The Role of Trading Halts in Monitoring a Specialist Market*

Roger Edelen
University of Pennsylvania
and
Simon Gervais
University of Pennsylvania

28 August 2001

*This paper evolved from a previous working paper entitled "Trading Halts in a Principal-Agent Model of an Exchange." Financial support by the Rodney L. White Center for Financial Research is gratefully acknowledged. The paper owes a great deal to George Sofianos for instrumental discussions at the inception of the project, and for assistance with its development and data collection. We have benefited from discussions with Utpal Bhattacharya, Peter DeMarzo James Gallagher, Diego Garcia, Gary Gorton, Joel Hasbrouck, Aditya Kaul, Ken Kavajecz, Richard Kihlstrom, John McMahon, Christopher Quick, Duane Seppi, David Shields, Nicholas Souleles, Matt Spiegel, Paul Weller, Donald Wickert, and Bilge Yılmaz. Also providing helpful comments and suggestions were Maureen O'Hara, two anonymous referees and seminar participants at the University of Washington, M.I.T., Ohio State University, the University of Oregon, New York University, Dartmouth College, Duke University, Temple University, McGill University, the Wharton School, the 1998 meetings of the Western Finance Association, the 1999 NASDAQ-Notre Dame Microstructure conference, the 2000 Market Microstructure conference at Yale University, and the 2000 Market Microstructure Meeting of the NBER. We would also like to thank Murray Teitelbaum for organizing our visit to the NYSE, and Benjamin Croitoru for his research assistance. The views and ideas expressed in this paper are those of the authors alone. Address correspondence to Simon Gervais, Finance Department, Wharton School, University of Pennsylvania, Steinberg Hall - Dietrich Hall, Suite 2300, Philadelphia, PA 19104-6367, (215) 898-2370.

Abstract

When a collection of specialists organize as an exchange, each can reap net private benefits at the expense of the exchange by quoting a privately optimal pricing schedule. Coordination makes all specialists and customers better off, but requires a system of monitoring and punishment, which breaks down when information asymmetries between the exchange and a specialist are high. The specialist may then seek a temporary trading halt to alleviate unjustified punishment, or the exchange may halt trading to prevent the quoting of damaging privately optimal pricing schedules. We test this theory on a sample of NYSE halts. As predicted, we find a significant increase in estimated information asymmetry immediately preceding trading halts.

1 Introduction

Many exchanges are organized and owned by members who perform most of the market making on the exchange. Such exchanges will seek to maximize the collective value of all market-maker activities. This study argues that trading halts play an important role in that maximization by mitigating an agency cost inherent in the organizational structure. Individual market makers, viewed as the agents operating the exchange, do not always have the incentive to act in the interest of the exchange, the principal in this context. This agency conflict arises whenever the exchange enjoys a reputation for trade execution quality that is shared by all market makers. As a result of this pooled reputation, gouging by one of the exchange's market makers (e.g., a substantial, temporary price fade prior to executing a customer's order) may result in a loss of order flow to all of the exchange's market makers. In some cases, gouging may also lead to sufficient trading profits to outweigh the offending market maker's expected loss of future order flow, yet insufficient profits to also cover the externality imposed on all other market makers. Thus, the exchange suffers an agency cost if market makers cannot be induced to internalize this externality.

The theoretical analysis in this paper presumes a specialist-organized market, where the exchange employs a single market marker, a specialist, for each stock. In this setting, exchange profits can increase when the specialists collectively agree to rules and monitoring designed to prevent opportunistic actions. However, given the tremendous amount of stock-specific information that is surely associated with market making, particularly in the case of a specialist market, this monitoring can never be perfect. The theoretical analysis goes on to show that when monitoring is sufficiently imprecise, punishment becomes a deadweight cost that fails to realign the incentives of specialists. Trading halts increase exchange profits by eradicating such no-win scenarios. This motive for halts can explain halts triggered by both the exchange and the specialist. Halts called by the exchange serve to protect all market makers' interests when it is anticipated that an upcoming event or announcement will compromise the exchange's ability to cost-effectively monitor that stock's specialist. Halts called by an individual specialist serve to protect that specialist's interests when monitoring is likely to result in punishment even when the specialist quotes the efficient pricing schedule given his information, i.e., that which maximizes expected exchange profits.

The two main implications of this analysis are as follows. First, trading halts should be called by the exchange on some occasions, and solicited by the specialist on other occasions.[1] This empirical fact, previously documented by Bhattacharya and Spiegel (1998) among others, has not been accounted for in the theoretical literature. Second, trading halts should be preceded by a substantial increase in the information asymmetry between the exchange and the specialist regarding the quoted, versus efficient, pricing schedule. Note that this does not refer to information asymmetry about the value of the stock, which plays no role in our theoretical analysis. This also is a novel prediction.

The empirical analysis in this paper tests these predictions. At most exchanges, monitoring of the market makers involves, among other things, a statistical analysis of trade, quote, and order flow data. From this analysis comes a hypothesized statement of the efficient price schedule for the market maker given existing conditions, and a test of whether the actual quoted schedule violates that norm. Researchers cannot know the statistical procedures employed in this evaluation process, indeed, market makers themselves likely do not know. However, simple statistical estimates of the price impact of transactions should provide a useful, observable proxy to the exchange's monitoring procedure. This proxy approach serves its purpose as long as the exchange tends to lose monitoring precision under the same circumstances that our simple estimates lose statistical precision. Since both operate on largely the same data (we use trading halt and Trade and Quote (TAQ) data from the New York Stock Exchange (NYSE) for our estimates), that condition seems likely to hold. We find that the statistical precision of price-impact estimates falls dramatically around NYSE trading halts, a relation that cannot be fully attributed to patterns in volatility, volume, or the level of the price-impact estimates. This suggests that monitoring uncertainty plays a central role in motivating trading halts.

A commonly held view of trading halts is that they are designed to reduce informational asymmetries between traders.[2] These asymmetries are reflected in excessive price volatility (Spiegel and Subrahmanyam, 2000), trading crowd uncertainty (Greenwald and Stein, 1991), or transactional risk (Kodres and O'Brien, 1994). Collectively, we refer to these arguments as the level playing-field hypothesis. Lee, Ready, and Seguin (1994) examine the effectiveness of halts in this regard by comparing the volatility of stocks that undergo a trading halt to the volatility of stocks with similar characteristics in similar circumstances but no halt. They find that non-halting stocks have lower ex post volatility. To the extent that volatility proxies for asymmetries across traders, this evidence is inconsistent with the level playing-field hypothesis. Instead, it suggests that trading halts have perverse effects, as in Grossman (1990) and Subrahmanyam (1994).

Halts in our model relate only to the degree of information asymmetry between the specialist and the exchange. Our model neither predicts nor is inconsistent with the more traditional level playing-field hypothesis. However, monitoring uncertainty, the key variable in our analysis, is likely correlated with those variables predicted to be important under the level playing-field hypothesis. To clarify inferences, we incorporate those variables into our empirical analysis. Our results confirm Lee, Ready and Seguin's (1994) finding that volatility (liquidity) is unusually high (low) both before and after trading halts. However, because the unusual monitoring uncertainty observed just before trading halts cannot be attributed to excess volatility, our empirical analysis suggests that the control sample constructed by Lee, Ready and Seguin (1994) does not fully replicate market conditions around actual trading halts. Our evidence suggests that a halt was called in one case, and not in the other, because there was a greater erosion in the exchange's ability to monitor in the one case.

There has been considerable debate about the role of trading halts and circuit breakers. This debate is part of a more general polemic in which lawyers and economists have made arguments for and against the self-regulation of securities exchanges. For example, Fischel and Grossman (1984), Fischel (1986), Kyle (1988), Miller (1991), and Mahoney (1997) all argue that exchanges have the proper economic incentives to adopt rules and regulations meant to benefit their customers (i.e., investors). These incentives revolve around the reputation effects of low-quality transaction services and disorderly markets, and the competition (for transaction volume) from other exchanges. In these authors' view, external (e.g., government) regulation will simply muddle these clear incentives. The opposite view is best characterized by Pirrong (1995) who argues that self-regulation is unlikely to result in efficient outcomes for two reasons: (i) exchanges cannot possibly internalize the full effects of disorderly markets; (ii) exchange competition is in reality less intense than what is required for the above incentive arguments to work. For these reasons, Pirrong (1995) advocates external regulation instead of self-regulation. Our view in this paper is best approximated by Kyle (1988) who advocates the use of trading halts in the self-regulation of an exchange. However, the underlying reason for halting in our study is profoundly different from that in Kyle's discussion. Our paper also relates to DeMarzo, Fishman and Hagerty's (2000) exploration of the policies that self-regulatory organizations implement to monitor their members. The possibility that halting can enhance monitoring, our principal concern in this paper, is not investigated by these authors.

Our empirical work differs considerably from most previous empirical work on trading halts, in the sense that it concentrates on the causes of trading halts and not their effects. For example, Hopewell and Schwartz (1978), Kryzanowski (1979), King, Pownall and Waymire (1992), and Wu (1998) all study the price discovery process before, during, and/or after trading halts. Our work is mostly about the triggers of such halts. In that respect, it is closely related to Lee, Ready and Seguin (1994), and to Bhattacharya and Spiegel (1998) but, as discussed above, it argues for a related but different halt trigger.

In terms of the forces at play, our paper more generally relates to two strands of the economics literature. First we can think of the confidence of an exchange's customers as a public good that can be "polluted" by any of the specialists for their own private benefit.[3] Following Pigou (1932), Buchanan (1969), and Baumol (1972) among many others, our objective is to find an optimal way for the exchange to control these negative externalities. The postulated mechanism is similar to those documented by the aforementioned authors in that a corrective tax is assessed on the deviating agents. Alternatively, we can think of an exchange's group of specialists as a team whose objective is to maximize total profits. Absent coordinating mechanisms however, the team members are tempted to deviate and free-ride on the goodwill of others. Like Groves (1973), Holmstrom (1982), and Kandel and Lazear (1992), we develop a system of incentives designed to enhance team behavior. Our approach differs slightly from the above literature in that the peculiar structure of specialist markets leads us to introduce the possibility for either the monitor or the monitoree to "pull the plug" (i.e., to call a halt) when the mechanisms in place do not perform as well as they should.

Our paper is organized as follows. In section 2, we develop a model in which an exchange consists of many stocks, the market in each of which is made by an assigned specialist. This section also shows how the incentives of these specialists can conflict with those of the exchange as a whole. Section 3 shows how an appropriate system of monitoring involving trading halts helps reconcile these conflicts. The empirical predictions of the model are developed in section 4. In section 5, we describe the data used in our empirical analysis, and present the tests of the model. Concluding remarks are presented in section 6.

2 The Model

2.1 The Basic Setup

Consider an exchange on which N stocks are traded for a large (possibly infinite) number of periods. We assume that the market-making activities for each stock $i = 1, \ldots, N$ are performed by a different specialist i. Thus, in every period, the specialist for stock i quotes a price schedule which specifies the price at which he is willing to accommodate trades of different sizes. Specialists are assumed to be risk-neutral, and so seek to maximize the expected profits that their own market-making activities generate. Throughout the paper, we refer to the exchange as the set of all N specialists gathered in this one trading venue. Alternatively, though, the exchange can be viewed as a third party sharing the specialists' profits.[4]

The customers of an exchange typically trade in many stocks, and they typically have alternative venues where they can execute those trades. We assume that the customers' choice of trading venue is not made solely at the level of individual stocks, but partly at the exchange level. That is to say, customers assess the performance of an exchange by computing the average performance of every specialist on the exchange. A bad experience with any one specialist negatively affects a customer's overall likelihood of choosing that exchange for any of his subsequent trades (including those of different stocks). The idea is that a bad fill of an order implies that the exchange has insufficient checks and monitors in place, which in turn implies that another bad fill is likely even if the order goes to a different specialist. In fact, this line of thought is well captured by the following passage from Fischel and Grossman (1984, p. 293):

> [...] the existence of fraud on the exchange will discourage customer participation with all members; it will lead customers to trade with other exchanges which have a better reputation for preventing member fraud.

In other words, the occurrence (or even perception) of gouging anywhere on the exchange will tend to alienate at least some investors against the exchange as a whole. This investor dissatisfaction could take many forms, for example a reduction of order flow from existing customers, or a loss of new customers or listings because of bad publicity.

More specifically, let us assume that, in a given period, specialist i can quote either a conservative price schedule or an aggressive price schedule. We denote the event in which specialist i quotes a conservative (aggressive) price schedule by $\tilde{\lambda}_i = G$ ($\tilde{\lambda}_i = B$), where the use of G (B) is meant to capture the fact that the conservative (aggressive) price schedule will be good (bad) from the exchange's point of view. The conservative and aggressive price schedules lead to expected profits of $a > 0$ and $2a$ respectively for the specialist. However, the aggressive price schedule is perceived as gouging by the customer and, as a result, imposes a negative externality of c on the whole exchange.[5] This externality is borne equally by each of the N specialists, leading to a reduction of $\frac{c}{N}$ in each's expected profits for the period. For reasons to be made clear later, we assume that

$$\frac{c}{N} < a < 2a < c. \tag{1}$$

We restrict the specialists' action space to only two price schedules in order to simplify the analysis. A larger action space would include price schedules that lead to many different specialist profits and negative externalities, but would leave the economic forces at play largely the same. Also, practically speaking, the conservative and aggressive price schedules are likely to change from period to period, as they both depend on the prevailing market conditions. For modeling purposes however, we can abstract from that: our results only depend on the fact that different price schedules are preferred by the exchange (the conservative schedule) and individual specialists left unsupervised (the aggressive schedule). Similarly, the decision not to specify the customers' motives for trading allows us to concentrate on the relationship between the exchange and the specialists, and not on the relationship between the specialist and the traders. Incidentally, this is why our model has little to say about the level playing-field hypothesis, which directly results from the interactions between specialists and traders.

Before we proceed to analyze the equilibrium to this model, we want to emphasize that our reduced-form model seeks to capture an aspect of an exchange that is often overlooked in microstructure models: market makers who band together to form an exchange do not operate independently. An essential characteristic of an exchange is the pooling of market-makers' reputations. This necessarily implies that a specialist's actions in one stock can affect the market for other, seemingly unrelated stocks.

