{-1}, \end{aligned} \tag{A23}$$

which proves the first part of (**??**). As $e_t = v_t - w_t$, we have $\Delta e_{t+1} = \Delta v_{t+1} - \Delta w_{t+1} = \Delta v_{t+1} - c_t(ke_t + \varepsilon_t)$, hence:

$$e_{t+1} \;=\; (I_N - c_t k)e_t - c_t\varepsilon_t + \Delta v_{t+1}. \tag{A24}$$

The terms e_t, ε_t, and Δv_{t+1} are uncorrelated, hence by taking the covariance of e_{t+1} we obtain:

$$\Sigma_{e,t+1} \;=\; (I_N - c_t k)\Sigma_{e,t}(I_N - c_t k) + c_t\Sigma_L c_t + \Sigma_v. \tag{A25}$$

Expanding the terms and using the formula for c_t in (A23), we obtain the second part of (**??**). □

Proof of Corollary ??. By definition, in a stationary equilibrium we have $\Sigma_{e,t+1} = \Sigma_{e,t}$. Hence, equation (**??**) implies that $\Sigma_{w,t}$ is constant and equal to Σ_v. From (**??**), we

have $\Sigma_w = \Sigma_e k(k\Sigma_e k + \Sigma_L)^{-1} k\Sigma_e$, hence:

$$\Sigma_v \;=\; \Sigma_e k(k\Sigma_e k + \Sigma_L)^{-1} k\Sigma_e. \tag{A26}$$

Inverting equation (A26), we obtain $k\Sigma_e\Sigma_v^{-1}\Sigma_e k = k\Sigma_e k + \Sigma_L$, which proves the formula for Σ_e in equation (**??**). Also, equation (**??**) implies $c = \Sigma_e k\big(k\Sigma_e k + \Sigma_L\big)^{-1}$, which proves the formula for c in equation (**??**). Thus, equation (A26) implies $c = \Sigma_v\Sigma_e^{-1}k^{-1}$, which finishes the proof of equation (**??**). Next, equation (**??**) follows directly from equations (1) and (A24).

Equation (8) implies that the price change at $t+1$ satisfies $\Delta p_{t+1} = \Delta w_{t+1} - \lambda\Delta x_t$ (we note the different time subscript notation for $\Delta x_t = x_{t+1} - x_t$). As $\varepsilon_t = \varepsilon_t^b - \varepsilon_t^s$, equation (A3) implies:

$$\Delta x_t \;=\; -k\lambda x_t - ke_t - \varepsilon_t, \tag{A27}$$

where we use the formula $\delta_t = \lambda x_t$ from equation (**??**). According to equation (**??**), the forecast change at $t+1$ satisfies $\Delta w_{t+1} = c_t(ke_t + \varepsilon_t)$. Thus, we obtain:

$$\Delta p_{t+1} \;=\; (c+\lambda)(ke_t + \varepsilon_t) + \lambda k\lambda x_t, \tag{A28}$$

which proves the first part of equation (**??**). Equation (A27) implies that $ke_t + \varepsilon_t = -\Delta x_t - k\lambda x_t$, which we substitute in equation (A28) to prove the second part of equation (**??**). This completes the proof.

□

Appendix B. Expected Profit under Liquidation

We consider the case in which there is a single risky asset, i.e., $N = 1$. For simplicity, we assume that the bid-ask spread is zero and trading in each period takes place at the same price. At each time $t \in \mathbb{N}$, the asset liquidates with a probability $\pi \in (0,1)$. If the asset does not liquidate at t, the investor trades $\Delta x_t = x_{t+1} - x_t$ at a price p_t. If the asset liquidates at t, the investor receives a cash amount of v_t for each of the x_t shares

and thus realizes the following profit:[13]

$$\Pi_t \;=\; x_t v_t - \sum_{s=0}^{t-1} p_s \Delta x_s \;=\; x_0 v_t + \sum_{s=0}^{t-1} (v_s - p_s)\Delta x_s. \tag{B1}$$

