Agency costs of institutional trading

Roger M. Edelen*
Carroll School of Management
Boston College
Chestnut Hill, MA 02467
rmedelen@bc.edu

Gregory B. Kadlec
Pamplin College of Business
Virginia Tech
Blacksburg, VA 24060-0221
kadlec@vt.edu

Abstract

Under the typical institutional trading arrangement a portfolio manager makes the trade decision and a trading desk executes the trade, with execution performance evaluated against a backward-looking benchmark such as the volume weighted average price (VWAP) or implementation shortfall (Perold, 1988). This arrangement gives the trader an incentive to maintain a relatively low ask (high bid) quote when valuations rise (fall) to expedite sell (buy) trades if the portfolio manager grants quantity discretion. We show that granting quantity discretion is generally optimal, despite the associated agency cost, and that the process results in price-adjustment delays because it inhibits information assimilation. We provide empirical support for the theory with several previously undocumented cross sectional and time series facts about price-adjustment delays.

This Version: July 25, 2007
First Version: March 2005

*Corresponding author. This paper was supported in part by a 2005 summer research grant from the Pamplin College of Business. We wish to thank the following for their helpful comments: Brad Barber, Tarun Chordia, Katrina Ellis, Simon Gervais, Huseyin Gulen, Craig Holden, Ning Jin, Robert Kosowski, Marc Lipson, Vijay Singal, A. Subrahmanyam, Bilge Yilmaz, and Seminar participants at the 2007 WFA meetings, Virginia Tech, UC Davis, Indiana University and Boston College. Errors are solely the responsibility of the authors.
1. Introduction

At most investment firms, the portfolio manager delegates trade implementation. This division of labor between portfolio managers and institutional traders follows from the distinct skill sets involved with each task. Expertise in portfolio management involves, for example, analysis of accounting statements; product and service markets; and portfolio efficiency – all relatively low frequency considerations. By contrast, expertise in trading involves high frequency considerations, such as information signalling and leakage; parsing order flow characteristics; and assessing market conditions across many trading venues. This paper develops and tests a model of the agency conflict that arises when the portfolio manager’s trade implementation needs are delegated to a separate economic entity, the trading desk. Our model explains several key features of institutional trading arrangements and provides novel predictions regarding information assimilation in security markets.

Perhaps the most common method used by portfolio managers to evaluate traders’ performance is to compare trade execution prices to the volume weighted average price (VWAP) of all trades during the trading day. We argue that performance metrics such as this (i.e., those using pre-trade ‘strike’ prices as a benchmark) can give the trader incentives that are at odds with the objectives of the portfolio manager. Specifically, the trader maintains a low ask quote (relative to fair value) when valuations rise to expedite sell trades, and a high bid quote (relative to fair value) when valuations fall to expedite buy trades. This bias in trading intensity inhibits the process of information assimilation in financial markets, resulting in predictable patterns in market prices. In particular, once the trader’s aggressive quoting behavior stops with order completion, the stock’s price
subsequently rises (or falls), reflecting with a delay the shift in valuation that triggered
the trader’s biased trading pattern. Thus, widespread use of delegated trade
implementation by institutional portfolio managers can cause equilibrium price
adjustment delays as traders systematically react to a common market environment.

Traders are willing to pay more than fair value when buying against falling
valuations, and to accept less than fair value when selling against rising valuations, to the
extent that their performance is evaluated relative to a backward-looking benchmark (e.g.
VWAP or implementation shortfall). As a result, the trader locks in positive
“performance” when price-to-VWAP is favorable even though price to fair value is
unfavorable. Likewise, traders seek to only partially fill orders when selling (buying)
against falling (rising) valuations even though the price is attractive relative to fair value.
This game of “selectively executing those [trades] that are most favorable” is referenced

It is important to note that our model employs a specific quantification of trading
performance (i.e., various modifications of VWAP) only for purposes of tractability. While Trade-Cost Analysis (or “TCA” - an algorithmic assessment of trader performance
using realized transaction prices) is in fact a very common practice, it is not our view that
it is the sole, or even dominant, determinant of trader compensation in practice. Rather,
our premise is that the PM’s qualitative assessment of trader performance is correlated
with the rigid quantification provided by TCA. For example, is the PM satisfied that she
is getting filled at approximately the price she expected (which, presumably, relates to the
price at the time the trade blotter was created)? Many PM’s may feel that their qualitative
assessment is more exacting than the rigid quantification of a TCA algorithm. Nevertheless, the incentives perceived by the trader align with our model.

A key consideration in this game is that the PM grants discretion to the trader to partially fill the order. Indeed, the portfolio manager could always simply mandate full execution. However, doing so would leave the trader with limited degrees of freedom to minimize trading costs. Thus, the PM faces a trade-off between the agency cost of granting discretion, and the agency benefit of the trader’s expertise. Our model solves for the optimum level of discretion, minimizing expected trade-implementation costs.

The core prediction of our model is that traders will tend to execute the PM’s order counter to both systematic and idiosyncratic shifts in valuation, given discretion. That is, they buy more (sell less) with decreases in valuation and sell more (buy less) with increases in valuation. Lipson and Puckett (2006) provide direct empirical support for this prediction in the context of systematic shifts in valuation. They show that institutions are net sellers when market valuations rise and net buyers when market valuations fall. Moreover, they find that the increase in institutional selling in rising markets, and buying in falling markets, is due to trade implementation rather than position decisions. This finding supports our model’s contention that it is the trader, not the portfolio manager, who alters execution quantity to run counter to valuation shifts.

Our model also predicts that the contrarian nature of delegated trade execution inhibits the assimilation of information into security prices, causing price-adjustment delays. While other theories of price adjustment delays exist, our model is unique in that it predicts price adjustment delays with respect to idiosyncratic stock returns, as well as
making several novel predictions regarding cross-sectional and time series properties of price adjustment delays.

Our empirical analysis parallels this theory of price adjustment delays. We first document real price-adjustment delays with respect to equity-index futures returns for a large sample of stocks, using a methodology that precludes nonsynchronous trading effects. We then test the conditional predictions of the model, specifically that price-adjustment delays are: (1) positively related to a stock’s price-VWAP ratio, (2) negatively related to lag buy-sell order flow imbalances, (3) negatively related to lag trading volume, and (4) positively related to the time of day. Our findings confirm each of these predictions.

Finally, the model also predicts that the degree to which the PM grants the trader discretion, and the extent of agency induced price adjustment delays, is related to the liquidity of the stock. For the most liquid stocks we predict that the PM will impose a full-execution constraint on the trader and grant no execution discretion. But for relatively illiquid stocks, where the trader’s expertise can lead to larger improvements in trade-execution costs, the PM grants a relatively high degree of discretion. Thus, our model predicts little price adjustment delay in large-cap liquid stocks but relatively large price adjustment delays in small-cap illiquid stocks.¹

Our model analysis of institutional trading has a number of parallels with other studies of trading. First, the agency conflict in our model is similar to that proposed by Harris and Schultz (1998) to explain the viability of SOES bandit trading. Second,

¹ The observation that nondiscretionary algorithmic trading strategies are most common with large-cap stocks is consistent with this prediction. See, for example, http://www.ftmandate.com/news/printpage.php/aid/655/Strategy_mapping.html, http://advancedtrading.com/goldbook2006/bui_nhan.jhtml
because the price-adjustment delays of our model arise from demand for liquidity, our analysis is consistent with the recognition of technical traders (arbitrageurs in our model) as providers of liquidity as opposed to exploiters of market inefficiencies [Kavejevecz and Odders-White (2004)]. Finally, our model demonstrates another way in which market participants can impede the adjustment of prices to market information. For example, Hasbrouck and Sofianos (1993) show how specialists or dealers may impede the adjustment of prices because of exchange stabilization obligations or inventory imbalances. Similarly, Admati and Pfleiderer (1988) and Foster and Vishwanathan (1992) show how public limit orders and trading strategies may impede the adjustment of prices.

The remainder of our paper proceeds as follows. Section 2 develops the framework for our model of the agency conflict inherent in delegating trade implementation. Section 3 explores the implications of the model. Section 4 discusses the sample selection, data sources, and methodologies used to test the empirical predictions of our model. Section 4 presents the empirical results. Section 5 summarizes our findings.

2. Model Framework and Solution

2.1 Set-up and overview

There are four economic entities in the model; the trading public, market makers, portfolio managers, and institutional traders. Trading occurs in two discrete intervals encompassing a trading day; a morning auction and an afternoon auction. This assumption greatly simplifies the model’s development but no generality is lost in considering more frequent intervals or longer trading horizons.
Let $V$ denote the fair value of the security traded, which follows the process

$$V_t = V_{t-1} \left(1 + \tilde{R}_M^t \right) + \tilde{\psi}_t,$$  \hspace{1cm} (1)

where $\tilde{R}_M^t$ is the market return in interval $t$ and $\tilde{\psi}_t$ is a publicly observed, idiosyncratic shift in value with prior variance $\psi^{-1}$ (precision $\psi$). The four economic entities in the model are described below.

2.1.1 The trading public

Public order flow has both an informed and a noise component, taking the form

$$Q_t^F = -\frac{1}{\lambda} \left( P_t - E(V_{t+1} | \tilde{\psi}_{t+1}) \right) + \tilde{\nu}_t = -\frac{1}{\lambda} \left( P_t - V_t - \tilde{\psi}_{t+1} \right) + \tilde{\nu}_t$$  \hspace{1cm} (2)

where $\tilde{\nu}_t$ is normally distributed noise trading with variance $\nu^2$. Public order flow constitutes the only source of private information about $\tilde{\psi}_t$: all other entities in the model learn about $\tilde{\psi}_t$ either when it is announced at time $t$, or by inference via equilibrium prices at time $t - 1$. Inelasticity ($\lambda > 0$) in the informed component of $Q_t^F$ could be motivated by risk aversion or adverse selection concerns, but for ease of exposition we take $\lambda$ as exogenous.

2.1.2 Market maker

The second entity in the model is a market maker (MM) who only trades against perceived deviations between price and fair value, with elasticity $\lambda$. In doing so, the MM
expends the effort required to maintain and apply a technology (e.g., “technical analysis”) that partially identifies the nature of order flow. This technology yields the signal\(^2\)

\[ \tilde{U}_t = \tilde{v}_t + \tilde{n}_t \]  

(3)

where precision(\(\eta\)) = \(h\). Let \(S_t\) denote the signal of \(\tilde{\psi}_{t+1}\) that the technology yields:

\[ S_t = P_t - V_t + \lambda \left( Q_t^F - zU_t \right) = \tilde{\psi}_{t+1} + \lambda \left( \tilde{U}_t - z\tilde{U}_t \right) \]  

(4)

where \(z \equiv \frac{h}{u+h}\). Using \(S\), the MM forms the conditional expectation (Appendix A)

\[ E(\tilde{v}_{t+1}|\tilde{S}_t, \tilde{U}_t) = V_t + w\tilde{S}_t \]  

(5)

where \(w \equiv \frac{u+h}{\lambda^2 y + u + h}\), and demands

\[ Q_{t,MM} = \frac{-1}{\lambda} \left( P_t - E(\tilde{v}_{t+1}|...) \right) = \frac{-1}{\lambda} \left( P_t - (V_t + wS_t) \right). \]  

(6)

2.1.3 Portfolio manager

The third entity in the model is a portfolio manager (PM) who formulates demands, denoted \(Q_{t,PM}\), based on long-horizon considerations. Plausible motives include “fundamental” forecasts of value (i.e., forecasts of \(\tilde{Y}_{t+j}\), \(j >> 1\)); portfolio rebalancing (Nagle, 2006); investor flow (Edelen, 1999); or noise trading (Dow and Gorton, 2000). Because \(Q_{t,PM}\) is based on distant (if any) forecasts of \(\tilde{Y}_{t+j}\), neither the realization \(\tilde{Y}_{t+1}\) nor the order flow in the market at time \(t\), alters the PM’s demands. Thus, we take the quantity \(Q_{t,PM}\) to be exogenous to the time \(t\) market.