2.2 Incentives and the Impossibility of Cooperation

In this model, all periods are similar but they are independent from one another. Our analysis therefore concentrates on just one of those periods. Let us denote the profits of specialist i in any given period by $\tilde{\pi}_i$. These profits come from direct market-making activities ($\tilde{\delta}_i$), which are reduced by the negative externalities imposed on specialist i by the other specialists ($\tilde{\varepsilon}_{ij}$, $j \neq i$):

$$\tilde{\pi}_i = \tilde{\delta}_i - \sum_{\substack{j=1 \\ j \neq i}}^{N} \tilde{\varepsilon}_{ij}.$$

We incorporate the negative externalities that specialist i imposes on himself into his direct profits $\tilde{\delta}_i$. The setup of section 2.1 means that

$$\begin{aligned} \mathrm{E}\left[\tilde{\delta}_i \mid \tilde{\lambda}_i = G\right] &= a, \\ \mathrm{E}\left[\tilde{\delta}_i \mid \tilde{\lambda}_i = B\right] &= 2a - \frac{c}{N}, \\ \mathrm{E}\left[\tilde{\varepsilon}_{ij} \mid \tilde{\lambda}_j = G\right] &= 0, \quad \text{and} \\ \mathrm{E}\left[\tilde{\varepsilon}_{ij} \mid \tilde{\lambda}_j = B\right] &= \frac{c}{N}. \end{aligned}$$

Specialist i's price schedule only affects his direct profits. His choice of a price schedule will therefore be the same irrespective of how he expects the other specialists to quote, and his decision to quote conservatively or aggressively comes down to a comparison of his expected direct profits under each alternative:

$$\begin{aligned} &\mathrm{E}\left[\tilde{\pi}_i \mid \tilde{\lambda}_i = B\right] - \mathrm{E}\left[\tilde{\pi}_i \mid \tilde{\lambda}_i = G\right] \\ &= \left(\mathrm{E}\left[\tilde{\delta}_i \mid \tilde{\lambda}_i = B\right] - \sum_{\substack{j=1 \\ j \neq i}}^{N} \mathrm{E}\left[\tilde{\varepsilon}_{ij}\right]\right) - \left(\mathrm{E}\left[\tilde{\delta}_i \mid \tilde{\lambda}_i = G\right] - \sum_{\substack{j=1 \\ j \neq i}}^{N} \mathrm{E}\left[\tilde{\varepsilon}_{ij}\right]\right) \\ &= \mathrm{E}\left[\tilde{\delta}_i \mid \tilde{\lambda}_i = B\right] - \mathrm{E}\left[\tilde{\delta}_i \mid \tilde{\lambda}_i = G\right] \\ &= a - \frac{c}{N} > 0. \end{aligned}$$

This last inequality, which obtains from (1), reflects the fact that the specialist only absorbs a fraction $\frac{1}{N}$ of the externalities that he imposes on the exchange. Since specialist i cannot affect the negative externalities imposed on him by the other specialists, he will choose to quote the aggressive price schedule, as will every other specialist. As a result, $\mathrm{E}\left[\tilde{\varepsilon}_{ij}\right] = \mathrm{E}\left[\tilde{\varepsilon}_{ij} \mid \tilde{\lambda}_j = B\right] = \frac{c}{N}$ for all $i \neq j$ and, in equilibrium,

$$\mathrm{E}\left[\tilde{\pi}_i\right] = \mathrm{E}\left[\tilde{\pi}_i \mid \tilde{\lambda}_1 = \cdots = \tilde{\lambda}_N = B\right] = 2a - \frac{c}{N} - (N-1)\frac{c}{N} = 2a - c. \tag{2}$$

Suppose instead that the specialists somehow coordinate so that all quote a conservative price schedule. In that case, no negative externalities are created, and the expected profits of each specialist are given by

$$\mathrm{E}\left[\tilde{\pi}_i\right] = \mathrm{E}\left[\tilde{\pi}_i \mid \tilde{\lambda}_1 = \cdots = \tilde{\lambda}_N = G\right] = a > 2a - c,$$

where the last inequality follows from (1). Therefore, all specialists would be better off if they could each agree to be conservative, but such coordination is impossible, given the incentives of each. In effect, the reputation of the exchange vis-à-vis its customers is a "public good" for the specialists. Without coordination incentives, every specialist is tempted to free-ride on the goodwill of the other specialists. Clearly the specialists and the exchange as a whole stand to gain from implementing mechanisms that will mitigate this free-rider problem.

3 Monitoring and Trading Halts

This section shows how a combination of monitoring and trading halts can help reduce the coordination problem of the exchange. The first-best scenario for the exchange and its specialists is perfect coordination, in which every specialist quotes a conservative price schedule. If the exchange is able to observe each specialist's price schedule perfectly, this first-best scenario obtains with a simple monitoring scheme involving large punishments for aggressive price schedules.

Perfect monitoring is unrealistic. With his unique position as an intermediary for every trade in a particular stock, the specialist has intimate knowledge of the order flow, trading crowd, and general activity in that stock.[6] That knowledge allows the specialist to assess both the price schedule that would satisfy the exchange's coordination requirements (in our model, the conservative price schedule) and the potential for opportunistic actions (i.e., the private gain from quoting an aggressive price schedule). It strains credibility to assert that the exchange could match this information set for each of the specialists, at any cost. Yet, to assess with perfect accuracy whether a specialist's actions on a particular trade are abusive, the exchange must have the same information as the specialist. Realistically, the exchange relies on statistics and surveys to evaluate the performance of each specialist. This information is a strict subset of the specialists', making any kind of monitoring imperfect.

We capture these effects in our reduced-form model by assuming that the exchange can only imperfectly observe the specialist's quoted price schedule.[7] Specifically, we assume that at the end of every period the exchange receives a signal $\tilde{\theta}_i$, which provides incomplete information about $\tilde{\lambda}_i$. The precision of this signal $\tilde{p}_i \in [0, \frac{1}{2}]$ is random and independently redrawn in every period:[8]

$$\tilde{\theta}_i \mid (\tilde{\lambda}_i = B) = \begin{cases} B, & \text{prob. } \frac{1}{2} + \tilde{p}_i \\ G, & \text{prob. } \frac{1}{2} - \tilde{p}_i, \end{cases} \tag{3}$$

$$\tilde{\theta}_i \mid (\tilde{\lambda}_i = G) = \begin{cases} G, & \text{prob. } \frac{1}{2} + \tilde{p}_i \\ B, & \text{prob. } \frac{1}{2} - \tilde{p}_i. \end{cases} \tag{4}$$

Notice that the exchange finds out the exact strategy adopted by the specialist when $\tilde{p}_i = \frac{1}{2}$, but is just as likely to be right as wrong when $\tilde{p}_i = 0$. We say that the exchange has relatively precise (imprecise) knowledge of the efficiency of specialist i's price schedule when $\tilde{p}_i$ is close to $\frac{1}{2}$ (zero).[9] For analytical convenience, we also assume that $\tilde{p}_i$ is uniformly distributed on the $[0, \frac{1}{2}]$ interval.

The precision $\tilde{p}_i$ is the (inverse-)level of information asymmetry between the exchange and specialist i. We assume that the specialist knows $\tilde{p}_i$ at the beginning of every period, whereas the exchange knows $\tilde{p}_i$ at the beginning of the period only with probability $q \in [0, 1]$. Allowing q to be smaller than one captures the fact that the exchange is sometimes unaware of how well it will be able to assess the specialist in the future. While this assumption enhances the applicability and generality of our model, it will become clear later that it does not affect any of our results.

3.1 Monitoring

We assume that monitoring takes the form of an ex post evaluation of the specialist's performance based on the exchange's information. More precisely, since the aggressive price schedule is detrimental to the exchange, we assume that the exchange systematically punishes specialist i if it observes $\tilde{\theta}_i = B$. The penalty $\tilde{\tau}_i$ is a fine of $\kappa > 0$ imposed at the end of the period, where κ is determined at the outset by the exchange.[10] In practice, this penalty is related to the overall performance of the specialist, and can take on many forms ranging from a dollar fine to a loss of the specialist's seat on the exchange. More commonly, the penalty results in the exchange not considering the specialist in the assignment of new stocks, a topic analyzed by Corwin (1999).

The monitoring scheme is not perfect given that the exchange does not have perfect information about the specialist's actions during the period. Indeed, it is possible that $\tilde{\theta}_i = B$ when $\tilde{\lambda}_i = G$, so that the specialist is punished even after quoting the efficient price schedule. Similarly, it is possible that $\tilde{\theta}_i = G$ when $\tilde{\lambda}_i = B$, so that the specialist can get away with customer gouging. Nevertheless, it is clear from (5) and (6) that the specialist is more likely to be punished when he quotes an aggressive price schedule than when he quotes a conservative schedule. Indeed,

$$\mathrm{E}\left[\tilde{\tau}_i \mid \tilde{\lambda}_i = B, \tilde{p}_i\right] = \kappa\left(\frac{1}{2} + \tilde{p}_i\right) \geq \kappa\left(\frac{1}{2} - \tilde{p}_i\right) = \mathrm{E}\left[\tilde{\tau}_i \mid \tilde{\lambda}_i = G, \tilde{p}_i\right].$$

We assume that any tax paid by a specialist to the exchange is a deadweight loss to both the exchange and the specialist, that is it is not possible to redistribute the tax. Realistically, monitoring is costly. For example, the NYSE incurs costs in the form of computer equipment and human resources for its electronic supervision system, StockWatch. In that sense, our assumption of a deadweight tax amounts to saying that the monitoring scheme is exactly financed by the taxes that it charges. This assumption also ties in with the work of Fischel and Grossman (1984), and Mahoney (1997), who point out that the value of self-regulation should be analyzed as a tradeoff between the benefits that it generates and the costs that it involves. As we discuss later, the choice of κ by the exchange should reflect that tradeoff.

Since the exchange is, ultimately, just a collection of specialists, it is relevant to ask whether specialists will agree to monitoring ex ante (that is prior to the exchange or any specialist learning $\tilde{p}_i$, $i = 1, \ldots, N$, at the beginning of the first period). As it turns out, the answer depends on the pro-rata externality that gouging creates, $\frac{c}{N}$. If that quantity is sufficiently small, then there can be cases where the specialists would decline to implement the monitoring scheme. This is because the specialists only incur a small fraction of the negative externality that they cause when quoting an aggressive price schedule, so that the extra profits that they can generate for themselves are enough to cover the punishments associated with the monitoring scheme. Increasing the potential punishment κ does not solve the problem, as large ex ante expected taxes would render the market-making activities of the specialists completely unprofitable: the specialists would simply refuse to be monitored.

At the beginning of every period, specialists decide on a price schedule for the period. Since punishment is more likely when they quote aggressively, it is no longer the case that they will always prefer aggressive price schedules, as in section 2.2. As before, no specialist can affect the externalities that others will impose on them. So every specialist decides on a price schedule based on the expected profits that directly result from their market-making activities *and* on the expected punishment that this schedule brings about. Before quoting his price schedule, the specialist also knows $\tilde{p}_i$, the precision of the exchange's ex post monitoring. Therefore, his quoting decision is based on a comparison between $\mathrm{E}\left[\tilde{\delta}_i - \tilde{\tau}_i \mid \tilde{\lambda}_i = G, \tilde{p}_i\right]$ and $\mathrm{E}\left[\tilde{\delta}_i - \tilde{\tau}_i \mid \tilde{\lambda}_i = B, \tilde{p}_i\right]$.

If specialist i quotes a conservative schedule, then his direct expected profits for the period are given by

$$\mathrm{E}\left[\tilde{\delta}_i - \tilde{\tau}_i \mid \tilde{\lambda}_i = G, \tilde{p}_i\right] = a - \kappa\left(\frac{1}{2} - \tilde{p}_i\right). \tag{5}$$

If instead he quotes an aggressive price schedule, then his direct expected profits for the period are

$$\mathrm{E}\left[\tilde{\delta}_i - \tilde{\tau}_i \mid \tilde{\lambda}_i = B, \tilde{p}_i\right] = 2a - \frac{c}{N} - \left(\frac{1}{2} + \tilde{p}_i\right)\kappa = a + \alpha - \kappa\left(\frac{1}{2} + \tilde{p}_i\right), \tag{6}$$

where we have defined $\alpha \equiv a - \frac{c}{N} > 0$ as the net personal gain to gouging. The following lemma results from a simple comparison of (5) and (6).

Lemma 3.1 *Under monitoring, specialist i will quote a conservative price schedule ($\tilde{\lambda}_i = G$) if*

$$\tilde{p}_i > \frac{\alpha}{2\kappa}, \tag{7}$$

and an aggressive price schedule ($\tilde{\lambda}_i = B$) otherwise.

The specialist will choose to be conservative when the chance of the exchange making an incorrect inference of his actions is slight (i.e., high $\tilde{p}_i$), when the punishment is high, or when the net benefit to gouging is small. Otherwise, the increase in expected punishment when an aggressive price schedule is quoted is not sufficiently large to deter the specialist from reaping the gains of gouging the customer.

Note that setting κ to a value smaller than α does not affect the incentives of the specialists. Since $\tilde{p}_i \in \left[0, \frac{1}{2}\right]$, it can then never exceed $\frac{\alpha}{2\kappa}$, as required by equation (7). This is why we always restrict κ to be larger than α in what follows. Also, while equation (7) seems to indicate that setting κ to a large value induces the specialist i to quote a conservative price schedule most of the time, a large κ is not the solution to the coordination problem presented in section 2.2. A large κ means that the specialist will be punished severely when the exchange observes $\tilde{\theta}_i = B$. Since this can happen even when the specialist's actions are conservative, the ex ante expected value of punishments might be sufficiently large that the specialist simply refuses to be monitored. Recall that ex ante expected profits are $2a - c$ without monitoring. An acceptable κ cannot cause ex ante expected profits to fall below this amount. The following result will help us assess the effectiveness of the monitoring scheme.