Define $\beta \in (0,1)$ by:

$$\beta \;=\; 1 - \pi. \tag{B2}$$

The investor's expected profit at $t = 0$ is:[14]

$$\begin{aligned} \Pi \;&=\; \mathsf{E}_0\Big(\pi\Pi_0 + (1-\pi)\pi\Pi_1 + \cdots + (1-\pi)^t \pi \Pi_t + \cdots\Big) \\ &=\; \mathsf{E}_0\Big(\sum_{t=0}^{\infty} (1-\beta)\beta^t \Pi_t\Big) \;=\; \mathsf{E}_0\Big(\sum_{t=0}^{\infty} \beta^t \Pi_t - \sum_{t=0}^{\infty} \beta^{t+1}\Pi_t\Big) \\ &=\; \mathsf{E}_0\Big(\Pi_0 + \sum_{t=0}^{\infty} \beta^{t+1}(\Pi_{t+1} - \Pi_t)\Big), \end{aligned} \tag{B3}$$

Equation (B3) shows that the expected profit can be considered as the expectation of a sum of discounted instantaneous profits: $\Delta\Pi_{t+1} = \Pi_{t+1} - \Pi_t$. From (B1), the instananeous profit at t satisfies:

$$\Delta\Pi_{t+1} \;=\; x_{t+1}v_{t+1} - x_t v_t - p_t \Delta x_t \;=\; x_{t+1}\Delta v_{t+1} + (v_t - p_t)\Delta x_t. \tag{B4}$$

Thus, the update in realized profit at $t+1$ has two sources: an inventory component, $x_t \Delta v_{t+1}$, and a trading component, $(v_{t+1} - p_t)\Delta x_t$. The inventory component measures the reevaluation of the existing inventory, and can be thought as a stochastic dividend paid at $t+1$. The trading component reflects the cash payment made for trading at t, but with a gain estimated at the value v_t.

We model risk aversion by having the dealer incur each period a penalty proportional to the variance of the instantanteous profit. Thus, the dealer's expected utility is of the form:

$$\sum_{t=0}^{\infty} \beta^t \Big(\mathsf{E}_t\big(\Delta\Pi_{t+1}\big) - \gamma\, \mathsf{Var}_t\big(\Delta\Pi_{t+1}\big)\Big), \tag{B5}$$

[13] We use the convention that a sum from a larger to a smaller index is zero, e.g., $\sum_{t=0}^{-1} p_t \Delta x_t = 0$.

[14] Note that the discount factor at t is β^{t+1}. Indeed, by convention, trading starts at $t = 0$, and even though the first instantaneous profit occurs at zero, it must be still discounted by β. This can be reconciled easily if the time before trading is denoted by $t = -1$.

where γ is the dealer's risk aversion coefficient. As in equations (A1) and (A2), we define:

$$h_t \;=\; \frac{a_t - b_t}{2}, \quad p_t \;=\; \frac{a_t + b_t}{2}, \quad \delta_t \;=\; v_t - p_t, \quad \eta_t \;=\; \varepsilon_t^s - \varepsilon_t^b. \tag{B6}$$

Equations (1) and (3) imply that:

$$\begin{aligned} \Delta x_t \;&=\; k\Big(\frac{a_t + b_t}{2} - v_t\Big) - \varepsilon_t^b + \varepsilon_t^s \\ &=\; -k\delta_t + \eta_t, \\ \Delta\Pi_{t+1} \;&=\; x_{t+1}\Delta v_{t+1} - (v_t - a_t)Q_t^b + (v_t - b_t)Q_t^s \\ &=\; x_{t+1}\Delta v_{t+1} - k\delta_t^2 - kh_t^2 + 2h_t\ell - (\delta_t + h_t)\varepsilon_t^s + (\delta_t - h_t)\varepsilon_t^b. \end{aligned} \tag{B7}$$