\(^2\) Correlated order flow suggests that the market return may also have a noise component, and that the market-maker signal’s could in fact provide information about the noise component to market return. In the interest of simplicity, we ignore this possibility.
This exogeneity assumption does not mean that the PM would necessarily want to have the quantity $Q^{PM}$ filled in period $t$ under any circumstance. For example, if the PM knew that the market price at time $t$ was adversely influenced by noise trading (e.g., temporarily pushed down in the case of a sell; temporarily pushed up in the case of a buy), then the PM might prefer to delay the order until a subsequent period when the pricing influence of noise traders is reversed, or at least not so severe. However, the PM cannot correctly interpret ‘market conditions’ without expending the time and effort required to develop and implement the MM’s technology. With limited information-processing capacity, this multi-tasking would undermine performance at the primary task of portfolio management. Hence, even though the PM would like to make $Q^{PM}$ endogenous to market conditions, as a practical matter she cannot.

Instead, the PM delegates the task of optimally reacting to the information in public order flow to an institutional trader – the fourth entity in the model. To the extent that the PM grants discretion to the trader to react to market conditions, this introduces an endogenous component to the PM’s demands, even though the PM’s instructions to the trader are exogenous.

2.1.4 Institutional trader

In delegating implementation to a trader, the PM’s objective is to achieve lower execution costs, which is feasible if the trader successfully implements the MM’s technology. The problem is, unlike the MM who profits from this activity, the trader has no natural incentive to expend the effort required in applying that technology because the
trader’s positions are dictated by the PM.\(^3\) Hence, the PM must compensate the trader in such a way that the trader chooses to acquire, and apply, the MM’s technology.

2.2. Price

The market clearing price satisfies

\[
Q^T_t + Q^{MM}_t + Q^F_t = 0
\]

(7)

where \(Q^T_t\) is the demands submitted by the trader in period \(t\). Using Eq. (2) and (6), market clearing implies

\[
\tilde{P}_t = V_{t-1} \left( 1 + \tilde{R}^M_t \right) + \tilde{\psi}_t + w \tilde{S}_t + \frac{1}{2} (\tilde{\psi}_{t+1} - w \tilde{S}_t) + \frac{\lambda}{2} \tilde{\nu}_t + \frac{\lambda}{2} Q^T_t.
\]

(8)

2.3. Trader incentives and monitoring by the portfolio manager

In filling \(Q^{PM}\), the PM would like the trader to react to the pricing distortions of noise trading. Since these are transient in nature, trading around them lowers execution costs. The PM cannot directly evaluate the trader’s performance on this dimension, however, because the PM cannot distinguish between price moves that the trader assesses to be noise (the \(U\) term) versus those assessed to be information (the \(S\) term). Instead, the PM seeks to induce the right behavior by compensating the trader according to execution performance relative to a benchmark. That benchmark is necessarily simple – it must be based on data observable to the PM. The most common benchmark in practice is a volume-weighted average of realized market prices during the trading window, i.e., ‘VWAP.’

\(^3\) Who owns the trade if it is dictated by both the PM and the trader? How are the profits split? If they are simply pooled, then the link between the trader compensation and effort is loosely speaking twice as noisy as with direct compensation. Likewise the PM’s incentives are weakened.
2.3.1 Compensation framework

We consider a generalized version of VWAP-based compensation:

\[ A - \phi(\text{TraderWAP} - \text{VWAP}) \cdot (Q^T_1 + Q^T_2) - \alpha(Q^{PM} - Q^T_1 - Q^T_2)^2, \]  

(9a)

where \( Q^{PM} \) is the target quantity specified by the PM; \( Q^T_1 \) and \( Q^T_2 \) are the trader’s execution quantities in each of the two trading rounds; \( A, \phi \) and \( \alpha \) are choice variables of the PM; and the adjusted weighted average prices for the trader and the market, respectively, are:

\[ \text{TraderWAP} = \frac{\hat{P}_1 Q^T_1 + \hat{P}_2 Q^T_2}{Q^T_1 + Q^T_2} \]  

(9b)

\[ \text{VWAP} = \frac{1}{2} \hat{P}_1 + \frac{1}{2} \hat{P}_2; \]  

(9c)

\( \hat{P}_1 \) and \( \hat{P}_2 \) are the market clearing prices in the morning and afternoon auction, respectively, possibly adjusted for factors observable to the PM.\(^4\) These adjustments are discussed below. For now, assume that no adjustment is applied, so performance is a standard VWAP measure with a penalty for non-execution.

The first term in Eq. (9a), \( A \), is a fixed payment to cover the trader’s participation cost; the second term is the trader’s execution performance relative to an adjusted VWAP; and the third term is a penalty for partial execution. Note that we make the simplifying assumption that volume is equal in the two periods. The third term gives the

\(^4\) As mentioned above, the portfolio manager gives the trader two rounds of trading to execute an order. In practice, VWAP benchmarks are typically applied on a daily basis (cite), so we think of the two periods in the model as one trading day, with a morning and an afternoon auction. The execution benchmark is the average clearing price in those two periods.
PM a lever to control execution by penalizing incomplete (or excessive) execution. Note that the PM can induce full execution by choosing $\alpha/\phi$ sufficiently high.\(^5\)

This compensation scheme grants the trader an option to asymmetrically, or partially, fill the order in volatile market conditions (e.g., more than half in one period, less than half in the other). This lowers execution costs to the extent the trader exploits transient supply-demand imbalances caused by noise traders, but it also introduces an agency cost. In particular, the trader can always claim “a difficult market” and partially fill the order when market prices fade away from the order for reasons unrelated to noise trading. For example, absent any adjustment to price, partially filling a buy order in a rising market causes $\text{TraderWAP}^\hat{}$ to tilt toward the lower first period price (assuming the stock has a positive beta). Thus, higher compensation is awarded. Likewise, partially filling a buy order when the stock experiences a positive idiosyncratic signal, $S$ causes $\text{TraderWAP}^\hat{}$ to tilt toward a lower price. Both the adjustment to price alluded to above, and the explicit penalty ($\alpha > 0$), serve to counteract this agency cost.

One final observation on this compensation framework: The trader chooses $Q_2^T$ and $Q_1^T$ to maximize:

$$\max_{Q_{2,1}^T} E \left[ A - \phi \left( \frac{\hat{P}_2 \cdot Q_2^T + \hat{P}_1 \cdot Q_1^T}{Q_2^T + Q_1^T} - \frac{1}{2} (\hat{P}_2 + \hat{P}_1) \right) \cdot (Q_2^T + Q_1^T) - \alpha (Q^{PM} - Q_2^T - Q_1^T)^2 \right]$$

which is equivalent to

$$\min_{Q_{2,1}^T} E \left[ \frac{\phi}{2} \cdot \Delta \hat{P}_2 \cdot \Delta Q_2^T + \alpha (Q^{PM} - Q_2^T - Q_1^T)^2 \right]. \tag{11}$$

\(^5\)This quadratic penalty function rather than a constrained linear penalty ($Q_2^T + Q_1^T \leq Q^{PM}$), makes the calculus far more tractible, as we don’t have to do the boundary checks and truncated expectations that
Thus, prices only enter the analysis by way of the adjusted return from period 1 to 2.

### 2.3.2 Adjustments to price

We assume that the performance evaluation either uses raw stock returns (with an indicator variable $I_{Ctrl}$ set to 0) or returns adjusted for market moves ($I_{Ctrl} = 1$). Thus:

$$\Delta \hat{P}_2 = \Delta P_2 - V_1 R^M_2 \cdot I_{Ctrl} \equiv \Delta P^{FV}_2 + \Delta P^{NT}_2 + \frac{\lambda}{2} \Delta Q^T_2 + \varepsilon^P_2,$$

(12)

where

$$\Delta P^{FV}_2 = V_1 R^M_2 (1 - I_{Ctrl}) + \frac{1}{2} (\psi_2 - wS_1) + wS_2,$$

$$P^{NT}_t = \frac{\lambda}{2} zU_t,$$

$$\varepsilon^P_2 = \frac{1}{2} ((\psi_3 - wS_2) + \lambda (\Delta \nu_2 - \Delta zU_2)).$$

In this expression, $\Delta P^{FV}_t$ indicates fair-value moves in price that are orthogonal to the trader’s information, but which impact the trader’s compensation; $\Delta P^{NT}_t$ indicates price changes due to noise traders. The trader cannot identify $\varepsilon^P_2$ as either informed $(\psi_3 - wS_2)$ or noise $(\Delta \nu_2 - \Delta zU_2)$. By contrast, the PM cannot even identify $\Delta P^{FV}_t$ versus $\Delta P^{NT}_t$. We later show that the larger the variance of $\Delta P^{FV}_t$, denoted $\sigma^2_{FV}$, the

---

Note that $I_{ctrl} = 0$ dominates in practice, i.e., trading performance is typically not evaluated relative to market-adjusted VWAP. One possible explanation is the difficulty of obtaining and processing the necessary data. For example, the typical stock trades hundreds of times during the day. The market return from the beginning of the trading window to the time of each trade would have to be recorded, and a proper beta applied, to construct the market-adjusted VWAP.

Note that both $\Delta P^{FV}_t$ and $P^{NT}_t$ are assessments made by the trader, using the signals $S$ and $U$. They do not represent the actual information and noise components to returns (unless the signals have infinite precision).
greater the agency costs of institutional trading. By contrast, the larger the variance of \( \Delta P_t^{NT} \), denoted \( \sigma_{NT}^2 \), the greater the incentive to hire and grant discretion to a trader.

The formulation of \( \Delta P_t^{FV} \) suggests a third specification: Control for market and idiosyncratic changes in fair value, i.e., \((\nu_2 - wS_1)\). This is not likely feasible, however, as the fair value implied by a public information event (e.g., an earnings announcement) is typically not well defined, nor is the unexpected component of that information.