Lemma 3.2 *Suppose that the exchange and the specialists adopt the monitoring scheme with some $\kappa > \alpha$. The ex ante expected profits of specialist i in any given period are then given by*

$$\mathrm{E}\left[\tilde{\pi}_i\right] = a\left(1 - \frac{\alpha}{\kappa}\right) + (2a - c)\frac{\alpha}{\kappa} - \left(\frac{\kappa}{4} + \frac{\alpha^2}{2\kappa}\right). \qquad (8)$$

This last result illustrates the monitoring tradeoff. As κ increases, more weight is put on the coordinated expected profits a, and less weight is put on the gouging outcome, $2a - c$. However, for $\kappa > \sqrt{2}\alpha$, better coordination comes at a price: the expected deadweight punishments, as represented by the last term in (8), are increasing in κ. The value of monitoring can therefore only be assessed by comparing the specialists' ex ante expected profits with and without it. If there exists a $\kappa > \alpha$ such that monitoring improves these expected profits, the exchange and the specialists will be willing to adopt monitoring. Otherwise, monitoring alone is not sufficient to solve the coordination problem of section 2.2. The following proposition assesses these possibilities.

Proposition 3.1 *The exchange and the specialists will agree ex ante to adopt monitoring if the net benefit to gouging, α, is sufficiently small. It is possible that the exchange and the specialists cannot agree to adopt monitoring if the self-imposed externality from gouging, $\frac{c}{N}$, is sufficiently small.*

This last result suggests that monitoring alone is not optimal in many plausible circumstances. In particular, since $\alpha = a - \frac{c}{N}$, N must be relatively small for monitoring alone to be both acceptable and effective. Intuitively, when N is large, specialists only absorb a small fraction $\frac{1}{N}$ of the negative externalities that they generate. As a result, a large punishment is needed to re-align their objectives with those of the exchange. Thus, with large N, the specialists face a choice of two losing propositions: incur a lot of externality costs, or incur rampant punishment. Although increasing κ reduces the negative externalities borne by the exchange through better specialist coordination, the deadweight taxes associated with monitoring exceeds these reductions, and turn monitoring into a losing proposition. In the remainder of the analysis we show that trading halts improve specialists (and exchange) profits in all situations, in particular in this all-too-likely situation where monitoring would not be acceptable.

3.2 Trading Halts

As discussed in section 3.1, monitoring alone may not be a very useful, or acceptable, approach to reducing the specialists' externality problem. There is a second problem with this approach. Even if the specialists do agree ex ante to the scheme of monitoring (alone) proposed in section 3.1, they (and the exchange) nevertheless face isolated periods where monitoring and punishing does more harm than good. Suppose for example that α is sufficiently small to make monitoring ex ante agreeable. In some periods (as we will soon show), upon learning $\tilde{p}_i$ specialist i may find himself in a position where any quoting decision leads to negative expected profits: an aggressive schedule does not produce sufficient gouging profits to recoup expected fines, yet a conservative quote also has a strong chance of generating a fine. In these circumstances, the option to halt trading rather than post a price schedule allows the specialist to credibly signal to the exchange that he cannot perform his market making duties diligently yet profitably.

A similar situation could arise for the exchange. Recall that, in our model, the exchange learns $\tilde{p}_i$ with probability q at the beginning of every period. In some periods, it will learn that its precision, $\tilde{p}_i$, is small. This might correspond to, for example, a listed company warning the exchange that, due to idiosyncratic events at the firm (e.g., a pending merger announcement), the potential for informed trading in that company's stock is currently extremely high. That information tells the exchange that it will be very difficult to distinguish gouging from efficient pricing ex post. When the specialist learns of the situation (i.e., observes $\tilde{p}_i$), he sees an opportunity to gouge with little fear of getting caught. The option to halt trading, and reopen with a call auction, lets the exchange protect its reputation in these vulnerable moments of trading that stock.

So let us assume that upon learning $\tilde{p}_i$ the exchange has the option to halt stock i before the specialist is allowed to quote a price schedule, and that specialist i has the option of halting instead of quoting a price schedule.[11] In both cases, the halt results in zero profits for specialist i with no externality imposed on the other specialists.[12] The specialist's problem is similar to that analyzed in section 3.1, except that he now has the option to halt. The specialist will exercise this option whenever his expected profits are negative with both the conservative and aggressive price schedule. The following lemma summarizes this.

Lemma 3.3 *Suppose that the exchange and the specialists adopt the monitoring scheme with some $\kappa > \alpha$. If halting is allowed for the specialist, he will choose to do so when $\kappa \geq 2a + \alpha$ and*

$$\tilde{p}_i \in \left(\max\left\{0, \frac{a+\alpha}{\kappa} - \frac{1}{2}\right\}, \frac{1}{2} - \frac{a}{\kappa} \right).$$

Otherwise, the specialist quotes the same price schedule as in Lemma 3.1.

Figure A shows the regions in which the specialist will find it optimal to call a halt. In that figure, the white (shaded) region indicates when the specialist, upon observing $\tilde{p}_i$, prefers quoting a conservative (aggressive) price schedule to an aggressive (conservative) one; this region is the same as that described in Lemma 3.1. The hashed region shows when the specialist prefers halting to quoting any price schedule. It is never optimal for the specialist to halt when the potential punishment is small ($\kappa < 2a + \alpha$) because his expected profits from quoting either a conservative (when $\tilde{p}_i$ is large) or an aggressive (when $\tilde{p}_i$ is small) price schedule exceed the expected punishment resulting in greater-than-zero profits. The specialist may choose to halt for $\kappa \geq 2a + \alpha$. Halting occurs when $\tilde{p}_i$ is such that the specialist cannot recoup his expected punishment with expected trading profits from either a conservative or aggressive price schedule. Small values of $\tilde{p}_i$ in the conservative range increase the likelihood of undue punishment, and large values in the aggressive range increase the likelihood that the specialist is caught gouging. Both imply negative profits.

[Figure A about here.]

With probability q, the exchange observes $\tilde{p}_i$ before the specialist makes his quoting decision. When this observation implies that the specialist will quote aggressively, the exchange halts trading in stock i (recall that aggressive price schedules are by definition detrimental to the exchange). The non-hashed shaded region of Figure A indicates the parameter combinations in which the exchange calls a halt, conditional on having observed $\tilde{p}_i$. Note that none of the results of our model depend on the value of q. Assuming that the exchange sometimes observes its monitoring precision before the specialist only serves to illustrate that halting has a dual purpose in the presence of monitoring: it allows the specialists to avoid excessive punishments, and it allows the exchange to prevent opportunistic actions by the specialist when monitoring is difficult.

Lemma 3.3 and the above discussion show that halting is at times optimal for the specialist or the exchange *in the presence of monitoring*. However, we have not shown that specialists will find the monitoring-halting scheme acceptable in the first place. This is far from obvious, as granting the exchange halting rights may prevent the specialist from making (potentially large) trading profits. The rest of this section shows that incorporating halts into the monitoring scheme causes monitoring to always be accepted by specialists, even in cases where they otherwise would not find it acceptable. The following lemma is the analogue to Lemma 3.2 when halting is allowed.

Lemma 3.4 *Suppose that the exchange and the specialists adopt the monitoring scheme with some $\kappa > \alpha$, and that halting is allowed for both the exchange and the specialist. The ex ante expected profits of specialist i in any period are then given by*

$$\mathrm{E}\left[\tilde{\pi}_i\right] = \tag{9}$$
$$\begin{cases} a\left(1-\frac{\alpha}{\kappa}\right)+(1-q)(2a-c)\frac{\alpha}{\kappa}-\frac{\kappa}{4}-\frac{\alpha^2}{2\kappa}+q\frac{\alpha}{2}\left(1+\frac{\alpha}{2\kappa}\right) & \text{if } \kappa \in (\alpha, 2a+\alpha) \\ \frac{2a^2}{\kappa}+(1-q)(2a-c)\left[\frac{2(a+\alpha)}{\kappa}-1\right]-\frac{a^2}{\kappa}-(1-q)\left[\frac{(a+\alpha)^2}{\kappa}-\frac{\kappa}{4}\right] & \text{if } \kappa \in [2a+\alpha, 2(a+\alpha)] \\ \frac{a^2}{\kappa} & \text{if } \kappa \in (2(a+\alpha), \infty). \end{cases}$$

We need to verify that the specialists' expected profits are improved when halting is possible. We establish this in two parts. First, we show that the halting option always improves the ex ante expected profits of the specialists under any monitoring scheme with $\kappa > \alpha$. We then show that there is always at least one monitoring scheme $\kappa > \alpha$ which the exchange and the specialists can agree upon ex ante. In fact, the existence of this monitoring scheme does not depend on any parameter of the model. The following proposition establishes the first part.

Proposition 3.2 *Suppose that the exchange and the specialists adopt the monitoring scheme with some $\kappa > \alpha$. The exchange and the specialists always prefer this monitoring scheme when it is* *complemented with the possibility for either to halt trading.*

This result captures the essence of the role for trading halts in this multi-specialist model of an exchange: in the presence of information asymmetries between the exchange and its specialists, the effect of monitoring is always improved by allowing the different parties to halt trading in a stock. Notice that, unlike other models of trading halts, our model does not require excess volatility or an extreme informational event for a halt to be called. Instead, trading halts are simply a mechanism to enhance the effectiveness of monitoring.[13] The next proposition shows that this enhancement is so great that it eliminates the situation where specialists would rather accept the externality than instigate monitoring. Also, trading halts in our model are the result of optimal decisions made by the exchange or the specialists: they are not alternatives to market breakdowns (see, e.g., Bhattacharya and Spiegel, 1991, and Spiegel and Subrahmanyam, 2000) or exogenous mechanisms to re-implement dominating equilibria (see, e.g., Kumar and Seppi, 1994).

Proposition 3.3 *As long as halting is allowed for both the exchange and the specialists, there exists a $\kappa > \alpha$ such that the exchange and the specialists agree ex ante to monitoring.*

This result is quite important. Proposition 3.1 shows that it may be the case that the deadweight loss of punishment through monitoring (without the possibility to halt) exceeds the reduction of externalities for any κ. Propositions 3.2 and 3.3 tell us that, not only is it the case that trading halts make monitoring better for both the exchange and the specialists, but it is also the case that trading halts *always* make monitoring optimal.

Although not shown in Proposition 3.3, the exchange and the specialists can agree to monitoring for many values of κ. These different values for specialist punishment will imply different frequencies of trading halts and punishment, as well as different expected deadweight punishments. As mentioned in section 3.1, assuming that punishments are deadweight losses is essentially equivalent to assuming that the cost of monitoring is covered by the punishments that are paid by the specialists. The fact that different values of κ can be agreed upon leaves an additional degree of freedom for the exchange to set up a monitoring scheme that it can finance.

4 Empirical Implications

The central insight from the model is that trading halts can play an important role in the monitoring process that is critical to a specialist-organized exchange. In particular, trading halts come into play when information asymmetries between the exchange and the specialists render monitoring ineffective. In our model, this information asymmetry is captured by $\tilde{p}_i$, the probability that the exchange will correctly monitor the specialist in a given period.

Proposition 4.1 *The expected value of $\tilde{p}_i$ conditional on a halt being called for stock i is smaller than its unconditional expected value.*

There is no direct empirical analog to $\tilde{p}_i$. However, its essence is clear. To make this proposition testable, we rephrase it as follows.

Implication 1 *Around a trading halt for a stock, the exchange will have more difficulty measuring the actions (and the appropriateness of those actions) taken by the stock's specialist than at other times.*

Practically speaking, given the number of specialists and stocks on an exchange, statistics gathered by the exchange throughout the trading process will play an important role in assessing the actions of the specialists. This sets up our main empirical test: we replicate this statistical analysis, greatly simplified, and use the precision from that statistical analysis as our test variable.

In the model, the exchange always prefers that the specialists quote a conservative price schedule, i.e., the efficient schedule for the exchange. This efficient schedule is fixed and common knowledge; it is the actual quoted schedule that is assumed known with asymmetric precision. In practice, asymmetric information about the efficient schedule itself seems at least as likely as asymmetric information about the specialist's quoted schedule. The efficient schedule depends on minutely detailed, constantly changing information about market conditions, trader identity, etc. Because specialists get to know the behavior of their customers over time, both through careful observation of the order flow and direct communication, their conditional knowledge of the efficient schedule is likely substantial. The exchange cannot possibly match the richness of the information sets possessed by the hundreds of specialists it must monitor. As a result, the exchange's conditional information about the efficient schedule is relatively sparse. In our tests of the model, we take the view that uncertainties about both the efficient schedule and the quoted schedule are relevant in measuring the precision with which the exchange can assess the actions of the specialists.

A second theoretical innovation of our model is that trading halts result from decisions made by the exchange or the specialists. This yields a second testable implication.

Implication 2 *Trading halts should be called by the exchange in some cases, and by specialists in other cases.*

Bhattacharya and Spiegel (1998) document in their empirical study of NYSE trading halts between 1974 and 1988 that about 50% of trading halts are news dissemination halts, and about 50% are order imbalance halts. Our data suggests a 25%-75% breakdown for 1993 and 1994.[14] Typically, news dissemination halts are called by the exchange upon learning that a news event is about to affect the value of a stock. On the other hand, order imbalance halts are usually called (or more precisely requested) by a specialist who observes unusual trading activity in his stock. Implication 2 is supported by these observations.

Proposition 4.1 and Implication 1 do not depend on whether the halt was called by the exchange or the specialist. Yet we know from section 3.2 that exchange-called halts will occur when the exchange expects the specialist to quote an aggressive price schedule. According to Figure A, this will happen when $\tilde{p}_i$ is close to zero. Specialist-called halts on the other hand can occur for intermediate values of $\tilde{p}_i$; indeed, for $\kappa \in [2a+\alpha, 2(a+\alpha)]$, the specialist will choose to quote an aggressive price schedule when $\tilde{p}_i$ is between $\frac{a+\alpha}{\kappa} - \frac{1}{2}$ and $\frac{1}{2} - \frac{a}{\kappa}$. This leads to the following result.

Proposition 4.2 *The expected value of $\tilde{p}_i$ conditional on a halt being called by the exchange for stock i is smaller than its expected value conditional on a halt being called by specialist i.*

Again, this result is not directly testable, but it can be rephrased according to the essence of $\tilde{p}_i$ to yield the following implication.

Implication 3 *The precision of statistical estimates of the specialist's actions (and the appropriateness of those actions given market conditions) is lower around exchange-called trading halts (e.g., news dissemination) than around specialists-called trading halts (e.g., order imbalance).*

In the remainder of the paper we test these implications using a sample of NYSE halts for the years 1993 and 1994, along with trade and quote data for the same two years.