At the optimum (see the proof of Proposition 1), we have:

$$\begin{aligned} \delta_t \;&=\; \lambda x_t, \quad \text{with} \quad \lambda \;=\; \frac{1}{2}\Big(-\gamma + \sqrt{\gamma^2 + \frac{4\gamma}{k}}\Big), \\ h_t \;&=\; h \;=\; \frac{\ell}{k}, \end{aligned} \tag{B8}$$

where λ is computed in the limit case when $\beta = 1$. Equation (B7) becomes:

$$\begin{aligned} x_{t+1} \;&=\; (1 - k\lambda)x_t + \eta_t, \\ \Delta\Pi_{t+1} \;&=\; x_{t+1}\Delta v_{t+1} - k\lambda^2 x_t^2 + \frac{\ell^2}{k} - \lambda x_t\eta_t - \frac{\ell}{k}(\varepsilon_t^s + \varepsilon_t^b). \end{aligned} \tag{B9}$$

As x_t and v_t are in the dealer's information set at t, so are $a_t = v_t + h$, $b_t = v_t - h$, and $\delta_t = v_t - (a_t + b_t)/2$. Therefore, the variance of $\Delta\Pi_{t+1}$ conditional on the information at t is:

$$\mathsf{Var}_t(\Delta\Pi_{t+1}) \;=\; \mathsf{Var}_t\Big(\big((1 - k\lambda)x_t + \eta_t\big)\Delta v_{t+1} - \lambda x_t\eta_t - \frac{\ell}{k}(\varepsilon_t^s + \varepsilon_t^b)\Big). \tag{B10}$$

If we define:

$$\sigma_\eta^2 \;=\; \mathsf{Var}(\eta_t) \;=\; \mathsf{Var}(\varepsilon_t^s - \varepsilon_t^b) \quad \text{and} \quad \sigma_v^2 \;=\; \mathsf{Var}(\Delta v_{t+1}), \tag{B11}$$

we have:

$$\begin{aligned}
\mathsf{Var}\Big(\big((1-k\lambda)x_t+\eta_t\big)\Delta v_{t+1}\Big) &= \big((1-k\lambda)^2x_t^2+\sigma_\eta^2\big)\sigma_v^2,\\
\mathsf{Var}\big(\lambda x_t\eta_t\big) &= \lambda^2x_t^2\sigma_\eta^2,\\
\mathsf{Var}\big(h(\varepsilon_t^s+\varepsilon_t^b)\big) &= \frac{\ell^2}{k^2}\sigma_\eta^2.
\end{aligned} \tag{B12}$$

Equations (B10) and (B12) imply:

$$\mathsf{Var}_t(\Delta\Pi_{t+1}) = (1-k\lambda)^2x_t^2\sigma_v^2+\sigma_\eta^2\Big(\sigma_v^2+\lambda^2x_t^2+\frac{\ell^2}{k^2}\Big)+ \text{ correlation terms.} \tag{B13}$$

We consider the case in which $k\gamma$ is very small:

$$1 \gg k\gamma. \tag{B14}$$

We compute:

$$\begin{aligned}
k\lambda &= \frac{-k\gamma+\sqrt{(k\gamma)^2+4k\gamma}}{2} = \sqrt{k\gamma}+o(k\gamma) \approx \sqrt{k\gamma},\\
\lambda^2 &\approx \frac{\gamma}{k}.
\end{aligned} \tag{B15}$$

Equations (B13) and (B15) imply:

$$\mathsf{Var}_t(\Delta\Pi_{t+1}) = \big(1-\sqrt{k\gamma}\big)x_t^2\sigma_v^2+\sigma_\eta^2\Big(\sigma_v^2+\frac{\gamma}{k}x_t^2+\frac{\ell^2}{k^2}\Big)+ \text{ correlation terms.} \tag{B16}$$

Therefore, if γ and σ_η are sufficiently small, we obtain:

$$\mathsf{Var}_t(\Delta\Pi_{t+1}) \approx x_t^2\sigma_v^2. \tag{B17}$$

This shows that when the dealer's risk aversion is sufficiently low, the expected utility in (B5) can be approximated by:

$$\mathsf{E}_0\sum_{t=0}^{\infty}\beta^t\Big(\mathsf{E}_t\,\Delta\Pi_{t+1}-\gamma x_t^2\Big), \tag{B18}$$

where we have replaced $\gamma\sigma_v^2$ by γ. Thus, we can replace the quadratic utility term,

$\gamma\,\mathsf{Var}(\Delta\Pi_{t+1})$, by the quadratic inventory penalty, γx_t^2.