2.4. Demands

The trader chooses demands \( Q^T_1 \) and \( Q^T_2 \) recursively, first choosing \( Q^T_2 \) to minimize Eq. (11) using Eq. (13):

\[
\min_{Q^{T,1}_2} E \left[ \frac{\phi}{2} \cdot \Delta \hat{P}_2 \cdot \Delta Q^T_2 + \alpha \left( Q^{PM} - Q^T_2 - Q^T_1 \right)^2 \right],
\]

which yields

\[
Q^T_2 = \Omega \left( Q^T_1 - \left( \frac{\Delta P^{FV}_2 + \Delta P^{NT}_2}{\lambda} \right) \right) + (1 - \Omega) \left( Q^{PM} - Q^T_1 \right),
\]

using \( \Omega = \frac{\phi \lambda}{\phi \lambda + 4\alpha} \) for notational convenience. \( \Omega \) indicates the relative emphasis that the PM places on trading profits versus quantity precision.

Next, the trader chooses period 1 demands, incorporating Eq. (16), yielding:

\[
Q^T_1 = \frac{1}{2} \left( Q^{PM} - \frac{P^{NT}_1}{\lambda} \right).
\]

Eq. (17) indicates that the trader chooses to deviate from the default period 1 quantity, \( Q^{PM}/2 \). Returning to the afternoon auction, Eqs. (16) and (17) yield:
\[ Q_2^T = \frac{1}{2} \left( Q_{PM} + \frac{P_1^{NT}}{\lambda} \right) - \frac{\Omega}{\lambda} \left( P_2^{NT} + \Delta P_2^{FV} \right). \] (18a)

\[ Q_2^T = \left( Q_{PM} - Q_1^T \right) - \frac{\Omega}{\lambda} \left( P_2^{NT} + \Delta P_2^{FV} \right). \] (18b)

where the first bracketed expression reflects the period 2 default quantity – the PM’s target demands less that filled in period 1 – and the second bracketed term reflects the trader-imposed quantity adjustment in period 2.\(^9\)

### 3. Model Analysis

This section examines the intuition behind the demand expressions derived in section 2, and gathers several predictions that arise from the model.

#### 3.1. Interpreting the demand equations

In both period 1 and period 2, the trader deviates from the prorated demands of the portfolio manager, \( Q_{PM} / 2 \). In period 1, this deviation reflects the trader’s attempt to counter his signal of noise trading, \( P_1^{NT} \), consistent with the PM’s intent. In period 2 this deviation reflects both the countering of noise trading, \( P_2^{NT} \), and fair-value price changes not accounted for in the trader’s performance evaluation, \( \Delta P_2^{FV} \).

From the PM’s perspective, fair-value changes in price should have no bearing on the quantity traded because \( Q_{PM} \) is motivated by liquidity; portfolio rebalancing; or long-

\(^8\) Calculations in appendix B.

\(^9\) The special case of \( \phi = 0 \) should be singled out, because the trader then has no incentive to react to noise trader distortions in period 1 when \( \phi = 0 \); his only concern is complete execution. Thus, he chooses \( Q_1^T + Q_2^T = Q_{PM} \) but otherwise period 1 demands are indeterminate. We therefore assume equal pro-roration across periods in this special case.
horizon information considerations – all of which are assumed to be independent of the high-frequency information contained in $\Delta P_2^{FV}$. Nevertheless, the trader manipulates order quantity in response to $\Delta P_2^{FV}$, buying less (or selling more) in period 2 when $\Delta P_2^{FV} > 0$ and buying more (or selling less) when $\Delta P_2^{FV} < 0$. Doing so mimics the behavior that the PM wants to see in response to transient price distortions caused by noise traders, and the PM cannot identify the true source of the price move. As a result, the PM incurs the agency cost of compensating the trader without any value added.

Note that fair-value moves in the period 1 price $(\Delta P_1^{FV})$ do not influence period 1 demands. These affect the benchmark $E_1[WAP\hat{P}]$ and performance $E_1[TraderWAP\hat{P}]$ equally, hence the trader sees no reason to skew execution quantities in response to $\Delta P_1^{FV}$. By contrast, fair value moves in period 2 can’t affect the price of trades already executed in period 1, so skewing trading in response to $\Delta P_2^{FV}$ is in the trader’s interest.

Another difference between period 1 and 2 demands is the $\Omega$ coefficient in period 2 demands. Recall that $\Omega$ indicates the relative emphasis that the PM places on trading profits versus complete execution of target demands ($Q^{PM}$). Because discretionary trading by the trader in period 1 (i.e., deviation from $Q^{PM}/2$) can be reversed in period 2, the trader fully pursues opportunities to improve execution performance in period 1 (if he receives some benefit for doing so, i.e., $\phi > 0$). By contrast, in period 2, the trader is punished for deviating from the target order quantity if $\Omega < 1$ because the trader has no chance to settle up the deviation from $Q^{PM}$.

3.2. Results on trader behavior
These demands suggest an explanation for the findings in Lipson and Puckett (2006), which shows that institutions are net sellers when markets are rising and net buyers when markets are falling. In particular, $R^M_2$ is a significant component of $\Delta P^{FV}_2$ when the PM monitors the trader with a straight VWAP procedure ($I_{cr} = 0$). According to Eq. (18), the volume of orders submitted by the trader to the market is negatively related to $\tilde{R}^M_2$. Our model is also consistent with a second result in Lipson and Puckett: The increase in institutional selling in rising markets, and buying in falling markets, can be attributed entirely to trade implementation (i.e. the discretion of the trader) rather than position decisions (the discretion of the PM). Recall that the PM requests a fill of $Q^{PM}$ irrespective of market conditions in our model. We summarize this result as

Proposition 1

To the extent that traders’ performance evaluation is based on straight VWAP ($I_{cr} = 0$), the volume of executed institutional buy (sell) orders will be negatively (positively) related to market returns, even if the portfolio manager’s order is independent of the market.

The evidence in Lipson and Puckett is consistent with the casual empirical observation that, in practice, straight VWAP is the dominant metric for evaluating traders. To the best of our knowledge, market adjusted VWAP is only a conceptual framework introduced in this paper and not widely practiced.

According to Eq. (18), period 2 fair-value moves relating to idiosyncratic public information, $\left(\psi_2 - wS_1\right)$, and idiosyncratic private information inferred from order flow, $wS^2$, also provide an opportunity to exploit the compensation scheme. In both cases, the compensation scheme does not distinguish from price moves due to noise trading. Thus,
**Proposition 2**

The volume of executed institutional buy (sell) orders is negatively (positively) related to permanent, idiosyncratic stock returns that can be attributed to both public and private information. Again, this effect occurs at the trading level, rather than the portfolio management level.

Proposition two is not examined in Lipson and Puckett.

Our final proposition is an existence result. Appendix C shows that the PM’s expected compensation cost (trader’s compensation received) is

\[
C^{\text{Comp}} = A + \frac{\phi}{4\lambda} \left( \Omega \cdot \sigma_{NV}^2 + (1 + \Omega) \cdot \sigma_{NT}^2 \right).
\]

(19)

Since \( A \) is the trader’s compensation under the no-effort strategy of completely filling the order at VWAP, and all terms in Eq. (19) are positive, this establishes

**Proposition 3**

The trader chooses to employ the MM’s technology and apply discretion in executing the order, rather than simply following VWAP exactly, if either

- there is some price fluctuation due to noise trading that the trader can identify and successfully exploit (i.e., \( \sigma_{NT}^2 > 0 \)), or
- there is an exogenous component to the stock’s return (driven by the market or idiosyncratic public information, i.e. \( \sigma_{NV}^2 > 0 \)).

The first bullet (price fluctuations that create opportunities for trading profits) represents the essence of the PM’s objective in employing the compensation scheme. The second bullet (price fluctuations that offer no opportunity for trading profits) represents an agency cost. By assumption, the PM does not have sufficient time or experience to discriminate between the two motives for discretion, so the trader’s performance is affected by both factors.
Given both a benefit and a cost to the compensation arrangement, it is not obvious that the PM will in fact choose to hire and grant discretion to the trader. The PM could alternatively simply execute according to a mechanical VWAP strategy using market orders, which eliminates both the compensation and agency cost of the trader. The decision to hire the trader is analyzed in the next section.

3.3. The portfolio manager’s decision

Define the total expected costs associated with the PM’s order as the sum of two components: the relative execution cost of completely filling the order versus a mechanical VWAP strategy, \( C^{\text{Exec}} \), and the trader’s compensation, \( C^{\text{Comp}} \) (Eq. xx). To evaluate \( C^{\text{Exec}} \) we introduce a third period, in which the PM executes the residual quantity \( \left(Q^{PM} - Q_2^T - Q_1^T\right) \) with a market order. We then compare the cash flows from this execution plus the cash flows from the trader’s executions to that which would have obtained had the PM submitted a market order for \( Q^{PM}/2 \) in each of the first two periods (i.e., VWAP trading). Appendix D shows that this yields

\[
C^{\text{Comp}} + C^{\text{Exec}} = A + \frac{5}{4} \lambda \Omega^2 \cdot \sigma_{FV}^2 - \left(1 + 3\Omega - 5\Omega^2\right)\frac{\sigma_{NT}^2}{4\lambda},
\]

and that the PM’s optimal choice for \( \Omega \), minimizing Eq. (21), is

\[
\Omega^* = \frac{3}{10} \cdot \frac{\sigma_{NT}^2}{\sigma_{FV}^2 + \sigma_{NT}^2}.
\]

According to Eq. (22), when \( \sigma_{FV}^2 \) is high the PM puts more weight on execution completion than trading profits (\( \Omega^* \) small in Eq. 10), because high \( \sigma_{FV}^2 \) means high agency costs of delegating trade implementation. Conversely, when \( \sigma_{NT}^2 \) is high, granting discretion is
preferred because the trader can generate substantial reductions in trading costs, so the PM chooses a high $\Omega^*$. Appendix D also establishes

**Proposition 4**

The PM is better off hiring a trader, rather than algorithmically trading according to VWAP, *if and only if* the trader’s fixed cost, $A$, satisfies

$$A < \frac{\sigma_{NT}^2}{4\lambda} \left( 1 + \frac{3}{2} \Omega^* \right).$$

However, if the PM does employ a trader, then the PM will grant the trader discretion to partially fill orders.

The second part of proposition 4 is a central result of the paper, as it establishes that the PM chooses to present the trader with a “gameable” compensation arrangement.

**Implementation shortfall and other alternatives**

As discussed in section 3.1, agency costs arise because in period 2 the trader can exploit the fact that half of his performance benchmark is set in period 1. Another popular benchmarking scheme is the implementation shortfall measure of Perold (1988). Under this approach 100% of the benchmark is set prior to the trader choosing both $Q^T_1$ and $Q^T_2$. Hence, both the incentive and the opportunity to manipulate performance is greater. Moreover, implementation shortfall is substantially more noisy than VWAP because *all* returns during the trading window figure in the trader’s compensation, which further accentuates the potential to manipulate. Thus, implementation shortfall does not change the basic dynamics of the problem.