5 Empirical Analysis

5.1 Data

The sample is an exhaustive list of trading halts on the NYSE for the period December 31, 1992, through December 31, 1994. We use the Trade and Quote (TAQ) dataset of transaction-level data, which begins December 31, 1992, to study these halts. After filtering the sample,[15] 1,775 halts remain over the two-year overlap between the halt and TAQ datasets, or about 3.5 halts per day. There are 690 stocks involved for an average of 2.6 halts per stock. Summary statistics for these halts are presented in Table 1. Delayed opens[16] are about twice as common as intraday halts, and order imbalance halts are about three times as common as news dissemination halts. As discussed in section 4, this last fact supports Implication 2, as news dissemination (order imbalance) halts are typically called by the exchange (the specialists).

For each halt we construct a control sample and a halt sample from the TAQ dataset. The control sample is used to specify the statistical characteristics of price schedules under normal circumstances. It is formed by collecting transaction and quote data for each stock that experienced a halt, for a number of consecutive trading days beginning with a randomly selected date between December 31, 1992, and June 30, 1994. Transactions are signed +1, −1, and 0 by ascertaining whether the price is higher than, lower than, or equal to the preceding quote midpoint, respectively. Before signing transactions, the time of transactions is set back five seconds relative to the time of quotes, as suggested by Lee and Ready (1991). We restrict the analysis to NYSE transactions and quotes. The number of trading days is chosen so that the number of transactions collected is greater than 2,500, subject to a minimum of twelve trading days. This yields an approximately equal amount of control data for each stock in the halt sample, although the number of days for which data is collected differs widely across stocks. As shown in Panel A of Table 2, the median number of days for which control data was gathered is 66, while the 10^{th} (90^{th}) percentile is 19 (223).

The basic unit of observation in all that follows is a trading interval, rather than an individual transaction or quote. This reduces the noise in estimating the price impact from individual transactions; for example, errors in the timing of quotes and trades become less of a factor. Furthermore, in a setting where large traders break up a position change into smaller trades, this is perhaps a more natural way to measure orders. Finally, it facilitates aggregation across stocks with different trade frequency. The interval length is calculated so that an average of five transactions occur in that stock's control sample.[17] Panel A of Table 2 shows that the length of the trading interval varies materially across stocks.[18]

The halt sample consists of data for the five trading days before and after each halt. We also split these data into trading intervals, using the interval length calculated in the stock's control data. These intervals are set so that the end (start) of the last (first) interval before (after) the halt coincides with the start (end) of the halt.

For every trading interval in both the control and halt samples, we define a number of variables used in testing the model. Trade count refers to the number of transactions within the trading interval. Share volume refers to the total number of shares traded within the trading interval. Trade imbalance refers to the signed share volume over the interval, using a positive (negative) sign for buy (sell) transactions. Both the price change and return variables are computed using the last quote midpoint of the current and preceding trading intervals. The price volatility over a trade interval can be measured using the product of return with trade imbalance, or using what we refer to as high-low volatility. This latter measure is computed as the difference between the highest and lowest quote midpoints in the interval, divided by the previous interval's last quote midpoint.

Panel B of Table 2 presents summary statistics of these variables in each of our two samples. Trading is more active than usual around trading halts, with about 50% more transactions per interval, a 50% greater trade imbalance, and about twice the share volume. This evidence is consistent with that provided by Lee, Ready and Seguin (1994), and by Corwin and Lipson (2000). Panel C presents more on the distribution of transaction count within these trading intervals. The fact that trading interval volatility is only 20-25% higher in the halt sample than in the control sample is due to the fact that we are including trading intervals that are far from trading halts when calculating the halt sample averages. Indeed, as we will see later, the volatility close to the halt is typically much larger than the control sample's volatility.

5.2 Estimating the Specialist's Actions

In a specialist-organized exchange, the price impact of a transaction is governed by the specialist. Thus, we use the statistical precision of price-impact estimates as a proxy for the precision with which the exchange can monitor the actions of the specialist. To meaningfully compare and pool across stocks, we normalize by using the interval return (rather than price change) and by scaling the interval trade imbalance by the average interval share volume in the stock's control sample. We refer to the latter as the scaled trade imbalance. Also, all price-impact regressions have an intercept to remove trends in stock prices.

As noted in other studies, the statistical specification of the relation between signed trades and price changes is nonlinear. For example, Peterson and Umlauf (1994) decompose order imbalance into three terms (the sign, the level, and the square), and demonstrate that each is significant with the first two positive and the third negative. This implies a concave relation between trade imbalance and price change. Similarly, Jones, Kaul, and Lipson (1994) show that the sign of a trade has no less explanatory power than the signed trade imbalance, again suggesting a concave relation.

To facilitate this concave relation we employ a log transformation of the scaled trade imbalance as the independent variable in our price-impact regressions. The transformed trade imbalance is computed as follows. First, we calculate the logarithm of the absolute value of each interval's scaled trade imbalance, and denote this quantity by L_{it} for stock i in trading interval t. Second, for every interval t and stock i, we compute $[L_{it} - \min_t L_{it}] \geq 0$, where the minimum is calculated over all trading intervals in the stock's control sample. Third, we sign that excess with the sign of the interval's trade imbalance. We denote the resulting variable by T_{it} and set it equal to zero if the interval's trade imbalance is zero. This transformation is centered on zero and fans out in both the positive and negative direction with log curvature. It leads to a parsimonious regression,

$$R_{it} = \alpha_i + \lambda_i T_{it} + \epsilon_{it}, \tag{10}$$

where the single parameter, λ_i, measures the price impact of trades, and where the return and transformed trade imbalance for stock i in trading interval t are denoted by R_{it} and T_{it} respectively. This regression is estimated separately for every stock $i = 1, \ldots, 690$ in the control data, over all available trading intervals t in the control sample. The estimated coefficient λ_i measures the incremental return associated with a ten-fold increase in the trade imbalance (we use base 10 log for this reason). In this simplified model of the monitoring process, the actions of the specialist are assessed with a single estimated regression coefficient. The standard error of the estimated coefficient is a natural proxy for the imprecision with which the exchange can assess the specialist.

The performance of the log-transformed specification is contrasted with the performance of a linear specification in Table 3. Not surprisingly, positive (negative) price changes tend to be associated with a positive (negative) trade imbalance, as reflected by the significantly positive λ_i coefficients. Both the average regression R^2 and the average precision of inferences (i.e., t-statistic) are materially higher with the concave specification.[19] In what follows, results are presented only using the log-transformed price-impact estimates.

5.3 Testing the Model

5.3.1 An Event Study of Trading Halts

Our model predicts that the information asymmetry between the exchange and the specialist will tend to be abnormally high around trading halts. Ideally, this could be tested by applying the technique of the previous section to trading intervals around the halts, and assessing whether the uncertainty from those regressions (i.e., the standard error of the λ_i estimate) is abnormal versus the control sample. However, there is simply not enough data concentrated around trading halts to run such an analysis.

An alternative procedure is to use the regression equations estimated in section 5.2 from the control sample, focusing on the relative or abnormal price impact and information asymmetry around the halt.[20] Consider a subseries of K consecutive trading intervals. The price impact over these K intervals can be written as

$$R_{jt} = \alpha_i + (\lambda_i + \gamma_j)T_{jt} + \varepsilon_{jt}, \quad t = 1, \ldots, K, \tag{11}$$

where ε_{jt} represents the disturbance to the price-impact relation in trading interval t for halt j, and γ_j represents the average price impact deviation (from the control sample λ_i) that obtains over those K trading intervals. Note that α_i and λ_i are not estimated in this regression: rather, they are the coefficients that were estimated over the stock's full control sample (in section 5.2). This specification captures the idea that, in normal conditions, we expect a return of $\alpha_i + \lambda_i T_{jt}$ in trading interval t, given a transformed trade imbalance of T_{jt}. Thus γ_j represents the average excess price impact coefficient over the subseries of trading intervals, and the standard error of the γ_j estimate indicates the precision with which that excess can be assessed over those intervals. Implication 1 predicts that this precision is lower when the subseries of trading intervals is close to a trading halt than when the subseries is randomly selected from the control sample.

Although our model does not offer any prediction about the size of γ_j, the level playing-field hypothesis predicts that liquidity dries up around trading halts. As pointed out by Kyle (1985), the effect of trade imbalance on prices essentially measures the inverse of liquidity. Therefore, to assess both our model and the level playing-field hypothesis, we perform an event study test around trading halts on both the γ_j estimates and the standard error of those estimates, denoted δ_j.

We use a simulated empirical distribution for the statistics γ_j and δ_j to evaluate whether halt observations are abnormal. For each halt, we randomly select 1,000 trading periods from the corresponding stock's control sample. We then run regression equation (11) over the six consecutive trading intervals ($K = 6$) beginning with the selected interval (i.e., six observations, one degree of freedom). This provides an estimate of the coefficient γ_j and its standard error δ_j over that subseries of intervals. These 1,000 randomly drawn estimates are then used to characterize the distribution of γ_j and δ_j under normal circumstances. Note that a separate distribution is estimated for each stock.

The event study uses the six consecutive trading intervals immediately preceding the halt, and three more sets of six consecutive intervals before that. Likewise we consider four subseries after the halt. For each halt, we run regression (11) in each of these halt subseries (again, each regression has six observations with the unit of observation being a trading interval). Finally, the γ_j and δ_j estimates are converted to a percentile value using the corresponding control-sample empirical distribution.[21]

Table 4 presents these percentiles aggregated across halts for each of the eight subseries characterizing the event study. The table also shows t-statistics against the null hypothesis that the mean halt-sample estimate is at the 50^{th} percentile of the corresponding control-sample empirical distribution. There is a statistically significant increase in the percentile ranking for the estimates of γ_j around a halt. This neither supports nor rejects our model, which does not offer any prediction about the steepness of the price schedule around trading halts. However, it supports the level playing-field hypothesis which argues that excessive information asymmetries and low depth levels trigger halts. Note that the γ_j estimates revert back to the 50^{th} percentile of the control-sample distribution relatively quickly after a halt. There is also a statistically significant increase in the percentile ranking for the uncertainty in monitoring precision, δ_j. This is consistent with Implication 1 of our model.

5.3.2 Controls

From Table 4, the abnormal uncertainty around a halt is both more extreme and more persistent than the abnormal price-impact coefficient around a halt. Nevertheless, it is difficult to disentangle the level playing-field hypothesis from the predictions of our model using these results, as the two variables are likely correlated. Similarly, it is well documented that volatility and trading volume tend to go up around trading halts (see, e.g., Lee, Ready and Seguin, 1994, and Corwin and Lipson, 2000). To better isolate the two hypotheses, we examine the marginal contributions in what follows.

The volatility and volume controls are structured similarly to the γ and δ variables. First a control-sample empirical distribution for the volatility (volume) within a subseries of six consecutive trading intervals is constructed. The volatility (volume) over such a subseries is computed as the mean of the high-low volatility (trade count) in each of the six trading intervals that make up the subseries. The volatility (volume) is similarly calculated over each of the event subseries, and a percentile value determined. The average of these percentiles, across halts, is presented in Table 5. Both volatility and volume are unusually large around trading halts. In fact, if anything, these two variables tend to be more abnormal around trading halts than γ and δ.

It is possible that the abnormal increases in γ and δ in Table 4 arise solely from their correlation with volatility and trading volume. Such a result would say nothing about whether halts are triggered by monitoring uncertainty or by an unlevel playing field. For example, the model does not assert that the costs associated with monitoring uncertainty are less important when volatility is abnormal. To evaluate whether the incidence of trading halts is consistent with the model, Table 4 is the relevant analysis. However, confidence in the model is higher if an effect is detected independently of confounding factors.

To this end we increment the analysis performed in section 5.3.1 with volatility and volume controls. The controls are constructed by regressing the 1,000 observations of γ and δ used to construct their control-sample empirical distributions on the corresponding subseries' volatility and volume. Thus, for each stock i, we estimate two regressions,[22]

$$\gamma_{it} = a_i^\gamma + b_{i1}^\gamma \sigma_{it} + b_{i2}^\gamma \sigma_{it}^2 + c_{i1}^\gamma \nu_{it} + c_{i2}^\gamma \nu_{it}^2 + \varepsilon_{it}^\gamma, \quad \text{and} \tag{12}$$

$$\delta_{it} = a_i^\delta + b_{i1}^\delta \sigma_{it} + b_{i2}^\delta \sigma_{it}^2 + c_{i1}^\delta \nu_{it} + c_{i2}^\delta \nu_{it}^2 + \varepsilon_{it}^\delta, \tag{13}$$

where σ_{it} is the volatility and ν_{it} is the volume in subseries t. The estimated coefficients, averaged across all stocks, are presented in Panel A of Table 6. That panel documents that about 20 to 25 percent of the variation in γ and δ across trading periods is explained by these controls. As suspected, volatility is positively correlated with both variables. Perhaps less intuitively, volume is negatively correlated with both, suggesting that market depth and estimation precision both rise with volume. The marginal significance of the squared terms indicates that the relationship is nonlinear.

Using the control-sample coefficient estimates we construct the unexplained γ and δ for the eight subseries surrounding a halt. We then assign a percentile using the distribution of residuals from the control-sample regressions (12) and (13). These percentile estimates are then aggregated across all stocks and presented in Panel B of Table 6. Controlling for volatility and volume has a damping effect on γ and, to a lesser extent, δ. This is particularly true in the post-halt periods. The post-halt effect is not surprising, as post-halt volatility and volume are highly abnormal.

Even after controls, the pre-halt uncertainty in price impact, δ, is still highly significant just before a halt, with a percentile of 54.7 and a t-statistics of 6.8. Thus, the tendency to find unusually poor monitoring conditions around a halt is not solely attributable to abnormal volatility and volume. Evidently, information asymmetries between the exchange and the specialist tend to be high right before a halt, volatility or not. This result helps distinguish the model from alternative explanations. The evidence on pre-halt γ is less clear, with most observations being at or near the 50^{th} percentile of the control-sample empirical distribution for γ after controls. Again, this does not rule out the level playing-field hypothesis: it just implies that price-impact coefficients provide no incremental information about the playing field beyond that provided with volatility.