Extending the trading window also does not alter the dynamics. The trader’s ability to profit from noise trading (i.e., lower trading costs) is proportional to
\[ \left( \frac{\sigma_{NT}}{\lambda} \right) \cdot \left( 1 + \frac{\Omega - 1}{N} \right) \], where \( N \) is the number of periods the trader is given to work the order. This increases with the trading window, up to a limiting value of \( \frac{\sigma_{NT}}{\lambda} \).

However, the trader’s incentive to manipulate the performance metric in response to fair-value price changes, proportional to \( \sigma_{FV} Q^{PM} \cdot \left( (N - 1)/2N \right) \), also increases with \( N \) up to limiting value \( \frac{\sigma_{FV} Q^{PM}}{2} \). Thus both the bad and the good associated with granting discretion to the trader increases when the trading interval is expanded.

A third alternative is a contingent extension of the trading window when the fill is incomplete. This allows the trader to manipulate VWAP performance along a second dimension. For example, suppose that the order is to buy and the fair-value price increases during the day, in which case the trader incompletely fills the order on average (see Eq. ). Extending the trading window causes the VWAP benchmark to increase: Since the fair-value price change is permanent, subsequent volume occurs, on average, at the higher period 2 price. Therefore, the longer the trader waits to fill the order, the better his performance relative to VWAP.

2.8. Price adjustment delays

Propositions 5 and 6 below summarize the model results on price adjustment delays, using the post-trading third period used to evaluate the PM’s decision.

**Proposition 5**

The return on the stock from period 0 (pre-trade window) to period 1 reflects a beta of 1 against the concurrent market return (equal to its true beta). Moreover, there is no covariance between the period 2 return on the stock and the period 1 return on the market. Thus, there are no price adjustment delays arising in the morning (period 1) auction.
Proposition 6

The return on the stock from period 1 to period 2 reflects a beta of \( 1 - \frac{\Omega^*}{2} \) against the concurrent market return, with \( \Omega^* \) as in Eq. (22). The period 3 return on the stock (from its period 2 value to its period 3 (fair) value) positively covaries with the period 2 market return, with a lag-beta coefficient of \( \frac{\Omega^*}{2} \). Thus, while the price eventually fully reflects its period 2 fair value, it does so only with a delay.

These results follow from the demand equations, (17) and (18), and the price equation, (8). The details are given in Appendix F.

There are no price-adjustment delays in the morning auction (Proposition 5) because the trader does not at that time have any information about his expected execution price versus VWAP.\(^{10}\) By contrast, in the afternoon auction, the trader knows whether the execution price (and roughly half of TraderWAP) will be higher or lower than VWAP. Thus, the trader manipulates TraderWAP by adjusting execution quantity in the afternoon auction. For example, with a sell in a rising market, overloading the afternoon auction puts more weight in TraderWAP on the (higher) afternoon price, causing TraderWAP to exceed VWAP. Likewise, with a sell in a falling market, underloading the afternoon auction tilts TraderWAP towards the higher morning auction price, causing TraderWAP to exceed VWAP.

An important corollary to our results on price adjustment delays relates to cross sectional variation. In particular, being a function of \( \Omega^* \), the severity of price adjustment delays in our model depend on \( \sigma_{NT}^2 \). Stocks with very deep markets, and little

\(^{10}\) This would not be the case with an implementation shortfall measure; with that price adjustment delays in the morning period would be predicted.
predictable price distortion from noise traders (e.g., large cap, heavily traded stocks) are not likely to exhibit material price adjustment delays, according to our theory. Moreover, PMs are not expected to give traders much discretion to manipulate execution relative to VWAP for these stocks. Thus, our model also yields the following

**Corollary 1**

The price adjustment delay converges to 0 as $\lambda \to 0$. Thus, institutions do not induce material price-adjustment delays in the most liquid securities. Moreover, the PM tends to use low-cost, mechanical VWAP strategies for these securities.

### 2.9. Other models of price adjustment delays

Prior studies attribute price-adjustment delays to various frictions in the trading process. While much of this predictability is an illusion caused by nonsynchronous trading, roughly half appears to reflect genuine price-adjustment delays [Kadlec and Patterson (2000)].

For example, with transaction costs, it is optimal for investors to accumulate “news bits” until their value exceeds the cost of transacting [Goldman and Sosin (1979), Cohen et al. (1983) Mech (1992)]. Also, specialists or dealers may impede the adjustment of prices because of exchange stabilization obligations or inventory imbalances [Hasbrouck and Sofianos (1993)]. Finally, public limit orders and trading strategies may impede the adjustment of prices [Kavajecz and Odders-White (2004), Admati and Pfleiderer (1988), Foster and Vishwanathan (1992)].

---

11 For additional studies of nonsynchronous trading and serial correlation see i.e., Fisher (1966), Lo and MacKinlay (1990), Boudoukh, Richardson, Whitelaw (1994), and Boudoukh, Richardson, Whitelaw (2003).

12 In addition to these market microstructure factors, information production by analysts and institutional investors affect stocks’ response to information [Brennan, Jagadeesh, Swaminathan (1993), Badrinath, Kale, and Noe (1995), Chordia and Swaminathan (2000)]. Finally, there is evidence that biases in human judgment may contribute to price-adjustment delays [Barberis and Thaler (2003)].
demonstrates another way in which market participants impede the adjustment of prices to market information.

3. Empirical Analysis

We test our model’s predictions regarding institutional trading and price-adjustment delays by regressing market-adjusted next-day stock returns following large informational events (equity index futures returns) on proxy variables characterizing the potential for agency conflict: price-VWAP ratio, buy-sell order flow imbalance, trading volume, and time of day. This section describes our sample selection, data sources, and the methodology used to conduct our tests.

3.1 Sample

Our sample is drawn from the universe of all U.S. domiciled common stocks traded on the New York Stock Exchange, American Stock Exchange, or Nasdaq during the period January 2001 through December 2001 -- roughly 7000 stocks. Because of the immense data requirements of our tests, we restrict our sample to stocks where our hypotheses regarding price-adjustment delays (PADs) are most relevant.

The first limitation imposed on the sample relates to a stock’s market beta. Because our proxy for information events is the return on equity index futures, the potential for a stock’s PAD is directly related to its beta (see Chalmers, Edelen, and Kadlec (2001)). Thus, we exclude stocks whose monthly beta estimate is less than 0.50. This screen eliminates 43% of the universe: 22% due to insufficient data to estimate beta (36 monthly returns), and 21% due to an estimated beta less than 0.5.
The second limitation on the sample relates to the typical holdings of institutional investors. Several studies document that institutions hold few micro-capitalization stocks [Falkenstein (1996)]. Thus, we exclude stock’s whose total market capitalization is less than $100 million (the lower bound for additions to the S&P 600 small cap index is $300 million). This screen eliminates 50% of the high-beta universe. Our final sample consists of 2200 stocks.

Panel A of Table 1 presents summary statistics of various characteristics of our sample stocks. The average (median) sample stock has a market capitalization of $5.2 billion ($0.7 billion), share price of $57 ($22), monthly beta of 0.85 (0.82), average daily return of 0.0008 (0.0007) and standard deviation of daily return of 0.037 (0.032). The proportion of sample stocks listed on the New York Stock Exchange, American Stock Exchange, and Nasdaq are 53 percent, 3 percent, and 44 percent, respectively.

3.2 Data

3.2.1 Futures transaction data

Price adjustment delays must be with respect to some information. Our proxy for information is the return on the “most relevant” index futures contract for each stock. More specifically, we estimate single-factor market model for each stock using monthly returns on the S&P 500, Nasdaq 100, and Russell 2000 indices, and choose the index with the highest $R^2$. We then use intra-day prices on the most active index futures contract on the Chicago Mercantile Exchange for the chosen index. The date comes from Tick Data, Inc. Because these contracts trade a portfolio of assets via a single security,

---

13 Those stocks removed due to market capitalization have an average beta of 1.35, thus the average beta of the remaining sample turns out to be less than 1 despite truncating at 0.5.
there is little reason to expect spurious predictability from nonsynchronous trading in this information source. The serial correlation of daily returns for the S&P 500, Nasdaq 100, and Russell 2000 futures contracts during 2001 are 0.04, 0.03, and 0.01, respectively.

### 3.2.2 Stock transaction & quote data

We collect transaction prices, volume, and bid and ask quotations for our sample stocks from the NYSE’s TAQ database. Following Hasbrouck (2003) we use only trades and quotes from the primary exchange for each stock. TAQ flags atypical trades and quotes by identifying trades that are batched; executed as part of a basket trade; or reported out of sequence. We exclude all such trades. We also exclude quotes that are identified as either non BBO eligible (best efforts basis), or immediately following a trading halt.

We apply filters to remove observations that may be subject to data entry errors (e.g., transposed and dropped digits). We eliminate transactions and quotes with reversals of greater than 10 percent over a sequence of three observations. Following Keim (1989), we eliminate quotes where the bid-ask spread is greater than 20 percent of the price for stocks priced over $10 dollars or greater than $2 for stocks priced under $10 dollars. Finally, we eliminate all transactions that occur following a quote that was eliminated.

To obtain a more accurate temporal ordering of trades and quotes, we adjust the time stamp for trades to correct for reporting delays in trades relative to quotes as documented in Lee and Ready (1991). However, we use a 1 second adjustment as
opposed to a 5 second adjustment due to the fact that the reporting delay is substantially smaller in our more recent sample period.

Panel B of Table 1 presents summary statistics of various trading characteristics for the sample stocks. The average (median) sample stock has a relative quoted bid-ask spread of 0.0068 (0.0047), a relative effective bid-ask spread of 0.0027 (0.0018), trades 437 (107) times per day with daily volume of 1.1 (0.2) million shares.

3.3 Methodology

3.3.1 Measurement Issues

Various studies calibrate the importance of nonsynchronous trading with the general conclusion that roughly half of the serial correlation of observed daily portfolio returns can be attributed to asynchronous time stamping of transactions. Because our focus is real price-adjustment delays, it is imperative that we employ a methodology that is free from nonsynchronous trading effects. A common means for correcting for nonsynchronous trading is to use the midpoint of the bid and ask quote in place of transaction prices. However, recent evidence suggests that bid and ask quotes are equally stale in the absence of a transaction [Chalmers, Edelen, and Kadlec (2001)]. To ensure proper temporal ordering of our dependent and independent variables, we measure the dependent variable (next-day returns of stocks) following a trade and the independent variables (index futures returns, price-VWAP ratio, buy-sell order imbalance, volume) prior to that trade.

Specifically, we define an “information event” as an economically meaningful return (larger than 60 basis points in magnitude) on the stock’s relevant index futures
contract over the preceding ninety minutes. After identifying an information event, we look for a trade in the stock during the interval 60-120 seconds after the event. If an “initiating” trade took place during this window, we have a valid observation -- that is, we can confirm that our variables have the proper temporal ordering. We then calculate the stock’s subsequent one-day return using the bid-ask midpoint prior to the “initiating” trade, and the bid-ask midpoint at market close on the following trading day. These subsequent one-day returns are then adjusted for concurrent market moves using the stock’s monthly beta times the return on the relevant futures contract over the exact same time interval.

Each of the explanatory variables used in the analysis is measured over the same ninety minute window used to identify “information events.” Even for a sample of 2000 stocks, space and computational requirements necessary to process all trades and quotes are immense. To reduce these requirements we summarize quotes and trades on a minute-by-minute basis, keeping only the last observed quote and trade for a given stock during a given minute of the trading day. Thus, when we refer to a trade or quote, we are referring to the last trade or quote of that minute.

3.3.2 Regression Specifications

To test our model’s predictions regarding institutional trading and price-adjustment delays we estimate the following pooled cross-sectional and time-series regression:

\[
ADJR_{jt} = a + b_1 \cdot \text{Index}_{jt} + b_2 \cdot \frac{P_{jt}}{VWAP_{jt}} + b_3 \cdot \text{BuySell}_{jt} + b_4 \cdot \text{Trading}_{jt} + b_5 \cdot \text{MSM}_{jt} + e_{jt}
\]  

(34)
where:

- $ADJR$ is the subsequent one-day return, adjusted for the concurrent market return.
- $Index$ is the stock’s relevant equity index futures return.
- $P/VWAP$ is the stock’s current midpoint price divided by its current $VWAP$ for the day.
- BuySell is the lag buy-sell order flow imbalance. We employ three alternative specifications for this measure: signed volume, signed share turnover, and signed number of trades.
- Trading is the unsigned lag trading activity measured three alternative ways: share volume, share turnover, and number of trades.
- MSM is the time of the event measured in minutes since midnight.

We estimate separate regressions for negative and positive information events (index futures returns) to allow for differences in the dynamics of price-adjustment delays following negative vs positive information.\(^\text{14}\)

Table 2 reports summary statistics for the variables used in the above regressions. Column 1 reports values for negative information events ($index < -0.0060$), while column 3 reports values for positive information events ($index > 0.0060$). For purposes of comparison, column 2 reports values for non-information events ($\text{abs}(index) < 0.0010$)).

A number of results are noteworthy. First, price-adjustment delays following negative information events are greater in magnitude than price-adjustment delays following positive information events. For example, the average $ADJR$ following a negative information event is 36 bp below the average non-information $ADJR$. By contrast, the

\(^{14}\) One might expect differences in price-adjustment delays due to factors such as short sale constraints.
average \textit{ADJR} following a positive information event is 24 bp above the average non-information \textit{ADJR}. Second, the order flow imbalance during negative information events is smaller than the order flow imbalance during positive information events. This is to be expected (buys outweigh sells during a rising market, and vice versa). Finally, trading activity is considerably higher for both negative and positive events than for non-information events.

3.4 Regression Results

Table 3 reports coefficient estimates and t-statistics (in parentheses) for the various regression specifications. Panel A reports estimates for regressions of negative information events while panel B reports estimates for regressions of positive information events. The columns of table 3 correspond to different specifications of buy-sell order flow imbalance and trading activity.

The coefficient estimates of table 3 are consistent with the predictions of our model. First, price-adjustment delays are positively related to the \textit{Price-VWAP} ratio. That is, next-day stock returns are more negative following negative market returns when \textit{Price-VWAP} is low, and stock returns are more positive following positive market returns when the \textit{Price-VWAP} is high. For example, from column 1 of panels A and B, the coefficients for \textit{Price-VWAP} are 0.044 (t = 6.40) and 0.039 (t = 5.84) for the negative and positive information event regressions, respectively. While many of our model’s predictions are consistent with other explanations of price-adjustment delays, this prediction appears to be is unique to our model.
Second, price-adjustment delays are negatively related to lag buy-sell order flow imbalance. From column 1 of panels A and B, the coefficients for buy-sell volume imbalance are -0.023 (t = -2.07) and -0.080 (t = -6.36) for the negative and positive event regressions, respectively.\textsuperscript{15} Third, price-adjustment delays are negatively related to frequency of trade. From column 1 of panels A and B, the coefficient on volume is positive in the negative market return regression, 0.669 (t =13.12), and negative in the positive market return regression, -0.234 (t = -4.40).

Finally, price-adjustment delays are positively related to the time of day. From column 1 of panels A and B, the coefficient on $MSM$ is negative and significant, -0.003 (t =-3.33), in the negative market return regression and positive and significant, 0.012 (t = 11.10), in the positive market return regression. As with $Price-VWAP$, this relation appears to be unique to our explanation of price-adjustment delays. Note that inferences from specifications using alternative measures of buy-sell order flow imbalance and frequency of trade (columns 2 and 3) are nearly identical to those discussed above.

A natural concern is that our proxies are capturing other forms of predictability in returns – as opposed to price-adjustment delays. For example, it is possible that $Price-VWAP$ is capturing some aspect of momentum – though momentum is typically defined with respect to much longer intervals. Similarly, Chordia and Roll, and Subrahmanyam (2004) also document a negative relation between stock returns and lag order flow imbalance – which they attribute to price-pressure reversals. To address this concern we estimate the same regression specifications using non-information events. That is,

\textsuperscript{15} This relation between stock returns and lag buy-sell order flow imbalance is also consistent with price reversals due to price pressures caused by order imbalances [Chordia, Roll, and Subrahmanyam (2002), and Chordia and Subrahmanyam (2004)]. We later provide evidence which attempts to isolate these two effects.
observations where the lag market return is close to zero. If the proxies are merely picking up predictability that is unrelated to lag market returns, they should be invariant to such partitioning.

Panel C of table 3 reports coefficient estimates (t-statistics) for regressions using non-information events (abs(lag index return) < 0.0010). As expected some of the coefficients are significant even when there is no discernable information event. For example, the coefficient for Price-VWAP is positive and significant in all regressions and the coefficient for buy-sell order imbalance is negative and significant for all regressions. Thus, some of the proxies appear to capture other forms of microstructure related predictability. However, the issue is not whether the proxies capture other forms of predictability but rather whether the specification of our model sharpens the predictive power of these variables in a manner that is consistent with our model. To that end the answer appears to be yes. In particular, the coefficients of most of our proxies are larger in magnitude (in the proper direction) and significance in the regressions using information events than in the regression using non-information events.

4. Summary and Conclusions

This paper develops and tests a model for (true) price-adjustment delays that arise from an agency conflict inherent to institutional investment managers. The central result of our model is that institutional traders will more aggressively fill sell orders in a rising market, and buy orders in a falling market. Likewise, they will tend to leave unfilled buy orders in a rising market, and sell orders in a falling market. Most
importantly, this behavior is completely attributable to the trader, not the portfolio manager. It is a manifestation of the agency relation inherent in institutional investing.

Our model of this agency conflict in institutional trading yields a number of testable implications regarding cross-sectional and time-series properties of price-adjustment delays. Using a large sample of stocks and transactions data, we document several conditional characteristics of price-adjustment delays that are consistent with our model. Specifically we find that price-adjustment delays are: (1) positively related to stocks’ price-VWAP ratio, (2) negatively related to lag buy-sell order flow imbalance, (3) negatively related to lag trading volume, and (4) positively related to time of day. However, the most compelling evidence for our model comes from the two principal findings in Lipson and Puckett (2006). They show that institutions are net sellers when markets are rising and net buyers when markets are falling, and that this pattern is due to trade implementation rather than position decisions.

Our model of price-adjustment delays has a number of parallels with prior studies of trading. First, the agency conflict of our model is similar to that proposed by Harris and Schultz (1998) to explain the viability of SOES bandit trading. Second, because the price-adjustment delays of our model arise from demand for liquidity, our analysis is consistent with the recognition of technical traders (short-term traders in our model) as providers of liquidity as opposed to exploiters of market inefficiencies [Kavejecz and Odders-White (2004)]. Our model also provides new insights on other empirical research. Third, our model demonstrates another way in which market participants impede the adjustment of prices to market information. For example, Hasbrouck and Sofianos (1993) show how specialists or dealers may impede the adjustment of prices
because of exchange stabilization obligations or inventory imbalances. Similarly, Admati and Pfleiderer (1988), Foster and Vishwanathan (1992) show how public limit orders and trading strategies may impede the adjustment of prices.
Figure 1  Binomial model of trading day

PreTrade  Morning Auction  Afternoon Auction

Node 3
- Execute $Q_1^T = U_1$
- Submit conditional demands $\{U_3, D_3\}$ to afternoon market

Node 2
- Execute $Q_2^T = U_2$
- Submit conditional demands $\{U_2, D_2\}$ to afternoon market

Node 1
- Submit conditional demands $\{U_1, D_1\}$ to morning market

Paths are determined by:

$$\tilde{Q}_1^N = \begin{cases} S \\ -S \end{cases}$$  \hspace{1cm} $$\tilde{Q}_2^N = \begin{cases} S \\ -S \end{cases}$$
Table 1

Descriptive Statistics for Sample Stocks

Table reports descriptive statistics for 2200 sample stocks drawn from the universe of 7000 U.S. domiciled common stocks listed on the American Stock Exchange, New York Stock Exchange, or Nasdaq during the period January 2001 through December 2001. The final sample (2200) excludes stocks with estimated betas less than 0.50 and market capitalization less than $100 million.

Panel A: Size, Risk, and Return Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>10th Ptile</th>
<th>90th Ptile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Cap (Millions)</td>
<td>5220</td>
<td>719</td>
<td>162</td>
<td>8519</td>
</tr>
<tr>
<td>Stock Price</td>
<td>57.28</td>
<td>21.97</td>
<td>7.75</td>
<td>47.58</td>
</tr>
<tr>
<td>Beta</td>
<td>0.85</td>
<td>0.82</td>
<td>0.64</td>
<td>1.20</td>
</tr>
<tr>
<td>Average Daily Return</td>
<td>0.0008</td>
<td>0.0007</td>
<td>-0.0013</td>
<td>0.0031</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.0370</td>
<td>0.0322</td>
<td>0.0177</td>
<td>0.0628</td>
</tr>
</tbody>
</table>

Panel B: Trading Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>10th Ptile</th>
<th>90th Ptile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily Volume (Thousands)</td>
<td>1071</td>
<td>239</td>
<td>27</td>
<td>2144</td>
</tr>
<tr>
<td>Daily Trades</td>
<td>219</td>
<td>55</td>
<td>9</td>
<td>338</td>
</tr>
<tr>
<td>Quoted Spread</td>
<td>0.0067</td>
<td>0.0048</td>
<td>0.0016</td>
<td>0.0139</td>
</tr>
<tr>
<td>Effective Spread</td>
<td>0.0025</td>
<td>0.0017</td>
<td>0.0004</td>
<td>0.0054</td>
</tr>
</tbody>
</table>
Table 2

Descriptive statistics of PAD regression variables

Table reports mean (median) values for variables used in price-adjustment delay (PAD) regressions. The dependent variable, ADJR, is the next-day market adjusted return for the stock following an “information event”. We define three types of information events based on the return of a relevant index futures over the preceding 90 minute period: negative information (Index<-0.0060), non information (abs(Index)<0.0010) and positive information (Index>0.0060). All independent variables are measured over the same 90-minute interval. Price-VWAP is the ratio of the midpoint of the current bid-ask quote divided by the volume-weighted average price. BuySell () is a measure of order flow imbalance based on the difference between buy and sell volume, buy and sell turnover, or buy and sell trades. Trading () is a measure of trading activity based on volume, turnover, or number of trades. MSM is the number of minutes since midnight.