Lee, Ready and Seguin (1994) find that trading halts do not dampen volatility. Harris (1998, p.35) conjectures that there must be something inherently different between the halts and pseudo-halts studied by Lee, Ready and Seguin. Table 6B is suggestive in this regard. Controlling for volatility and volume, it is still the case that it is unusually difficult to monitor in the period preceding a halt. However, after the halt monitoring is no more difficult than it usually is, given volatility and volume.

As noted above, it is difficult to disentangle the level playing-field hypothesis from the predictions of our model since γ and δ are likely correlated. We attempt to shed some light on the incremental effect of δ given γ by augmenting equation (13) to include γ_{it} as a regressor. We then repeat Table 6B using these residuals. While the abnormal δ just prior to halting declines somewhat (from percentile 54.7 to percentile 53), adding the cross effect does not change the result that the uncertainty in price impact increases significantly in advance of the halt.

5.3.3 Types of Trading Halts

Implication 3 predicts that δ_j and ε_j^δ are higher around halts called by the exchange than halts called by the specialist. Because news dissemination halts are typically called by the exchange whereas order imbalance halts are called by the specialists, this partition provides a means to test Implication 3. However, there is another differences between these halt types that tends to obfuscate the analysis. News dissemination halts are usually triggered by a company official notifying the exchange of corporate matters likely to lead to unusual market conditions. In response, trading is halted within minutes, even seconds. Conversely, order imbalance halts are usually preceded by an extended period of unusual order flow and trade dynamics, by definition. This difference in the duration with which "halt-like conditions" precede the halt itself goes against the predictions of Implication 3: monitoring imprecision several intervals before the halt is more likely with halts called by specialists.

Nevertheless, there is weak support for Implication 3 in Table 7, which examines the exchange's ability to monitor the specialist first without, and then with volatility and volume controls (i.e., δ as in Table 4, and ε^δ as in Table 6B). In Panel A, without controlling for volatility and volume, monitoring is more difficult at news dissemination halts in the subseries immediately preceding the halt, but not in the earlier subseries. However, this difference is not significant. Support is somewhat stronger in Panel B, where the difference is statistically significant and begins long before the halt is triggered.

5.3.4 Predicting Halts

Our analysis has so far concentrated on whether the monitoring conditions around observed trading halts is consistent with the predictions of the model. An alternate way to empirically assess the model is to use it to predict the occurrence of trading halts. We conduct such a test by constructing a pooled sample from the halt and control subseries. For each halt, we combine the subseries of six trading intervals immediately preceding the halt (i.e., subseries -1) with 100 randomly selected subseries of six trading intervals from the corresponding stock's control period. For each of these subseries, we calculate γ, δ, and average high-low volatility (σ), and convert them into percentiles against the corresponding empirical distributions (as in the preceding analyzes). These percentiles are then used in a probit regression with a dependent indicator variable that takes on a value of one when a trading period precedes a halt. Thus, the dependent variable is zero 100 times more often than it is one.

Two versions of this regression are considered: one with the γ and δ percentiles as independent variables, and the other adding the volatility percentile variable. The resulting coefficients are presented in Table 8. The first regression, without the volatility term, shows that both γ and δ provide independent predictive content for a subsequent halt. In particular, as goes the model, when monitoring uncertainty is high, a halt is more likely, even after controlling for concurrent abnormal price-impact. Likewise, the converse holds: abnormal price-impact also predicts a subsequent halt after controlling for monitoring uncertainty. Thus, there is evidence for both the monitoring and level playing-field motives for a halt. Adding a volatility term yields an interesting result: γ no longer helps to predict the occurrence of trading halts. However, δ still comes in significantly. Again, this does not preclude the level playing-field motive: it simply confirms that volatility alone is sufficient to account for that motive, that is liquidity conditions do not provide any incremental information.

6 Conclusion

Specialist-organized exchanges bear a principal-agent relationship: the exchange seeks to maximize the overall profits of all specialists, but individual specialists take the actions that generate those profits. Absent coordination, overall exchange profits are lowered by a negative externality that one specialist's pricing schedule imposes on all other specialists. With perfect monitoring, the exchange can motivate each specialist to quote the overall profit-maximizing, efficient schedule. However, information asymmetries between the specialists and the exchange compromise the effectiveness of monitoring. Indeed, when those asymmetries are particularly large, the expected costs associated with errors in inference on the one hand, versus customer gouging on the other hand, can imply negative profits for the exchange. Trading halts allow the exchange to avoid such a situation. Trading halts also allow the specialists to avoid situations where their expected individual profits are negative because of a high likelihood that even diligent actions will be punished, due to faulty monitoring caused by information asymmetries.

Our model of this principal-agent relationship demonstrates the central role that information asymmetry between the exchange and the specialists plays regarding trading halts. We test the model by constructing an empirically observable proxy for this information asymmetry. Our proxy for monitoring precision is the statistical precision with which price impact can be estimated using transaction level data. In a specialist-organized exchange, the price impact of transactions is governed by specialists who have a natural monopolistic position, so inferences about price impact are de facto inferences about the specialist's actions. We compare this measure around trading halts to the same measure in normal circumstances, and show that trading halts are associated with a substantial decline in the precision of inferences about the specialists' actions.

The tests jointly consider the more traditional level playing-field hypothesis, which predicts that informational asymmetries across traders will be unusually large around trading halts. The evidence is consistent with this hypothesis playing a role. However, the level playing-field hypothesis cannot explain the empirical finding that information asymmetries between the exchange and the specialists play an independent role. That is, after controlling for variables associated with the level playing-field hypothesis (volatility, volume, price-impact measures), our proxy for information asymmetry is still abnormally high around halts. Moreover, it still helps predict a halt.

Reputational concerns should be no less important to the profitability of dealer-organized exchanges, such as the NASDAQ, than they are to the profitability of specialist-organized exchanges like the NYSE. As with a specialist market, the overall reputation of a dealer market is driven by the actions of the individual dealers operating in that market. To the extent that a common reputational link to the exchange introduces a negative externality across dealers, the principal-agent structure that we model again applies. However, successful gouging of the customer in a dealer market requires the coordination of all dealers making a market in the affected stock. Absent such coordination, the offending dealer does not gouge the customer, he just takes himself out of the market. Thus, competition across dealers lowers the incentive to deviate from the exchange-optimal pricing schedule. Similar competition (from the trading crowd) is surely a factor at specialist-organized exchanges, but it seems plausible that forces of competition act stronger in the competing dealer arena. Thus, proxies for monitoring uncertainty may play a weaker role at dealer-market halts than at specialist-market halts.

The recent consolidation of specialist firms at the NYSE may be related to the agency problems considered in this paper. When operating independently, specialists are sometimes tempted to quote privately optimal price schedules, even though these schedules create negative externalities for the other specialists. In this context, the formation of specialist firms and the consolidation of existing firms may be a product of the ability of these firms to reduce the frequency and impact of such damaging actions through more efficient contracting. Indeed, if the incentives of the specialists operating within a firm are properly realigned, these specialists will avoid actions that reduce the firm's surplus. The extent to which the consolidation process can help resolve agency problems on the exchange then depends on the relative ability of the exchange and of specialist firms to realign their members' incentives.

Finally, our model does not explicitly describe the process by which trading after a halt is any less susceptible to monitoring problems than trading before the halt. It is worth noting, however, that the reopening mechanism is a call auction run by the specialist, who is allowed to participate only so as not to "upset the public balance of supply and demand as reflected by market and limited price orders."[23] Any trader concerned about poor monitoring conditions can avoid agency problems by participating in the reopening call auction. Madhavan and Panchapagesan (2000) find that specialists frequently participate in daily stock openings, which are subject to the same restriction. However, there is reason to think that the restriction is more binding for a reopening following a halt. Daily openings occur 700,000 or so times a year, compared to less than a thousand halts per year. Moreover, halts trigger additional rules on the reopening mechanism, e.g., opening indications and order imbalances must be posted without change for some time before the market can reopen. Thus, a halt exposes the specialist to more attention and scrutiny than that found in the daily opening. We conjecture that these factors imply less agency cost on the resumption of trading than existed prior to the halt.

Consistent with that conjecture, a substantial number of shares trades in the reopening auction: 0.98 of the average daily control-sample volume for the halted stock. Moreover, in the spirit of the level playing-field hypothesis, the halt serves as an explicit warning to market participants that monitoring conditions are poor. This can induce caution on the part of liquidity seekers (trading customers) and competitive entry by liquidity providers (the trading crowd). Indeed, Table 6B suggests that, after controlling for the extreme volatility and volume after a halt, abnormal monitoring problems no longer obtain. The mechanisms for resolution of the poor market conditions that triggered the halt in the first place are not well understood and represent an area in need of future research.

Appendix A

Proof of Lemma 3.1

It is straightforward to verify that the expression in (5) is strictly larger than that in (6) if and only if (7) is satisfied. ∎

Proof of Lemma 3.2

Notice that the total expected profits generated by all the specialists of the exchange in a given period (which we denote by $\tilde{\pi}^e$ for exchange profits) can be written as

$$\tilde{\pi}^e = \sum_{i=1}^{N} \tilde{\pi}_i = \sum_{i=1}^{N} \left(\tilde{\delta}_i - \sum_{\substack{j=1 \\ j \neq i}}^{N} \tilde{\varepsilon}_{ij} - \tilde{\tau}_i \right) = \sum_{i=1}^{N} \left(\tilde{\delta}_i - \sum_{\substack{j=1 \\ j \neq i}}^{N} \tilde{\varepsilon}_{ji} - \tilde{\tau}_i \right) \equiv \sum_{i=1}^{N} \tilde{\pi}_i^e,$$

where the last equality defines the contribution (after negative externalities) of specialist i to the exchange's profits. In equilibrium, since all the specialists are identical, they adopt the same strategy (i.e., they react the same way when they learn $\tilde{p}_i$), and so

$$\mathrm{E}\left[\tilde{\pi}_i\right] = \frac{1}{N}\mathrm{E}\left[\tilde{\pi}^e\right] = \mathrm{E}\left[\tilde{\pi}_i^e\right]. \tag{14}$$

When specialist i quotes $\tilde{\lambda}_i = G$, his expected contribution to the exchange's profits is

$$\begin{aligned} \mathrm{E}\left[\tilde{\pi}_i^e \mid \tilde{\lambda}_i = G, \tilde{p}_i\right] &= \mathrm{E}\left[\tilde{\delta}_i - \sum_{\substack{j=1 \\ j \neq i}}^{N} \tilde{\varepsilon}_{ji} - \tilde{\tau}_i \,\middle|\, \tilde{\lambda}_i = G, \tilde{p}_i \right] \\ &= a - \sum_{\substack{j=1 \\ j \neq i}}^{N} 0 - \kappa\left(\frac{1}{2} - \tilde{p}_i\right) = a - \kappa\left(\frac{1}{2} - \tilde{p}_i\right), \end{aligned} \tag{15}$$

whereas it is

$$\begin{aligned} \mathrm{E}\left[\tilde{\pi}_i^e \mid \tilde{\lambda}_i = B, \tilde{p}_i\right] &= \mathrm{E}\left[\tilde{\delta}_i - \sum_{\substack{j=1 \\ j \neq i}}^{N} \tilde{\varepsilon}_{ji} - \tilde{\tau}_i \,\middle|\, \tilde{\lambda}_i = B, \tilde{p}_i \right] \\ &= 2a - \frac{c}{N} - \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{c}{N} - \kappa\left(\frac{1}{2} + \tilde{p}_i\right) = 2a - c - \kappa\left(\frac{1}{2} + \tilde{p}_i\right) \end{aligned} \tag{16}$$

when he quotes $\tilde{\lambda}_i = B$. As shown in Lemma 3.1, specialist i will quote $\tilde{\lambda}_i = G$ when $\tilde{p}_i > \frac{\alpha}{2\kappa}$, and $\tilde{\lambda}_i = B$ otherwise. Therefore,

$$\begin{aligned} \mathrm{E}\left[\tilde{\pi}_i^e\right] &= \mathrm{E}\left\{\mathrm{E}\left[\tilde{\pi}_i^e \mid \tilde{\lambda}_i = G, \tilde{p}_i\right] \,\Big|\, \tilde{p}_i > \frac{\alpha}{2\kappa}\right\} \mathrm{Pr}\left\{\tilde{p}_i > \frac{\alpha}{2\kappa}\right\} \\ &\qquad + \mathrm{E}\left\{\mathrm{E}\left[\tilde{\pi}_i^e \mid \tilde{\lambda}_i = B, \tilde{p}_i\right] \,\Big|\, \tilde{p}_i \leq \frac{\alpha}{2\kappa}\right\} \mathrm{Pr}\left\{\tilde{p}_i \leq \frac{\alpha}{2\kappa}\right\} \\ &= a\left(1 - \frac{\alpha}{\kappa}\right) + (2a - c)\frac{\alpha}{\kappa} - \left(\frac{\kappa}{4} + \frac{\alpha^2}{2\kappa}\right). \end{aligned}$$

This last expression, along with (14), completes the proof. ∎

Proof of Proposition 3.1

The exchange and the specialists can only agree to adopt the monitoring scheme if the ex ante expected profits of the specialists are increased as a result of monitoring with some $\kappa > \alpha$. Without monitoring, the ex ante expected profits of every specialist are $2a - c$ (see equation (2)); with monitoring, they are given by equation (8). This means that monitoring is only possible when

$$a\left(1 - \frac{\alpha}{\kappa}\right) + (2a - c)\frac{\alpha}{\kappa} - \left(\frac{\kappa}{4} + \frac{\alpha^2}{2\kappa}\right) > 2a - c,$$

which is equivalent to

$$(c - a)\left(1 - \frac{\alpha}{\kappa}\right) - \frac{\kappa}{4} - \frac{\alpha^2}{2\kappa} > 0. \tag{17}$$

For $\alpha = 0$, this inequality is satisfied for all $\kappa \in (0, 4a)$ since

$$(c - a) - \frac{\kappa}{4} > c - 2a > 0.$$

This establishes the first part of the proposition. For $\alpha \approx a$ (i.e., for $\frac{c}{N}$ close to zero), the inequality in (17) becomes

$$(c - a)\left(1 - \frac{a}{\kappa}\right) - \frac{\kappa}{4} - \frac{a^2}{2\kappa} > 0,$$

which can never be satisfied if c is close enough to a. This establishes the second part of the proposition and completes the proof. ∎