<table>
<thead>
<tr>
<th></th>
<th>Negative</th>
<th>Neutral</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADJR</td>
<td>-0.0018</td>
<td>0.0018</td>
<td>0.0042</td>
</tr>
<tr>
<td></td>
<td>(-0.0016)</td>
<td>(0.0014)</td>
<td>(0.0027)</td>
</tr>
<tr>
<td>Index</td>
<td>-0.0099</td>
<td>0.0000</td>
<td>0.0103</td>
</tr>
<tr>
<td></td>
<td>(-0.0082)</td>
<td>(0.0000)</td>
<td>(0.0081)</td>
</tr>
<tr>
<td>Price-VWAP</td>
<td>0.994</td>
<td>1.000</td>
<td>1.006</td>
</tr>
<tr>
<td></td>
<td>(0.996)</td>
<td>(1.000)</td>
<td>(1.004)</td>
</tr>
<tr>
<td>BuySell (volume)</td>
<td>931</td>
<td>7740</td>
<td>25797</td>
</tr>
<tr>
<td></td>
<td>(-100)</td>
<td>(1000)</td>
<td>(2700)</td>
</tr>
<tr>
<td>BuySell (turnover)</td>
<td>-0.023</td>
<td>0.56</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>(-0.040)</td>
<td>(0.022)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>BuySell (trades)</td>
<td>-6</td>
<td>7</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>(-1)</td>
<td>(2)</td>
<td>(7)</td>
</tr>
<tr>
<td>Trading (volume)</td>
<td>307</td>
<td>150</td>
<td>324</td>
</tr>
<tr>
<td></td>
<td>(51)</td>
<td>(38)</td>
<td>(51)</td>
</tr>
<tr>
<td>Trading (turnover)</td>
<td>0.0016</td>
<td>0.0010</td>
<td>0.0017</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td>(0.0005)</td>
<td>(0.0009)</td>
</tr>
<tr>
<td>Trading (trades)</td>
<td>174</td>
<td>112</td>
<td>174</td>
</tr>
<tr>
<td></td>
<td>(69)</td>
<td>(54)</td>
<td>(71)</td>
</tr>
<tr>
<td>MSM</td>
<td>803</td>
<td>807</td>
<td>799</td>
</tr>
<tr>
<td></td>
<td>(835)</td>
<td>(799)</td>
<td>(817)</td>
</tr>
<tr>
<td>Observations</td>
<td>138946</td>
<td>159036</td>
<td>123276</td>
</tr>
</tbody>
</table>
Table 3

Regressions of next-day stock returns on lagged trading characteristics

Coefficient estimates (t-statistics) from pooled cross-sectional and time series regressions of market-adjusted next-day stock returns on lagged index futures returns and trading characteristics. The dependent variable, ADJR, is the next-day market adjusted return for the stock following a “information event”. We define three types of information events based on the return of a relevant index futures over the preceding 90 minute period: negative information (Index<-0.0060), non information (abs(Index)<0.0010) and positive information (Index>0.0060). All independent variables are measured over the same 90-minute interval. Price-VWAP is the ratio of the midpoint of the current bid-ask quote divided by the volume-weighted average price. BuySell () is a measure of order flow imbalance based on the difference between buy and sell volume, buy and sell turnover, or buy and sell trades. Trading () is a measure of trading activity based on volume, turnover, or number of trades. MSM is the number of minutes since midnight.

Panel A: Negative Information Regressions

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index</td>
<td>0.203</td>
<td>0.239</td>
<td>0.207</td>
</tr>
<tr>
<td></td>
<td>(10.78)</td>
<td>(12.61)</td>
<td>(10.96)</td>
</tr>
<tr>
<td>Price-VWAP</td>
<td>0.044</td>
<td>0.043</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>(6.40)</td>
<td>(6.35)</td>
<td>(6.76)</td>
</tr>
<tr>
<td>BuySell (volume)</td>
<td>-0.023</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-2.07)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BuySell (turnover)</td>
<td></td>
<td>-0.244</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.77)</td>
<td></td>
</tr>
<tr>
<td>BuySell (trades)</td>
<td></td>
<td>-0.004</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.08)</td>
<td></td>
</tr>
<tr>
<td>Trading (volume)</td>
<td>0.669</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(13.12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trading (turnover)</td>
<td></td>
<td>0.744</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(15.18)</td>
<td></td>
</tr>
<tr>
<td>Trading (trades)</td>
<td></td>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(13.74)</td>
</tr>
<tr>
<td>MSM</td>
<td>-0.003</td>
<td>-0.003</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td>(-3.33)</td>
<td>(-2.71)</td>
<td>(-3.36)</td>
</tr>
<tr>
<td>Adj. R-square</td>
<td>0.0027</td>
<td>0.0031</td>
<td>0.0028</td>
</tr>
</tbody>
</table>
Table 3

Regressions of next-day stock returns on lagged trading characteristics

Coefficient estimates (t-statistics) from pooled cross-sectional and time series regressions of market-adjusted next-day stock returns on lagged index futures returns and trading characteristics. The dependent variable, ADJR, is the next-day market adjusted return for the stock following a “information event”. We define three types of information events based on the return of a relevant index futures over the preceding 90 minute period: negative information (Index<-0.0060), non information (abs(Index)<0.0010) and positive information (Index>0.0060). All independent variables are measured over the same 90-minute interval. Price-VWAP is the ratio of the midpoint of the current bid-ask quote divided by the volume-weighted average price. BuySell () is a measure of order flow imbalance based on the difference between buy and sell volume, buy and sell turnover, or buy and sell trades. Trading () is a measure of trading activity based on volume, turnover, or number of trades. MSM is the number of minutes since midnight.

Panel B: Positive Information Regressions

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index</td>
<td>0.388</td>
<td>0.378</td>
<td>0.395</td>
</tr>
<tr>
<td></td>
<td>(22.37)</td>
<td>(21.58)</td>
<td>(22.79)</td>
</tr>
<tr>
<td>Price-VWAP</td>
<td>0.039</td>
<td>0.027</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>(5.84)</td>
<td>(4.10)</td>
<td>(6.68)</td>
</tr>
<tr>
<td>BuySell (volume)</td>
<td>-0.080</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-6.36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BuySell (turnover)</td>
<td></td>
<td>-1.063</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-6.01)</td>
<td></td>
</tr>
<tr>
<td>BuySell (trades)</td>
<td></td>
<td></td>
<td>-0.199</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-4.13)</td>
</tr>
<tr>
<td>Trading (volume)</td>
<td>-0.234</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-4.40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trading (turnover)</td>
<td></td>
<td>-0.194</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-3.77)</td>
<td></td>
</tr>
<tr>
<td>Trading (trades)</td>
<td></td>
<td></td>
<td>-0.438</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-5.77)</td>
</tr>
<tr>
<td>MSM</td>
<td>0.012</td>
<td>0.012</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(11.10)</td>
<td>(11.64)</td>
<td>(10.90)</td>
</tr>
<tr>
<td>Adj. R-square</td>
<td>0.0063</td>
<td>0.0061</td>
<td>0.0064</td>
</tr>
</tbody>
</table>
Regressions of next-day stock returns on lagged trading characteristics

Coefficient estimates (t-statistics) from pooled cross-sectional and time series regressions of market-adjusted next-day stock returns on lagged index futures returns and trading characteristics. The dependent variable, ADJR, is the next-day market adjusted return for the stock following a “information event”. We define three types of information events based on the return of a relevant index futures over the preceding 90 minute period: negative information (Index<-0.0060), non information (abs(Index)<0.0010) and positive information (Index>0.0060). All independent variables are measured over the same 90-minute interval. Price-VWAP is the ratio of the midpoint of the current bid-ask quote divided by the volume-weighted average price. BuySell () is a measure of order flow imbalance based on the difference between buy and sell volume, buy and sell turnover, or buy and sell trades. Trading () is a measure of trading activity based on volume, turnover, or number of trades. MSM is the number of minutes since midnight.

Panel C: Non Information Regressions

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index</td>
<td>0.394</td>
<td>0.406</td>
<td>0.375</td>
</tr>
<tr>
<td></td>
<td>(2.43)</td>
<td>(2.50)</td>
<td>(2.30)</td>
</tr>
<tr>
<td>Price-VWAP</td>
<td>0.059</td>
<td>0.056</td>
<td>0.058</td>
</tr>
<tr>
<td></td>
<td>(8.69)</td>
<td>(8.36)</td>
<td>(8.26)</td>
</tr>
<tr>
<td>BuySell (volume)</td>
<td>-0.022</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-2.70)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BuySell (turnover)</td>
<td></td>
<td>-0.298</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-2.21)</td>
<td></td>
</tr>
<tr>
<td>BuySell (trades)</td>
<td></td>
<td></td>
<td>-0.047</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-2.03)</td>
</tr>
<tr>
<td>Trading (volume)</td>
<td>0.007</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trading (turnover)</td>
<td></td>
<td>0.289</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5.80)</td>
<td></td>
</tr>
<tr>
<td>Trading (trades)</td>
<td></td>
<td></td>
<td>-0.175</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-3.13)</td>
</tr>
<tr>
<td>MSM</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(1.03)</td>
<td>(0.80)</td>
<td>(1.22)</td>
</tr>
<tr>
<td>Adj. R-square</td>
<td>0.014</td>
<td>0.016</td>
<td>0.0014</td>
</tr>
</tbody>
</table>
References


Appendices

A. Market Maker Expectations

Let $X$ denote the transformation

$$X_t = P_t - V_t + \lambda Q^F_t = \tilde{\psi}_{t+1} - \lambda \tilde{\upsilon}_t.$$  \hfill (a1)

The random variables $\tilde{\psi}_{t+1}$, $\tilde{U}_t$, and $\tilde{X}_t$ are joint normal

$$
\begin{bmatrix}
\tilde{\psi}_{t+1} \\
\tilde{U}_t \\
\tilde{X}_t \\
\end{bmatrix} =
\begin{bmatrix}
\tilde{\psi}_{t+1} \\
\tilde{\upsilon}_t + \tilde{\eta}_t \\
\tilde{\psi}_{t+1} - \lambda \tilde{\upsilon}_t \\
\end{bmatrix} \sim N
\begin{bmatrix}
0 \\
0 \\
0 \\
\end{bmatrix}
\begin{bmatrix}
y^{-1} & 0 & y^{-1} \\
0 & y^{-1} & \lambda u^{-1} \\
y^{-1} & \lambda u^{-1} & y^{-1} + \lambda^2 u^{-1} \\
\end{bmatrix}, \hfill (a2)
$$

so

$$E[\tilde{\psi}_{t+1} | \tilde{U}_t, \tilde{X}_t] = 0 + [\tilde{U}_t \tilde{X}_t] \Sigma_{22}^{-1} \Sigma_{21}$$

$$= [\tilde{U}_t \tilde{X}_t] \begin{bmatrix}
h^{-1} + u^{-1} & \lambda u^{-1} \\
\lambda u^{-1} & y^{-1} + \lambda^2 u^{-1} \\
\end{bmatrix}^{-1} \begin{bmatrix}
0 \\
y^{-1} \\
\end{bmatrix}$$
\[
\begin{bmatrix}
\frac{uh}{u+h} + \frac{\lambda^2 y h}{(u+h)(\lambda^2 y + u + h)} & -\lambda y h \\
-\frac{\lambda y h}{(\lambda^2 y + u + h)} & \frac{y(u+h)}{(\lambda^2 y + u + h)}
\end{bmatrix}
\begin{bmatrix}
0 \\
y^{-1}
\end{bmatrix}
= \begin{bmatrix}
-\frac{\lambda h}{(\lambda^2 y + u + h)}
\end{bmatrix} \cdot \bar{U}_t + \begin{bmatrix}
\frac{u+h}{\lambda^2 y + u + h}
\end{bmatrix} \cdot \bar{X}_t
= \begin{bmatrix}
\frac{u+h}{\lambda^2 y + u + h}
\end{bmatrix} \begin{bmatrix}
-\frac{h}{u+h} \cdot \bar{U}_t + \bar{X}_t
\end{bmatrix} = w(\bar{X}_t - \lambda z \bar{U}_t) = w \tilde{S}_t
= w(\bar{X}_t - \lambda z \bar{U}_t) = w \tilde{S}_t.
\]

\[\text{Var}[\tilde{y}_{t+1} | \bar{U}_t, \bar{X}_t] = \Sigma_{11} - \Sigma'_{21} \Sigma_{22}^{-1} \Sigma_{21}\]

\[
\begin{bmatrix}
\frac{uh}{u+h} + \frac{\lambda^2 y h}{(u+h)(\lambda^2 y + u + h)} & -\lambda y h \\
-\frac{\lambda y h}{(\lambda^2 y + u + h)} & \frac{y(u+h)}{(\lambda^2 y + u + h)}
\end{bmatrix}
\begin{bmatrix}
0 \\
y^{-1}
\end{bmatrix}
= \begin{bmatrix}
y^{-1} - \frac{y^{-2} y(u+h)}{(\lambda^2 y + u + h)} = y^{-1} \left(1 - \frac{(u+h)}{(\lambda^2 y + u + h)}\right) = y^{-1} \left(\frac{\lambda^2 y}{(\lambda^2 y + u + h)}\right) = \frac{\lambda^2}{(\lambda^2 y + u + h)}
\end{bmatrix}
\]
B. Demands in period 1

First note (see Eq. (16)):

\[
\Delta Q_2^T = -\frac{\Omega}{\lambda} \left( \Delta P_2^{NT} + \Delta P_2^{FV} \right) + (1 - \Omega) \left( Q^{PM} - 2Q_1^T \right) \quad \Rightarrow \quad \frac{\partial \Delta Q_2^T}{\partial Q_1^T} = -2(1 - \Omega) \quad (b1)
\]

\[
Q_1^T + Q_2^T = \Omega \left( 2Q_1^T - \frac{\left( \Delta P_2^{FV} + \Delta P_2^{NT} \right)}{\lambda} \right) + (1 - \Omega) Q^{PM} \quad \Rightarrow \quad \frac{\partial (Q_1^T + Q_2^T)}{\partial Q_1^T} = 2\Omega \quad (b2)
\]

The trader’s period 1 objective is determined by Eq. (14), which has first order condition w.r.t. \( Q_1^T \):

\[
-2(1 - \Omega) \frac{\phi}{2} \left( E_1 \left[ \Delta P_2^{NT} + \Delta P_2^{FV} \right] + \lambda E_1 \left[ \Delta Q_2^T \right] \right) - 2\alpha \cdot 2\Omega \cdot \left( Q^{PM} - E_1 \left[ Q_2^T + Q_1^T \right] \right) = 0 \quad (b3)
\]

Note that \( E_1 \left[ \Delta P_2^{FV} \right] = 0 \) since \( wS_1 = E_1 \left[ \tilde{\psi}_2 \right] \) (see Appendix A). Also, \( E_1 \left[ P_2^{NT} \right] = 0 \). Thus,

\[
-2(1 - \Omega) \frac{\phi}{2} \left( \lambda(1 - \Omega) Q^{PM} - \lambda(1 - \Omega) Q_1^T - (1 - \Omega) P_1^{NT} \right) - 2\alpha \cdot 2\Omega \cdot \left( \Omega Q^{PM} - 2\Omega Q_1^T + \Omega P_1^{NT} \right) = 0 \quad (b4)
\]

\[
\left( -2\alpha \cdot 2\Omega^2 - \phi \lambda (1 - \Omega)^2 \right) \left( Q^{PM} - 2Q_1^T - \frac{P_1^{NT}}{\lambda} \right) = 0 \quad \Rightarrow \quad Q_1^T = \frac{1}{2} \left( Q^{PM} - \frac{P_1^{NT}}{\lambda} \right) \quad (b5)
\]

Prop 1
increases the trader’s incentive to manipulate trading in response to fair-value price moves. For example, with two periods the
trader’s manipulation incentive is, on average, proportional to 0.25 $Q^{\text{PM}}$ (manipulate in period 2 (½ the time), in response to the
quantity filled in period 1 ($Q^{\text{PM}}/2$). With three periods, the incentive is proportional to (1/9 + 2/9 =) 1/3 $Q^{\text{PM}}$. As the number of periods
grows, the incentive approaches proportionality to 1/2 $Q^{\text{PM}}$.

$$\frac{1}{N} \left(0 + \frac{1}{N} + \frac{2}{N} + \frac{N-1}{N} + \ldots\right) = \lim_{N \to \infty} \frac{1}{N} \frac{N \cdot N - 1}{2} = \frac{1}{2}$$
C. Expected compensation

First note:

\[
\Delta Q_2^T = \frac{P_1^{NT}}{\lambda} - \frac{\Omega}{\lambda} \left( P_2^{NT} + \Delta P_2^{FV} \right) \quad (c1)
\]

\[
Q_1^T + Q_2^T = Q^{PM} - \frac{\Omega}{\lambda} \left( P_2^{NT} + \Delta P_2^{FV} \right) \quad (c2)
\]

Where \( \Gamma = \frac{\Delta P_2^{FV}}{\lambda} \) this is a change, from previous

Expected compensation is

\[
A - E_0 \left[ \frac{\phi}{2} \Delta \hat{P}_2 \cdot \Delta Q_2 + \alpha \left( Q^{PM} - Q_2^T - Q_1^T \right)^2 \right] - \frac{\phi \lambda}{4} E_0 \left[ \Delta Q_2^{FV} \right]^2 - \frac{\phi \lambda}{4} \frac{\Omega^2}{\lambda^2} \left( \sigma_{NT}^2 + \sigma_{FV}^2 \right) \quad (c3)
\]

Components:
\[ E_0 [\Delta P_{2}^{FV} \cdot \Delta Q_2] = E_0 \left[ \Delta P_{2}^{FV} \cdot \left( -\frac{\Omega}{\lambda} \Delta P_{2}^{FV} \right) \right] = -\frac{\Omega}{\lambda} \cdot \sigma_{FV}^2 \]  
\[ (c4) \]

\[ E_0 [\Delta P_{2}^{NT} \cdot \Delta Q_2] = E_0 \left[ \Delta P_{2}^{NT} \cdot \left( \frac{P_1^{NT}}{\lambda} - \frac{\Omega}{\lambda} (P_{2}^{NT} + \Delta P_{2}^{FV}) \right) \right] = -\frac{\sigma_{NT}^2}{\lambda} (1 + \Omega) \]  
\[ (c5) \]

\[ E_0 [\Delta Q_2^2] = E_0 \left[ \left( \frac{P_1^{NT}}{\lambda} - \frac{\Omega}{\lambda} (P_{2}^{NT} + \Delta P_{2}^{FV}) \right)^2 \right] = \frac{\sigma_{NT}^2}{\lambda^2} (1 + \Omega^2) + \frac{\sigma_{FV}^2}{\lambda^2} \Omega^2 \]  
\[ (c6) \]

Compensation

\[ A - \left( -\frac{\phi \Omega}{2\lambda} \cdot \sigma_{FV}^2 \right) - \left( -\frac{\sigma_{NT}^2}{2\lambda} \cdot \phi (1 + \Omega) \right) - \left( \frac{\sigma_{NT}^2}{4\lambda} \cdot \phi (1 + \Omega^2) + \frac{\sigma_{FV}^2}{4\lambda} \cdot \phi \Omega^2 \right) - \left( \alpha \cdot \frac{\Omega^2}{\lambda^2} \left( \sigma_{NT}^2 + \sigma_{FV}^2 \right) \right) \]  
\[ (c7) \]

\[ A + \sigma_{FV}^2 \left( \frac{\phi \Omega}{2\lambda} - \frac{\phi \Omega^2}{4\lambda} - \alpha \cdot \frac{\Omega^2}{\lambda^2} \right) + \sigma_{NT}^2 \left( \frac{\phi (1 + \Omega)}{2\lambda} - \frac{\phi (1 + \Omega^2)}{4\lambda} - \alpha \cdot \frac{\Omega^2}{\lambda^2} \right) \]  
\[ (c8) \]

\[ A + \sigma_{FV}^2 \left( 2\phi \Omega - 4\phi \Omega^2 - 4\alpha \cdot \frac{\Omega^2}{\lambda^2} - 4\alpha \cdot \frac{\Omega^2}{4\lambda^2} \right) + \sigma_{NT}^2 \left( \frac{2\lambda \phi (1 + \Omega)}{4\lambda^2} - \frac{\lambda \phi (1 + \Omega^2)}{4\lambda^2} - 4\alpha \cdot \frac{\Omega^2}{4\lambda^2} \right) \]  
\[ (c9) \]

\[ A + \frac{\phi}{4\lambda} \cdot \left( \Omega \cdot \sigma_{FV}^2 + (1 + \Omega) \cdot \sigma_{NT}^2 \right) \]  
\[ (c10) \]
D. Relative cost of order completion

First we show that

\[
C^{\text{Exec}} = E_0 \left[ \frac{1}{2} \Delta P_2 \cdot \Delta Q_2 + (P_3 - VWAP) \cdot (Q^{\text{PM}} - Q^T_1 - Q_2^T) \right],
\]