Proof of Lemma 3.3

Upon learning $\tilde{p}_i$, the specialist prefers halting to quoting a conservative price schedule if

$$0 > a - \left(\frac{1}{2} - \tilde{p}_i\right) \quad \Leftrightarrow \quad \tilde{p}_i < \frac{1}{2} - \frac{a}{\kappa}. \tag{18}$$

Similarly, the specialist prefers halting to quoting an aggressive price schedule if

$$0 > a + \alpha - \left(\frac{1}{2} + \tilde{p}_i\right) \quad \Leftrightarrow \quad \tilde{p}_i > \frac{a+\alpha}{\kappa} - \frac{1}{2}. \tag{19}$$

Halting is preferred to quoting any price schedule when both (18) and (19) are satisfied. For $\kappa < 2a + \alpha$, this is impossible. For $\kappa \in [2a + \alpha, 2(a + \alpha)]$, this occurs when

$$\frac{a+\alpha}{\kappa} - \frac{1}{2} < \tilde{p}_i < \frac{1}{2} - \frac{a}{\kappa}. \tag{20}$$

Finally, for $\kappa > 2(a + \alpha)$, $\frac{a+\alpha}{\kappa} - \frac{1}{2} < 0$ and (19) is always satisfied, so that (20) reduces to (18). This completes the proof. ∎

Proof of Lemma 3.4

We use the same strategy and notation as in the proof of Lemma 3.2. First, suppose that $\kappa \in (\alpha, 2a + \alpha)$. As shown in Lemma 3.3, the specialist never halts when κ is in this range. His decision to quote a conservative or an aggressive price schedule therefore depends on whether (7) is satisfied or not. When the exchange learns $\tilde{p}_i$ at the beginning of the period, it halts whenever it expects the specialist to quote an aggressive price schedule (i.e., when $\tilde{p}_i \leq \frac{\alpha}{2\kappa}$), and then $\tilde{\pi}_i^e = 0$. Thus, for this range of κ, we can use (15) and (16) to get

$$\begin{aligned}
\mathrm{E}\left[\tilde{\pi}_i^e\right] &= \mathrm{E}\left\{\mathrm{E}\left[\tilde{\pi}_i^e \mid \tilde{\lambda}_i = G, \tilde{p}_i\right] \,\middle|\, \tilde{p}_i > \frac{\alpha}{2\kappa}, \text{no halt}\right\} \Pr\left\{\tilde{p}_i > \frac{\alpha}{2\kappa}, \text{no halt}\right\} \\
&\quad + \mathrm{E}\left\{\mathrm{E}\left[\tilde{\pi}_i^e \mid \tilde{\lambda}_i = B, \tilde{p}_i\right] \,\middle|\, \tilde{p}_i \leq \frac{\alpha}{2\kappa}, \text{no halt}\right\} \Pr\left\{\tilde{p}_i \leq \frac{\alpha}{2\kappa}, \text{no halt}\right\} \\
&= \mathrm{E}\left\{a - \kappa\left(\frac{1}{2} - \tilde{p}_i\right) \,\middle|\, \tilde{p}_i > \frac{\alpha}{2\kappa}\right\} \Pr\left\{\tilde{p}_i > \frac{\alpha}{2\kappa}\right\} \\
&\quad + \mathrm{E}\left\{2a - c - \kappa\left(\frac{1}{2} + \tilde{p}_i\right) \,\middle|\, \tilde{p}_i \leq \frac{\alpha}{2\kappa}\right\} (1-q) \Pr\left\{\tilde{p}_i \leq \frac{\alpha}{2\kappa}\right\} \\
&= \int_{\frac{\alpha}{2\kappa}}^{\frac{1}{2}} \left[a - \kappa\left(\frac{1}{2} - p\right)\right] 2\,dp + (1-q) \int_0^{\frac{\alpha}{2\kappa}} \left[2a - c - \kappa\left(\frac{1}{2} + p\right)\right] 2\,dp \\
&= a\left(1 - \frac{\alpha}{\kappa}\right) - \kappa\left(\frac{1}{4} - \frac{\alpha}{2\kappa} + \frac{\alpha^2}{4\kappa^2}\right) + (1-q)(2a-c)\frac{\alpha}{\kappa} - (1-q)\kappa\left(\frac{1}{4} + \frac{\alpha^2}{2\kappa^2}\right).
\end{aligned}$$

This expression yields the first line of (9). Now suppose that $\kappa \in [2a+\alpha, 2(a+\alpha)]$. Lemmas 3.1 and 3.3 tell us that the specialist, when the exchange does not halt, will quote an aggressive price schedule for $\tilde{p}_i \in \left[0, \frac{a+\alpha}{\kappa} - \frac{1}{2}\right]$, will halt for $\tilde{p}_i \in \left(\frac{a+\alpha}{\kappa} - \frac{1}{2}, \frac{1}{2} - \frac{a}{\kappa}\right)$, and will quote a conservative price schedule for $\tilde{p}_i \in \left[\frac{1}{2} - \frac{a}{\kappa}, \frac{1}{2}\right]$. When the exchange does observe $\tilde{p}_i$, it calls a halt for $\tilde{p}_i \in \left[0, \frac{a+\alpha}{\kappa} - \frac{1}{2}\right]$. Thus,

$$\begin{aligned}
\mathrm{E}\left[\tilde{\pi}_i^e\right] &= \mathrm{E}\left\{\mathrm{E}\left[\tilde{\pi}_i^e \mid \tilde{\lambda}_i = G, \tilde{p}_i\right] \,\middle|\, \tilde{p}_i > \frac{\alpha}{2\kappa}, \text{no halt}\right\} \Pr\left\{\tilde{p}_i > \frac{\alpha}{2\kappa}, \text{no halt}\right\} \\
&\qquad + \mathrm{E}\left\{\mathrm{E}\left[\tilde{\pi}_i^e \mid \tilde{\lambda}_i = B, \tilde{p}_i\right] \,\middle|\, \tilde{p}_i \le \frac{\alpha}{2\kappa}, \text{no halt}\right\} \Pr\left\{\tilde{p}_i \le \frac{\alpha}{2\kappa}, \text{no halt}\right\} \\
&= \mathrm{E}\left\{a - \kappa\left(\frac{1}{2} - \tilde{p}_i\right) \,\middle|\, \tilde{p}_i > \frac{1}{2} - \frac{a}{\kappa}\right\} \Pr\left\{\tilde{p}_i > \frac{1}{2} - \frac{a}{\kappa}\right\} \\
&\qquad + \mathrm{E}\left\{2a - c - \kappa\left(\frac{1}{2} + \tilde{p}_i\right) \,\middle|\, \tilde{p}_i \le \frac{a+\alpha}{\kappa} - \frac{1}{2}\right\} (1-q) \Pr\left\{\tilde{p}_i \le \frac{a+\alpha}{\kappa} - \frac{1}{2}\right\} \\
&= \int_{\frac{1}{2}-\frac{a}{\kappa}}^{\frac{1}{2}} \left[a - \kappa\left(\frac{1}{2} - p\right)\right] 2\,dp + (1-q)\int_0^{\frac{a+\alpha}{\kappa}-\frac{1}{2}} \left[2a - c - \kappa\left(\frac{1}{2} + p\right)\right] 2\,dp \\
&= \frac{2a^2}{\kappa} - \kappa\frac{a^2}{\kappa^2} + (1-q)(2a-c)\left[\frac{2(a+\alpha)}{\kappa} - 1\right] - (1-q)\left[\left(\frac{a+\alpha}{\kappa}\right)^2 - \frac{1}{4}\right],
\end{aligned}$$

which yields the second line of (9). Finally, suppose that $\kappa \in (2(a+\alpha), \infty)$. In this case, Lemmas 3.1 and 3.3 imply that the specialist never quotes an aggressive price schedule, so that the exchange never halts trading. The specialist quotes a conservative price schedule when $\tilde{p}_i > \frac{1}{2} - \frac{a}{\kappa}$ and halts otherwise. We then have

$$\begin{aligned}
\mathrm{E}\left[\tilde{\pi}_i^e\right] &= \mathrm{E}\left\{\mathrm{E}\left[\tilde{\pi}_i^e \mid \tilde{\lambda}_i = G, \tilde{p}_i\right] \,\middle|\, \tilde{p}_i > \frac{\alpha}{2\kappa}, \text{no halt}\right\} \Pr\left\{\tilde{p}_i > \frac{\alpha}{2\kappa}, \text{no halt}\right\} \\
&\qquad + \mathrm{E}\left\{\mathrm{E}\left[\tilde{\pi}_i^e \mid \tilde{\lambda}_i = B, \tilde{p}_i\right] \,\middle|\, \tilde{p}_i \le \frac{\alpha}{2\kappa}, \text{no halt}\right\} \Pr\left\{\tilde{p}_i \le \frac{\alpha}{2\kappa}, \text{no halt}\right\} \\
&= \mathrm{E}\left\{a - \kappa\left(\frac{1}{2} - \tilde{p}_i\right) \,\middle|\, \tilde{p}_i > \frac{1}{2} - \frac{a}{\kappa}\right\} \Pr\left\{\tilde{p}_i > \frac{1}{2} - \frac{a}{\kappa}\right\} + 0 \\
&= \int_{\frac{1}{2}-\frac{a}{\kappa}}^{\frac{1}{2}} \left[a - \kappa\left(\frac{1}{2} - p\right)\right] 2\,dp \;=\; \frac{2a^2}{\kappa} - \kappa\frac{a^2}{\kappa^2} \;=\; \frac{a^2}{\kappa},
\end{aligned}$$

and this yields the third line of (9). ∎

Proof of Proposition 3.2

To prove this result, we need to show that the expression for $\mathrm{E}\left[\tilde{\pi}_i\right]$ in Lemma 3.4 is bigger than that in Lemma 3.2 for any $\kappa > \alpha$. First, suppose that $\kappa \in (\alpha, 2a+\alpha)$. The difference between the two expressions is then

$$\left[a\left(1-\frac{\alpha}{\kappa}\right)+(1-q)(2a-c)\frac{\alpha}{\kappa}-\frac{\kappa}{4}-\frac{\alpha^2}{2\kappa}+q\frac{\alpha}{2}\left(1+\frac{\alpha}{2\kappa}\right)\right]$$
$$-\left[a\left(1-\frac{\alpha}{\kappa}\right)+(2a-c)\frac{\alpha}{\kappa}-\frac{\kappa}{4}-\frac{\alpha^2}{2\kappa}\right]=q(c-2a)+q\frac{\alpha}{2}\left(1+\frac{\alpha}{2\kappa}\right)>0.$$

Now suppose that $\kappa\in[2a+\alpha,2(a+\alpha)]$. The difference is now

$$\left\{\frac{2a^2}{\kappa}+\overbrace{(1-q)}^{\leq 1}\overbrace{(2a-c)}^{<0}\overbrace{\left[\frac{2(a+\alpha)}{\kappa}-1\right]}^{\geq 0}-\frac{a^2}{\kappa}-(1-q)\left[\frac{(a+\alpha)^2}{\kappa}-\frac{\kappa}{4}\right]\right\}$$
$$-\left\{a\left(1-\frac{\alpha}{\kappa}\right)+(2a-c)\frac{\alpha}{\kappa}-\frac{\kappa}{4}-\frac{\alpha^2}{2\kappa}\right\}$$

$$\geq\ \left\{\frac{2a^2}{\kappa}+(2a-c)\left[\frac{2(a+\alpha)}{\kappa}-1\right]-\frac{a^2}{\kappa}-\left[\frac{(a+\alpha)^2}{\kappa}-\frac{\kappa}{4}\right]+q\kappa\overbrace{\left[\left(\frac{a+\alpha}{\kappa}\right)^2-\frac{\kappa}{4}\right]}^{\geq 0}\right\}$$
$$-\left\{a\left(1-\frac{\alpha}{\kappa}\right)+(2a-c)\frac{\alpha}{\kappa}-\frac{\kappa}{4}-\frac{\alpha^2}{2\kappa}\right\}$$

$$\geq\ \frac{a^2}{\kappa}+\overbrace{(2a-c)}^{<0}\overbrace{\left[\frac{2a+\alpha}{\kappa}-1\right]}^{\leq 0}-\frac{a^2+2a\alpha+\alpha^2}{\kappa}+\frac{\kappa}{2}-a+\frac{a\alpha}{\kappa}+\frac{\alpha^2}{2\kappa}$$

$$\geq\ \frac{1}{\kappa}\left[-a\alpha-\frac{\alpha^2}{2}+\overbrace{\frac{\kappa^2}{2}}^{\geq\frac{\kappa}{2}(2a+\alpha)}-a\kappa\right]\ \geq\ \frac{1}{\kappa}\left[-a\alpha-\frac{\alpha^2}{2}+\overbrace{\alpha\kappa}^{\geq\alpha(2a+\alpha)}\right]\ \geq\ \frac{\alpha}{2\kappa}(2a+\alpha)\ >\ 0.$$

Finally, suppose that $\kappa\in(2(a+\alpha),\infty)$. In this case, the difference is

$$\frac{a^2}{\kappa}-\left[a\left(1-\frac{\alpha}{\kappa}\right)+\overbrace{(2a-c)}^{<0}\frac{\alpha}{\kappa}-\frac{\kappa}{4}-\frac{\alpha^2}{2\kappa}\right]\ >\ \frac{a^2}{\kappa}-a+\overbrace{\frac{a\alpha}{\kappa}}^{>0}+\frac{\kappa}{4}+\overbrace{\frac{\alpha^2}{2\kappa}}^{>0}$$
$$>\ \frac{1}{\kappa}\left(a-\frac{\kappa}{2}\right)^2\ >\ 0.$$

This completes the proof. ∎

Proof of Proposition 3.3

We need to find a $\kappa>\alpha$ such that $\mathrm{E}\left[\tilde{\pi}_i\right]$ in Lemma 3.4 is greater than $2a-c$. Take $\kappa=2(a+\alpha)$. Then, using equation (9), we have

$$\mathrm{E}\left[\tilde{\pi}_i\right]=\frac{a^2}{2(a+\alpha)}>0>2a-c. \qquad ∎$$