(20)

where \(C^{\text{Exec}} < 0\) indicates that the trader improves execution performance. Then (Appendix E) shows

\[
\tilde{P}_1^{\text{HYPO}} - \tilde{P}_1 = \frac{\lambda}{4} \left( Q^{\text{PM}} + \frac{\delta_1}{2} \right) - \frac{\lambda}{4} Q^{\text{PM}} = \frac{\lambda}{4} \frac{\delta_1}{2}
\]

\[
\tilde{P}_2^{\text{HYPO}} - \tilde{P}_2 = \frac{\lambda}{4} \left( Q^{\text{PM}} - \frac{\delta_1}{2} \right) + \Omega \frac{\lambda}{2} \left( \frac{\delta_2}{2} - \frac{\Pi_2}{\lambda} \right) - \frac{\lambda}{4} Q^{\text{PM}} = ... = \frac{\lambda}{4} \frac{\delta_1}{2} + \Omega \frac{\lambda}{2} \left( \frac{\delta_2}{2} - \frac{\Pi_2}{\lambda} \right)
\]

Relative cost of executed orders

\[
\left( \frac{\hat{P}_2 \cdot Q^T_2 + \hat{P}_1 \cdot Q^T_1}{Q^T_2 + Q^T_1} - \frac{1}{2} \left( \hat{P}_1^{\text{HYPO}} + \hat{P}_2^{\text{HYPO}} \right) \right) \cdot (Q^T_2 + Q^T_1)
\]

(10)

\[
\left( \frac{\hat{P}_2 \cdot Q^T_2 + \hat{P}_1 \cdot Q^T_1}{Q^T_2 + Q^T_1} - \frac{1}{2} \left( \hat{P}_1 + \hat{P}_2 \right) - \frac{1}{2} \left( \hat{P}_1^{\text{HYPO}} - \hat{P}_1 + \hat{P}_2^{\text{HYPO}} - \hat{P}_2 \right) \right) \cdot (Q^T_2 + Q^T_1)
\]

(10)
\[
\left( \frac{\hat{P}_2 \cdot Q_2^T + \hat{P}_1 \cdot Q_1^T}{Q_2^T + Q_1^T} - \frac{1}{2} (\hat{P}_1 + \hat{P}_2) \right) \cdot (Q_2^T + Q_1^T) - \left( \frac{1}{2} (\hat{P}_1^{HYPO} - \hat{P}_1 + \hat{P}_2^{HYPO} - \hat{P}_2) \right) \cdot (Q_2^T + Q_1^T)
\]

Relative cost of unexecuted orders

\[
\left( P_3 - \frac{1}{2} (\hat{P}_1 + \hat{P}_2) \right) \cdot (Q_{PM}^T - Q_2^T - Q_1^T) - \left( \frac{1}{2} (\hat{P}_1^{HYPO} - \hat{P}_1 + \hat{P}_2^{HYPO} - \hat{P}_2) \right) \cdot (Q_{PM}^T - Q_2^T - Q_1^T)
\]

Add the two residuals ---

\[
- \left( \frac{1}{2} (\hat{P}_1^{HYPO} - \hat{P}_1 + \hat{P}_2^{HYPO} - \hat{P}_2) \right) \cdot (Q_2^T + Q_1^T) - \left( \frac{1}{2} (\hat{P}_1^{HYPO} - \hat{P}_1 + \hat{P}_2^{HYPO} - \hat{P}_2) \right) \cdot (Q_{PM}^T - Q_2^T - Q_1^T)
\]

\[
- \left( \frac{1}{2} (\hat{P}_1^{HYPO} - \hat{P}_1 + \hat{P}_2^{HYPO} - \hat{P}_2) \right) \cdot Q_{PM}
\]

Which is zero in expectation. So, using hypo is the same as using straight VWAP.

**E. Cost to the portfolio manager**

Relative trading costs = unconditional expectation of
\[ E_0 \left[ \frac{1}{2} \Delta P_2 \cdot \Delta Q_2 + (P_3 - V W A P) \cdot \left( Q^{PM} - Q_1^T - Q_2^T \right) \right] \]

\[ = E_0 \left[ \frac{1}{2} \Delta P_2 \cdot \Delta Q_2 + \left( \Delta P_3 + \frac{1}{2} \Delta P_2 \right) \cdot \left( Q^{PM} - Q_1^T - Q_2^T \right) \right] \]

\[ = E_0 \left[ \frac{1}{2} \Delta P_2 \cdot \left( \frac{P_{1NT}}{2\lambda} - \frac{\Omega}{\lambda} (P_{2NT} + \Delta P_{2FV}) \right) + \left( \Delta P_3 + \frac{1}{2} \Delta P_2 \right) \cdot \left( \frac{\Omega}{\lambda} (P_{2NT} + \Delta P_{2FV}) \right) \right] \]

\[ = E_0 \left[ \Delta P_2 \cdot \left( \frac{P_{1NT}}{2\lambda} \right) + \Delta P_3 \cdot \left( \frac{\Omega}{\lambda} (P_{2NT} + \Delta P_{2FV}) \right) \right] \quad (e1) \]

First term

\[ = E_0 \left[ \Delta P_2 \cdot \frac{P_{1NT}}{2\lambda} = E_0 \left[ \left( \Delta P_{2FV} + \Delta P_{3NT} \frac{\lambda}{2} \Delta Q_2^T + \epsilon_{3}^p \right) \cdot \frac{P_{1NT}}{2\lambda} + \Delta P_{3NT} \frac{\lambda}{2} \frac{\sigma_{NT}^2}{2\lambda} \right] = \frac{\sigma_{NT}^2}{4\lambda} \right] \quad (e2) \]

(negative cost means a benefit)

Second term

\[ = E_0 \left[ \left( \Delta P_{3FV} + \Delta P_{3NT} \frac{\lambda}{2} \Delta Q_3^T + \epsilon_{3}^p \right) \cdot \frac{\Omega}{\lambda} (P_{2NT} + \Delta P_{2FV}) \right] = \]

where \( \Delta Q_3^T = \left( Q^{PM} - Q_1^T - Q_2^T \right) - \left( \frac{2\Omega}{\lambda} (P_{2NT} + \Delta P_{2FV}) \right) - \frac{O^{PM}}{2} - \frac{P_{1NT}}{2\lambda} \)
dropping terms with no correlation,

\[ E_0 \left[ \left( \Delta P_{3}^{NT} + \Omega P_{2}^{NT} + \Omega \Delta P_{2}^{FV} \right) \left( \frac{\Omega}{\lambda} \left( P_{2}^{NT} + \Delta P_{2}^{FV} \right) \right) \right] = \]

\[ = -\sigma_{NT}^{2} \frac{\Omega}{\lambda} \left( 1 - \Omega \right) + \frac{\Omega^2}{\lambda} \sigma_{FV}^2 \]

Both first and second terms

\[ = -\sigma_{NT}^{2} \frac{1}{\lambda} \left( \frac{1}{4} + \Omega(1 - \Omega) \right) + \frac{\Omega^2}{\lambda} \sigma_{FV}^2 \]

\[ = \Omega^2 \lambda E_0 \left( \Gamma^{-2} \right) - \left( 1 + 4(1 - \Omega) \Omega \right) \frac{\lambda}{16} E_0 \left( z U_1^2 \right) \]

A.2.3 Total cost to the PM

\[ = -\sigma_{NT}^{2} \frac{1}{\lambda} \left( \frac{1}{4} + \Omega(1 - \Omega) \right) + \frac{\Omega^2}{\lambda} \sigma_{FV}^2 + \left( A + \frac{\phi}{4\lambda} \left( \Omega \cdot \sigma_{FV}^2 + (1 + \Omega) \cdot \sigma_{NT}^2 \right) \right) \]

\[ = A - \frac{\sigma_{NT}^2}{4\lambda} \left( 1 + 4\Omega(1 - \Omega) - \phi(1 + \Omega) \right) + \left( 4\Omega^2 + \phi\Omega \right) \frac{\sigma_{FV}^2}{4\lambda} \]
Optimum choice?

Let \( \alpha = \frac{\lambda}{4} (1 - \phi) \). Then \( \Omega = \frac{lf}{lf + 4l(1 - f) / 4} = f \)

\[
\min \Omega = \alpha - \sigma_{NT}^2 \left( 1 + 3\Omega - 5\Omega^2 \right) + \frac{5\Omega^2}{4\lambda} \sigma_{FV}^2
\]

\[
0 = -\frac{\sigma_{NT}^2}{4\lambda} (3 - 10\Omega) + \frac{10\Omega}{4\lambda} \sigma_{FV}^2
\]

\[
\Omega^* = \frac{3}{10} \cdot \frac{\sigma_{NT}^2}{\sigma_{FV}^2 + \sigma_{NT}^2} < 1.
\]

Test: compensation less than net benefit to PM?

\[
A + \frac{\phi}{4\lambda} \cdot \left( \Omega \cdot \sigma_{FV}^2 + (1 + \Omega) \cdot \sigma_{NT}^2 \right) < \frac{\sigma_{NT}^2}{\lambda} \left( \frac{1}{4} + \Omega(1 - \Omega) \right) - \frac{\Omega^2}{\lambda} \sigma_{FV}^2
\]

\[
4\lambda A < \sigma_{NT}^2 \left( 1 + 4\Omega(1 - \Omega) - \Omega \cdot (1 + \Omega) \right) - 5\Omega^2 \sigma_{FV}^2
\]

\[
4\lambda A < \sigma_{NT}^2 \left( 1 + 3\Omega - 5\Omega^2 \right) - 5\Omega^2 \sigma_{FV}^2
\]

\[
4\lambda A < \sigma_{NT}^2 \left( 1 + 3\Omega \right) - 5\Omega^2 \left( \sigma_{FV}^2 + \sigma_{NT}^2 \right)
\]
\[ 4 \lambda A < \sigma_{NT}^2 (1 + 3 \Omega) - 5 \Omega \cdot \frac{3}{10} \sigma_{NT}^2 \]

\[ A < \frac{\sigma_{NT}^2}{4 \lambda} \left( 1 + \frac{3}{2} \Omega^* \right) \]
A.3. Price adjustment delays

Market clearing prices are

\[ P_1 = V_1 + \lambda (Q_1^T + Q_1^N) \]

\[ P_2 = V_2 + \lambda (Q_2^T + Q_2^N). \]  \hspace{1cm} (a24)

Using Eqs.(15) and (16),

\[ P_1 = P_0 \left( 1 + \tilde{R}_1^M \right) + \lambda \left( \frac{1}{2} Q_{1}^{PM} + \frac{3}{4} Q_{1}^{N} \right) \]

\[ P_2 = V_2 + \lambda \left( \frac{Q_{2}^{PM}}{2} + \frac{Q_{1}^{N}}{4} - \phi \left( Q_{2}^{N} + \frac{V_1 \tilde{R}_2^M}{\lambda} \right) + Q_{2}^{N} \right) \]

\[ = V_1 \left( 1 - \frac{\phi}{2 \lambda} \tilde{R}_2^M \right) + \lambda \left( \frac{Q_{2}^{PM}}{2} + \frac{Q_{1}^{N}}{4} \right) + Q_{2}^{N} \left( 1 - \frac{\phi}{2} \right). \]  \hspace{1cm} (a25)