Proof of Proposition 4.1

We need to show that $\mathrm{E}\left[\tilde{p}_i \mid \text{halt}\right] < \frac{1}{4} = \mathrm{E}\left[\tilde{p}_i\right]$. For $\kappa \in (\alpha, 2a+\alpha)$, this is trivial, since halts are only called by the exchange when $\tilde{p}_i \in \left[0, \frac{\alpha}{2\kappa}\right] \subset \left[0, \frac{1}{2}\right]$. This is also trivial for $\kappa \in (2(a+\alpha), \infty)$, since halts are then only called by the specialist when $\tilde{p}_i \in \left[0, \frac{1}{2} - \frac{a}{\kappa}\right] \subset \left[0, \frac{1}{2}\right]$. So it only remains to show that this is true for $\kappa \in [2a+\alpha, 2(a+\alpha)]$. In that case,

$$
\begin{aligned}
&\mathrm{E}\left[\tilde{p}_i \mid \text{halt}\right] \\
&= \frac{q\mathrm{E}\left[\tilde{p}_i \mid \tilde{p}_i \le \frac{1}{2} - \frac{a}{\kappa}\right] \Pr\left\{\tilde{p}_i \le \frac{1}{2} - \frac{a}{\kappa}\right\} + (1-q)\mathrm{E}\left[\tilde{p}_i \mid \frac{a+\alpha}{\kappa} < \tilde{p}_i \le \frac{1}{2} - \frac{a}{\kappa}\right] \Pr\left\{\frac{a+\alpha}{\kappa} - \frac{1}{2} < \tilde{p}_i \le \frac{1}{2} - \frac{a}{\kappa}\right\}}{q \Pr\left\{\tilde{p}_i \le \frac{1}{2} - \frac{a}{\kappa}\right\} + (1-q) \Pr\left\{\frac{a+\alpha}{\kappa} - \frac{1}{2} < \tilde{p}_i \le \frac{1}{2} - \frac{a}{\kappa}\right\}} \\
&= \frac{q \int_0^{\frac{1}{2} - \frac{a}{\kappa}} 2p\, dp + (1-q) \int_{\frac{a+\alpha}{\kappa} - \frac{1}{2}}^{\frac{1}{2} - \frac{a}{\kappa}} 2p\, dp}{q\left(1 - \frac{2a}{\kappa}\right) + (1-q)\left(1 - \frac{2a}{\kappa} - \frac{2(a+\alpha)}{\kappa} + 1\right)} \\
&= \frac{q\left(\frac{1}{2} - \frac{a}{\kappa}\right)^2 + (1-q)\left[\left(\frac{1}{2} - \frac{a}{\kappa}\right)^2 - \left(\frac{a+\alpha}{\kappa} - \frac{1}{2}\right)^2\right]}{q\left(1 - \frac{2a}{\kappa}\right) + (1-q)2\left(1 - \frac{2a+\alpha}{\kappa}\right)} \\
&= \frac{1}{4} \frac{q(\kappa - 2a)^2 + (1-q)4\alpha(\kappa - 2a - \alpha)}{q\kappa(\kappa - 2a) + (1-q)2\kappa(\kappa - 2a - \alpha)}.
\end{aligned}
$$

This last quantity will be smaller than $\frac{1}{4}$ if and only if

$$
q(\kappa - 2a)^2 + (1-q)4\alpha(\kappa - 2a - \alpha) < q\kappa(\kappa - 2a) + (1-q)2\kappa(\kappa - 2a - \alpha).
$$

Since this inequality is linear in q, it is sufficient to verify that it holds for $q = 0$ and $q = 1$. For $q = 0$, we need to verify that $\kappa > 2\alpha$, which is true since $\kappa > 2a + \alpha > 2\alpha + \alpha > 2\alpha$. For $q = 1$, we need to verify that $\kappa - 2a < \kappa$, which is obviously true. This completes the proof. ▮

Proof of Proposition 4.2

This result only applies to values of $\kappa \in [2a+\alpha, 2(a+\alpha)]$, since this is the only range in which halts can be called by either the exchange or the specialist. In that range, the exchange calls a halt when $\tilde{p}_i \in \left[0, \frac{a+\alpha}{\kappa} - \frac{1}{2}\right]$, whereas the specialist calls one when $\tilde{p}_i \in \left(\frac{a+\alpha}{\kappa} - \frac{1}{2}, \frac{1}{2} - \frac{a}{\kappa}\right)$. The result easily follows. ▮

Footnotes

[1]We say that the specialist solicits, rather than calls, a halt to conform to the rules in effect at the NYSE (our data source). At the NYSE, only a floor official can call a halt, not the specialist.

[2]Halts in individual-stock trading differ from market-wide circuit breakers. Halts are called subjectively in response to that stock's market conditions; circuit breakers go into effect objectively, generally as a result of excessive price moves. This difference has often been neglected by the literature on trading halts; an exception is the work by Subrahmanyam (1995).

[3]This analogy was pointed out by an anonymous referee.

[4]For example, the NYSE collects fees from its specialists in exchange for the right to make a market. These fees come in the form of a fraction of the total trade volume handled by each specialist, in addition to a fixed annual fee. Modeling these fees explicitly would add little to the economics of the paper.

[5]This negative externality can be interpreted as the present value of all the costs imposed by traders on the exchange following a period of aggressive behavior by a specialist.

[6]For example, the relationship between the specialist and the exchange's customers is at the heart of Benveniste, Marcus and Wilhelm's (1992) market-making model.

[7]In our model, price schedules are either conservative or aggressive, and it is common knowledge that the exchange prefers the conservative schedule. The only uncertainty for the exchange is whether or not the specialist quoted it. In practice, the problem of determining the "conservative" (i.e., efficient) price schedule is itself probably just as difficult as that of inferring whether or not the specialist quoted it. We suspect that a simple modification to the model would facilitate this extra layer of uncertainty without altering our basic results. However the added complexity (two sources of uncertainty as opposed to just one) would not add any intuition. We will come back to this issue in our empirical analysis.

[8]The assumption that $\tilde{p}_i$ is independent across periods is made for tractability purposes: it allows us to treat one period at a time. In general, we should expect the precision process to

be autocorrelated. As our analysis will make clear, the main predictions of the model would be unaffected with this more general structure.

[9]In what follows, we assume that this is all the information that the exchange can use, that is we interpret $\tilde{\theta}_i$ as the total ex post information of the exchange. In particular, we assume that this information includes the prior distribution for the price schedule $\tilde{\lambda}_i$ quoted by the specialist in equilibrium.

[10]We also refer to this penalty as a tax, following Pigou (1932).

[11]We assume that halting without knowing $\tilde{p}_i$ is never optimal for the exchange. Otherwise, closing the exchange altogether would be the best alternative.

[12]Of course, in reality, the halt itself may impose some externality as well (Grossman, 1990). The important issue is that halting causes less of an externality than continuous trading in these situations. Without loss of generality, the model simply assumes that no externality are caused by halts.

[13]Of course, excess volatility or extreme information events would probably increase the information asymmetries between the exchange and the specialists, and so would be associated with more halts than regular situations. But they are certainly not necessary for halts to occur.

[14]This will be documented later in Table 1.

[15]These include the removal of any quotes or transactions with: price below \$1 or above \$200; change in quotes over \$2; specially coded trades; opening and closing trades; quotes with sizes of zero or 100 shares; percentage price moves greater than 50%; spreads greater than 20% of the previous spread midpoint. These filters are taken principally from Blume and Goldstein (1997). The largest cut came from restricting attention to stocks which have on average at least 5 trades per day. Taken together, these filters accounted for the removal of about 1.5% of the available data.

[16]Stocks on the NYSE must open within 20 minutes. Any stock that fails to do so becomes an officially delayed open, and must open according to the procedures outlined with a halt.

[17]We have repeated all of our analysis with ten trades and twenty trades, with no meaningful difference in results. We only present the results with five trades as they allow a finer partition around trading halts.

[18]The average number of trades per trading interval turns out to be smaller than 5.0 in the control sample because we did not allow any trading interval in that sample to overlap the overnight period. In other words, because the six and a half hour trading days do not contain an exact integer number of periods that have the required length, we end up truncating the last such period.

[19]In results not presented here, multiple regressions adding higher powers of trade imbalance or transformed trade imbalance did not significantly improve explanatory power.

[20]We are grateful to an anonymous referee for suggesting this approach.

[21]We also performed the analysis by comparing the actual regression coefficients in the halt and control samples, that is without the use of percentiles. The results were qualitatively similar.

[22]We considered many specifications for these regressions: with and without square terms, with and without cross-controls. The results were essentially the same with all specifications.

[23]See rule 104¶2104.11 of the New York Stock Exchange Guide.

References

Baumol, W. J, 1972, "On Taxation and the Control of Externalities," *American Economic Review*, 62, 307-322.

Benveniste, L. M., A. J. Marcus, and W. J. Wilhelm, 1992, "What's Special About the Specialist?," *Journal of Financial Economics*, 32, 61-86.

Bhattacharya, U., and M. Spiegel, 1991, "Insiders, Outsiders, and Market Breakdowns," *Review of Financial Studies*, 4, 255-282.

Bhattacharya, U., and M. Spiegel, 1998, "Anatomy of a Market Failure: NYSE Trading Suspensions (1974-1988)," *Journal of Business and Economic Statistics*, 16, 216-226.

Blume, M. E., and M. A. Goldstein, 1997, "Quotes, Order Flow, and Price Discovery," *Journal of Finance*, 52, 221-244.

Buchanan, J. M., 1969, "External Diseconomies, Corrective Taxes, and Market Structure," *American Economic Review*, 59, 174-177.

Corwin, S. A., 1999, "Specialist Portfolios, Specialist Performance, and New Listing Allocations on the NYSE," working paper, University of Georgia.

Corwin, S. A., and M. L. Lipson, 2000, "Order Flow and Liquidity around NYSE Trading Halts," *Journal of Finance*, 55, 1771 - 1805.

DeMarzo, P. M., M. J. Fishman, and K. M. Hagerty, 2000, "The Enforcement Policy of a Self-Regulatory Organization," working paper, University of California at Berkeley.

Fischel, D. R., 1986, "Regulatory Conflict and Entry Regulation of New Futures Contracts," *Journal of Business*, 59, S85-S102.

Fischel, D. R., and S. J. Grossman, 1984, "Customer Protection in Futures and Securities Markets," *Journal of Futures Markets*, 4, 273-295.

Greenwald, B. C., and J. C. Stein, 1991, "Transactional Risk, Market Crashes, and the Role of Circuit Breakers," *Journal of Business*, 64, 443-462.

Grossman, S. J., 1990, "Institutional Investing and New Trading Technologies," in *Market Volatility and Investor Confidence*, New York Stock Exchange.

Groves, T., 1973, "Incentives in Teams," *Econometrica*, 41, 617-631.

Harris, L. E., 1998, "Circuit Breakers and Program Trading Limits: The Lessons Learned," in Robert E. Litan and Anthony M. Santomero (ed.), *Papers on Financial Services*, Brookings Institution Press, Washington, D.C.

Holmstrom, B., 1982, "Moral Hazard in Teams," *Bell Journal of Economics*, 13, 324-340.

Hopewell, M. H., and A. L. Schwartz Jr., 1978, "Temporary Trading Suspensions in Individual NYSE Securities," *Journal of Finance*, 33, 1355-1373.

Jones, C. M., G. Kaul, and M. L. Lipson, 1994, "Information, Trading, and Volatility," *Journal of Financial Economics*, 36, 127-154.

Kandel, E., and E. P. Lazear, 1992, "Peer Pressure and Partnerships," *Journal of Political Economy*, 100, 801-817.

King, R., G. Pownall, and G. Waymire, 1992, "Corporate Disclosure and Price Discovery Associated with NYSE Temporary Trading Halts," *Contemporary Accounting Research*, 8, 509-531.

Kryzanowski, L., 1979, "The Efficacy of Trading Suspensions: A Regulatory Action Designed to Prevent the Exploitation of Monopoly Information," *Journal of Finance*, 34, 1187-1200.

Kodres, L. E., and D. P. O'Brien, 1994, "The Existence of Pareto-Superior Limits," *American Economic Review*, 84, 919-932.

Kumar, P., and D. J. Seppi, 1994, "Information and Index Arbitrage," *Journal of Business*, 67, 481-509.

Kyle, A. S., 1985, "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315-1335.

Kyle, A. S., 1988, "Trading Halts and Price Limits," *Review of Futures Markets*, 7, 426-434.

Lee, C. M. C., and M. J. Ready, 1991, "Inferring Trade Direction from Intraday Data," *Journal of Finance*, 46, 733-746.

Lee, C. M. C., M. J. Ready, and P. J. Seguin, 1994, "Volume, Volatility, and New York Stock Exchange Trading Halts," *Journal of Finance*, 49, 183-214.

Madhavan, A., and V. Panchapagesan, 2000, "Price Discovery in Auction Markets: A Look Inside the Black Box," *Review of Financial Studies*, 13, 627-658.

Mahoney, P. G., 1997, "The Exchange as a Regulator," *Virginia Law Review*, 83, 1453-1500.

Miller, M. H., 1991, *Financial Innovations and Market Volatility*, Blackwell, Cambridge, MA.

New York Stock Exchange, 2001, *New York Stock Exchange Guide*, Volume 2: Constitution and Rules, Commerce Clearing House, Inc., Chicago, IL.

Petersen, M. A., and S. Umlauf, 1994, "Quote Revisions and the Quoted Size: An Empirical Analysis," working paper, Northwestern University.

Pigou, A. C., 1932, *The Economics of Welfare* (4th ed.), MacMillan, London.

Pirrong, S. C., 1995, "The Self-Regulation of Commodity Exchanges: The Case of Market Manipulation," *Journal of Law and Economics*, 38, 141-206.

Spiegel, M., and A. Subrahmanyam, 2000, "Asymmetric Information and Disclosure Rules," forthcoming *Journal of Financial Intermediation.*

Subrahmanyam, A., 1994, "Circuit Breakers and Market Volatility: A Theoretical Perspective," *Journal of Finance*, 49, 237-254.

Subrahmanyam, A., 1995, "On Rules Versus Discretion in Procedures to Halt Trade," *Journal of Economics and Business*, 47, 1-16.

Wu, L., 1998, “Market Reactions to the Hong Kong Trading Suspensions: Mandatory Versus Voluntary,” *Journal of Business Finance and Accounting*, 25, 419-437.

Figure Legends

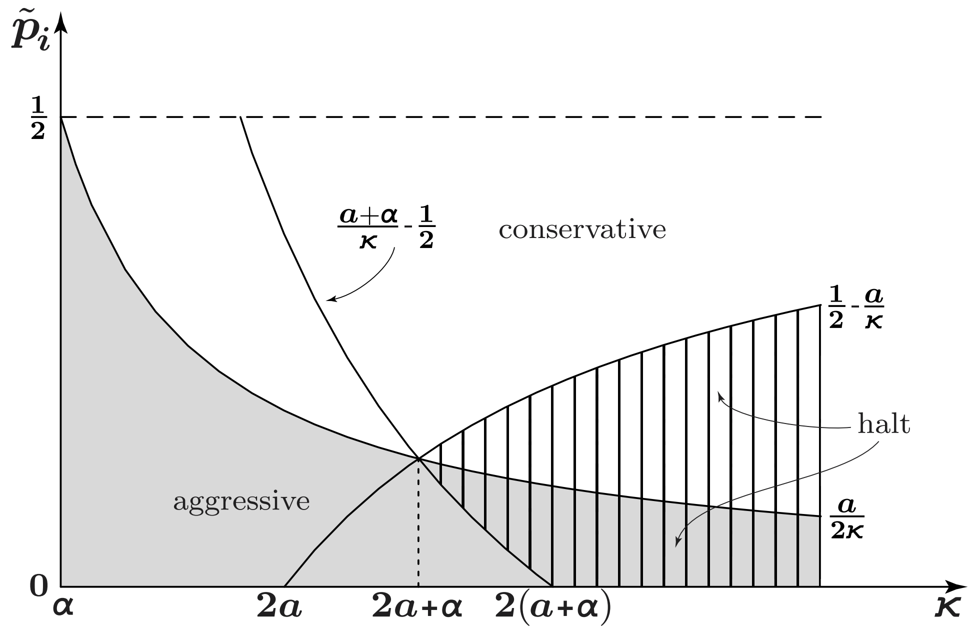

Figure A
Halting decisions by the specialist

Table 1
Frequency characteristics of NYSE halts

Halts per stock:	1	2-4	> 4
# stocks	303	290	97
Duration (mins.):	< 30	31-90	> 90
# halts	776	813	186

	News dissemination	Order imbalance	Total
Delayed open	182	989	1,171
Intraday halt	278	326	604
Total	460	1,315	

The sample of trading halts consists of 1,775 halts on the New York Stock Exchange between December 31, 1992 and December 31, 1994. These trading halts involve 690 different stocks, which we break down into stocks that experience one, two to four, and more than four halts over the two years. The trading halts are also broken down according to their length (in minutes) and their type. "News dissemination" halts are called when a major corporate announcement is either in progress or imminent, whereas "order imbalance" halts are called based on trading characteristics. A "delayed open" occurs when a stock fails to open by 9:50 AM, and an "intraday halt" means that the stock was trading when it was halted.

Table 2
Summary characteristics of the TAQ control sample and TAQ halt sample

Panel A: Control sample characteristics, distribution across stocks

	Percentile		
	10th	50th	90th
Number of trading days in control sample	19	66	223
Length of trading intervals (in minutes)	14	43	104

Panel B: Characteristics of trading intervals

	Control sample	Halt sample
Trade count	4.8	6.9
Volume (in shares)	12,715	26,581
Absolute value of trade imbalance (in shares)	7,715	12,988
Return × Sign of trade imbalance	0.24%	0.29%
High-low volatility	0.54%	0.70%
Standard deviation of interval return	0.60%	0.79%

Panel C: Distribution of the transaction volume in trading intervals

	Number of transactions					
	0	1 – 3	4 – 6	7 – 9	10 – 25	> 25
Control sample	3.3%	37.5%	32.9%	15.5%	10.3%	0.5%
Halt sample	3.9%	27.7%	27.3%	17.0%	20.9%	3.2%

A control sample from the Trade and Quotes (TAQ) data set is produced for every stock that experiences a halt. Each involves all transaction and quote data for consecutive trading days starting with a randomly selected date between December 31, 1992, and June 30, 1994, and each contains not less than 2,500 transactions and twelve trading days. The halt sample consists of the five days before and after the halt. Trading intervals are defined of length equal to the average time it takes for five transactions to occur in the stock's control sample. The variables in Panel B are first averaged across intervals for a stock, then averaged across stocks. Trading variables are computed using all observations within a trading interval. Trade imbalance refers to the sum of the interval's signed trades using the Lee and Ready (1991) algorithm. Returns are computed using successive intervals' ending quote midpoint. High-low volatility refers to the highest quote midpoint in a trading interval minus the lowest, divided by the previous interval's last quote midpoint.

Table 3
Specification of the statistical assessment of specialists in the control sample

Regressor	Mean slope (λ_i) coefficient	Mean t-statistic	R^2
Scaled trade imbalance	0.0019	8.4	13.2%
Log transformed trade imbalance	0.00087	11.5	20.9%

Trading intervals are defined for each stock so that they contain five transactions on average over the control sample (as in Table 2). A trading interval's return is defined using the last quote midpoint in that and the previous intervals. An interval's trade imbalance is defined as the sum of the interval's signed trades using the Lee and Ready (1991) algorithm, and an interval's scaled trade imbalance is defined as its trade imbalance divided by the stock's average volume per interval (averaged over the stock's control sample). Two regressions are run, separately for each stock: the return on the scaled trade imbalance, and the return on a log transformation of the scaled trade imbalance (as described in section 5.2). These regressions are run with an intercept (not presented). The table presents the mean statistics across stocks.

Table 4
Event study on the statistical assessment of specialists and the precision of that assessment around trading halts

	Subseries before the halt				(halt)	Subseries after the halt		
	-4	-3	-2	-1	1	2	3	4
γ percentile	50.9	51.1	53.2	52.9	60.7	54.1	50.5	50.1
	(1.2)	(1.5)	(4.4)	(3.8)	(14.1)	(5.8)	(0.9)	(0.2)
δ percentile	52.4	51.4	53.1	57.0	74.1	57.7	54.8	52.9
	(3.4)	(2.3)	(5.2)	(10.3)	(40.1)	(13.6)	(6.3)	(4.1)

Trading intervals are defined for each stock so that they contain five transactions on average over the control sample (as in Table 2). For each halt, subseries of six consecutive trading intervals are used to estimate γ in

$$R_t = \alpha_i + (\lambda_i + \gamma)T_t + \varepsilon_t, \quad t = 1, \ldots, 6,$$

where α_i and λ_i are the intercept and coefficient estimates from the corresponding stock's log-transformed trade imbalance regressions in Table 3. The standard error of the γ estimate is denoted δ. For each halt, γ and δ are estimated in each of the four subseries before a trading halt, and in each of the four subseries after a halt (i.e., the event study spans the 48 trading intervals around each halt, since each subseries is six consecutive trading intervals.) We similarly estimate γ and δ in 1,000 randomly selected subseries (each containing six consecutive trading intervals) in each stock's control sample. A percentile value is then assigned to each of the γ and δ estimates around a halt using the corresponding distribution of control-sample estimates. These percentile values are then pooled across trading halts, by relative subseries. The table presents the sample mean percentile, and a t-statistic (in parentheses) for the difference between the mean percentile and the 50th percentile.

Table 5
Event study on volatility and volume around trading halts

	Subseries before the halt				(halt)	Subseries after the halt		
	-4	-3	-2	-1	↓ 1	2	3	4
Volatility percentile	51.8 (2.5)	50.7 (1.0)	52.6 (3.4)	57.1 (9.1)	86.7 (81.7)	71.4 (34.6)	64.5 (22.1)	60.7 (14.5)
Volume percentile	53.0 (2.4)	53.0 (2.5)	55.2 (8.1)	62.0 (15.3)	91.4 (97.8)	83.0 (56.1)	76.8 (37.6)	73.1 (31.1)

Trading intervals are defined for each stock so that they contain five transactions on average over the control sample (as in Table 2). Subseries of six consecutive trading intervals are used, as in Table 4, to compute volatility and volume. Volatility is defined as the average high-low volatility (as in Table 2) over the six trading intervals in the subseries, and volume is defined as the average trade count per interval. For each halt, volatility and volume are estimated in each of the four subseries before a trading halt, and in each of the four subseries after a halt (i.e., the event study spans the 48 trading intervals around each halt, since each subseries is six consecutive trading intervals). Volatility and volume are similarly calculated in 1,000 randomly selected subseries (each containing six consecutive trading intervals) in each stock's control sample. A percentile value is then assigned to each of the volatility and volume computed around a halt using the corresponding distribution of control-sample estimates. These percentile values are then pooled across trading halts, by relative subseries. The table presents the sample mean percentile, and a t-statistic (in parentheses) for the difference between the mean percentile and the 50th percentile.

Table 6
Event study on the statistical assessment of specialists and the precision of that assessment around trading halts, controlling for volatility and volume

Panel A: Average coefficient estimates and R^2 from control-sample regressions

	b_1	b_2	c_1	c_2	R^2
γ regression	0.13	-1.17	-81.2	3.6	18.2%
	(58.1)	(-1.6)	(-15.0)	(2.0)	
δ regression	0.07	-1.13	-33.2	3.8	24.4%
	(51.3)	(-3.0)	(-6.4)	(2.0)	

Panel B: Event study on ε^γ and ε^δ (i.e., unexplained γ and δ) around trading halts

	Subseries before the halt				(halt) ↓	Subseries after the halt		
	-4	-3	-2	-1	1	2	3	4
ε^γ percentile	50.2	50.7	52.2	50.2	47.0	48.8	48.2	48.5
	(0.2)	(1.0)	(3.0)	(0.2)	(-3.3)	(-1.4)	(-2.2)	(-1.9)
ε^δ percentile	52.2	51.1	51.4	54.7	52.2	48.0	48.8	48.8
	(3.1)	(1.6)	(2.0)	(6.8)	(2.3)	(-2.3)	(-1.4)	(-1.5)

Trading intervals are defined for each stock so that they contain five transactions on average over the control sample (as in Table 2). Subseries of six consecutive trading intervals are used to compute volatility (denoted σ) and volume (denoted ν) as in Table 5, and estimates of γ and δ as in Table 4. For each halt, these four variables are estimated in each of the four subseries before a trading halt, and in each of the four subseries after a halt (i.e., the event study spans the 48 trading intervals around each halt, since each subseries is six consecutive trading intervals). Likewise, the four variables are estimated in each of 1,000 randomly selected subseries (each containing six consecutive trading intervals) in each stock's control sample. Using these control-sample estimates, the following regressions are estimated by stock:

$$\gamma_t = a^\gamma + b_1^\gamma \sigma_t + b_2^\gamma \sigma_t^2 + c_1^\gamma \nu_t + c_2^\gamma \nu_t^2 + \varepsilon_t^\gamma,$$
$$\delta_t = a^\delta + b_1^\delta \sigma_t + b_2^\delta \sigma_t^2 + c_1^\delta \nu_t + c_2^\delta \nu_t^2 + \varepsilon_t^\delta.$$

The resulting coefficients and R^2 are averaged across stocks and presented in Panel A (t-statistics are shown in parentheses). The unexplained γ (δ) for any subseries of six consecutive trading intervals is calculated as that subseries' γ (δ) less the predicted γ (δ) from the above control-sample regressions. The unexplained γ (δ) is converted to a percentile value using the empirical distribution of ε^γ (ε^δ) residuals from these regressions. Panel B presents the mean unexplained γ (δ) percentile values for the four subseries before and after a halt, where the average is taken across all halts. The t-statistics (in parentheses) correspond to the difference between the mean percentile and the 50th percentile.

Table 7
Event study around trading halts by halt types

Panel A: Event study on δ around trading halts

	Subseries before the halt				(halt)	Subseries after the halt		
	-4	-3	-2	-1	1	2	3	4
News dissemination	49.5	51.2	52.3	57.4	73.7	58.9	53.8	53.5
	(-0.3)	(0.9)	(1.6)	(5.0)	(17.5)	(6.5)	(2.6)	(2.4)
Order imbalance	53.4	50.9	52.8	55.6	74.8	57.4	55.5	52.2
	(4.1)	(1.1)	(3.1)	(6.5)	(34.3)	(9.5)	(6.7)	(2.7)

Panel B: Event study on ε^{δ} (i.e., unexplained δ) around trading halts

	Subseries before the halt				(halt)	Subseries after the halt		
	-4	-3	-2	-1	1	2	3	4
News dissemination	53.2	53.8	53.0	56.4	55.7	49.3	49.3	51.4
	(2.2)	(2.7)	(2.1)	(4.1)	(3.1)	(-0.4)	(-0.5)	(0.9)
Order imbalance	52.3	50.7	51.0	53.8	51.9	47.4	48.7	47.5
	(2.6)	(0.8)	(1.1)	(4.4)	(1.7)	(-2.8)	(-1.5)	(-2.9)

The analysis for δ and ε^{δ} performed in Tables 4 and 6B is repeated separately for the two halt types: news dissemination halts and order imbalance halts. The figures presented are the percentile values from the corresponding control-sample distribution. The t-statistics, reported in parentheses, test for the difference between the mean percentile and the 50th percentile.

Table 8
Probit regressions predicting a halt

	γ percentile	δ percentile	σ percentile
Regression 1			
Coefficient estimate	6.2	30.6	
Standard error	(2.9)	(3.0)	
Regression 2			
Coefficient estimate	-0.1	19.3	26.5
Standard error	(3.0)	(3.3)	(3.3)

Trading intervals are defined for each stock so that they contain five transactions on average over the control sample (as in Table 2). Periods of six consecutive trading intervals make up a subseries. A pooled sample of subseries is constructed from the control sample and the halt sample for each stock. For each halt, one subseries is taken from the halt sample — the subseries just preceding the halt. 100 randomly selected subseries are taken from the corresponding control sample. Three variables are calculated in each subseries: the percentile value of the price impact variable, γ (as computed in Table 4), the percentile value of the monitoring imprecision variable, δ (as computed in Table 4), and the percentile value of volatility, σ (as computed in Table 6). A fourth variable is set equal to one if the subseries precedes a halt, and zero if it comes from the control sample. The table presents the results of a probit regression of this indicator variable on γ, δ, and σ. All figures are in units of 10^{-4}. Standard errors are in parentheses.